Al-Romaimi, Khamis (2024) Asset Management Power Transformer Health Index Development and Analysis for Driving Asset Maintenance Strategy for Electrical Utilities. Doctoral thesis, The University of Sunderland.

Downloaded from: http://sure.sunderland.ac.uk/id/eprint/18664/

Usage guidelines

Please refer to the usage guidelines at http://sure.sunderland.ac.uk/policies.html or alternatively contact sure@sunderland.ac.uk.

Al-Romaimi, Khamis Salim Hamad (169137898), Asset Management Power Transformer Health Index Development and Analysis for Driving Asset Maintenance Strategy for Electrical Utilities. Doctoral thesis, University of Sunderland.

Usage guidelines

Please refer to the usage guidelines at http://sure.sunderland.ac.uk/policies.html or alternatively contact sure@sunderland.ac.uk.

Asset Management Power Transformer Health Index Development and Analysis for Driving Asset Maintenance Strategy for Electrical Utilities.

By

KHAMIS SALIM HAMAD AL-ROMAIMI

For PhD by Published and Creative Works

A thesis submitted in partial fulfilment of the requirements of the University of Sunderland for the degree of Doctor of Philosophy by Published and Creative works in Engineering.

FACULTY OF TECHNOLOGY

October 2024

DECLARATION

I, Khamis S H. Al-Romaimi declare that the thesis for the award of a Doctor of Philosophy Degree by published and creative works at the University of Sunderland, Faculty of Engineering and Advanced Manufacturing, hereby submitted by me, has not been previously submitted for a degree at this or any other university and that it is my original work in design and execution. Above all, the co-authors in the published works that are consolidated to form this thesis were my promoters and did not provide any technical intellectual contributions. Furthermore, all the reference materials contained therein have been duly acknowledged. Therefore, I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

.....

Khamis S H Al-Romaimi (Researcher)

.....

Prof. David Baglee (Director of studies)

.....

Dr. Derek Dixon (Co-Supervisor)

DEDICATION

Thanking continuously to Almighty God, who gave me the blessing, the knowledge, time, a good family and the strength to fulfil my dreams despite all the challenges during my PhD study journey.

I dedicate my work to my mother and my brother, Zayed Salim Hamad Al-Romaimi, who has raised me to be strong, and to my good family and children, who have supported me during the years of this research and study.

ABSTRACT

The Power Transformer Health Index (PTHI) has become an increasingly suitable condition evaluation method for driving maintenance strategies for electrical utilities. Many researchers considering PTHI lack practical field experience and Asset Management (AM) practice. AM practice has not been considered or reported widely as a vital factor for improving PTHI.

The thesis aims to integrate the researcher's technical field experience besides the AM standard ISO55000 framework and obtain a new PTHI health condition model. Developing a good analysis and understanding of the actual data, the PTHI model, and the benefits of driving PT asset management maintenance strategy is vital. This research employs a triangulation research methodology, a combination of qualitative, descriptive studies, and quantitative analysis, strengthening the overall validity, quality, reliability, accuracy and findings of the research PTHI model. This research uses detail condition assessment using conventional HI, Fuzzy Logic HI and eight Machine Learning (ML) algorithms, including Binary Logistic Regression (BLR), K-Nearest Neighbors (KNN), Naive Bayes (NB), Random Forest (RF), Decision Trees (DT), Artificial Neural Networks (ANN), and Support Vector Machines (SVM) for developing PTHI health condition models as binary and multi-class classification problems and in addition, using linear and polynomials regression models.

The research contributed to multiple publications and a better understanding of ML evaluation metrics for the imbalanced dataset. For the first time, the F1 score was introduced in this field in this research area compared with other researchers who used only overall accuracy. This research has also updated the transformers diagnostic condition assessment HI factors data interpretation list. The detailed analysis and the study led to two research impacts: the technical impact of using the PTHI web app tool to support AM and field maintenance engineers and the financial impact of the PTHI model. The financial analysis shows a vast saving impact for implementing the ML PTHI model, with results of 59.9% for considering early replacement plan decisions. The research recommended developing the international best practices of AM standard ISO 55000 framework and culture for the electrical utility. This development will support the effective PTHI model for implementing the data-driven decision-making approach in diverse situations and breaking the silo among departments.

TABLE OF CONTENTS

DEDICATION ii
ABSTRACT.. iii
TABLE OF CONTENTS iv
ACKNOWLEDGEMENTS ix
PUBLICATIONS x
LIST OF ABBREVIATIONS xi
LIST OF ACRONYMS xix
PREFACE..... xx
LIST OF TABLES xxi
LIST OF FIGURES i
Chapter 1: Introduction 2
1.1 Recent Development in Power Transformer Health Index Technology 4
1.2 Rationale Behind the Research 7
1.2.1 Studying the Health Condition of Power Transformers. 7
1.2.2 Utilisation of the Field Electrical Engineering Experience and Asset Management Knowledge 8
1.2.3 Integrate a Relationship Between the PTHI and the AM Framework 8
1.2.4 Authority Electricity Regulators (AER) Challenges to Electrical Utilities 10
1.2.5 Critical Role of Data Asset Management 11
1.3 Problem Statement 11
1.4 Research Questions 12
1.5 Thesis Format 12
1.6 Summary 15
Chapter 2: Literature Review 17
2.1 Introduction 17
2.2 Asset Management Practice 19
2.2.1 Asset Management Process for Power Transformers 24
2.2.2 Asset Management Global Data Problems for PTHI Development 28
2.3 Power Transformer System 30
2.3.1 Power Transformer Active Components Impacting Transformer Health 31
2.3.2 Diagnostic Parameters for PTHI 32
2.3.3 Power Transformer Condition Assessment Methods 35
2.4 The Importance of the Health Index 41
2.5 Power Transformer Health Index Development 44
2.5.1 Purposes of HI Condition Assessment 47

2.5.2 Risk Management for Failure Modes and Ageing Mechanism 48
2.5.3 Developing Failure Mode Assessment Score for Health Index 49
2.5.4 Power Transformer Health Index Calculation 52
2.5.5 The Role of PTHI in Driving Maintenance Strategy 56
2.6 The Research Power Transformer Health Index Approaches 61
2.6.1 Conventional Health Index Approach 61
2.6.2 Fuzzy Logic Health Index Approach 61
2.6.3 Artificial Intelligent Health Index Approach 62
2.7 Conclusion 65
2.7.1 Gaps in Knowledge 68
Chapter 3: Research Methodology 72
3.1 Introduction 72
3.2 Research Aim 72
3.3 Research Objectives 72
3.4 Research Approach 73
3.5 Research Framework 74
3.6 Data Management Process and Challenges 77
3.6.1 Data Collection Phase 77
3.6.2 Dataset Analysis and Preparation Phase 80
3.7 Research Data Model 85
3.8 Summary 86
Chapter 4: Case Study 1 Conventional Health Index Approach 89
4.1 Introduction 89
4.2 Data Workflow 90
4.3 Health Index Formulation 91
4.3.1 Scoring and Weighting Factors Using Standards Ranking Method 92
4.3.2 Data Interpretations 94
4.4 Case Study Investigation 97
4.4.1 Conventional Health Index (Singh and Swanson, 2018) 97
4.4.2 Conventional Health Index (Kadim et al.,2018) 98
4.4.3 Conventional Health Index Models Benchmark 100
4.5 Summary 106
Chapter 5: Case Study 2 Fuzzy Logic Health Index Approach 108
5.1 Introduction 108
5.2 An Experiment of Building a Simple Fuzzy Logic Model in MATLAB 108
5.2.1 Membership Functions Model 109
5.3 Expert Rules Conversion Results 115

5.4 Summary 117
Chapter 6: Case Study 3 Machine Learning Health Index Approach for 30 Electrical Power Transformers 120
6.1 Introduction 120
6.2 Principles of developing HI Machine Learning models 121
6.2.1 Objectives and end goal 121
6.2.2 Content of HI Model 121
6.3 Learn by Experiments and Implementing Machine Learning Models 123
6.3.1 The dataset description 123
6.3.2 Machine Learning Models 126
6.3.3 Evaluation Metric Used 131
6.4 Summary 137
Chapter 7: Case Study 4, Health Index for Electrical Utility dataset, using Machine Learning and Web App Tool 139
7.1 Introduction 139
7.2 Case Study Investigation 139
7.2.1 Data Collection 140
7.2.2 Data Preparation and Validation 140
7.2.3 Explanatory Data Analysis 142
7.2.4 Further Data Preparation and Steps Taken for Modelling. 148
7.3 Using Power Transformer Health Index Scoring Model in Excel 149
7.4 Comparison: HI Calculated by Electrical Utility and PTHI Scoring Model 150
7.5 Analysis of Electrical Utility Data Using the Flex Dashboards 151
7.6 Data Imbalance Problem 151
7.7 Implementation of More ML Algorithms on the Electrical Utility Under 69kV Transformers Data (Descriptive Quantitative Comparative Analysis in Spreadsheet) 156
7.7.1 Sensitivity 157
7.7.2 Accuracy 157
7.7.3 Specificity 158
7.7.4 F1 Score for Bad and Very Bad Scores 158
7.7.5 Research Result 161
7.8 Summary 162
Chapter 8: Case Study 5 Ageing Correlation with Health Index and Research Technical and Financial Impacts 164
8.1 Introduction 164
8.2 Hazard Plotting Approach for Failure Rate Evaluation 165
8.3 Case Study 165
8.3.1 Failures Rate (Ageing) Analysis 165

8.3.2 Understanding The Transformers Age Population 167
8.3.3 The Calculation of Apparent Age for Risk Assessment 168
8.3.4 Calculating Adjusted Failure Rate Based on Apparent Age Utilising a Regression Model. 168
8.3.5 Simple Linear Regression 169
8.3.6 Second-order polynomial (Quadratic) Regression 172
8.3.7 Comparison Between the Performance of a Linear First-Order and Second-Order Polynomial (Quadratic) Regression Model 173
8.3.8 Fitting a More Complex Cubic Regression Model 174
8.3.9 Regression Model's Summary 175
8.3.10 Evaluation of the Regression Models Performance 177
8.3.11 Calculating Remaining Life by Condition Utilising Health Index Model 177
8.4 Research Impacts 179
8.4.1 Technical Impact of PTHI Web App Tool 179
8.4.2 Financial Impact of the Effect of PTHI Asset Management Predictive Maintenance Decisions 187
8.5 Summary 195
Chapter 9: Conclusion and Recommendations 197
9.1 Introduction 197
9.2 Power Transformer Asset Management Health Index Modelling Analysis 198
9.3 Response to Research Questions 200
9.4 Conclusion 203
9.5 Recommendation for Maintenance Management in Electrical Utilities 207
9.5.1 Enhance Power Transformer Health Index Assessment. 207
9.5.2 Enhance Data Management Practices in the Electrical Utility. 209
9.5.3 Improve the Maintenance Strategy Tactics. 210
9.6 Research limitation 212
9.7 Research Contribution to Knowledge and Basis for Originality 213
9.8 Further Research 214
References- 216
Appendix A | Publications 264
Appendix B | Research Milestones 270
Appendix C | Recent Development in Power Transformer HI 275
Appendix D | Data Management Processing and Issues 274
Appendix E | Health Index Formulation 294
Appendix F | Failure Modes and Degradation Mechanisms 307
Appendix G | Scoring Assessment Matrix's 327

Appendix H | Modelling and Stimulation Results 331

ACKNOWLEDGEMENTS

I present one for the most of my sincere gratitude to Prof. David Baglee (Director of Studies) for supervising my research. His welcome and inspiration before I started the research journey and continued support and advice helped me to complete my research.

Also, words cannot express my gratitude to Dr. Derek Dixon and Dr. Michael Knowles (Co-Supervisors) for their support and advice during the research period.

Thanks to the industrial contact for their support, time, and data, as the data played a huge role in completing this PhD PTHI investigation.

This endeavour would not have been possible without the family, the children, and the wife's understanding, patience and support.

Acknowledging the University of Sunderland library staff for their support is a must.

Although sincere gratitude goes to the academic tutor module leader, Helen Scott, for her mentoring, encouragement, and support during the completion of this PhD thesis, the time she provided to learn her modules and improve the student's academic teaching skills is honoured and valuable, and it has made students interested in academic teaching and tutoring.

Finally, the project management MSc students I tutored during the last two years of this research are worth mentioning for their positive support and nomination for the student-nominated teaching awards. The winning prize is the nomination by the student for the years 2022/2023 and 2023/2024, proud of the nomination in the following categories: Most Digitally Engaging Tutor, Most Inspirational Member of University Staff, Outstanding Personal Academic Tutor, Outstanding Project Supervisor and Tutor Who Provides Outstanding Feedback.

Huge thanks go to all the staff of the University of Sunderland, including the GRS team, especially Miss Stephanie Brockett, Professor John Fulton, and Dr Mark Proctor.

PUBLICATIONS

The following publications are developed for this research and listed in Appendix A:

Research Journals:

Khamis Salim, David Baglee and Derek Dixon (2023) 'Health Index Assessment for Power Transformer Strategic Asset Management in Electrical Utilities', Int. J. of Strategic Engineering Asset Management. **DOI:** 10.1504/IJSEAM.2023.10059373

Khamis Salim, David Baglee and Derek Dixon (2023) 'Strategic Asset Management Health Index for Predicting Power Transformer Health Conditions', Int. J. of Strategic Engineering Asset Management.

Khamis Salim, David Baglee and Derek Dixon (2024) 'Risk-Stratify Power Transformers Based on Health Index Condition Utilizing Asset Management', Int. J. Risk Assessment and Management.

Research Conference:

Khamis Salim (2023) Power Transformer Asset Management Health Index in Electrical Utilities maintenance system. 3rd Annual Postgraduate Research Conference, Sunderland University, 14 September 2023.

Khamis Salim, David Baglee and Derek Dixon (2023) "Power Transformer Failure Rate Investigation Using Historical Maintenance Inspection Data", 12th IMA International Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) 4-6 July 2023 the University of Nottingham, Jubilee Campus, UK.

Khamis Salim (2022) Power Transformer Asset Management Health Index in Electrical Utilities maintenance system. 2nd Annual Postgraduate Research Conference, Sunderland University, 15 September 2022.

Reviewed Papers:

Chao, D. (2023) 'An Improved Detection Method for Key Components and Defects of Transmission Lines', IET Research Journals, Generation, Transmission and Distribution.

LIST OF ABBREVIATIONS

2FL	2-Furfuraldehyde content
ABB	Asea Brown Boveri
ACBDV	Alternating Current Breakdown Voltage
AEI	Ancillary Equipment Index
AER	Authority Electricity Regulators
AI	Artificial Intelligence
AIHI	Artificial Intelligent Health Index
AM	Asset Management
AMHA	Asset Management Health Assessment
AMS	Asset Maintenance Strategy
AMSP	Asset Management Strategic Planning
ANN	Artificial Neural Network
ASTM	American Society for Testing and Materials standards
BDV	Breakdown Voltage
BGR	Basic Gas Ratio
BLR	Binary Logistic Regression
BPEU	Best Performa Electrical Utilities
BSI	British Standards Institution
C	Chemical
CAPEX	Capital Expenditure

CBM	Condition-Based Maintenance
CBRCM	Condition-Based and Reliability-Centred Maintenance
CDKPI	Condition Data Key Performance Indicators
CEA	Canadian Electricity Association
CER	Continuous Emergency Rated
CI	Composite Index
CMMS	Computerized Maintenance Management Systems
CMR	Compact Modular Recloser
CSV	Comma Separated Value
DA	Detailed Assessment
DCG	Dissolved Combustible Gases
DDF	Dielectric Dissipation Factor,
DF	Dissipation factor
DGA	Dissolved Gas Analysis
DGAF	Dissolved Gas Analysis Factor
DP	Degree of Polymerisation
DQS	Data Quality Score
DRLF	Degradation Remaining Life Function
DSO	Distribution System Operator
DT	Decision Trees
E	Electrical

EDA	Exploratory Data Analysis
EPRI	Electrical Power Research Institute
ERP	Enterprise Resources Planning
ETL	Extract Transform Load Process
FAT	Factory Acceptance Tests
FIS	Fuzzy Inference System
FL	Fuzzy Logic
FN	False Negative
FP	False Positive
FRA	Frequency Response Analysis
FSA	Fleet Screening Assessment
GF	Grading Factors
GIS	Geographic Information System
GRNN	General Regression Neural Networks
HCGA	Health Check and Gap Analysis
HI	Health Index
HV	High Voltage
HVDC	High Voltage Direct Current
IAM	Institute of Asset Management
ID	Identification
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers

IF	Influencing Factors
IFT	Inter-Facial Tension
IMA	Institute of Mathematics and its Applications
ISO	International Organisation for Standardization
ITOMS	International Transmission Operation and Maintenance Study
KNN	K-Nearest Neighbors
kVA	kilovolt-Ampere
LF	Load Factor
LTPHI	Long-Term Plant Health Index
LV	Low Voltage
M	Mechanical
MATLAB	MATrix LABoratory
MCS	Monte Carlo Simulation
MDPI	Multidisciplinary Digital Publishing Institute
MHEW	Ministry of Housing Electricity and Water
MI	Mitigation Index/Score
MIMAR	Modelling in Industrial Maintenance and Reliability
ML	Machine Learning
MO	Mineral Oil,
MP	Master Plan
MR	Maschinenfabrik Reinhausen

MS	Microsoft
MVA	megavolt-amperes
NA	Not Available
NB	Naïve Bayes
NE	Natural Ester liquid
NF	Neural-Fuzzy
O&M	Operation and maintenance
OEM	Original Equipment Manufacturer
OIP	Oil Impregnated Paper
OLTC	On Load Tap Changer
ONAF	Oil Natural Air Forced
ONAN	Oils Natural Air Natural
OPEX	Operating Expenditure
OQ	Oil Quality
OQA	Oil Quality Analysis
OQF	Oil Quality Factor
PAM	PAS 55 Assessment Methodology
PAS	Publicly Available Specification
PASS	Publicly Available Specification Standard
PCB	Poly Chlorinated Biphenyl
PD	Partial Discharge

PDCA	Plan, Do, Check and Act
PF	Power Factor
PHI	Plant Health Index
PM	Planned Maintenance
PPM	Parts Per Million
PT	Power Transformer
PTHI	Power Transformer Health Index
R&D	Research and Development
RACI	Responsibility, Accountability, Consultancy and Informed
RBM	Risk-Based Maintenance
RBP	Resin Bonded Paper
RCA	Root Cause Analysis
RCM	Reliability-Centered Maintenance
RF	Random Forest
RHI	Reliability Health Index
RLI	Refurbishment Lifetime Index
RM	Risk Management
ROI	Return on Investment
$RS_{Mitigated}$	Refurbishment Score Mitigated
SAM	Strategic Asset Management
SAMPDM	Strategic Asset Management Planning Decision-Making
SCADA	Supervisory Control And Data Acquisition

SE	Synthetic Ester liquid
SELV	Separated Extra Low Voltage
SFRA	Sweep Frequency Response Analysis
SPA	Strategic Planning Assessment
SPSS	Statistical Platform Software Suite
Stds	Standards
SVM	Support Vector Machine
T	Thermal
TAI	Transformer Assessment Index
TBM	Time-Based Maintenance
TC	Tap Changer
TDCG	Total Dissolved Combustible Gases
TE	Total Expenditure
TEC	Transformer Electronic Control
TICM	Transformer Intelligent Condition Monitoring
TIP	Total Investment Plan
TN	True Negative
TOTEX	Total Expenditures
TP	True Positive
TR	Technical Regulators
TX	Transformer

txt	text
U	Unites
UHF	Ultra High Frequency
UK	United Kingdom
UoS	University of Sunderland
US	United States
VG	Very Good
vs	versus
WF	Weighting Factors
YoM	Year of Manufacture

LIST OF ACRONYMS

CIGRE	Conseil International Des Grands Réseaux éLectriques which means “International Council on Large Electric Systems”
C_2H_2	Acetylene
C_2H_4	Ethylene
C_2H_6	Ethane
CH_4	Methane
CO	Carbon Monoxide
CO_2	Carbon Dioxide
F-statistic	is the mean value obtained from regression analysis between two populations.
H_2	Hydrogen
mgKOH/g	milligrams of potassium hydroxide per gram
O_2	Oxygen
p-value	Probability of Standard Distribution value in R programming language

PREFACE

This research thesis uses the student's past work experience in the electrical industry sector at the Sultanate of Oman, Ministry of Housing Electricity and Water (MHEW) before privatisation, which covers generation, transmission, and distribution sectors. It also includes student technical work experience after the electrical sector privatisation of Oman Electricity Transmission Company. The research considers the researcher's deep knowledge of Asset Management (AM) and international best practices using the ISO55000 AM standard. The researcher's interest is to study in depth the integration of the technical and strategic roles of the AM in novel research model investigations that include real case studies. This HI requires considering the AM standard “ISO55000” framework as a strategic role towards developing the power transformer HI as a condition evaluation method as a technical role of the AM maintenance strategy.

The thesis incorporates various data types, including a literature review, simulated data, and available data from the electrical utility. The research supervisor is Professor David Baglee, Head of the Engineering Department. Initially, the second supervisor was Dr. Michael Knowles, and during the final year, Dr. Derek Dixon was assigned as a second supervisor after the earlier co-supervisor left the university.

LIST OF TABLES

Table 2.1: Fault measurements using gas ratios as in IEC 60599 (2007)..... 32
Table 2.2: The research PTHI variables. 33
Table 2.3: PTHI indicators (Picher et al., 2014). 35
Table 2.4: Health index scoring using conventional method (Khamis et al., 2023). .. 55
Table 3.1: The terms ‘data’, ‘information’ and ‘knowledge’ (Henke et al., 2016). 77
Table 3.2: PT subsystem and testing (Leauprasert et al., 2020). 80
Table 4.1: Comparison of conventional HI models (Singh and Swanson, 2018)..... 101
Table 4.2: Interpretation of HI numerical scoring value. 102
Table 4.3: Comparison of calculated HI models using Kadim et al. (2018) data. 103
Table 4.4: Confusion matrix between Kadim et al. (2018) and this research. 103
Table 4.5: Contrast between the PTHI model and other researchers’ models. 104
Table 4.6: HI Index scoring factor (Singh and Swanson (2018). 105
Table 5.1: PTHI ranges for 2FL and DP. 110
Table 5.2: HI ranges for PT health states. 114
Table 6.1: Demonstration of Binary Classification confusion matrix. 132
Table 6.2: Confusion Matrix comparison for the experimented ML models. 134
Table 7.1: Statistical summary of the electrical utility data*. 143
Table 7.2: Transformer oil colour score and weight factors (Jahromi et al., 2009).. 150
Table 7.3: Interpretation of transformer oil colour. 150
Table 8.1: PTs population based on system voltage. 165
Table 8.2: Failure analysis based on (CIGRE Technical Brochure 642, 2015). 166
Table 8.3: Case study samples data. 170
Table 8.4: The case study results for the linear first-order regression model. 170
Table 8.5: Case study results for second-order polynomial (quadratic) model. 172
Table 8.6: Comparison between the regression models' performance 177
Table 8.7: Cost calculation for the planned PT versus (vs) unplanned replacements after failure. 189
Table 8.8: Transformers in “Bad” condition. 192
Table 8.9: Nominal value ranges, conditions and scores for oil transformer. 192
Table 8.10:Oil replacement cost vs transformer replacement after breakdown..... 194

Table B-1: The research stages over the four years. 271

Table C-1: Recent developments in power transformer HI. 276

Table C- 2: Power Transformer Active Components. 283

Table D- 1: Main Data Issues..... 275

Table D- 2: Data validation methodologies (CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761, 2019)..... 276

Table D- 3 shows various solutions for controlling data unobtainability during PTHI development(CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761, 2019)..... 277

Table D- 4: Gas limit recommendations [PPM] based on various standards (Leauprasert et al., 2020). 279

Table D- 5: IEC/IEEE codes for DGA analysis techniques (Notingher et al.,2008) . 279

Table D- 6: Transformer rating based on DGA factor analysis (Jahromi et al., 2009)..... 280

Table D- 7: scoring and weight factors for gas levels [PPM] (Jahromi et al., 2009)..... 280

Table D- 8: Gas limit references assessment (Jahromi et al., 2009; IEEE Std, C57.104, 2008; IEEE Std, C57.106, 2006; CIGRE Working Group 05 1983). 280

Table D- 9: Gading technique for oil analysis parameters (Jahromi et al., 2009).... 281

Table D- 10: Furfural test or age ratings (Jahromi et al., 2009)..... 281

Table D- 11: Transformer components failure tests (Raj, 2019; Wattakapaiboon and Pattanadech, 2016; Hernanda et al., 2014; Naderian et al., 2008, Khamis et al., 2023). 282

Table D- 12: 30 Transformer's Medium Dataset (Abu-Elanien et al.,2012; Zuo, W. et al.,2016). 286

Table D- 13: 15 Transformer's small Dataset (Kadim et al.,2018) 288

Table D- 14: 3 Transformer's small Dataset (Singh and Swanson, 2018) 288

Table D- 15: Health Index scoring calculation and required decision (Jahromi et al., 2009)..... 289

Table D- 16: Health Index scoring interpretation (Naderian et al.,2008). 289

Table D- 17: Equations for calculating the final health index score. 290

Table D- 18: Scoring of DP assessment..... 291

Table D- 19: Scoring of Age Assessment. 291

Table D- 20: Scoring of TDCG assessment. 291

Table D- 21: Scoring of Basic Gas Ratio Assessment 291

Table D- 22: C2H2 Rating Scores 292
Table D- 23: H2 Rating Scores 292
Table D- 24: C2H6 Rating Scores 292
Table D- 25: C2H4 Rating Scores 292
Table D- 26: CH4 Rating Scores 292
Table D- 27: CO Rating Scores 293
Table D- 28: Scoring of moisture in paper 293
Table D- 29: Scoring of Moisture in Oil Assessment 293
Table D- 30: Scoring of Acidity Assessment 293
Table E- 1: HI design principles 295
Table E- 2: Types of Health Indexs (HIs) 298
Table E- 3: The two decisions are managed in these circumstances as follows: 303
Table E- 4: HI scoring methods (CIGRE Technical Brochure 761, 2019). 304
Table F- 1: List of failure modes (CIGRE Technical Brochure 761, 2019) 308
Table F- 2: Ageing mechanisms list (CIGRE Technical Brochure 761, 2019). 309
Table F- 3: State evaluation methods for transformer based on failure mode (CIGRE Technical Brochure 761, 2019). 312
Table F- 4: The various processes for power transformer data validation (CIGRE Technical Brochure 761, 2019). 318
Table F- 5: Power transformer component primary and sub-failure modes. 319
Table F- 6: Transformer active parts component's function, failure causes, and failure sources (Murugan and Ramasamy, 2019). 321
Table F- 7: List power transformer sub-systems affected by failure modes and ageing mechanisms. 324
Table G- 1: Categorised failure modes of the scoring matrix (CIGRE Technical Brochure 761, 2019) 328
Table G- 2: Thermal Assessment Scoring Matrix (CIGRE Technical Brochure 761, 2019) 329
Table G- 3: Isolating Fluids Scoring Matrix (CIGRE Technical Brochure 761, 2019). 330
Table H- 1: Grading Method for DGA Factor based FOR U<=69 kV 332
Table H- 2: Grading Method for Oil Test Parameters based FOR U<=69 kV 333
Table H- 3: Transformer Rating based on DGA Factor 334
Table H- 4: Transformer Rating based on OQF Factor. 334

Table H- 5: Age Rating 334
Table H- 6: Furan Rating 334
Table H- 7: Interpretation of Output 335
Table H- 8: The electrical utility Interpretation of HI value 335
Table H- 9: Fuzzy-Logic simple rules for the PTHI model. 336
Table H- 10: The formulas for the quantitative measures (Caret: Classification and regression training - the comprehensive R, 2023). 353
Table H- 11: ANN Confusion Matrix 356
Table H- 12: ANN Statistics By Class/Condition 356
Table H- 13: SVM Confusion Matrix 358
Table H- 14: SVM Statistics By Class/Condition 359
Table H- 15: RF Confusion Matrix 361
Table H- 16: RF Statistics by Class/Condition 362
Table H- 17: KNN Confusion Matrix 367
Table H- 18: KNN Statistics by Class/Condition 367
Table H- 19: NB Confusion Matrix 369
Table H- 20: NB Statistics by Class/Condition 369
Table H- 21: DT Confusion Matrix 371
Table H- 22: DT Statistics by Class/Condition 371
Table H- 23: Probability assessment using age function. 389
Table H- 24: Transformer HI condition monitoring and the remaining life ranges. .. 390
Table H- 25: The electrical utility interpretation of HI value according to the HI model. 398
Table H- 26: Grading Method for DGA Factor based FOR U<=69 kV 401
Table H- 27: Grading Method for Oil Test Parameters based FOR 69 kV < U < 239 kV. 402
Table H- 28: Transformer Rating based on DGA Factor 403
Table H- 29: Transformer Rating based on OQF Factor 403
Table H- 30: Age Rating 403
Table H- 31: Furan Rating 403
Table H- 32: transformer oil colour scoring and weighting. (Jahromi et al., 2009). . 404
Table H- 33: Interpretation of the electrical utility data oil colour into nominal values. 404

LIST OF FIGURES

Figure 1.1: Asset Management PTHI Development Process Khamis et al. (2024). 6

Figure 1.2: Asset management strategic planning subject group (The Institute of Asset Management, 2015). 9

Figure 2.1: Research Framework. 17

Figure 2.2: Categories of assets based on AM standard (PAS55, 2008). 20

Figure 2.3: Asset management system hierarchy (PAS55, 2008). 21

Figure 2.4: AM system (The Institute of Asset Management, 2015). 22

Figure 2.5: AM lifecycle activities (PAS55, 2008). 23

Figure 2.6: Research framework for PTHI development. 24

Figure 2.7: Ownership and electrical utilities use asset health indices (The Canadian Electricity Association, Deloitte, 2015) 25

Figure 2.8: Health check and gap analysis. 26

Figure 2.9: Siemens power transformer (Sermsukroongsakul et al., 2019). 31

Figure 2.10: The PTHI layers (Li et al., 2018). 34

Figure 2.11: HI processes (CIGRE Technical Brochure 761, 2019) 36

Figure 2.12: Preliminary FSA (ABB Review 1 Transformer Challenges, 2008). 39

Figure 2.13: Transformer FSA (ABB Review 1 Transformer Challenges, 2008). 40

Figure 2.14: PT failure modes and faults (Murugan and Ramasamy, 2019). 48

Figure 2.15: PT active components (Murugan and Ramasamy, 2019). 49

Figure 2.16: Failure modes, assessment indices and diagnostic result relationship (CIGRE Technical Brochure 761, 2019). 52

Figure 2.17: Development stages of HI calculation method (The Canadian Electricity Association, Deloitte, 2015; Dumitran et al., 2014). 53

Figure 2.18: Maintenance approaches for data management, decision making, and PTHI research asset management philosophy. 57

Figure 2.19: Continuous AM improvement learning cycle (Khamis et al., 2024). 60

Figure 3.1: HI and AM using the Plan, Do, Check and Act (PDCA) concept. 74

Figure 3.2: Health Index methodology (Khamis et al., 2023). 76

Figure 3.3: Data sources for Health Index assessment (Khamis et al.,2024) 78

Figure 3.4: Workflow to achieve the classification of HI Condition (Győrödi et al., 2016). 81

Figure 3.5: Data science research model (The data science puzzle revisited, 2017). 85

Figure 4.1: Research workflow for the conventional mathematical HI model. 91

Figure 4.2: Health Index scoring calculation approach for power transformer (Khamis et al.,2023). 96

Figure 4.3: Priority chart for HI scoring parameters. 97
Figure 4.4: Methodology of the HI (Singh and Swanson, 2018). 98
Figure 4.5: Factors for HI Calculation (Kadim et al.,2018). 99
Figure 4.6: Health index design (Kadim et al.,2018). 99
Figure 4.7: Weighting factors for DGA and OQA (Singh and Swanson, 2018). 100
Figure 4.8: Comparison of calculated PTHI and Kadim HI (2018). 102
Figure 6.1: Main Principles for HI development diagram. 122
Figure 6.2: AI HI development process diagram. 123
Figure 6.3: Research Methodology using AI in R programming language. 123
Figure 6.4: Statistical summary based on R programming analysis. 124
Figure 6.5: The relationship between the individual factors. 124
Figure 6.6: Acidity Data Analysis. 125
Figure 6.7: DCG Data Analysis. 125
Figure 6.8: BDV Data Analysis. 125
Figure 6.9: Furfuraldehyde Data Analysis. 125
Figure 6.10: Binary Logistic Regression model for the transformer HI. 126
Figure 6.11: Artificial Neural Network model design for the transformer HI. 128
Figure 6.12: SVM algorithm method for the transformer HI. 128
Figure 6.13: PTHI uses the Bayesian Theorem based on the NB algorithm. 129
Figure 6.14: Decision Tree (DT) method for the transformer HI. 130
Figure 6.15: Random Forest (RF) method for the transformer HI. 130
Figure 6.16: K-Nearest Neighbours (KNN) for the transformer HI. 131
Figure 6.17: Comparison of ML Models' Accuracy. 133
Figure 6.18: Comparison between seven ML models' sensitivity and specificity rates. 136
Figure 6.19: Binary Logistic Regression model (Zuo et al., 2016). 136
Figure 7.1: The workflow methodology for the actual data investigation. 139
Figure 7.2: PT Age Histogram. 140
Figure 7.3: Number and percentage of transformers by voltage. 142
Figure 7.4: Correlation metric of the electrical utility data. 144
Figure 7.5: Multiple scatterplots of available parameters. 145
Figure 7.6: PT Histogram, Year of Installation. 146
Figure 7.7: The electrical utility data scatter plot showing the age and HI relationship. 147
Figure 7.8: The electrical utility data density plot for age (left) and BDV (right). 147
Figure 7.9: Density plot of the electrical utility HI (left) and its scoring rule (right). 148

Figure 7.10: Comparison between the electrical utility HI and PTHI model. 151
Figure 7.11: Original electrical utility records based on condition. 153
Figure 7.12: Result of Summary function using R programming language. 154
Figure 7.13: SVM model description. 155
Figure 7.14: SVM model description with an accuracy of 86%. 156
Figure 7.15: SVM model with an accuracy of 93%. 156
Figure 7.16: Sensitivity measure for ANN, NB, KNN, SVM, DT and RF. 157
Figure 7.17: ML model's accuracy. 158
Figure 7.18: Specificity measure for ANN, NB, KNN, SVM, DT and RF. 158
Figure 7.19: F1- Score Comparison. 159
Figure 7.20: Statistics of ML algorithms by class (details are in Appendix H Part 3). 160
Figure 7.21: PTHI model decision based on the transformer health conditions. 161
Figure 8.1: Studying transformers ageing correlation with HI. 164
Figure 8.2: Number of faults in years by voltage (2008-2022). 166
Figure 8.3: Failure rate analysis by voltage(left) by location(right). 167
Figure 8.4: Number of transformers classified by age group. 168
Figure 8.5: The PTHI calculation for apparent age T1 and T2 (Li et al., 2020). 169
Figure 8.6: HI first order simple linear regression model. 171
Figure 8.7: Transformers in fair condition. 172
Figure 8.8: HI second-order polynomial regression model curve. 173
Figure 8.9: Linear and quadratic regression model analysis. 174
Figure 8.10: Second-order polynomial regression model curve. 174
Figure 8.11: Cubic regression model analysis. 175
Figure 8.12: Linear regression model summary. 175
Figure 8.13: Quadratic regression model summary. 176
Figure 8.14: Cubic regression model summary. 176
Figure 8.15: Risk matrix based on PTHI condition assessment results. 178
Figure 8.16: PTHI risk matrix for prioritising annual maintenance plan decisions. 179
Figure 8.17: Transformer condition in "Good" or normal using DGA inputs. 180
Figure 8.18: Transformer condition in "Good" or normal range using BDV inputs data. 180
Figure 8.19: Transformer condition as Good. 181
Figure 8.20: Transformer condition in moderate range using DGA inputs. 182
Figure 8.21: Transformer condition in moderated range using BDV inputs. 182
Figure 8.22: Transformer condition as moderate. 183

Figure 8.23: Transformer condition in the Bad range using DGA inputs. 184
Figure 8.24: Transformer condition in the Bad range using BDV inputs. 184
Figure 8.25: The transformer condition is Bad. 185
Figure 8.26: Dataset description for the PTHI web application tool. 186
Figure 8.27: The technical impact of producing the PTHI web application tool. 186
Figure 8.28: Overview of the PTHI web application tool. 187
Figure 8.29: Economic optimum study (The Institute of Asset Management, 2015). 191
Figure 9.1: PTHI practical AM process for assessing transformer lifecycle. 211
Figure F- 1: The process of core and paper insulation damage. 309
Figure F- 2: The process occurring on the damage of paper insulation. 310
Figure F- 3: The process of winding failure that leads to replacing the whole PT. .. 310
Figure F- 4: The damage process due to the tap changer. 311
Figure F- 5: Copper corrosion process that led to membrane expansion. 311
Figure F- 6: The overheating process that led to winding failure. 311
Figure F- 7: Phase to Earth Fault in fluid-filled 11kV Cable Box. 313
Figure F- 8: Degradation and Crystallisation of insulating liquid in 11kV and 33kV Cable Boxes (CIGRE Technical Brochure 625, 2015). 313
Figure F- 9: 33kV Dry Type Cable Box Failure due to PD Activity (CIGRE Technical Brochure 625, 2015). 313
Figure F- 10: Forms of bushing failure modes (CIGRE Technical Brochure 625, 2015). 316
Figure F- 11: Overheating of the above station via application of IR/sensitive camera (CIGRE Technical Brochure 625, 2015). 316
Figure F- 12: Silver Sulphide specimen's development on OLTC acquaintances (CIGRE Technical Brochure 625, 2015). 316
Figure F- 13: Particle and sledge development in water cooler (CIGRE Technical Brochure 625, 2015). 316
Figure F- 14: IR image demonstrating obstructed radiators (CIGRE Technical Brochure 625, 2015). 317
Figure F- 15: Infrared imaging displaying hot spot (U.S. Department of the Interior, 2000). 317
Figure H- 7: BLR Models Result 341
Figure H- 8: ANN Model Results. 342
Figure H- 9: SVM Model Result. 342
Figure H- 10: NB Model Result 343

Figure H- 11: Original Count NPG records based on condition. 343
Figure H- 12: Number of electrical utility records stimulated-based fuzzy rules after adding stimulated data. 344
Figure H- 13: Output of the SVM model showing accuracy. 344
Figure H- 14: Using the DataSetUnder69kVUpdate.csv file, reading through the CSV in Rstudio 345
Figure H- 15: Understanding the structure of the data frame using the str() function 345
Figure H- 16: The above figure shows the first six records of the data frame using the head() function 345
Figure H- 17: Creating a data frame 346
Figure H- 18: Checking the data frame structure and data types. 346
Figure H- 19: Convert the columns to factor so the SVM classifier can handle them. 346
Figure H- 20: Checking the data summary 347
Figure H- 21: Replacing the null value with the column's average 347
Figure H- 22: Checking the data dimensions 347
Figure H- 23: Output of the SVM model showing the accuracy after experiment 2. 348
Figure H- 24: Number of the records per class/condition. 348
Figure H- 25: Numer of electrical utility records with the stimulated. 349
Figure H- 26: Structure of the data frame. 349
Figure H- 27: Summary of the data frame columns 350
Figure H- 28: Dimension of training and testing set. 350
Figure H- 29: Output of the SVM second model showing the accuracy after experiment 3. 350
Figure H- 30: Output of the SVM first model for the accuracy during experiment 4. 351
Figure H- 31: Output of the SVM first model showing the final accuracy during experiment 4. 351
Figure H- 32: Error screen during experiment 4. 351
Figure H- 33: Final SVM model accuracy after experiment 4. 352
Figure H- 34: Illustration of confusion Matrix 353
Figure H- 35: McFadden value of the binary logistic regression model. 355
Figure H- 36: Artificial Neural Network model. 356
Figure H- 37: Screenshot from R studio showing ANN Model Output (1). 357

Figure H- 38: Screenshot from R studio showing ANN Model Output (2). 357
Figure H- 39: Screenshot from R studio showing ANN Model Output (3). 358
Figure H- 40: Screenshot from R studio showing SVM Model Output (1). 359
Figure H- 41: Screenshot from R studio showing SVM Model Output (2). 360
Figure H- 42: Screenshot from R studio showing SVM Model Output (3). 360
Figure H- 43: Screenshot from R studio showing SVM Model Output (4). 361
Figure H- 44: Screenshot from R studio showing RF Model Output (1). 363
Figure H- 45: Screenshot from R studio showing RF Model Output (2). 363
Figure H- 46: Screenshot from R studio showing RF Model Output (3). 364
Figure H- 47: Screenshot from R studio showing RF Model Output (4). 364
Figure H- 48: Screenshot from R studio showing RF Model Output (5). 365
Figure H- 49: Screenshot from R studio showing RF Model Output (6). 365
Figure H- 50: Screenshot from R studio showing RF Model Output (7). 366
Figure H- 51: Screenshot from R studio showing RF Model Output (8). 366
Figure H- 52: Screenshot from R studio showing RF Model Output (9). 367
Figure H- 53: Screenshot from R studio showing KNN Model Output (1) 368
Figure H- 54: Screenshot from R studio showing KNN Model Output (2). 368
Figure H- 55: Screenshot from R studio showing KNN Model Output (3). 369
Figure H- 56: Screenshot from R studio showing NB Model Output (1). 370
Figure H- 57: Screenshot from R studio showing NB Model Output (2). 370
Figure H- 58: Screenshot from R studio showing DT Model Output (1). 372
Figure H- 59: Screenshot from R studio showing DT Model Output (2). 372
Figure H- 60: Screenshot from R studio showing DT Model Output (3). 373
Figure H- 61: Decision Tree Model Plot 373
Figure H- 62: Web Application workflow. 375
Figure H- 63: Screenshot for the PTHI analysis using flex dashboard. 375
Figure H- 64: Number of Transformers by Condition. 376
Figure H- 65: Percentage of Current transformers condition. 376
Figure H- 66: Transformers Moisture Levels. 377
Figure H- 67: Transformer Acidity Levels. 377
Figure H- 68: Breakdown Voltage Levels. 378
Figure H- 69: Data tables showing the electrical utility records. 378

Figure H- 70: PTHI model in shiny web application login screen. 379

Figure H- 71: PTHI model in shiny web application tabs..... 380

Figure H- 72: Health Index web application screenshot. 381

Figure H- 73: Health Index web application screenshot showing acidity and Health Index distribution. 381

Figure H- 74: Health Index web application screenshot showing density distribution of all continuous variables. 382

Figure H- 75: Health Index web application screenshot showing density plot of acidity and health Index. 382

Figure H- 76: Health Index web application screenshot showing a quantile-quantile plot of all continuous variables (part 1). 383

Figure H- 77: Health Index web application screenshot showing a quantile-quantile plot of all continuous variables (Part 2). 383

Figure H- 78: Health Index web application screenshot showing plots of continuous variables by transformer condition (Part 1). 384

Figure H- 79: Health Index web application screenshot showing plots of continuous variables by transformer condition (Part 2). 384

Figure H- 80: Health Index web application screenshot showing overall correlation heatmap. 385

Figure H- 81: Health Index web application screenshot showing bivariate continuous distribution based on transformer condition (Part 1). 385

Figure H- 82: Health Index web application screenshot showing bivariate continuous distribution based on transformer condition (Part 2). 386

Figure H- 83: Health Index web application screenshot showing scatterplots of Health Index columns with other continuous variables (Part 1). 386

Figure H- 84: Health Index web application screenshot showing scatterplots of Health Index columns with other continuous variables (Part 2). 387

Figure H- 85: Studying PT failures using the Age and Health Index correlation. 388

Figure H- 86: PT failure rate using Hazards Function. 389

Figure H- 87: Age grouping of the transformers fleet by HI. 390

Figure H- 88: The regression line according to HI by age for the transformer population shows lower and upper confidence intervals. 391

Figure H- 89: Regression line for calculating the apparent age. 392

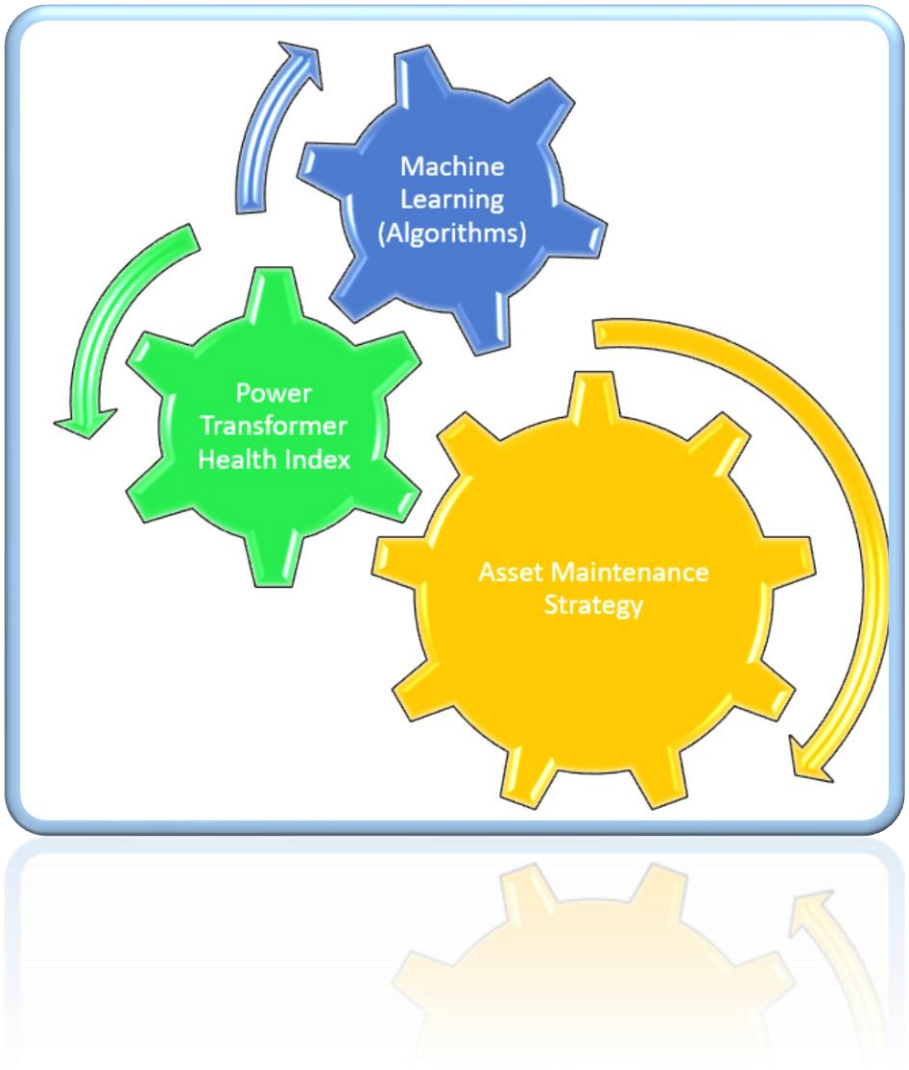
Figure H- 90: Summary of the initial regression model using R programming language. 393

Figure H- 91: initial regression model plot. 393

Figure H- 92: Regression model built for 132kV transformers using R programming language..... 394
Figure H- 93: Regression model plot for 132kV transformers. 394
Figure H- 94: Regression model for 132kV transformers after removing the outliners. 395
Figure H- 95: Regression model for 66kV transformers. 395
Figure H- 96: Regression model utilising the electrical utility HI System. 396
Figure H- 97: Enhanced scatterplot to better understand the model for all voltage. 396
Figure H- 98: Enhanced scatterplot to better understand the model by voltage. 397
Figure H- 99: Transformer HI as a function of age. 397
Figure H- 100: Diagnostic plots for the regression model. 398
Figure H- 101: Residual vs Fitted for linear (left) and second-order polynomial (right). 400

Chapter 1

Introduction



Chapter 1: Introduction

This research, which is of significant importance, focuses on studying, reviewing, and developing total health condition investigation and analysis of small, medium, and large fleets of electrical Power Transformer (PT) using the Asset Management (AM) Health Index (HI) terminology for driving electrical utilities asset maintenance strategy (Soni and Mehta, 2022). The findings of this research would serve as a crucial guide for electrical utilities, Asset Management Strategic Planning (AMSP) decision-making for corrective, preventive short-term annual maintenance planning, and assessing long-term power transformers lifecycle planning.

According to (ISO 55000 series, 2014), the AM strategic role is vital in ensuring efficient and economical operation, maintenance, and asset lifecycles planning. Health index data are complex and are accommodated from different resources (Dong et al., 2019). However, it is also important to note the significant challenges in building a trustworthy continuous monitoring condition model using noisy data for the large fleet of PTs (ABB Review 1 Transformer Challenges, 2008). These challenges include data unavailability and inconsistency, extensive data management and processing challenges, and high dependency on experts' knowledge, which requires a multidisciplinary approach and massive investment to address the issues (ABB Transformer Services (TRES), 2018).

The AM operation and maintenance system investigation is an important field and asset lifecycle (The Institute of Asset Management, 2015) to develop HI that can add the required value during 1) maintenance predictive planning, 2) Operating Expenditures (OPEXs), 3) Capital Expenditures (CAPEX), 4) Total Expenditures (TOTEXs), 5) Strategic Asset Management Planning Decision Making (SAMPDM) which practised by few world-class AM institutions based on the planning for the required significant data portfolio (CIGRE Technical Brochure 706, 2007; Jahromi et al., 2009; CIGRE Technical Brochure 735, 2018). A sound TOTEX system is good when it incorporates actual data for making the right decision that serves the industry and achieves robust investment master planning for the whole network and business based on the AM standard framework (ISO 55000 series, 2014).

The research topic is structured based on the holistic picture of the AM standard framework (ISO 55000 series, 2014), maintenance data, maintenance strategy, ERP AM system experience, best practices of HI application, and IEC and IEEE international standards for data interpretation.

The HI is an asset planning tool (Bertrand, H., 2010). HI considers condition-based maintenance to replace equipment based on action (Allard, L. et al. 2010). The maintenance strategy improves PT performance. In addition, it decreases unnecessary costs (Lorin et al. 2019). According to (Lundgaard L. et al., 2014), accomplishing maintenance strategy further improves PT performance. HI calculations are considered by (Naderian et al.,2008) to evaluate PTHI. Understanding PT's health condition is vital (Saha et al., 2014). PT health index score requires finding the faulty transformer that requires further care than others (Jahromi et al., 2009). The HI will then prioritise the required maintenance programme. The ranking method of the HI averages is quantising against health conditional ranges for different types of parameters indicated in IEC, IEEE, CIGRE, and other electrical industry standards.

According to (Francheck and Woodcock, 2009; IEC 60076-7, 2018; CIGRE Technical Brochure 443, 2010), loading information records based on the performed maintenance history is an important technique that can add HI accuracy. The monthly load data justify PT's health condition for the HI calculation scores. The HI's computation techniques differ depending on the HI developer (Linhell and Lundgaard, 2021) and network operator, mainly based on the power network ageing challenge required to achieve electrical utilities business requirements (Foros and Istad, 2020).

Besides the conventional linear approaches, programming approach and several Artificial Intelligence (AI) methods are explored as Fuzzy Logic (Abu-Elanien et al., 2012; Dorji et al., 2022), Support Vector Machine (SVM) (Yi Cui and Saha, 2013; Rediansyah et al., 2021) and Binary Logistic Regression (BLR) (Zuo et al., 2016) to combine the inspection and diagnostic parameters into HI quantitative score. However, (Naderian et al., 2008; Jahromi et al.,2009; Wang et al.,2022; Jin et al., 2022) stated that the most attractive method is the weighting and scoring factor method (Chakravorti, 2013) due to its simplicity, flexibility, and practicality. This research also considered other methods, including Fuzzy Logic and Artificial Neural Networks (ANN) (Khamis et al., 2023). The resulting HI score based on (Ashkezari et

al. 2012; Ashkezari et al., 2013; Basuki and Suwarno, 2018; Hernanda et al., 2014) can be different ranges within 0 and 1, 0 and 10, or 0 and 100 based on the linguistic expression to define a HI score. It was declared by (Jin et al., 2022) that using distinct PT monitoring input parameters can lead to contradictory health index scores as weighting the input parameters differently and subjectively by the experts influences the health index score (Padmanaban et al., 2023). The HI whole process can be entirely automated for daily practice (Haema and Phadungthin, 2012), and the necessary code to be written in the agreed system using computer language using R statical language, Python, or MATLAB scripts (Heathcote and Arnold, 2016).

In (ISO 18095, 2018), the condition assessment and monitoring automation process objective is to be capable of reading and using asset measurements directly from a health index tool for enabling health condition decision-making for the electrical utility, then adding more of the required assets probability failure for the coming years as the automation process can take more than expected time frame (Li and Dong, 2019). The CIGRE international survey working group (CIGRE Technical Brochure 659, 2016) mentioned that active components are the most significant factor that increases the PT probability of failure of 536 records. The research applied the HI concept to obtain practical and trustworthy condition identification, including challenges in data management. For a successful HI, it is vital to combine actual data, including historical records and inspection data, and integrate it with PT failure modes that could impact PT performance (CIGRE Technical Brochure 659, 2016).

1.1 Recent Development in Power Transformer Health Index Technology

This section presents the findings in the PTHI research area. Appendix C Table C-1 provides a high-level summary of the recent research literature review findings. Chapter two of the literature review offers thorough, deep details and literature reviews.

This strategic section includes critical aspects of data science product, design, and development needs for PTHI that are examined respectfully in the following subsections and more over each of the thesis chapters.

PTHI application recently is considering artificial intelligence and machine learning (ML) technology (Li, S. et al. 2017) for power transformers predictive maintenance

planning. HI predicts the PT condition and probability of failure to operate the mandatory quality of AM service, avoids future failure of PT, and avoids the relevant downtime as it is very costly for electrical companies (Malik et al., 2012). HI is considered the building block of the broader AM process (The Canadian Electricity Association, Deloitte, 2015). This HI development requires actual data updates, especially for critical incidents. Focusing on business objectives leads to the big picture of the AM. Hence, this will build robust preventive and corrective maintenance decision-making for assessing asset lifecycle planning. HI can envisage future failure before its occurrence (Nitu et al., 2017). HI objectives must support daily operation and maintenance switching schedules by including the power transformers that need urgent response and avoid unnecessary maintenance costs (Khamis et al., 2023), as well as extending transformer life by taking the right actions at the right time that keep these units running longer and reduce new investments plan (Murugan and Ramasamy, 2019).

In fact, (Madavan and Balaraman, 2016) raised that PT preventive and corrective maintenance HI should be contingent on accurate statistics by using datasets that reflect the PT health condition to consider the relevant decision (Rausand et al., 2020). Also, this offers perceptions to respond to severe doubts of the maintenance crew and helps them decide to start with cheaper maintenance before thinking of significant replacement or repair based on visual transformer inspection. The core intention of the Artificial Intelligent Health Index (AIHI) is to give a quantitative interpretation of transformers' conditions to the maintenance team using big data. AIHI automates the decision-making course utilising computer science (Gurung et al., 2020). HI shall be incorporated with current systems. This HI needs more investigation and consideration of connecting the analytical tools. Figure 1.1 presents the seven fundamental steps for the Asset Management Power Transformer Health Index Development Process as a technology prototype that can benefit electrical utility.

Figure 1.1: Asset Management PTHI Development Process Khamis et al. (2024).

PTs are essential assets in electrical transmission grid stations (CIGRE Brochure 422, 2010). This asset plays a vital part in the reliability and energy quality of the electrical grid and its delivery to end users. The fact is that these in-operation transformers are getting old and coming to an end of life. Hence, transformer maintenance ought to ensure 100% availability and reliability of in-service transformers for better energy quality. In addition, transformer health conditions and risk assessment (Yuan et al., 2021) are essential to improve the performance of the in-service transformer, expand the transformer end life, and decrease unnecessary costs. Hence, considering optimal decisions over transformer life cycle operation and maintenance management is beneficial. HI meant assessing and categorising transformers' health conditions from critical to normal. This health condition will reflect the required decision-making regarding routine repair and emergency replacement (Martins, 2014). Then, the decisions made in TOTEX for future long-term master planning.

1.2 Rationale Behind the Research

The rational motivation behind this research is to serve industry, society, and the national economy and contribute to research knowledge and articles. Therefore, it requires studying the HI data of local electrical utilities objectively and confidently while assessing a large fleet of power transformers' health conditions. This study contributes to reducing the reinvestment long-term replacement costs through a deferral replacement long-term budget. Moreover, it requires improving the short-term annual maintenance plan and daily operation and maintenance decisions (Khamis et al., 2023). Furthermore, this helps justify spending plans before including the need for any new investment in pre-investment appraisal documents. Then, the required decision will include the long-term master planning investment. In similar practice, the age expectancy of the power transformers enhances transformer performance before failure occurs.

In contrast, unhealthy power transformers recognise it as an issue, and a decision can be considered ahead of failure (Jin et al., 2022). Mainly, based on (Leauprasert et al., 2020), PTHI is a valuable tool for assessing large fleet of transformers and simplifying the identification of power transformers' health, which requires proper and urgent consideration (Pompili and Scatiggio, 2015). A mandatory principle for a useful PTHI model is the availability of proper data to build a reliable model.

1.2.1 Studying the Health Condition of Power Transformers.

This research considers PTHI intended to investigate the data collected from the literature and the electrical utility. This research customises the PTHI using the actual and available datasets collected from the literature and electrical utility based on challenges of small, medium, and big datasets. The rationale is also to discover suitable evaluation techniques for the PTHI model. This evaluation requires reviewing a few existing PTHI models and running various experiments, as shown in the related chapters 6 – 9. Based on these experiments, findings, and the data used, a suitable PTHI model will be proposed. The critical parameters utilised in PTHI evaluation are oil analysis, maintenance history, load history, temperature records, specific design data, and dissolved gas analysis (DGA). According to (Raj, 2019), the PTHI evaluation model handles the available data, which assesses the condition of all transformers.

The PTHI evaluation model must reflect the PT deterioration process appropriately and clearly. It should also recognise the ageing mechanisms and the possible PT failure modes.

Also, the proposed PTHI model performance uses different KPIs or measures, including accuracy, sensitivity, specificity, and F1 score. These measures are experimented on several datasets based on this research literature review (Arshad et al., 2021; Krishnamoorthi et al., 2022; Khamis et al., 2023). The outcomes must demonstrate the PTHI model's capability to differentiate the diverse health conditions of power transformers. The PTHI value must be accurate and consistent, indicate the actual individual transformers' condition, and provide the required action for further decision-making.

1.2.2 Utilisation of the Field Electrical Engineering Experience and Asset Management Knowledge

The research uses the students' electrical engineering work experience in the electrical industry, generation, transmission, and distribution. It also includes strategic asset management (SAM) knowledge and international best practices based on ISO55000 and ITOMS. This interest is to study how to integrate technical experience with the strategic roles of the AM in a novel research model investigation that includes five case studies. The PTHI requires considering the AM standard “ISO55000” framework as a strategic role towards developing the PTHI as a condition evaluation method as a daily technical role of the AM that meets the annual maintenance plan for driving the electrical utility maintenance strategy.

1.2.3 Integrate a Relationship Between the PTHI and the AM Framework

Figure 1.2 shows the critical relationship between the main clauses (1-7) of Asset Management (AM) ISO 55000 series (2014) and PAS55 (2008) standards. This knowledge and practical experience considered in formatting the PTHI, which considers capital investment decision-making, Operation and Maintenance (O&M) decision-making, and lifecycle value realisation and avoid shutdowns due to unplanned outages of faulty transformers in a large fleet that can lead to total or partial blackout (The Institute of Asset Management, 2015). In SAM, it is difficult to examine

each matter freestanding, as there are complex interrelationships between most of them (Tang and Wu, 2011).

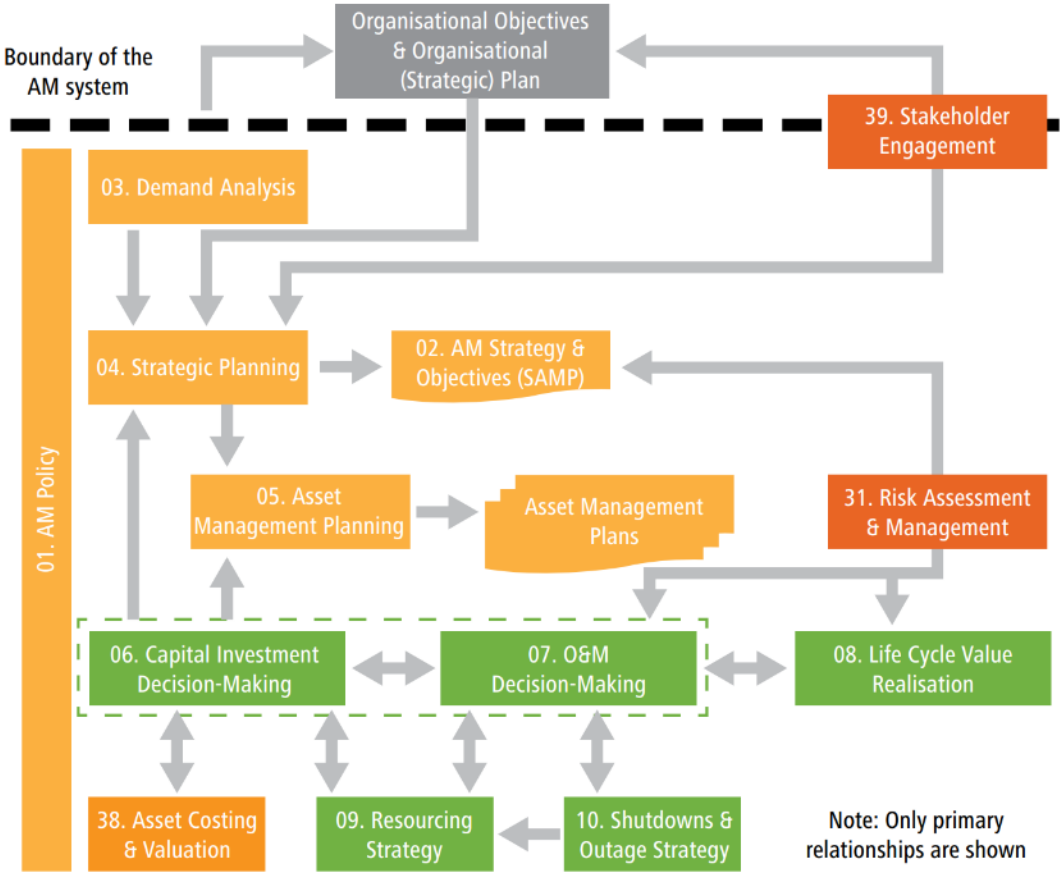


Figure 1.2: Asset management strategic planning subject group (The Institute of Asset Management, 2015).

The strong interaction and philosophy shown in (The Institute of Asset Management, 2015) among AM strategic planning subject groups and operation and maintenance (O and M) decision-making include PTHI since PTHI is implemented based on asset maintenance strategy.

The PTHI ensures that the right decisions on maintenance planning are considered (Rediansyah et al., 2021) and that the long-term investment plan identifies technical consequences and justifications. ISO 18095 (2018) stated that electrical utility regulators are requesting the submission of a structured annual budget explanation from electrical utilities that meets the required quality level using PT monitoring

programmes. This action requires assessing asset lifecycle planning observations on crucial investments, maintenance planning activities and system reliability reinforcement. The regulators also require pre- and post-investment appraisal for each PT or asset health.

These are essential inputs to AM strategy and planning, which keeps engaging the role of each stakeholder, managing risks, evaluating health conditions, and examining individual condition classes. The concept of the ISO 55000 series (2014) framework stated that it is a challenge to prove how the PTHI is an intelligent tool for making efficient maintenance decisions and how this can support the holistic picture of AM implementation.

1.2.4 Authority Electricity Regulators (AER) Challenges to Electrical Utilities

Technical Authority Electricity Regulator (AER) of electrical utilities are challenging and pressurising the network operators by requesting a comprehensive investigation and explanation for a long-term financial plan provided by electrical utilities every year, three years, five years, seven years based on related subsidies whether its generation, transmission, or distribution. AER is demanding evidence about each single asset's health condition. However, AER requires a long-term plan in the form of a Master Plan (MP), including the required maintenance actions and network reliability. Electrical utilities play a tremendous role in achieving the technical and regulatory requirements and securing the required power system expansion and reinforcement funding. Health Indices (HI) is an intelligent, practical, and auditable tool that can assist electrical utilities and offer a complete technical and financial analysis and justifications that explain the reasons behind any penny for short, medium, or long-term funding requirements (Khamis et al., 2023; Foros and Istad, 2020; Lundgaard L. et al., 2014). Hence, electric utilities can capitalise on weighty expanses of time and funding approval and archive data elements based on the results of PTHI. Using the relevant data elements in Asset HI assessment and AM activities places the data in a decent practice position. It enhances a valuable Return on Investment (ROI) for the early electrical system charges.

1.2.5 Critical Role of Data Asset Management

According to Khamis et al. (2023), world-class electrical utilities protect their staffing explicit and tacit information, which is considered one of the asset types based on (ISO 55000 series, 2014). This research addressed the importance of integrating PTHI with the big picture of the AM standard framework to break the silo among different technical departments within electrical utilities to improve the health index actual data, enable the right decision, and protect both explicit and tacit information and knowledge. At any time, an ageing staff, in particular, means tacit knowledge can leave the company. Khamis et al. (2024) stated that PTHI works with operating data-driven decisions to include expert knowledge before departing the company. The ageing and loss of technical experts are leading to a growing quantity of "emergency" alarms. This increase in alarms leads to more time and workforce activities, which means additional costs that are difficult to control. In this research, operating the electrical network on proactive or predictive maintenance planning allows for restoration of the cost, time, and human resources, which will manage and decrease the total expenses for operating the power system.

1.3 Problem Statement

Health condition assessment techniques are decision solution tools, especially for large fleets of power transformers. CIGRE, Technical Brochure 630 (2015) confirms that electrical utilities invest in developing a good PTHI as a monitoring system for managing asset needs.

Power transformers by Tenbohlen et al. (2016) are considered critical assets for the transmission grid and considered one of the costliest equipment. Failures of the strategic power transformers based on (CIGRE Technical Brochure 248, 2004) can lead to total area blackout and significant economic impact. Evaluating the operated and non-operated PT health condition without a PT specialist is vital (CIGRE Technical Brochure 323, 2007; CIGRE Technical Brochure 436, 2010; CIGRE Technical Brochure 494, 2012). HI evaluation enables the maintenance and reinvestment within the required time frame.

PTHI must be an intelligent tool that can assess and process in-service PTHI using actual condition data, followed by data interpretation using relative importance scores

and weight factors that explain the total PT health condition. The development of PTHI requires thorough knowledge and a power transformer's technical speciality. Especially in representing PT ageing mechanisms and the likely deterioration reasons behind its failure. AI is an increasingly popular tool for confirming a correct PT HI evaluation and utilising varieties of methods to control and evaluate its health condition using actual data of diagnostic tests for oil, DGA, frequency response, partial discharge, and degree polarisation (Saha and Purkait, 2004; Shintemirov et al.,2010; Wang et al.,2022).

1.4 Research Questions

The research sought to answer the following question: How can the health index identify PT health conditions for electrical utility driving annual maintenance plan and PT lifecycle planning?

Which will fulfilled by answering the below sub-questions:

1. What are the active parts of the power transformer that impact the power transformer's health condition?
2. What data and which parameters are vital to study power transformer health conditions using the health index study, and how is the analysis processed?
3. How to Develop A PTHI Conventional Model Using the Identified Parameters?
4. What ML algorithms to develop PTHI models and asset management decision-making tools?
5. How can the PTHI condition be used to define the right asset management decision?

1.5 Thesis Format

The research thesis contains nine chapters, which are stated in the research milestone plan in Appendix B and described briefly below:

- Chapter One introduces the research topic ‘Electrical Power Transformer Asset Management Health Index Development’ with a general overview that stresses the research rationale, problem statement, and research questions. It also includes a list of the thesis format.

- Chapter two includes a literature review and details recent findings from other articles and documents in this research area of PTHI asset management driving asset maintenance strategy, annual maintenance plan, and asset lifecycle planning for electrical utilities. It also focuses on PT system understanding. It discusses relevant main active parts, the overall effect of transformer components, and operating performance issues in electrical utility energy management and power supply systems. It also focuses on bringing and discussing together several PT condition assessment methods.
- Chapter three contains the research methodology, aim, objectives, method and framework, and data collection process.
- Chapter four presents a critical analytical review and discussion for comparing HI conventional scoring results with the research HI results using similar small data from authors' literature reviews utilising the case study approach. This analysis uses an Excel spreadsheet that includes three PTHI models.
- Chapter five includes a preliminary study experimenting with developing fuzzy logic HI simulation in MATLAB to study the available membership functions. The analysis categorises the five PTHI classes into “Very Good”, “Good”, “Moderate”, “Bad”, and “Very Bad”. Health condition classes integrate with PT conditions in the following chapters for developing AI and ML PTHI models and the HI web app tool.
- Chapter six presents literature data analysis using AI PTHI for the collected data from the literature reviews. This work helped to decide which AI algorithms are convenient for extensive data analysis in the next chapter. This work includes seven models based on different ML algorithms in R studio for medium data of 30 PT records collected from literature reviews.
- Chapter seven presents a detailed analysis of electrical utility data, including EDA and ML data modelling process, including analysis of electrical utility transformers data using the flex dashboards, implementation of ML model, and then discusses the deployment of the SVM ML model in PTHI Shiny web application as a tool that helps in short term maintenance decision-making based on daily operation and maintenance engineers switching schedules of the main annual maintenance plan. This assessment also serves the asset lifecycle of long-term robust master planning, which can benefit long-term budget planning and a detailed PTHI case study analysis. This work includes six ML models for the electrical utility data.

- Chapter eight studies the correlation between the health index value and PT's remaining life. This chapter also presents the hazard PTHI approach, PTHI variation with age, calculation of failure modes based on operation and maintenance records, calculation of failure rate based on apparent age, and a detailed case study. These works include two regression models, one in Excel spreadsheet and the other in R programming language. The findings present the correlation between age and PTHI. The PTHI investigation concluded using a risk matrix decision-base model that considers the right decision and avoids unnecessary spending.
- Chapter nine includes the research conclusion and recommendations.
- Appendix A lists the research publications.
- Appendix B shows the research milestones.
- Appendix C shows a high level of recent development in PTHI that focuses on using AI, ML, and algorithm coding to support short-term daily maintenance activities and issues.
- Appendix D presents data management issues and processes.
- Appendix E explains the power transformer's health index formulation.
- Appendix F explains the power transformer's failure modes and degradation mechanisms.
- Appendix G includes the scoring assessment matrixes.
- Appendix H includes PTHI modelling and simulation results.

This research includes three conventional HI models in an Excel spreadsheet as the primary platform for the first analysis. It also covers a literature review of small, medium, and big actual data records collected from the electrical utility.

The research covers eighteen models considered research strength, and the PTHI design creativity approach investigated the electrical PT health condition analysis.

1.6 Summary

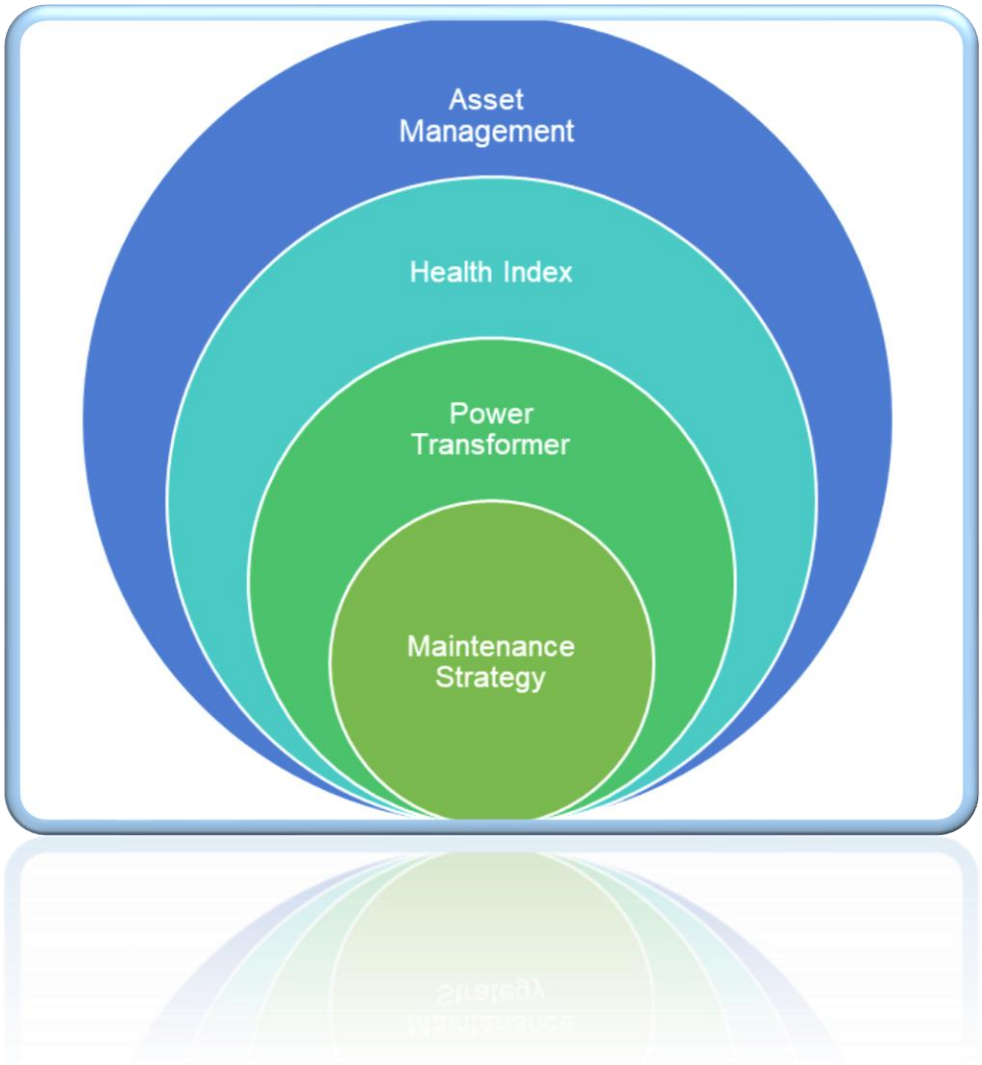
This research uses the term “PTHI,” which employs the same terminology for the health index employed by researchers in the reference papers and applied in this research investigation, including electrical utility and industry practice. All researchers agree upon the fundamental principles and usefulness of the PTHI despite the facts of various analysis and modelling techniques.

This research studies, examines, and develops the PTHI tool as an essential block of the overall AM process. HI forms a building block to the broader ISO55000 AM standard framework development. PTHI in this research is a PT condition evaluation method that identifies the risky PT that needs more attention than others. This need is by prioritising the decision-making based on the new health condition of each PT unit using the scoring results and criticality across PT classes or individual power transformers. The scoring results rating range from 0 to 100, where 100 describes power transformers in “Very Good” new conditions, and 0 describes power transformers in “Very Bad” health conditions that are about to fail and need attention before failure occurs. Hence, a quick response by a practical decision for intervention and urgent action: how soon and what actions are required. This PTHI work can be regular maintenance, repair, refurbishment, or replacement to keep running the momentum of the asset lifecycle and future continuous improvement.

The next chapter will discuss recent findings from journal articles, documents, and white papers published about this PTHI research area. It also focuses on PT system understanding. It discusses relevant main active parts and the overall effect of transformer components on operating performance for electrical utility energy management and power supply systems monitoring. It also discusses several PT condition assessment methods and presents research PTHI approaches.

Chapter 2

Literature Review



Chapter 2: Literature Review

2.1 Introduction

Chapter 1 presented the thesis topic, discussed the latest advancements in power transformer HI, and the reasons for conducting this research. The research topic focuses on three significant areas: asset management practice, power transformer systems, and electrical utilities' health index driving maintenance strategy. These fundamental principles are integrated in-depth and directly into the research questions.

This research considers both the theory and practice of PTHI development. It looks at recent health condition assessments and systems developments to understand the learning lessons, especially for missing data and data uncertainty (Prasojo et al., 2021). This research considers examples of PT health condition assessments, focusing on the failure modes (CIGRE Technical Brochure 761, 2019). The PTHI results must indicate the required action for the health condition (Qiu et al., 2015).

Figure 2.1 presents the research focus on three areas that required more attention during the literature review analysis. The proposed PTHI design is for only power transformer assets; the related data must be collected only for this asset. The PTHI design meets the objectives of asset management, power transformers management, and health indexes for driving maintenance strategies for electrical utilities.

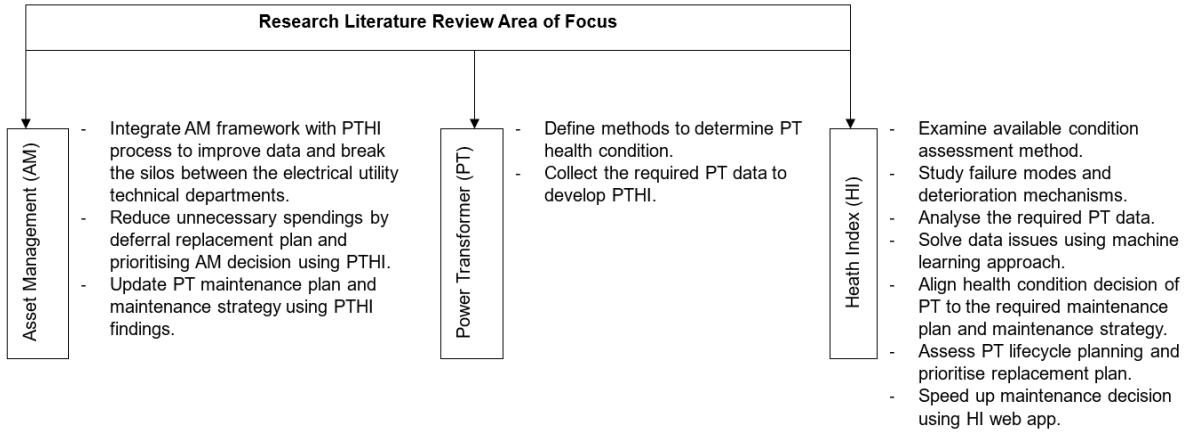


Figure 2.1: Research Framework.

Electrical utilities are keen to control future funding and transformer ageing challenges while introducing international best practices, AM concepts and standards, and AM culture. All are important in serving and linking together the organisation's

strategic goals. Hence, it became essential to implement a good AM, as most electrical utilities work by complying with technical regulations. This monitoring philosophy is part of electrical utility master planning, government obligations, and price control funding needs.

However, in recent years, the evaluation of the health condition of electrical assets, including transformers, has gained significant attention with the development of various techniques and methods (Bigdeli and Aghajanloo, 2016). Accurately assessing PT reliability and effectiveness is essential for the electricity supply chain. This assessment allows informed decisions for PT operation and maintenance practice and ensures the decision is performed optimally and efficiently (Dong et al., 2019).

This chapter examines the existing literature, manufacturing technical documents, and study group documents published about this research area. It will also explore the methodologies that have raised concerns in these journals on asset management, power transformers, health index development, and their implications for maintenance strategies in electrical utilities. The literature review chapter considers the following structure:

Section 2.2 will discuss the international best practices of asset management, which link to in-service power transformers as important physical assets that require frequent attention during assessing asset lifecycle planning and prioritising replacement plans. This action enables the necessary efficiency, effectiveness, cost-optimal decision-making, and management of harmful risks before transformer failure occurs, leading to system failure and unnecessary spending.

Section 2.3 will include the practical and technical details to enable a good understanding and comprehensive information of different PT active components that can be the source of failure modes and deterioration mechanisms expected to occur for in-service power transformers. This understanding is necessary to gain the required knowledge and learn about power transformers' active components before further studying individual failure modes and deterioration mechanisms as part of the development of PTHI.

Section 2.4 will examine PTHI development, including the concept and purpose of condition assessment. It will also review the existing condition assessments and the implemented approaches, including the scoring ranking method, fuzzy logic, and machine learning approaches.

Section 2.5 will detail the required process and health condition indicators, failure modes, and mechanisms that reflect PT's health condition.

Section 2.6 will discuss the approaches implemented in this research for developing the PTHI models: the conventional health index approach, fuzzy logic and artificial intelligence approaches.

Section 2.7 will summarise the key findings from the literature review, address the knowledge gap, and link it to this research aim and methodology in the next chapter.

2.2 Asset Management Practice

This section comprises an overview of asset management practices and standards, as well as the purpose and benefits of AM. It then goes through and explains the AM framework. Then, it discusses the Asset condition assessment process. This section concludes with a discussion of the need for change in the AM process, relying on the conclusion from the literature reviews.

According to Hastings (2021), asset management is a recognised practice and requirement due to its complicated technical nature in the current PT system. Hastings (2021) also pointed out in his book that over the past decade, those working on asset management have made a concerted effort to build a formal approach to asset management systems, knowledge, and education. A vital component of these efforts is the release of standards.

AM standards are ISO 55001, 2014 series and PAS55 (2008). AM became well known in the 1980s when the AM physical asset concept became an expensive asset in different governments and private sectors in Western European countries (Hughes, 2003). Asset management standards (PAS55, 2008; ISO 55001, 2014) define the concept of AM as methodical and matched actions and performs over which a company optimally and sustainably achieves its resources and asset systems, their

related performance, hazards and expenses over their lifespans to reach its company tactical plan.

PAS55, 2008 defines five categories of assets. These categories include data, physical, intangible, cash, and human assets, as depicted in Figure 2.2.

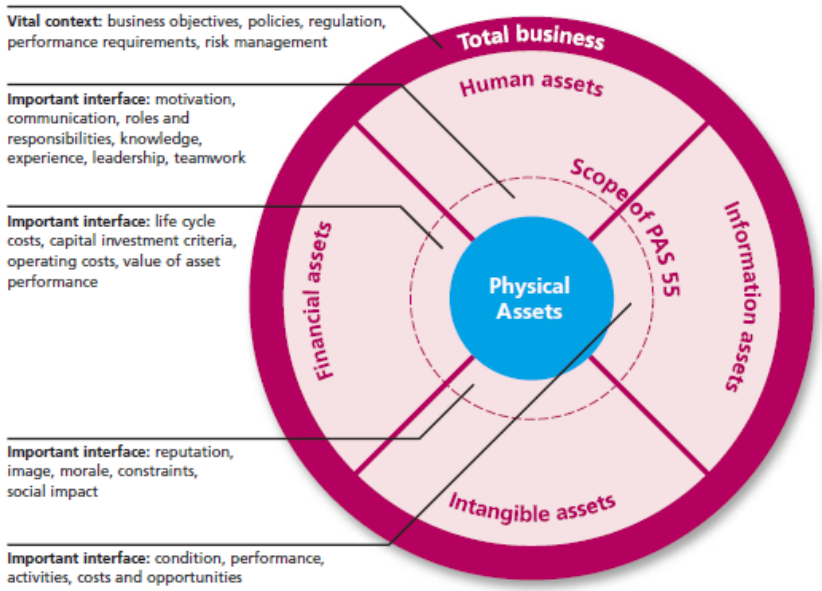


Figure 2.2: Categories of assets based on AM standard (PAS55, 2008).

The PAS55 (2008) has been improved and became the international standard ISO55000 series (2014). ISO55000 series (2014) specifies the big picture of an AM system, as indicated in Figure 2.3. However, there are no details about how to outline the system. This finding can only detailed by experts in the area of AM and field experience (The Institute of Asset Management, 2015).

ISO55000 series (2014) details AM implementation journey requirements. Electrical utilities are known as good AM utilities once the company has fully complied with all PAS55 (2008) / ISO55000 series (2014) clauses and criteria and is certified as PAS55 (2008) / ISO55000 series (2014). This AM standard for this research considers the following definitions.

- Asset: is considered any priced valuable item.
- Health: it describes the asset condition required to achieve a certain level of service for the required period.
- Index: is a sign of the health condition of the asset to be addressed.

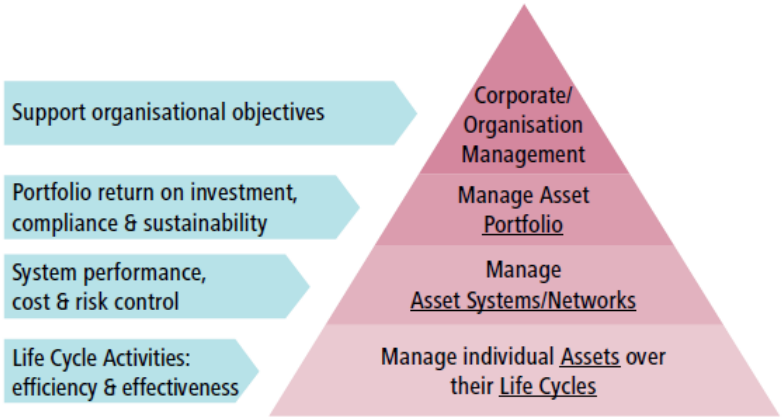


Figure 2.3: Asset management system hierarchy (PAS55, 2008).

Asset management bridges the gap between technology and business. The asset manager's responsibility is to bring together a combination of technical and business knowledge to successfully and efficiently address the asset-related needs of the overall organisation (Hastings, 2021).

ISO55000 series (2014) states that the primary purpose of AM is to derive value from assets for the organisation and its stakeholders (Almeida et al., 2022). AM framework based on AM standards covers the whole organisation, “everyone and everything”, with value for money. Efforts are required from all institutions either to reach at least the minimum practice and awareness of the value of practising best practices of AM or to reach the maximum to sustain with continuous improvement of AM culture, which is hardly complex and costly over the short term (Psarommatis et al., 2020).

The Institute of Asset Management (2015), as illustrated in Figure 2.4, states that asset condition assessment encompasses many procedures, including monitoring, analysis, and evaluation. It is thus a continuous and complex lifecycle process that necessitates good data transformation into helpful information via higher quality accurate data to evaluate in-service PT performance, manage risks, increase transformer end life, reduce unnecessary spending during lifecycle delivery, and enable effective optimal investment decisions (Jahromi et al.,2009; Algudsi and El-Hag, 2019; Yuan et al., 2021).

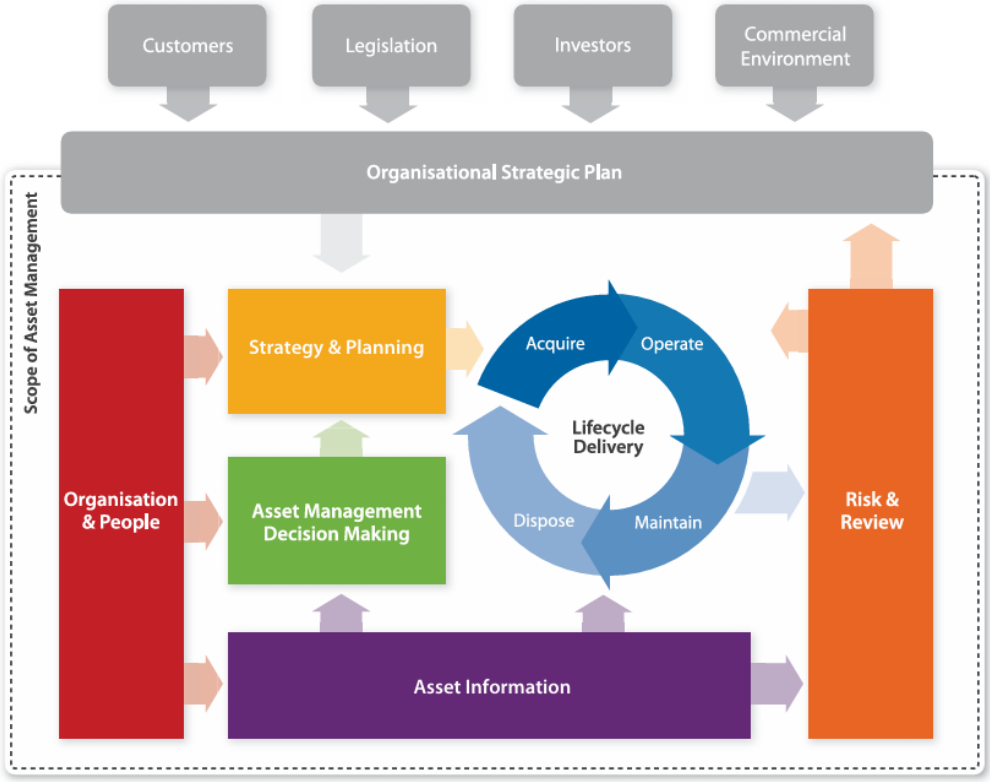


Figure 2.4: AM system (The Institute of Asset Management, 2015).

AM benefits assist the electrical utility or any organisation to gain the required return on investment and achieve the organisational objectives, which require considering risk, performance, and cost to operate the company at a high level of service and less cost (ISO 55001/2, 2014).

The Institute of AM (IAM) is a global consortium that is not-for-profit or academic. It provides tools to progress the AM journey. The IAM was established in 1994 and has over 22,000 members from 158 countries (The Institute of Asset Management, 2015). Recently, various chapters from the IAM have been representing the IAM in different regions. The IAM provides AM professionals and others interested in the area with additional knowledge and best practices for managing infrastructure assets.

Asset management plays a critical role in enhancing data asset management and breaking down data silos in practice and academia (Khamis et al., 2015). These findings can create a valuable and auditable health index for power transformer assets (Heywood and McGrail, 2016). Furthermore, (The Canadian Electricity Association, Deloitte, 2015) stated that the asset health index is a building block in the big picture of asset management. HI is a condition assessment tool requiring actual and correct

data to enable good decision-making in asset maintenance prioritisation and assessing the asset lifecycle of power transformers (Abbasi and Malik, 2016). The ISO55001 and PASS55 Asset Management standards define the lifecycle of a physical asset as project, operation, maintenance, and disposal, as shown in Figure 2.5. HI condition assessment tool considers these asset lifecycle phases during planning and considerations for the required decision that needs accountability, especially since implementing HI decision-making conclusions is very costly (The Institute of Asset Management, 2015). Optimal equipment asset management must consider reducing the life cycle costs associated with these tasks while ensuring that building performance levels such as condition, serviceability, safety, and capacity are fulfilled (Grussing,2014).

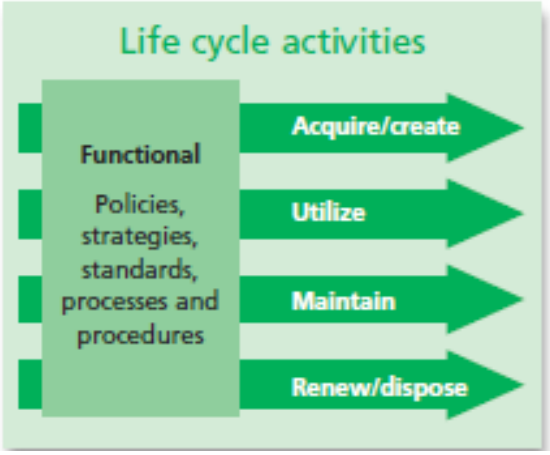


Figure 2.5: AM lifecycle activities (PAS55, 2008).

While the international ISO 55000 series of asset management standards have been successfully adopted by various organisations (Alsyouf et al., 2018), Diop et al. (2021) emphasise that electrical utilities must manage their replacement plan before the expiry lifecycle, become obsolete as a result of technological advancements, or transition to more efficient and carbon-free power alternatives.". Khamis et al. (2023) concluded the importance of the development asset management frame for the PTHI.

The overall electrical power system performance efficiency for in-service power transformers will be improved, reducing the enormous costs of power transformers and managing unnecessary risks during peak loads (Khamis et al., 2014). Petchrompo and Parlikad (2019) recommend utilising data-driven ways to reduce subjectivity in fleet and portfolio asset management decision-making, which this research considers

data-driven decision PTHI model as further development in the research area. Furthermore, Khamis et al. (2023) introduced the AM standard ISO 55000 framework as a best practice for implementing PTHI data-driven decision-making.

To conclude, it appears, based on the literature review analysis a gap in the data-driven approach is the need for investigation of change in asset management processes which will support breaking the silo among the technical departments and enable a good development of the PTHI tool.

Figure 2.6 stresses the interconnection of these research areas in driving effective asset management practice and the development of a robust PTHI tool that drives electrical utility annual maintenance plan and maintenance strategy.

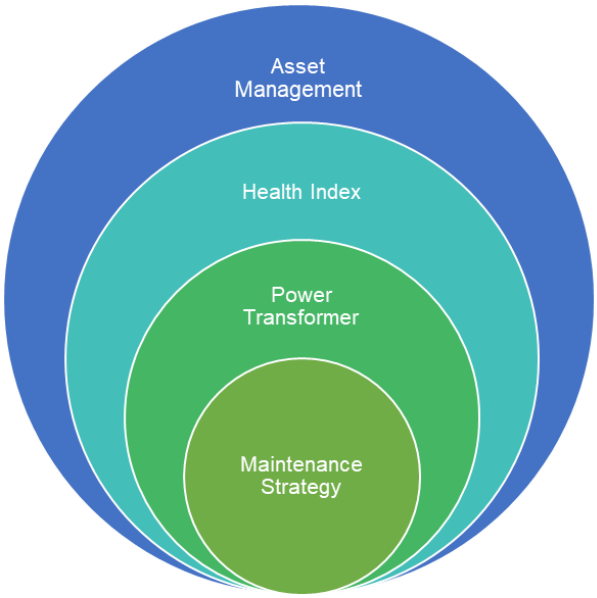


Figure 2.6: Research framework for PTHI development.

2.2.1 Asset Management Process for Power Transformers

Electrical energy remains a precious resource in humankind's growth, which needs prompt industrial expansion (The Canadian Electricity Association, Deloitte, 2015). It is necessary to guarantee reliable and efficient electrical energy for end-users, as the energy deliverable must be of the best quality and low cost since AM is about quality of service by managing risk and cost and improving system reliability and performance. Moreover, it is necessary to guarantee continual, reliable, and efficient power supply to end customers (CIGRE Technical Brochure 642 Transformer

Reliability Survey, 2015). The Canadian Electricity Association, Deloitte (2015) declared that global government bodies own electrical utility systems with profits of over 60% of scattered assets of over billions ($115) from the value of whole assets, as shown in Figure 2.7. PT forms part of the electrical system's backbone, classifying these assets as the most expensive and reliable. Therefore, AM must ensure the in-service PTs are running without failure to secure the power supply to end users. This necessity requires frequent online monitoring, including using the SCADA system and offline health condition analysis, including HI (Jin et al., 2022).

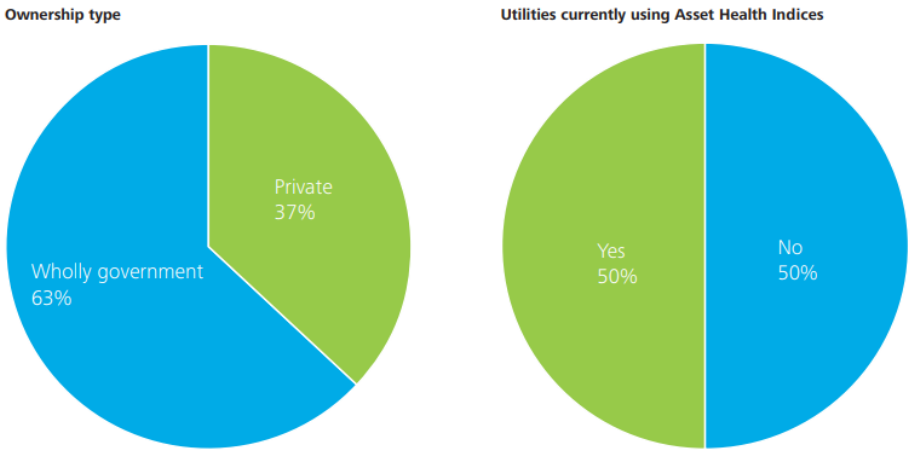


Figure 2.7: Ownership and electrical utilities use asset health indices (The Canadian Electricity Association, Deloitte, 2015).

In addition to the above findings, Abbasi and Malik (2016) stated that there is a need to change the asset management process from the perspective of power transformers. The process involves many activities to guarantee transformers' optimal performance and long life. Foros and Istad (2020) and Khamis et al. (2023) declared three main areas to focus on maintenance strategies: corrective, preventive, and predictive maintenance. Power transformer condition monitoring incorporates multiple methods, including Dissolved Gas Analysis (DGA), which provides critical data on transformer health and is essential for developing health indices (Basuki and Suwarno, 2018). In addition, risk management is vital to understanding failure modes and failure impact and what maintenance strategies should be used to reduce these risks (CIGRE Technical Brochure 761, 2019).

The Electrical Power Research Institute (EPRI) survey and research by (Kurtz, 2005; Abbasi and Malik, 2016) 100% assumed the requirement is to focus on the asset management process, which is strategically accurate as considered in this research to break the silo for the health index and asset management since the health index is a building block of the overall big picture of the asset management (Deloitte, 2015). Hence, the novelty of the research is in integrating the health index with the asset management standard framework (ISO 55000 series, 2014). This novelty will be a practical key that serves the overall reason for long-term budgeting planning for assessing PT lifecycle planning.

This investigation summarises that according to AM standards (PASS 55,2008; ISO 55000 series, 2014), the need to conduct a Health Check and Gap Analysis (HCGA) by the electrical utility due to its importance in defining the required AM building process for developing the needed PTHI.

Figure 2.8 shows this research's initial thoughts and proposal. The diagram is self-explanatory and can be considered and used by electrical utility companies.

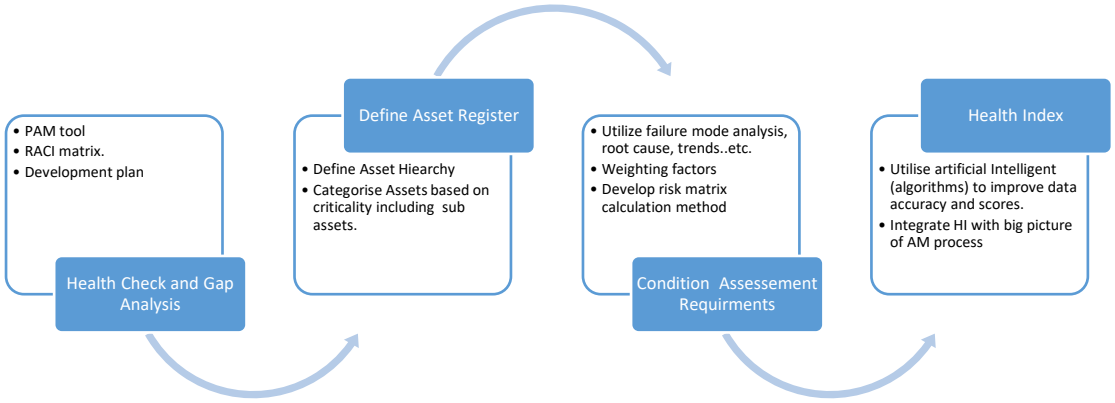


Figure 2.8: Health check and gap analysis.

ISO standards for Asset Management (AM) and Risk Management (RM) are ISO55001 and ISO30001. These standards offer the implementation criteria that benefit electrical utilities in day-to-day operations or business context activities (Gaha et al., 2021; Bertling et al., 2005). The electrical utilities are supposed to cascade to the standards to implement operation and maintenance procedures.

Best practices are AM electrical utilities that utilise the line-of-sight theory to drive the big picture of organisational vision, mission, and objectives into working policy,

strategy, plan and daily procedures (ISO 55001, 2014). Asset Management, an Anatomy (2015) confirms that the AM's holistic role connects organisational network strategic planning with operational and maintenance technical field activities. Likewise, the decision-making procedure maintains the physical equilibrium between numerous participating activities, issues, risks, costs, performance, business continuity, objectives, and benefits (PAS 55 1/2,2008). The risk management framework states that risks can be opportunities for PT health condition judgement, and positive risks are opportunities for finding faulty transformers and further development in PTHI (Scatiggio et al., 2016).

Furthermore, electrical utilities tackle these disputes, utilising a range of simulations and applications that utilise various spreadsheet models and other condition assessment tools to identify PT health conditions based on various inspections and tests (EA Technology, 2018; DNV, 2013). Vermeer et al. (2016) agreed that PTHI improves lowering the ambiguities and quantifies risks in an improved methodology for electrical PT AM decision-making means. Usually, these simulations (Patel et al., 2021) and applications are developed reliant on conventional techniques (Jahromi et al., 2009), which indicates limited efficiency in managing data complications and ambiguities (Picher et al., 2014). Inadequate efforts by Heywood and McGrail (2016), which illustrates and detail the way of the PTHI development to a level illustrates simultaneously information and insights acquired from measurable computational simulations to the decision maker's requirements, which widely discussed in various CIGRE technical brochures, including CIGRE Technical Brochure 343 (2008) and Technical Brochure 576 (2014).

Besides, the influence of new unquantifiable aspects, which can hardly be quantifiable as municipal insight, diplomatic influence and organisation repute, might hardly turn into prominent decisions (ABB Brochure 2750, 2014). However, managing data uncertainty is a highly difficult and complicated technical ecosystem to adequately describe a method for planning a robust PTHI application which counts on international standards and best practices (CIGRE Technical Brochure 761, 2019). PTHI continues as a structural part of the overall AM (Duran Tovar et al., 2015) that serves electrical utilities to manage the electrical power transmission grid at high service and low cost.

PTHI does not consider the AM framework and does not widely report using AM as a vital factor for improving HI (Khamis et al.,2023). This research integrates the technical field understanding, besides the asset management standard ISO55000 framework, to obtain the required PTHI model. This AM is to add the practical concept to the HI as a building block of the big picture of the asset management process that covers strategic planning asset management. It also includes daily operation, maintenance, and investment master planning roles that require robust dataset validation.

2.2.2 Asset Management Global Data Problems for PTHI Development

Although implementing the PTHI technique in transformer maintenance provides significant benefits, several challenges still need to be addressed. Currently, global electrical utilities offer no reliable public data for the HI study, including at least DGA, Oil Quality (OQ), and Furans test results (Khamis et al.,2023). The availability of this data will enable further research investigation using PTHI (Kinnunen et al., 2016). Unfortunately, the published papers' literature reviews share high-level diagnostic test records (Zuo et al., 2016). The direct and indirect measurement tests are costly (Islam et al., 2017), require dedicated manpower experience, are time-consuming, specialist, and shall be conducted periodically based on PT manufacture guidance (Singh and Swanson, 2018). A literature review published by Abu-Elanien et al. (2012) utilised accurate data from Middle East Electrical Utility for 90 oil transformers, of which the literature shares only 30 records. Kadim et al. (2018) shared 15 total records of transformers presented in the literature. The research applied two data groups; the first data is 33 kV transformers data provided by the electrical utility, whereas the second data was stimulated data obtained from Monte Carlo Simulation (MCS). This data relies on a repetitive arbitrary assortment approach as practical for complex statistical measurement. The insufficient data volume requires stimulating the numerical values data to an acceptable level to achieve the required investigation. At the same time, Zuo et al. (2016) applied thirty transformers' data from another literature review (Abu-Elanien et al., 2012). Also, Islam et al. (2018) gathered the Australian electrical utility dataset for the PTHI, which did not share any dataset. In contrast, Singh and Swanson (2018) shared three records in the case study demonstration. The data management process is considered a challenge in research

by (Khamis et al.,2023), which is an essential process for academia and electrical utilities.

Still, no reliable dataset reveals the findings for sufficient analysis within the literature review papers. Still, data volume remains a concern (Jin et al., 2022) for ML investigation. The data can be stimulated to an acceptable level to design the projection simulation. The input parameters and scoring study are built-in statistical values, and alphabetic and colour coding is enhanced (CIGRE Technical Brochure 761, 2019). Commonly, electrical utilities appoint a third party to perform PTHI (DNV, 2013) and present the results statement to the concerned department to enable accurate future decisions and include the required budgeting plan before considering transformer maintenance, replacement or repair into the overall master investment master plan (Dong et al., 2019). In this research, PTHI is a tool for PT specialists and engineers and proposes the correct decision for short-term maintenance switching schedules and assessing PT lifecycle planning. PTHI can add more value using AI simulation to conduct the needed health condition assessment (Leauprasert et al., 2020; Abu-Elanien et al., 2011). The necessity of analysis and design for developing the PTHI tool is expressed based on the recent reviews and findings.

Power transformers are the most critical asset in grid stations. Transformers are vital in electrical systems' reliability and energy quality (Martins, 2014). Once in-service PT gets old, more care is vital for available and reliable PT analysis (CIGRE Technical Brochure 755, 2019). Transformer risk and condition evaluation based on technical standards (CIGRE Technical Brochure 625, 2015) are critical to improving performance, expanding operational life, and optimising investments considering robust solutions during the PT lifecycle (Lundgaard et al., 2014). Also, PTHI is recognised as a solution to evaluate transformers based on existing conditions and support the required cost-effective decision for maintenance or replacement (Sun et al., 2016). Khamis et al. (2023) stated the importance of the data management process and having a master dataset as an asset register for the proposed PTHI model based on the asset management standard framework (ISO 55000 series, 2014).

2.3 Power Transformer System

This section discusses and provides comprehensive information about the PT unit, considering this as the problematic health condition investigation for this asset management HI development research.

Power transformers are considered a strategic, vital and costly asset in electrical networks (Zeinoddini-Meymand et al., 2024; Kadim et al., 2018; The Canadian Electricity Association, Deloitte, 2015). These transformers perform a weighty task in the electrical system: generation, transmission, and distribution of electricity networks (Metwally, 2011).

The unit is expensive in terms of its design specifications, transportation, installation, operational and maintenance complexity. It is also costly if in-service PT fails and time-consuming if it is out of service for an emergency outage repair, replacement or non-emergency shutdown, including annual maintenance based on critical routine maintenance before failures occur that could lead to enormous costs for unit refurbishment, replacement, repair, or unit upgrade as system reinforcement.

Electrical utilities operate power systems comprising large PT fleets, which vary with capacity rates, starting with some kilovolt-ampere (kVA) to hundreds of megavolt-amperes (MVA). These transformers are reliable assets in terms of their operation and can be running with good maintenance and operating conditions for more than 40 years and sometimes can reach over 60 or 80 years, which is the ultimate goal of each electrical utility (Metwally, 2011; Sefidgaran et al., 2011; Dumitran et al., 2014)

The paper by Madavan and Balaraman (2016) stated that the essential components for power transformers and their subsystems' active parts encompass bushing, core, winding, tap changer, tank and liquid or solid insulation that includes paper and oil, which constantly risk a variety of operating strains during its long-term service producing deterioration. Functioning PT for long service can generate degradation to transformer subsystems, which may result in catastrophic breakdown following significant revenue losses in the same unit or power supply interruption to strategic industry or commercial customers. In addition, some PT failures consume much time, and in some places, it can take over 12 months to get this unit back in service.

2.3.1 Power Transformer Active Components Impacting Transformer Health

Over the past ten years, much research has shown different active components of power transformers that significantly impact their health and operational lifespan. This section shows the main active components of PT, as shown in Figure 2.9. This knowledge is vital before starting to develop an HI assessment. In addition, Rao et al. (2019) highlighted the significance of dissolved gas analysis (DGA) in overseeing the chemical alterations in transformer oil as a crucial indicator of internal defects and degradation. Table 2.1 shows the ranges considering different gases based on faults (IEC 60599, 2007). Furthermore, Table C-2 lists the power transformer's active component and describes how it impacts its health.

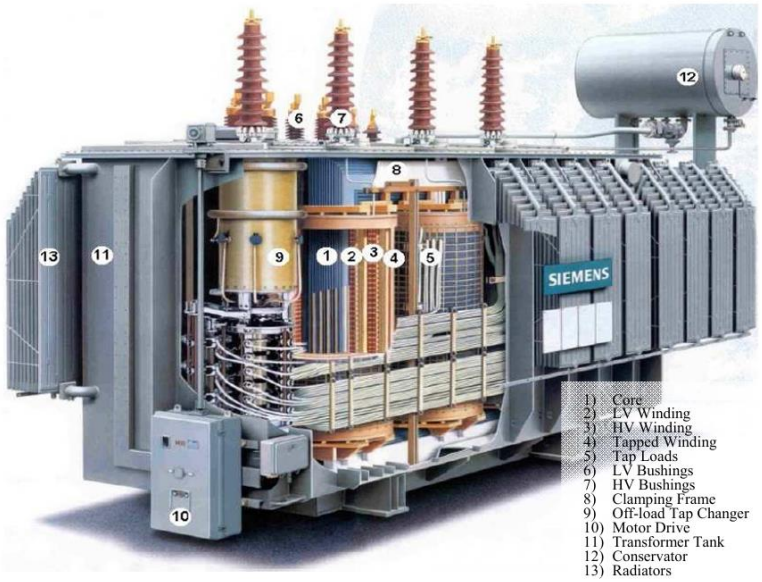


Figure 2.9: Siemens power transformer (Sermsukroongsakul et al., 2019).

Table 2.1: Fault measurements using gas ratios as in IEC 60599 (2007).

Probable Fault	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
Partial discharges	*	<0.1	<0.2
Low energy discharges	>1	0.1-0.5	>1
High energy discharges	0.6-2.5	0.1-1	>2
Thermal fault (T1)	*	>1	<1
Thermal fault (T2)	<0.1	>1	1-4
Thermal fault (T3)	<0.2	>1	>4

* Not significant T1 fault: T<300°C, T2 fault: 300°C<T<700°C, T3 fault: T>700°C

The following section will discuss the necessary parameters for developing the health index model.

2.3.2 Diagnostic Parameters for PTHI

This section will discuss the necessary parameters for developing the PTHI model. Khamis et al. (2023) used age, OQ (moisture, acidity, BDV), DGA and Furfuraldehyde. A literature review by (Naderian et al., 2008; and Jahromi et al.,2009) used DGA, Moisture in paper, OQA, furan, age, and dissipation factor parameters. The literature by Singh and Swanson (2018) used the DP factor, DGA factor, OQ factor and moisture in paper factor. Kadim et al., 2018 used four factors: DGA factor, OQ factor, Furan and age. This examination is in chapter four of the conventional HI model. Zuo et al. (2016) present Oil breakdown voltage (BDV), total acidity of Oil, 2-Furfuraldehyde content, and dissolved gas analysis (DGA) as the HI input parameters. This consideration is in chapter five of the fuzzy logic PTHI model. Benhmed et al. (2018) found that acidity, Breakdown Voltage (BDV), Furfuraldehyde (FFA) and water content are the most essential features in estimating the transformer HI. This scrutiny is in chapter seven of the ML PTHI model. This parameter selection justifies the parameters shown in Chapter 3, Figure 3.3, as sufficient to determine the health condition of PT.

Table 2.2 shows the relevant input data variables for the oil paper transformer insulation collected from the electrical utility for the large fleet of transformers for this investigation health index model as per the earlier literature review findings. The outcomes can be proposed for future maintenance programmes and compared with existing utility practices for additional improvement based on the international best

practices and standards from the literature reviews. This data was a contribution and support from electrical utility for this research study, which is appreciated by the academia and mentioned in various publications by this research.

Table 2.2: The research PTHI variables.

x1	water content in oil
x2	total acidity of oil
x3	oil breakdown voltage (BDV)
x4	dissipation factor (DF)
x5	Dissolved combustion gases (DCG)
x6	2-Furfuraldehyde content

Based on (The Canadian Electricity Association, Deloitte, 2015; Dumitran et al., 2014), the selection of the influencing factors, the grading and the weighting are worked out by expert's judgement based on the outcome of various experiences.

Each electrical utility, researcher or developer of PTHI places different values on the various factors involved within a utility, academia or industry. Scoring and weighting factors must agreed upon within the same country. Multiple stakeholders must define what is essential and how the calculations should result in a consistent approach across an asset class (CIGRE Technical Brochure 761, 2019; The Canadian Electricity Association, Deloitte, 2015). Therefore, the proposed limit from electrical utilities' agreed scores and weighting factors to structure the HI score classes. The condition boundary put into consideration based on several international studies and standards published by the International Electrical and Electronics & Engineers (IEEE), the International Electrotechnical Commission (IEC), and the Conseil International des Grands Réseaux Electriques (CIGRE). These include (IEEE C57.106, 2015; IEC 60505, 2011; CIGRE Technical Brochure 761, 2019; Suwanasri et al., 2013; Pompili and Scatiggio, 2015).

It is vital to the whole method that data on degradation and failure from operation and maintenance practices is available. The most valuable data is from postmortem physical investigations of transformer scraps. Finding the range of degradation for

utmost scrapped transformers is conceivable by noticing and testing critical parts, most significantly the solid insulation. If solid insulation fails operation, the failure mechanism must be known. In any situation, the data confirm and calibrate the state assessment methods to find any earlier unknown failure mechanisms (CIGRE Technical Brochure 735, 2018). The literature reviews agreed on the PTHI layers for developing the asset register, which consists of the data layer, factor layer and health index layer (Jürgensen et al., 2016; Li et al., 2018; CIGRE Technical Brochure 761, 2019; Rediansyah et al., 2021), as shown in Figure 2.10.

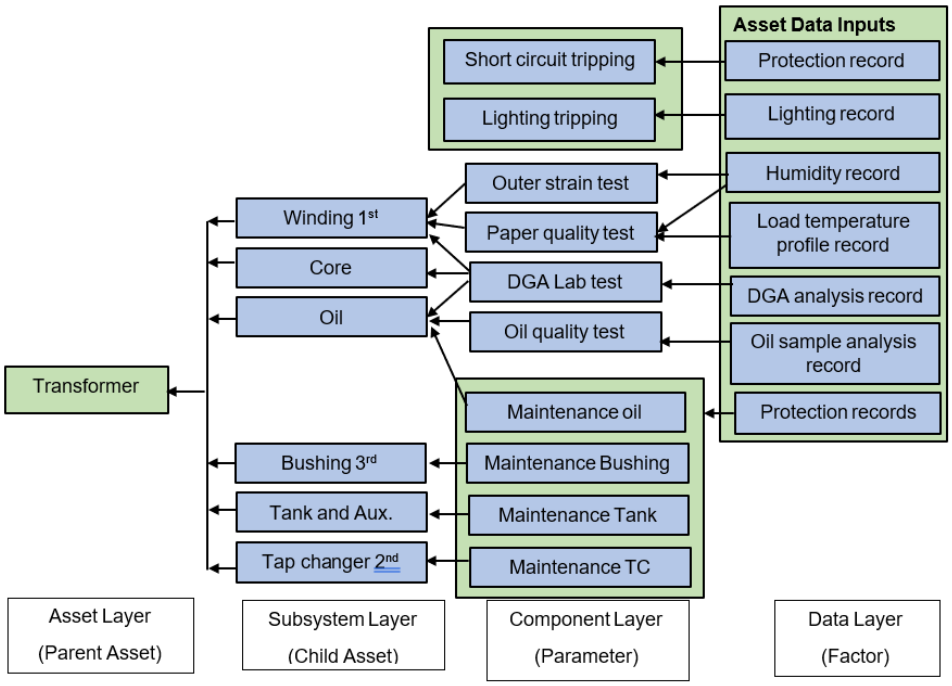


Figure 2.10: The PTHI layers (Li et al., 2018).

Evaluating the most related data defines the critical indicators for designing the HI Foros and Istad (2020). These indicators require considered data accessible for the enormous transformer fleet to assess specific transformers on similar ground based on key Indicators (Picher et al., 2014), as shown in Table 2.3. A good understanding of these indicators and how linking with failure modes and mechanisms is vital for good PTHI development and determining PT health conditions, as shown in section 2.5.3.

Table 2.3: PTHI indicators (Picher et al., 2014).

Indicator	Description
Dissolved Gas Analysis (DGA)	DGA indicator is analysed built on the absolute values of the numerous gases assessed, the progress of these gases and the rating coefficients for all gases.
Oil tests	Oil quality is considered reliant on acidity, interfacial tension, dielectric strength, and power factor.
Moisture content	Moisture content in oil is assessed to assess the water content in the paper.
Solid insulation	Paper ageing indicators assessed in PT oil for instance methanol, furans, and CO/CO_2 content (subject on data accessibility).
Aging Mechanism	The concept aging mechanism refers to the process that can lead over a period of a time to reduce the quality of service for the indicated transformer subsystem. This presents the age of transformer since the year of installation.
Failure Mode of similar PT	Failure mode is defined as the present and detectable abnormality condition within the transformer. Failure mode data is applied to detect less reliable PT groups. PT groups are built for PT with similar specifications, manufacturer, similar serial numbers and years of manufacture.

2.3.3 Power Transformer Condition Assessment Methods

There are various power transformer condition assessments. This section provides in-depth learning on the available health condition assessments prior PTHI development processes and before conducting a detailed analysis and comparison among each PTHI model that is available in the academia from various researchers including (Zuo et al., 2016; Zeinoddini-Meymand and Vahidi, 2016; Scatiggio and Pompili, 2013) and within market vendors including (Alstom Grid, Power Transformers Vol 2, 2011; Hydro-Quebec TransEnergie, 2019; DNV, 2013; EA Technology, 2018; ABB, 2008).

Various researchers, electrical utility efforts, and published papers consider the importance of transformer condition assessment as it provides the PT's actual health indications (ABB Review 1 Transformer Challenges, 2008). Some referred to the condition indices as the “Health Index” (Foros and Istad, 2020; CIGRE Technical Brochure 761, 2019; The Canadian Electricity Association, Deloitte, 2015), whether based on practical experience or the methods of the applied investigation or considering it from the prospect of the global manufacturer terminology.

The academic works, theories, practices, and terminology described in these periodicals have all been under review, studied, and improved from how it was tackled and produced to a new contribution that focuses on solving today's short-term operation and maintenance activities, which consider a long-term solution for future problems.

This research's quality and accuracy deliver the required detailed narrative, commendations, and observations, contributing to other previously published methodologies. It is beneficial to start with the content of the references for further investigation, especially the HI condition assessment, which is a relative tool that requires further improvement. This research idea is undertaken for further investigation, consensus, findings and achievement, as shown in chapters 6, 7, 8, and 9, and the preferred method is declared using the indicated ML measures. (CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761 (2019) proposed the HI process, which consists of numerous steps, as shown in Figure 2.11.

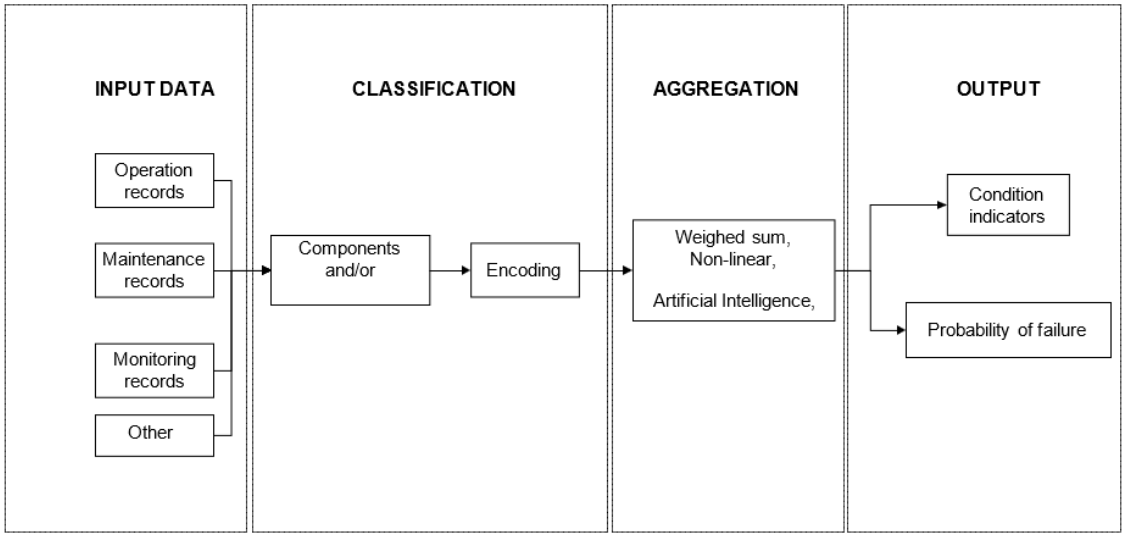


Figure 2.11: HI processes (CIGRE Technical Brochure 761, 2019).

This research covers the above process and contributes to the identified research gaps based on the published literature.

AM decisions from HI assessments require attention to prevent any risks due to the impact of the masking score, which may occur during the scoring calculation or the designing process of the HI development (Sukrislismono and El Nur, 2015). However, HI scoring calculation does not intend to be used as “Alarm management” for a “Bad”

score requiring immediate action. Therefore, a detailed investigation must be considered for any score issue and solved despite the HI score. AM must respond correctly to set alarms and act upon bad or suspected scores, specious diagnostic results, or dangerous observations. Neglecting a “Bad” score can cause catastrophic failure shortly. The unpredicted failures of bushings and on-load tap changers (OLTC) might regularly cause devastating transformer breakdowns. AM must minimise the full condition assessment time frame and respond appropriately by inspecting flashover or smoke from transformers during site inspection (CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761, 2019).

However, it is agreed by (Jürgensen et al., 2016) that all parts and subparts of transformers should go under thorough PTHI assessment layers. However, the HI assessment used for transformers assesses only the active part. HI, assessment for transformer active parts conducted for the following failure assessments (CIGRE Technical Brochure 761, 2019):

1. Thermal failure: This assessment is quicker than predicted ageing because of the high temperatures from metal hot spots, hot metal defects isolated using paper, and hot winding spots.
2. Dielectric failure: this is the role of the deterioration method based on liquid isolation/solid isolation due to extreme thermal or strain of electrical fault, dielectric liquid/paper, factory faults, and contamination of solid particles.
3. Mechanical failure: This is based on the collapse of the composition of windings, solid insulation and the core because of various factors, including solid electromagnetic forces from winding deformation, over-winding short circuit current flow because of the shocks over the carriage, erection through the fault of the system and solid insulating material deterioration of solid insulating material.

The diagnostics techniques use laboratory tests to conduct DGA. The online DGA screens analyse the signs of an early fault emerging in transformers, which is the best example of FSA and relies on IEEE/IEC standards. Using related standards, including ISO18095 (2018), relating to condition monitoring and diagnostics of power transformers (CIGRE Technical Brochure 761, 2019). The condition assessments form the following two methods:

2.3.3.1 Fleet Screening Assessment (FSA).

The objective of a Fleet Screening Assessment (FSA) or Strategic Planning Assessment (SPA) for a population of a large fleet of PT is to find the weakest units by quick analysis (ABB Review 1 Transformer Challenges, 2008). AM considers HI a detailed assessment after conducting initial FSA, a method that uses readily available information to analyse each large fleet transformer quickly. Hence, the maintenance or replacement activities for the power transformers prioritise the actions that undergo detailed assessment. The scoring calculation method is the basis for ranking the transformers. The analysis process can be automated or semi-automated to simplify the FSA process. The power transformers identified through the FSA process would most probably be subject to extra detailed analysis before considering any budget decision to ensure the FSA powerement's outcome with the right and the required investment evidence (ABB Review 1 Transformer Challenges, 2008).

The FSA task backs the strategic decisions associated with single units and larger populations. The HI information collection from online monitors or offline testing, design, operation, maintenance of family history records, diagnostic tests, and technical analysis. Regular inspections of the critical units and visual inspections to assess transformers and their subparts. The SPA or FSA considers the following two essential approaches:

1. Statistical analysis: to estimate transformer life as a primary independent variable.
2. A unit-oriented method is obtaining the state or resist the ability of each specific subsystem within the transformer. Although statistical analysis frequently assists as a vital first step, the specific and unit-oriented tactic is essential in most situations.

Transformer screening algorithms are reliant on found and accurate data. Based on the ABB example, as one of the largest manufacturers of the transformers industry, the data is to be stated as scores between 0 and 100 (ABB Review 1 Transformer Challenges, 2008). Where the 100 score shows the worst and most severe state for the considered factor, some influencing factors are more serious than others. Allocating the weighting value to each influential factor becomes necessary. The

general sign for the practical risk of a specific unit is by joining the specific scores – either as a weighted sum or via the highest value of the specific scores. The influential factor analysis relies on existing limits based on ABB Figure 2.12. Some are related to historical general degradation, and others are related to thermal wear, likely unexpected cases, repair, and experience.

In summary, as shown in Figure 2.13, FSA is used for large transformer fleets to detect transformer weakness for more diagnostic tests and assessments before making significant conclusions by a detailed condition assessment. These conclusions are repair, refurbishment, replacement, or regular maintenance. Preliminary FSA relies on initial site visit inspection, a few tests, or other reasons, such as loading conditions, age, historical faults, and factory records. It requires the assessments to rely only on a few accessible data. Conducting this type of assessment is always better than nothing. It leads to understanding any severe issues that developed a while ago.

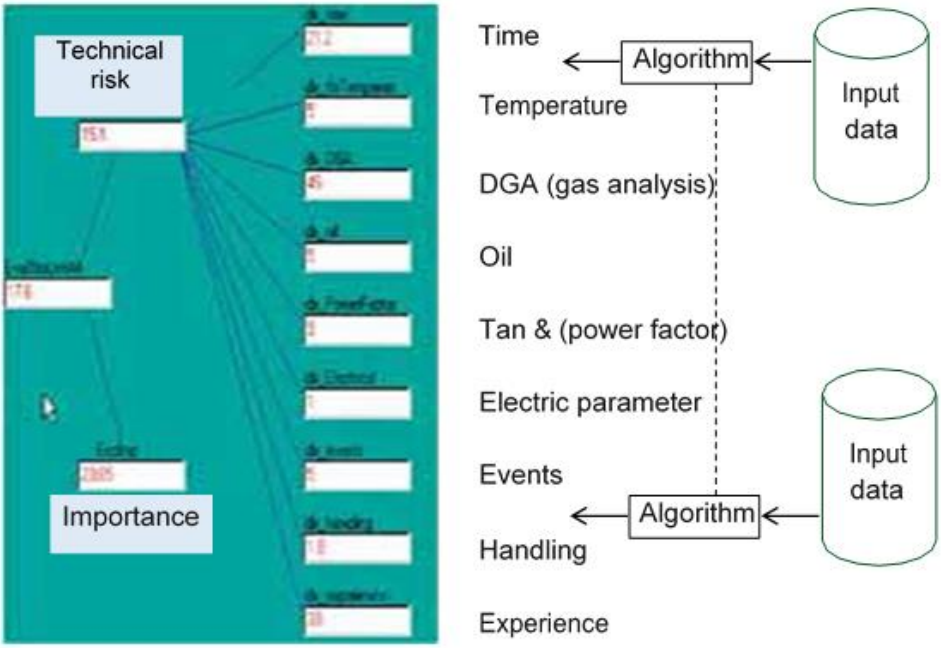


Figure 2.12: Preliminary FSA (ABB Review 1 Transformer Challenges, 2008).

1 Example of a fleet screening survey

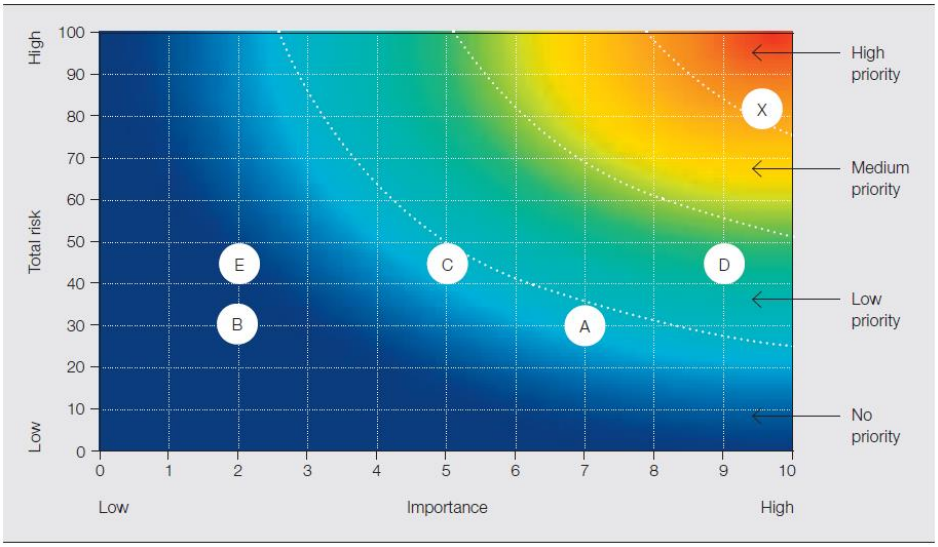


Figure 2.13: Transformer FSA (ABB Review 1 Transformer Challenges, 2008).

2.3.3.2 Detailed Assessment (DA)

AM uses transformer specialist experience to perform a detailed and thorough analysis of each transformer before assigning transformer HI assessment scores. Significant decisions are reliant on transformers' HI scoring calculation. Suppose the implementation of a proper and practical scoring calculation. In that case, a score deriving through a more exhaustive method can assist in backing up the justification of the AM team to the leadership team and regulator to agree on repair, replacement or refurbishment transformer decisions. Since a calculation for each transformer in the large fleet, the allocated scores will rate the transformer's health condition. If the transformer scores calculation via efficient/effective methods with timely, accurate data, then the AM can effortlessly recognise the transformers that most instantly need care or decision. Also, many numeric scoring calculations result in a value that permits transformers to be rated.

Nevertheless, the scores themselves are often arbitrary. In these circumstances, the AM must not undertake a transformer with a score of 84, which is twice as good as a transformer with a score of 42. Nor must undertake the residual life or the probability of failure of one transformer is twice that of the other. In most scenarios, the numeral provides a means of rating the transformers for a large fleet.

Furthermore, a detailed statistical assessment is applied using the ranking method to approximate the failure probability or transformer residual life. This detailed assessment is similar to this research's conventional, fuzzy logic and ML PTHI, which require the involvement of various AM teams to include the outcomes within the lifecycle planning investment. Hence, HI is scaled to represent the parameter factors. A good HI and scoring calculations are the ones that easily find these transformers that require a quick response. The period for decision relies on the period defined in the scoring metric, which is applied to evaluate the failure modes/mechanisms mentioned in Appendix G Tables G-1, G-2, and G-3, which are parts of the HI development process.

During the detailed assessment, it is crucial and the responsibility of the AM to transfer the knowledge about the aim of the HI to all other technical departments, including the detailed technical analysis discussion. It is considered a reliability HI, prioritising transformers about to collapse and assessing PT lifecycle planning. The AM must utilise more than the transformer replacement budget to replace the top 10 transformers on the list. This decision is because some of the transformers on the top list can effortlessly and economically go for repair. In contrast, other transformers on the bottom of the list require an urgent replacement, which needs more attention in the scoring calculation metric.

However, the proper way to decide is to implement a detailed technical analysis to find those transformers that require a decision. Based on the cost-value analysis, these implementations effectively link the total investment plan. Therefore, a detailed analysis will consider criticality and other matters.

2.4 The Importance of the Health Index

PTHI is an AM decision-making tool that generally provides maintenance and AM teams with a detailed evaluation of the health condition of a large fleet of assets or particular assets (PAS55, 2008). It shall support making accurate decisions at the right time (CIGRE Technical Brochure 761, 2019; Kinnunen et al., 2016; Ibrahim et al., 2016). PTHI shall include the main asset as ‘PT’ and other ‘PT’ sub-assets within the same group of assets because the collapse of the main transformer can be due to any small component within the transformer subsystem Islam et al. (2016).

HI, scoring classes based on the PT health condition agreed by the HI developer and researchers vary from 2 to 5 classes based on the following recently defined health condition:

- Two classes, 'Healthy' and 'Unhealthy' (Zuo et al., 2016).
- Three classes, 'Good', 'Moderate' and 'Bad' (Bohatyrewicz, Płowucha and Subocz, 2019).
- Five classes agreed by academia and practice; 'Very Good', 'Good', 'Moderate', 'Bad' and 'Very Bad'. These are the most used and detailed classes in both academia and practice, according to (Jin et al. 2022; Wattakapaiboon and Pattanadech, 2016; CIGRE Technical Brochure 761, 2019; Dias and Cheim, 2021; Gorgan et al.,2010; Naderian et al. (2008); Jahromi et al., 2009; Christina et al., 2018, Da Silva et al., 2021; Khamis et al., 2023).

Although the PTHI must have an annual calibration, similar to other tools, it enables checking the algorithm codes, scoring and weighting calculations and conducting additional updates based on new PT parameter condition classes (Jürgensen et al., 2017).

Foros and Istad (2020) declared that for transformer quantitative health condition assessment and residual lifetime assessment, three standard primary methodologies are available within the academia: degradation modelling, data online monitoring system, and statistical ageing analysis based on PT in-service end-of-life modelling. (Abbasi and Malik, 2016) stated statistical analysis and health evaluation, which the researcher agreed with the manufacturer's findings (ABB Review 1 Transformer Challenges, 2008). This finding remains similar to Foros and Istad (2020), but it is only defined differently based on the CIGRE Technical Brochure 761 condition assessment of power transformers (2019) and (Khamis et al., 2023). Lundgaard et al. (2015) mentioned that PTHI in practice differs from academia, making integrating the two more critical. Therefore, Khamis et al., 2023) promoted the integration field experience AM framework with the development of PTHI as a vital solution and improvement to the PTHI.

The PTHI methodology considers condition monitoring data obtained from oil tests, visual inspection, and historical records and then characteristically accumulates these records into a quantitative rate of current condition (Gorgan et al., 2012). PTHI

methodology development requires innovative statistical models. However, PTHI usually necessitates a certain level of transformer specialist judgement.

There are several gains and some drawbacks in conjunction with the PTHI methodology. It empowers the calculation of residual life expectancy. Also, it empowers the whole existing condition monitoring record by integrating it with the assessment of residual life expectancy. It can be tricky because it requires practical experience that can add value to PTHI, leading to better PT ageing for lifecycle planning. The logic of this methodology varies on the condition and representability of the accessible statistical records. The PTHI method supports the AM decisions and details the advantages and disadvantages of the assessment decisions (Vermeer et al., 2013). These decisions allow AM, transformer engineers and specialists to determine and develop the most suitable HI for large power transformer fleets (Vermeer et al., 2015). Transformer criticality or other circumstances must be considered before making a replacement or repair decision with any additional significant interferences (Vermeer et al., 2016).

On the other hand, the drawback is the complexity of the mathematical model (Milosavljevic and Janjic, 2020), as well the conventional PTHI decision-making without AI depends on power transformers experts' judgements (Jahromi et al., 2009; Kumar Bhowmik, 2015).

The PTHI model rates a fleet of electrical transformers reliant on the possibility of the predicted failure (Arias Velásquez and Mejía Lara, 2018) at a particular time and degraded from its novel condition (Murugan and Ramasamy, 2019). The PTHI rates all transformers needing any decision that relies on the logic of technical insistency (Metwally, 2011). Operators' prerequisite to recognise the most suitable decision for the transformers is acknowledged and demanding care (Martins, 2014). The trustworthiness of the PTHI will not usually result in a score that directly relates to the trustworthiness or the fitness of the transformer. If the PTHI is fit calculated, it will grade the transformers that are less fit or less trustworthy, which need urgent action. Hernanda et al. (2014) mentioned that AM would consider the transformer for either replacement refurbishment or repair decision. The cost evaluation against the proposed decision is to achieve the required optimum performance and avoid further risks to the transformer. It also confirms that PTHI can assist in identifying the bad

condition of the transformers, which likely requires an urgent decision. Other reasons for measurement are before conclusive tactics and urgencies of each transformer replacement repair or refurbishment decision (Karthik et al., 2013).

PTHI users must understand each evaluation process before ranking a transformer calculation score for each transformer, especially if the data is unavailable for some large fleet transformers, which can lead to the wrong score (IEEE Std C57.140, 2017). Hence, the user must confirm that the scoring calculation relies on good data quality and that low data quality must be indicated and solved. If the data used for scoring calculation needs to be revised or corrected, the transformer which requires immediate action must not be indicated by the HI (Jagers et al., 2009).

It is essential to be careful with the proposed purpose (repair, replacement, regular maintenance) for the HI development. A poorly designed HI can hide and mask a “bad” result, including a transformer performing well but whose bushing is in bad condition. The HI that yields a simple numeric result, the overall score can show the PT is healthy. This finding occurred because bad bushing conditions were hidden and masked by the good scores allocated to the other failure modes (ABB Review 1 Transformer Challenges, 2008). Therefore, this can still managed by using weighting factors. At the time, the bushing score, weight, and a bad score would result in a transformer with a bad bushing.

Furthermore, the other allocated failure modes have good scores (Varoquaux, 2018). The issue is that the masking impact overstatement for different failure modes, which must be deeply weighted. Problems not linked to the bushing can yield failure and cause immediate responses hidden and masked by the weighted system. The immediate linked issues can cause a poorly designed HI (CIGRE Technical Brochure 761, 2019).

2.5 Power Transformer Health Index Development

Over time, the PTHI concept has developed for electrical utilities, as noted by Fauzi et al. (2020). HI is a sophisticated tool used to rank the power transformer using a numerical scoring system to reflect its condition (Wang et al., 2022). The purpose of this section is to review the concept of PTHI through a detailed assessment of recent literature.

Health indices play a crucial role in determining maintenance strategies for power transformers. The international economic changes made clear, and the desire to use the PTHI condition assessment tool to prevent early-time malfunctions due to its simplicity of HI (Alqudsi and El-Hag, 2019), which can lead to managing unnecessary spending to both network operators and government bodies (Istad et al., 2018). Zuo et al. (2016) mentioned that PTHI is an excellent approach to offering overall asset health conditions. PTHI intends to implement maintenance strategies and plan routine maintenance needs, signifying the failure rate, empowering electrical network operators, and even determining the asset's end of life (Alstom Grid Power Transformers Vol 2, 2011).

Initially, the maintenance management of power transformers revolved around evaluating each unit's physical state and reliability (Bohatyrewicz, Płowucha and Subocz, 2019). As Senoussaoui et al.'s research in 2021 has shown, a comprehensive approach is necessary for a thorough evaluation of transformer health. It has become increasingly evident that a more holistic approach is needed over time (Benhmed et al., 2018). This approach leads to developing a suitable PTHI method to evaluate the power transformer's situation (Tenbohlen et al., 2016). Therefore, creating an effective method to assess PT health conditions is crucial (Azmi et al.,2017), facilitating decision-making on transformer asset lifecycle planning (ISO 55001, 2014), including daily operation and maintenance activities.

The improvement of HI in line with AM planning and strategies became a decent best practice for operating electrical networks at high service and low cost (National Grid, 2019) that relief electrical utilities in more control in the CAPEX and OPEX and in projecting daily, monthly, even yearly assets maintenance and replacement plans which is this research interest. This scenario lowers PT's risk of failures and energy supply breaks due to major system outages, as stated by Kadim et al. (2018).

PTHI operates to achieve standard operation maintenance and asset lifecycle planning (Dumitran et al., 2014). PTHI can be built on various algorithms (Nhu et al., 2020), which are associated with meeting a particular requirement that yields highly competent decision-making products for handling large amounts of data for electrical utilities (San, 2013). These HI products are mainly to improve in-service PT performance (Khan and Hussain, 2019).

Davidenko and Kuzina (2017) mentioned that HI as a condition assessment metric is a critical invention that aims to aid electrical companies in becoming more agile and economical in coping with the daily rising volumes of records. HI construction to promptly enhance the execution of a company's operation and data-centric enterprise activities by computerising the developments that are very pricy, time-consuming, or complicated to scale as stated by (The Canadian Electricity Association, Deloitte, 2015) and agreed by Kadim et al. (2018). The PTHI method can be established using computer and data science to enhance capital and maintenance actions (Fauzi et al., 2020; Gao et al., 2013; Henke et al., 2016; Jan et al., 2015). The electrical network operators see PTHI as an innovative preventive maintenance tool based on their desire to prevent earlier failures, which could be costly Kadim et al., 2018). A good PTHI allows electrical utilities to plan asset maintenance or replacement rather than focusing only on future asset lifecycle planning problems (Khamis et al., 2023).

Technically, the risky transformer failure modes are categorised as internal or external, decreasing the effect on a large fleet's transformer lifetime to optimise the investment and avoid future power supply disturbances (Martin et al., 2017). It also stated that most electrical companies moved from time-based to condition-based management approaches (Hashemnia et al., 2021). It agrees that the Condition Base Maintenance (CBM) approach is a method that could ensure highly reliable in-service PT (Hernandez and Labib, 2017). Jan et al. (2015) confirmed that this dramatically affects electrical PT downtime when a failure occurs. Also, it is mentioned in (CIGRE Technical Brochure 445, 2011) that PT is a reliable asset that can be difficult to break.

Consequently, when this occurs, normalisation takes a very long time, especially bringing another redundant PT unavailable within the area, not just the site, which is challenging and depends on network operation and maintenance policies, cost, resources, and grid code obligations. PTHI designer uses practices based on IEC 62975 (2021), IEEE Std, C57.19.00 (2018) and CIGRE Technical Brochure 445 (2011). This practice will aid the decision that PTHI is the best strategy for transformer life management and is a confidence practice by electrical utilities (Suwnansri, 2014).

To conclude, developing the PTHI is a vital solution to the demand for a reliable and uniform tool that can integrate various condition status data and deliver a singular outcome (Gouda et al., 2021).

The following section will discuss the purpose of developing HI condition assessment for power transformers, including managing the risk of failure mode and ageing mechanism and creating a failure mode assessment score.

2.5.1 Purposes of HI Condition Assessment

This section discusses the purpose of developing HI condition assessment for power transformers. HI techniques are essential for predictive maintenance and driving maintenance strategy for electrical utilities. According to Jahromi et al. (2009) and CIGRE Technical Brochure 761 (2019), the purposes of PTHI condition assessment vary based on the defined purpose, including repair, replacement and refurbishment. Hence, it is significantly required to know during the health condition assessment analysis the difference in purposes of PTHI CIGRE Technical Brochure 761 (2019). Understanding the differences between the questions regarding which index/score is required is vital because it determines the scope of the PTHI development.

AM team also understands the variance between which power transformers on the network are about to fail and which power transformers go first for replacement. The first question considers all the thinkable failure mechanisms. Still, the second question dismisses mechanisms modified by different solutions rather than replacement purposes, meaning they are only interested in using replacement index rather than in purposes for repair and refurbishments. Therefore, it is required to focus only on selecting the scores related to the purposes of replacement indices.

To conclude, HI development will study and prioritise transformer intervention, including repair and replacement. However, similar methods research and prioritise other interventions. Each HI would utilise a subgroup of the available data state that is unlike the other.

The following section will discuss the management of the risk of failure mode and ageing mechanism in power transformers.

2.5.2 Risk Management for Failure Modes and Ageing Mechanism

This section presents the management of the risk of failure mode and ageing mechanism in power transformers. Risk management in power transformer asset management involves identifying, assessing, and mitigating risks associated with transformer operation. This risk mitigation includes understanding failure modes and their impacts and implementing strategies to reduce these risks (IEEE, 2019).

According to CIGRE Technical Brochure 761 condition assessment of power transformers (2019) and the United States (US) Bureau of Reclamation, Transformers Basic Maintenance and Diagnostics (2005), the failure mode concept describes the detected current abnormal condition with the transformer. In contrast, the ageing mechanism concept refers to the process that can lead after a specific time to reduce the performance quality of the indicated transformer subsystem. The ageing mechanisms will not be visible or occur at first. Still, with sufficient time, situations will lead to failure mode as various faults of PT components are considered for failure analysis, expressed in Figure 2.14.

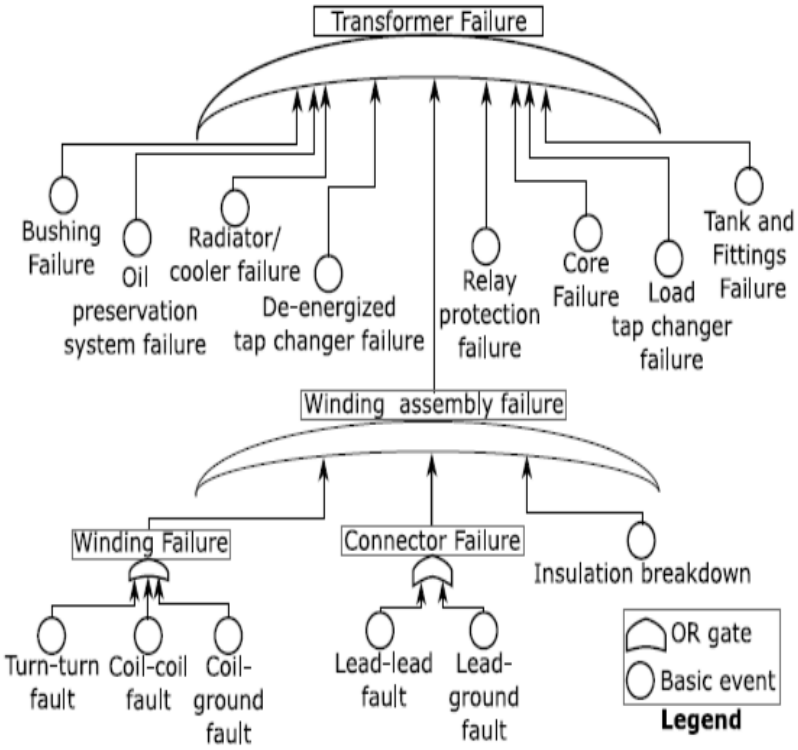


Figure 2.14: PT failure modes and faults (Murugan and Ramasamy, 2019).

Figure 2.15 shows the transformer's active components, which predict the source of failures. The AM team and the maintenance and operation department in the electrical utilities work closely and daily to ensure these components are well maintained.

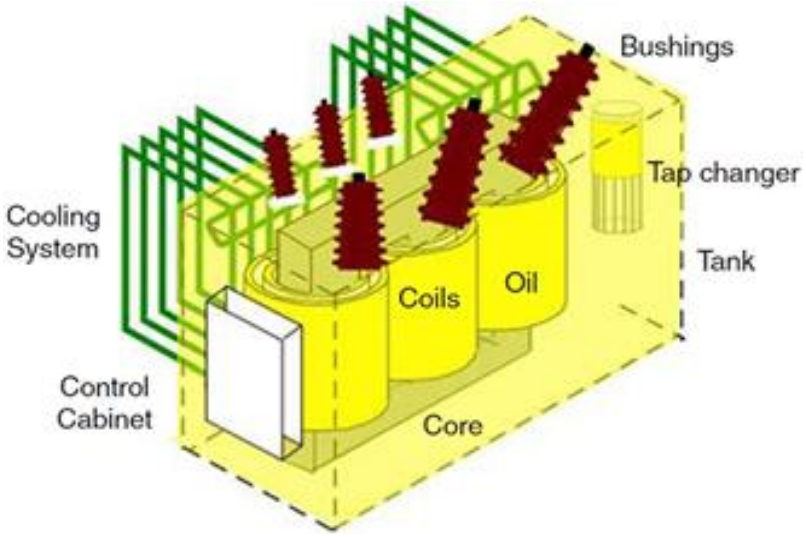


Figure 2.15: PT active components (Murugan and Ramasamy, 2019).

2.5.3 Developing Failure Mode Assessment Score for Health Index

It is vital to determine the purpose of the HI, including the related failure modes or mechanisms, and use test data and other available data to evaluate the failure mode or mechanisms. It is also vital to differentiate between failure mechanisms that can lead to failure due to partial discharge and the indicators applied for assessment. Moreover, it considers the level of dissolved gas in the dielectric fluid. Several transformer assessment methods were derived directly from indicator to score. It is difficult to properly understand and accurately weigh the signs without seeing the failure mechanisms signs. Assessments must be reliable to build a scoring metric. Hence, scores are assigned consistently to each failure mechanism of each transformer.

This section addresses the development mechanism for the power transformer failure mode. According to CIGRE Technical Brochure 761 (2019), the indicators are allocated to related failure mechanisms and applied to define the probability of failure as the actions assigned to the selected timeframe. Then, link the performance of various failure mechanisms with the complete transformer assessment score.

The HI scores are developed based on the related condition assessment index and the related failure mechanisms. The condition assessment scores are required and linked in a way that permits a transformer with a single failure mechanism near failure emphasised before failure occurs (Tran et al., 2020; Wani and Ranga, 2023). However, for a transformer with various minor issues, emphasis is placed on the top of a transformer with fewer minor issues. This listing process is a complicated task because of different levels of data and perceived rank for unlike signs (Zeinoddini-Meymand and Vahidi, 2016; Rediansyah et al., 2021), considering the facts of applying and using several methods for calculating HI scores (Prasojo et al., 2021; Vasović et al., 2021).

The most excellent demanding technique of grouping is using the probability of failure from each failure mechanism and then multiplying the survival rates from each failure mechanism collected to reach the total survival rate (Ghasemi et al.,2023; Gjorgiev and Sansavini, 2022). This technique is equivalent to a group's failure, showing each group's failure rate. Regrettably, the data on the failure mechanisms and the quality of the signs is not enough to use this way with confidence, although the overall tactic is actually to satisfy the above grouping principles (AJ et al., 2018; Csanyi, 2020; Fengjun Yi et al., 2015).

The weighted average method is applied to join scores. The prime of weightings is hard, and random weightings can cause critical bias without the PT expert judgement (CIGRE Technical Brochure 761, 2019). The averaging method shows that the more failure mechanisms are involved, the harder it is to let specific “Bad” factors have the importance that justifies the total impact on a healthy PT (Foros and Istad, 2020).

Once the direct/indirect measurement parameters are determined, this requires knowing how critical these tests (diagnostics/inspection) are to PT health conditions that also indicate the relative importance (Picher et al., 2014; IEEE Std C57.140, 2017). The relative importance assessment follows the process of finding the weighting and scoring factors (CIGRE A2_101, 2014), as shown in Appendix F.

The calculated scores are decent as the data is accurate when assessing the transformer. Unfortunately, the used failure mode data for scoring is not always new CIGRE Technical Brochure 761 (2019). This ignorance of failure modes data quality

will result in low-quality scores, especially if the data are unavailable entirely. It can also make a minor variance to the total score. It can also make a “Good” transformer look “Bad”, which leads to the transformer going unnoticed for severe repair or replacement (Jahromi et al., 2009; Chen et al., 2016). Likewise, the low quality of failure mode data can make a “Bad” transformer look good. Applying a “Bad” score or “Bad” data means misunderstanding the quality of failure mode data and the reliability or data confidence level of a score (Wang et al., 2022).

(Khamis et al., 2023; Jin et al., 2022; Abbasi and Malik, 2016); stated that good AM must use various numeric scoring methods that deliver a score suitable only for the status of the transformers. The score's total value can mislead the transformer's total state or the immediate action to report transformer issues (Jahromi et al. 2009). Nevertheless, when a HI is applied to indicate the worst transformers in a large fleet, the applied calculation to obtain the calculated score must be revised to ensure the correct response is within the required timeframe (Fengjun Yi et al., 2015).

All PTHI assessments depend on the failure modes/mechanisms analysis methods, and evaluating them using the utmost detailed and innovative health conditions is vital, according to CIGRE Technical Brochure 761 Condition Assessment of Power Transformers (2019). However, this is not practically or economically conceivable in several health states. The choice must be made based on the method of evaluating a failure mode/mechanism. There are various methods of evaluation, and each evaluation method has an equivalent grade of correctness and ambiguity. Several evaluation methods are offered to reduce ambiguity levels (Arias Velásquez and Mejía Lara, 2018).

CIGRE, Technical Brochure 761 (2019) presents diagnostic findings showing various failure mechanisms and the allocated scores for each failure mechanism, collected for unlike indices. Figure 2.16 describes the relationships between diagnostics indications, failure modes/mechanisms and assessment index. When developing HI, AM must link the related failure modes/mechanisms with the related HI purpose. Data availability for the required condition assessment is necessary because it will lead to accurate results.

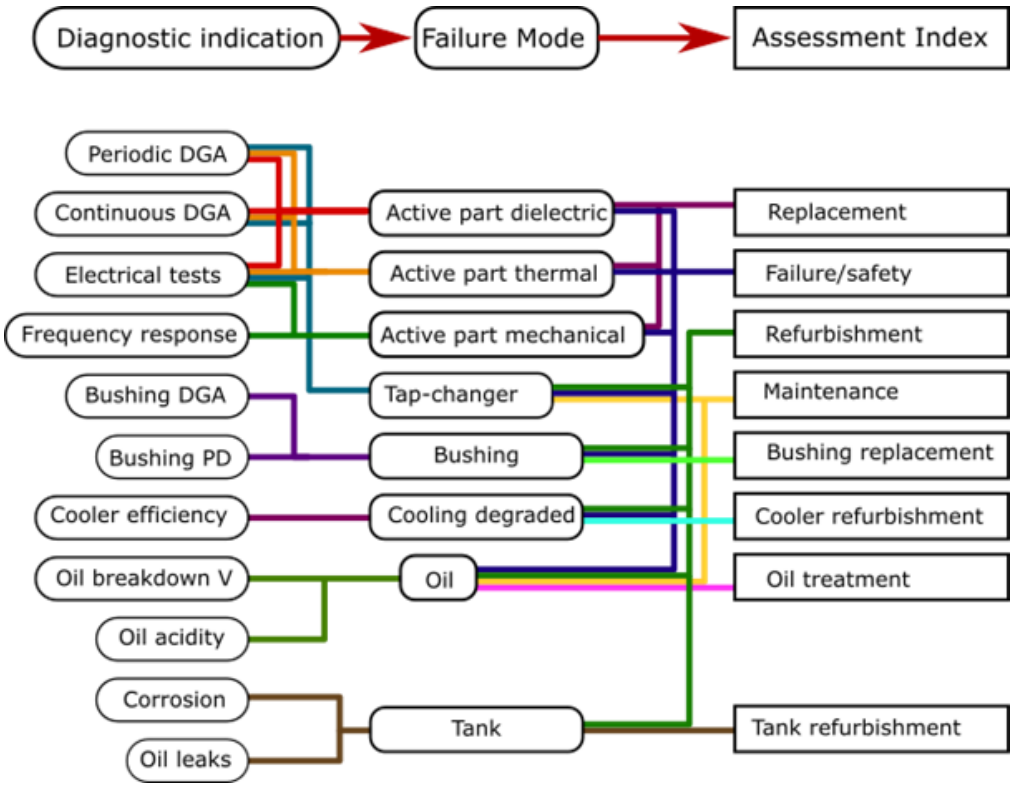


Figure 2.16: Failure modes, assessment indices and diagnostic result relationship (CIGRE Technical Brochure 761, 2019).

To conclude, failure modes and mechanisms play a crucial role in determining the health condition of power transformers. These indicators assign the score and weighting factors for calculating the overall PTHI results that describe the PT health condition.

The following section will present the development process and the approach to health index calculation.

2.5.4 Power Transformer Health Index Calculation

This section presents the development process and the approach to health index calculation. According to (The Canadian Electricity Association, Deloitte, 2015 and Dumitran et al., 2014), asset HI developers follow the process for the development of PTHI: identify, analyse, then develop, as shown in Figure 2.17.

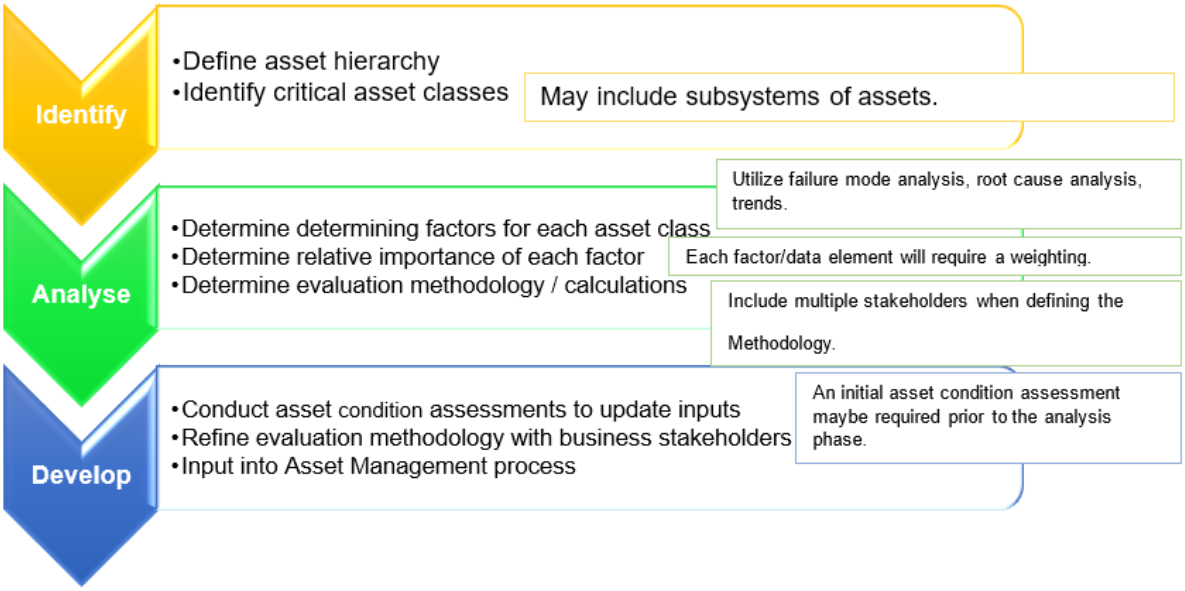


Figure 2.17: Development stages of HI calculation method (The Canadian Electricity Association, Deloitte, 2015; Dumitran et al., 2014).

This research is considered the HI equation in Chapter 4 (4.2) based on the researcher (Jahromi et al., 2009; Murugan and Ramasamy, 2019), which enables the correct and precise measurement of a transformer's condition. In addition, (Puluhem et al.,2011) agreed that the HI equation is derived using the following simple equation concept:

$$HI = \Sigma [Wfi \times Ifi] \quad (2.1)$$

Where HI represents the final HI metric for each failure mode and mechanism of degradation, Wfi is the weighting factor for the relative importance of the diagnostic test parameters, and Ifi is the influencing grading factor for all PT health conditions obtained from the diagnostic test. Tables D-17 and E-4 present the scoring methods for calculating HI based on academia and practice as agreed by (Jahromi et al., 2009; Puluhem et al., 2011; Murugan and Ramasamy, 2019; The Canadian Electricity Association Deloitte, 2015; CIGRE Technical Brochure 761, 2019).

The number of diagnostic factors used to calculate the total HI scoring model varies within the research area and practice. The research paper by Naderian et al. (2008) considered 20. In contrast, research by Wattakapaiboon and Pattanadech (2016) considered 15 factors. Research by Dias and Cheim (2021) considered 15 factors.

Hernanda et al. (2014) considered fault types of possibility on transformer 23 factors, stating that the selection of the diagnostic measurement factors on experience, empirical data, investigating literature, or expert knowledge. The research paper by (Gorgan et al.,2010) discusses 23 factors.

Table 2.4 presents a new list for assessing transformer conditions, improving previous conventional HI literature of PT condition assessment factors. This research shows the newly developed HI table of 28 diagnosis inputs to compute the required HI (Khamis et al., 2023). Each diagnostic's scoring and weighting factors are presented for each input and considered within Chapter 4, and their levels are in Tables D-6 and D-7.

Table 2.4: Health index scoring using conventional method (Khamis et al., 2023).

#	Transformer (Tx) Diagnostic Tests Parameters	W_i Weighting Factors (Relative Importance of each test)	Rating Condition (Green: Very Good, Blue: Good, C: Fair, Orange: Bad, Red: Very Bad)					S_i Scoring Factors (Transformer rating health condition)				
1	DGA	10	A	B	C	D	E	4	3	2	1	0
2	Load History	10	A	B	C	D	E	4	3	2	1	0
3	Power Factor	10	A	B	C	D	E	4	3	2	1	0
4	Infrared thermography	10	A	B	C	D	E	4	3	2	1	0
5	Tx Oil Quality	6	A	B	C	D	E	4	3	2	1	0
6	Overall Tx Condition	8	A	B	C	D	E	4	3	2	1	0
7	Furans Compound Content or Age	5	A	B	C	D	E	4	3	2	1	0
8	Turns ratio	5	A	B	C	D	E	4	3	2	1	0
9	Leakage reactance	8	A	B	C	D	E	4	3	2	1	0
10	Winding resistance	6	A	B	C	D	E	4	3	2	1	0
11	Core-to-ground	2	A	B	C	D	E	4	3	2	1	0
12	Bushing Condition	5	A	B	C	D	E	4	3	2	1	0
13	Main Tank Corrosion	2	A	B	C	D	E	4	3	2	1	0
14	Cooling Equipment	2	A	B	C	D	E	4	3	2	1	0
15	Oil Tank Corrosion	1	A	B	C	D	E	4	3	2	1	0
16	Foundation	1	A	B	C	D	E	4	3	2	1	0
17	Grounding	1	A	B	C	D	E	4	3	2	1	0
18	Gaskets, seals	1	A	B	C	D	E	4	3	2	1	0
19	Connectors conservator tank	1	A	B	C	D	E	4	3	2	1	0
20	Oil Leaks	1	A	B	C	D	E	4	3	2	1	0
21	Oil Level	1	A	B	C	D	E	4	3	2	1	0
22	DGA of LTC	6	A	B	C	D	E	4	3	2	1	0
23	LTC Oil Quality	3	A	B	C	D	E	4	3	2	1	0
24	Overall LTC Condition	5	A	B	C	D	E	4	3	2	1	0
25	Conductivity factor (k_c)	10	A	B	C	D	E	4	3	2	1	0
26	Polarization index (k_p)	10	A	B	C	D	E	4	3	2	1	0
27	Loss factor tg8 at (f = 1mHz)	10	A	B	C	D	E	4	3	2	1	0
28	Others*	3	A	B	C	D	E	4	3	2	1	0

*Vator tank, PT/CT, cable box, manufacture, protection equipment, neutral grounding reactor.

2.5.5 The Role of PTHI in Driving Maintenance Strategy

This section presents the role of PTHI in driving maintenance strategy for electrical utilities. According to (Khamis et al., 2023), the maintenance strategy improves the secure operation and maintenance of the power transformer and cuts its operating expenses. The accomplishment of these two objectives leads to the development and implementation of a good HI as a condition assessment tool for an AM power transformer that serves electrical utilities in any investment decision-making. Improving HI results will require action for the estimated power transformer health condition, which should have accurate scoring and weighting calculations as precisely as possible, such as doctors giving the correct medication to patients. If the assessment of the conditions is incorrect, then the cost savings might disappear, extra costs may arise, and time normalisation, utilising the best labour and resources, may be difficult.

PTHI aims to evaluate transformer health conditions using diagnosis methods to enable effective maintenance strategy before running an illustration about the fleet screening and a detailed condition assessment, which earlier discusses this approach (ABB Service Handbook, 2007; CIGRE Technical brochure 642, 2015). Failure mode and mechanism indicators support the development of detailed condition assessments to drive the daily maintenance and annual plan and update the overall maintenance strategy.

Figure 2.18 presents different maintenance strategy approaches for securing data management, decision-making, and research philosophy for the best practices of asset management implementation.

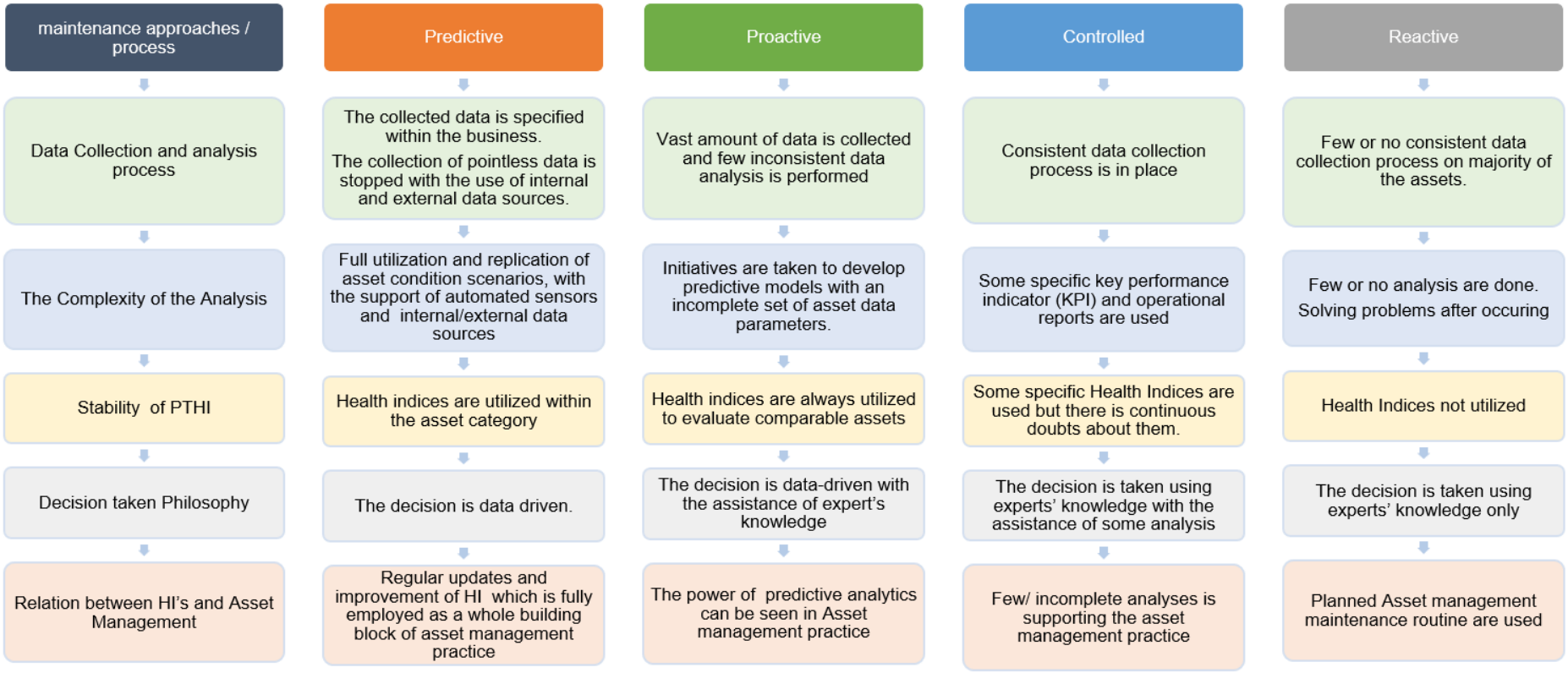


Figure 2.18: Maintenance approaches for data management, decision making, and PTHI research asset management philosophy.

In addition, over the last decades, substantial pressure from the technical regulator, Authority Electricity Regulator (AER), customers, and stakeholders to keep tariffs down and manage unnecessary spending has increased. This challenge requires further prioritising the daily tasks, annual maintenance plan, deferral transformer replacement plans, and lifecycle investment assessment.

Over recent years, the well-established AM-certified electrical utilities such as ISO55000 or Publically Available Specifications (PAS55) have become reliant on AM technologies, including Enterprise Resources Planning (ERP), work orders tracking system, daily information recording system, and health index for managing momentous maintenance jobs which built on the need of achieving planning for the daily, weekly, monthly and annual maintenance plan and maintenance strategy (The Canadian Electricity Association, Deloitte, 2015).

Electrical utilities face power transformer ageing challenges that require substantial replacement costs. Once assets age and need financing for the maintenance strategy, the utilities must equalise these challenging factors by working more innovatively on condition maintenance assessment needs because replacing individual transformer components as single parts might improve the total health of the entire transformer.

Money has become the main driver in assessing the power transformer lifecycle and prioritising replacement plans. Electrical utilities must use various scientific solutions to prioritise their total investment plan based on high replacement needs, deliver a good-quality energy supply, and manage their assets, including prioritising replacement plans, assessing the PT lifecycle, and preventing power transformer failure before it occurs.

However, the above challenges will support the need for change in the asset management process mentioned earlier by (Abbasi and Malik, 2016) and play a critical role in breaking the silo among asset management technical departments, then will improve the HI outcomes for evaluating power transformer health condition, considering the data-driven approach, assessing lifecycle and prioritisation of daily maintenance tasks for the required annual maintenance plan and maintenance strategy.

Leveraging analytics besides a driven approach, asset HI can achieve a reasonable price, risk, and performance basis, improving HI as a reliability-centred maintenance tool for the electrical utility. Technically, the HI can manage the Total Expenditure (TE) of a long-term Master Plan (MP) (Khamis et al., 2023).

Moreover, the maintenance system requires a justified maintenance strategy that improves the secured in-service PT and helps to decrease unnecessary operation and maintenance expenditures. Accomplishing maintenance strategy goals leads to realising the value of a good HI (CIGRE Technical Brochure 761, 2019).

Applying the PTHI ML model in the production stage requires following an agile methodology that contributes to using a continuous retraining strategy (Baumann et al., 2022). This strategy monitors the ML performance continuously and continuously retrains the ML model without compromising the importance of the data validation process to ensure that the PTHI ML model keeps updated with the changes in PTHI according to the maintenance task frequency. This type of management will enable the PTHI to capture the changes in the maintenance task's nature and frequency (Khamis et al., 2023).

In practice, this should apply an automated data pipeline process, as the Maintenance Asset Register should be an online system that is updated regularly to keep track of PT technical details and history information (Hastings, 2021). Kavikondala et al. (2019) stated that the ML model should be retrained automatically on the actual data, including the changes in PTHI. According to (Khamis et al., 2023), the maintenance task frequency should be automated through an automated data pipeline to ensure that the quality of the model is up to date.

Figure 2.19 shows how the PTHL ML model keeps performing well in the production cycle after deployment, even with changes in actual data. This deployment ensures high confidence in data quality, as it should follow a continuous improvement and learning approach according to Asset Management principles (Khamis et al., 2023).

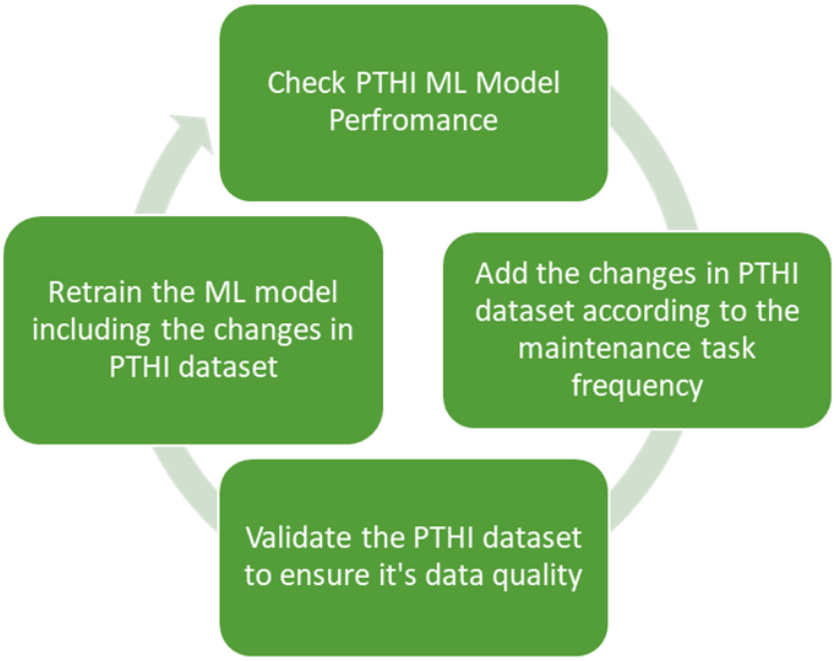


Figure 2.19: Continuous AM improvement learning cycle (Khamis et al., 2024).

Asset management must ensure that asset maintenance drives the annual maintenance plan and strategy through good daily operation and maintenance switching schedules. The maintenance plan and strategy must updated based on PTHI condition results and findings. Nevertheless, AI supports the accuracy of required new maintenance decision data by utilising and solving actual data issues that lead to optimising maintenance strategy.

To conclude, there are no articles journals by academia published regarding applying the asset management framework for PT to enhance maintenance strategy and overcome HI data issues based on short-term annual maintenance plans (Abbasi and Malik, 2016; Khamis et al., 2023). This research will study the health index and overcome data challenges using artificial intelligence and machine learning. This analysis using the ML approach will contribute to the electrical utility data analysis and improve the accuracy of the HI calculation.

2.6 The Research Power Transformer Health Index Approaches

2.6.1 Conventional Health Index Approach

The conventional method depends on the ranking method developed based on the IEC and IEEE scores and weighting factors. The scores and weights agreed by transformer experts (Jahromi et al., 2009; Foros and Istad, 2020) proposed the use of the HI method, which aims to deliver a comprehensive assessment of electrical PT performance, taking into account various factors due to the difference between PTHI development in academia and utilities.

The health index calculation equation developed by Jahromi et al. (2009) quantitatively measures transformer status. A paper published by CIGRE A2_101 (2014) declared that the PTHI is a crucial indicator describing total PT health status. This index provides valuable insight into a PT's well-being and performance by considering various critical factors (Chen et al., 2016). CIGRE Technical Brochure 761, 2019) confirmed that by analysing the index, experts could assess the transformer's reliability, longevity, and efficiency, thus enabling the PTHI to make informed decisions from maintenance, replacement, or repair (Ghoneim and Taha, 2020).

The method remains less costly and easy to use. Still, it has a complicated mathematical scoring and weighting model. It also takes time for the initial detailed condition assessment development, which requires a thorough reading about these international standards to maintain the condition boundaries for each parameter or diagnostic factor and fully requires PT expert judgement (Jahromi et al., 2009; Wang et al., 2022). In this research, the conventional health index model is developed in Excel in Chapter 4 for further details.

2.6.2 Fuzzy Logic Health Index Approach

According to Singh et al. (2013), Fuzzy logic, once considered an obscure mathematical curiosity, is now used in many engineering and scientific works. Fuzzy Logic-Set theory is considered an extension and gross oversimplification in the context of set membership. This set theory allows partial membership with varying degrees of membership. A classical set contains elements that satisfy precise membership

properties, while a fuzzy set contains elements that satisfy imprecise membership properties (Machado et al., 2016). Using a fuzzy rule-based expert system in evaluating intellectual capital assists the asset management team in understanding and assessing the level of each asset created through scholarly activities (Singh et al., 2013).

Fuzzy logic (FL) affirmed its competence to process ambiguity and fuzziness in transformer conditions for 30 transformer datasets (Abu-Elanien et al., 2012). A paper by Chantola et al. (2018) utilised the Fuzzy logic optimisation approach. It stated that it is a very effective method for PT maintenance projection to rank the PT's health condition. Also, according to (Abu-Elanien et al., 2012; Arshad et al., 2014), Fuzzy logic (FL) has been frequently used due to its ability to handle uncertainty and imprecision in transformer condition data.

However, it still has a high dependency on expert judgments. Therefore, based on recent literature reviews, as stated by (Kadim et al., 2018), the AI ML approach solves data issues and overcomes high dependency on transformer experts. In addition, Fuzzy Logic is not a good solution for solving significant data issues as big data and AI have become the most powerful tools with high sensitivity, availability, intelligence and cheapness (Tran et al., 2020) for condition identification to make PTHI model.

Kadim et al. (2018) stated that conventional methods are insufficient in solving significant data issues and not a good choice due to their high dependency on PT experts' knowledge and judgments. Wang et al. (2022) agreed with Kadim et al. (2018), as stated the difference between the conventional traditional method and Fuzzy logic are both not a good choice for extensive data investigation of large fleet of power transformers. In this research, the investigation of the Fuzzy Logic model in MATLAB is available in Chapter 5.

2.6.3 Artificial Intelligent Health Index Approach

According to (Wang et al., 2022; Prasojo et al.,2021; Alqudsi and El-Hag, 2019; Ahmad et al. (2022), in recent years, artificial intelligence has become the best extensive data analysis tool, solving data issues and improving the calculation accuracy of designing and developing an effective and efficient HI. Computer science and data processing technology have gained rapid development. Various algorithms,

including pattern recognition algorithms, were developed for better HI results (Saha et al., 2014). Many artificial intelligence algorithms use neural network algorithms for failure prediction and PT condition monitoring (Islam et al., 2017). Other algorithms were used, including evolving wavelet networks (Huang and Huang, 2002) and SVM (Yi Cui and Saha, 2013).

Based on several journal publications over the last few years have demonstrated the effectiveness of various machine learning and artificial intelligence for classifying the health index of power transformers and overcoming HI issues such as high reliance on PT's expert judgement or missing data based on (Pereira et al., 2009; Kadim et al., 2018; Tang and Wu, 2011; Zuo et al., 2016; Rediansyah et al., 2021).

In more detail, Zuo et al. (2016) stated that Binary Logistic Regression (BLR) is effective and reliable in assessing PT health state and has been generally used in many research areas, including medical and science, as it has the benefit that it provides a straight evaluation of the class possibilities in comparison to support vector machines and ANN. In addition, Support vector machines (SVM), as demonstrated by (Alqudsi and El-Hag (2019) and Rediansyah et al. (2021), have been recognised for their durability while processing high-dimensional data as well as their efficacy in recognising complex borders between health problems. Also, Random Forests (RF) and Decision Trees (DT) are vital algorithms for their interpretability and competency in managing non-linear relationships and feature interactions, with notable contributions (Alqudsi and El-Hag, 2019; Ahmad et al., 2022).

Furthermore, according to Mitchell (2015), Naive Bayes (NB) is respected for its simplicity and computational efficiency, particularly in resource-limited circumstances. Moreover, K-Nearest Neighbours (KNN) is known for its simplicity and effectiveness in capturing local data structures, making it suitable for finding health patterns based on historical data, as testified by Benhmed et al. (2018).

Together, these techniques provide a comprehensive toolkit for dealing with the complexities of PTHI classification. This technique reflects the wide range of capabilities required to manage uncertainty, non-linear relationships, interpretability, and computational efficiency in this domain.

Adding several calculated parameters for the PT subsystems allows assessing the health condition of all in-service or scraped transformers. Due to the large data size that needs detailed analysis, AI for PTHI is essential for a large fleet of power transformers. AIHI is used to assess the condition of faulty transformers in detail. Therefore, ML studies require sufficient data to build models (Arbabshirani et al. 2017).

Of course, using the most available diagnostic test parameters will distinguish the health condition of the transformer's subcomponents and map the conditions and actions into numerical scores, which present the knowledge gained about power transformers' condition (Gorgan et al.,2010).

Using the PTHI to assess the health condition of PT provides an acceptable estimation of the residual life considering the vital features of the transformer insulation system (Tanasescu et al., 2016). According to (Xian et al., 2018), the dielectric losses within the electro-insulating system, including paper polarisation and conductivity factor (Singh and Swanson, 2018), affect the PTHI firmly. Therefore, combining and utilising all relevant data from maintenance is crucial, strategically, and technically economical at low cost and high performance. The PTHI contribution is improving electrical utility practice, and the integration with AM makes the HI more valuable and effective in providing efficient decision-making for a diverse scenario for the categorised assets. In addition, the correct practical direction of the PTHI framework should be proposed based on the research's detailed investigation and supportive insights from industry practitioners (Kinnunen et al., 2016). The PTHI investigation will be studied based on asset life cycle principles and methodology (PASS 55,2008; ISO 55000 series, 2014).

In addition, electrical utilities can combine the direct data, indicator effect, and risk matrix model to calculate the effectiveness of failure modes/mechanisms. Based on the efficacy of failure mode/mechanisms, the usefulness life remaining is estimated (CIGRE Technical Brochure 642, 2015; CIGRE Technical Brochure 659, 2016; CIGRE Working Group 05, 1983).

Other AM system operators consider combining FSA and DA methods, where the basic FSA leads to the first readily available information. Still, the combination may also use intelligent algorithms and additional data when available. However, before

making a significant decision, this ML detailed condition assessment method (PTHI) can yield a better result than the basic FSA of HI and the conventional detailed HI evaluation. Therefore, this research will study and explore the exceptional capability of using AI in developing PTHI.

The influence of ML HI modelling and investigation ensures that the outputs significantly contribute to early PT condition prediction in electrical utilities and consider the optimum decision at the required time frame by solving data issues and adding a better value to the AM lifecycle assessment and replacement plan decisions.

Chapters 7 and 8 of this research discuss further examination and investigation of using AI and ML HI models in R programming and Python.

2.7 Conclusion

In conclusion, based on earlier findings from the literature review, the health index approach and implementation are game-changing. The PTHI depend on the developer, researcher, and electrical utility agreeing on scoring and weighting factors within a specific country or region and considering power transformer expert judgments. Technically, the actual data plays a critical role in investigating PT health conditions and improving asset management decision-making.

The literature review shows a need to change the asset management process, and the research investigation uses electrical engineering field experience with the strategic asset management standard framework. This experience and knowledge would cover the required PTHI development and ensure a data-driven approach is in place.

This section summarises the literature review, highlighting the key points informing the drivers for establishing a valuable and auditable PTHI, considering and utilising machine learning to develop the HI model for PT health condition investigation, as presented in chapters 6 and 7. It is also developing the technical and financial analysis to demonstrate the technical and financial impacts of this research, which was introduced in Chapter 8 by creating a costing model analysis and the web application that supports the site engineers to determine the actual health condition of power transformers to consider a further detail assessment while doing site inspection visits.

Ultimately, this chapter's literature review findings identified the gap in knowledge and understanding that underscores the necessity of this research for developing PTHI that integrated with the AM standard framework.

Section 2.3 discussed many practical and technical details to enable a good understanding and comprehensive information of different PT active components as the cause of any failure modes and deterioration mechanism for in-service power transformers.

Expert knowledge and speciality of PT engineers' complaints become crucial for optimum PT asset management conclusions. Without this knowledge, skills, and speciality, it is impossible to comply with minor transformer health condition investigations that include studying the required actions for replacement within the required time frame, repair, or time-based maintenance decisions. This knowledge is essential for establishing reasonable condition assessment criteria using efficient and reliable HI tools. HI applies input parameters to the main prominent ageing mechanisms and failure modes, which can positively impact in-service PT performance. This knowledge and understanding are crucial before examining and developing HI.

The PTHI technique comprehensively evaluates transformer health, considering various parameters, including moisture, DGA, oil quality, temperature, frequency response analysis, and partial discharge factors.

This research improved the diagnostic measurement list from earlier literature reviews. Also, it is essential to remember that relying solely on the HI approach could not uncover the full extent caused by PT failure.

Combining the health index approach with other risk assessment techniques that use risk analysis and maintenance strategies is crucial and described in exemplary asset management practice and implementation. This combination of methods presents a dependable and accurate evaluation of PT's health condition, empowering timely decision-making in transformer operation and maintenance and reducing the involvement of power transformer specialists.

Good HI is about a good understanding of the AM framework, international best practices, good field experience and knowledge about electrical power transformers, reasonable data interpretations based on related standards for using good actual data and assigning a mix of background engineers (O and M, AM, Projects, and Control) for the HI implementation process.

There are various power transformer condition assessments. FSA is a fast and practical approach to assessing the conditions of large fleets of power transformers. Detailed condition assessment depends on the quality of the maintenance records calculated using the PTHI value, which represents individual PT health conditions.

The PTHI approach enables AM to identify faulty power transformers and ensure better risk decision-making within the required time. In addition, AI has become the key player in developing effective and efficient HI. This ML approach presents an alternative super method to estimate the values for the missing data uncertainty and detect PT health conditions.

The chapter also focused on identifying the latest developments in PTHI and using AI, ML, and algorithm coding. This development will support short-term daily maintenance activities and issues that can solve long-term PT health condition planning.

Based on the detailed literature review analysis, this research considers a thorough condition assessment for investigating power transformer health conditions, utilising the conventional Fuzzy Logic and AI ML algorithms for HI model development.

The need for more research implies a change in PTHI's asset management process, as detailed in section 2.2, based on the findings from the literature reviews.

The following section addresses the knowledge gap by utilising the asset management standard framework to provide context and application for developing good asset management PTHI models to drive electrical utilities' annual maintenance plans and strategies.

2.7.1 Gaps in Knowledge

The literature review in this research has identified in Chapter 2 that many researchers in the field of electrical utilities PTHI lack a focus on practical experience, especially in academia, which differs from practice. The importance of considering the facts of the electrical field experience and the knowledge about the strategic asset management framework implies the change in the PTHI development and overlooks the related improvement.

Furthermore, the literature review lacks the implementation of a combination of extensive data statistical evaluation, technical health condition evaluation impact, and financial analysis impact, considering the AM framework by finding optimal value for prioritising replacement plans and assessing life cycles and data-driven decisions. Integrating the analysis of these combinations by producing the PTHI web app to serve the on-site maintenance engineers and add future investment within the long-term master plan.

Consequently, the following points summarise the identified gap in the literature:

- There is a lack of integration of the fieldwork experience, interdependencies, strategic AM business decisions, and risk management connections as it is a very complex task that underlying these uncertainties has not been explored among researchers.
- The existing imbalanced data and data uncertainty issues have been a primary challenge in developing HI by academia and practice, as the data management and handling process for extensive data portfolios is vital and considered in this research.
- The existing literature lacks an AI data-driven portable PTHI app that speeds up the determination of the health condition of transformers during site health inspections.
- There is no health index considering the short-term planning for the annual maintenance plan, which is why this research intends to focus on the annual maintenance plan and prioritise daily operations and maintenance activities by knowing the immediate health decisions driven by the health condition. Therefore,

short-term daily maintenance activities are of interest to this research as these form a permanent solution to future problems.

This research will now address the above gap in the research area by:

- The HI model accuracy and data management process review through the development and implementation methodology using the asset management standard framework and field experience. This research will impact the utility practice by implementing AM framework that breaks the silo between the technical department to improve the PTHI data. This study will enable a robust HI that requires electrical utilities to integrate the HI with the AM standard framework ISO55000 through an exemplary implementation of the maintenance strategy, which a good HI must drive for the annual maintenance plan. Usually, the data silo is due to resistance of some departments based on the data being distributed in different technical departments and by implementing good asset management practices and culture, the data becomes an available asset to the whole business. This research will focus on implementing performance evaluation matrices, risk metrics, confusion matrices, and exploratory data analysis to explain and discover the data correlation and compare this research model's accuracy with other researchers.

- This research will solve the PTHI real data imbalance issue using computer and data science techniques. These techniques are required to ensure the data is balanced with high confidence, leading to more accurate decisions. The decision must be correct and integrated with cost-effectiveness and optimal investment value. This achievement will support managing the operation and maintenance spending when it becomes higher than the initial installation cost of the related transformer. The actual data is balanced among transformer health condition classes as data correctness plays a critical role in the algorithms to identify each class equally and correctly quickly. Also, various maintenance planning scenarios will be integrated with a detailed financial analysis, enabling a better maintenance plan and strategy.

- This research will also provide a PTHI web app that will show the technical impact of this research. This tool verifies the transformer's health condition

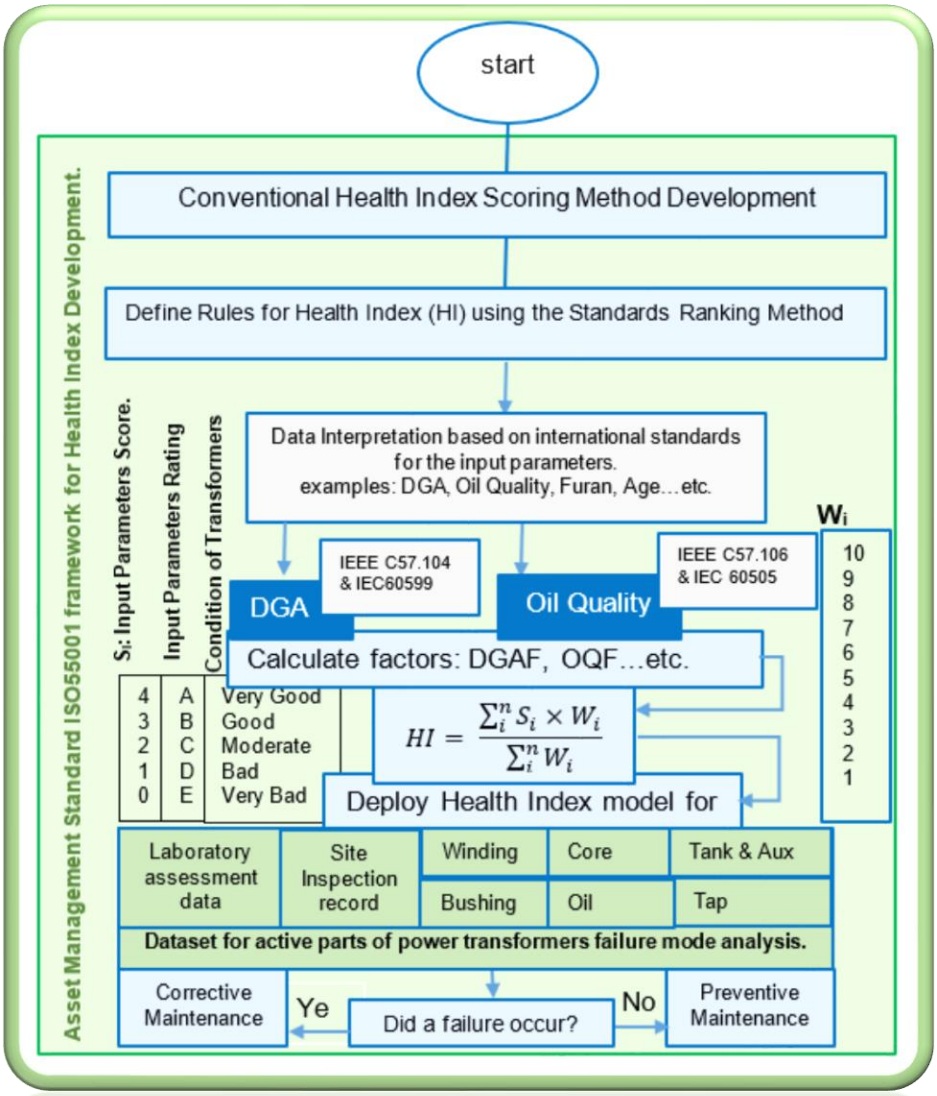
while at the site for further detailed health index condition assessment. Then, to be included in the required actions and manage the unplanned outages (risks) within the annual maintenance plan for driving the electrical utility maintenance strategy.

- This research will appraise transformer health conditions to prioritise daily operation and maintenance switching schedules by knowing which transformers need immediate health decisions. The study aims to solve today's operation and maintenance problems, leading to solving future asset lifecycle long-term problems rather than estimating long-term asset lifecycle condition issues, as widely studied by scholars.

The next chapter will present the research methodology and a detailed analysis of the steps of developing PTHI models for electrical power transformers.

Chapter 3

Research Methodology



Chapter 3: Research Methodology

3.1 Introduction

The last chapter defined the research findings and the recent developments in the research area from several works of literature for the way forward. As a result, it has narrowed the research questions and added the required knowledge for the data collection process, data processing and modelling. Moreover, learning asset management standards (ISO55000 series) and other international best practices has supported the research methodology through better data interpretation and developing a new method considering the ISO55000 framework for the health index electrical PT condition assessment. Furthermore, this has created a difference between this research outcome and other researchers within the same research area. Understanding the difference between practical experience and academic learning has added significant value to the research outcomes. Combining knowledge and experience has made the required difference and contributed to academia by adding AM practical knowledge and experience to the research methodology, inspiring future research and innovation in the field.

This chapter presents the research methodology promoted for the electrical PTHI investigation. It includes aims, methods, approaches, milestones, and presenting considered data collection and analysis procedures.

3.2 Research Aim

The research aims to develop a PTHI tool that integrates with the role of the asset management standard ISO55000 framework. This AM framework is a vital factor for in-service PT performance and reliability, especially since AM is about quality of service. Furthermore, according to ISO 3001, which is related to condition risk-based decision-making, the PTHI model as a health condition evaluation tool must consider risk-based decisions.

3.3 Research Objectives

In order to accomplish the research aim, the below objectives give the required answers to the research questions:

1. Study the PT system and its main subcomponents that impact the unit's failure.
2. Investigate ageing mechanisms and failure modes within the in-service operation of electrical power transformers.
3. Develop a conventional health index model as the primary data correction and validation platform.
4. Collect, analyse, and prepare proper PT assessment data for modelling based on international standards and utility practices.
5. Examine and recognise PTHI to ensure maintenance decisions are considered and will not negatively impact the annual cost budget.
6. Embrace the outstanding capability of ML algorithms to identify faulty transformers in a large fleet of PT records by developing ML health index models for robust HI decision-making.
    a. Develop effective dataset decision-making driven AI prognosis and adaptive learning control algorithms for accommodating vital PTHI app that uses failure modes to identify PT failure conditions.
    b. Validate the algorithms by benchmarking the model results with other researchers.
    c. Validate the algorithms by using actual data from the related industry.
7. Determine the relationship between asset management standard ISO55000 framework and PT health condition modelling to improve power transformers performance for health index design, consider the availability of the correct actual dataset and the ISO5500 framework is implemented in daily practice by the electrical utility.

3.4 Research Approach

The research applied the literature review findings alongside past electrical engineering experience in AM, control, planning, design, project, daily operation, and maintenance schedules. This research supported the academic analysis during the interview conducted for data collection presented in Chapter Two, Table 2.2, while implementing the past AM knowledge and working experience in the electricity and water sectors. That includes conducting several experiments on developing HI models using conventional and ML algorithms. This approach improves the research area

findings and adds further novelty and contribution. Implementing conventional, fuzzy logic and AI in developing PTHI is insufficient without considering the AM ISO 55001 standard framework since the AM covers the whole business. Consequently, implementing this approach calls for additional study and research on finding faulty transformers in a large fleet of transformers within an electrical utility system (Rediansyah et al., 2021). This work requires using actual data collected from the electrical utility and scientifically studied.

3.5 Research Framework

The excellent experience of the AM standard ISO55000 framework is being explored and used within the research framework, as discussed in Figure 3.1 and where applicable. The framework supports the research's aim to improve operation and maintenance decision-making by developing the required HI, which is considered a building block from the big picture of the AM.

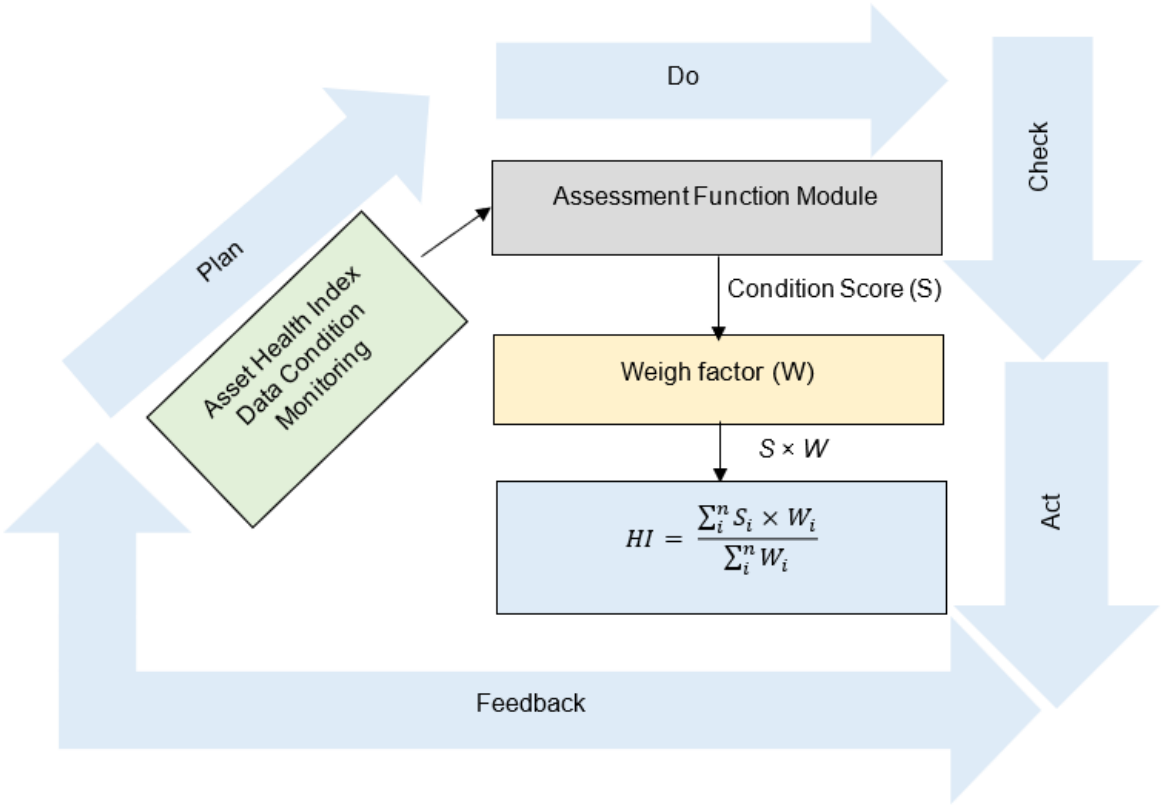


Figure 3.1: HI and AM using the Plan, Do, Check and Act (PDCA) concept.

The research framework is based on the research methodology described in the flowchart of Figure 3.2, which is based on the HI conventional method and explained in further detail in Chapter 4.

1. Develop the Asset Management HI Excel worksheet analysis model as a platform for the algorithm's simulation purposes.
2. Simulate the platform and its findings using AI available tools for effective maintenance data-driven real-time condition monitoring decision-making.
3. Develop a new technique that serves robust maintenance investment planning for the electrical industry using PTHI.
4. Implement the maintenance whole life cycle cost analysis for the HI to ensure optimum investment planning integration of OPEX and CAPEX.
5. Establish a Key Performance Indicator (KPI) to ensure the success of the research for maintaining the required quality, reliability, and robust maintenance investment services.
6. Ensure the quality of services of the HI is in line with the international benchmark models, best practices, and standards, including IEEE, IEC and CIGRE.

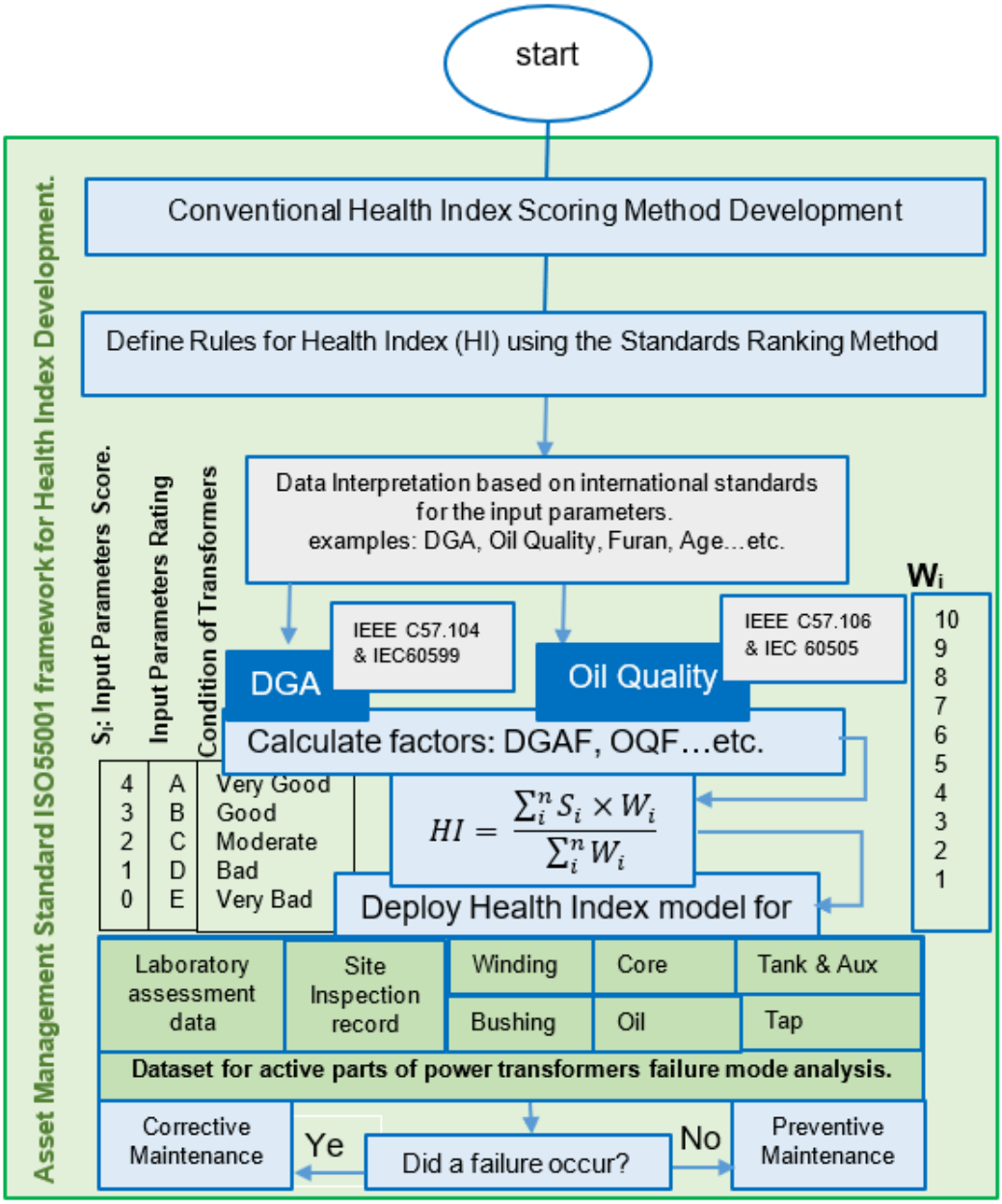


Figure 3.2: Health Index methodology (Khamis et al., 2023).

3.6 Data Management Process and Challenges

This section discusses electrical utilities' data collection, analysis, and preparation methodologies. It is the most challenging section because the research model results rely on how the data is processed and handled. This research requires linking much information, international best practices, and standards together in a simple and friendly way that covers using international best practices, processes and methodology, which manages the data collection, analysis, validation, preparation, and data availability from the local utility. This data management requires skills, understanding, and knowledge for data collection, analysis, and preparation methods based on anomaly failure detection that complies with ML algorithms, which is the hardest part.

3.6.1 Data Collection Phase

The data collection process followed during the research is through the literature review, face-to-face interviews with the local electrical utility asset management team, and follow-up correspondence.

It is essential to identify the required maintenance data/information/knowledge as defined in the audit commissioning report (Henke et al., 2016), as shown in Table 3.1, before using the required process, interviews, awareness, and training needs for the “Development Power Transformer Asset Management Health Index Using AI”.

Table 3.1: The terms ‘data’, ‘information’ and ‘knowledge’ (Henke et al., 2016).

Data	Data are numbers, words or images that have yet to be organised or analysed to answer a specific question.
Information	Produced through processing, manipulating, and organising data to answer questions, adding to the knowledge of the receiver.
Knowledge	What is known by a person or persons? Involves interpreting information received, adding relevance and context to clarify the insights the information contains.

3.6.1.1 Data Source

The data for power transformer condition assessments are from technical studies and inspections within any electrical utility's operation and maintenance department, except loading data from the control centre's Supervisory Control And Data Acquisition (SCADA) system within the electrical utility. The PTHI input data must be actual data because it will evaluate the health condition of the related assets, conduct a lifecycle investment costing model, and enable technical decision-making for the required annual maintenance actions. These inputs of actual data from different resources are laboratory tests, online monitoring systems (SCADA), conditional boundaries, criteria based on IEEE standards, IEC Coding Standard, CIGRE research investigation and findings, regular inspections, and maintenance historical records.

Figure 3.3 summarises the above findings and shows the data types from various electrical utilities data sources relying on the literature review findings in section 2.4.1, which can benefit this research investigation.

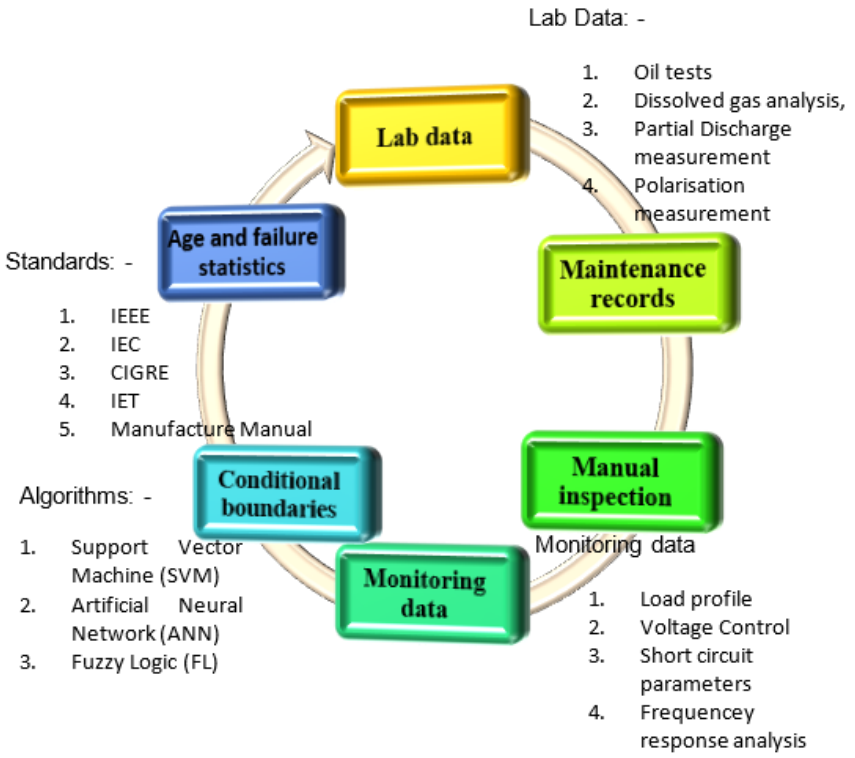


Figure 3.3: Data sources for Health Index assessment (Khamis et al.,2024).

3.6.1.2 Inputs Data

The PT data inputs must include data linked with oil characteristic tests, DGA, measurement of partial discharge, measurement of polarisation, frequency response analysis, maintenance records, load profile, parameters for measuring short circuit withstand capability, lighting withstand capability, age and failure statistics as shown in Tables F-1 and F-2 in Appendix F (CIGRE Technical Brochure 761, 2019).

3.6.1.3 Test Plans

PT test plans vary from one utility to another, as shown in Appendix D, Table D-11, but the conventional, perfectly recognised tests, according to (Raj, 2019; Wattakapaiboon and Pattanadech, 2016; Hernanda et al., 2014; Naderian et al., 2008; Khamis et al., 2023), which are effective as diagnostic practices incorporate :

- DGA.
- Tap changer tests and inspection include motor current, contact resistance, and oil tests.
- Bushing tests that cover capacitance.
- Insulation resistance (Megger)
- Oil quality tests
- Routine visual inspection of tank, radiator, fan, and bushing.
- Dissipation factor test.
- Capacitance assessment.
- Infrared thermography test.
- Turns-ratio test.
- Transformer winding resistance test.
- Leakage reactance test.
- Excitation current test and core-to-ground test.

Additional tests that are utilised by a few utilities still, possibly, will not remain a widespread practice by the other utilises, which incorporate:

- Frequency response analysis
- Furan analysis
- Partial discharge assessment.

Further than these tests, a few practical data are commonly obtainable to assess the long-term transformer requirements using load and maintenance history records. Numerous tests and visual inspections accomplish preserving the transformer state and increasing fundamental transformer mechanisms' life cycle, as shown in Table 3.2.

Table 3.2: PT subsystem and testing (Leauprasert et al., 2020).

Subsystems	Measurement
Active Part; Winding and Core	Turn ratio test, DC winding resistance test, short circuit impedance test, capacitance test
Bushing	Capacitance test, tan delta test
DGA	CO_2, C_2H_4, C_2H_2, C_2H_6, CH_4, CO, C_3H_6, C_3H_8, H_2, O_2, N_2
Oil Insulation	Oil dielectric strength test, tan delta test
Paper Insulation	Polarisation index test
Surge Arrester	Capacitance test, tan delta test, watt loss test

3.6.2 Dataset Analysis and Preparation Phase

The research utilised three datasets, which consisted of:

- Small dataset, 3 and15 transformers (Literature Review Data),
- Medium dataset 30 transformer records (Literature Review Data).
- Extensive dataset (Actual Electrical Utility Data)

The classification of six input parameters includes 2-Furfuraldehyde, BDV, DCG, Water, Dissipation factor (DF) and Acidity to predict PT health condition. The output of health index analysis for the PT failure modes presents five categories: “Very Bad”, “Bad”, “Moderate”, “Good”, and “Very Good”. The “Good” and “Very Good” classes represent a positive transformer health state. “Moderate” class is a representation of a condition that needs caution. In contrast, the “Bad” and “Very Bad” represent a “negative” transformer health state, indicating that actions are immediately required. Figure 3.4 expresses the classification process proposed by the paper (Győrödi et al., 2016).

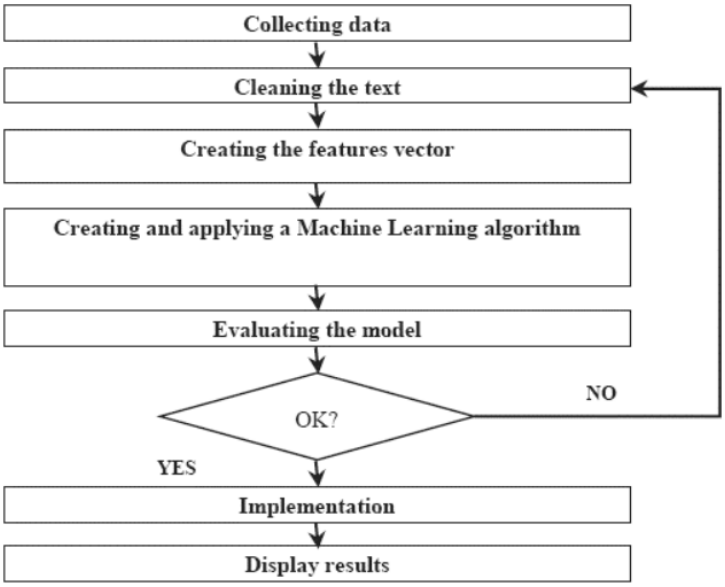


Figure 3.4: Workflow to achieve the classification of HI Condition (Győrödi et al., 2016).

During the analysis phase, the PT data will have five main elements (The Canadian Electricity Association, Deloitte, 2015):

1. PT identification (ID) classification
2. Health Circumstances
3. Performance of in-service PT
4. Breakdown causes
5. Importance based on unit risk information

These above five sections shape the analysis parameters to determine the PTHI scores.

Further innovative asset HI calculations use the data elements in several ways and comprise multifaceted data analytics and curving outcomes, conceivably leveraging AI and ML investigation and other innovative practices to reach the asset HI score with the required decision within the required time response.

Data analytics specialists study PTHI model trends to identify breakdown causes before PT failure occurs and enable the consideration of the required decision. The breakdown causes differ based on PT type, area, performance, and importance. According to (CIGRE Technical Brochure 761, 2019), the details of HI calculation score methods are presented in Appendix E, Table E-4, as it shows the data elements

introduced within some electrical companies. This finding is the primary source for calculating PTHI scores for PT failure modes. The individual transformer parts are targeted for repair and replacement, while the whole PT will continue in service after the required action. PT subcomponent can impact the total health condition of PT; hence, this is clearly shown in PTHI scores to consider the required decision.

3.6.2.1 Health Index Data Outputs

These results were obtained from the scoring methods and weighting factors from the HI calculation. The score determines the remaining life of the assets as the score describes the actual PT health condition to make the right decision at the right time. The results depend on the data interpretation process and analysis techniques collected from the electrical utility.

The value of the PTHI calculated score drives the data accuracy, design, wholeness, and timelines of the data (Hughes, 2003). The calculated HI score from the collected electrical utility data requires good data quality, scoring, and weighting information to evaluate the transformer.

In most situations, the data presented for scoring the transformer are not current, as some are from the literature review and others from the electrical utility, which is more than three to five years old.

The diagnostic tests consider PT distribution over various ages for large transformer fleets. The impact of this missing or incorrect data on the score must be understood. In some situations, this impact can be insignificant to the total score. In other situations, it might show an unhealthy transformer looking healthy, causing it to be unnoticed for severe repair or replacement.

A transformer with a decent calculated score founded on incorrect data can have a severe issue. Likewise, missing or inaccurate data can make a healthy transformer look unhealthy.

Executing a “quality score” or some other suggestion of the quality of the implemented data (typically called “significance”) can assist in knowing the quality of data and the reliability or confidence level of a score, as shown in Appendix F, Table F-4.

3.6.2.2 Data Measuring Critical Factors

Each diagnostic parameter factor is either a direct or calculated measure, depending on several diagnostic tests and how data are collected. The data is categorised as follows:

- Transformer oil tank: direct measurement
- PT external and internal strains, the worked-out measures based on PT's various sub-sections of data calculation.
- Area and climate condition maps present the data, which are worked out and collected from a geographical information system (GIS). It includes the previous area climate records that have impacted overall PT performance.

The PTHI calculation uses traditional methods and ML, depending on a steady range between 0 and 1, 1 and 10, or 1 and 100 (CIGRE A2_101, 2014).

3.6.2.3 Determining Relative Importance Using Specialist Weighting Factors

After the primary analysis of the critical factors that impact PT performance, it is necessary to find the relative importance of the direct and indirect measures in representing PT health conditions. According to CIGRE A2_101 (2014), failure modes and root cause analysis will identify the key factors contributing to and representing the relative importance. Relative importance assessment follows the process of finding the weighting factors. The first factors represent the direct measures that include the individual parameters of the condition of the PT. Meanwhile, the PT potential health condition indicators represent the main factors. Therefore, using the visual inspection data, which means the direct condition of PT, helps avoid using diagnostic tests that present indicator factors. The indirect measurements became in use once the direct measures were impossible to collect through site inspections and regular maintenance performance, which the PT maintenance engineers and specialists classified as more reliable.

Direct indicators are input parameters that do not require direct interaction using PT active parts; this includes station area, age, soil, and temperature data. The indirect indicators are input parameters that present the warning indicator about a PT's health condition through diagnostic tests. This cause means the old PT is not always the same as the new PT; its age can cause the worst PT health condition due to the

mechanism deterioration of PT parts. This cause does not always mean this assumption is reliable for all individual PTs in a large fleet. In this condition, the indirect measurements will be less reliable than direct measurements. Hence, this information becomes part of the rational reasoning mentioned in the earlier section for improving PTHI data quality and corrective maintenance.

3.6.2.4 Data Condition Function

The data condition, function, and PTHI analytical performance can extend the asset lifetime approximations using the statistical analysis and evaluation method. This function depends on the data inputs that occur through inspections, diagnostic testing and maintenance, and this method provides the most accurate and latest outcomes of the three analysis and evaluation methods (CIGRE A2_101, 2014). The condition function operates by correcting the statistical approximation of residual lifetime reliant on the condition of an asset. Assuming condition analysis and evaluation data expose that the asset is in good condition, the approximation residual lifetime is extended depending on the group of predefined rules. This approximation means the estimated residual life reduction is correct if the condition analysis and evaluation methods expose a poor asset condition.

Furthermore, there is a likelihood that the data condition exposes an undesirable condition for PT (CIGRE A2_101, 2014). An alert will generate a proper indicator to sign the “Bad” transformer for urgent replacement or maintenance. The change in the estimated residual lifetime is contingent on the PT condition and whether the condition level is good or bad. In this HI model, this is managed based on a set of predefined rules. Therefore, the data obtained from the site inspections, maintenance, and diagnostics equation is structured and grouped correspondingly with the threshold values.

Different methods propose building the required model, applying conventional score-weighting mathematical and artificial intelligence methods using artificial neural network algorithms (Islam et al., 2018). Figure 3.5 shows how ML algorithms, as a subset of artificial intelligence methods integrated with PT maintenance data, can be employed for an extensive dataset of a large fleet of power transformers to enable data mining and drive insights about PT health conditions.

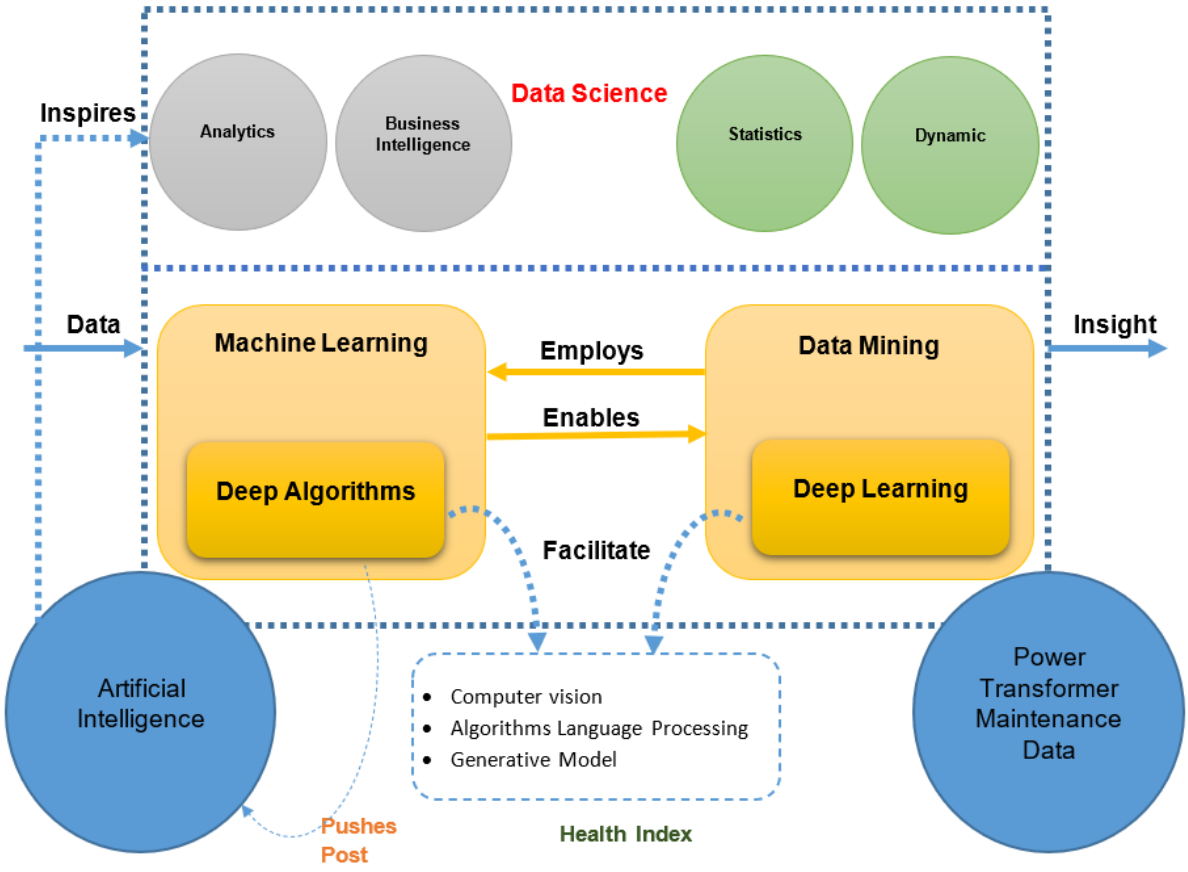


Figure 3.5: Data science research model (The data science puzzle revisited, 2017).

3.7 Research Data Model

The research proposed AI and ML for data preparation and modelling. Two tactics promote robust PTHI development. The first tactic builds a conventional predictable scoring technique, with PTHI calculation using scoring and weighting algorithms. The second tactic builds PTHI calculation using an ML approach.

ML includes the implementation of multiple algorithms that can integrate the AM framework with the robust PTHI model, which initially uses the preliminary design of the conventional model and the fuzzy logic model, which involves a specific setting of input/output data for data preparation. However, (Kadim et al., 2018; Wang et al. (2022); Jin et al., 2022) considered ML more suitable for handling extensive data analysis. This analysis is achieved by solving data issues and reducing the independence of PT expert judgements for further HI investigation that requires detailed algorithms data training intention. This focus is an essential data

management process for developing the PTHI model using the ML approach and AM framework best practice.

3.8 Summary

This research employs a triangulation methodology, combining qualitative, descriptive, and quantitative studies (Faryadi, 2019). According to Thompson (2004), using a mixture of methodologies is more scientific and beneficial.

Multiple approaches will be used, including a detailed descriptive statistical analysis of the dataset collected (Al-Raqqad, 2017) from the structured interview with the electrical utility staff and a quantitative analysis of the dataset and the HI models experiments result. It will include the knowledge using the observation from previous researcher working experience in various electrical utility departments (design, projects, planning, and asset management) to demonstrate the knowledge obtained from understanding the data and the benefits of developing the PTHI model in driving PT asset management maintenance strategy.

This chapter also discussed the research aims and objectives. It also discussed the research method followed to develop the HI model and the asset management ISO 550000 framework applied to break the silo for HI data requirement and importance. The research became more interesting after collecting the actual data, and then it went into a detailed ML investigation, mainly benefiting from the PTHI model. This development in the Excel platform initially uses the literature review data and then electrical utility data. Comparations with other international research students improve the accuracy of the PTHI model. This research shows a positive impact through using electrical utility datasets to investigate PT health conditions for the large fleet of power transformers. The big data transition is required to become fully available for the research asset management PTHI, flex dashboard, and PTHI web application assessing PT health condition.

The data management and analysis stated in Appendix D, Tables D-1, D-2, and D-3 are demonstrated very well during this study, which covers data collection, preparation, validation, analysis, and PTHI findings. The investigation included 1231 data records from the local electrical utility, as discussed in Chapter 8. Considering all voltage level records, the data were classified based on transformer voltage level

using relevant IEC and IEEE standards for the related diagnostics parameters.

The data quality is considered the main challenge for a better outcome from the HI calculation scoring methods. This PTHI also relies on the analytics functions, including the weighting factors. These evaluation techniques rely on accurate service data, which process PTHI into a score that describes PT's overall condition to support a quick response. Then, the HI scores will provide the needed practical decision for the intervention: what and how soon, whether a maintenance or refurbishment or planned replacement, to keep running the momentum of the PT whole life cycle cost-effective and ensure the required future performance by each PT is continuously achieved and managed. The most important features influencing PTHI calculation methods are:

- Failure modes.
- Diagnostic techniques.
- User’s experience.
- Condition monitoring data.

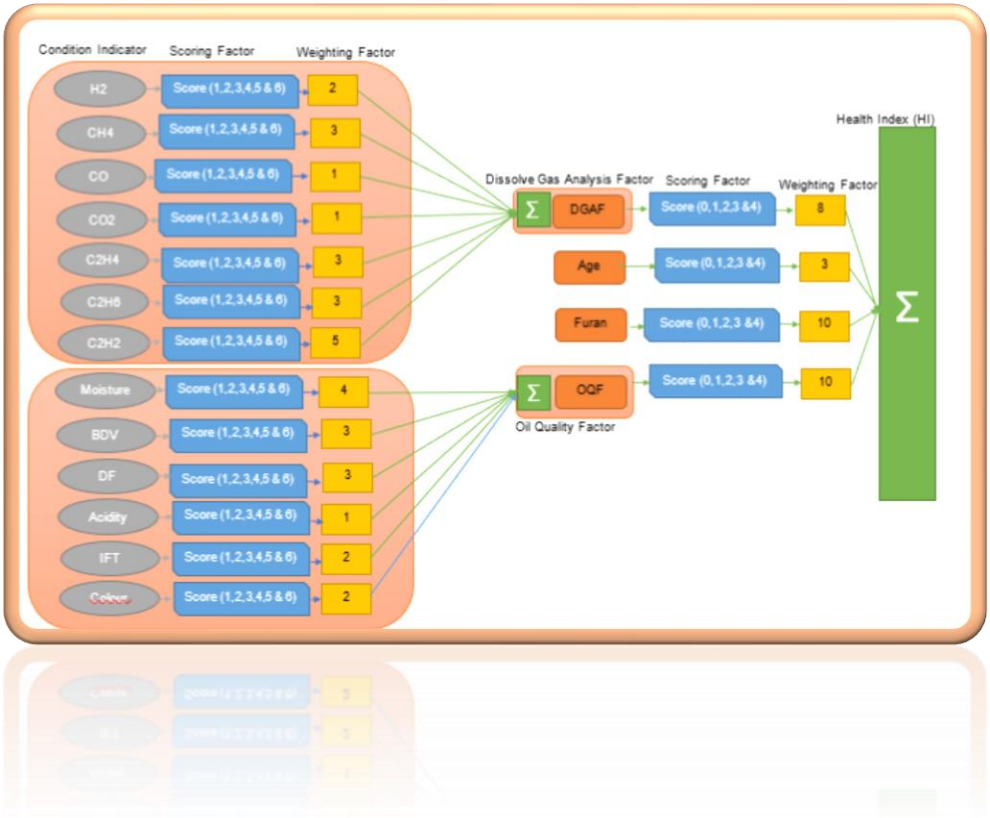
All the information sources led to the importance of master database development to improve long-term data reliability and accuracy, which requires frequent updates by a dedicated field team. The HI is considered a good AM tool despite the differences among electrical utilities. The electrical utilities operate in-service PT effectively and efficiently and postpone the replacement plan of old PT through the decision-making that must continuously seek ways to extend the lifetime of PT and defer unnecessary costs. However, PTS must not be in operation to the point where the transformers begin to affect the energy supply or possibly be a threat to the environment.

The HI method is critical as an AM tool for electrical utilities and complex industrial automation processes. Although it can be time-consuming and costly initially, the method's payback makes significant savings on long-term master planning investments.

The next chapter will utilise the knowledge and data obtained from the literature and implement the formulation of conventional HI scoring as a starting point for transformer condition assessment. It will further compare their PTHI output result with this research output results and findings in section 4.4.3.3.

Chapter 4

Case Study 1 Conventional Health Index Approach



Chapter 4: Case Study 1 Conventional Health Index Approach

4.1 Introduction

Power transformers are high-priced equipment in any grid station. Evaluating its actual health condition is an exciting subject for most utilities, especially since there is a growing need to enhance optimal financial and technical operation. To accomplish the best possible balance for price control, master plan budget between operation and maintenance costs, assess lifecycle asset performance reinforcements and deliver economic and technical justifications for reliable engineering evaluations and capital replacement plans.

To assess the condition and life expectancy of PT, various measurement and monitoring parameters might be considered including DGA, moisture, oil quality, partial discharge, Partial Discharge (PD), Frequency Response Analysis (FRA), recovery voltage records, Tap Changer (TC) tests, PT bushing tests, Furfural, Power Factor (PF), Load Factor (LF), DP and visual inspection maintenance (Naderian et al., 2008; Jahromi et al.,2009; CIGRE Technical Brochure 761, 2019). These factors are available in Appendix F, Tables F-5, F-6 and F-7. Utilities may carry out some of these tests regularly to assess the condition of the power transformers.

HI is an effective method to create a judged capital and maintenance plan, relying on various results collected from laboratory tests and field inspection tests to analyse the condition of a large fleet of transformers. However, formulation of HI that carries all the factors could lead to unreasonable facts for various situations due to data unavailability and some utility decisions not to perform some tests on in-service transformers due to many reasons, including being expensive, time-consuming, lack of experience force to long term planning outage for in-service transformer which effects badly into the energy delivery commitment. Also, several essential testing parameters are ambiguous and unclear or understood by these utilities, so best practices and international standards recognise these tests.

The conventional HI scoring method could help start transformer condition analysis and strategic maintenance planning for transformers as it could reduce operational costs and can be adapted. This chapter includes the formulation of the HI scoring method based on the collected data from the literature review that described and

compared the HI output result with this research output results and findings in section 6.4.3.3.

Four main input parameters, DGA, oil-immersed paper life, auxiliary component condition, and electrical diagnostic tests, are collected from the literature review articles, as shown in Appendix D Table D-12, and are considered the basis of transformer HI formulation and assessment.

This research concluded with 28 inputs weighted according to importance from the above research, 20 inputs by Naderian et al. (2008) and Jahromi et al. (2009). The Furan test has less weighting factor than DGA, PF, and load records, as shown in Chapter 2, Table 2.4. The PTHI considers a very healthy transformer condition as 100%. It further relates the PT residual lifetime necessities judged in PTHI for maintenance or replacement.

Kadim et al. (2018) used the Naderian et al. (2008) approach for the HI scoring calculation. The researcher acknowledged that most utilities use the scoring method since it can appropriately assess transformers' condition. Still, each utility uses a different score system and weighting factors. Kadim et al. (2018) research removed the tap changer section and the percentage ratio of the transformer and “Tap Changer” from the HI calculation due to the unavailability of the data.

4.2 Data Workflow

Figure 4.1 describes the workflow for this case study investigation. Two HI mathematical models were developed in Microsoft Excel Software to study the health condition of electrical transformers. Both models consider the conventional scoring method.

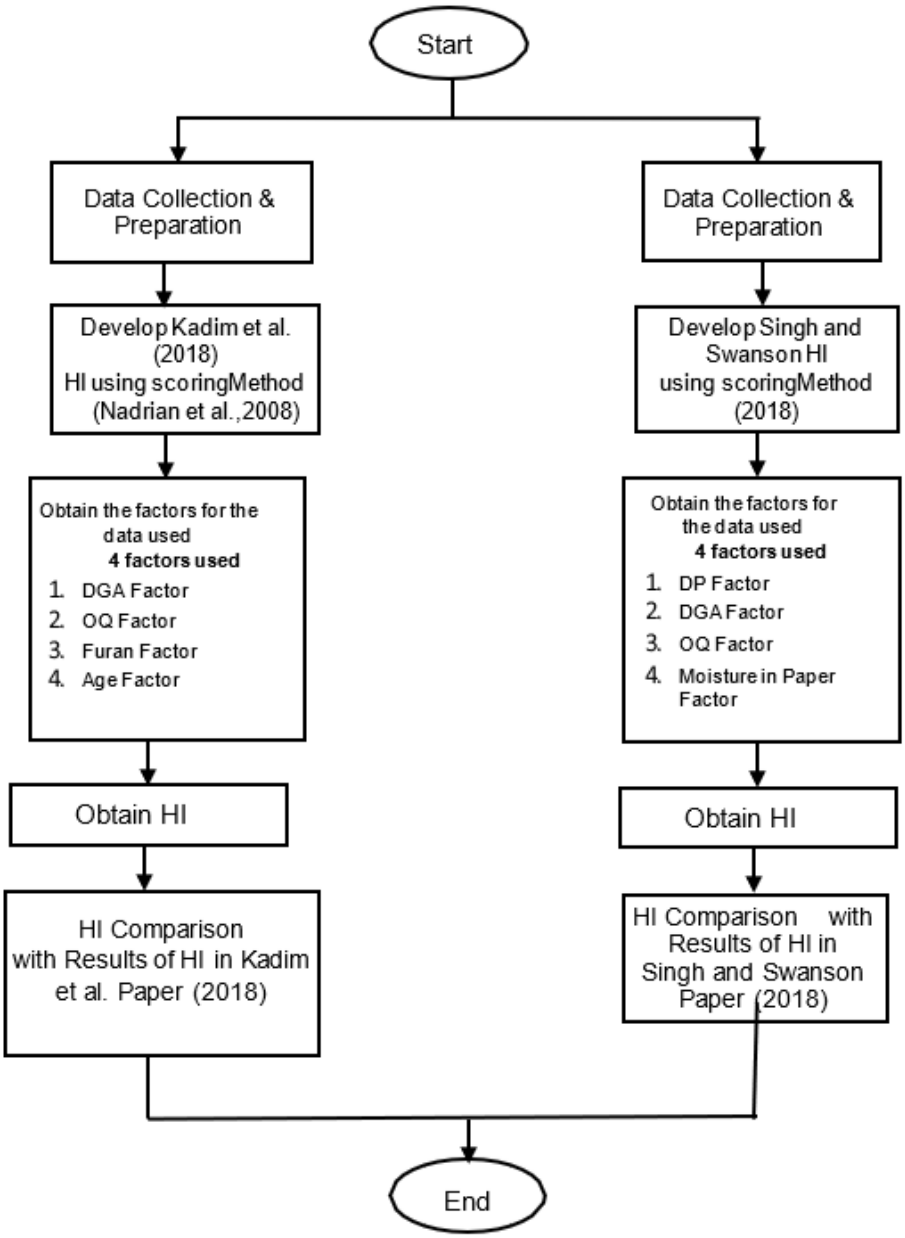


Figure 4.1: Research workflow for the conventional mathematical HI model.

4.3 Health Index Formulation

A measured scoring and weighting approach correctly identifies PT's health condition. This approach comprises the following measures:

1. Designated scores or existing situation deterioration measurements form health scores ranging from “normal health situation” to “poor health situation.”
2. The importance of the relative weighting factor for every diagnostic parameter allocation is from “very low importance” to “very high importance”.

3. Next, the highest feasible score is linearly assessed by totalling numerous measures 1 and 2 for every situation diagnostic factor (Duval, 2002).
4. Then, the total HI scoring calculation by using formula (4.1), which is considered by (Murugan and Ramasamy, 2019):

$$\mathrm{HI} = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} S_{max} \times W_i} \times 100\% \qquad (4.1)$$

Where HI represents the final HI metric, Si is the score of each diagnostic test's parameters, as shown in Chapter 2, Table 2.4. The Wi is the weight of the relative importance of individual diagnostic test parameters used to calculate DGA and Oil Quality Analysis (OQA) factors, and S_{max} signify the weight and maximum score of all diagnostics, respectively (Murugan and Ramasamy, 2019), n signifies the number of diagnostic tests involved in PTHI calculation. PTHI calculation involves dividing its total condition score by its maximum condition score, then multiplying by 100 because of getting small numerical fractional values as per the standards, which needs to be multiplied by 100 to lead to the correct condition of the identified condition of PT.

Health indexing measures transformer condition based on several criteria correlated to lengthy deterioration features, which can cause PT life termination over time. Signs of PT's condition, including its sub-parts, can be conceivable using PT's actual health records linked with the electric network operation of the in-service transformer. The total HI scores of a PT use different weighting factors that are unique to every utility, agreed upon per country, and can differ from academia (Lundgaard L. et al., 2014; Foros and Istad, 2020).

4.3.1 Scoring and Weighting Factors Using Standards Ranking Method

Arithmetical scoring and weighting methods are practical for all the conditions of diagnostic test outcomes, which occur in the shape of multicriteria assessment to compute PTHI for a large fleet of power transformers. Considering scoring/weighting factors, the factual PT condition weighing is in the shape of percentage indicators. The ranking technique is practical for sorting the PT condition into various ranks based on numerical scores, which are presented as Score ‘1’ for “Very Good” level, Score ‘2’ for “Good” level, Score ‘3’ for “Moderate”, Score ‘4’ for “Bad” level and Score

‘5’ for “Very Bad”. The score levels definition by applying the proposed limit from electrical utilities agreed on various international standards considering IEEE, IEC, and CIGRE guides (IEEE Std C57.104, 2008; Leauprasert et al., 2020; Suwanasri et al., 2013; Pompili and Scatiggio, 2015). Appendix H, Part 1 indicates the scoring, weighting, and interpretation rules.

These linguistic provisions are streamlined manifestations representing the PT's health condition. These provisions apply weighting factors to rating the degree of relative importance or impact of individual diagnostic factors that impact the PT's health condition. The weighting factors are diverse, agreeing with the present electrical utility practice. The scoring code begins with Weight ‘1’ for very low importance (“Very Good”), Weight ‘2’ for low importance “Good”, Weight ‘3’ for moderate importance “Caution”, Weight ‘4’ for high importance “Bad”, Weight ‘5’ for very high importance “Very Bad” (Murugan and Ramasamy, 2019; Flint et al., 2017). Resolving each diagnostic factor's weighting or scoring factors needs a PT specialist. It requires multiplying the scoring approach for all ranking measures by its weighting factor. Then, the total score for the present situation or degradation status of the PT under investigation is measured. The anticipated scoring and weighting factors for every diagnostic factor are described in Chapter 2, Table 2.4 (Khamis et al., 2023).

In a multicriteria analysis approach, various factors combine into an idealised condition-based Health Index. This approach requires grouping the numerous factors, constructing the mathematical and logical formulation, and determining the importance weightings of all aspects to permit combining the numerical values into a single HI score.

This solution is in two steps of procedures for some factors that depend on applying DGA and OQA. The first step is determining the weighting and scoring factors impacting input parameter ranking analysis by considering scoring tables of individual parameters, as shown in Appendix H Part 1. Then the factors of OQA or DGA can be computed according to the Equation (4.2) in (Naderian et al., 2008):

$$\text{OQA or DGA Factor} = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} W_i} \quad (4.2)$$

S_i is the score of individual parameters, Wi is the weight of individual parameters used to calculate that factor, and n represents the number of diagnostic tests.

The typical HI is standardised to the highest requirement measure of 100 and 0 for the least score of requirement measure, where < 30% signifies the “Very Bad” PT health condition and 100% signifies the “Very Good” PT health condition.

Practically, a mix of factors and not all of them will use diagnostic parameters according to the data availability and its limitation to categorise PT health into a score that shows its actual health condition as shown in Appendix D, Table D-15 and D-16. The HI score is the sum of the weighted DGA score, oil quality score, Furan score, and Age score oil to categorise the transformer mentioned in the paper by Kadim et al. (2018), or it can use the sum of the weighted oil quality score, moisture score, DP score and DGA score, which these parameters to categorise the transformer condition as discussed in the paper by Singh and Swanson (2018). Despite various practices and equal methodologies, all HI calculations signify similar results with the same intent of proving PT's health condition using the conventional HI scoring method.

4.3.2 Data Interpretations

Interpretation of input data must correctly use relevant standards to interpret the HI results. Table D-16 in Appendix D gives categories of Health Indexes and associates these with an expected lifetime and required decision. HI values are sorted into discrete categories from “very good” to “very bad”. The required accumulation for various discrete classes using PTHI needs to understand the relation between health condition scoring practice, which in some cases can be improved and fine-tuned to a suitable level based on relative importance. This interpretation ensures the relative importance of impacts testing parameters due to “dominant” factors and those due to generalised deterioration issues that occurred over time for the in-service PT. The PTHI development process must consider actual coherence, practice, and rationale.

The result of HI is not the same as maintenance testing since it is highlighted based on discovering defects and deficiencies that require correction or remediation to maintain the asset operating for some time. In line with IEEE standard C57.104, the interpretation of transformer conditions is classified into four conditions as shown in Appendix D, Table D-15; consider IEEE gas level interpretations (Jahromi et al., 2009) as well it can be interpreted as shown in Appendix D, Table D-16 following the approach by Naderian et al. (2008). Figure 4.2 indicates the main condition factors utilised in this literature review analysis, including the related scoring and weighting factors for PTHI assessment. The PTHI is the sum of chosen available testing parameters considered as factors, including OQA, DGA, Age and Furan. These factors define two levels: the first presents individual parameters, and the second presents the main factors. The scores are defined according to specific ranges defined by international standards and then multiplied by weights defined by utility experts, as shown by yellow boxes in the middle right of Figure 4.2. This range includes OQA factors, one of the testing parameters that provide six individual parameters for oil quality measurements (Moisture, BDV, DF, Acidity, IFT, and Colour). All have specific scores according to a defined range by international standards and defined weights by utility electrical experts. Similar to the calculation of the OQA, the DGA factor by applying the same method for its seven individual parameters (H_2, CH_4, C_2H_2, C_2H_6, C_2H_4, CO and CO_2), and an overall HI is calculated by applying the same methods, including all available factors as shown in Appendix H, Table H-1 to H-6.

- Table H-1 shows the scores according to the defined range at the individual level to calculate the DGA factor.
- Table H-2 shows the scores according to the defined range at the individual level for calculating the Oil Quality factor.
- Table H-3 shows the scores according to the defined range to measure DGA at the factor level.
- Table H-4 shows the scores according to the defined range to measure Oil Quality at the factor level.
- Table H-5 shows the scores according to the defined range for the age factor.
- Table H-6 shows the scores according to the defined range for the Furan factor.

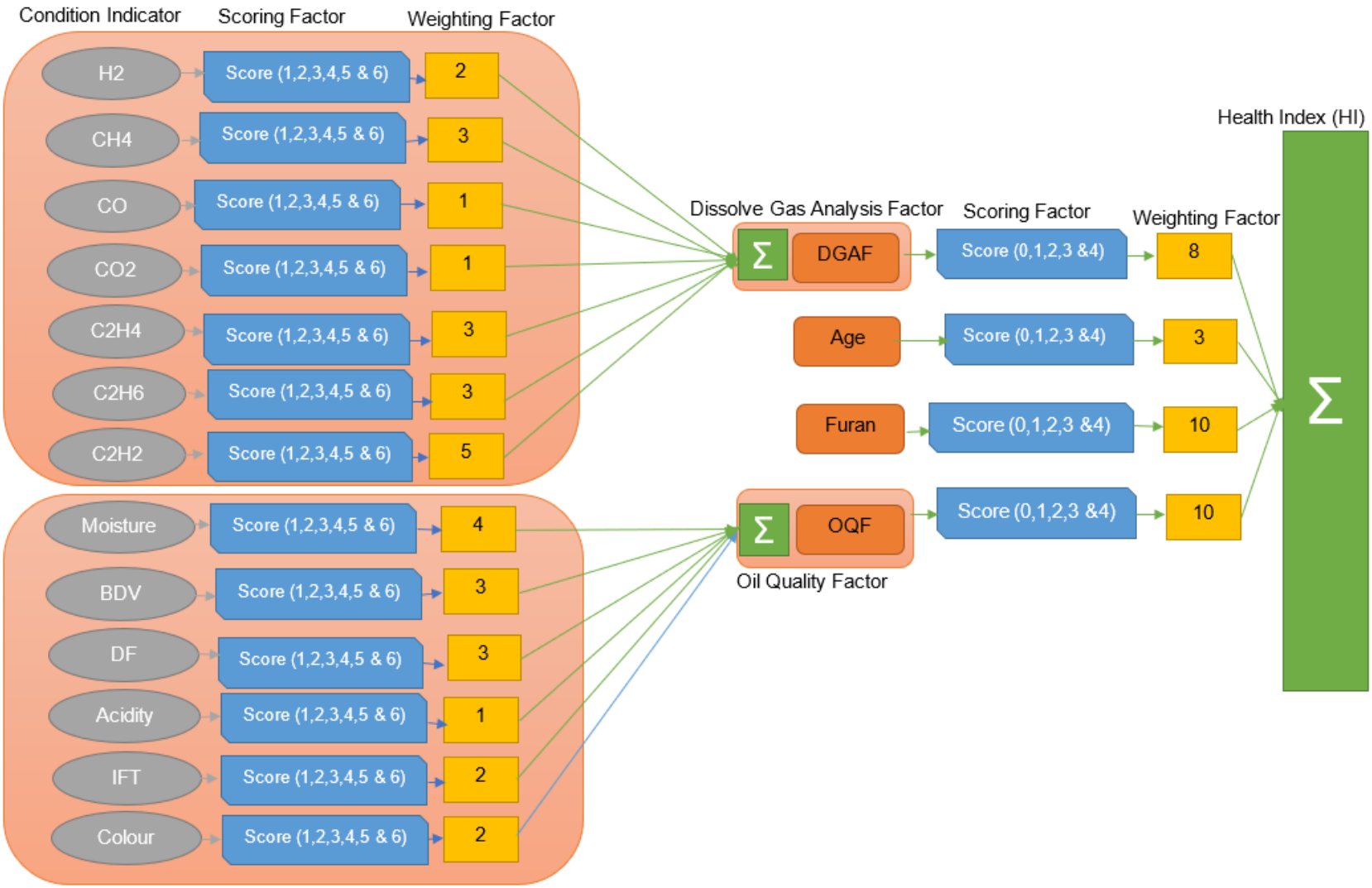


Figure 4.2: Health Index scoring calculation approach for power transformer (Khamis et al.,2023).

4.4 Case Study Investigation

This case study presents the extracted data from the literature review by Kadim et al. (2018) and Singh and Swanson (2018), available in Appendix D, Table D-13 and D-14.

4.4.1 Conventional Health Index (Singh and Swanson, 2018)

This HI development uses collected sample data from the above literature review. The overall HI is the percentage of calculated testing parameters weightings and scoring factors. Four parameters were considered due to the importance and availability, namely OQA, Moisture in Paper, DP and DGA test, as shown in Figure 4.3.

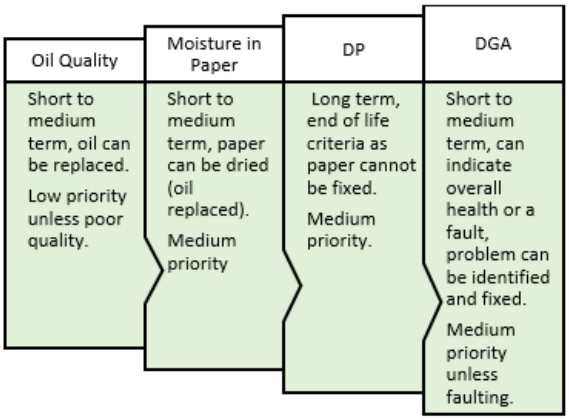


Figure 4.3: Priority chart for HI scoring parameters.

4.4.1.1 HI Calculation Method (Singh and Swanson, 2018)

Figure 4.4 shows the calculation design used to calculate the HI using the earlier input parameters.

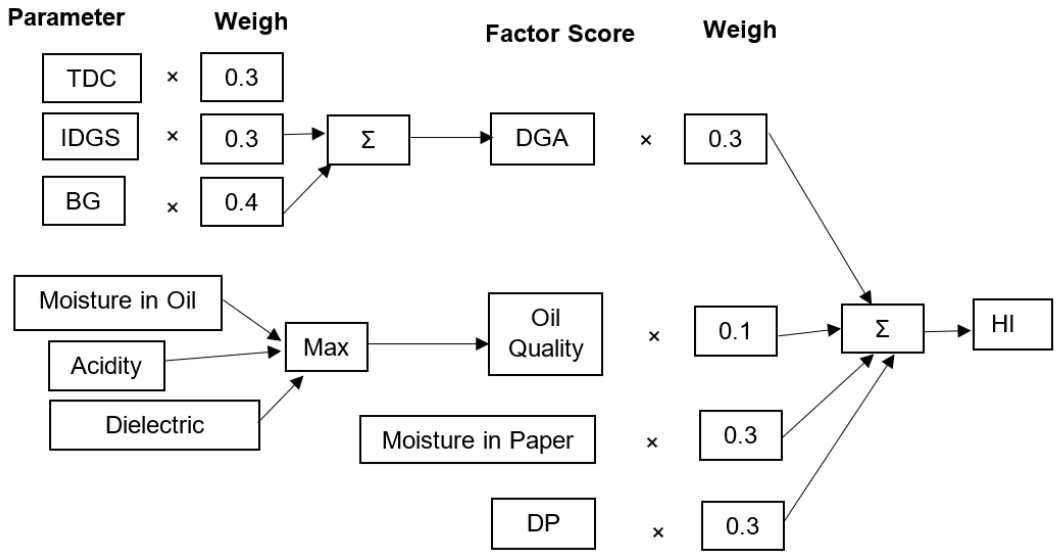


Figure 4.4: Methodology of the HI (Singh and Swanson, 2018).

4.4.2 Conventional Health Index (Kadim et al.,2018)

This HI development uses the above literature review, which collected sample data from 15 transformers in Appendix D, Table D-13. This HI considers four parameters due to their relative importance and availability: Furan, Age, OQA, and DGA test, as shown in Figure 4.5. The relevance of the small circles and the spacing represent the regular tests based on visual inspection as direct measure tests, while the big circles are expensive diagnostic tests that require laboratory analysis. The distance of the circle and the big HI circle represent the time frame for conducting this test. As per the manufacturer's warranty for the transformer, the age only became an issue after 25-30 years as it shows the longer distance, which is not an issue at the early transformer age. Furan is an expensive test and time-consuming. Most electrical utilities conduct this test at some point in the unit lifecycle and must take the transformer out of service before the end of its life and scrap the transformers (Singh and Swanson, 2018).

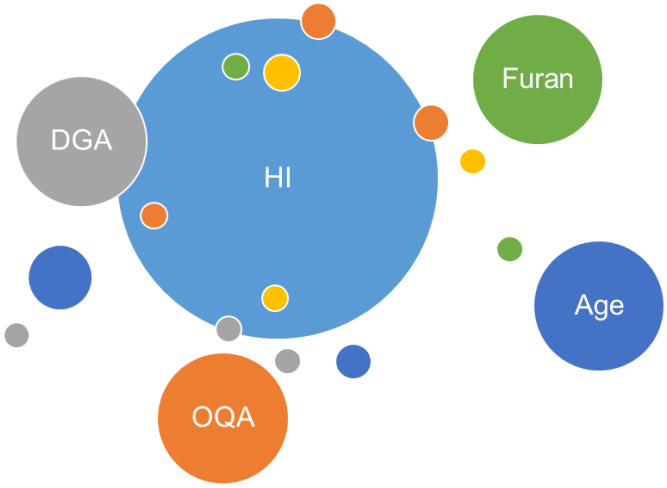


Figure 4.5: Factors for HI Calculation (Kadim et al.,2018).

Figure 4.6 shows the HI design utilised by Kadim et al. (2018). It states that a few adjustments to Naderian et al.'s (2008) approach include removing the tap changer section and the percentage ratios of transformers and tap changers.

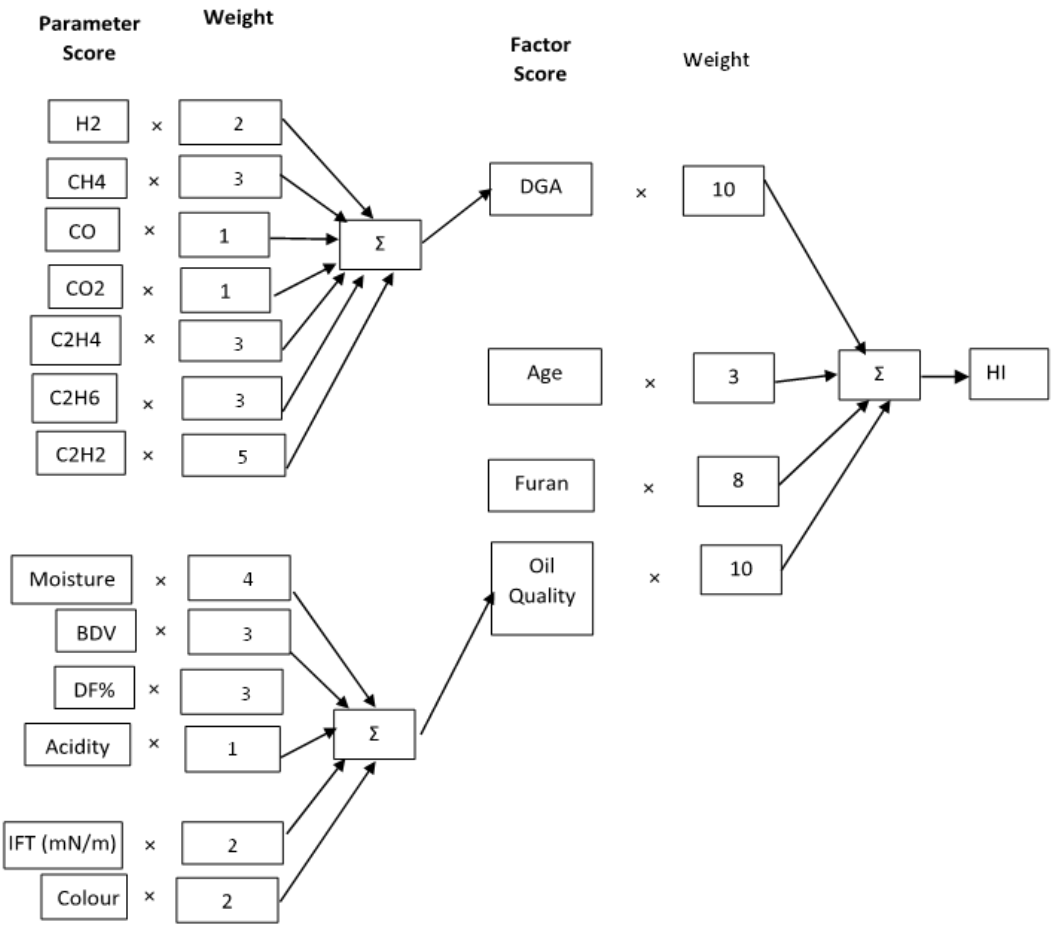


Figure 4.6: Health index design (Kadim et al.,2018).

Singh and Swanson (2018) made further adjustments to construct the HI scoring model, considering weights for individual parameters to determine the DGA and OQA. As shown in Figure 4.7, using the weights listed in the Naderian et al. approach (2008) did not result in similar outputs.

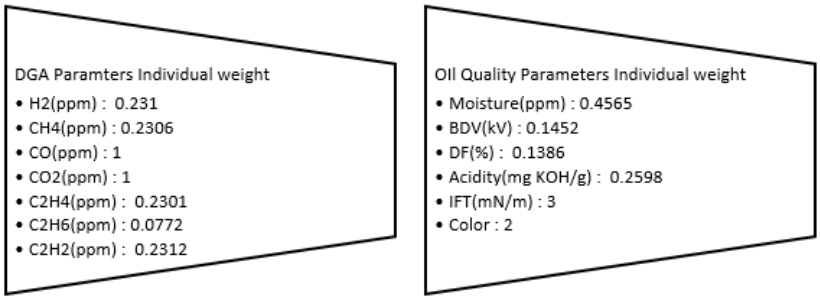


Figure 4.7: Weighting factors for DGA and OQA (Singh and Swanson, 2018).

4.4.3 Conventional Health Index Models Benchmark

4.4.3.1 Comparison 1: Sings Score using his approach.

Singh and Swanson's (2018) paper used moisture in paper, DP, and the DGA and OQA factors to calculate the health index. The model quantifies the results of the health categories into four categories: A for healthy, B for moderately healthy, C for moderately unhealthy, and D for unhealthy.

Table 4.1 contrasts the HI from Naderian et al. (2008), Taengko et al. (2013), and Singh and Swanson (2018) with the HI of this research. The HIs suggested by Naderian et al. (2008) and Taengko et al. (2013) have more elements based on Furan, IFT, and colour than those used in this comparison. However, using these parameters is not possible as intended by the authors. Nevertheless, this proves that an HI development must follow the actual data.

Table 4.1: Comparison of conventional HI models (Singh and Swanson, 2018).

	Health Index* (HI) by (A: Very Good, B: Good, C: Moderate, D: Bad, E: Very Bad)			
	Naderian	Taengko	Sing	PTHI
	Researcher 1	Researcher 2	Researcher 3	This research
TX1	88% (D)	2 (Risk=D)	5.5 (D)	5.49 (D)
TX2	64% (C)	2.67 (Warning=C)	3.11 (D)	2.684 (C)
TX3	64% (C)	2.33 (Risk=D)	1.92 (B)	2.184 (C)

*HI data interpretation (0-100%) varies among the researchers (0-100% or 0-10).

Transformer TX1: The PTHI model declared TX1 unhealthy. The OQA, in this case, is so bad that it shows TX1 as unhealthy. It shows a 5.55 score based on Singh and Swanson (2018) HI. On the other hand, in research by Naderian et al. (2008), the HI showed a healthy transformer. Furthermore, in research by Taengko et al. (2013), the HI showed an unhealthy transformer. This weighting factor for the OQA was not an adequate indicator to show a problem.

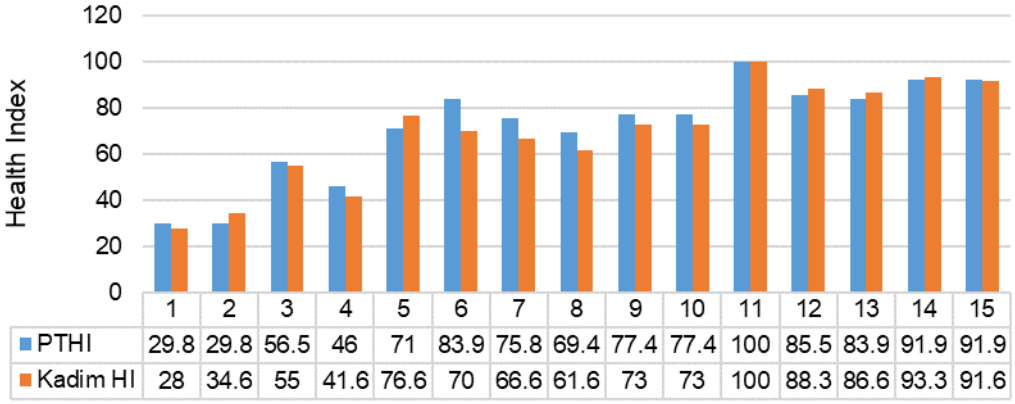
Transformer TX2: The PTHI model suggested that the transformer was in a moderately unhealthy condition due to the high levels of C_2H_2, which revealed discharges in the transformer. The Singh and Swanson (2018) HI calculated score is 3.11, suggesting the transformer is unhealthy. Utilising a per-gas indicator and the Basic Gas Ratio (BGR) caused a reasonable score classification. On the other hand, research by Naderian et al. (2008) and Taengko et al. (2013) demonstrated that HI demonstrated the same transformer as moderately healthy. However, both researchers utilised similar per-gas indicators.

Transformer TX3: The PTHI model suggested that TX3 is moderately unhealthy. These findings show BGR at low temperatures on a thermal fault due to extreme levels of C_2H_6 and TDCG. Naderian et al. (2008) and Taengko et al. (2013) suggested it is moderately healthy. Furthermore, Singh and Swanson (2018) HI found the same transformer condition moderately healthy and reasoned that based on dry paper due to extreme DP, OQ was found “Good”.

4.4.3.2 Comparison 2: Kadim et al. (2018) Scores using Naderian et al. (2008) approach.

This section considers the accuracy metric and confusion matrix to assess the scoring model performance on sample data from the paper by Kadim et al. (2018). Figures 4.8 and 4.9 compare the HI score of Kadim et al. (2018) and the HI calculated in this research chapter. In the model shown in Figure 4.8, the accuracy rate is 93% due to the weight of the unmentioned factors used in the Kadim et al. (2018) paper. This correction is after checking and correcting the error by the related researcher, as initially, the accuracy calculation was 60%.

$$\text{Accuracy rate} = \frac{\text{Number of correct predictions}}{\text{Total number of prediction}} = \frac{14}{15} = 0.933333$$



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
PTHI	29.8	29.8	56.5	46	71	83.9	75.8	69.4	77.4	77.4	100	85.5	83.9	91.9	91.9
Kadim HI	28	34.6	55	41.6	76.6	70	66.6	61.6	73	73	100	88.3	86.6	93.3	91.6

Figure 4.8: Comparison of calculated PTHI and Kadim HI (2018).

Table 4.2: Interpretation of HI numerical scoring value.

Health Index Range	Condition	Maintenance Action
0-29	Very Bad	Replacement
30-49	Bad	Replacement
50-69	Fair	Repair & Refurbishment
70-84	Good	Inspection
85-100	Very Good	Inspection

Table 4.3: Comparison of calculated HI models using Kadim et al. (2018) data.

HI Score (PTHI)	Transformer Condition (PTHI)	HI Score (Kadim et al. HI)	Transformer Condition (Kadim et al. HI)
29.84	Very Poor	28	Very Poor
29.84	Very Poor	34.6	Poor
56.45	Fair	55	Fair
45.97	Poor	41.6	Poor
70.97	Good	76.6	Good
83.87	Good	70	Good
75.81	Good	66.6	Fair
69.35	Fair	61.6	Fair
77.42	Good	73	Good
77.42	Good	80	Good
100	Very Good	100	Very Good
85.48	Very Good	88.3	Very Good
83.87	Good	86.6	Very Good
91.94	Very Good	93.3	Very Good
91.94	Very Good	91.6	Very Good

Table 4.4 shows the confusion matrix, which compares the scoring results of Kadim et al. (2018) and the PTHI calculation used in this research. The white cells present misprediction and the blue colour presents correct prediction. The model correctly predicted four transformers as “Very Good”, whereas one “Very Good” transformer as “Good”, and one “Poor” transformer as “Very Poor”, which is not an issue because both lead to the same maintenance recommendation action. On the other hand, it misclassified one “Fair” transformer as “Good”, which is an issue as it will lead to a different set of actions than the required one.

Table 4.4: Confusion matrix between Kadim et al. (2018) and this research.

Result Predicted	Kadim et al. Results				
KA-HI model	Very Good	Good	Fair	Poor	Very poor
Very Good	4				
Good	1	4	1		
Fair			2		
Poor				1	
Very poor				1	1

4.4.3.3 Overall Comparison

Based on an earlier analysis that used eighteen transformers, fifteen records from Kadim et al. (2018) and three from Singh and Swanson (2018) were used to evaluate the performance of the formulated HI model.

The HI model scoring method was developed based on IEEE standards (Appendix H Part 1). The percentage of score difference to assess the condition of transformers from the conventional HI table (Naderian et al.,2008; Jahromi et al.,2009). The testing parameters are DGA, OQA and age. In contrast, Furan data was unavailable because it is an expensive test that most utilities do not carry very often (Singh and Swanson, 2018).

Table 4.5 contrasts the scoring factors and overall HI results between the PTHI and other researchers' HI models. The table shows different calculated scoring factors and HI scores determined by this research KA-HI for the Kadim et al. (2018) and Singh and Swanson (2018) models. This analysis includes three HI models linked to Kadim et al. (2018) individual and factor parameters.

Table 4.5: Contrast between the PTHI model and other researchers' models.

Transformer Number	Scoring Factors* for the Final Health Index Based on Industry Standards.						KA-HI	Kadim et al. (2018) HI	Singh and Swanson (2018) HI
	DGAF	OQF	Furan	Age	Moisture in paper	DP			
TX1	3.3	30			4	1	5.49		5.5
TX2	4.5	2			2.78	1	2.68		3.11
TX3	5.1	0			1.18	1	2.18		1.92
TX4	20	0	4	3			29	28	
TX5	20	0	4	3			29	34.6	
TX6	20	0	4	4			56	55	

*The numerical values represent the scoring factors, not the input parameters.

This analysis differs from the other two research models because it uses different individual parameters and concepts for different scoring factors. The scoring factors of DGAF, OQF, Moisture in paper and DP for transformers TX1, TX2 and TX3 are based on a paper by Singh and Swanson (2018), as shown in Appendix D, Table D-18 to D-30. Whereas the DGAF, OQF, Furan scoring factor and age for Transformers TX4, TX5 and TX6 are based on the paper by (Jahromi et al., 2009) as shown in Appendix H, Table H-1 to H-6. Kadim et al. (2018) proposed using the DGAF, OQF, Furan and Age to study PT health conditions, while Kadim did not consider Moisture in paper and DP were not used. In contrast, Singh and Swanson (2018) used DGAF, OQF, and moisture in paper and DP data, and Singh did not utilise furan and age data for HI calculation.

The comparison between Nederian and Taengko is in a paper by Singh and Swanson (2018) in Table 4.1 to verify the accuracy of his model. Then it is utilised the three models, including Kadim as a fourth researcher, to verify this research HI model accuracy using two different scoring factors concepts (Singh and Swanson, 2018) and (Jahromi et al., 2009).

This research investigation continued to include PT health condition decision-making base by benchmarking with Kadim et al. (2018) model using the same scoring factor concept this researcher used by using the scoring factors of the researcher (Jahromi et al.,2009) as shown in Appendix D, Table D-16.

This research investigation is oppositely benchmarking with Singh and Swanson (2018) model by utilising their own scoring factors concept in Table 4.6 for calculating final HI scores.

Table 4.6: HI Index scoring factor (Singh and Swanson (2018).

Category	Descriptions	Score
A	Healthy	score < 1.01
B	Moderately healthy	$1.01 \leq$ score < 2.01
C	Moderately unhealthy	$2.01 \leq$ score < 3.01
D	Unhealthy	$3.01 \leq$ score

The above summary emphasises that HI is a valuable tool that mixes various data, considering that regular inspection diagnostic tests align with comprehensive health investigations for PTs. It also shows that the conventional HI scoring method helps start transformer condition analysis and strategic maintenance planning, as it could reduce operational costs and be easily adapted. From the economic aspect of any organisation, correct asset HI is the ultimate priority when studying any electrical utility PT health condition. Therefore, PTHI will assist in replacing corrective with a proactive annual maintenance plan. An example of an HI list produced by many researchers is available in Appendix D, Table D-17.

To conclude, the earlier results developed by KA-HI models prove an excellent model accuracy using the conventional HI model approach.

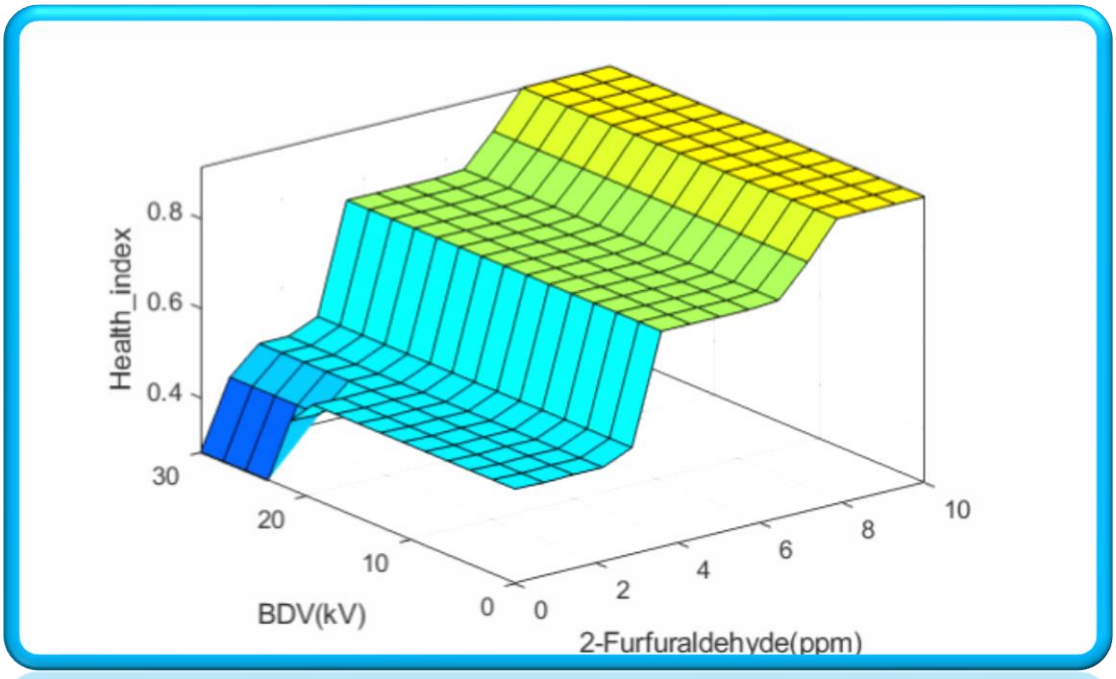
4.5 Summary

A conventional mathematical model is a widely traditional method considered and used by many HI developers and electrical utilities for developing power transformers Health Index. This chapter investigated conventional mathematical scoring and weighting factor methods for health index model development. The accuracy of the PTHI model with four researchers using conventional scoring and weighting factor methods led to acceptable PTHI results. This PTHI model accuracy is discussed in detail earlier, especially in the case study on the six failed transformers. This analysis proved that the HI scoring calculation algorithm can recognise transformers that require intervention before failure occurs. In the second comparison, the HI proved to correctly classify data in its condition with an accuracy of 93%. However, the accuracy rate improvement is by using the same weight used for that utility, which was not mentioned in the same papers as it is unique for each electrical utility, as stated by many authors. PTHI has to be an essential tool for asset management electrical utility, which benefits operation, maintenance, strategic asset planning and AM teams. The benefit of this research is that conventional PTHI model accuracy forms a reasonable basis as a condition assessment platform that can check and compare with future PTHI model results and findings from chapters 6, 7 and 8 because it is evaluated with other researchers' HI models and also evaluated with electrical utility HI model in chapter 7. It appears that the conventional method depends on the IEC and IEEE standards, which classify the condition score and weight into values linked to various health condition categories (“Very Good”, “Good”, “Fair”, “Bad”, and “Very Bad”) that require a decision. Initially, this traditional method requires understanding the complex mathematical model based on standards condition boundaries before it becomes friendly and easy to use. A national survey within the United Kingdom can be of interest for further development in the health index tool, scoring, weighting factors and related standards.

The next chapter will discuss the development of a fuzzy logic model in MATLAB. This literature review investigates the method used for HI development. It also evaluates the health condition of oil-immersed transformers using fuzzy rules and similar indicators to determine the health condition of power transformers and justify the required decision.

Chapter 5

Fuzzy Logic Health Index Approach



Chapter 5: Case Study 2 Fuzzy Logic Health Index Approach

5.1 Introduction

Imagine a mathematical approach that can handle uncertainty and vagueness, producing a partial truth value between 0 and 1. This approach is the essence of fuzzy logic, a non-conventional mathematical approach introduced in 1965 by Zadeh. It uses linguistic variables to process and explain information without certainty (Trillas & Eciolaza, 2015). This chapter explores the development of the fuzzy logic model in MATLAB and the use of 30 transformer records found in the literature. It demonstrates how fuzzy logic is an excellent method for demonstrating expert speciality for the related specific application. It generates decisions that can be remarkable for a better, effective and efficient result. Fuzzy logic is precious in HI evaluating PT deterioration. An expert system that utilises fuzzy logic can encode experience-based diagnoses and produce even more effective HI assessments. This assessment ensures that transformers' reliability, maintainability, and availability are considered (Arshad et al., 2014).

5.2 An Experiment of Building a Simple Fuzzy Logic Model in MATLAB

The available test data from the literature forms six membership functions, including Oil (2-furfuraldehyde, water content, DCG, DF, BDV, acidity) and DGA (CH_4: methane, C_2H_2: acetylene, C_2H_4: ethylene, C_2H_6: ethane, H_2: hydrogen). PTHI considers individual DGA gases as an ideal solution for identifying specific faults in a transformer that explain the transformer's health condition. These DGA gases determine the total health condition of the power transformer.

MATLAB is considered a six-membership function for developing a fuzzy model for the 30-transformer dataset, according to Abu-Elanien and Salama (2012) and Jürgensen et al. (2016). The seventy-nine fuzzy rules in the fuzzy model indicate one linguistic attribute for each transformer case. The final result signifies the transformer's health condition.

Specialised membership functions in MATLAB help find PT's health conditions. These functions assist in defuzzification, resulting in a precise value for the health index.

5.2.1 Membership Functions Model

The membership functions' model development follows IEEE standards, literature, and the knowledge of transformer field engineers, specialists, and operators. These functions are created explicitly for 69 kV or lower voltage levels, the operating voltages of all test transformers mentioned in this chapter.

1) 2-Furfuraldehyde content (2FL) - membership functions: This measurement is costly. It required the transformer to be out of service to find the degree of polymerisation (DP) in paper solid insulation. The test is one of the best indicators for evaluating PT health conditions and requires exceptional engineering experience. This measurement considers the solid insulation of PT for evaluating PT ageing (Abu-Elanien and Salama, 2010; Singh and Swanson, 2018). 2FL increases in PT oil due to moisture entrance inside winding paper insulation. This test is costly and needs PT's speciality. The PT health condition is categorised based on five states: “Very Bad”, “Bad”, “High Moderate”, “Low Moderate”, and “Good”. It appears that the 2FL membership function for this solid insulation paper in electrical PT is in a good state as the insulating paper DP is more than 700 units equal to 2FL 0.13 parts per million (PPM) based on a similar above standard (Abu-Elanien and Salama, 2010). In “Very Bad” condition, the DP of solid insulation transformer paper is lower than 200 units, equal to 2FL 7.3 parts per million (PPM). The five PT health states are estimated to be good to very bad. These identify the health condition of power transformers, as summarised in Table 5.1 and expressed in Figure 5.1.

Table 5.1: PTHI ranges for 2FL and DP.

PT Health State	2FL Level in (PPM)	DP units in (PPM)
Very bad	7.3	< 200
Bad	2.8 - 7.3	320 – 200
High moderate	0.8 - 3.55	560 – 290
Low moderate	0.13 - 1.2	700 – 425
Good	0.13	> 700

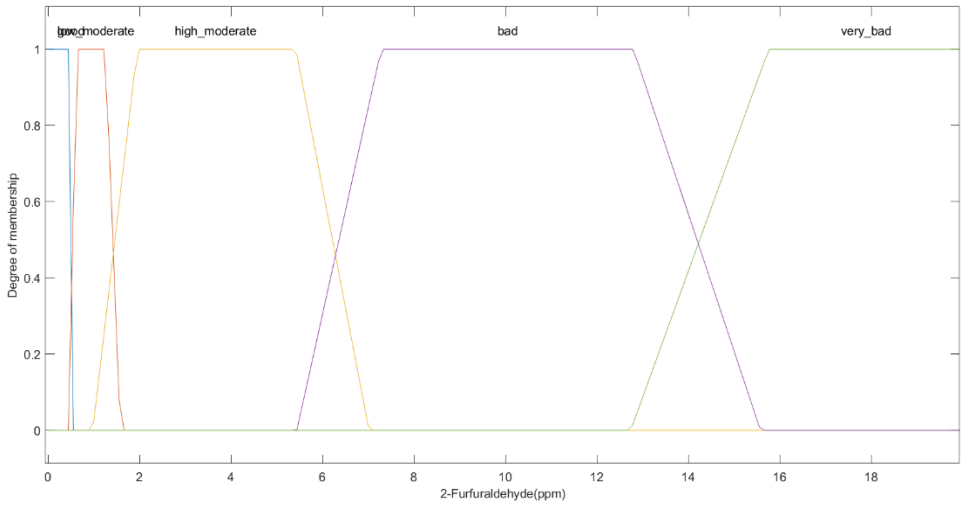


Figure 5.1: 2FL express five membership functions.

2) Break Down Voltage (BDV) - membership functions: This measurement shows the strength of PT oil insulation resistance to the impact of external voltage strains (IEEE Standard C57. 106, 2006). This BDV record assessment is conducted based on 1-mm interval assessment methods proposed by IEEE Standard C57. 106 (2006) and American Society for Testing Materials (ASTM) D18l6 (2019). Hence, BDV assessment is considered one of the most accurate methods for testing oil quality for PT health conditions. When the oil's ability to resist decreases, it will result in a rise in Partial Discharges (PD) and some arcs, speeding up individual power transformers' ageing. PT's health condition due to BDV uses similar health condition categories as “Bad”, “Moderate”, and “Good”. Figure 5.2 is developed based on IEEE Standard C57. 106 (2006), as it shows the three BDV membership functions of PT health conditions.

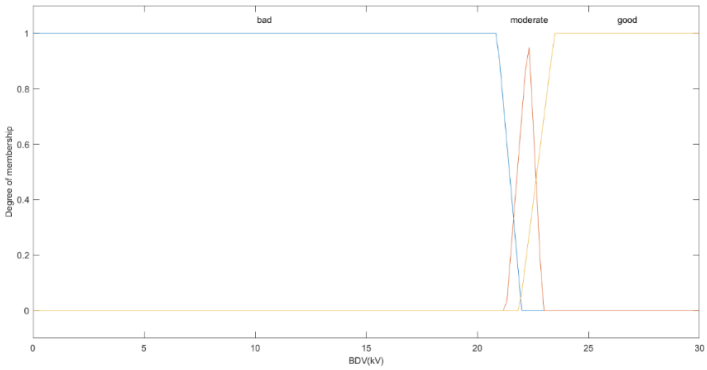


Figure 5.2: Membership functions for oil BDV.

3) Dissolved Compositable Gases (DCG) - Membership functions: This measurement follows regular in-service PT manoeuvre procedure. As the PT produces various DCG gases, the density will rise in the form of deviation within the PT and increase PT ageing (Abu-Elanien and Salama, 2010; IEEE Standard C57. 104, 2008). The DGA gas expression by the Rogers or Dornenberg ratios per IEC C57. 104 (2008) standard for the reclassification of the critical values of faulty power transformers. DGA is a strategic player in evaluating PT health conditions by utilising all DCG gases, identified as CH_4: methane, C_2H_2: acetylene, C_2H_4: ethylene, C_2H_6: ethane, and H_2: hydrogen. Similarly, the DCG membership function categorises the PT health condition as “Bad”, “Good”, or “Moderate”, as expressed in Figure 5.3 for describing the amount of DCG gases within a faulty power transformer.

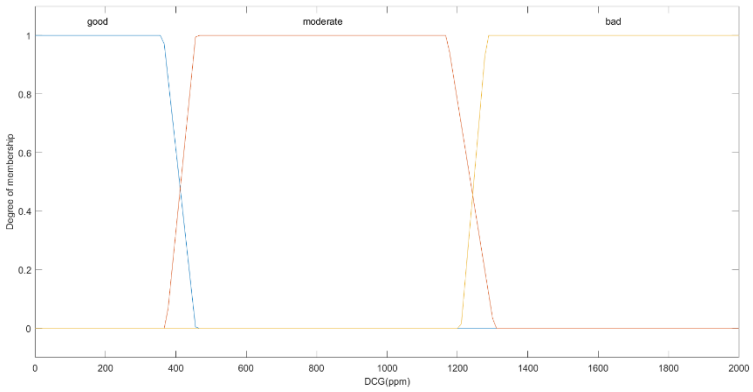


Figure 5.3: Membership functions for DCG.

4) Water content - membership functions: A significant proportion of ageing transformer contamination to varying degrees with water. The impact of

transformer oil water content on the power transformer's health is widely known, including the acceleration of ageing rates (Lelekakis et al., 2012; Nasrat et al., 2013), and at high saturation levels, the oil's dielectric strength decreases. (IEEE Standard C57. 106, 2006; Garcia et al., 2018); and issues resulting from bubbling at elevated temperatures (Perez, 2012). The water content level in transformer oil can be a reliable sign of PT's overall health. Also, as the paper insulation gets old, it will negatively impact the transformer oil water content, further reflecting the thermal insulation's PT condition to a certain degree. The water content level in transformer oil is split into three linguistic terminologies ("Good", "Moderate", and "Bad") as it can create related membership functions. This model uses three membership functions, as illustrated in Figure 5.4. The boundaries for the three water content conditions were primarily sourced from the IEEE Standard C57.106 (2006) for voltage levels up to 69 kV. If the voltage is high, the permissible amount of water in the insulating oil for ongoing usage is less than indicated in Figure 5.4.

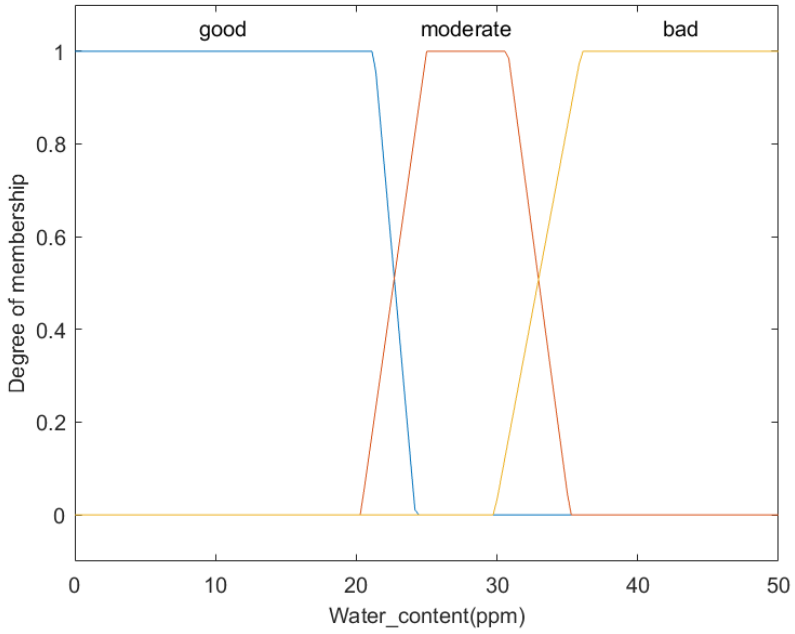


Figure 5.4: Membership functions for water content.

5) Acidity - membership functions: The acidity measurement is in (mgKOH/g) for transformer oil. This measurement will estimate the deterioration of the health

condition of transformer insulation. Over time, the level of transformer oil acidity increases with prolonged usage, which causes further deterioration in PT's health condition (Gouda and El Dein, 2019). According to Nasrat et al. (2013), PT paper insulation in transformers can cause a further rise in the PT oil acidity. Figure 5.5 demonstrates the use of three membership functions, “Bad”, “Moderate”, and “Good”, to model acidity levels similar to PT water content. The limits for the three conditions are primarily from the IEEE Standard C57.106 (2006).

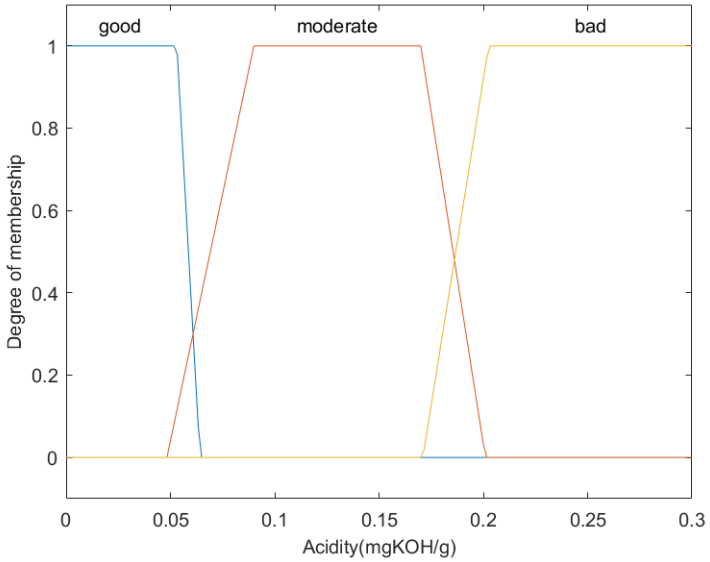


Figure 5.5: Membership functions for acidity content.

6) Dissipation factor (DF) - Membership functions: This measurement shows energy lost due to an in-service power transformer. The measurement to assess the characteristics and pollution of PT insulating oil as the PT oil deteriorates and the dissipation factor rises. Once the dissipation factor increases, it will relocate to a heat energy source for the PT oil. This DF will raise the oil temperature of the power transformer. Finally, it will lead to speeding up PT ageing. Similarly, the PT health condition utilising this membership function is categorised as “Bad”, “Good”, and “Moderate”, as expressed in Figure 5.6, which is developed based on the information found in IEEE Standard C57. 106 (2006).

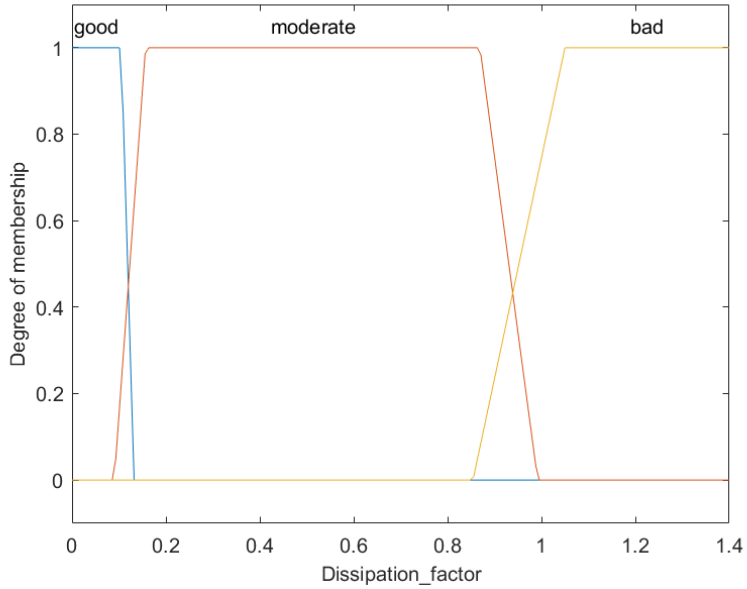


Figure 5.6: Membership functions for the dissipation factor.

7) Membership functions for the output PTHI: PT health condition expressed based on HI five health states: “Very Bad”, “Bad”, “Moderate”, “Good”, and “Very Good”.
The HI results are summarised in Table 5.2 and expressed in Figure 5.7.

Table 5.2: HI ranges for PT health states.

PT Health State	HI Range
Very Bad	0.85 - 1
Bad	0.65 – 0.9
Moderate	0.35 – 0.7
Good	0.2 – 0.4
Very Good	0 – 0.25

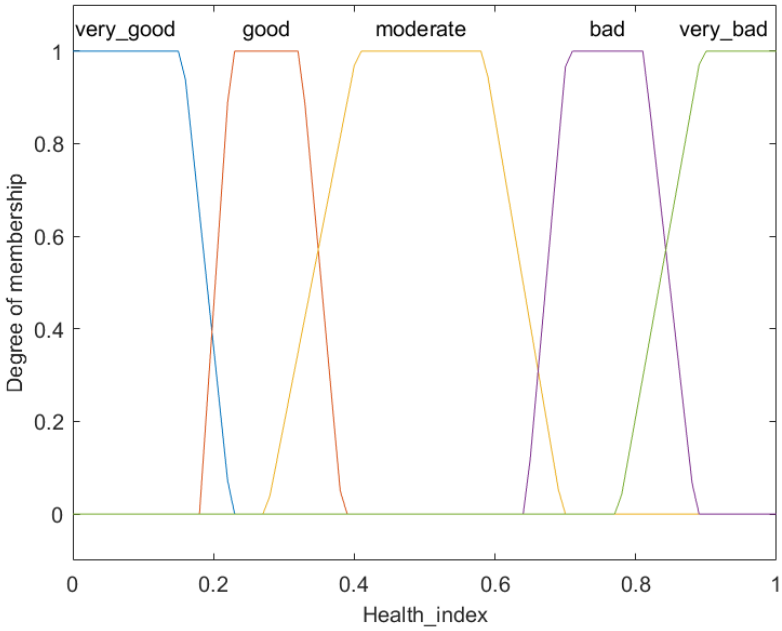


Figure 5.7: HI expresses PT health state.

5.3 Expert Rules Conversion Results

This section states that six input membership functions (2FL, BDV, DCG, Water Content, Acidity, and Dissipation Factor) are used in MATLAB, as shown in Figure 5.8 for the 30 PT records. The last column represents the HI value for the PT health condition.

.

Figure 5.8: The expert rules applied for calculating the health index based on various parameters.

Table H-9 in Appendix H lists seventy-nine fuzzy rules utilised to build the PTHI model for this research chapter. The novelty involved converting complex rules to simple ones used in this HI model. The input variables are given different priorities, with 2-Furfuraldehyde content being the greatest urgency based on the reliable output in evaluating PT solid insulation health state. The findings also show that DCG output is a second urgency. On the other hand, DF, BDV, acidity, and moisture (water content) express the slightest urgency for maintenance decisions.

PT moisture based on water content factor indicated the importance of this factor utilising MATLAB surface plots as expressed in Figure 5.9 a (left) -b (right).

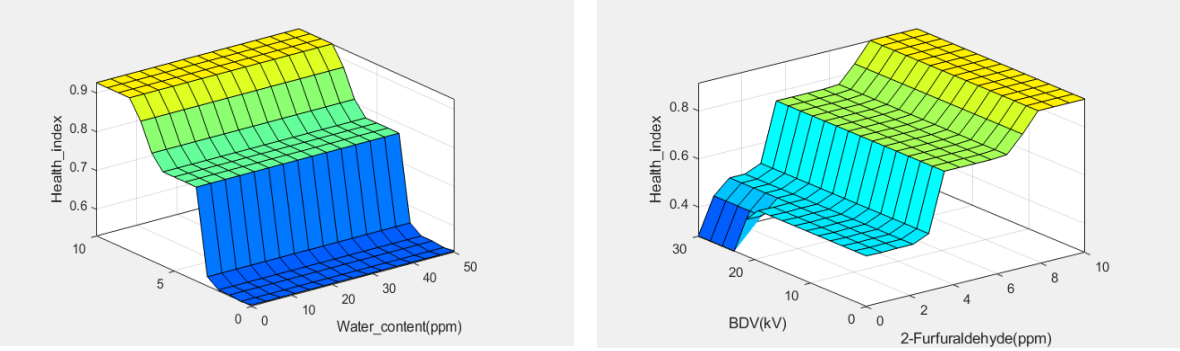


Figure 5.9: Fuzzy Logic Model Surface graphs.

5.4 Summary

This chapter presents a technique to evaluate PTHI using a MATLAB fuzzy model for PT oil parameters to calculate the PT health state by employing seventy-nine fuzzy rules. These rules assist in mapping the membership function parameters to obtain the required results that can benefit the maintenance decision. The MATLAB fuzzy model investigating 30 PT records and the six-membership function assessed PT's health condition. The novelty of this chapter's work is converting complex rules to simple rules that consider PTHI model development.

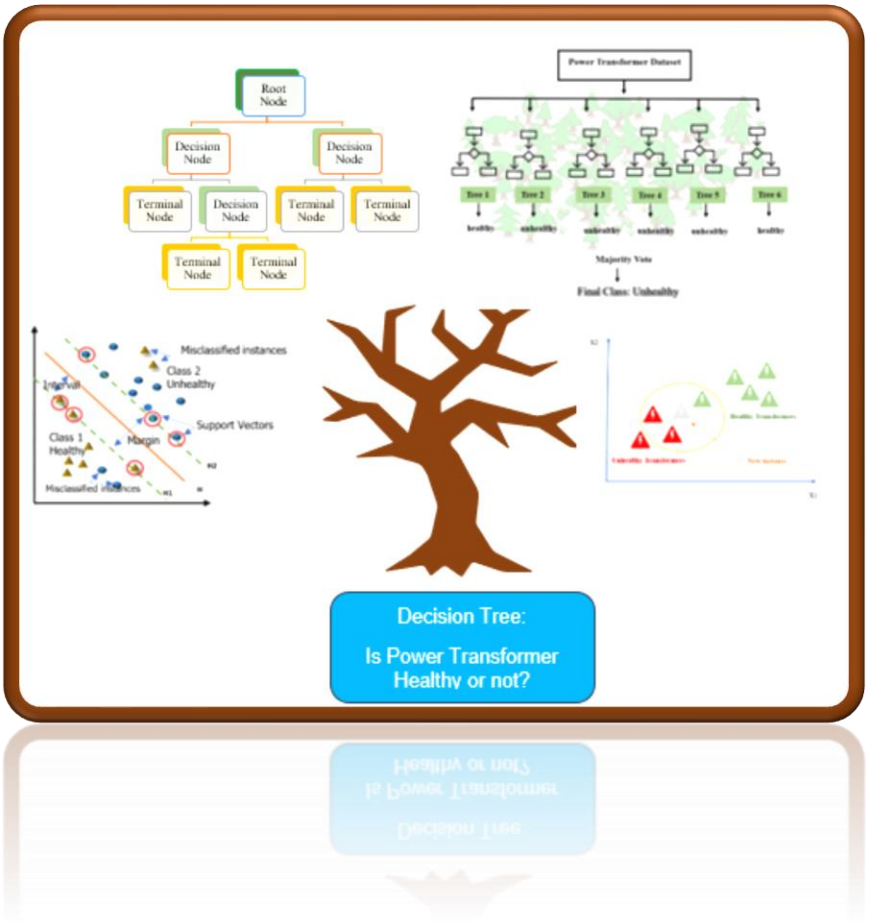
PTHI introduced the benchmark approach as a general method based on membership and transparent calculation rules. This method is disadvantageous because utilities can apply it only directly to define the fleet of transformer health indices. Based on other researchers' findings in the result section, FL is not a good choice for big data or solving data issues and is highly dependent on power transformer expert judgment.

Further study will be carried out in the following chapters, 6, 7, and 8, using the medium datasets collected from the literature review and the actual big data collected from the electrical utility, using ML algorithms, in line with the recent rapid development of the importance of using AI and ML for the development of PTHI and solving significant data issues. Therefore, future research to integrate fuzzy logic with ML for PTHI is an opportunity for further investigation.

The next chapter will use the learning lessons of Fuzzy logic and conventional HI methods based on small datasets to build AI and ML HI models that use computer science and implement a data science approach for solving data issues for medium datasets consisting of 30 electrical power transformers.

Chapter 6

Case Study 2 Machine Learning Health Index Approach for 30 Electrical Power Transformers



Chapter 6: Case Study 3 Machine Learning Health Index Approach for 30 Electrical Power Transformers

6.1 Introduction

Utilising AI ML algorithms in R programming language offers various solutions to overcome HI issues in terms of high reliance on expert judgment or missing data (Tang and Wu, 2011). Most developed studies focus on studying the HI for critical, expensive, and reliable assets similar to power transformers or overhead lines using MATLAB, Monte Carlo Simulation, R programming language or Python. Still, this research went deeper into something more that suits statistical analysis using R programming language. The conventional HI method is less efficient, neglecting AI algorithms and deep learning methodology, proving that AI is vital for improving HI for electrical utilities (Kadim et al., 2018). AI leads to intelligent HI by developing the required algorithms, especially once the required dataset is validated and available. AI solves dataset issues, predicts the related numerical missing data, and improves data uncertainty caused by expensive diagnostic testing or lack of transformer experts' judgements. Therefore, this will improve the estimation accuracy and balance imbalanced data even though the dataset will remain as estimated data. The HI outcomes will remain helpful in finding any faulty transformers within a large fleet of power transformers, and it will keep the required goal of decision-making achieved within the required time frame.

According to the literature review finding in section 2.4.2.3, the effectiveness and reliability of the proposed seven algorithms in assessing PT health state in this research area. This chapter presents an assessment for building these seven ML models in R statistical programming language, namely Binary Logistic Regression (BLR), Artificial Neural Network (ANN), Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbours (KNN), Decision Tree (DT) and Random Forest (RF) algorithms. The PTHI model is a binary classification problem, and it uses available literature review data to start HI modelling analysis. Finally, the model's assessment considered utilising various metrics, including model accuracy percentages, confusion matrix, sensitivity, and specificity of the model, which will be later helpful for a large fleet of power transformers' health condition investigation.

6.2 Principles of developing HI Machine Learning models

The four points below consider the principles of developing a HI model using ML algorithms.

- Define the objectives and end goal.
- Define the content of the HI model.
- Learn by experiments from implementing ML models.
- Evaluate, study, and improve the results to a better accuracy.

6.2.1 Objectives and end goal

The objective is to design a PTHI model implementing AI using ML models as, over the past decade, the HI has become an excellent condition evaluation method that drives power transformers' asset management maintenance strategy for electrical utilities. The goal is to provide a tool built with AI using an ML model that can benefit preventive maintenance and AM investment decision-making.

6.2.2 Content of HI Model

HI is available for electrical utility maintenance, AM engineers, consultants, manufacturers, and academia to investigate the needed HI structure and process, which is presented in Figure 6.1 that drives the electrical utility maintenance strategy and life cycle planning. Electrical utilities employ the conventional ranking method as it gives the necessary asset health condition review and in-house implementation. Nevertheless, the methodology is unique for the electrical utility concerning the mathematical models based on scoring and weighting factors regardless of the various analysis techniques in developing the HI calculation methods.

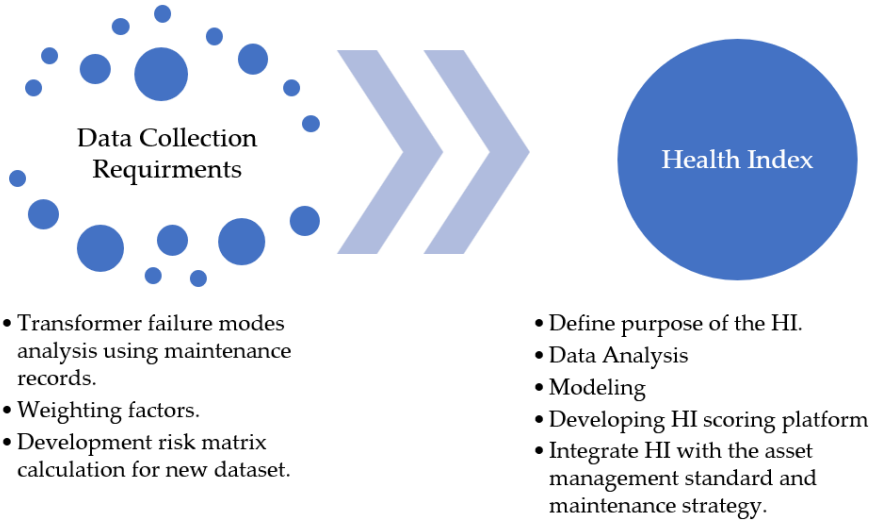


Figure 6.1: Main Principles for HI development diagram.

Electrical utilities and academia are focusing when designing HI on tackling critical assets that are costly and a threat to the leading network's performance and operation. According to Li et al. (2018), PT is an essential asset because of its high cost and importance in transmission, generation and distribution networks. Velasquez and Lara (2016) assess the actual condition of the overhead lines due to their importance. The overhead lines and underground cables play a strategic role in energy supply transfer to end users from one network to another.

Therefore, an electrical utility's primary interest drives good decision-making to commit to, or possibly defer, multimillion pounds of new investment for asset refurbishment, replacement, or maintenance programs. This research uses these judgements for a detailed investigation to support the deferral of funding challenges and consider optimum investment solutions. This judgement requires using different types of computer languages and improving the old mathematical model, tactic and HI calculation scoring methods in MATLAB.

HI developers, like consultants or manufacturers, make a lot of business and profit through day-to-day improvement on the required technologies and enterprise software that present the HI tool. The HI condition scoring solution shall combine all electrical assets in one asset register for asset health condition investigation. Therefore, academia can bring system operators, researchers, consultants, and manufacturers together to work on agreed HI practices to serve the environment, society, and the

local and national economies, which can control asset ageing funding challenges and ensure the best HI solution is in place. Figure 6.2 shows the processes of HI model development.

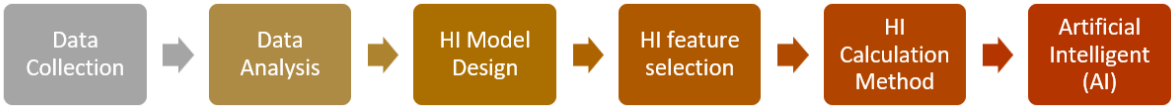


Figure 6.2: AI HI development process diagram.

6.3 Learn by Experiments and Implementing Machine Learning Models.

Various experiments are implemented in this chapter using ML algorithms that utilise the following research methodology.

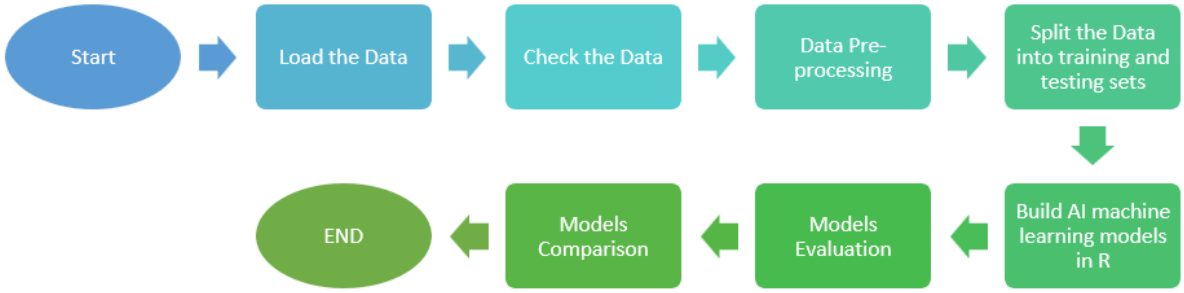


Figure 6.3: Research Methodology using AI in R programming language.

6.3.1 The dataset description

The research uses R programming language and visualisation techniques to understand the dataset. First, Figure 6.4 is a statistical summary of the four input parameters used in this chapter's experiments (Acidity, BDV, DCG, and Furfuraldehyde) that shows each parameter's minimum, maximum, mean point, and outliners.

```
> summary(healthIndexData)
       HI          Acidity             BDV              DCG        
 Min.   :0.0   Min.   :0.0100   Min.   : 4.05   Min.   :  8.00  
 1st Qu.:0.0   1st Qu.:0.0335   1st Qu.:33.67   1st Qu.: 31.25  
 Median :0.0   Median :0.0900   Median :47.25   Median : 74.50  
 Mean   :0.3   Mean   :0.0996   Mean   :48.13   Mean   :129.57  
 3rd Qu.:1.0   3rd Qu.:0.1542   3rd Qu.:66.30   3rd Qu.:163.25  
 Max.   :1.0   Max.   :0.2290   Max.   :75.00   Max.   :504.00  
 Furfuraldehyde
 Min.   : 0.0200  
 1st Qu.: 0.2175  
 Median : 0.6700  
 Mean   : 2.5617  
 3rd Qu.: 4.2600  
 Max.   :15.7000
```

Figure 6.4: Statistical summary based on R programming analysis.

Figure 6.5 presents 2D scatter plot illustrations of the transformer records in this chapter's experiments. The dataset is available in papers (Zuo et al., 2016; Abu-Elanien et al., 2012) and Appendix D, Table D-12.

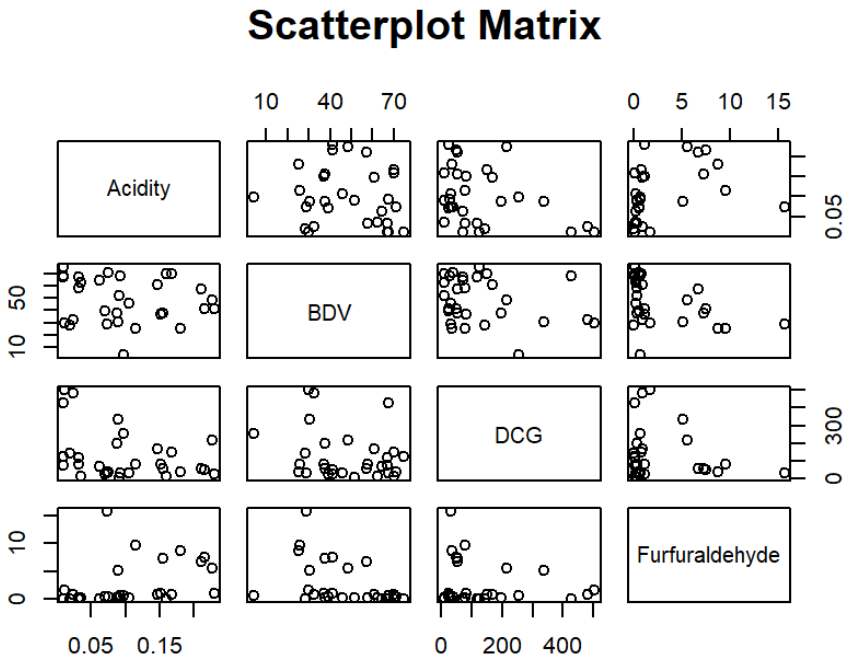


Figure 6.5: The relationship between the individual factors.

Figure 6.5, the 2D scatter plot shows the records of transformers in this experiment as 21 transformers were healthy in class (0) and nine transformers binary classification as unhealthy in class (1) axis. The scatterplot shows a matrix with the four input parameters along the diagonal boxes. The other boxes in the matrix display the association between each pairwise combination of 4 input parameters. The box in the

top right corner of the matrix displays a scatterplot of values for furfuraldehyde on the x-axis versus acidity on the y-axis. Whereas the box left to it displays a scatterplot of values for DCG on the x-axis versus acidity on the y-axis, respectively. The figure shows no correlation between each pairwise combination of 4 input parameters.

Furthermore, the box plots of the individual parameters are in the figures below. First, the acidity boxplot in Figure 6.6 shows the value range of acidity as it ranges between 0.05 ppm to 0.15 ppm with a median and a mean value of 0.09 ppm. Secondly, DCG data values range from 8 to 170 ppm with three outliners which show high DCG values in the range of 400 – 500 ppm with a median value of 74.50 ppm and a mean value of 129.57 ppm, which is affected by the outliners in the dataset as shown in Figure 6.7. Thirdly, the BDV boxplot in Figure 6.8 shows that it ranges between 4.05 ppm and 75 ppm with a mean value of 48.13 ppm. Fourthly, the Furfuraldehyde boxplot in Figure 6.9 ranges between 0.02 ppm and 15 ppm with a mean value of 2.56 ppm.

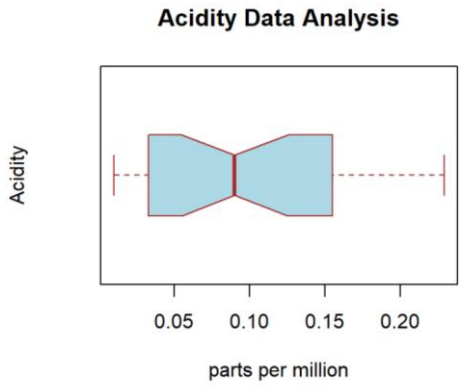


Figure 6.6: Acidity Data Analysis.

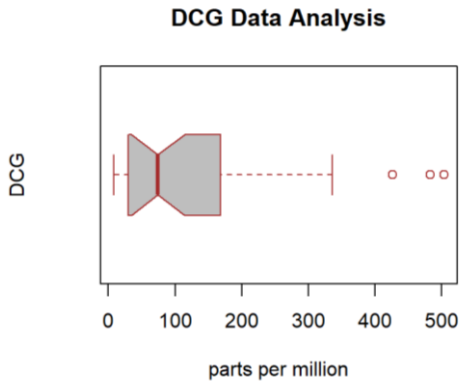


Figure 6.7: DCG Data Analysis.

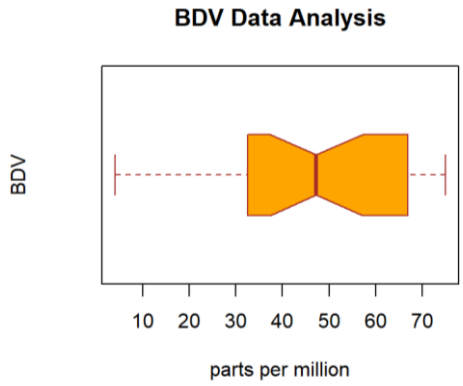


Figure 6.8: BDV Data Analysis.

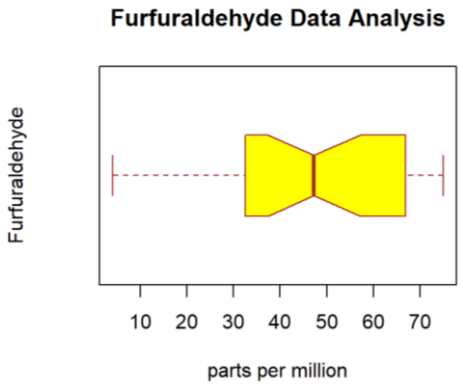


Figure 6.9: Furfuraldehyde Data Analysis.

6.3.2 Machine Learning Models

In this section, seven ML algorithms, which are Artificial Neural Network (ANN), Binary Logistic regression (BLR), Decision Tree (DT), K-Nearst Neighbours (KNN), Random Forest (RF) and Naïve Bayes (NB). Universally, those algorithms are applied for binary classification purposes, as explained below during the investigation. This AI model is to identify accurate patterns identifying transformers' health condition characteristics that specify healthy or unhealthy transformers. Also, paying attention to those individual ML algorithms, especially their advantages and disadvantages.

6.3.2.1 Binary Logistic Regression (BLR)

It is discussed earlier in the literature review chapter in a few works where the BLR algorithm is deployed and experimented with in a statistical hypothesis model to validate the correlation concerning dependent and independent variables. The BLR model application occurs during probabilistic statistical classification. It is also mentioned in a research paper by (Zuo et al., 2016) that BLR calculation accuracy is corrected based on individual essential characteristics which are considered during BLR implementation while the algorithm eliminates the unrequired variables. Nevertheless, the BLR model is not suitable for investigating multifaceted correlations. Hence, it is replaced by more multifaceted algorithms, the same as ANN, which certainly exceeds the BLR algorithm. During the implementation of the BLR model for big dimensional datasets, an overfit issue can occur on the training dataset, which, if not avoided, will lead to wrong HI calculations and findings. In this chapter, the BLR model will study assembling the asset measurements into a numeric PTHI value. Figure 6.10 shows the proposed relationship using the BLR model for designing transformer HI.

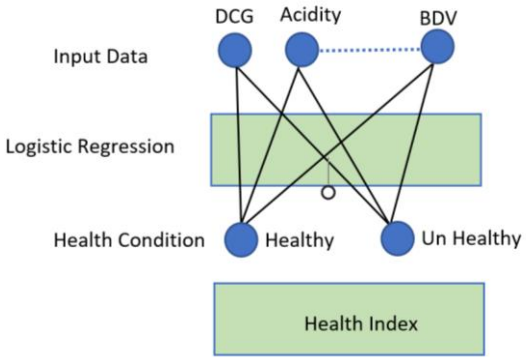


Figure 6.10: Binary Logistic Regression model for the transformer HI.

6.3.2.2 Artificial Neural Network (ANN) Algorithm

The ANN algorithm combines the assessment of various transformer components into a numeric HI value. This chapter develops the ANN model based on the literature review and IEEE standards using a training set derived from four conditional boundaries: acidity, breakdown voltage, dissolved combustible gases, and furfuraldehyde. The proposed method's performance is comparable with expert classification.

In the literature review chapter, the ANN algorithm is another AI solution for binary classification modelling based on various literature, as the algorithm performs effectively, as shown in the research paper (Islam et al., 2018).

ANN algorithm is considered in this simulation and modelling reliant on wide use, recommendation, and feedback from various academic researchers, especially a paper by (Islam et al., 2018). The algorithm directs all the unspecified intelligence hidden for the implemented dataset. Also, theoretically, ANN model developers consider linear, nonlinear, individual and multilayer approaches according to the research (Samanta et al., 2003). The strength of ANN in finding faulty transformers in complicated cases differs from other algorithms. The ANN configuration is interconnected circular nodes named artificial neurons. An artificial neuron signifies all circular nodes. The arrows signify a link of the output of one artificial neuron towards the input of the first circular nodes, which consist of 3 layers: input, hidden, and output, as illustrated in Figure 6.11. The input layer signifies input parameters experimented using the ANN model. The output signifies the ANN results identifying healthy (0) or unhealthy (1) power transformers. During ANN modelling, sigmoid or logistic activation functions play a part in the output and the hidden layers. The logistic function considers every actual value as input and output values in the range of (0) to (1). The higher the input signifies, the more positive inputs, and the nearer the output value is determined as (1). Conversely, the more trivial the input signifies, the more negatives, and the nearer the output is specified as (0). Therefore, the ANN model signifies plotting the weight of individual input parameters in input circular nodes holding “X” values and output parameters in the output circular node holding “Y” values because each gap will represent multidimensional (Islam et al., 2017).

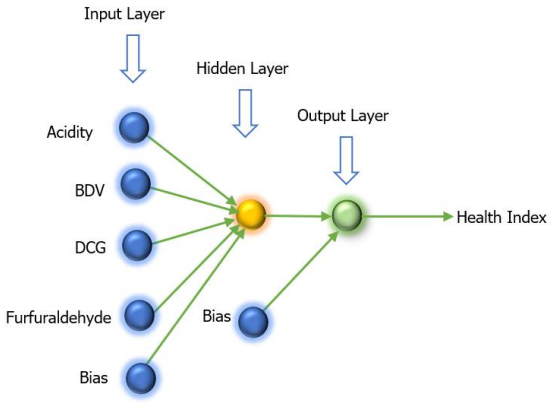


Figure 6.11: Artificial Neural Network model design for the transformer HI.

6.3.2.3 Support Vector Machine (SVM)

SVM is considered an algorithm that solves classification/or regression problems. It solves the problem by establishing a hyperplane or a line within the classes in the dataset. SVM works more efficiently in multidimensional datasets that include separate health classes. A dataset intends to signify the class healthy as (0) and unhealthy as (1) power transformers with distinct forms. As illustrated in Figure 6.12, the SVM model positions a linear boundary between the unlike classes and position input parameters concerning which the boundary is expanded (Samanta et al., 2003). However, the SVM algorithm struggles to establish a hyperplane for nonlinear datasets. Hence, the Kernel function is the substitute for resolving this issue (Patle and Chouhan, 2013) and establishing a third-dimensional boundary.

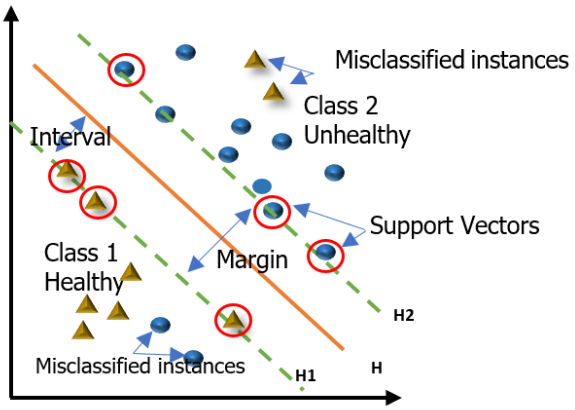


Figure 6.12: SVM algorithm method for the transformer HI.

6.3.2.4 Naïve Bayes (NB)

NB is another classifier considered a practical ML algorithm for sorting responsibilities that consider text arrangement into different classes. NB models the distribution of input parameters for a specific dataset classification. NB is also called a probabilistic classifier reliant on Bayes theorem (Dinov, 2018). Moreover, the Bayes theorem is known as the Bayes rule, which grants the ability to "invert" uncertain likelihood. NB algorithm relies on the hypothesis that the predictor variable is tentatively independent, assuming the category of the statistics model is in doubt. This category means the posterior probability of the experiment for a specific class granted the predictor variables is calculated applying Bayes theorem (Mitchell, 2015).

It is an algorithm used in classification as a supervised ML technique that classifies the instance based on its nearest neighbours in the feature space. It finds how many training samples are nearby in the distance to a given instance to predict the instance class (Alqudsi and El-Hag, 2019). Figure 6.13 demonstrates the application of the Bayesian Theorem to PT health condition classification.

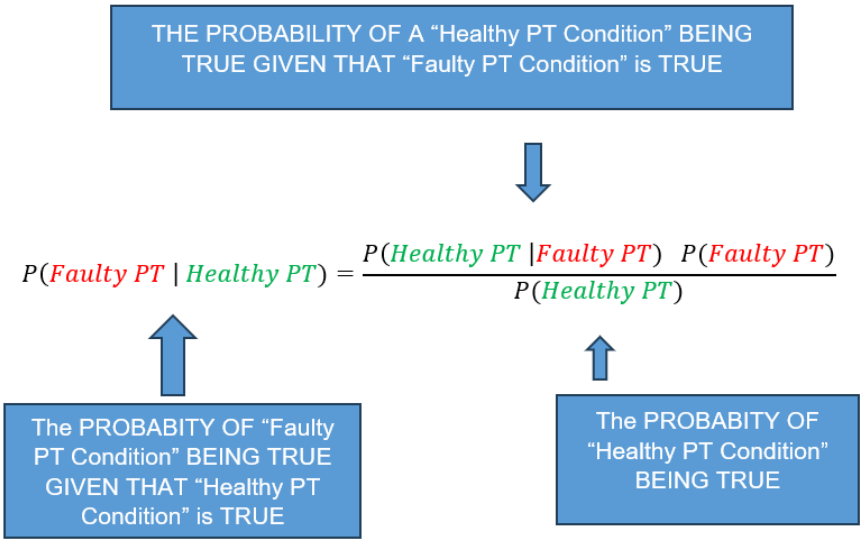


Figure 6.13: PTHI uses the Bayesian Theorem based on the NB algorithm.

6.3.2.5 Decision Tree (DT)

The DT algorithm works by assessing the information obtained from the variables. The maximum normalised details are then employed to separate the instances into

divisions of the same class label. This activity repeatedly continues for all the separated divisions as children of a node. Lastly, the decision tree trimming eliminates the branches that do not assist in the classification process (Alqudsi and El-Hag, 2019), as shown in Figure 6.14.

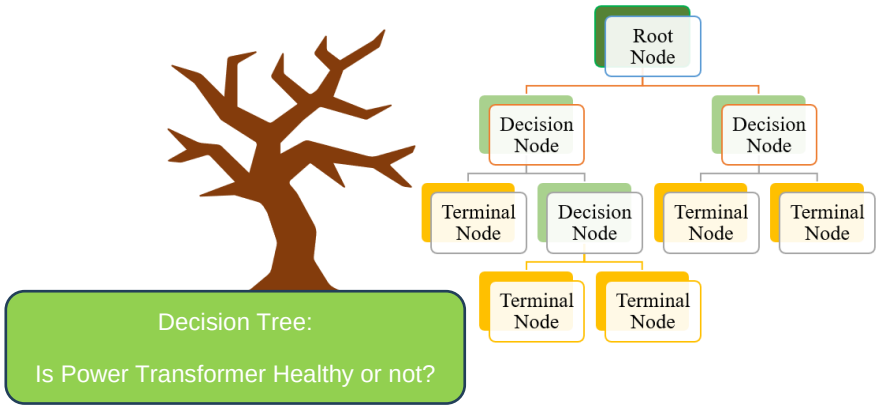


Figure 6.14: Decision Tree (DT) method for the transformer HI.

6.3.2.6 Random Forest (RF)

RF is an ensemble algorithm that uses multiple decision trees, as shown in Figure 6.15. The algorithm's result is the vote agreed by most of the decision trees (Alqudsi and El-Hag, 2019). RF does not face overfitting issues because adding trees reduces generalisation errors. According to Prasojo et al. (2021), RF quickly calculates the relative importance of the input features to the output variable.

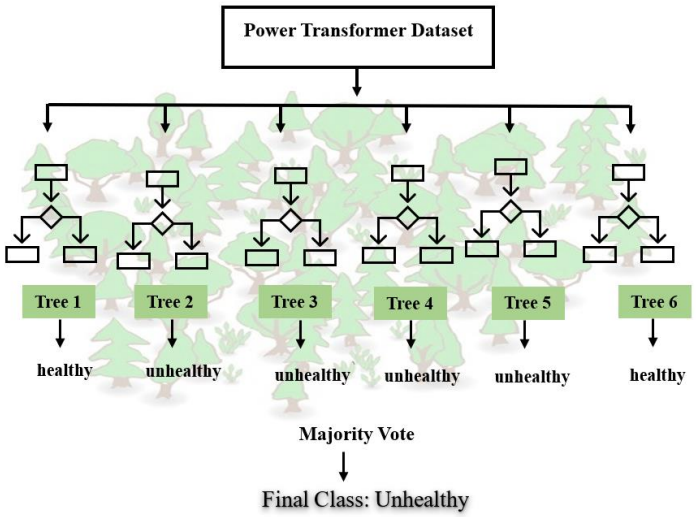


Figure 6.15: Random Forest (RF) method for the transformer HI.

6.3.2.7 K-Nearest Neighbours (KNN)

In the featured space, consider three neighbours according to Euclidean distance. The new instance will allocate the class with most neighbours. The new instance in Figure 6.16 has three neighbours in the circle drawn and assigned as the unhealthy transformer class. As shown, two neighbours are in the unhealthy transformer class and 1 in the healthy class.

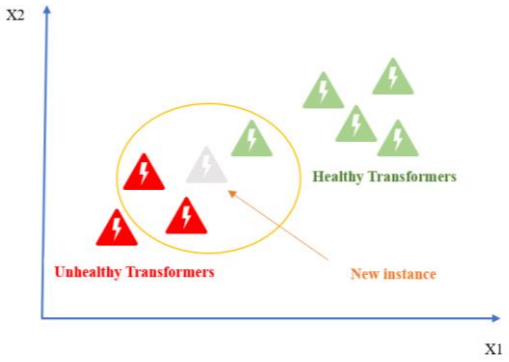


Figure 6.16: K-Nearest Neighbours (KNN) for the transformer HI.

6.3.3 Evaluation Metric Used

This section chose multiple evaluation metrics based on (M and M.N, 2015), including model accuracy, confusion matrix, sensitivity and specificity as described below.

6.3.3.1 Accuracy

$$Model\ Accuracy = \frac{TP+TN}{(TP+FP+TN+FN)} \times 100\% \quad (6.1)$$

To calculate the accuracy of the ANN model.

$$ANN\ PTHI\ Model\ Accuracy = \frac{4+4}{4+0+4+0} \times 100\% = \frac{8}{8} \times 100\% = 100\% \quad (6.2)$$

6.3.3.2 Confusion Matrix

The model can present the values of False Positive (FP), True Positive (TP), True Negative (TN), and False Negative (FN) and illustrate them in a matrix format.

Table 6.1: Demonstration of Binary Classification confusion matrix.

Prediction / Actual	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Where:

- The healthy transformers are considered the positive class or (class 0).
- True Positive (TP) / True Negative (TN): It is “True” because the model predicted it correctly. As it was healthy (0), the model predicted it as healthy (class 0), or it was unhealthy (class 1), and the model predicted it as unhealthy or as (class 1).
- False Positive (FP): False prediction of the positive class, which is healthy or as (class 0).
- False Negative (FN): False prediction of the “Negative” class, which is unhealthy or as (class 1).
- False Positive (FP) / False Negative (FN) are false predictions or mistakes the model makes in predicting the right class type.

6.3.3.3 Sensitivity

It can describe the model's ability to predict the positive class correctly, also known as the TP rate. In the PTHI problem, the positive class is the healthy transformers (class 0).

6.3.3.4 Specificity

It can describe the model's ability to correctly predict the negative class, also known as TN rate, where the negative class in our problem is the unhealthy transformers or (class 1).

6.3.3.5 Results and Evaluation: Models Comparison

This section compares the ML models on the available 30-transformer dataset with algorithms, namely ANN, BLR, DT, KNN, RF, SVM, and NB models. Using model accuracy as an evaluation metric, the Artificial Neural Network and Random Forest models perform the best with 100% accuracy, followed by the Naïve Bayes and Decision Tree model at 92%, and both Binary Logistic Regression (BLR) and Support Vector Machine (SVM) perform equally similarly with a value range of around 88%. At the same time, KNN performed the least, with an accuracy rate of 75%.

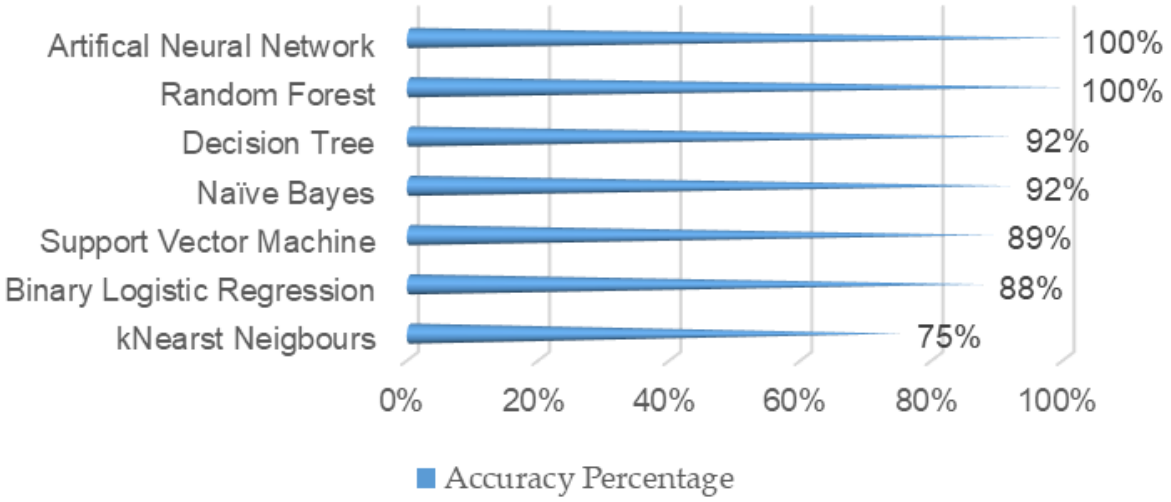


Figure 6.17: Comparison of ML Models' Accuracy.

Figure 6.17 compares machine learning model accuracy based on 30% of the 30 transformer records. The dataset considers 70% training and 30% testing sets.

Table 6.2 shows a consequence demonstration of the model's ability to predict correct classes. The seven models present quality and accuracy insights explained in the remarks section within the same table.

Table 6.2 presents a detailed self-explanatory insight based on the findings using the second metric for evaluating the model's performance and accuracy. This finding shows each ML model's accurate and false predictions based on healthy (0) and unhealthy (1) classes. The classes justify the required decision to be considered by the asset management team. Indeed, this will lead to attention to the faulty transformers that require urgent action.

Table 6.2: Confusion Matrix comparison for the experimented ML models.

ML Model	Confusion Matrix	Remarks
Binary Logistic Regression	predicted 0 1 0 4 0 1 1 3	- The model correctly predicted the healthy transformers as healthy four times and did not make a mistake as it predicted healthy transformers (0) as unhealthy (1), zero times. - The model has predicted unhealthy transformer (1) as healthy (0) one time and correctly predicted unhealthy (1) as unhealthy (1) three times. - The model accuracy is 87%.
Artificial neural network	predicted 0 1 0 4 0 1 0 4	- The model correctly predicted healthy transformers (0) as healthy (0) four times and did not make a mistake as it predicted healthy transformers (0) as unhealthy (1), zero times. - The model has predicted unhealthy transformers (1) as healthy (0) zero times and correctly predicted unhealthy transformers (1) as unhealthy (1) four times. - The accuracy of the model is 100%.
Support Vector Machine	predicted 0 1 0 5 0 1 1 3	- The SVM model correctly predicted healthy transformers (0) as healthy (0) five times and did not make a mistake as it predicted healthy transformers (0) as unhealthy (1) zero times. - The model predicted unhealthy transformers (1) as healthy (0) one time and correctly predicted unhealthy transformers (1) as unhealthy (1) three times. - The accuracy of the model is 88%.

ML Model	Confusion Matrix	Remarks
Naïve Bayes	predicted 0 1 0 8 1 1 0 3	- The model correctly predicted the healthy transformer (0) as healthy (0) eight times and mistakenly predicted the healthy transformer (0) as unhealthy (1) one time. - The model has predicted unhealthy transformers (1) as healthy (0), zero times and unhealthy transformers (1) as unhealthy (1), three times. - The accuracy of the model is 92%.
Decision Tree (DT)	predicted 0 1 0 8 1 1 0 3	- Similarly to the above Naïve Bayes Model, the model correctly predicted the healthy transformer (0) as healthy (0) eight times and mistakenly predicted the healthy transformer (0) as unhealthy (1) one time. - The model has predicted unhealthy transformers (1) as healthy (0), zero times and unhealthy transformers (1) as unhealthy (1), three times. - The accuracy of the model is 92%.
Random Forest (RF)	0 1 0 9 0 1 0 3 predicted	- The model correctly predicted the healthy transformer (0) as healthy (0) nine times. - The model has predicted unhealthy transformers (1) as healthy (0), zero times and unhealthy transformers (1) as unhealthy (1), three times.
K Nearest Neighbours (KNN)	predicted on 0 1 0 6 2 1 0 0	- The KNN model has predicted unhealthy transformers (1) as healthy (0) two times but did not correctly predict any unhealthy transformers (1) as unhealthy (1).

Furthermore, sensitivity and specificity are employed as evaluation matrices, as shown in the results provided in Figure 6.18. Regarding sensitivity of the models, which is the model's ability to correctly predict the healthy transformer as healthy. RF, ANN, DT, KNN, and NB have the highest sensitivity value of 1, equivalent to 100%. On the other hand, specificity, the model's ability to correctly identify the unhealthy transformers as unhealthy, is known as the “true” negative rate. All the ML models of RF, BLR, ANN, and SVM have a high specificity value of 100%, whereas the DT and NB have a 75% specificity percentage and KNN has 0% as there are none of the faulty transformers as shown in Figure 6.18.

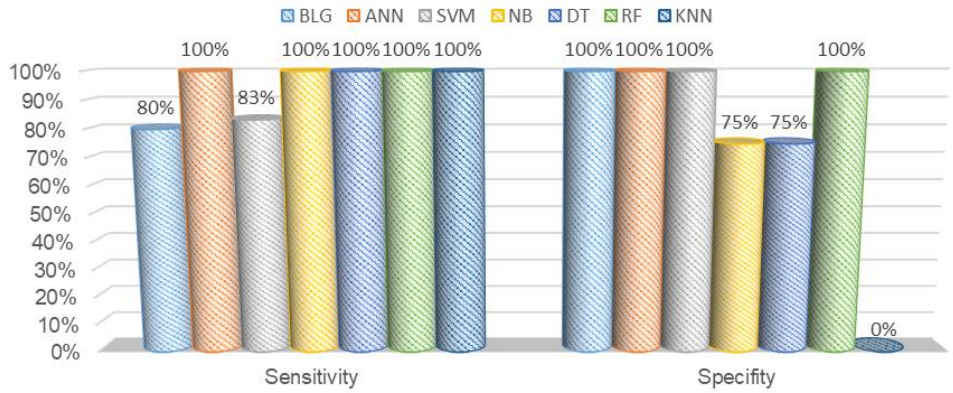


Figure 6.18: Comparison between seven ML models' sensitivity and specificity rates.

Comparison between the result of (Zuo et al., 2016) and the results presented in this chapter is impossible because the researcher used the records of 30 transformers as input and the use of the (SPSS) platform to perform the mathematical calculations of BLR. The dataset is unavailable in training and testing datasets. This concept does not follow a typical data science practice for the PTHI model in this chapter's experiments. This research considers a testing dataset of 30% of the whole dataset of 30 records, which was divided randomly into the R statistical programming language.

Actual group	Predicted		Actual group size	Percentage correctly classified
	0	1		
0 (health)	20	1	21	95.2%
1 (unhealthy)	1	8	9	88.9%
Predicted group size	21	9	30	93.3%

Figure 6.19: Binary Logistic Regression model (Zuo et al., 2016).

However, the proposed ML models' effectiveness in predicting the unhealthy transformers from the above results is similar to that of the earlier conventional mathematical models.

6.4 Summary

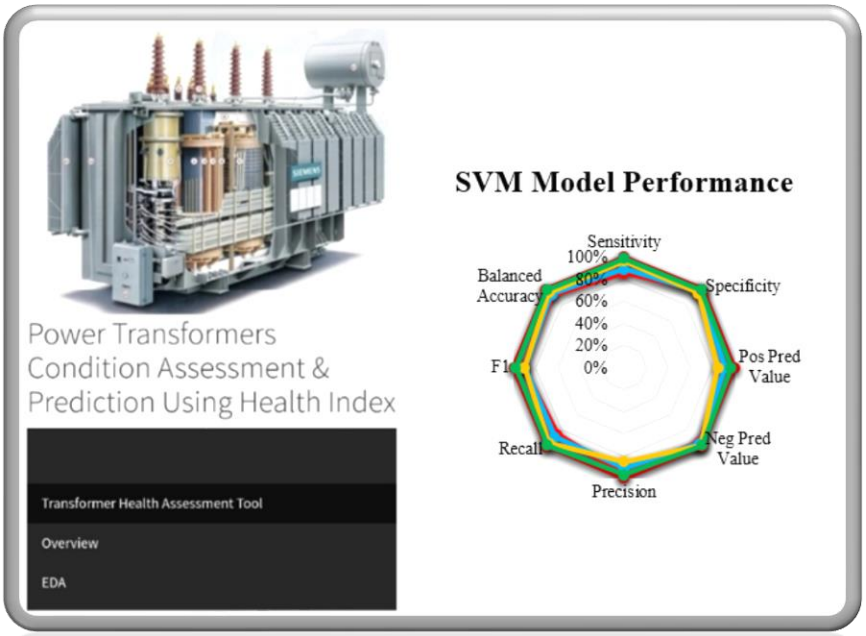
This chapter is an assessment of multiple AI ML models constructed in R programming language, which consist of seven models, namely Artificial Neural Network (ANN), Binary Logistic regression (BLR), Decision Tree (DT), K-Nearst Neighbours (KNN), Random Forest (RF) and Naïve Bayes (NB). This assessment is a new contribution in the research area as earlier conducted studies were carried out in Fuzzy Logic using MATLAB by (Abu-Elanien et al., 2012) and BLR using SPSS by (Zuo et al., 2016), and this research using ML in R programming language. It uses available literature review data to start HI analysis. Features concluded in the transformers HI analysis data include oil liquid (BDV, Acidity, DCGs) and paper solid (2FL contents) isolations.

In recent years, AI using ML models has been vital in improving the calculation accuracy of designing and developing effective and efficient HI models. It is an alternative advanced method to estimate missing data values and uncertainty. From this chapter's experiment results, the Artificial Neural Network algorithm performed the best with a 100% accuracy rate, followed by the Naïve Bayes 91%, and both Support Vector Machine and Binary Logistic Regression models perform equally similarly with a value range of around 88%. In addition, ML models are very reliable in identifying faulty transformers as the specificity rate for 4 out of 7 models is 100%, namely for Artificial Neural Networks, Binary Logistic Regression, Random Forest, and Support Vector Machines.

The next chapter will discuss implementing ML models in R programming language with actual field data collected from the electrical utility. The PTHI model analysis uses a large fleet of power transformers and tested ML algorithms that provide better results when evaluating matrices, including model accuracy, sensitivity, specificity, and confusion metrics.

Chapter 7

Case Study 3 Health Index Support Tool For Driving Maintenance Strategy For Electrical Utility



Chapter 7: Case Study 4, Health Index for Electrical Utility dataset, using Machine Learning and Web App Tool

7.1 Introduction

This chapter presents the electrical utility's data collection, preparation, planning, and analysis before formulating HI. It also includes the decision taken and the process used to prepare a dataset from the received electrical utility files and actions undertaken, the ML modelling algorithm chosen, and the app developed. Good knowledge of the failure mode data was essential in terms of applying the required data science solutions for handling data issues, especially conducting data cleaning, data type casting from text to numbers, removing symbols, deleting or replacing missing data with average, and carrying out a detailed data analysis and modelling. Figure 7.1 represents the workflow based on the research HI assessment investigation. The analysis presents this research's technical and financial impacts for implementing the PTHI tool in assessing asset management effectiveness and efficient decision-making for driving electrical utility daily switching schedules, annual maintenance plans, maintenance strategy and avoiding unnecessary spending for PT lifecycle planning. The electrical utility proposed PTHI analysis workflow:

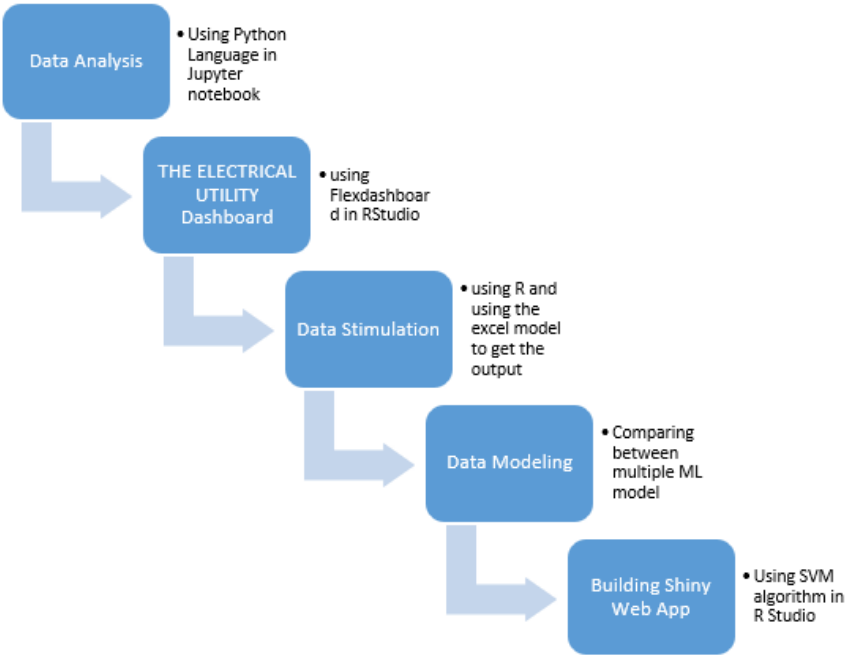


Figure 7.1: The workflow methodology for the actual data investigation.

7.2 Case Study Investigation

This section presents a detailed analysis of the data received.

7.2.1 Data Collection

The data presented in this chapter is from the electrical utility. Two main data files were received. One includes historical laboratory files; the other is the most recent file, including HI results scoring factors according to the utility as nationally agreed scoring factors. The data for this case study consists of OQA analysis, DGA analysis, Furan test, and ageing information. As shown in Figure 7.2, the oldest transformer was installed in 1952 and is 70 years old, whereas the youngest transformer was installed in 2017 at five years old, considering the year 2022 as the reference year.

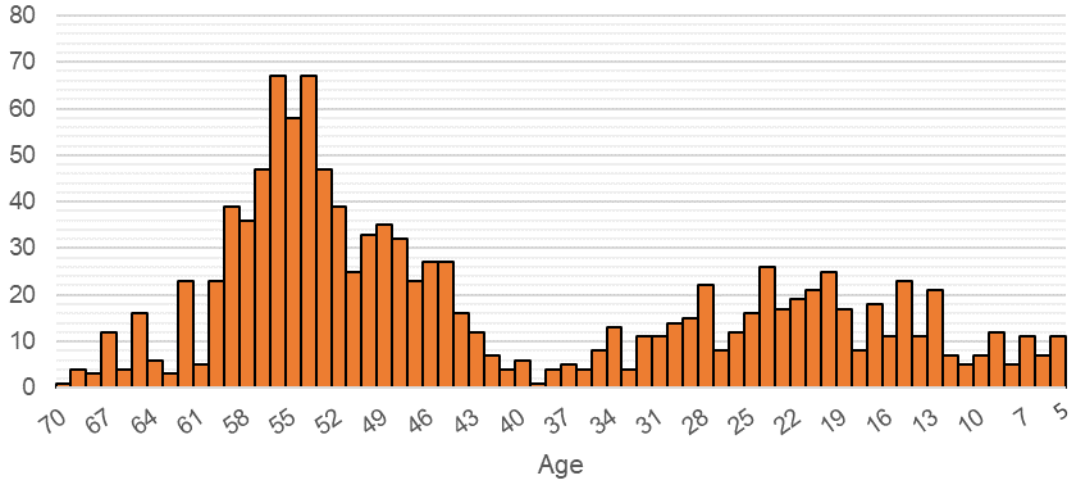


Figure 7.2: PT Age Histogram.

7.2.2 Data Preparation and Validation

This subsection represents the process for validating and preparing data of the 132kV transformers from the second file received from the electrical utility, which means “Asset Data”. Practically, the null values of parameters are to be replaced with the average or deleted. Here, 900 records of DGA values out of the 1231 records are missing, which requires deleting these values. In addition, typecasting performance was vital for some variables, as saving numbers as the text requires changing the text format to float. Also, some empty cells contain a dash symbol, replaced by a blank to calculate the unavailable data in each column. A description of the data main sheet is as follows:

- The “Asset Summary” sheet has data about the transformer (TX) ID, Substation Name, Voltage, Age, and HI Score. After filtering the 132 kV and saving it in an

Excel file, 217 records are available, so the research investigation requires including all 132 kV.

- Obtaining DGA values from the “TX DGA Modifier” sheet, removing the rows without values, and saving them in an Excel file resulted in 285 records for the DGA test. However, these records do not show the voltage range, so the 132kV could not be filtered.
- After obtaining the Oil test result from the “TX Oil Modifier” sheet, removing the rows with no values, and saving the file as an Excel file, 1185 records for the Oil test are available. Again, the voltage range is unavailable, so the 132kV could not be filtered.

However, using Python programming is vital to merge the Excel files into a new file utilising the TX ID as a common field. However, an expectation of getting 217 complete records for all 132kV transformers, but only 41 records are available after the merge operation in Python, showing (ID, Age, Moisture, BDV, Acidity, and DGA (as total not breaking down into individual gases that include H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, CO, CO_2), HI score, which is suspected to be because of the way of writing and defining the TX ID in the three excel files might be different. As explained, from the 132kV 41 records, only 33 fully detailed records were found, as there are null values in other records. Also, the PTs ID finding was previously available in a different format from the second file. Moreover, the age column is not available. In addition, Furan values were hardly included in both files, as six records from the 33 records only have Furan data.

Therefore, manual data validation ensures that the data is correctly defined. This manual validation took around nine weeks, going through the files, identifying the transformers and joining the data. This validation work includes reviewing the first file received, focusing on merging the records, filtering the transformer based on the voltage levels, and saving the files in separate Excel spreadsheets according to voltage levels. Hence, a detailed ML analysis is necessary for the way forward into a further AM framework implementation with this researcher's previous engineering field experience and the earlier learning lessons from literature reviews to develop robust and suitable PTHI tools for prioritising transformer replacement plans and assessing transformer lifecycle health conditions.

7.2.3 Explanatory Data Analysis

An Explanatory Data Analysis (EDA) initially uses Python to explore the electrical utility data. This analysis helped in understanding the dataset characteristics by describing the actual data to enable us to consider the proper action towards solving the issue in the dataset that consists of the missing data and how to conduct the modelling process. This analysis also helped to evaluate the data scientifically as received as described by the data correlation metric. The outcomes there were a few challenges at the beginning related to the actual research data collection, preparation, and analysis from the electrical utility. This analysis considers the time spent collecting most of the data as a whole set. Then, once the data was in numerical value, further data understanding was vital, which was needed to go back to the electrical utility for data interpretation. Still, this was not sufficient as large amounts of data were missing, and the company was not able to provide the missing data, which then resulted in the research with a current challenge to overcome the data issue using data science methods to solve the missing data issues that the model can predict. This data analysis led to the findings from the earlier literature review of previous PTHI approaches. The Fuzzy Logic model in section 5.4 and the conventional model in section 4.4.3.3 are unsuited for significant data investigation.

Figure 7.3 shows the number of transformers based on the voltage classification. The total number of transformers from the large fleet is 1231 units, with the voltages of 132kV,33kV, and 66kV identified. Parameters extracted from routine laboratory checks mainly focus on examining dielectric characteristics through oil quality checks and insulation properties through DGA.

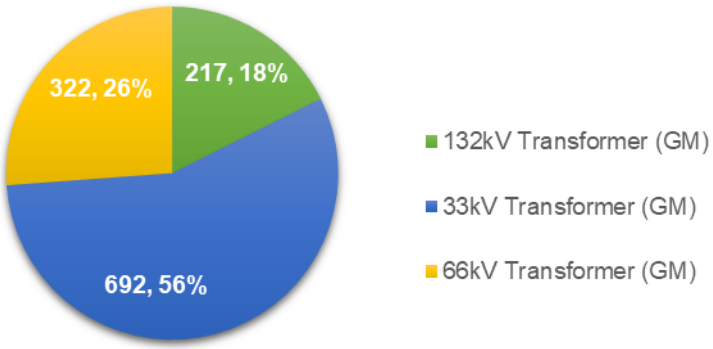


Figure 7.3: Number and percentage of transformers by voltage.

Table 7.1 shows the statistical analysis of the parameters concerning the minimum, maximum, standard deviation, and average for each column. This dataset shows that the transformers' age ranged from 5 to 70 years and averaged 41.6 years.

Table 7.1: Statistical summary of the electrical utility data*.

	Age	Moisture	Acidity	BDV	DGA	Furfuraldehyde	Health Index
Mean	41.669198	21.957598	0.132945	63.772151	267.298246	0.187878	3.153503
Std	17.164061	170.009150	0.905780	12.619900	301.143476	0.445652	1.532753
Min	5.000000	5.000000	0.010000	14.000000	0.000000	0.000000	0.620434
Max	70.000000	5861.000000	23.050000	302.000000	2600.000000	6.876670	5.500000

*The utility equation is not provided; the research applied HI equation 4.1 in Chapter 4.

Figure 7.4 shows a correlation metric result obtained automatically using Python. This result helps to explain the relationship between each variable and how the variable's correlation with the HI score is vital for more analysis. The age parameter shows a high positive correlation with HI, with a value of 0.91.

The negative values imply a contrary relationship to the transformer condition as follows: BDV (-0.31), DGA (-0.12), and Furfuraldehyde (-0.0076). Applying a value limit above 0.2, most parameters in this dataset show a low correlation with the HI score parameter except for BDV (-0.31). The value of 1 illustrates that the dataset parameter has an evenly balanced relationship with itself, as shown in the relation of Age with Age. Additionally, realising the correlation between the dataset parameters can facilitate cutting the number of parameters without drastically compromising the accuracy of results.

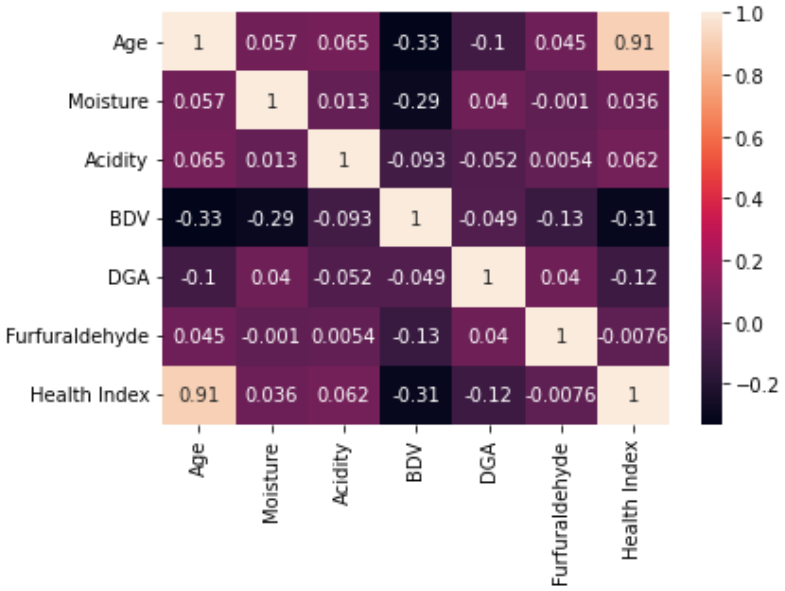


Figure 7.4: Correlation metric of the electrical utility data.

Figure 7.5 shows the relation between each parameter and the HI. It also shows the outliners that form issues and must deal with them before feeding the data to the HI ML model. This collection of plots shows the relationship between the input parameters among themselves and the output variable (HI). It will also show the outliers who are far from the common points. Hence, a pair plot, in other words, shows the visualisation of the relationship between various plots related to the HI and input parameters. In this study, age correlates more with HI as it shows a positive linear relationship. In the scattered plot, dots in some plots are linearly, and a few outliers indicate the condition of the transformers impacted by these test parameters. The histogram plot shows the transformers categorised based on the age, BDV, DGA, and Furan group concerning the HI scores that indicate the condition of these transformers. The histogram summarised the dataset distribution analysis obtained from Data Exploratory Analysis processed in Python to show the visualisation of the collected datasets. The data distribution is on an interval scale, whether discrete or continuous.

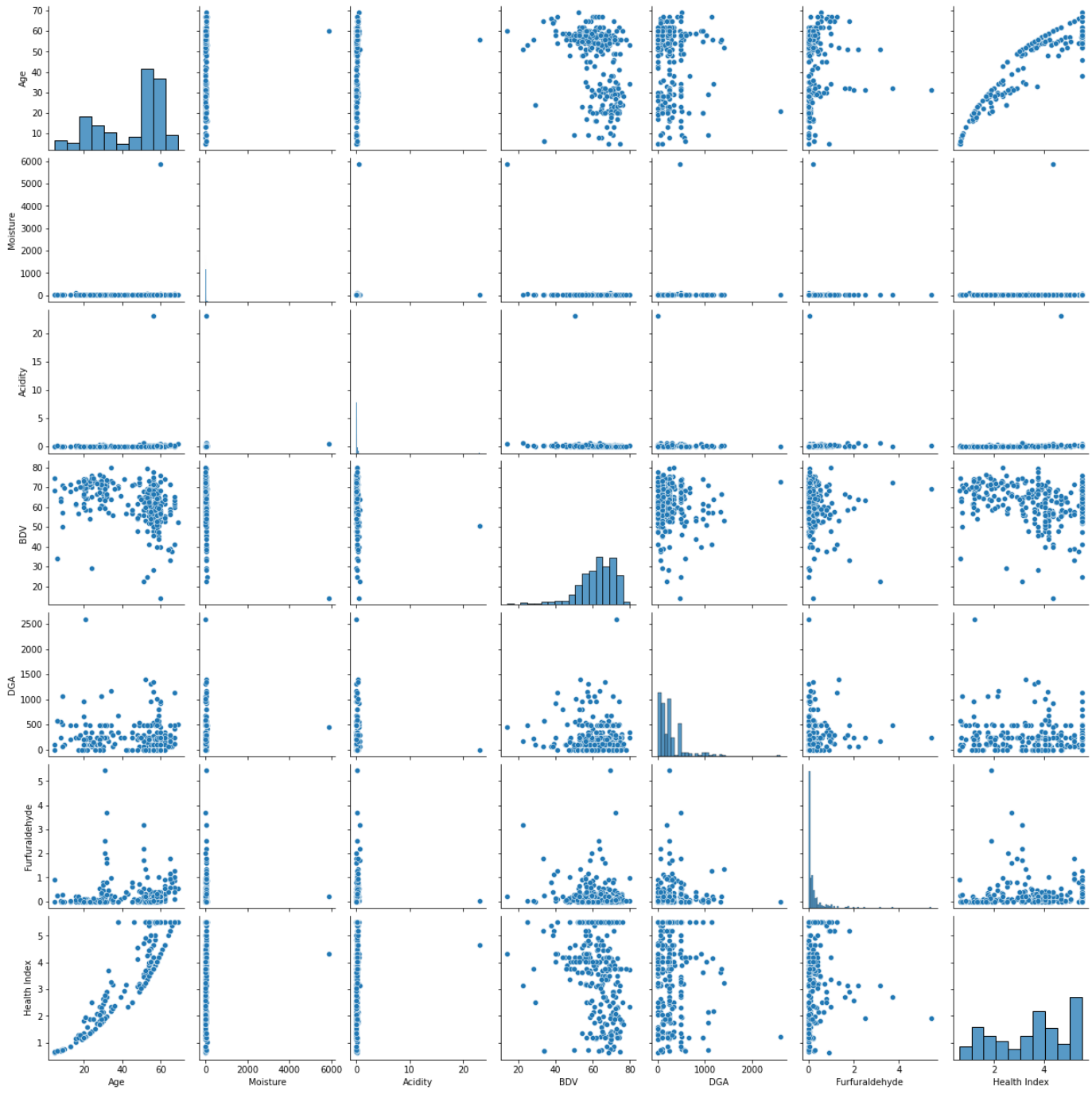


Figure 7.5: Multiple scatterplots of available parameters.

Foros and Istad (2020) stated that the Asset management team should prioritise asset transformer replacement plans. Figure 7.5 shows the effect of prioritising the transformer replacement plan due to good asset management technology and improvements in maintenance, technology and condition assessment techniques.

Figure 7.6 shows a reduction in the electrical utility's investment plan for new transformers, enabling a better condition assessment. This HI improves the maintenance decision and extends the transformer life cycle. For clarity, huge investments led to the installation of new transformers between 1952 and 1980, and based on these findings, the installation of new transformers was reduced from 1984 to 2017.

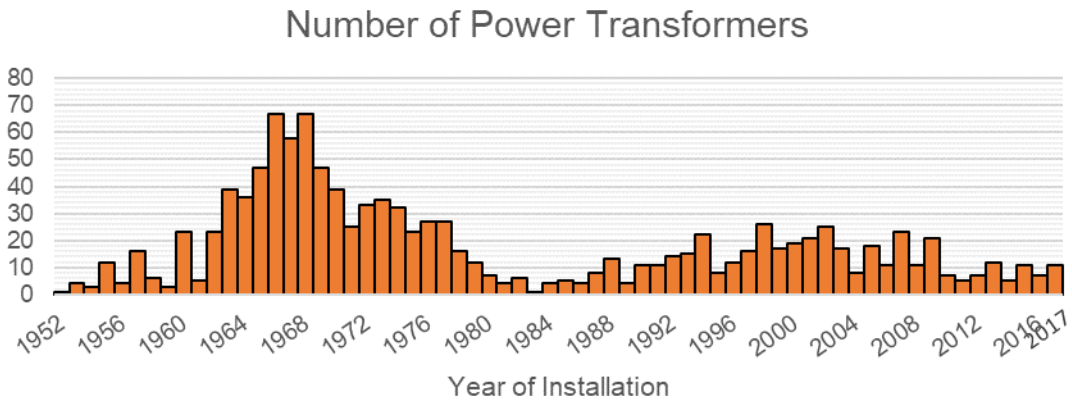


Figure 7.6: PT Histogram, Year of Installation.

Figure 7.7 visualises the statistical relationships among HI score and transformer age parameter analysis as the curve changes drive transformer age, similar to the earlier correlation matric findings in Figure 7.4. This finding clearly shows a strong positive correlation between the HI score and age score, which could state that in the electrical utility, as the transformer ages, the HI increases, which means that as the transformer ages, the transformer's health condition worsens. Undoubtedly, the findings of the analysis of the collected data could be different, but future changes in the data could enable different findings.

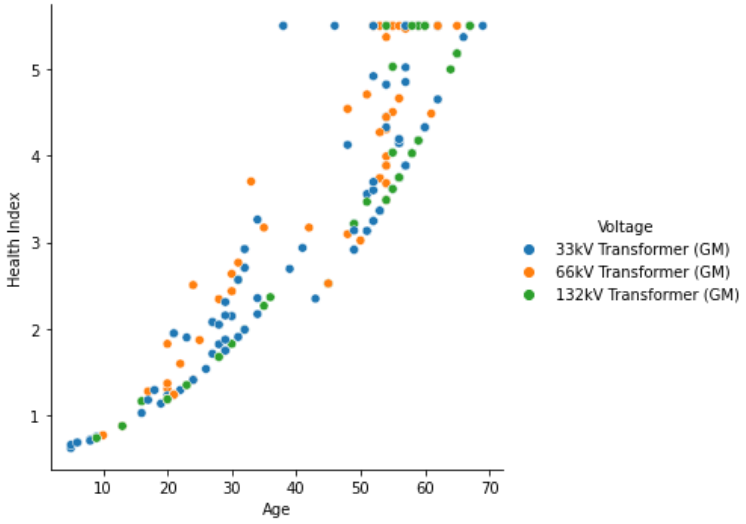


Figure 7.7: The electrical utility data scatter plot showing the age and HI relationship.

Figure 7.8 presents density plots for the electrical utility dataset, including density plots for Age and BDV, which examine the distribution of some variables in a dataset.

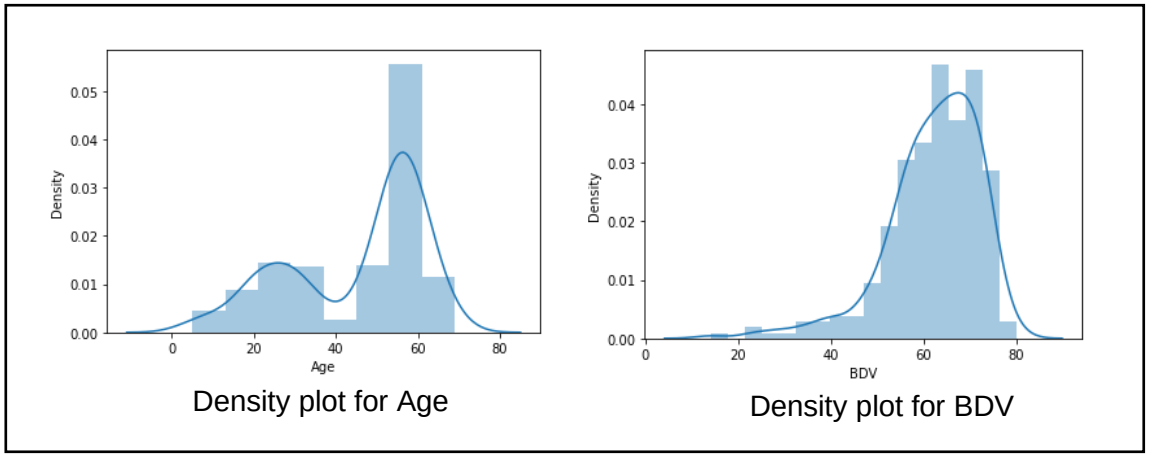
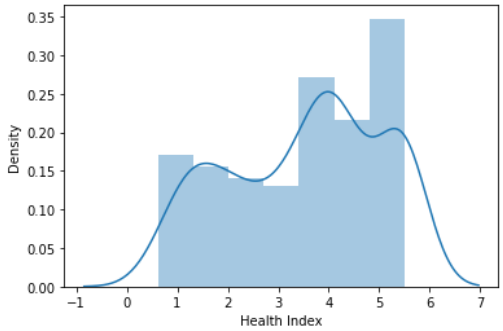


Figure 7.8: The electrical utility data density plot for age (left) and BDV (right).

As shown in Figure 7.8 above, the age is around 50, which means that most of the transformers are getting older. At the same time, the BDV concentration is around 50 -78, which means that most transformers have good BDV values based on the IEC 60156 (2018) standard for describing the dielectric breakdown.

Figure 7.9 presents another density plot for HI scoring parameters. According to the scoring rules, most electrical utility transformers are in “Very Good” and “Good” conditions.



The HI Density plot

Range	Condition
0 - 0.5	Very Good
0.5 - 5.4	Good
5.5 - 9	Fair
10 - 100	Poor
>100	Very Poor

The Electrical Utility Scoring Rules

Figure 7.9: Density plot of the electrical utility HI (left) and its scoring rule (right).

7.2.4 Further Data Preparation and Steps Taken for Modelling.

An analysis plan is a solution for handling the electrical utility dataset by considering two data sources from the electrical utility. The first source is the historical laboratory data, including the most recent test data per the (lims file). The second source is the electrical utility “Asset Data”, with HI already calculated per some stations' utility scoring rules.

In the laboratory data file are data for OQA, DGA in detail, Furans, historical records, and no age. In this case, the question is whether developing a model of three factors is practical since the company is not considering the Furan test for all transformers. This test is expensive and not considered unless it is necessary. Hence, other options within the electrical utility second file are the solution. The best decision to consider is the following parameters from the lims file:

1. Oil analysis parameters.
2. Detailed parameters of DGA analysis.
3. Furan.
4. Historical maintenance record: this requires developing a new model.
5. Age: It was not considered temporary, but within the Excel model, a column with empty data is before getting the final PTHI calculation values. Likewise, the electrical utility and this research are both considered to develop the age data from “Asset Data” files based on the required fitting values. Otherwise, the age importance will be less and require deletion from the Excel formulation. This age

justifies the literature review data that age is one of the data sources for developing good HI, as stated in section 3.6.1.1.

The first process is to use the detailed parameters of three primary analyses (OQA, DGA and Furan) in a mathematical rule-based model in Excel to produce the PTHI as an output variable on the historical laboratory maintenance records (lims file) for each voltage level separately.

The second process was to develop the ML data model for each voltage level separately, considering the primary variables, including individual gas values and oil parameter values as independent variables and the PTHI output as dependent variable, without considering the mathematical rule-based model utilising the scoring and weighting factors.

7.3 Using Power Transformer Health Index Scoring Model in Excel

After filtering the 132kV transformer data from the electrical utility and lims file, a weighted average scoring model was used in Excel and saved as "AstMgmt132kVData.xlsx". Also, the Furan value of values <0.1 was replaced with a value of 0 because the less than sign was causing an error in the formula. In addition, after creating the scoring model in Excel. The rules were checked and modified according to those based on (Jahromi et al., 2009) as the oil parameters modification is according to the (IEEE C57.106, 2015) ranking method for oil tests for voltage levels between 69kV and 239kV.

Furthermore, the interpretation of colour was a minor issue that needed correct handling using the conventional mathematical model. According to (Jahromi et al., 2009), following the IEC recommendation by CIGRE to follow ISO 2049, the transformer oil colours are converted into scores as presented in Table 7.2.

Table 7.2: Transformer oil colour score and weight factors (Jahromi et al., 2009).

	Range	Category	Score	Weight
Color	<=1.5	A	1	
	1.5-2.0	B	2	
	2.0-2.5	C	3	
	>=2.5	D	4	2

However, the classification by colour name is vital. Hence, the “Amber” indicates “Good” oil condition, and the “Yellow” indicates “Bad” oil (IEEE C57.106, 2015). Therefore, the following table classifies the soil conditions into five categories. Table 7.3 shows the electrical utility oil colour interpretation into nominal values. The same applies to transformers rated 66kv and 33kv using the rules based on (Jahromi et al., 2009), oil standards (IEC 60599, 2022; IEEE C57.106, 2015; and DGA standards (IEC 60505, 2011; IEEE C57.104, 2019).

Table 7.3: Interpretation of transformer oil colour.

Colour	Nominal Value Range	Condition	Score
Light Amber	0 ≤ colour < 1.2	Very Good	1
Amber	1.2 ≤ colour < 1.5	Good	2
Pale Yellow	1.5 ≤ colour < 2.0	Moderate	3
Yellow	2.0 ≤ colour < 2.5	Bad	4
Dark Yellow	colour ≥ 3.0	Very Bad	5

7.4 Comparison: HI Calculated by Electrical Utility and PTHI Scoring Model

The PTHI score evaluates transformers' end-of-life, a sign of transformer failure, and risk expectancy. This comparison consequently aids in applying preventive maintenance plans. The primary conventional method used in this study is assigning a weighting and a scoring factor to each diagnostic test according to IEC and IEEE standards to calculate the PTHI value. This value interpretation is simplified linguistic terms that describe the transformer condition, as Table 4.2 in Chapter 4 shows.

The electrical utility HI stated that 208 transformers rated 132kV were all in “Very Good” condition. Figure 7.10 shows the result of the scoring model and the electrical utility HI. The result shows the model built in Excel has split the 208 transformers, which were classified as “Very Good” into 39 transformers in “Very Good”, 153 transformers as “Good”, and 16 transformers as “Fair”.

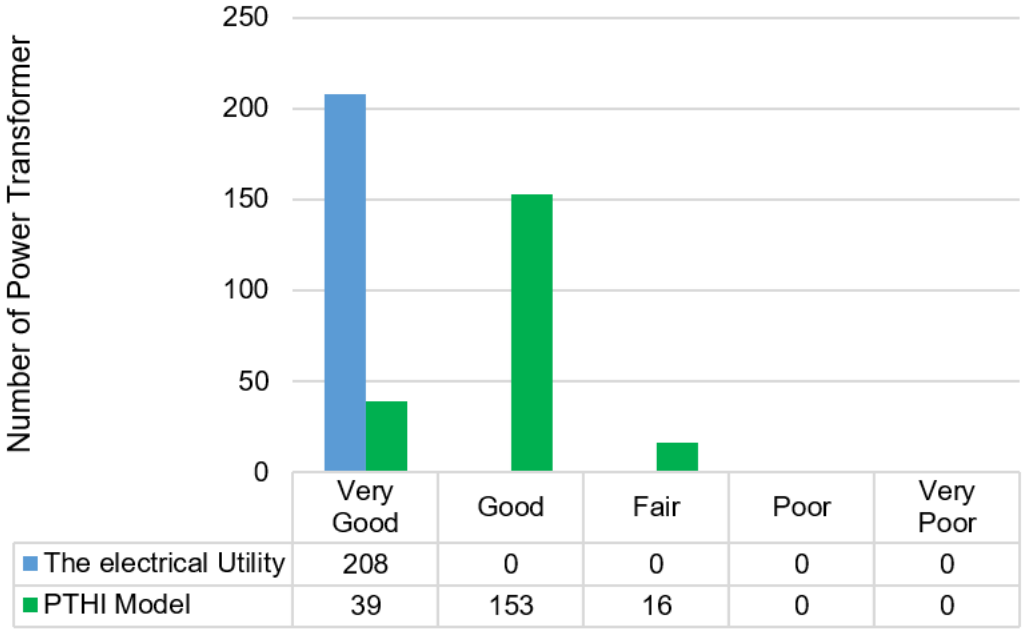


Figure 7.10: Comparison between the electrical utility HI and PTHI model.

7.5 Analysis of Electrical Utility Data Using the Flex Dashboards

This section represents an analysis of the electrical utility transformers' data using the flex dashboards in R statistical language and the experiment conducted to develop an ML model, which is used later in a tool as a web application. The built ML PTHI model will not predict the HI scores. The HI scores are well known as a number that could vary according to electrical utility agreed scoring factors. Therefore, setting the optimal annual maintenance action is the primary goal of the maintenance strategy. The ML algorithm primarily aims to predict the transformer and the optimal annual maintenance action classes. The ML tool provides the electrical utility with the health condition class of a given transformer. It considers the drivable decision-making process, whether preventive or corrective actions, to ensure proper transformer health condition management of a large fleet during in-service optimal lifecycle operation.

To investigate the HI of transformer analysis has been developed, a flex dashboard is vital to visualise the result and deliver valuable information to AM decision makers.

7.6 Data Imbalance Problem

SVM (Support Vector Machine) is a well-known and approved algorithm for its high accuracy. SVM algorithms are considered good, supervised ML algorithms. SVM can

solve problems related to classification and regression line problems. It draws a hyperplane with the largest side-line in the multidimensional space (Arshad et al., 2021). The longer distance of the margin between two categories shows a better expectation level for categorisation.

However, the class issue is that the dataset is imbalanced and does not represent equally all data classes. In the PTHI model, the class is in a transformer condition since no records are available for “Bad” and “Very Bad” conditions. It is considerable that when a dataset does not represent all classes of data equally, the model usually might be overfitted to the “Good”, “Very Good”, and “Fair” classes, which are represented more in the dataset, and the model becomes unaware of the existence of the minority class. This problem may even give a model a good accuracy level but fail miserably in real life (Raj, 2019).

The simulation of records is vital to solving the imbalance of data. A simulation for the transformer missing data classes of “Bad” and “Very Bad” per IEEE rules for DGA and the Oil grading method for transformers under 69kV are available in Appendix H, Table H-1 for DGA and Table H-2 for OQA. Also, perform the same stimulation idea for 132kV to stimulate “Bad” and “Very Bad” scenarios using IEEE rules and oil grading methods of transformers greater than 69 and less than 230kV, which are shown as well in Appendix H 132kV Scoring rules Tables H-26 to H-31. After that, these simulated data records were injected into the PTHI Excel model to get the HI and consequently obtain a dataset with all scenarios that will be available for later building an ML model.

Due to the dataset imbalance problem in Figure 7.11, an oversampling technique is better for the minority classes by stimulating more data to ensure the predictive model can predict the cases of the minority class correctly. The count of the real dataset for under 69kV transformers includes both records for 66 And 33kV transformers since the simulated dataset record applies the same rules of IEC and IEEE standards as shown in Appendix H, Tables H-1 and H-2.

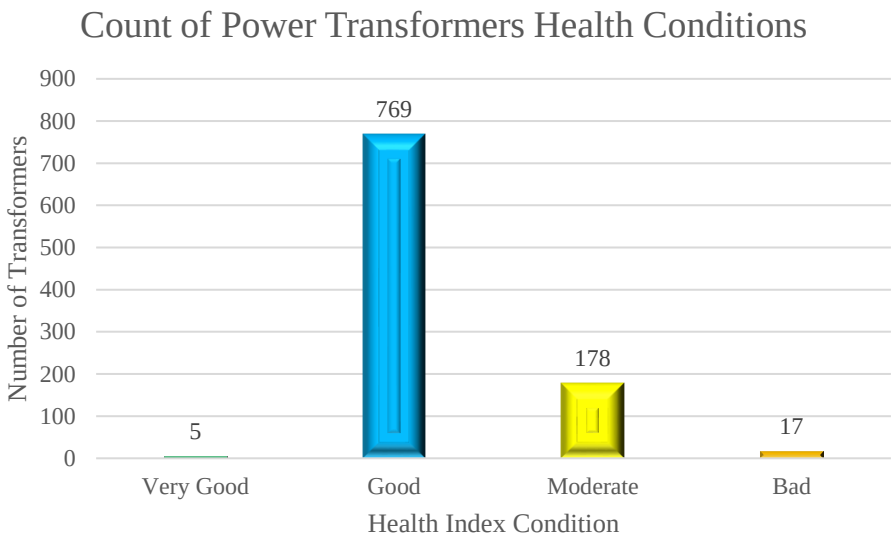


Figure 7.11: Original electrical utility records based on condition.

Initially, fuzzy rules and parameters range according to the IEEE and IEC standards to specify the inputs and output class of stimulated data for the Very Good, Moderate, Bad, and Very Bad classes. However, the dataset used resulted in a low accuracy of 24%, which is not accepted.

As a result of low SVM model accuracy, the decision to double-check the output columns using the conventional model built in Excel earlier resulted in a different output. The output parameters were considered from the Excel ranking model as it has earlier proven high reliability when comparing PTHI with the electrical utility HI results and comparing PTHI with other researchers' journal papers. Then, the newly updated dataset is available, merged with the actual dataset, and saved as DataSetUnder69kVUpdate.csv.

Understanding the structure of the data frame is essential. This understanding can help remove the unwanted columns and prepare the data for modelling. It creates a data frame utilising only the columns needed for the SVM model's prediction. Also, converting the columns to factor. Hence, this is considerably by the SVM classifier.

A part of understanding the data frame is using the summary function to find the minimum and maximum mean for each column and how much data is missing from each column, as shown in Figure 7.12. The Furan has the most columns with missing

data, with 265 records. In contrast, other columns, including DF, have 23 NAs, and Acidity and BDV have 9 NAs filled by the column's mean value.

```
> summary(TCAdata)
  Health_Index    condition         Maintenance_Action
 Min.   :15.32   Length:3560        Length:3560
 1st Qu.:38.71   Class :character   Class :character
 Median :62.10   Mode  :character   Mode  :character
 Mean   :51.57
 3rd Qu.:68.55
 Max.   :91.94

      H2                 CH4                 CO
 Length:3560        Min.   :   0.0427   Min.   :   1.353
 Class :character   1st Qu.:   4.7417   1st Qu.: 113.846
 Mode  :character   Median :  62.3886   Median : 313.547
                    Mean   : 213.8219   Mean   :1496.253
                    3rd Qu.: 198.5339   3rd Qu.: 892.526
                    Max.   :2207.0000   Max.   :9991.405

     CO2                C2H4               C2H6
 Min.   :    0.01   Min.   :   0.010   Min.   :  0.000
 1st Qu.: 1005.69   1st Qu.:   6.645   1st Qu.:  3.803
 Median : 2129.80   Median :  43.122   Median : 50.596
 Mean   : 3150.03   Mean   : 164.216   Mean   :163.889
 3rd Qu.: 3995.86   3rd Qu.:  99.549   3rd Qu.:100.341
 Max.   :12210.00   Max.   :4230.000   Max.   :999.946

      C2H2              Moisture            BDV
 Min.   :  0.0094   Min.   :    0.25   Min.   :  0.026
 1st Qu.:  0.4317   1st Qu.:   24.00   1st Qu.: 44.056
 Median :  3.5342   Median :   51.68   Median : 70.000
 Mean   :136.9799   Mean   :   60.96   Mean   : 85.726
 3rd Qu.: 34.3417   3rd Qu.:   76.07   3rd Qu.:121.692
 Max.   :994.9239   Max.   :17366.00   Max.   :610.000
                                       NA's   :9

       DF                Acidity              Age
 Min.   : 0.00004   Min.   :  0.00002   Min.   :  0.01362
 1st Qu.: 0.12100   1st Qu.:  0.14000   1st Qu.: 13.13593
 Median : 6.90453   Median :  3.50882   Median : 45.68978
 Mean   :26.83645   Mean   : 25.65603   Mean   : 43.03884
 3rd Qu.:53.11472   3rd Qu.: 50.52656   3rd Qu.: 59.00000
 Max.   :99.97525   Max.   : 99.99730   Max.   :119.89017
 NA's   :23         NA's   :9
     Furan
 Min.   :0.00024
 1st Qu.:0.04000
 Median :0.09099
 Mean   :0.34678
 3rd Qu.:0.69374
 Max.   :6.20000
 NA's   :265
```

Figure 7.12: Result of Summary function using R programming language.

After rebuilding the SVM Classifier to estimate the PT health condition, the model has 414 support vectors and obtained an accuracy of 79%. The model accuracy improved by learning from previous experiments. Hence, further investigation is vital repeatedly and considerably. This time, the class “Very Good” had only 26 records, a low number comparable to other classes. As a result, there are fewer records for the “Very Good” class for balancing the dataset, and more records are needed to continue the ML investigation so the algorithm can learn the condition classes. The updated model of “svm_ConditionClassifier” shows that it has 1755 support vectors and is used to predict five classes. It achieved an accuracy of 83%, which is quite good and acceptable, as shown in Figure 7.13.

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  1755

 ( 137 239 327 582 470 )


Number of Classes:  5

Levels:
 Bad Good Moderate Very Bad Very Good

10-fold cross-validation on training data:

Total Accuracy: 83.95568
Single Accuracies:
 84.36214 84.01639 84.42623 84.77366 83.60656 86.47541 81.06996 83.60656 81.9
6721 85.2459
```

Figure 7.13: SVM model description.

Following the previous experiment, another predictor classifier for the same dataset and a new investigation. This new classifier output intends to link the condition with related decisions and primarily three actions: inspection, repair or refurbishment, and replacement. This output reflects the condition specified using the ML algorithm. This experiment is considerably successful as it achieved an accuracy of 86%, as shown in Figure 7.14.

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  1721

 ( 343 796 582 )


Number of Classes:  3

Levels:
 Inspection Repair & Refurbishment Replacement

10-fold cross-validation on training data:

Total Accuracy: 86.78703
Single Accuracies:
 86.00823 86.47541 84.83607 87.65432 85.65574 87.29508 87.65432 87.70492 84.83607
 89.7541
```

Figure 7.14: SVM model description with an accuracy of 86%.

Furthermore, during the modelling experiment, an error occurred, and a solution that ensured the data pre-processing was performed correctly by applying a normalisation function to each column and replacing empty cells with 0. This function solves the error and improves the model accuracy to 93%, as shown in Figure 7.15.

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  470

 ( 140 162 168 )


Number of Classes:  3

Levels:
 Inspection Repair & Refurbishment Replacement

20-fold cross-validation on training data:

Total Accuracy: 93.43186
Single Accuracies:
 96.69421 93.44262 95.90164 90.98361 91.80328 91.73554 94.2623 96.72
131 90.98361 93.44262 90.08264 95.08197 90.98361 95.08197 94.2623 9
5.04132 91.80328 91.80328 95.90164 92.62295
```

Figure 7.15: SVM model with an accuracy of 93%.

7.7 Implementation of More ML Algorithms on the Electrical Utility Under 69kV Transformers Data (Descriptive Quantitative Comparative Analysis in Spreadsheet)

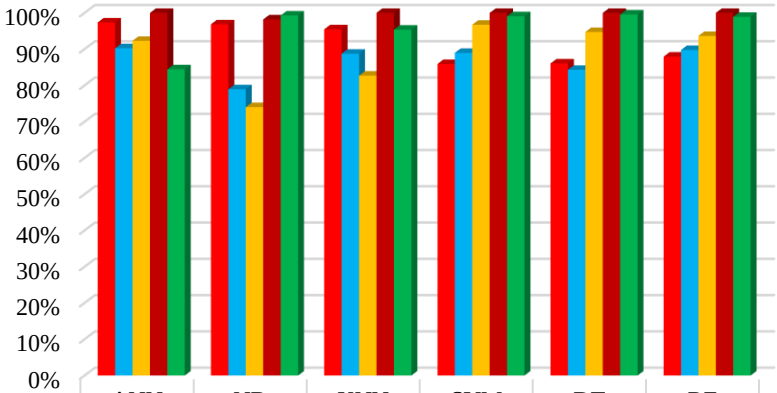
This section answers the question about applying four KPI measures to evaluate different model performances: accuracy, sensitivity, specificity, and F1 score. It

includes implementing the ML models applied in the previous chapter, except Binary logistic Regression, which applies with different principles. The ML problem is treated in this chapter as a multi-class classification, not as a binary classification.

This chapter utilised six machine models for the electrical utility data for transformers under 69kV. Details are included in Appendix H Part 3. The results of the four measures are available below:

7.7.1 Sensitivity

Sensitivity measures for SVM are equally “Good” for the “Moderate”, “Very Bad”, and “Very Good” classes, which are above 96%.



	ANN	NB	KNN	SVM	DT	RF
Class: Bad	97%	97%	95%	86%	86%	88%
Class:Good	90%	79%	89%	89%	84%	90%
Class: Moderate	92%	74%	83%	97%	95%	94%
Class:Very Bad	100%	98%	100%	100%	100%	100%
Class: Very Good	84%	99%	95%	99%	100%	99%

Figure 7.16: Sensitivity measure for ANN, NB, KNN, SVM, DT and RF.

7.7.2 Accuracy

The RF and SVM models performed the best, with an accuracy of around 93.45% and 93.72%. In contrast, NB performed with the lowest accuracy of around 87.7%. The models' results are comparable with other literature review findings (Alqudsi and El-Hag, 2019), as ANN was not the best in their research. This finding depends on the data types, their nature, and the conclusion within the research area of the previous researcher.

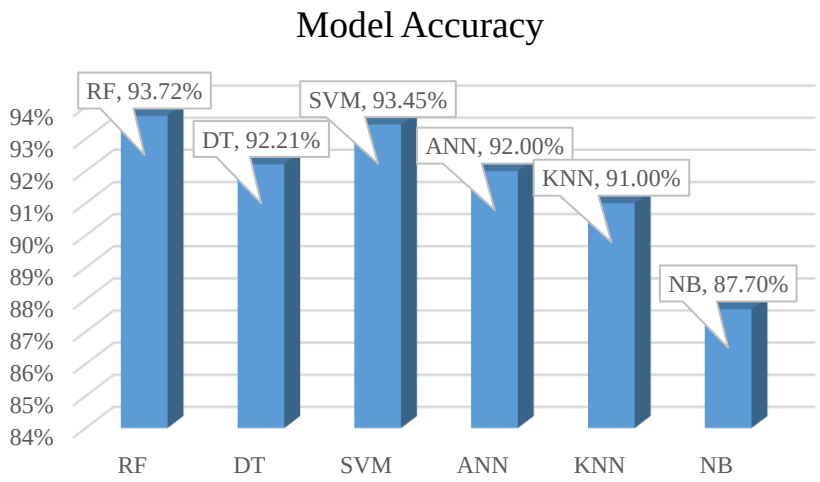
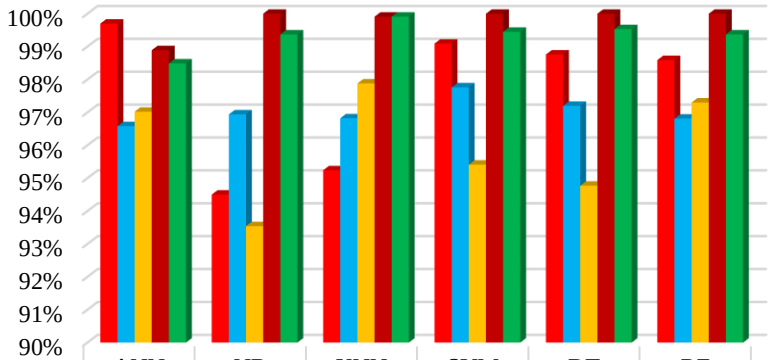


Figure 7.17: ML model's accuracy.

7.7.3 Specificity

The highest specificity measures for class Very Bad are all developed models, with values above 98%, followed by Class Very Good. Where ANN has the highest specificity for class Bad, SVM performs well, and the SVM for classes Bad, Good, Very Bad, and Very Good is above 97.7%.



	ANN	NB	KNN	SVM	DT	RF
Class: Bad	100%	95%	95%	99%	99%	99%
Class:Good	97%	97%	97%	98%	97%	97%
Class: Moderate	97%	94%	98%	95%	95%	97%
Class:Very Bad	99%	100%	100%	100%	100%	100%
Class: Very Good	98%	99%	100%	99%	100%	99%

Figure 7.18: Specificity measure for ANN, NB, KNN, SVM, DT and RF.

7.7.4 F1 Score for Bad and Very Bad Scores

The F1 score is an accuracy metric to show the percentage of the correct prediction for each class. This identification is a harmonic means of precision and recall. It was

considered for class “Bad” and “Very Bad” because if the transformer condition is “Bad” and the model mispredicts it as “good”, the effect will be catastrophic if neglected and the necessary action is incorrect, which can lead to not including the required budget for the related action in a real practical case. Figure 7.19 shows that.

- For the “Very Bad” class, SVM, RF, and DT have an F1 score of 100%, which means that the model made no errors in predicting the “Very Bad” class.
- For class: Very Good, DT performed the best with a score of (99%), followed by SVM, NB, and RF with an F1 score of 98%, and KNN performed 97%. In contrast, ANN performed the least, scoring 89%.

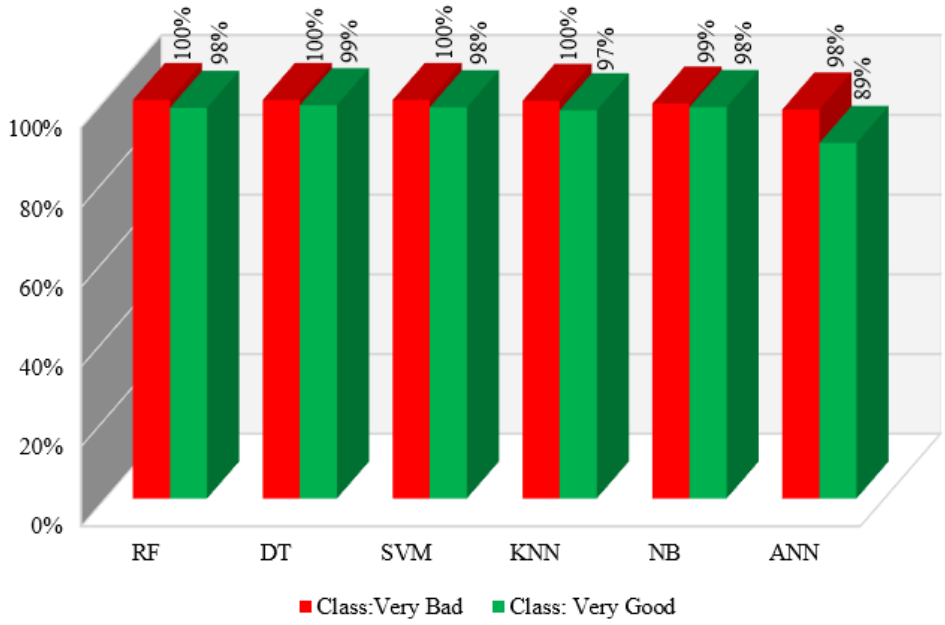


Figure 7.19: F1- Score Comparison.

Figure 7.20 shows the KPI measures used to check the performance of the different classes in all six algorithms implementation using electrical utility data. The KPIs considered are the evaluation metrics of balance accuracy, “positive” prediction (pos pred) value, negative predictive (neg pred) value, precision, specificity, sensitivity, and F1 score. The score shows good model performance, especially when the evaluation matrices are close to 1 or near 100%, as shown in Figure 7.20.

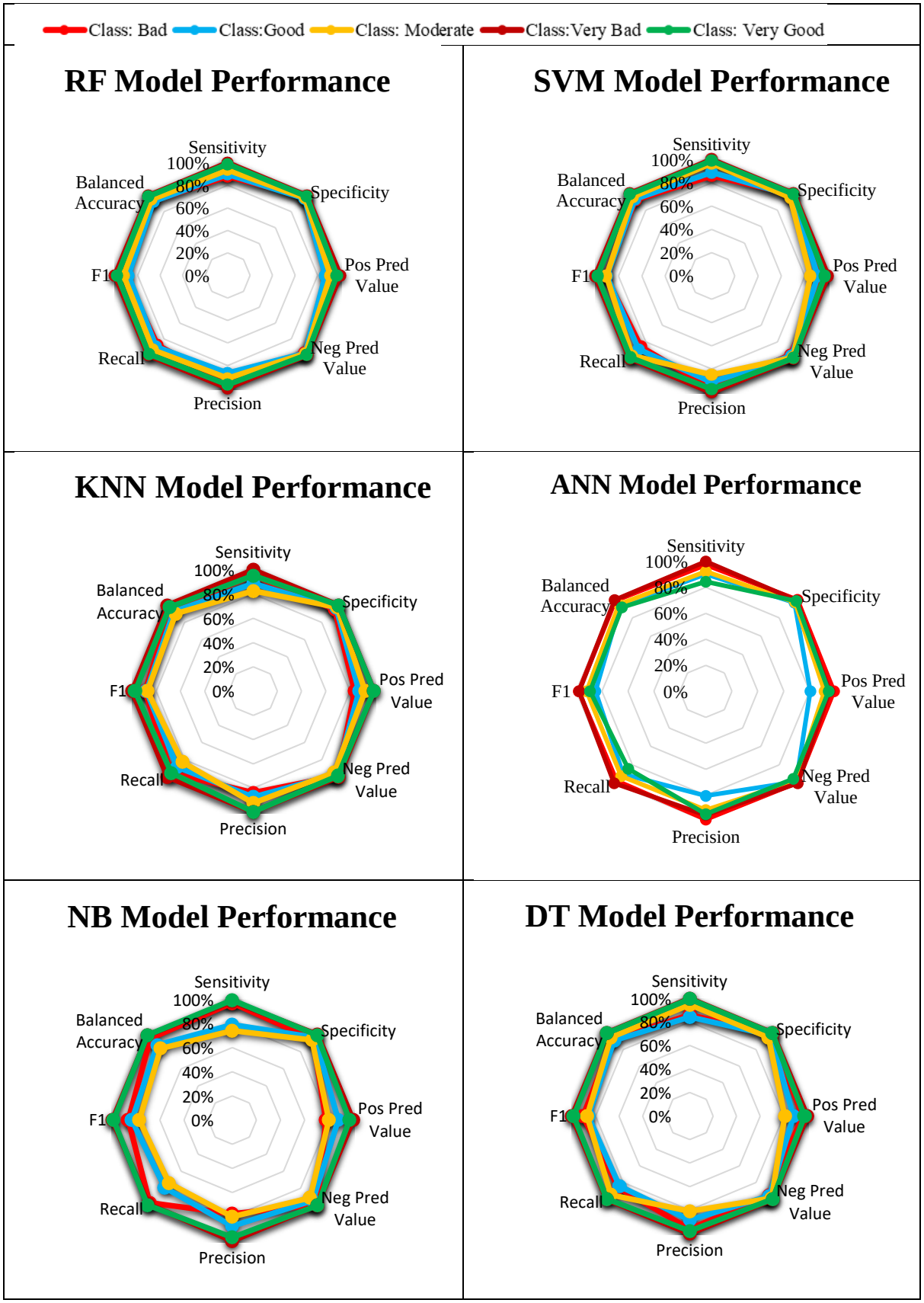


Figure 7.20: Statistics of ML algorithms by class (details are in Appendix H Part 3).

7.7.5 Research Result

The PTHI model development for under 69kV transformers as the PTHI will drive the decision-making process. The result from deploying the PTHI model was shown earlier in Figure 7.11 as five transformers were in “Very Good” condition, 769 power transformers were in “Good” condition, 178 power transformers were in “Moderate” condition, and 17 were in “Bad” condition.

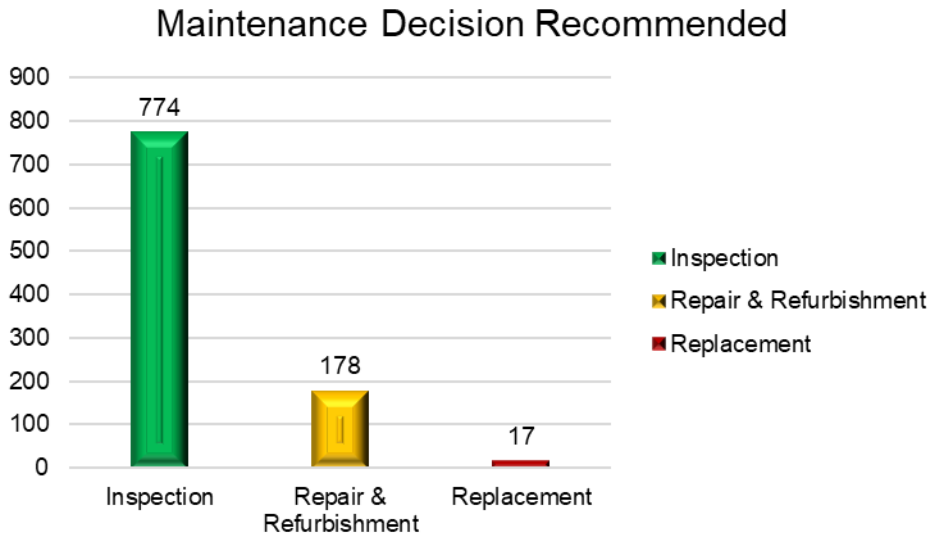


Figure 7.21: PTHI model decision based on the transformer health conditions.

Figure 7.21 describes the findings regarding the transformer condition and the considerable decision. Seven hundred seventy-four power transformers were in good and “Very Good” condition and were just due for a routine inspection. One hundred seventy-eight transformers are in Moderate condition and due for Repair and Refurbishment. In contrast, seventeen transformers are in bad condition and are due for replacement.

7.8 Summary

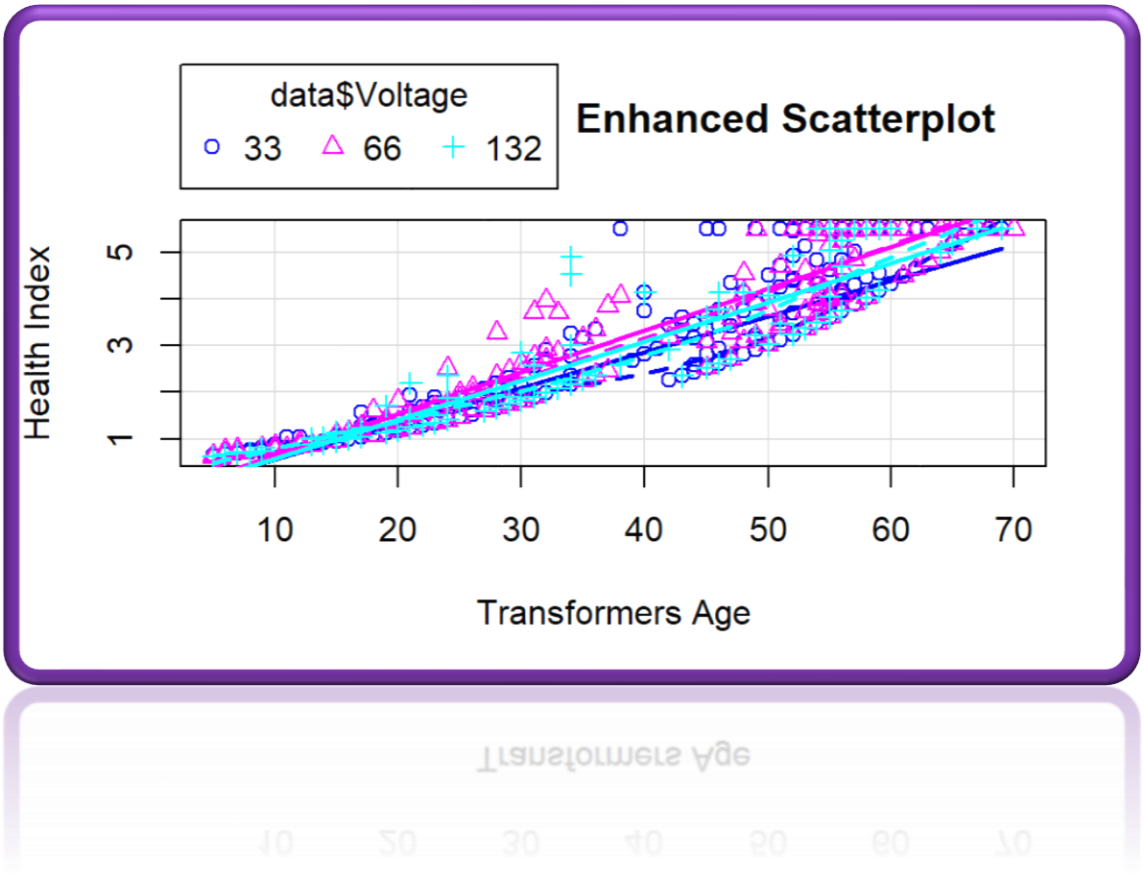
This chapter presented a case study investigation of the collected data from the electrical utility. It included a detailed analysis of the files received as Explanatory Data Analysis (EDA) to explore these datasets using Python programming.

Moreover, in this chapter, six machine models are the critical investigation for using the electrical utility data for transformers under 69kV. The RF and SVM models performed the best, with an accuracy of 93.7% and 93.5%, respectively, as shown in Figure 7.17. In contrast, the NB performed the least with 87% accuracy, as shown in Figure 7.17. Furthermore, since SVM was one of the best performers and therefore it was utilised further in developing an Asset Management decision-making tool that uses this SVM model to classify transformer conditions and recommend a corresponding action that can help in automating the decision-making process, especially adding a positive impact for the electrical utility large fleet of power transformers for the health condition prediction.

The next chapter will carry out a detailed investigation of the failure mode analysis due to PT ageing, which was discussed briefly earlier in this chapter in section 7.2.3. Figure 7.4 shows that the PTHI are more correlated to the ageing parameter.

Chapter 8

Power Transformer Ageing Correlation with Health Index



Chapter 8: Case Study 5 Ageing Correlation with Health Index and Research Technical and Financial Impacts

8.1 Introduction

The results in the previous chapter on PT ageing being more correlated with PTHI are interesting for further investigation. This chapter investigates the correlation between ageing and HI.

Hence, this chapter includes a study of the ageing of the large fleet of electrical PT using the actual electrical utility data. It includes a practical regression model development to calculate the adjusted age or apparent age to support PT replacement and investment decisions through the HI failure rate analysis.

To study how the PTHI can affect the ageing of PT, the following process, as shown in Figure 8.1, is followed to study how HI can be used to predict the failure probability utilising regression analysis:

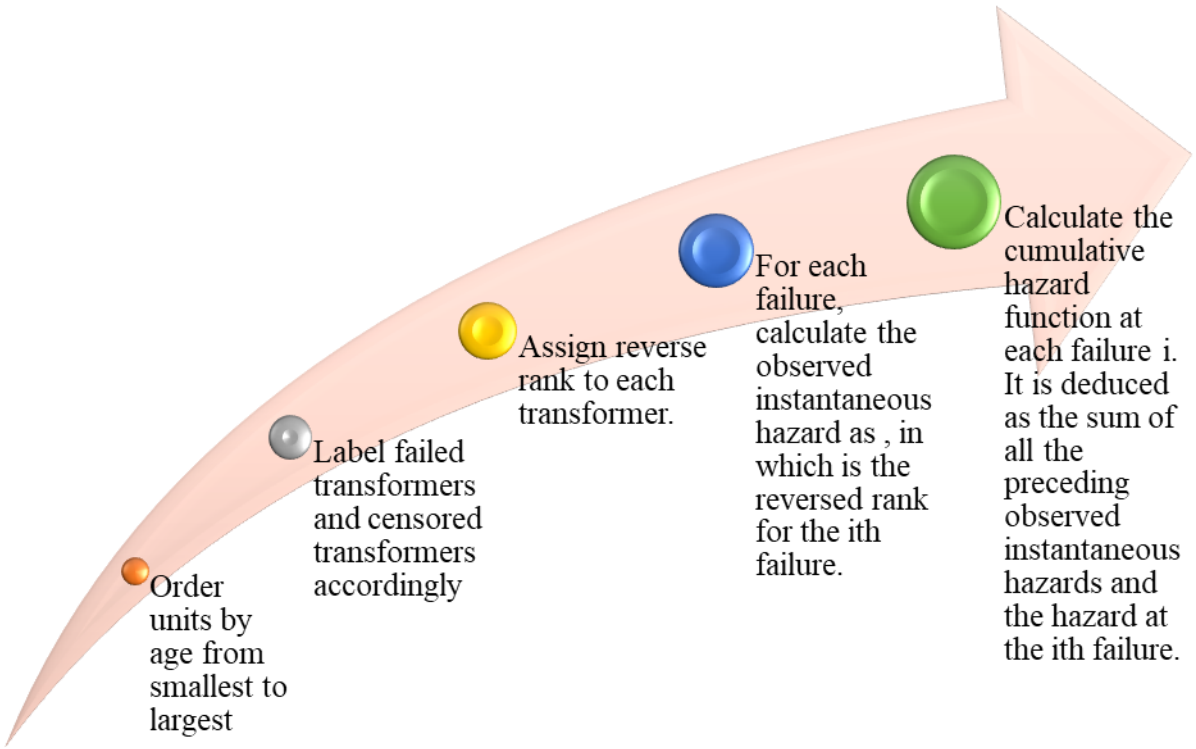


Figure 8.1: Studying transformers ageing correlation with HI.

8.2 Hazard Plotting Approach for Failure Rate Evaluation

In addition, it is well known that the “ageing system's failure rate is best modelled by Weibull distribution”, as stated by (Fengjun Yi et al., 2015). The actual electrical dataset contains 1231 transformers closely monitored over 14 years. Generally, the statistical analysis procedure for PT lifetime includes gathering data, choosing the best model of transformer population, selecting the models’ parameters, and placing the data into the model.

8.3 Case Study

This section includes a case study incorporating failure analysis and developing a regression model that aids in defining the apparent age from the HI value. A practical demonstration of how apparent age can is vital to define the remaining life by asset condition HI and helps prioritise decision-making for electrical PT.

8.3.1 Failures Rate (Ageing) Analysis

This section presents the failure data of PT with operating voltages of 132kV, 66kV, and 33kV. The failure analysis was counted from 2008 to 2022 for 14 years. Data populations are received at the end of 2022, as Table 8.1 shows the number of transformers per system voltage. The investigation contains 34 failures that occurred in the electrical utility network.

Table 8.1: PTs population based on system voltage.

Voltage Level	132 kV	66 kV	33 kV
Number of Transformer	217	322	692

As the apparent age shows the actual condition of the transformer, it can help to calculate the adjusted failure rate according to the failure rate formula. This formula is available within the research and the CIGRE study group (CIGRE Technical Brochure 642 Transformer Reliability Survey (2015) and CIGRE Working Group 05 (1983). The failure rate was calculated based on Bossi’s (1983) equation:

$$\lambda = \frac{\sum_1^i n_i}{\sum_1^i N_i} . 100\% \qquad (8.1)$$

Where:

n_i = the number of transformers that failed in the i^{th} year.

N_i= the number of transformers in service during the i^{th} year.

Based on the above equation (8.1), Table 8.2 shows the number of failures by voltage. Figure 8.2 indicates that most faults occur on 33kV PT. It also shows the failure rate calculated in this analysis by dividing the number of PT failures by transformer years.

Table 8.2: Failure analysis based on (CIGRE Technical Brochure 642, 2015).

	Failure Analysis by Voltage			Failure Analysis by Location/Causes		
	132kV	66kV	33kV	TC	TX	TX(CER/CMR)
No. of failures	5	3	26	18	6	10
Total of Failure	34			34		
Transformer years*	7939	12850	28424	Total of Transformers		1231
No. of Transformers	217	322	692			
Failure rate (%)	0.06%	0.02%	0.09%	52.94%	17.65%	29.41%

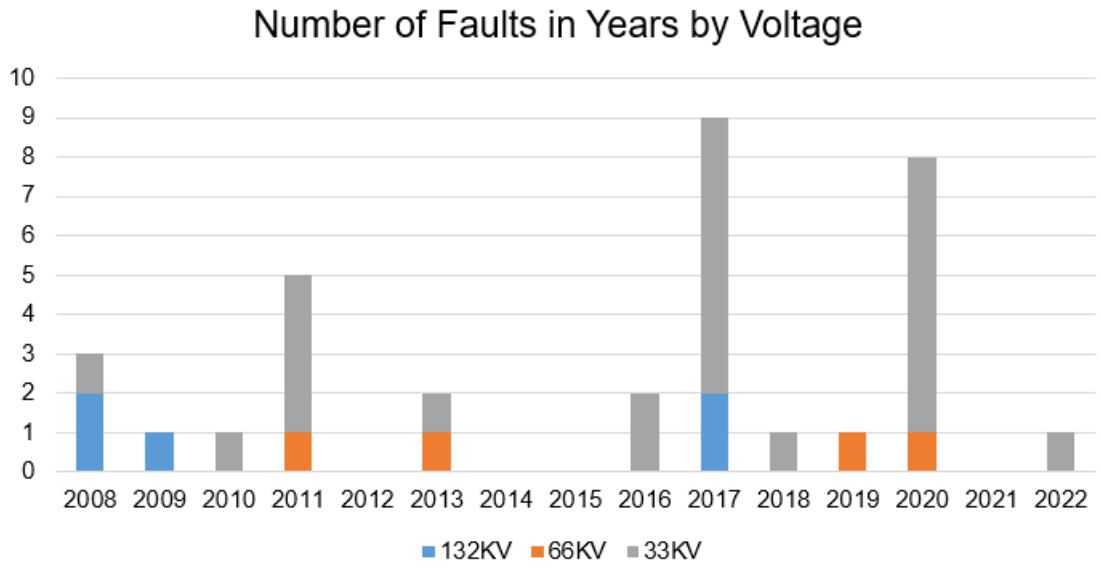


Figure 8.2: Number of faults in years by voltage (2008-2022).

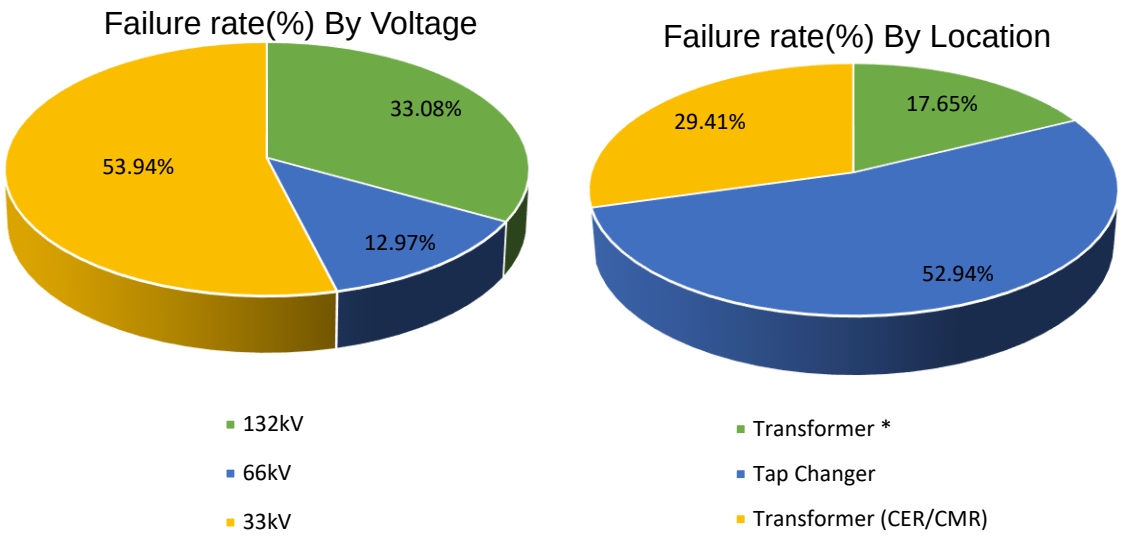


Figure 8.3: Failure rate analysis by voltage(left) by location(right).

*These transformers are 132kV and 33kV, tripped on Bucholz of an electrical fault, transient fault, fire, cable burns out of thermal fault, exterior flashover of an electrical fault, and stuck contact of a mechanical fault.

The failure analysis results show that 53.94%, 12.97%, and 33.08% of failures occurred at 33kV, 66kV, and 132kV voltages, respectively. Also, 52.94% of failures are due to tap changers (33kV, 66kV,132kV) transformers, and CER/CMR transformer faults are in Figure 8.3.

8.3.2 Understanding The Transformers Age Population

Figure 8.4. presents the distribution of the transformers' age population in the electrical utility. This histogram allows the understanding of the transformer age population to understand the age demographics within the utility infrastructure. The x-axis represents the age group in five-year intervals, while the y-axis shows the number of transformers within each group.

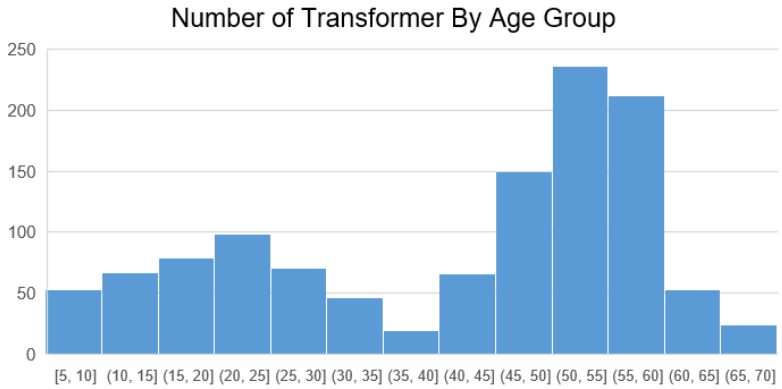


Figure 8.4: Number of transformers classified by age group.

8.3.3 The Calculation of Apparent Age for Risk Assessment

This chapter used the apparent age method using the electrical utility's data. The “apparent” age is the modified age that relates to all transformer conditions and a population using the regression line. PT's “apparent” age is obvious for maintenance decisions, including asset replacement and repairs. It is not viable to precisely relate the HI score to the probability of failure.

8.3.4 Calculating Adjusted Failure Rate Based on Apparent Age Utilising a Regression Model.

The apparent age is used as a measurement system to assess the health state of a single transformer unit among the large transformer fleet. Given the many different types of transformer health conditions, this system uses this measurement approach to surpass the standard transformer age. This analysis covers approximately all the consequences that a transformer may withstand.

1. Firstly, the regression model was used by conducting several experiments to find the best model that fits the age failure data.
2. Secondly, the regression line for the best-fitted model determines the apparent age.
3. The adjusted failure rate is calculated based on the transformer’s “apparent” age.
4. Finally, the adjusted failure rate is the accurate action that asset managers can use within the electrical utility to ensure PT system reliability and operation at the required AM quality of service.

Figure 8.5 demonstrates how the regression line originates the apparent age from the average of the whole population. This action is done by drawing a horizontal line from the observation point to the regression line and then a vertical line from the abscissa point to the x coordinates. The transformer's apparent age revelation is by utilising the regression line. This option is chosen based on earlier findings that HI is more correlated with age with a high positive correlation value of 0.91, as shown in correlation matrix analysis in chapter 7, section 7.2.3 Figure.

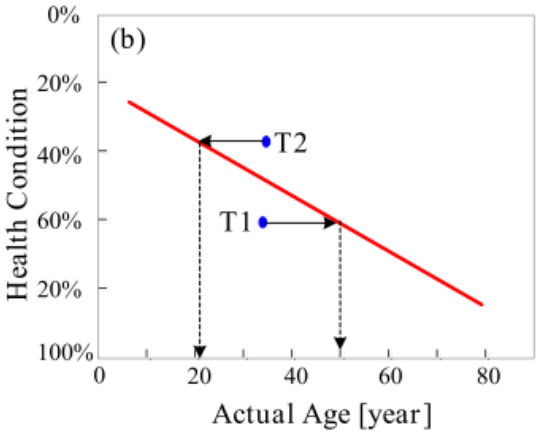


Figure 8.5: The PTHI calculation for apparent age T1 and T2 (Li et al., 2020).

The reliability of power transformers is a significant concern to all electrical utilities. This concern is because the cost of the outage and its restoration are very high, and it concisely contributes to higher system operating costs and increases user bills. According to CIGRE Technical Brochures 642 (2015), failure is defined very broadly in IEC 60050-191(1990) and IEEE C57.117 (1986). In contrast, the failure definition in Bossi's (1983) is more focused as it includes those situations that require the transformer to be out of service for replacement or repair (CIGRE Technical Brochure 227, 2003).

8.3.5 Simple Linear Regression

This section shows a sample of seven transformers using apparent age to calculate the adjusted failure rate. The apparent age calculation uses the technique illustrated in the above section, employing the best-fitted regression line. Table 8.3 shows the

apparent age calculated for the sample and its expected PT health condition according to its HI value.

Table 8.3: Case study samples data.

Sample	Health Index	Condition	Actual Age
TX1	3.9	Good	57
TX2	2.7	Good	47
TX3	3.6	Good	45
TX4	5.5	Fair	57
TX5	1.9	Good	25
TX6	3.4	Good	53
TX7	1.9	Good	29

The case study results above confirm the significance of studying the transformer's apparent age, which indicates its actual health condition. Also, the actual age does not always indicate the real transformer condition, as some young transformers might be in “Bad” condition, whereas some old transformers might be in “Very Good” condition. The result of utilising the regression model and the hazard function to study the apparent age and probability of failure is in Table 8.4.

Table 8.4: The case study results for the linear first-order regression model.

Sample	Actual Age	PoF Based on Actual Age	Apparent Age	PoF Based on Apparent Age
TX1	57	Medium	51	Medium
TX2	47	Low	37	Low
TX3	45	Low	47	Low
TX4	57	Medium	70	High
TX5	25	Very Low	24	Very Low
TX6	53	Medium	49	Low
TX7	29	Very Low	26	Very Low

In Figure 8.6, a case study sample shows that TX1 is better than TX4 as it is below the average line. The arrow line originated from TX1, or TX4, and has an intersection point with the regression line. The value in abscissa is 51 and 70, which means that even though TX1 and TX4 have an actual age of 57 years when studying the transformer condition, TX1's apparent age is 51 years old, whereas TX4's apparent age is 70 years old. Also, TX4 and other transformers above the line are considering transformers with accelerated ageing since their apparent age is more than their actual age.

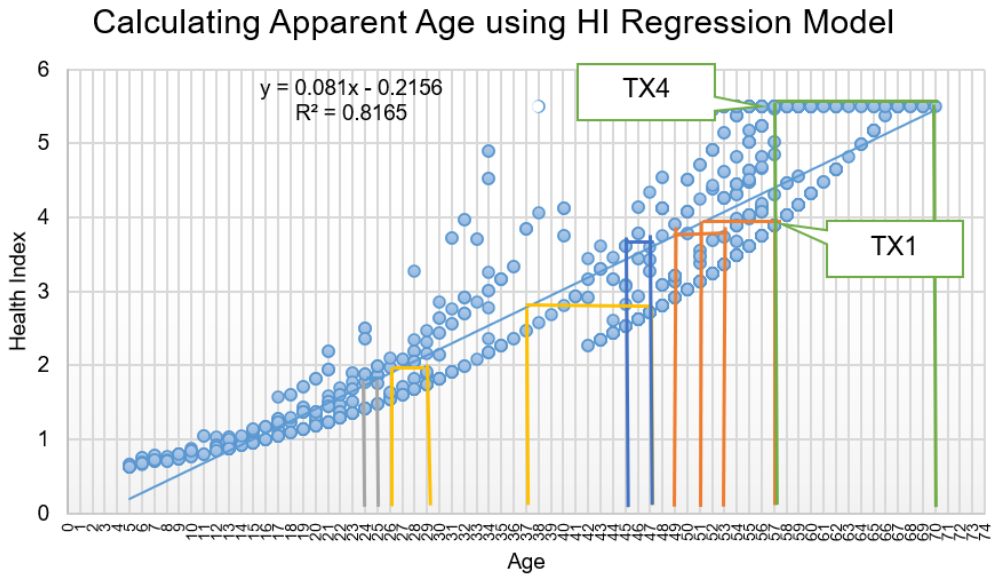


Figure 8.6: HI first order simple linear regression model.

Going through all records and focusing on those transformers whose HI score is equal to 5.5, a total of 160 transformer records, which are in fair condition, were found. They are all above the regression line, as shown in Figure 8.7. All of those have an apparent age of 70, as measured using the regression line, which requires attention and action for repair and refurbishment.

Studying the population of transformer records in depth, Figure 8.7 shows a result of studying the transformer identified in fair condition by examining the apparent age as a sign of accelerated ageing. The ID of the transformer is unavailable for confidentiality, but the orange bars indicate the transformer's actual age compared to its apparent age (blue bars) due to PT's health condition. The result is a list of those transformers affected by the accelerated ageing issue that need further focus and care. The PT Age study experiments are included in Appendix H Part 5.

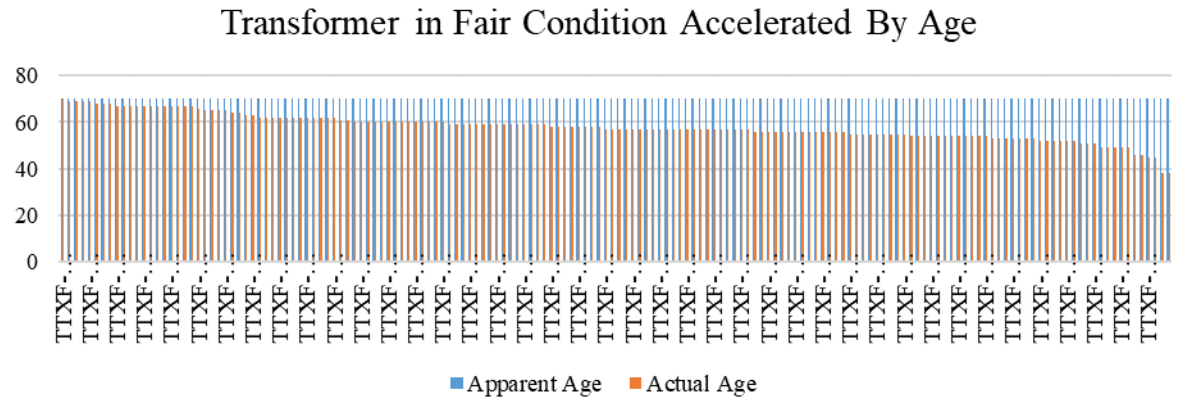


Figure 8.7: Transformers in fair condition.

8.3.6 Second-order polynomial (Quadratic) Regression

This experiment is conducted based on the knowledge obtained from the book (Faraway, J.J., 2002). This section shows the same sample of seven transformers considered in the previous section to show the use of apparent age in calculating the adjusted failure rate. The apparent age calculation uses the mentioned technique, employing the second-order polynomial regression model (Maulud & Abdulazeez, 2020; Jie & Zheng, 2019; Niu et al., 2018). Table 8.5 shows the results of the second-order polynomial regression model, the probability of failure, the actual age, and the apparent age.

Table 8.5: Case study results for second-order polynomial (quadratic) model.

Sample	Health Index	Actual Age	PoF Based on Actual Age	Apparent Age	PoF Based on Apparent Age
TX1	3.9	57	Medium	52	Medium
TX2	2.7	47	Low	42	Low
TX3	3.6	45	Low	50	Low
TX4	5.5	57	Medium	67	High
TX5	1.9	25	Very Low	20	Very Low
TX6	3.4	53	Medium	50	Medium
TX7	1.9	29	Very Low	27	Very Low

In Figure 8.8, a case study sample shows that TX1 is better than TX4 as the position is below the average line. The arrow line originated from TX1, or TX4, and has an intersection point with the regression line. The value in abscissa is 51 and 70, which means that even though TX1 and TX4 have an actual age of 57 years when studying the transformer condition, TX1's apparent age is 52 years old, whereas TX4's apparent

age is 67. Also, TX4 and other transformers above the line are transformers with accelerated ageing since the apparent age is more than the actual age.

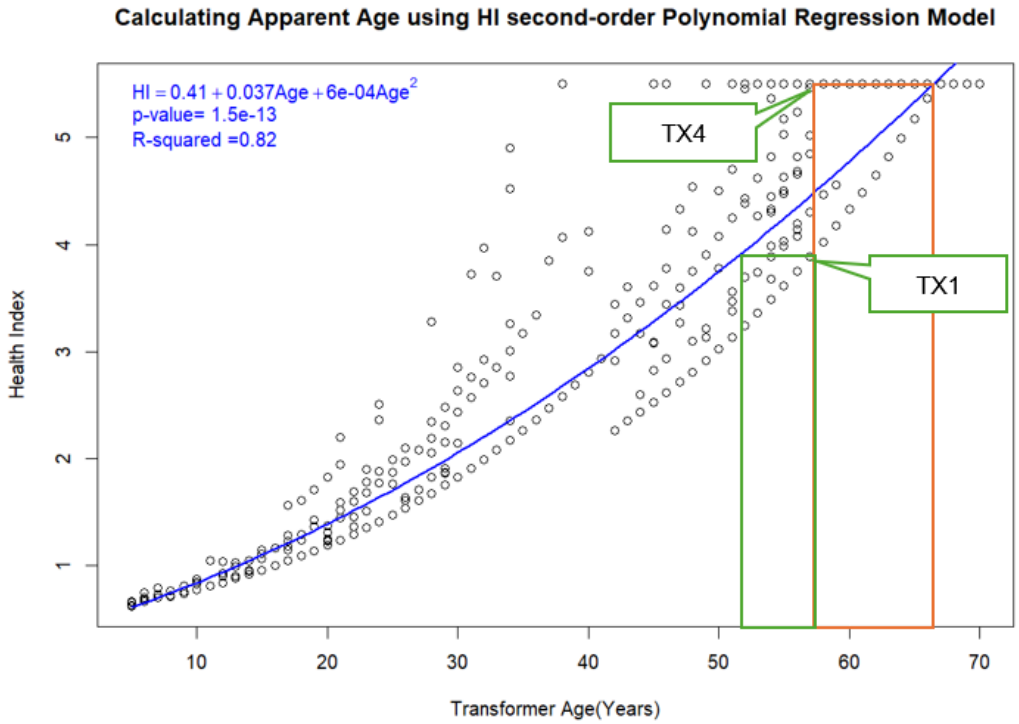


Figure 8.8: HI second-order polynomial regression model curve.

8.3.7 Comparison Between the Performance of a Linear First-Order and Second-Order Polynomial (Quadratic) Regression Model

Fitting a more complicated quadratic mode, Figure 8.9 shows the analysis of variance performed using a breakdown of the quadratic model with type 1 sum of squares by default. The p-value of 1.52e-13 explains significant variability in the HI above and beyond what age explains, which justifies the extra term of the age square in the model. Figure 8.10 shows the effect of taking a slight curve when using the polynomial second-order relationship compared to the earlier simple linear first-order, which was a straight line.

```
> anova(model_linear)
Analysis of Variance Table

Response: HI
            Df  Sum Sq Mean Sq F value    Pr(>F)
Age          1 2258.70 2258.70  5224.8 < 2.2e-16 ***
Residuals 1174  507.53    0.43
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(model_quadratic)
Analysis of Variance Table

Response: HI
            Df  Sum Sq Mean Sq  F value    Pr(>F)
Age          1 2258.70 2258.70 5468.886 < 2.2e-16 ***
AgeSQ        1   23.07   23.07   55.856  1.52e-13 ***
Residuals 1173  484.46    0.41
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 8.9: Linear and quadratic regression model analysis.

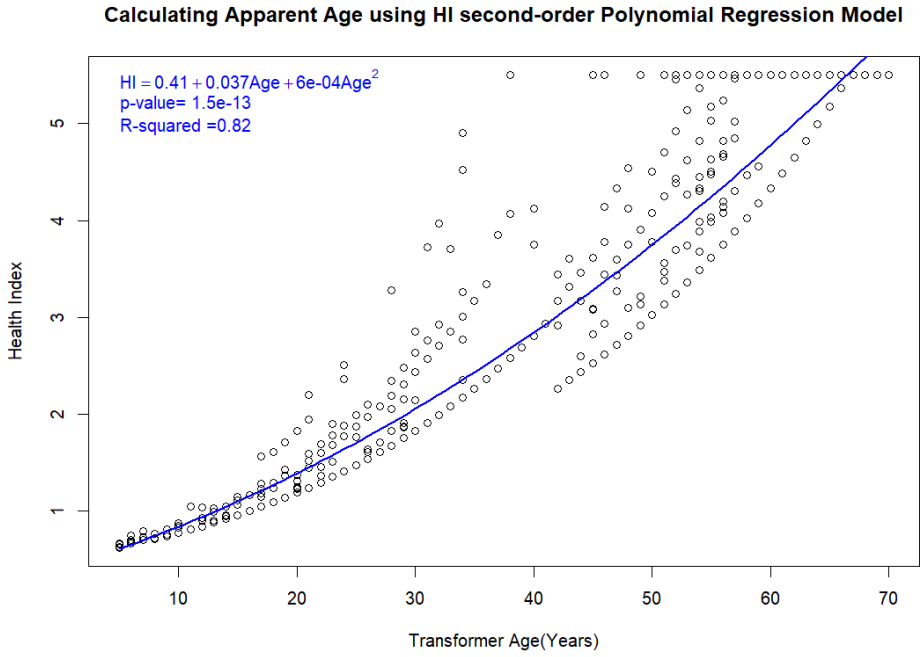


Figure 8.10: Second-order polynomial regression model curve.

8.3.8 Fitting a More Complex Cubic Regression Model

Furthermore, fitting a more complicated cubic model is a vital experiment. Figure 8.11 shows analyses of the variance breakdown of the cubic model. The age cub does not significantly improve the regression model, as indicated by the p-value of 0.1156, above the typical significance standard level value of 0.05. Therefore, the second polynomial (quadratic) model fits better than the cubic and the simple linear first-order models in capturing the relationship between the age and health index variables.

```
> anova(model_cubic)
Analysis of Variance Table

Response: HI
            Df  Sum Sq Mean Sq   F value    Pr(>F)
Age          1 2258.70 2258.70 5475.7839 < 2.2e-16 ***
AgeSQ        1   23.07   23.07   55.9269  1.47e-13 ***
AgeCUB       1    1.02    1.02    2.4794    0.1156
Residuals 1172  483.44    0.41
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 8.11: Cubic regression model analysis.

8.3.9 Regression Model’s Summary

Figure 8.12 presents a summary of the first-order, simple linear regression.

```
> summary(model_linear)

Call:
lm(formula = HI ~ Age)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9211 -0.5161 -0.1198  0.3614  2.6380

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.21562    0.05061  -4.261  2.2e-05 ***
Age          0.08099    0.00112  72.282  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6575 on 1174 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8165,    Adjusted R-squared:  0.8164
F-statistic:  5225 on 1 and 1174 DF,  p-value: < 2.2e-16
```

Figure 8.12: Linear regression model summary.

Figure 8.13 presents a summary of the second-order polynomial (quadratic) model.

```
> summary(model_quadratic)

Call:
lm(formula = HI ~ Age + AgeSQ)

Residuals:
    Min      1Q  Median      3Q     Max
-0.7554 -0.5330 -0.1455  0.3016  2.8277

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.084e-01  9.705e-02   4.208 2.77e-05 ***
Age         3.677e-02  6.018e-03   6.110 1.35e-09 ***
AgeSQ       6.002e-04  8.031e-05   7.474 1.52e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6427 on 1173 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8249,    Adjusted R-squared:  0.8246
F-statistic:  2762 on 2 and 1173 DF,  p-value: < 2.2e-16
```

Figure 8.13: Quadratic regression model summary.

Figure 8.14 presents a summary of the third-order polynomial (cubic) model.

```
> summary(model_cubic)

Call:
lm(formula = HI ~ Age + AgeSQ + AgeCUB)

Residuals:
    Min      1Q  Median      3Q     Max
-0.7392 -0.5312 -0.1646  0.3142  2.8252

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.051e-01  1.615e-01   1.270 0.204325
Age          6.049e-02  1.622e-02   3.729 0.000202 ***
AgeSQ       -1.337e-04  4.729e-04  -0.283 0.777520
AgeCUB       6.633e-06  4.213e-06   1.575 0.115615
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6423 on 1172 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8252,    Adjusted R-squared:  0.8248
F-statistic:  1845 on 3 and 1172 DF,  p-value: < 2.2e-16
```

Figure 8.14: Cubic regression model summary.

8.3.10 Evaluation of the Three Regression Models Performance

Table 8.6 presents an overall comparison between the three models. The second-order polynomial model fits better than the linear model, as shown in the adjusted R squared value of 82% compared to the earlier 81%. However, the third-order polynomial model does not provide a statistically significant improvement over the quadratic model, even though the R squared value is improved (82%) compared to the linear. However, the p-value is 0.1156, which shows no statistical significance as it is above the standard significance value of 0.05. Therefore, based on the adjusted R-squared and p-value, the second-order polynomial appears most appropriate for capturing the relationship in the dataset, which indicates the relationship by a slight curve rather than a straight line (Faraway, 2002).

Table 8.6: Comparison between the regression models' performance.

Regression Model	Adjusted R-squared	p-value
First-order polynomial (Linear)	0.8164	<2.2e-16
Second-order polynomial (Quadratic)	0.8246	1.47e-13
Third-order polynomial (Cubic)	0.8248	0.1156

8.3.11 Calculating Remaining Life by Condition Utilising Health Index Model

The remaining life is defined by expert judgment (Jahromi et al., 2009) using the approximate expected lifetime, according to the calculated PTHI and transformer condition. This section uses the electrical utility data for all voltage levels to build a remaining life model that presents a risk assessment and helps prioritise decision-making. After calculating the PTHI and specifying the transformer's condition using the conventional model in Chapter 4, the remaining life calculation uses the "vlookup" function in the spreadsheet. The result is utilising the pivot table to sum the transformer condition by its remaining life group and prioritising the related actions of the maintenance plan. A simple colour scheme will represent the transformer condition. The first step will be dealing with the transformer identified as Very Bad, which has a remaining life of 0-2 years. Then, the second step is the analysis of the 18 transformers in "Bad" conditions with a survival expectation for 2-3 years.

Remaining Life	Very Good	Good	Moderate	Bad	Very Bad
0-2 years	0	0	0	0	1
2-3 years	0	0	0	18	0
3-7 years	0	0	200	0	0
7-15 years	0	949	0	0	0
>15 years	9	0	0	0	0

Figure 8.15: Risk matrix based on PTHI condition assessment results.

According to Figure 8.15, the PTHI condition assessment matrix above, the PT health condition leads the future risk base model decision, and as a result, the decision matrix in Figure 8.16 to drive asset management maintenance strategy and prioritise the short-term annual maintenance plan for the transformers that needs urgent attention or replacement.

- The TX1 transformer in Very Bad condition (pink zone) with a remaining life of 0-2 years is given the highest priority and is due for replacement.
- Eighteen transformers in Bad condition (red zone) with an expected remaining life of 2-3 years are given the second highest priority and should go for refurbishment. This decision is because the PTs have more than one component in the same transformer/s within the current year.
- Two hundred transformers are in “Moderate” condition (orange zone) with an expected remaining life of 3-7 years and should go for repair, including repairing their single component in the same transformer/s within the next two years.
- Nine hundred forty-nine transformers are in “Good” condition (yellow zone) and are set for regular maintenance within the next five years, including regular inspections.
- The remaining nine transformers (green zone) need routine inspection as the expected remaining life is over 15 years.

Remaining Life	Inspect	Regular Maintenance	Repair	Refurbishment	Replacement
0-2 years	0	0	0	0	1
2-3 years	0	0	0	18	0
3-7 years	0	0	200	0	0
7-15 years	0	949	0	0	0
>15 years	9	0	0	0	0

Figure 8.16: PTHI risk matrix for prioritising annual maintenance plan decisions.

8.4 Research Impacts

This research analysis of the PTHI model development led to the following technical and financial impacts:

8.4.1 Technical Impact of PTHI Web App Tool

This section presents the technical impact that supports, improves, and speeds up site engineer decision-making in determining the actual power transformer's health condition. This impact is done by deploying the SVM model in the Shiny web application and testing the system. Shiny server is software used to install the PTHI web application and can be used to deploy multiple applications. R Studio Connect is a server-based solution that can deploy the Shiny application, Markdown documents, and other kinds of content and serve both the Shiny application and Markdown documents as ordinary websites.

This tool's primary function is to assist in knowing the conditions of transformers and recommend an action according to the main parameters entered through the form. After that, the application deployment on the website https://www.shinyapps.io/. The testing of the PTHI tool uses an Android smartphone to show the ability of the shiny web application to be portable for the onsite maintenance engineers. The system is accessible through a web browser on any smartphone/tablet. The PTHI assessment

tool allows the electrical utility user to enter the parameters to check the PT's health condition. The following section will test different scenarios to ensure the PTHI tool shows the correct condition for the transformer's health condition.

8.4.1.1 Testing Scenario 1: Transformer Parameters in Good or Normal Range

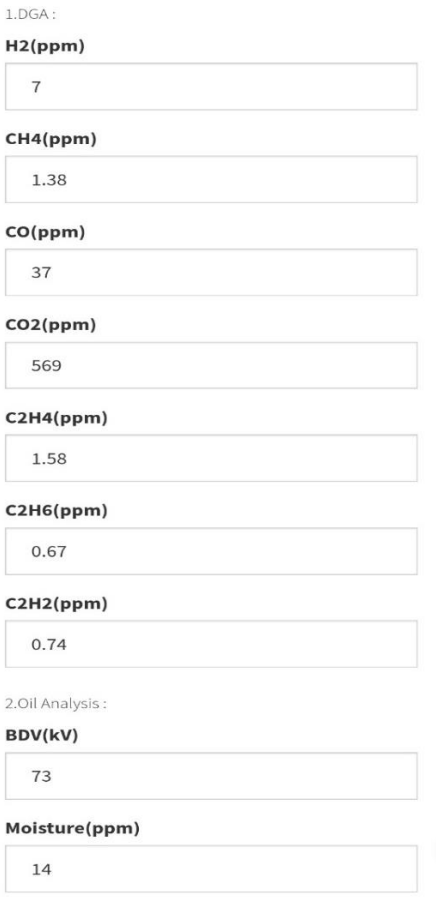


Figure 8.17: Transformer condition in “Good” or normal using DGA inputs.

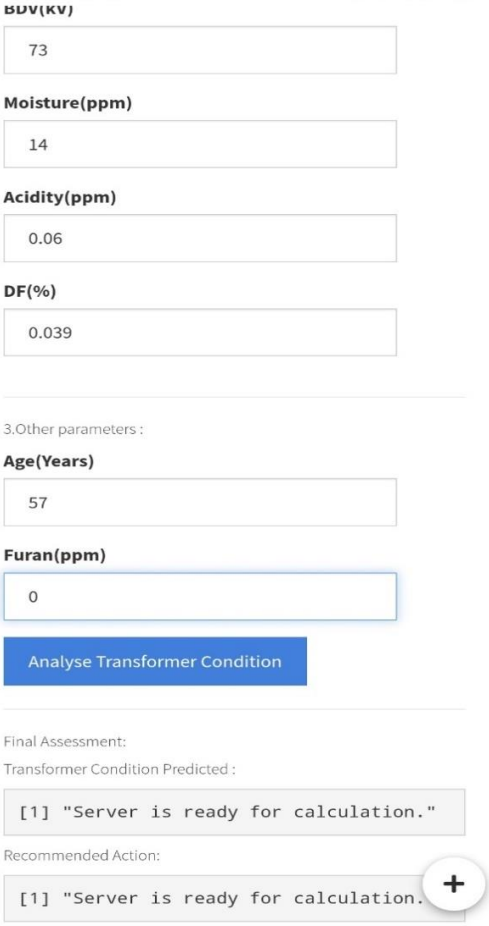


Figure 8.18: Transformer condition in “Good” or normal range using BDV inputs data.

8.4.1.1.1 Output:

The AM PTHI tool, as shown in Figure 8.19, successfully identified the condition as good.

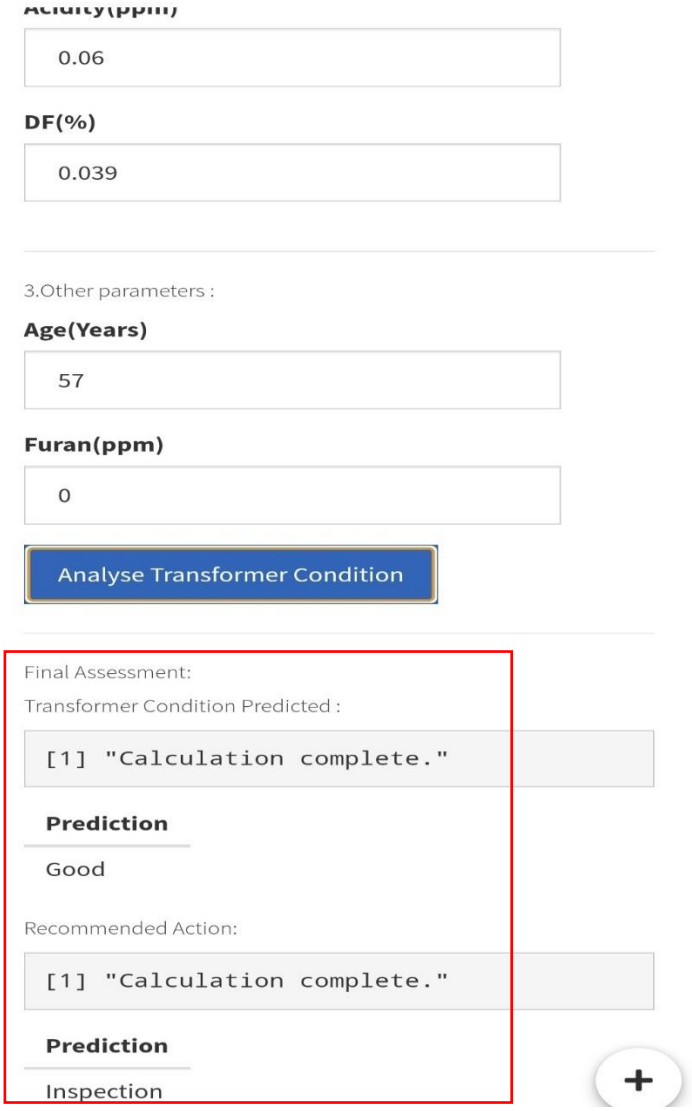


Figure 8.19: Transformer condition as Good.

8.4.1.2 Testing Scenario 2: Transformer parameters in Moderate range

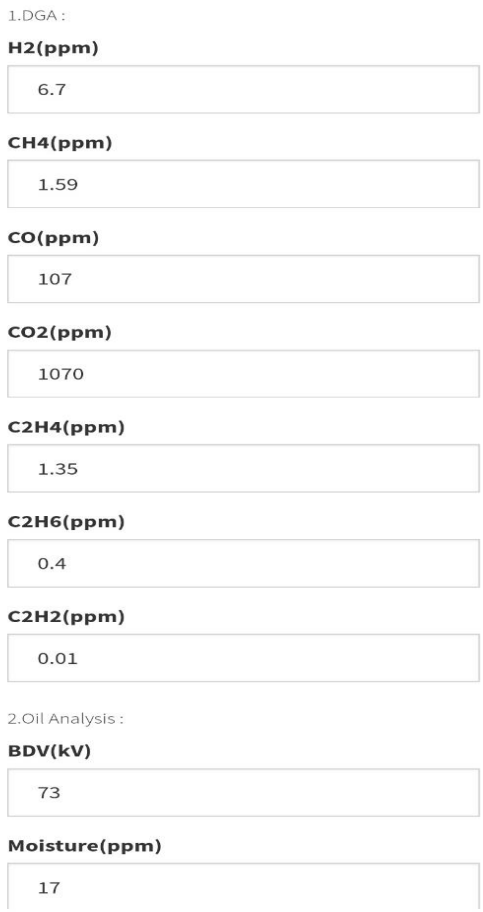


Figure 8.20: Transformer condition in moderate range using DGA inputs.

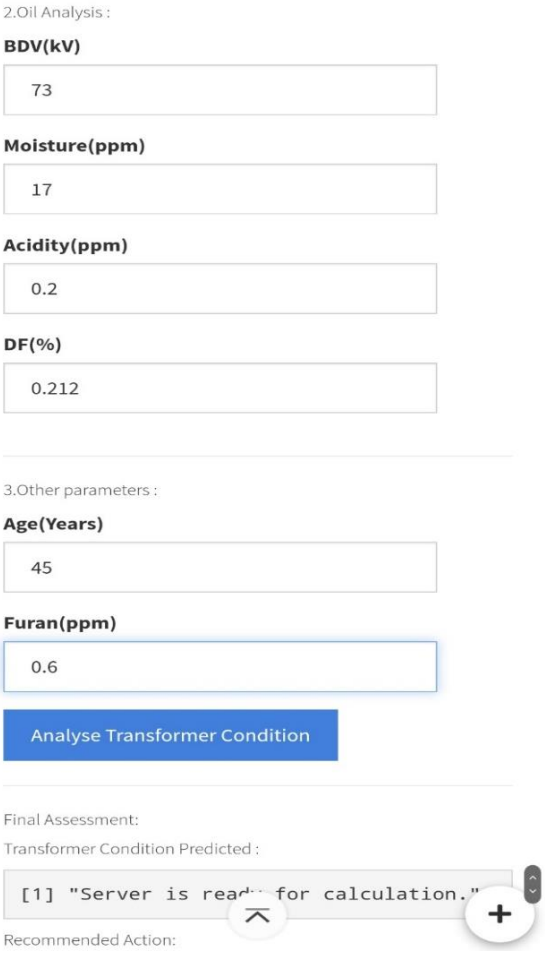


Figure 8.21: Transformer condition in moderated range using BDV inputs.

8.4.1.2.1 Output:

The AM PTHI tool, as shown in Figure 8.22, successfully identified the condition as moderate and recommended a repair.

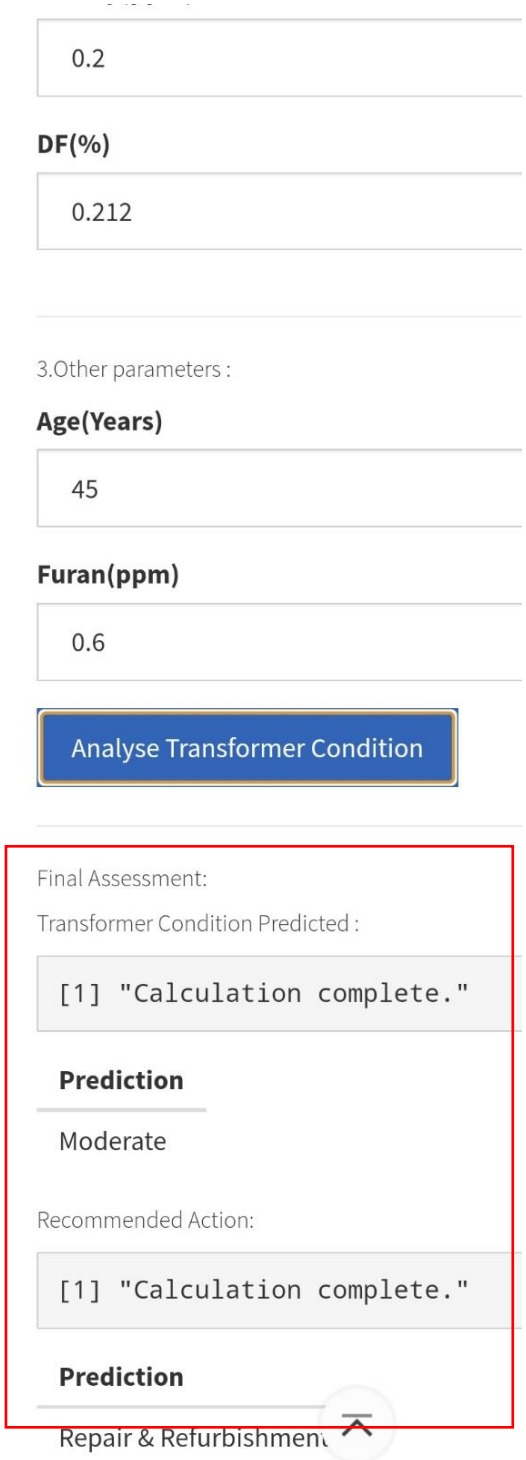


Figure 8.22: Transformer condition as moderate.

8.4.1.3 Testing Scenario 3: Transformer with parameters in Bad range

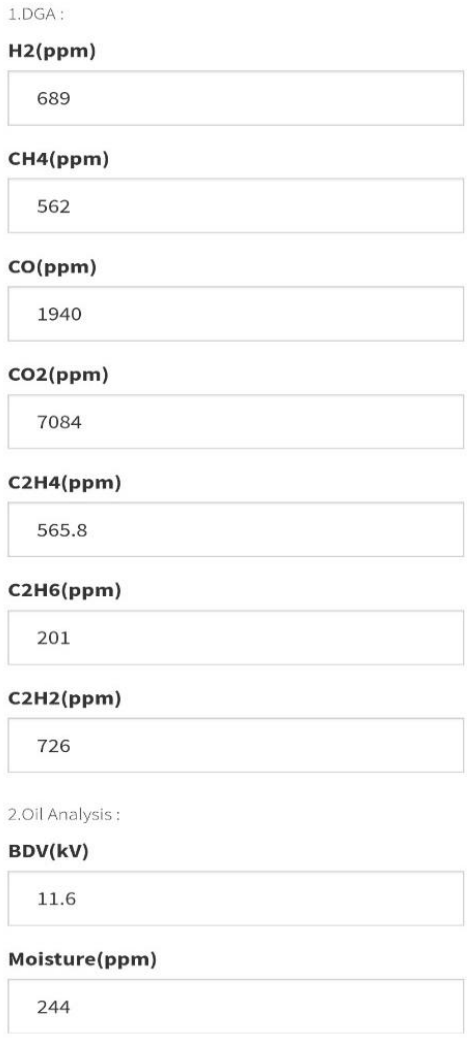


Figure 8.23: Transformer condition in the Bad range using DGA inputs.

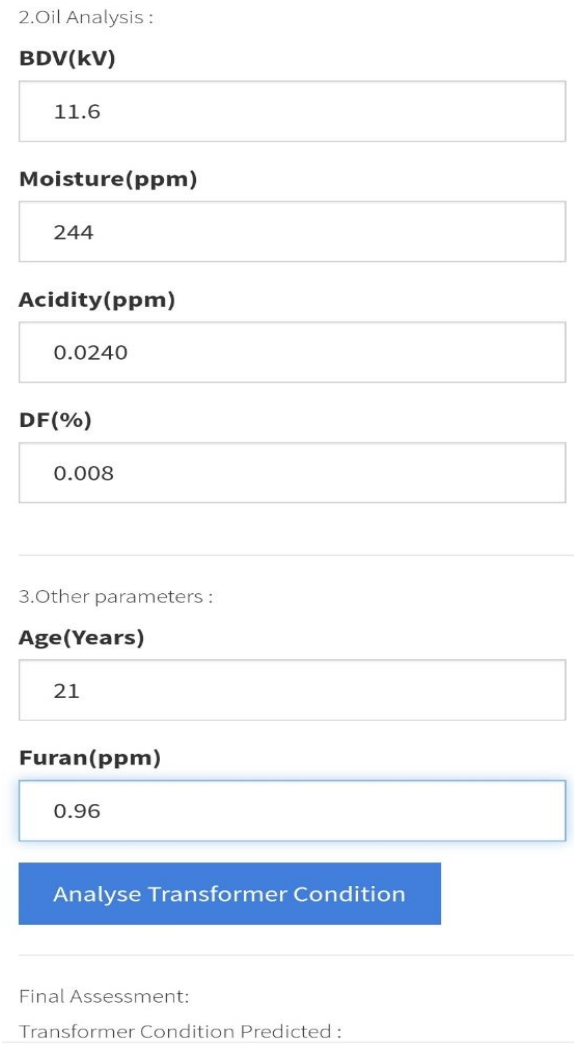


Figure 8.24: Transformer condition in the Bad range using BDV inputs.

8.4.1.3.1 Output:

The AM PTHI tool, shown in Figure 8.25, successfully identified the condition as bad and recommended replacing the transformer.

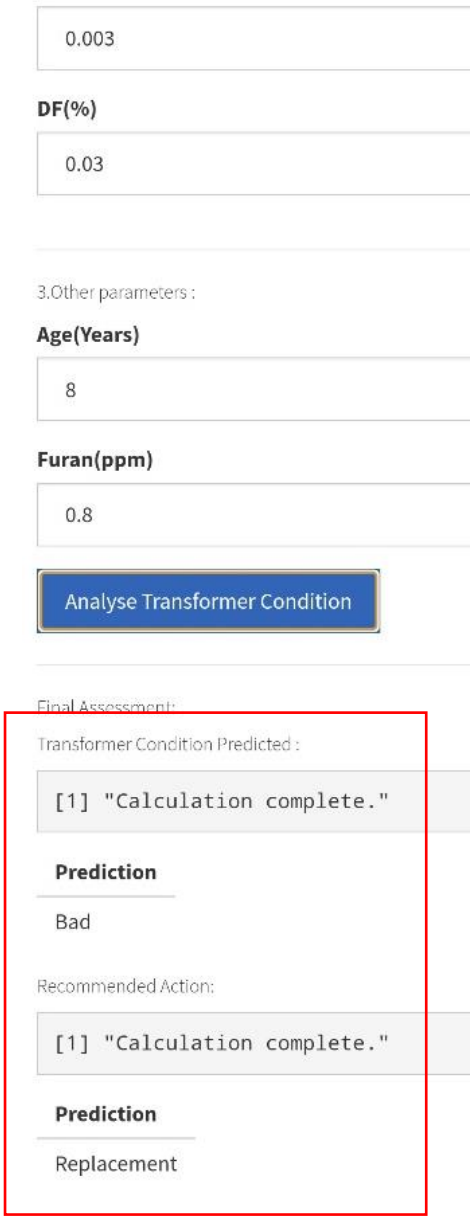


Figure 8.25: The transformer condition is Bad.

8.4.1.4 Exploratory Data Analysis (EDA)

The Exploratory Data Analysis (EDA) tab in the web application shows EDA to identify the main characteristics of the dataset used for modelling. Further figures are available in Appendix H, Part 4 Mobile and Website Application.

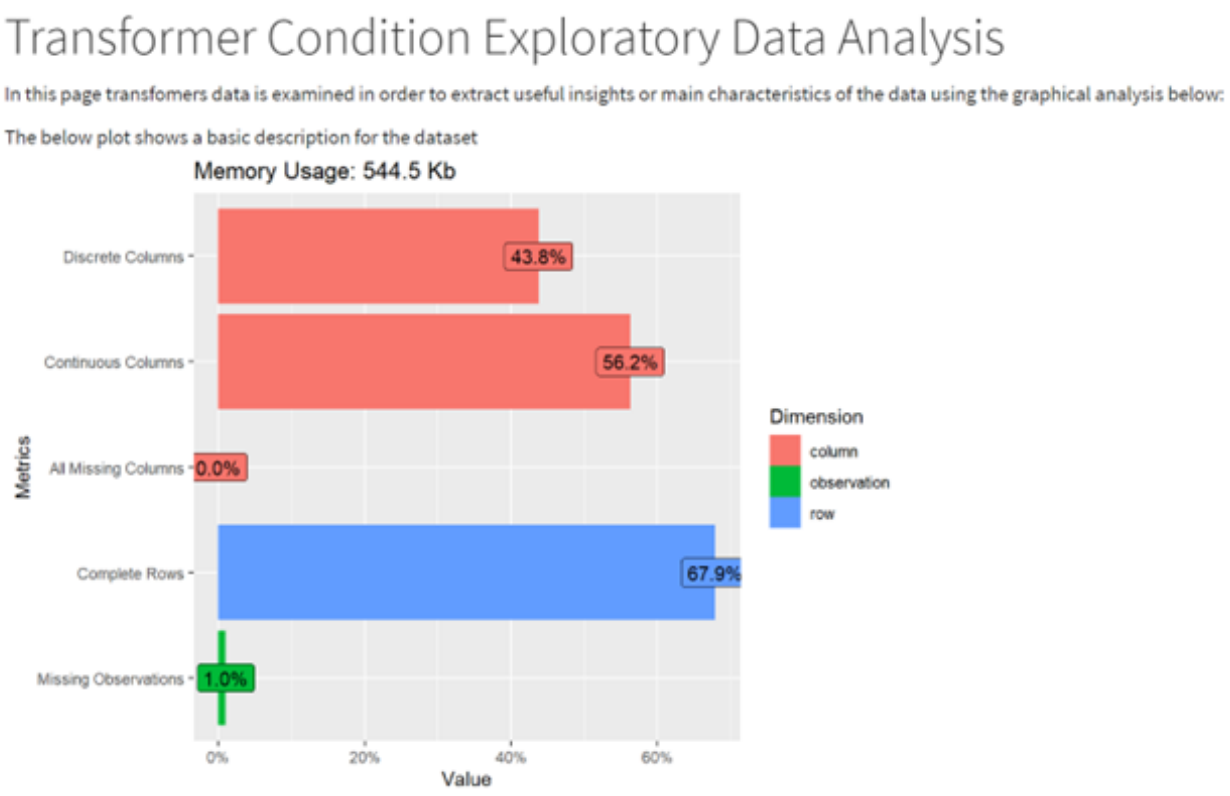


Figure 8.26: Dataset description for the PTHI web application tool.

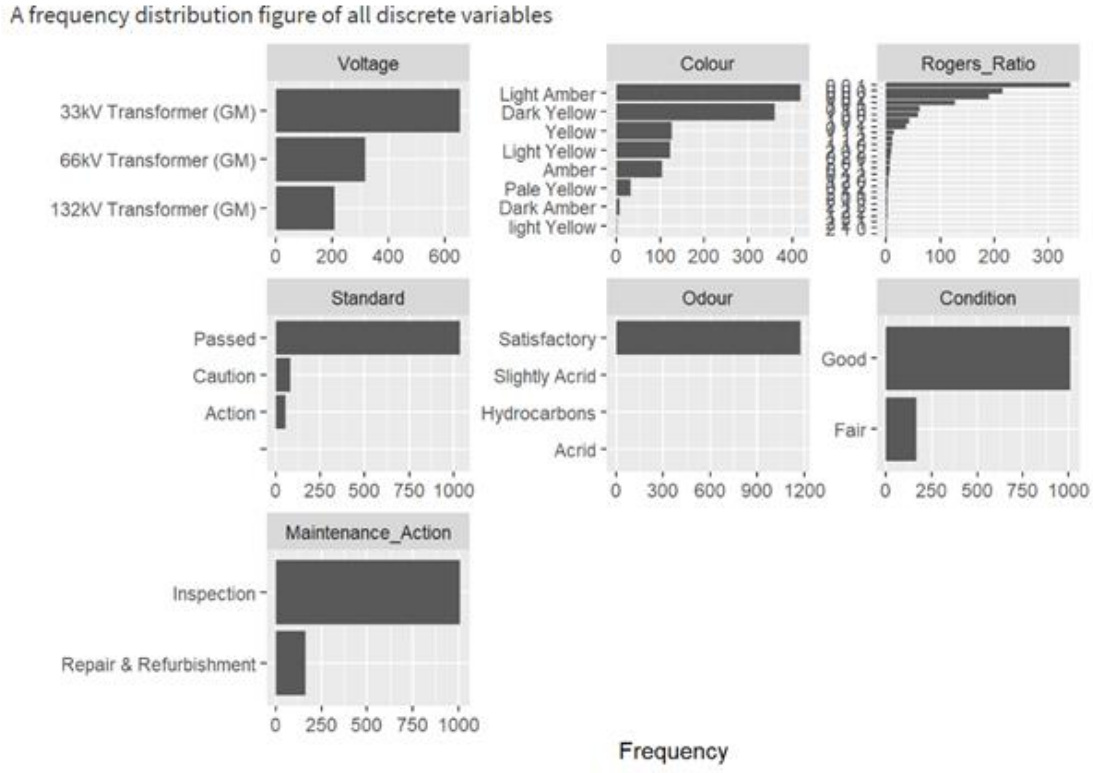


Figure 8.27: The technical impact of producing the PTHI web application tool.

8.4.1.5 Mobile and Web Application Overview Tab

This section presents the mobile or web application and shows some information about the mobile or web application and the contact email address.

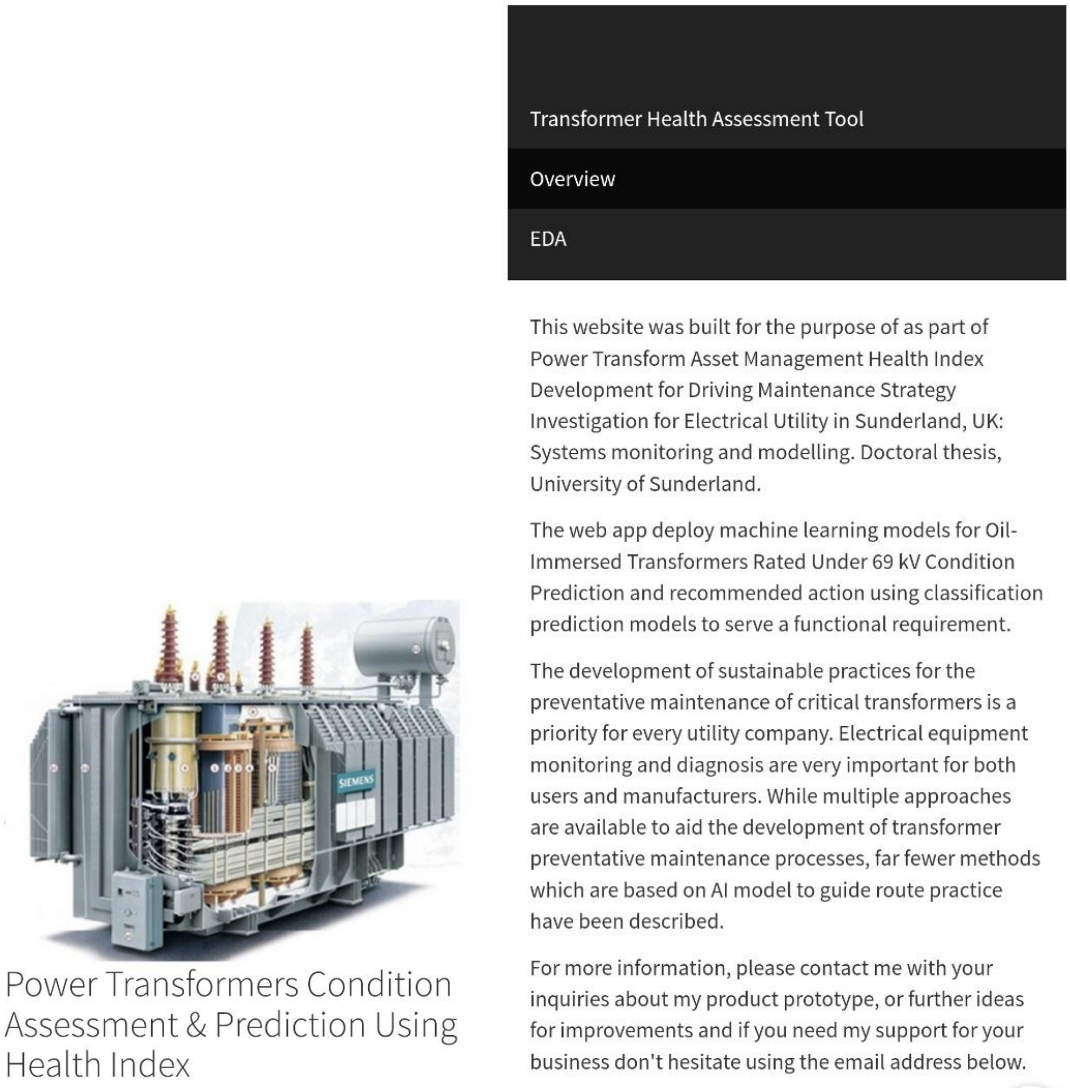


Figure 8.28: Overview of the PTHI web application tool.

8.4.2 Financial Impact of the Effect of PTHI Asset Management Predictive Maintenance Decisions

Alvarez Quiñones et al. (2022) conducted a study by developing a tool for maintenance scheduling for distribution transformers that relied on a classification predictive model to discover transformers that are susceptible to failure. The study demonstrated a 13% reduction in maintenance cost by applying the proposed model using actual data from the Cauca Department of Colombia.

In this research, the cost analysis is developed based on solving short-term annual maintenance plans of the overall asset maintenance strategy utilising the transformer cost values provided by the electrical utility. This analysis includes the strategic lifecycle asset planning, which supported using all the required AM decisions in the Total Investment Plan (TIP):

- **Option 1: Do Nothing (Unplanned Outage) for the PT unit.**

This option means waiting for the failure to occur, then starting corrective maintenance and accepting all the consequences. This option is costly for electrical utility if the failure is not prevented by preventive maintenance before failure occurs. It has several impacts, including substantial financial costs and damaging the company's reputation. It ends up sacking the leadership team and replacing the company with new execution management that complies with the government, regulatory, and safety guidelines. This option can lead to losing the power supply to many areas (partial or total system blackout) for a very long time. The power supply restoration can take longer, especially without the required experience and the time to recruit or hire them. In some cases, restoring power supply through portable generators is not technically and economically a solution as the required arrangements are not possible due to design issues within the station, which consumes time and is costly.

- **Option 2: Replacement Decision (Planned Outage) for the PT unit**

The ML PTHI model results indicated 18 transformers in “Bad” condition and 1 in “Very Bad” condition due for replacement. The sample collected data for 2019 consists of 12 transformer records in “Bad” condition with remaining life of 2 to 3 years. This comparison highlights the financial impact of considering replacement decisions versus corrective maintenance after a failure occurs. Table 8.7 compares the transformer replacement and failure costs, which caused a loss of power supply due to failure.

In this theoretical case analysis, the projection for the replacement cost of one 33kV transformer is about £537,000, involving purchasing, transportation, site preparation, installation, testing, and any necessary surplus equipment. In contrast, transformer failure costs involve interruption costs, which accompany the loss of power supply through the failure. In this scenario, the projection is £875,000 for the cost of the 33kV

transformers. The variation between the two scenarios is £338,000, which signifies the potential cost savings of 62.94% for ten transformers by replacing these transformers before the failure occurs; in contrast, not managing the transformer failure led to financial and operational consequences by not acting within the required timeframe.

Table 8.7: Cost calculation for the planned PT versus (vs) unplanned replacements after failure.

An assumption cost for the replacement before vs after failure			
Cost of one 33kV transformer replacement vs failure cost			
Cost component	**Replacement cost (£) ***	**Failure Cost (£) ****	**Difference (£)**
New transformer purchase	500,000	500,000	0
Transportation	5,000	5,000	0
site preparation	10,000	10,000	0
Installation and Labor	15,000	15,000	0
Testing and commissioning	5,000	5,000	0
Downtime costs (Loss of Power)	0	140,000	-140,000
Misc. (Others)	2,000	200,000	-198,000
Total Cost for 1 transformer	537,000	875,000	-338,000
cost x 10 (No of Bad transformer)	5,370,000	8,750,000	-3,380,000
Difference in % reduction in Cost for bad transformer(33kV)			62.94%
Cost of one 66kV transformer replacement vs failure cost			
Cost component	**Replacement cost (£) ***	**Failure Cost (£) ****	**Difference (£)**
New transformer purchase	700,000	700,000	0
Transportation	5,000	5,000	0
site preparation	10,000	10,000	0
Installation and Labor	15,000	15,000	0
Testing and commissioning	5,000	5,000	0
Downtime costs (Loss of Power)	0	200,000	-200,000
Misc. (Others)	2,000	200,000	-198,000
Total Cost for 1 transformer	737,000	1,135,000	-398,000
cost x 1 (No of Bad transformer)	737,000	1,135,000	-398,000
Difference in % reduction in Cost for bad transformer(66kV)			54.00%
Cost of one 132kV transformer replacement vs failure cost			
Cost component	**Replacement cost (£) ***	**Failure Cost (£) ****	**Difference (£)**
New transformer purchase	900,000	900,000	0
transportation	5,000	5,000	0
site preparation	10,000	10,000	0
Installation and Labor	15,000	15,000	0
Testing and commissioning	5,000	5,000	0
Downtime costs (Loss of Power)	0	250,000	-250,000
Misc. (Others)	2,000	200,000	-198,000
Total Cost for 1 transformer	937,000	1,385,000	-448,000
cost x 1(No of Bad transformer)	937,000	1,385,000	-448,000
Difference in % reduction in Cost for bad transformer(132kV)			47.81%
Summary of Cost Reduction			
Total Reduction in Cost (£)	7,044,000	11,270,000	-4,226,000
Total Reduction in Cost (%)			59.99%

* Planned replacement
** Unplanned replacement due to failure

Considering the accumulative percentage of cost-saving reduction for all 12 transformers in varying with the three voltage levels as one transformer is 132kV, one transformer is 66kV, and the remaining ten are 33kV, is equal to £4,226,000 which is cost reduction by 59.99% comparing to the first option. The cost analysis focuses on

the cost-effectiveness of ML PTHI model results, which will lead to better maintenance efficiency and the implementation of optimal cost value that should reduce the operation and maintenance costs before the new cost becomes higher than the project's initial cost. Figure 8.25 presents AM standard practice for correct decision-making based on the difference between the initial costs during project construction and the increased operation and maintenance costs. Suppose the system performance costs increase with the increment of operation maintenance costs activities due to various transformer failures, supply interruptions, and routine works. In that case, the quality of AM services management is achieved by managing costs, improving asset performance, and managing the electrical utility business using international best practices. Urgent actions and mitigations must be taken within the required time frame to achieve optimal savings. Therefore, balancing planned and unplanned electrical risks improves power system performance by improving the annual maintenance plan and managing risks caused by faults, frequent outages, and unnecessary spending. This analysis will strengthen the AM system as good AM helps electrical utilities operate their power system at high performance and low cost, as it is the best economic annual maintenance decision for optimal asset lifecycle planning investment (The Institute of Asset Management, 2015). Figure 8.29 shows the economic optimum lowest point where the initial project costs reduce as the maintenance costs increase, which means operating the power system at a high cost as operation and maintenance costs are higher than the initial project cost. PTHI analysis justifies the required decision for better AM investment and provides the AER with the funding justification.

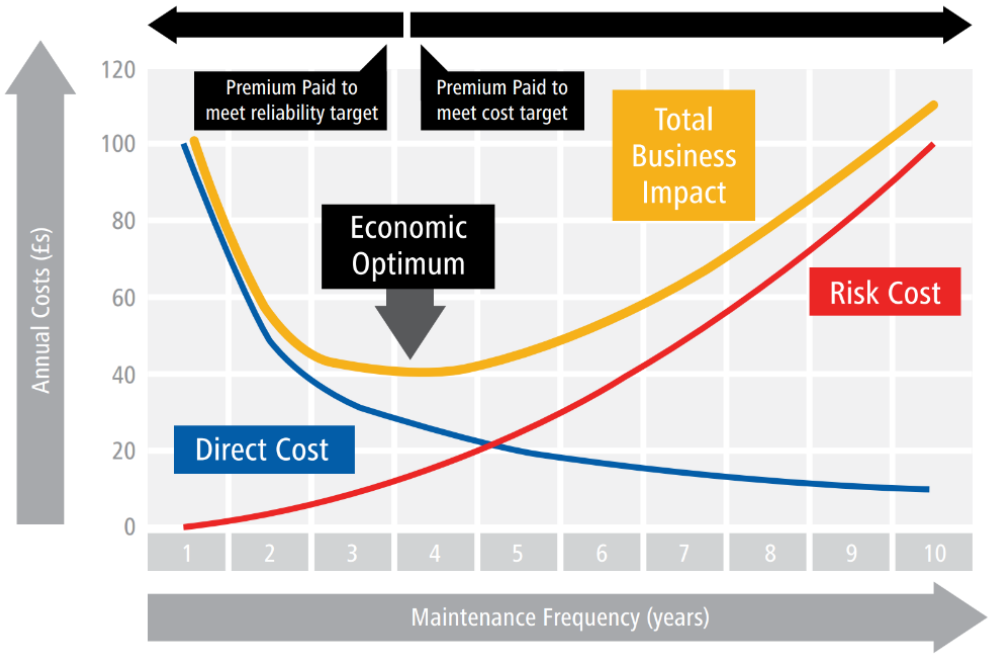


Figure 8.29: Economic optimum study (The Institute of Asset Management, 2015).

- **Option 3: Refurbishment Decision for Active Components**

The PTHI results require further analysis using electrical engineering experience and judgment. Table 8.8 shows the records of the 12 transformers identified earlier as in bad condition. The table is coloured using conditional formatting in a spreadsheet, which helped to focus on the parameters that show unnormal ranges using the conditional rules constructed from the knowledge obtained from the international standards (IEC 60599, 2022; IEC 60505, 2011; IEEE C57.104, 2019; IEEE C57.106, 2015) which vital for this research.

Table 8.8: Transformers in "Bad" condition.

TX ID	Voltage	Age	Colour	H2 ppm	CH4 ppm	CO ppm	CO2 ppm	C2H4 ppm	C2H6 ppm	C2H2 ppm	Moisture	BDV	DF	Acidity
1	33kV	69	Light Amber	70	17.1	461	5674	22.5	40.3	0.01	33	42	0.207	0.46
2	66kV	65	Amber	11.2	7.32	299	2827	8.4	6.56	4.55	13	60	0.267	0.24
3	33kV	60	Dark Yellow	38.7	58.9	183	3384	62.2	37.5	0.15	17366	12		0.5
4	33kV	60	Light Amber	88.1	93.4	268	3857	88.5	65.6	0.01	51	27	1.614	0.56
5	33kV	56	Dark Yellow	687.54	2207	186	1239	4230	702.1	36.15	21	39	0.081	0.04
6	33kV	55	Dark Yellow	18.7	25.3	110	2624	78.6	28.6	283.2	57	8	2.505	0.03
7	33kV	51	Amber	15.8	14	522	4272	40.6	9.75	0.01	44	23	1.181	0.66
8	33kV	51	Amber	20.4	5.69	437	4350	22.7	4.95	2	30	38	0.784	0.76
9	33kV	45	Amber	11.1	3.4	232	6078	4.93	2.19	0.14	21	62	0.619	0.14
10	33kV	34	Dark Yellow	72	7.2	95	654	8.16	1.41	63	24	56	0.018	0.06
11	33kV	32	Light Amber	14.4	2.46	140	740	2.7	0.05	12.2	19	74	0.19	0.07
12	132kV	20	Light Amber	15.6	74.5	160	2749	270	62.8	1.72	15	51	0.089	0.1

*Interpret DGA and OQA colour codes according to the scores in Appendix H, Tables H-1, H-2, H-5, H-32 and H-33.

Table 8.9: Nominal value ranges, conditions and scores for oil transformer.

Colour	Nominal Value Range	Condition	Score
Light Amber	0 ≤ colour < 1.2	Very Good	1
Amber	1.2 ≤ colour < 1.5	Good	2
Pale Yellow	1.5 ≤ colour < 2.0	Moderate	3
Yellow	2.0 ≤ colour < 2.5	Bad	4
Dark Yellow	colour ≥ 3.0	Very Bad	5

Transformers 3, 5, 6 and 10 have bad oil quality, as indicated by the dark yellow colour shown in Table 8.8 and 8.9 above. This finding led to the study of the decision on oil replacement cost and its effect on increasing the life span of a transformer. Table 8.10 shows a study of oil replacement cost vs transformer replacement after breakdown occurs. This accumulative percentage of cost-saving reduction for all 12 transformers equals £11,200,000, a cost reduction of 99.38%. This finding focuses on the cost-effectiveness of ML PTHI model results for driving better annual maintenance strategy efficiency.

The following finding could be discussed in further new publications or postdoctoral research to explore other possible solutions and suggest other rehabilitation techniques:

- The 132KV TX No.12 has high ethylene (C_2H_4) values of 270ppm.
- The 33kV TX No.1 has a high acidity of 0.46 ppm.

- The 66 TX No.2 has a high acidity value of 0.24 ppm.
- The 33kV transformers No.4 and No.7 have high moisture, break-down voltage strength of 1.614 and acidity of 0.56 ppm.
- TX No.8 has a “Bad” acidity level of 0.66 ppm, “Moderate” DF of 0.784, moderate moisture level of 30 ppm, and “Very Bad” CO_2 level of 4350 ppm.
- TX No.9 had a moderate DF level of 0.619 and a “Bad” acidity of 0.14 ppm.

Table 8.10:Oil replacement cost vs transformer replacement after breakdown.

An assumption cost for the replacement after failure vs oil replacement cost			
Cost of one 33 kV transformer replacement vs Oil replacement			
Cost Component	**Failure Cost (£) ****	**Oil Replacement (£)**	**Difference (£)**
Oil replacement	0	5,000	-5,000
New transformer purchase	500,000	0	500,000
Transportation	5,000	0	5,000
site preparation	10,000	0	10,000
Installation and Labor	15,000	0	15,000
Testing and commissioning	5,000	0	5,000
Downtime costs (Loss of Power)	140,000	0	0
Misc. (Others)	200,000	0	200,000
Total Cost for 1 transformer	875,000	5,000	730,000
cost x 10 (No of Bad transformer)	8,750,000	60,000	8,690,000
The difference in % reduction in cost for oil replacement for 33kV PT			99%
Cost of one 66kV transformer replacement vs Oil replacement			
Cost Component	**Failure Cost (£) ****	**Oil Replacement (£)**	**Difference (£)**
Oil replacement	0	5,000	-5,000
New transformer purchase	700,000	0	700,000
Transportation	5,000	0	5,000
site preparation	10,000	0	10,000
Installation and Labor	15,000	0	15,000
Testing and commissioning	5,000	0	5,000
Downtime costs (Loss of Power)	200,000	0	0
Misc. (Others)	200,000	0	200,000
Total Cost for 1 transformer	1,135,000	5,000	930,000
cost x 1 (No of Bad transformer)	1,135,000	5,000	1,130,000
The difference in % reduction in cost for oil replacement for 66kV PT			100%
Cost of one 132kV transformer replacement vs Oil replacement			
Cost Component	**Failure Cost (£) ****	**Oil Replacement (£)**	**Difference (£)**
Oil replacement	0	5,000	-5,000
New transformer purchase	900,000	0	900,000
Transportation	5,000	0	5,000
site preparation	10,000	0	10,000
Installation and Labor	15,000	0	15,000
Testing and commissioning	5,000	0	5,000
Downtime costs (Loss of Power)	250,000	0	0
Misc. (Others)	200,000	0	200,000
Total Cost for 1 transformer	1,385,000	5,000	1,130,000
cost x 1 (No of Bad transformer)	1,385,000	5,000	1,380,000
The difference in % reduction in cost for oil replacement for 132kV PT			100%
Summary of Cost Reduction			
Total Reduction in Cost (£)	11,270,000	70,000	11,200,000
Total Reduction in Cost (%)			99.38%

** Unplanned replacement due to failure

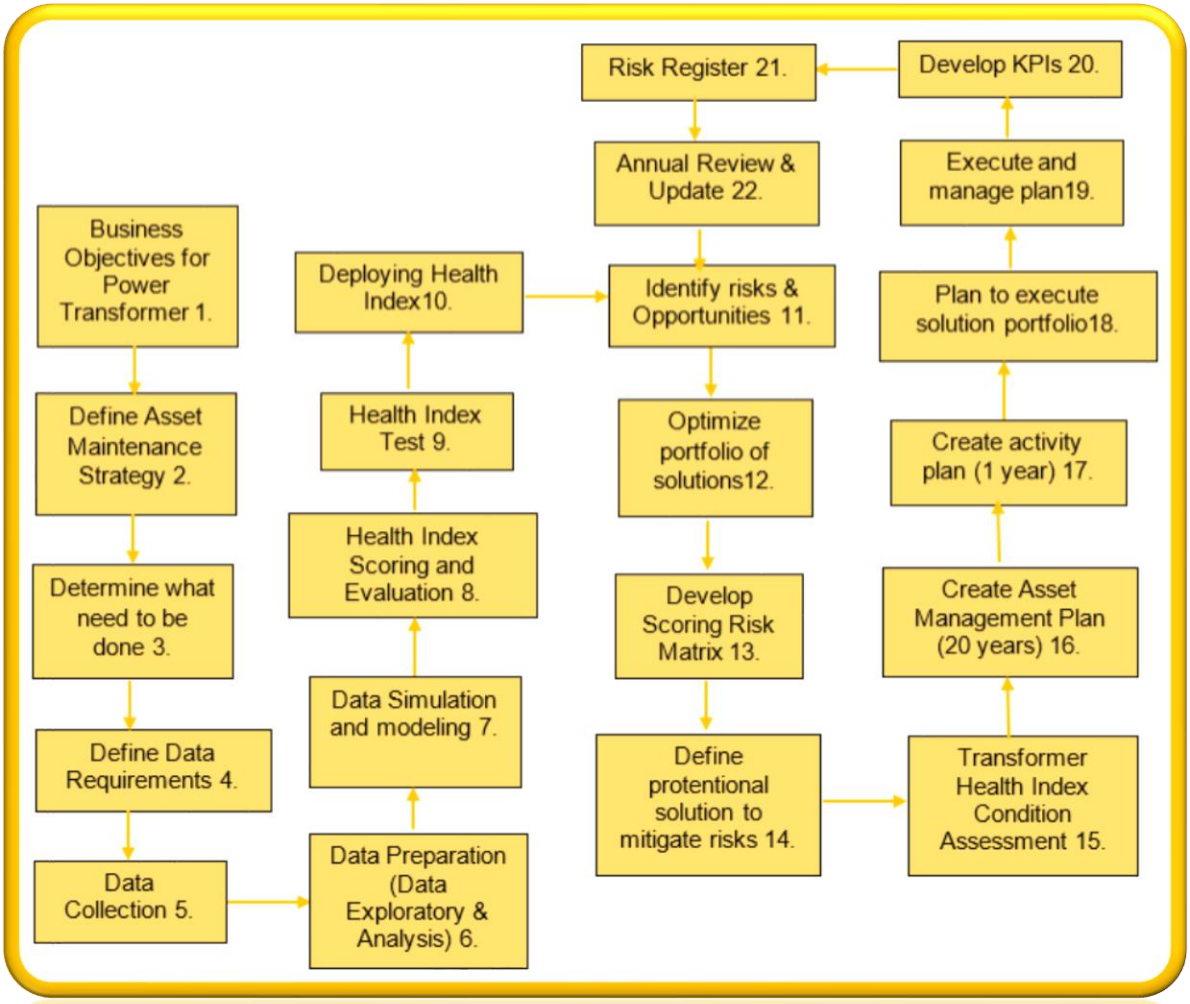
8.5 Summary

In this chapter, the age function is part of the HI for the electrical utility transformers to define the adjusted failure rate and consider accordingly the right decision. The research studied a large fleet of transformers' failure rate and the ageing of 1231 transformers from the electrical utility. This study will enable robust decision-making for transformer replacement plans, manage annual maintenance plan operating expenditure of the maintenance strategy and capitalise economically long-term asset lifecycle planning budgets, which adds an excellent contribution to the local and the national economy by the electrical utility. When assessing PT states, firstly, it requires recognising both failure modes or failure mechanisms, as well as the outcomes and involvements, to manage the sources of failure before it occurs. According to the analysis performed to analyse a large fleet of transformers' health indices, the average population utilising regression analysis with a coefficient correlation of R^2 equal to 0.8165 shows an accuracy of 82%. The apparent age plots a line from the transformer using the PTHI value to the required regression line. In this research, the transformer's apparent age indicates the health condition since the actual age does not always reveal the real transformer condition. Transformers with accelerated ageing are those who are older than their actual age. This process intends to assist asset managers in decision-making regarding maintenance in electrical utilities to maintain the transformer's reliability and avoid sudden transformer failures. The financial analysis shows the research's positive impact using ML PTHI model results and studying three scenarios for possible cost reduction and savings. The result showed a cost reduction of 59.9%, considering the planned outage for the replacement decision. Also, there is a cost reduction of 99.38% considering the oil replacement decision; however, further postdoctoral research can be crucial in further improving this research area.

Next, chapter 9 presents the conclusion and recommendations chapter, which includes the research outcomes, the learning lessons, limitations, and future improvements. It also presents an illustrative AM process for considering the HI for the annual maintenance and long-term lifecycle planning.

Chapter 9

Conclusion and Recommendations



Chapter 9: Conclusion and Recommendations

9.1 Introduction

Electrical power transformers are considered high-priced assets within electrical energy delivery networks. To provide a good-quality asset management service, a PT should have a robust annual maintenance plan to ensure its reliability and performance at high service and low cost, especially over the long run. Solving today's short-term annual maintenance plan and maintenance strategy is a solution to long-term maintenance lifecycle planning.

Managing ageing PTs is the core concern, and it is challenging for electrical utilities based on the actual case, as shown in Figure 8.4. Ageing concerns are mainly about ensuring in-service PT at the required performance and reliability for the estimated time and ensuring PT insulation strength in oil-infused paper insulation as per standard. As per the manufacturer catalogue, industrial power transformers are around 30 to 40 years old, and PT will possibly survive up to 60 to 80 years with an excellent annual maintenance plan. Nevertheless, some power transformers may fail before their time due to improper maintenance handling and wrong daily operation practices. This research considered PTHI findings to complete the PTHI investigation, including the apparent age study to estimate the PT failure rate and that PTHI scores correlate more to PT age, as discussed in Chapter 7, Section 7.2.3 (Figure 7.4) and Chapter 8.

This research presents a practical process of evaluating PT failure rates through evaluating PT health conditions using diagnostic laboratory test parameters and maintenance data. A regression model in R studio considers obtaining the finest model by investigating the PTHI relationship with transformer age information. Then, the required process is determined to find the adjusted probability of failure, considering the health condition assessment. The PTHI model facilitates short-, medium-, and long-term risk investment AM decision base. These strategic AM-based decisions for annual maintenance planning categorisation follow preventive and corrective maintenance practices. The study of failed-in-operation PT parameters presents a prevention decision that practically fits the proposed PTHI process.

This research investigation played a strategic role in implementing the learning for

improving the national economy and future government funding price control as a long-term budget planning pathway for electrical utility by utilising PTHI as a good decision-making approach. The PTHI model can shift the capital plan, reduce the costs of the replacement plan and manage the required targets for the large fleet of PT that suits best predicted PT health condition, which means considering international best practices for budget deferral up 15 to 20 years of the capital plan and reducing PT future replacement costs.

The research PTHI model examined most of the earlier academic investigation methods of several researchers' works, approaches, and rules considered in this research area. This testing used the available data in the literature, and the outcomes obtained were acceptable. The comparison and benchmark have improved the accuracy of the PTHI model calculation, and the results were 93.7% (RF) and 93.5% (SVM) for the collected actual dataset from the electrical utility. This research includes the development of the ML and conventional mathematical PTHI model for electrical utility, and for dataset confidentiality, the company is named only as an electrical utility. The research used the electrical utility data and implemented the PTHI model for studying PTs' health condition, which follows the categories of “Very Bad”, “Bad”, “Moderate”, “Good”, and “Very Good”. The research impacts include transferring technical PTHI analysis to a web app tool and developing the financial model, as well as studying worst-case scenario analysis that requires better monitoring of the annual maintenance plan, maintenance strategy, and risk decision base for lifecycle budgeting deferral.

9.2 Power Transformer Asset Management Health Index Modelling Analysis

PTHI models consider using different analysis approaches based on the conventional and AI ML algorithms on three datasets: small, medium, and big data. The datasets cover the input parameters for PT health condition analysis, which include DGA, Furan, Age, and Oil (Water, BDV, Acidity, DF, 2FL, and DCG). This approach forms the research novelty strength, creativity, and contribution of using multiple PTHI approaches. PTHI calculation considers the AM framework standard of ISO55000 for the electrical utility annual maintenance plan and lifecycle planning. The 18 analysis models are three PTHI conventional models in the spreadsheet for the three datasets and seven PTHI ML models for the 30 transformers dataset. Six PTHI

ML models for an extensive dataset and two apparent age prediction models that depend on HI condition using linear regression algorithm were used to map apparent age with failure rate analysis for the extensive actual dataset of the large fleet of transformers of a total of 1231 records. These records were experimented with in Excel spreadsheets and R programming language.

The data preparation process includes data cleaning and typecasting transformer health records based on applying the researcher's electrical engineering experience and academic research knowledge to design the PTHI model and international best practices.

The SVM algorithm development for the PTHI model web application was built in R studio to show the technical impact of this research on the decision-making process required for PT health conditions. PTHI web application is designed based on the related international standards; for oil test parameters analysis, it uses IEEE C57.106 (2006), and for DGA test parameters analysis, it is considered IEC60599 (2007). This consideration for the ranking method parameters of transformer voltage rules: Unites (U) ≤ 69 kV, 69 kV < U < 239 kV and 230 kV < U.

It was learned from the clash between the two web application forms and this repetition problem that the functionality division of the web application introduced for representing PTHI of this research needs. The new design using the PTHI web application included only the HI tool that will utilise the PTHI model to forecast the health condition of the PTHI for the PT estimated health condition. This problem has led to the introduction of Flex Dashboard. The purpose was to demonstrate data records visualisation for the new balanced dataset that includes bad transformer records for correctly training the SVM algorithm. Flex dashboard provides a simple data visualisation tool to build interactive dashboards in the R programming language. The practical implementation requires using R markdown to present data visuals for a large fleet of PT conditions in a dashboard format and including them in the PTHI as a recent development in this research area.

The dashboard of PTHI includes report export functions to other formats in HTML or PDF, which is more practical for sharing information, and there is no need to deploy it on a server for the electrical utility.

9.3 Response to Research Questions

The research answers to the following research main question which is:

"How can Health Index identify health conditions for electrical utility PT annual maintenance plan and assessing lifecycle planning?"

by completing the following sub-questions:

1. **What are the active parts of the power transformer that impact the power transformer's health condition?**

These are the core, winding, tap changer, bushing, oil, tank, and auxiliary equipment. The practical and technical knowledge about different PT active components is available in detail in Chapter 2 to better understand the PT failure modes. This knowledge must be gained and learned before further recognising PT deterioration mechanisms and failure modes and developing PTHI.

2. **What data and which parameters are vital to study power transformer health conditions using the health index study, and how is the analysis processed?**

Laboratory data, maintenance records, site inspection, monitoring data, condition boundaries, ageing and failure statistics form the data requirements for the HI study. In order to process the data probably, it is first essential to understand the business objective, collect data from the available source, understand the data by analysing it, prepare the data for modelling using multiple data engineering techniques, and implement within various case studies chapters (4, 5, 6, 7 and 8). The actual data investigation took up to 16 months for data collection, preparation, and analysis as it required a thorough study because, within the collected data, the names of substations, voltages, and input parameters were not identical. It is also necessary to organise the records to be efficient. The records prepared are based on industrial experience and academic research needs that are best suited. Considering the facts that this PTHI development covers the following analysis input parameters that consider investigation PT health condition as factor level:

1. OQA Sample Tests.
2. DGA Sample Tests.

3. Furan Sample Test.
4. Age.

The above PTHI parameters analysis is available in Chapter 6 and Appendix D. The data is handled manually row by row because the names of substations, voltages, and input parameters need better analysis within the electrical utility files. Understanding the data and why some sheets were available while others were missing was challenging. Hence, the records analysis is essential for designing the PTHI research simulation model.

The parameters used for the PTHI study are collected from the literature review and the electrical utility to reflect PT's health condition. For the actual data, it is vital to use the sample number, serial number, asset ID, substation name, and transformer ID to develop the PTHI model. Initially, the HI mathematical model formulation platform used conventional HI in Excel. Later, the learned experience was crucial in developing a practical ML PTHI prediction model and web app tool that aids decision-making for PT annual maintenance and life cycle planning.

In addition, it is an achievement to identify the required patterns that describe and reflect transformer classes. Then, transformer issues can be solved during the annual maintenance plan using preventive planning maintenance or replacing the PT if the maintenance cost and time are more than the actual cost value of the defective power transformer. A detailed process is available in section 4.3.2 Figure 4.2.

3. How to Develop A PTHI Conventional Model Using the Identified Parameters?

By identifying the available parameters, it is assigning the weights and scores and building formulas based on the weight score mathematical model in MS Excel as shown in Figure 4.2 and in Appendix H, Table H1-H6 (Table H-1, Table H-2, Table H-3, Table H-4, Table H-5, Table H-6). This conventional model accuracy is measured by benchmarking with earlier research and studies in section 4.4.3 Conventional Health Index Models Benchmark.

4. What ML algorithms to develop PTHI models and asset management decision-making tools?

Multiple ML algorithms can be applied. This research uses RF, DT, SVM, ANN, NB, and KNN to develop PTHI models as multi-class classification problems for PT HI condition classification.

In this research, applying ML algorithms requires building two ML classifiers that use ML algorithms to classify PT health conditions and decisions. The second ML model classifies the recommended decision to inspect, refurbish, or replace PT, considering the predicted health condition of the concerned power transformer.

Chapters 6 and 7 show that the SVM algorithm plays a good role in predicting future PT failures. This prediction will achieve preventive maintenance objectives and decide to prevent failures before they occur by conducting the required maintenance plan within the required time scale. Hence, using computer science, ML algorithms, and AM strategy based on PTHI calculation scores forms the technical driver for the electrical utility maintenance strategy. Furthermore, the PTHI web application tool design through the deployment of the SVM model in Shiny web application using R programming language is available in Chapter 8.

5. How can the PTHI condition be used to define the right asset management decision?

The PTHI model defines the required decision for AM to drive PT's annual maintenance plan and lifecycle planning by mapping the maintenance actions based on transformer health condition into the annual maintenance plan and lifecycle planning that drive AM and the maintenance strategy for the electrical utility. Also, the failure rate can be mapped based on apparent age, which is dependent on the transformer's health condition, using data engineering techniques. In addition, to ensure the right decision is considered by accurate evaluation of model performance using multiple evaluation techniques, including model accuracy, sensitivity, specificity and F1 score, and risk-based, available in chapters 7 and 8, including failure mode indicators.

9.4 Conclusion

The “Power Transformer Health Index” is successfully applied as a good AM tool to assess and categorise transformers’ health conditions for future enhancement, including preventive and corrective maintenance work orders. PTHI is also an accountable tool for a new investment that covers end-of-life PT replacement decisions. Different diagnostic tests, factors, and indicators play strategic and technical roles in customising a good PTHI to optimise the required annual maintenance and capital plans. The discussed PTHI method is an approach towards a successful PTHI formulation where both the accuracy of the data and the formulation process play a difference. As a result, PTHI has become an increasingly suitable condition evaluation method in driving strategic planning asset management for improving maintenance plans and extending the PT lifecycle. This direction requires managing the in-service PT, potentially focusing on daily operation and maintenance planning lifecycle. PTHI plays a strategic role in PT's preventive maintenance planning and decision-making on health conditions for electrical utilities. In this research, PTHI considered the AM standard framework ISO55000 for improving operation and maintenance programmes that require linking the whole business together for HI decision-making to be considered based on the fundamental asset management factors such as performance, risk, and life cycle cost decisions. Therefore, the PTHI as a decision tool will contribute to the total asset management long-term investment master plan. The research PTHI formulation uses electrical engineering field experience, international best practices, and CIGRE, IEC, IEEE, AM, and ISO standards. During participation with the International Transmission Operation and Maintenance Study (ITOMS) group between 2007 and 2013, which includes around 33 electrical utility companies, including the UK national grid, best practices have developed as ITOMS offers international best practices by benchmark international electrical utilities which are interested in operating their electrical network at low cost and high services. These best practices will support the control of their future funding and asset ageing challenges, specifically for reliable transformers.

Recently, AI became word of mouth based on this research literature review findings as an excellent innovative solution for big data that improves data accuracy and reduces high dependency on power transformer experts' judgments and recommendations for PTHI development. This research has considered it a primary

driver for the HI maintenance strategy decision-making and solving missing data. AI is becoming an alternative approach to estimating the value of the missing data using data science, ML, and deep data learning technologies as AI can overcome the existing conventional condition evaluation method issues that include data redundancy, data uncertainty, missing data as well as manage the high need for PT expert knowledge and experience. The investigation of the PTHI model using the electrical utility collected this data has shown that PT ageing is the only parameter showing a high positive correlation between transformer health condition and HI calculation. This finding has led to further analysis deploying the AI for the PTHI calculation method in different analytical scenarios as it plays a long-term solution in extending PT life, managing spare parts, managing failure, deferral unnecessary spending by managing and prioritise frequency tasks of daily maintenance switching schedules as considering today solutions as an excellent permanent solution for future problems and solving maintenance strategy issues. Significantly, the utility has reduced its investment in the new transformer replacement plan. This finding could be due to improving its maintenance strategy by introducing better condition assessment, including PTHI and CBRM, which occurred from 1984 onwards and has avoided unnecessary spending and deferral of a replacement plan.

Since PTHI is vital for electrical utilities, coding and algorithms form the right angle for data science strength and consider for the data calculation PTHI risk matrix by employing four measures that include accuracy, specificity, sensitivity and F1 score that consider identifying the suitable health condition classes to consider correct decision making. AI is considered a key factor for data issues utilising related parameter indicators. The actual and estimated data play a role in improving and ensuring data-driven decisions. The PTHI risk decision condition is assessment leads the maintenance frequency task base that replaces the equipment base reliant on real data quality, which reflects actual unit conditions on a large fleet of transformers. However, good integration between PTHI and probability of failure is a must to reflect with the available actual datasets and provide a reasonable justification for the considered decision by electrical utilities for their number of frequency maintenance tasks, which can be increased or decreased through a good optimisation of the annual maintenance plan and maintenance strategy that rely entirely on the formulation of PTHI as it is an exciting process even if it consumes much time.

This research investigated techniques to develop PTHI, bringing innovative practices that can contribute to academia and electrical utility daily asset management planning for robust maintenance decision-making. Electrical utilities can gain complete access to this purported PTHI methodology, which must integrate with the AM standard framework and practical technical electrical engineering experience.

PTHI results pinpoint those power transformers which are near to end life or deterioration state point. This result segregates PTs with a higher failure rate. This knowledge improves preventive and proactive maintenance decision-making for electrical utilities. Hence, the maintenance team's attention draws to which transformer should remain in operation or which should go for repair, refurbishment or replacement. PTHI theory involves input parameters, a mathematical formula used in the HI calculation, and results considered for a robust investment master plan that mediates extending the PT maintenance lifecycle and reflects on budget deferral of any nearby investment.

PTHI calculation parameters state the measurement analysis for considering the active parts of power transformers, causes of failures, and allocating scores and weights for the ranking methods using electrical system perception, shown in a series of parts over the research chapters.

The research declared based on this researcher's practical experience, and according to (Khamis et al., 2023), the essential facts are the importance of daily maintenance activities as solving today's maintenance problems is a solution for future maintenance permanent problems. The chance of solving maintenance issues can be actioned based on three factors:

- Kits and spare parts accessibility
- Understanding PT specification
- Good level of technical experience among maintenance crews
- Implying the change in frequency and nature performed on PTs maintenance strategy and capturing changes in PTHI if the maintenance task frequency increases or decreases or the impact of cancelling or adding a new maintenance task based on prioritising the annual maintenance plan and strategy.

PTHI indicates the transformers are in a poor state, signifies the relevant transformers can fail at any moment, and replacement decision is the best option that covers the maintenance actions, assuming this mechanism is irreparable or while the transformer functioning life ends, or physical state is not within a technical and optimal performance that shows transformer will no longer sustain in operation over the short term. Hence, replacement decisions are made for specific parts or as complete parts that rely on transformers' techno-economical knowledge.

Each PT primary and sub-failure mode for each sub-system classification into four main classes:

- Thermal
- Mechanical
- Chemical
- Electrical

There are numerous reasons for the breakdown versus all existing mechanisms to the extent of insight into the maintenance planning team. However, the recommended method is for the benefit of the maintenance planning team to:

- Calculate PTHI scores based on actual transformer data.
- Extend the residual life of the power transformer.

The features of scores and weighting vary on the specific decision of every electrical utility situation to alert the transformer probability of failure level. PTHI with AI to:

1) Determine the PT's remaining life.

2) Assess the risk to the organisation of power transformers as ‘risky assets’ that impact network performance due to risky strategic transformers.

3) Identify top priority capital replacement for total investment plan ‘total expenditure’.

4) Enable optimising maintenance strategy on the ground with the overall direction of the business mission, vision, AM policy, objective, strategy, plan and procedures

PTHI index score values help recognise risky PTs that require more consideration than other transformers and aid in prioritising maintenance plans through the best use

of PTHI. The best international measurement is the average quantisation against conditional ranges for different measurements.

The conclusion is that there are various methods to calculate HI based on research and practice. Every researcher within the research area followed different approaches, which are all valid. Still, there is a need to manipulate the rules based on manufacturer reference, transformer type, voltage limitations, standards used, and data interpretation approach by the electrical utility. PTHI justifies PT's annual maintenance plan for the Asset Maintenance Strategy (AMS). It also improves the secured PT performance and decreases unnecessary spending.

9.5 Recommendation for Maintenance Management in Electrical Utilities

PTHI methods are critical condition assessments for electrical utilities (Khamis et al., 2023). Therefore, PTHI integrates with the strategic AM standard framework (ISO55000 series) for electrical utility business practices to improve the data portfolio and break the silo within the utility department for developing good PTHI.

The research recommendations focus on improving three areas: the PTHI technique using an AI approach, integrating PTHI with AM strategic and field experience, and developing a balanced data-driven maintenance strategy.

9.5.1 Enhance Power Transformer Health Index Assessment.

9.5.1.1 Improve Health Index Assessment Analytics Using Artificial Intelligent Approach

PTHI using AI, machine learning, computer, and data science will improve the PTHI assessment technique to manage significant data uncertainty issues, missing data, PT health condition classification, and dependency on expert judgements for electrical systems. PTHI design requires ensuring the constructed model derives a probabilistic representation of the HI, seeing unexpected data acquisition, interpretation, and modelling uncertainties. This model will lead to preventing failure before failure occurs and taking the required countermeasures.

9.5.1.2 Extend Long-Term Transformer Lifecycle Planning and Deferral Unnecessary Spending Utilising the Relationship Between the Power Transformer Health Index and the Probability of Failure.

PT failure rates must use AI to estimate the predicated effective age of power transformers compared to the actual age. Based on the correlation, it is necessary to link the health condition with the probability of failure. Failure occurs when the transformers are stressed internally or externally and exposed to exceed their rating strength. The probability of stress exceeding this strength leads to failure. This prevention involves calculating health conditions and the probability of failure, which can assist in the necessary HI decision-making.

The transformer residual life presents the remaining age between “life expectancy” and “effective age”. It utilises the PTHI score by investigating the probability of failure. The condition of PT is not possible only by considering particular statistical analysis, whether age, visual inspection, laboratory test or international standard. Using the comprehensive analysis for PTHI, laboratory tests and ageing will extend transformer life and defer unnecessary spending, including the replacement plan. As a result, this will support the decision-makers by knowing the time that system performance costs from operating expenditures are more expensive than the actual initial capital cost for each specific transformer. Hence, this will lead to knowing the optimal economic cost of the PT replacement decision.

9.5.1.3 Integrate HI with Asset Management Standard Framework.

Integrating HI with AM standards within the electrical utility helps to break the silo by developing asset management practices and culture. Breaking the silo between technical departments will solve data issues for developing a good HI. Furthermore, this will improve the decision-making process by considering good-quality data.

Electrical utilities must start to leverage the output of the PTHI using strategic AM standards with better daily maintenance operational planning since solving today’s maintenance problems is a permanent solution for future problems caused by wrong maintenance at first time. The transformer health parameters assist in developing PTHI that justifies the reasons behind each decision and new spending. Electrical utilities must maintain the technical PT speciality while developing PTHI and support

the earlier recommendation for the required deferral of unnecessary spending. Doing this means PTHI will enable asset owners and AM (operation and maintenance) to agree with the required decision, including scheduling retirement, operating and maintaining the transformers more reasonably and operating PTs at low cost and high performance by optimising maintenance plan and strategy to increase or decrease the maintenance frequency tasks in a day or month or a year which assumed will solve future long-term problems. Doing so will conclude by applying a justified PTHI approach that supports the holistic, big picture of AM based on AM's best-performing electrical utilities.

PTHI is more efficient with the AM standard framework as the idea is to ensure the total value of a good PTHI recognis by the whole business. The AM standard and framework improve network performance, minimise risks, and allow unnecessary spending in optimal PTHI. It is imperative to learn the usefulness of implementing AM culture, which needs to be developed for electrical utility by understanding and implementing AM ISO55000. Supporting these arguments requires conducting a PASS 55 health check and gap analysis using the PAS 55 Assessment Methodology (PAM) tool to analyse and prioritise the AM implementation journey for the related electrical utility organisation.

9.5.2 Enhance Data Management Practices in the Electrical Utility.

9.5.2.1 Improve Data Management Processes

Data management is the biggest challenge for PTHI development in this research. It was necessary to develop the correct dataset using the detailed data preparation, validation, and analysis process. This thorough process is the key to good PTHI results. Using the most actual data helped to select related indicators to calculate the final PTHI. In standard practice or research, the data analysis depends on direct/indirect measurements from site inspection and laboratory diagnostic tests. The failure data relating to a failure mode and mechanisms will identify the PTs' health condition that needs more attention than others. PTHI classified a large fleet of PTs based on the same set of manufacturers, specifications, serial numbers, and ages that needed to be studied together. A typical PTHI must consist of five elements:

1. The asset ID includes the manufacturer's name, serial number, installation date, initial capital expenditure, and area information covering the weather, soil, and temperature records.
2. Failure mechanisms are a collection from regular maintenance, inspection, and deterioration analysis records.
3. Fault cause investigation: investigate breakdown causes, plots, and the relationship between dataset parameters.
4. Cruciality and risk decision: Individual PTs within the large fleet of transformers depend on the transformer's location, service level, and rating capacity.
5. Usage: Current usage includes in-service transformers, loads considering the voltage level for step-up or step-down transformers, rating capacity, and supplied area.

9.5.2.2 Improving Data Science Knowledge and Skillset of Electrical Utility Electrical Engineers.

This skill will assist in efficient data handling and analysis and reduce the cost associated with specific data collection projects and resources, whether contract-based or new staffing. PTHI investigation needs a practical focus that relies on accurate data from the electrical system end users, including PT expert judgements, knowledge, speciality, and agreed-upon PTHI national scoring and weighting practice.

9.5.3 Improve the Maintenance Strategy Tactics.

The big picture of AM requires improving the maintenance strategy tactics and culture by moving from reactive, proactive, and controlled to data-driven maintenance strategy PTHI decision. It starts with improving short-term maintenance plans to predict PT's health conditions for better AM services. This data quality utilises good PTHI that integrates failure and mechanism modes using good quality data. The quality of data improvement is by developing a good asset register that maintains a regular update by gathering all information needed for robust PTHI that masters both short-term maintenance plans and asset lifecycle planning. World-class electrical utilities are always in advance to fulfil AM practices for better implementation of asset maintenance strategy, which serves predictive and analytical maintenance decisions.

PTHI principles help avoid misleading international AM best practices, typically applied to avoid unnecessary decisions that improve preventive maintenance management and investment for operating expenditure projects. PTHI reinforce and improves PT speciality through the required details and explanation of the health condition analysis and failure mode investigation results obtained from PTHI and gives a better defence for 30 30-year master plan. Figure 9.1 presents a research conclusion initially developed to study and examine an essential aspect of the overall AM process, the PTHI, which is considered for actual implementation by the electrical utility.

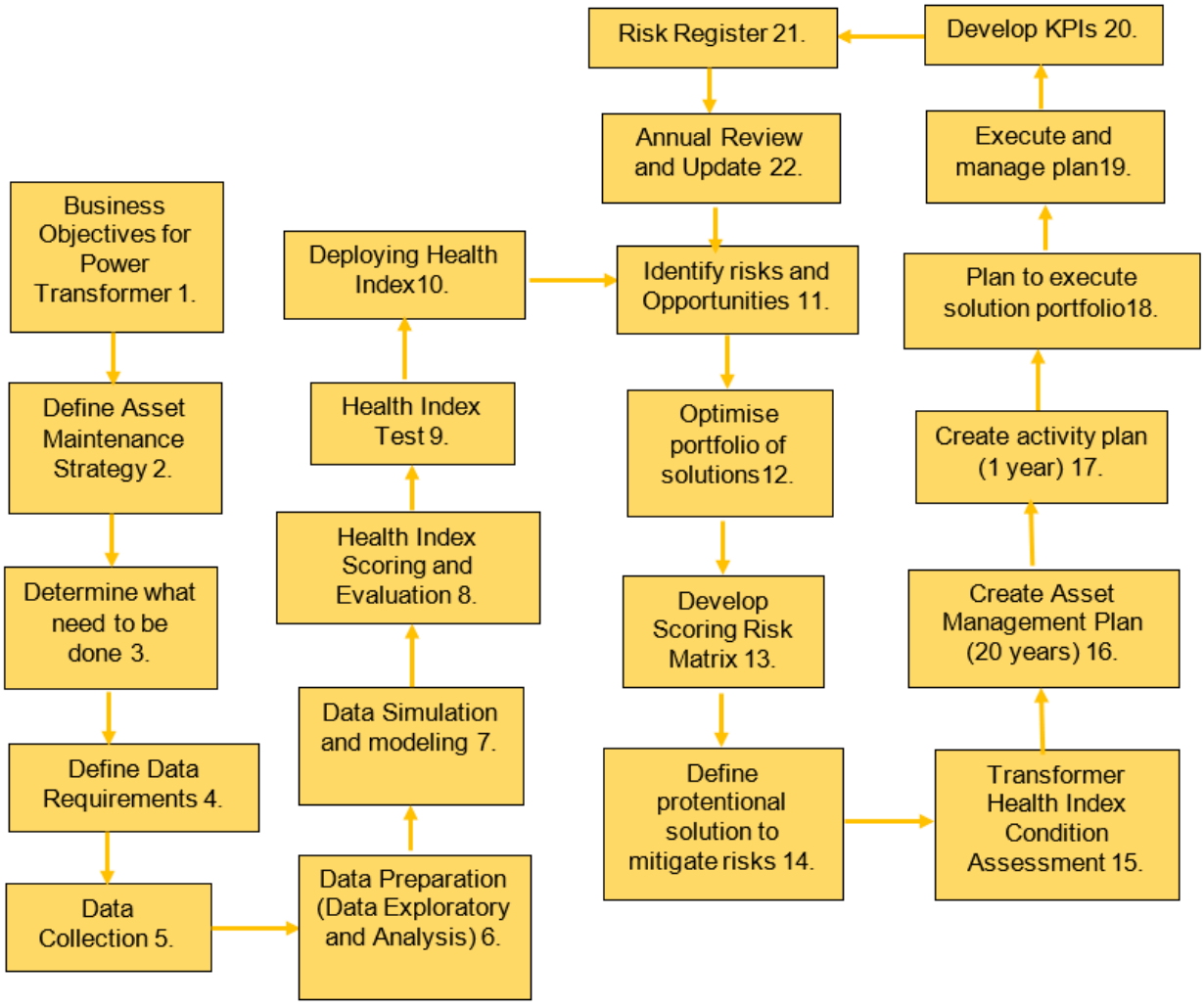


Figure 9.1: PTHI practical AM process for assessing transformer lifecycle.

9.6 Research limitation

The research was limited to the following issues and resources:

- Time constraints during the extended year: the time between the data collection enquiries periods and the study completion time is very short. It would be better if the data collection from the electrical utility were during the first or second year, not at the extended year, at the end of semester 1. The time for data collection, preparation, cleaning, structure, analysis, and HI design was the most challenging constraint. This time constraint is why the HI benchmark approach with other researchers in the research area was the key to ensuring good HI development.
- A partnership between the electrical utility and the University will facilitate students with data collection. The need for such a partnership is urgent. It would be better for the university engineering faculty to partner with the local electrical utilities Research and Development (R and D) team. This study will enable the university researchers and the electrical utility department to work together and conduct studies and research for AM, operation, maintenance, and maintenance planning activities, including improving PTHI. The development of PTHI by electrical utility companies on academic findings and daily utility practice. It built the PTHI model as a valuable and auditable tool, as the PTHI customisation to the needs of the electrical utility. This research could be a combined project between the university and the electrical utility and structured in a return of investment way that can be more beneficial than now to the electrical utility and long-term to the research area. The electrical utility can benefit from this research ML algorithm investigation. This benefit must contrast further automation in the current process with a sufficient level of utility PTHI and a data-driven decision-making process. This process concludes that this research is limited to the absence of an electrical utility specialist, maintenance engineers' interaction, practical knowledge, and experience sharing with the university's area of research.
- PTHI evaluation was limited to the provided data: the data was not available for all of the power transformers HI study parameters. The data provided were limited to laboratory diagnostic tests, which were still not provided in the required details.

- PTHI research is limited to PT's speciality, knowledge, and resources within the university. However, this was explored widely by academia and other researchers through several publications.
- Data management challenges are the biggest hurdle for developing research HI. It necessitates the development of the correct dataset. This challenge requires a meticulous process for data cleaning, preparation, validation, interpretation, and analysis, which is successful because the actual data is the key to better HI results. Hence, the most relevant actual data is to select a key indicator for developing the final HI, ensuring the thoroughness and validity of the research.

9.7 Research Contribution to Knowledge and Basis for Originality

The research achievement contributed to a few publications and participations that presented the challenges in several case studies, as discussed in detail in the four technical chapters. The research includes a contribution to the journal Int. J. of Strategic Engineering Asset Management with the paper 'Health Index Assessment for Power Transformer Strategic Asset Management in Electrical Utilities'. Similarly, two other contributions to the Sunderland University research conference on 15 September 2022 and a further publication with a paper entitled "Principles of Asset Management Health Index Development in Driving Maintenance Strategy for Electrical Utilities". Also, another contribution at the 12th IMA International Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) on 4-6 July 2023, Nottingham University, Jubilee Campus, United Kingdom.

A web app HI tool utilising an ML model approach, the electrical utility dataset, power transformers condition classification, and the recommended research action using a shiny web application in R programming language. The app ran and tested all the data, showing that the electrical utility transformers are in good, fair condition and that there are no "Bad/Faulty" transformers based on the provided data. This HI tool reflects the significant technical impact of this research for site engineers to speed up finding the health condition of the power transformer by using the developed mobile web app to make the required decision before going for a detailed PTHI assessment.

Furthermore, this research understands the ML evaluation metrics for imbalanced datasets in PT health condition evaluation, specifically the achievement of the F1

score for each health condition class. In this PTHI research, the F1 score is considered and confirmed for the first time in this field compared to other researchers' literature reviews that used only overall accuracy as the best evaluation metric.

The contribution of improving AM to the business knowledge and culture will support effective PTHI data decision-making, especially within diverse situations for the categorised assets level and according to international best practices and standards. This research introduced the AM standard ISO 55000 framework as a best practice for implementing PTHI data-driven decision-making (Khamis et al.,2023). PTHI will support any electrical utilities or power system network operators with numerous stakeholder objectives and seek to provide coherent and adequate energy to connected consumers by integrating a good quality of AM service within the whole business and the country. This implementation of AM framework will break the silo among technical departments, which will ensure the PTHI decision made is on data-driven decisions. PTHI is according to the researcher's electrical engineering field experience and ITOMS international best practices, IEC, IEEE, CIGRE, ISOs and AM standards.

The financial impact of this research in actual practice that considered worse case scenarios for electrical utility based on the electrical utility provided prices for power transformers of 132kV, 66kV and 33kV. The prices are rolled out within the research financial analysis and show a considerable saving impact for implementing ML PTHI model results of 59.9% considering the planned replacement decision.

9.8 Further Research

To support the development of the testing technology area, more economical solutions to solve PT failure modes and deterioration mechanisms issues by investigating specific reasons within the PT system and its active components, including parameters like acidity and DGA analysis.

Also, investigating the lifecycle costing model with health conditions is another space for improvement within the research area. This model requires studying the total expenses fortitude for transformer components spare. There are still many gaps, and efforts must automatically improve long-term network master planning using online PTHI. A better collaboration between the engineering department within the university

and the local electricity utility R and D team will support further development and add more value to the local society, industry, academia, and economic growth in the national economy using PTHI. Additional assessment is crucial, along with a sensible aspect. Hopefully, this will increase the number of researchers.

To investigate integrating the Fuzzy and ML models and conducting a detailed PTHI analysis for new datasets. Further research investigating maintenance task frequency is required, as it should be automated through an automated data pipeline to ensure the quality of the model is up to date based on the continuous AM improvement learning cycle integrated into this new research with field experience and starting online monitoring PTHI, which requires investigating differently the available FSA technologies.

Furthermore, further research is required to investigate other expensive main electrical assets, including overhead lines (OHL) and Gas-Insulated Switchgear (GIS).

Further postdoctoral research to explore other possible solutions and develop a solid foundation can be the key to attracting new researchers.

References

A review of Equipment Aging Theory and Technology (2020) in SearchWorks catalog. Available at: https://searchworks.stanford.edu/view/7097006 (Accessed: 04 August 2023).

A. Bossi, J. Dind, J. Frisson, U. Khoudiakov, H. Light and e. al (1983) 'An international survey on failures in large power transformers in service', Electra, no. 88, pp. 21-48, 1983.

A. Patle and D. S. Chouhan(2013) 'SVM kernel functions for classification', International Conference on Advances in Technology and Engineering (ICATE), Mumbai, India, 2013, pp. 1-9, doi: 10.1109/ICAdTE.2013.6524743.

ABB Brochure 2750 (2014). Bushings Diagnostic and Conditioning [Brochure] Available at: https://library.e.abb.com/public/60d37661a03145b993bad1795a43bfe5/2750%20515-142%20en%20Rev%202.pdf?x-sign=Prm9OA6mLWfvTJlSxrx0dE0zwKb3F%20F9dzqENHhfGlU066N/7KzP%20QRq6dzFjj7P

ABB Management Services Ltd. (2007). Service Handbook for Transformers (3rd ed.). Zurich, Switzerland: ABB Management Services Ltd.

ABB Review 1 (2008). Transformer challenges. Available at: https://library.e.abb.com/public/5d00c795b313f965c1257427002d50e9/29-33%201M808_ENG72dpi.pdf

ABB Transformer Services (TRES) (2018) 'Bushing replacement Upgrade to the latest technology and mitigate risk', ABB.

Abbasi, E. and Malik, O.P. (2016) 'Failure rate estimation of power transformers using Inspection Data', '2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS) [Preprint].' doi:10.1109/pmaps.2016.7764150.

Abdelfatah, M., EL-Shimy, M. and Ismail, H.M. (2011) 'Reliability Analysis of 220kV Power Transformers in Egypt', Ain Shams Engineering Journal, 2(3–4), pp. 183–194. doi:10.1016/j.asej.2011.09.005.

Abu-Elanien, A.E., Salama, M.M. and Ibrahim, M. (2011) 'Determination of transformer health condition using artificial neural networks', 2011 International Symposium on Innovations in Intelligent Systems and Applications [Preprint]. doi:10.1109/inista.2011.5946173.

Abu-Elanien, A.E., Salama, M.M. and Ibrahim, M. (2012) 'Calculation of a health index for oil-immersed transformers rated under 69 kv using fuzzy logic', IEEE Transactions on Power Delivery, 27(4), pp. 2029–2036. doi:10.1109/tpwrd.2012.2205165.

Abu-Elanien, A.E.B. and Salama, M.M.A. (2010) 'Asset management techniques for transformers', Electric Power Systems Research, 80(4), pp. 456–464. doi:10.1016/j.epsr.2009.10.008.

Ahmad et al. (2022) 'Artificial Intelligence-Based Approach for Prediction of Power Transformer Health Index', 'International Conference on Power, Energy and Innovations (ICPEI), Pattaya Chonburi, Thailand, 2022, pp. 1-4, doi: 10.1109/ICPEI55293.2022.9986545'.

AJ, C. et al. (2018) 'Causes of transformer failures and diagnostic methods – A Review', Renewable and Sustainable Energy Reviews, 82, pp. 1442–1456. doi:10.1016/j.rser.2017.05.165.

Alarie, S. at al. (2019) 'Development of an Asset Management Decision Support Tool for Hydro-Quebec TransEnergie'. CIGRE Canada Conference1.Available at: https://www.researchgate.net/profile/Stephane-Alarie/publication/340133438_Development_of_an_Asset_Management_Decision_Support_Tool_for_Hydro-Quebec_TransEnergie/links/5e7a733b4585152fc0ec7735/Development-of-an-Asset-Management-Decision-Support-Tool-for-Hydro-Quebec-TransEnergie.pdf

Allard, L. et al. (2010) 'Paper A2-202 - Vibro-Acoustic Diagnostic Contributing to an Optimised Maintenance Strategy', in CIGRÉ 2010 Session.

Almeida, N., Trindade, M., Komljenovic, D. and Finger, M. (2022) 'A conceptual construct on value for infrastructure asset management', Utilities Policy, 75, p.101354.

Alqudsi, A. and El-Hag, A. (2019) 'Application of Machine Learning in Transformer Health Index Prediction', Energies, 12(14)', p. 2694. doi:10.3390/en12142694.

Al-Raqqad, H. K., Al-Bourini, E. S., Al Talahin, F. M., and Aranki, R. M. E. (2017) 'The Impact of School Bullying on Students' Academic Achievement from Teachers Point of View', International Education Studies, 10, 44-50. https://doi.org/10.5539/ies.v10n6p44

Alstom Grid, Power Transformers - Expertise, vol. Volume 2, p. Section 1.6. *Power transformers vol.2 expertise pdf* (2011) *Scribd*. Available at: https://www.scribd.com/document/397399413/Power-Transformers-Vol-2-Expertise-pdf (Accessed: 03 August 2019).

Alsyouf, I., Alsuwaidi, M., Hamdan, S. and Shamsuzzaman, M. (2021) 'Impact of ISO 55000 on organisational performance: evidence from certified UAE firms', Total Quality Management & Business Excellence, 32(1-2), pp.134-152.

Altman, D.G., Bland, J.M. (1994) "Diagnostic tests 1: sensitivity and specificity," British Medical Journal, vol 308, 1552),

Altman, D.G., Bland, J.M. (1994) "Diagnostic tests 2: predictive values," British Medical Journal, vol 309, 102), Velez, D.R., et al. (2008) "A balanced

Alvarez Quiñones, L.I., Lozano-Moncada, C.A. and Bravo Montenegro, D.A. (2022) 'Machine learning for predictive maintenance scheduling of distribution transformers', 'Journal of Quality in Maintenance Engineering, 29(1)', pp. 188–202. doi:10.1108/jqme-06-2021-0052.

Amruthnath, N. and Gupta, T. (2018) 'Fault class prediction in unsupervised learning using model-based clustering approach'. In 2018 International Conference on Information and Computer Technologies (ICICT) (pp. 5-12). IEEE.

Arbabshirani, M.R. et al. (2017) 'Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls', NeuroImage, 145, pp. 137–165. doi:10.1016/j.neuroimage.2016.02.079.

Arias Velásquez, R.M. and Mejía Lara, J.V. (2018) 'Bushing failure in power transformers and the influence of moisture with the spectroscopy test', Engineering Failure Analysis, 94, pp. 300–312. doi:10.1016/j.engfailanal.2018.08.006.

Arias Velásquez, R.M. and Mejia Lara, J.V. (2018) 'Circuit breakers 500 kv degradation in substation reactors caused by Inductive Current', Engineering Failure Analysis, 90, pp. 64–81. doi:10.1016/j.engfailanal.2018.03.002.

Arias Velásquez, R.M. and Mejía Lara, J.V. (2018) 'Current transformer failure caused by electric field associated to circuit breaker and pollution in 500 kv substations', Engineering Failure Analysis, 92, pp. 163–181. doi:10.1016/j.engfailanal.2018.05.017.

Arshad, M., Islam, S. and Khaliq, A. (2014) 'Fuzzy logic approach in power transformers management and decision making', *IEEE Transactions on Dielectrics and Electrical Insulation*, 21(5), pp. 2343–2354. doi:10.1109/tdei.2014.003859.

Arshad, W. et al. (2021) 'Multi-class text classification: Model Comparison and selection', 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) [Preprint]. Available at: https://doi.org/10.1109/icecce52056.2021.9514108.

Arumugam, S. (2021) 'Failure diagnosis and root-cause analysis of in-service and defective distribution transformers', Energies, 14(16), p.4997.

Ashkezari, A.D. et al. (2012) 'Multivariate analysis for correlations among different transformer oil parameters to determine transformer health index', 2012 IEEE Power and Energy Society General Meeting [Preprint]. doi:10.1109/pesgm.2012.6345231.

Ashkezari, A.D. et al. (2013) 'Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers', IEEE Transactions on Dielectrics and Electrical Insulation, 20(3), pp. 965–973. doi:10.1109/tdei.2013.6518966.

ASTM D1816 - 12 (2019) Standard Test Method For Dielectric Breakdown Voltage of Insulating Liquids Using VDE Electrodes. Available at: https://tuxdoc.com/download/astm-d1816-12-2019-standard-test-method-for-dielectric-breakdown-voltage-of-insu_pdf.

ASTM, A938-07(2013), Standard Test Method for Torsion Testing of Wire, ASTM International, West Conshohocken, PA, 2013.

ASTM, D1275 (2006) 'Test method for corrosive sulphur in electrical insulating oils'.

ASTM, D3487-09 (2016) 'Standard Specification for Mineral Insulating Oil Used in Electrical Apparatus'.

ASTM, D6871-03 (2008) 'Standard Specification for Natural (Vegetable Oil) Ester Fluids Used in Electrical Apparatus'.

Atanasova-Hoehlein, I. et al. (2013) 'Assessment of Oil Corrosivity', in CIGRÉ A2-C4 Joint Colloquium. Zurich.

Atmathew (2015) 'Evaluating Logistic Regression Models': R-bloggers, R. Available at: https://www.r-bloggers.com/2015/08/evaluating-logistic-regression-models/ (Accessed: April 4, 2022).

Azmi, A. et al. (2017) 'Evolution of transformer health index in the form of the mathematical equation', Renewable and Sustainable Energy Reviews, 76, pp. 687–700. doi:10.1016/j.rser.2017.03.094.

Basuki, A. and Suwarno (2018) 'Online dissolved gas analysis of Power Transformers based on Decision Tree Model', 2018 Conference on Power Engineering and Renewable Energy (ICPERE) [Preprint]. doi:10.1109/icpere.2018.8739761.

Baumann, N. et al. (2022) 'Dynamic data management for continuous retraining', Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings [Preprint]. doi:10.1145/3550356.3561568.

Benhmed, K. et al. (2018) 'Feature selection for effective health index diagnoses of Power Transformers', IEEE Transactions on Power Delivery, 33(6), pp. 3223–3226. doi:10.1109/tpwrd.2017.2762920.

Bertling, L., Allan, R. and Eriksson, R. (2005) 'A reliability-centred asset maintenance method for assessing the impact of maintenance in power distribution systems', IEEE Transactions on Power Systems, 20(1), pp. 75–82. doi:10.1109/tpwrs.2004.840433.

Bertrand, H. (2010) 'Advanced monitoring and modelling methods for power transformer asset management', European Journal of Electrical Engineering, 13(5–6), pp. 591–620. doi:10.3166/ejee.13.591-620.

Besharatifard, H. et al. (2022) 'Acoustic based localization of partial discharge inside oil-filled transformers', IEEE Access, 10, pp. 55288–55297. doi:10.1109/access.2022.3177603.

Bigdeli, M. and Aghajanloo, J. (2016) 'Condition assessment of transformer insulation using dielectric frequency response analysis by artificial bee colony algorithm', Archives of Electrical Engineering, 65(1), pp. 45–57. doi:10.1515/aee-2016-0004.

Bochenski, B.M. (2012) 'An advanced analysis of a root cause of a power transformer failure', 2012 IEEE Electrical Power and Energy Conference [Preprint]. doi:10.1109/epec.2012.6474948.

Bohatyrewicz, P., Płowucha, J. and Subocz, J. (2019) 'Condition assessment of power transformers based on health index value', Applied Sciences, 9(22), p. 4877. doi:10.3390/app9224877.

Boulenouar, Julia (2014). Infrastructure asset management: a key building block for sustaining rural water services. Loughborough University. Conference contribution. https://hdl.handle.net/2134/31026

Bouzeghoub, M. (2004) 'A framework for analysis of data freshness', Proceedings of the 2004 international workshop on Information quality in information systems [Preprint]. doi:10.1145/1012453.1012464.

Bureau of Reclamation, U.S.D. of the I. (2005) 'Transformers: Basics, maintenance, and Diagnostics - Bureau of Reclamation', Available at: https://usbr.gov/tsc/techreferences/mands/mands-pdfs/Trnsfrmr.pdf (Accessed: 12 Mar 2019).

Canadian Electricity Association (CEA), Deloitte's Asset Management team. (2015). Asset Health Indices utility industry necessity [White paper]. http://www.pubmanitoba.ca/exhibits/mh_gra_2015/coalition-10-3.pdf

Caret: Classification and regression training - the comprehensive R (2023), https://cran.r-project.org/web/packages/caret/caret.pdf

Chakravorti, S., Dey, D. and Chatterjee, B. (2013) 'Recent trends in the condition monitoring of Transformers', Power Systems [Preprint]. doi:10.1007/978-1-4471-5550-8.

Chantola et al. (2018) 'Integrated Fuzzy Logic Approach for calculation of health index of power transformer', 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) [Preprint]. doi:10.1109/icicct.2018.8473316.

Chen, W. et al. (2016) 'Analysis of furfural dissolved in transformer oil based on confocal laser Raman spectroscopy', IEEE Transactions on Dielectrics and Electrical Insulation, 23(2), pp. 915–921. doi:10.1109/tdei.2015.005434.

Christina et al. (2018) 'Causes of transformer failures and diagnostic methods–A review'. Renewable and Sustainable Energy Reviews, 82, pp.1442-1456.

CIGRE A2_101 (2014) 'Use of Health Index and Reliability Data for Transformer Condition Assessment and Fleet Ranking'. Hydro-Québec Canada.

CIGRE Technical Brochure 227 (2003) 'Life management techniques for power transformer' (WG A2.18). Available at: https://www.e-cigre.org/publications/detail/227-life-management-techniques-for-power-transformers.html

CIGRE Technical Brochure 755 (2019) 'Transformer Bushing Reliability'.

CIGRE Technical Brochure 761 (2019) 'Power transformers and reactors, condition assessment of power transformers'.

CIGRE Technical Brochures 642 (2015) 'Transformer Reliability Survey' CIGRE Working Group 05 (WG A2.37). Available at: https://e-cigre.org/publication/642-transformer-reliability-survey (Accessed: 22 July 2023).

CIGRE Working Group 05. (1983). An international survey on failures in large power transformers in service. Electra, (88).

CIGRE Working Group A2.18 (2003) 'Guide for Life management techniques for power transformer', International Council on Large Electric Systems: Paris, France, Brochure 227.

CIGRE, "Electra Reliability Survey," 1988.

CIGRE, Technical Brochure 248 (2004) 'Economics of transformer management'.

CIGRE, Technical Brochure 323 (2007) 'Ageing of cellulose in mineral-oil insulated transformers'.

CIGRE, Technical Brochure 343 (2008) - Recommendations for Condition Monitoring and Condition Assessment Facilities for Transformers.

CIGRE, Technical Brochure 378 (2009) ‘Copper Sulphide in Transformer Insulation’.

CIGRE, Technical Brochure 393 (2010) ‘Thermal Performance’.

CIGRE, Technical Brochure 413 (2010) ‘Insulating Oil Regeneration and Dehalogenation’.

CIGRE, Technical Brochure 420 (2010) - Generic Guidelines for Life Time Condition Assessment of HV Assets and Related Knowledge Rules”.

CIGRE, Technical Brochure 422 (2010) Working Group C1.16. Transmission asset risk management.

CIGRE, Technical Brochure 436 (2010)- Experiences in Service with New Insulating Liquids.

CIGRE, Technical Brochure 443 (2010) ‘DGA in Non-Mineral Oils and Load Tap Changers and Improved DGA Diagnosis Criteria’.

CIGRE, Technical Brochure 445 (2011) ‘Guide for Transformer Maintenance’.

CIGRE, Technical Brochure 453 (2011) ‘Glossary of Terms Used in the Electric Industry’.

CIGRE, Technical Brochure 494 (2012) ‘Ageing of cellulose in mineral-oil insulated transformers’.

CIGRE, Technical Brochure 576 (2014) ‘Guide for IT Strategies for Asset Management of Substations - General Principles’.

CIGRE, Technical Brochure 625 (2015) ‘Copper Sulphide Long Term Mitigation and Risk Assessment’.

CIGRE, Technical Brochure 630 (2015) ‘Guide for Transformer Intelligent Condition Monitoring (TICM) Systems’.

CIGRE, Technical Brochure 659 (2016) - Transformer thermal modelling.

CIGRE, Technical Brochure 673 (2016) ‘Guide on Transformer Transportation’.

CIGRE, Technical Brochure 706 (2007) - Guidelines for the use of Statistics and Statistical Tools on Life Data”.

CIGRE, Technical Brochure 735 (2018) - Transformer Post-Mortem Analysis.

CIGRE, WG A2.30 TB 349 (2008) ‘Moisture Equilibrium and Moisture Migration within Transformer Insulation Systems’.

Csanyi, E. (2020) ‘Better understanding of transformer failures and maintenance necessity’, EEP - Electrical Engineering Portal. Available at: https://electrical-engineering-portal.com/transformer-failures-maintenance (Accessed: 24 Aug 2020).

Da Silva et al. (2021) ‘Enhanced health index for power transformers diagnosis’, Engineering Failure Analysis, 126, p.105427.

Davidenko, I.V. and Kuzina, T.S. (2017) ‘Analysis of the Modern Methods of the Power Transformers Health Index Calculation’, 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) [Preprint]. doi:10.1109/eiconrus.2017.7910856.

Denos et al. (2010) ‘Ageing diagnosis by chemical markers influence of core-type and shell-type technology’. In CIGRE, Paris, France. Available at https://cigreindia.org/CIGRE%20Lib/CIGRE%20Session%202010%20paper/A2_201_2010.pdf

Densley, R.J. and Gupta, B.K. (2001) ‘Effect of temperature on sensitivity of diagnostic tests on oil-impregnated paper insulation’, Proceedings: Electrical Insulation Conference and Electrical Manufacturing and Coil Winding Conference (Cat. No.01CH37264) [Preprint]. doi:10.1109/eeic.2001.965765.

Diaby, M. and Stiegemeier, C. (2017) ‘Intelligent transformers keep the plant running and reduce maintenance expenditures’, ABB Technical Paper,

https://https://search.abb.com/library/Download.aspx?DocumentID=9AKK107045A4678&DocumentPartId=

Dias and Cheim (2021) 'Transformer assessment using health index – Part II: Sensitivity analysis and critical discussion', Transformers Magazine, 8(1), p.58.

Difference between core form and Shell Form XFMR (2021) AllumiaX. Available at: https://www.allumiax.com/difference-between-core-form-and-shell-form-power-transformers-by-generalpac (Accessed: 29 July 2021).

Dinov, I.D. (2018) 'Probabilistic learning: Classification using naive bayes', *Data Science and Predictive Analytics*, pp. 289–305. doi:10.1007/978-3-319-72347-1_8.

Diop, I., Abdul-Nour, G. and Komljenovic, D. (2021) 'Overview of strategic approach to asset management and decision-making', International Journal of Engineering Research & Technology, 10, pp.64-89.

DNV. (2013). Utility asset management | condition-based maintenance | Cascade. [online] Available at: <https://www.dnv.com/services/utility-asset-management-and-condition-based-maintenance-cascade-5156> [Accessed 21 February 2021].

Dominelli, N. (2004) 'Equipment Health Rating of Power Transformers', Conference Record of the 2004 IEEE International Symposium on Electrical Insulation [Preprint]. doi:10.1109/elinsl.2004.1380501.

Dong, M. et al. (2019) 'A Novel Maintenance Decision Making Model of Power Transformers Base on Reliability and Economy Assessment', IEEE Access, 7, pp. 28778–28790. doi:10.1109/access.2019.2897606.

Dorji, T. et al. (2022) 'Fuzzy logic approach for power transformer condition assessment using health index', 2022 International Conference on Power, Energy, and Innovations (ICPEI) [Preprint]. doi:10.1109/icpei55293.2022.9986526.

Duarte, E. et al. (2010) 'A Practical Approach to Condition and Risk Base Power Transformer Asset Replacement', 2010 IEEE International Symposium on Electrical Insulation [Preprint]. doi:10.1109/elinsl.2010.5549580.

Dumitran, L.M. et al. (2014) 'Method for lifetime estimation of power transformer mineral oil', Fuel, 117, pp. 756–762. doi:10.1016/j.fuel.2013.10.002.

Duran Tovar, I.C. et al. (2015) 'Comparison of maintenance scheme effects on power transformer reliability performance', Ingenieríae Investigation, 35(1Sup), pp. 73–81. doi:10.15446/ing.investig.v35n1sup.53435.

Duval, M. (2002) 'A review of faults detectable by gas-in-oil analysis in transformers', IEEE Electrical Insulation Magazine, 18(3), pp. 8–17. doi:10.1109/mei.2002.1014963.

Duval, M., (2016) 'Use of pentagons and triangles for the interpretation of DGA in electrical equipment', In Proceedings of the TechCon North America Conference.

E. S. W Lampe (1976) 'The Oxygen-Free Transformer, Reducing Ageing by Continuous Degassing', CIGRE Paper 12.05, 1976.

EA Technology (2018) Condition Based Risk Management (CBRM) [online]. Available at: cbrm-brochure.pdf (eatechnology.com).

Ekanayake. C (2003) 'Application of dielectric spectroscopy for estimating moisture content in power transformers', Dept. Elect. Power Eng., Chalmers Univ. Technol., Gothenburg, Sweden, Tech. Rep. 465 L for Tech. Lic. Degree.

Eliyan, F.F. (2021) 'Corrosivity of hydrocarbon flows in oil pipelines', Proceedings of 1st Corrosion and Materials Degradation Web Conference [Preprint]. doi:10.3390/cmdwc2021-09923.

Faraway, J.J. (2002) 'Practical regression and ANOVA using R' (Vol. 168). Bath: University of Bath.

Faryadi, Q. (2019) 'PhD thesis Writing process: A systematic approach—how to write your methodology, results and conclusion', Creative Education, 10(04), pp. 766–783. doi:10.4236/ce.2019.104057.

Fauzi, N.A. et al. (2020) 'Fault prediction for power transformer using optical spectrum of transformer oil and data mining analysis', IEEE Access, 8, pp. 136374–136381. doi:10.1109/access.2020.3011504.

Fengjun Yi et al. (2015) 'Simulation of HVDC transmission system failure rate bathtub curve based on Weibull distribution', 2015 IEEE 2nd International Future Energy Electronics Conference (IFEEC) [Preprint]. doi:10.1109/ifeec.2015.7361599.

Flint, N. et al. (2017) 'An ecosystem health index for a large and Variable River Basin: Methodology, challenges and continuous improvement in Queensland's fitzroy basin', Ecological Indicators, 73, pp. 626–636. doi:10.1016/j.ecolind.2016.10.007.

Fofana, I., Bouaïcha, A. and Farzaneh, M. (2010) 'Characterization of aging transformer oil-pressboard insulation using some modern diagnostic techniques', European Transactions on Electrical Power, 21(1), pp. 1110–1127. doi:10.1002/etep.499.

Foros, J. and Istad, M. (2020) 'Health index, risk and remaining lifetime estimation of Power Transformers', IEEE Transactions on Power Delivery, 35(6), pp. 2612–2620. doi:10.1109/tpwrd.2020.2972976.

Francheck, M.A. and Woodcock, D.J. (2009) 'Life-Cycle Considerations of Loading Transformers Above Nameplate Rating'. Weidmann Technical Services.

G. I. Webb, J. Boughton, Z. Wang, (2005) ' Not So Naive Bayes: Aggregating One-Dependence Estimators. Machine Learning', Ed. Springer, 58 (1): 5–24. doi:10.1007/s10994-005-4258-6.

Gafvert U, Frimpong G, Fuhr J. (1998) 'Modelling of dielectric measurements on power transformers'. International Council on Large Electric Systems (CIGRE), Paris (France), 15–103.

Gaha, M. et al. (2021) 'Global methodology for electrical utilities maintenance assessment based on risk-informed decision making', Sustainability, 13(16), p. 9091. doi:10.3390/su13169091.

Galkina, O. and Yachenko, V. (2014) 'Application of iterative software development methodologies for reducing service quality gaps', Proceedings of the 2014 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference [Preprint]. doi:10.1109/elconrusnw.2014.6839195.

Gao, Z., Saxen, H. and Gao, C. (2013) 'Guest editorial: Special section on data-driven approaches for complex industrial systems', IEEE Transactions on Industrial Informatics, 9(4), pp. 2210–2212. doi:10.1109/tii.2013.2281002.

Garcia, B., Cespedes, A. and Garcia, D. (2018) 'Moisture Analysis for Power Transformers', Power Transformer Condition Monitoring and Diagnosis, pp. 87–123. doi:10.1049/pbpo104e_ch3.

Garton, C.G. (1941) 'Dielectric loss in thin films of insulating liquids', Journal of the Institution of Electrical Engineers - Part III: Communication Engineering, including the Proceedings of the Wireless Section of the Institution, 88(1), pp. 23–40. doi:10.1049/ji-3-1.1941.0006.

Ghasemi et al. (2023) 'Equipment failure rate in electric power distribution networks: An overview of concepts, estimation, and modelling methods', Engineering Failure Analysis, 145, p.107034.

Ghoneim, S.S. and Taha, I.B. (2020) 'Comparative study of full and reduced feature scenarios for health index computation of Power Transformers', IEEE Access, 8, pp. 181326–181339, doi:10.1109/access.2020.3028689.

Gilbert, R. et al. (2008) 'Kinetics of the production of chain-end groups and methanol from the depolymerization of cellulose during the ageing of paper/oil systems.

part 1: Standard Wood Kraft Insulation', Cellulose, 16(2), pp. 327–338. doi:10.1007/s10570-008-9261-1.

Gilbert, R. et al. (2009) 'Kinetics of the production of chain-end groups and methanol from the depolymerization of cellulose during the ageing of paper/oil systems. Part 2: Thermally-upgraded insulating papers', Cellulose, 17(2), pp. 253–269. doi:10.1007/s10570-009-9365-2.

Gjorgiev and Sansavini (2022) 'Identifying and assessing power system vulnerabilities to transmission asset outages via cascading failure analysis', Reliability Engineering & System Safety, 217, p.108085.

Glover, E.J. et al. (2001) 'Improving category specific web search by learning query modifications', Proceedings 2001 Symposium on Applications and the Internet [Preprint]. doi:10.1109/saint.2001.905165.

Glover, E.J. et al. (2002) 'Using web structure for classifying and describing web pages', Proceedings of the 11th international conference on World Wide Web [Preprint]. doi:10.1145/511446.511520.

Gorgan, B. at al. (2010). 'Calculation of power transformers health indexes'. Annals of the University of Craiova, Electrical Engineering Series, 34, pp.13-18.

Gorgan, B. et al. (2012) 'Calculation of the remaining lifetime of Power Transformers Paper Insulation', 2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) [Preprint]. doi:10.1109/optim.2012.6231792.

Gouda, O.E. and El Dein, A.Z. (2019) 'Prediction of aged transformer oil and paper insulation', Electric Power Components and Systems, 47(4–5), pp. 406–419. doi:10.1080/15325008.2019.1604848.

Gouda, O.E. et al. (2021) 'Enhancing the diagnostic accuracy of DGA techniques based on IEC-TC10 and related databases', IEEE Access, 9, pp. 118031–118041. doi:10.1109/access.2021.3107332.

Graham, J.S. (2013) 'High-voltage bushings', High Voltage Engineering Testing, pp. 467–494. doi:10.1049/pbpo066e_ch12.

Grussing, M.N. (2014) 'Life cycle asset management methodologies for buildings', Journal of Infrastructure Systems, 20(1), p.04013007.

Gupta, S. et al. (2021) 'Data-driven preprocessing techniques for early diagnosis of diabetes, heart and liver diseases', 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT) [Preprint]. doi:10.1109/icecct52121.2021.9616835.

Gurung, G., Shah, R. and Jaiswal, D.P. (2020) 'Software development life cycle models A comparative study', International Journal of Scientific Research in Computer Science, Engineering and Information Technology, pp. 30–37. doi:10.32628/cseit206410.

Gutman, I., Ahlholm, C., Åkesson, U., Holmberg, A., Wu, D. and Jonsson, L.,(2013) 'Long-term service experience and inspection results of HV equipment made of silicone rubber insulators'. In CIGRE Symposium (Vol. 412).

Gyorodi, C. et al. (2010) 'Full-text search engine using mysql', International Journal of Computers Communications; Control, 5(5), p. 735. doi:10.15837/ijccc.2010.5.2233.

Győrödi, R. et al. (2016). 'A System of Categorization and Classification Based on Certain Criteria'. In 6th International Conference on Computers Communications and Control (ICCCC) (pp. 156). Department of Computer Science and Information Technology, University of Oradea.

Haema, J. and Phadungthin, R. (2012) 'Condition Assessment of the Health Index for Power Transformer', 2012 Power Engineering and Automation Conference [Preprint]. doi:10.1109/peam.2012.6612413.

Hashemnia, N. et al. (2021) 'Using machine learning to predict and avoid malfunctions: A revolutionary concept for condition-based Asset Performance

Management (APM)', 2021 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia) [Preprint]. doi:10.1109/isgtasia49270.2021.9715635.

Hastings, N.A.J. (2021) 'Asset Management Information Systems. In: Physical Asset Management, Springer, Cham. https://doi.org/10.1007/978-3-030-62836-9_13

Hastings, N.A.J. (2021) 'Introduction to asset management. Physical Asset Management: With an Introduction to the ISO 55000 Series of Standards,' pp.1-24.

Heathcote, M. J. (2007). *The J and P Transformer Book* (13th ed.).

Heathcote, M.J. and Arnold, R. (2016) J and P Transformer Book. Amsterdam: Newnes.

Heathcote, M.J., (1998). The J and P Transformer Book.

Henke, N. et al. (2016) The age of analytics: Competing in a data-driven world, McKinsey Company. Available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world (Accessed: 13 Aug 2020).

Hernanda, I.G. et al. (2014) 'Application of Health Index Method for Transformer Condition Assessment, TENCON 2014 - 2014 IEEE Region 10 Conference' [Preprint]. doi:10.1109/Tencon.2014.7022433.

Hernandez, M.D. and Labib, A. (2017) 'Selecting a Condition Monitoring System for Enhancing Effectiveness of Power Transformer Maintenance', Journal of Quality in Maintenance Engineering, 23(4), pp. 400–414. doi:10.1108/jqme-07-2015-0027.

Heywood, R.J. and McGrail, T. (2015) 'Clarifying the link between data, diagnosis and asset health indices', Asset Management Conference 2015 [Preprint]. doi:10.1049/cp.2015.1748.

Heywood, R.J. and McGrail, T. (2016). Generating Asset Health Indices Which Are Useful and Auditable. [online] Available at:

https://www.cce.umn.edu/documents/CPE-Conferences/MIPSYCON-Papers/2016/GeneratingAssetHealthIndiceswhicharebothUsefulandAuditable.pdf [Accessed 31 Jul. 2023].

Hjartarson, T. and Otal, S. (2006) 'Predicting future asset condition based on current health index and Maintenance Level', ESMO 2006 - 2006 IEEE 11th International Conference on Transmission; Distribution Construction, Operation and Live-Line Maintenance [Preprint]. doi:10.1109/tdcllm.2006.340747.

Houhanessian, V.D. (1998) Measurement and analysis of dielectric response in oil-paper insulation systems. PhD thesis. Swiss Federal Institute of Technology, ETH No. 12832.

Huang, Y.C. and Huang, C.M. (2002) 'Evolving Wavelet Networks for power transformer condition monitoring', IEEE Power Engineering Review, 22(2), pp. 63–63. doi:10.1109/mper.2002.4312016.

Hughes, D. (2003) 'The use of health indices to determine end of life and estimate remnant life for distribution assets'. In Proceedings of the 17th International Conference on Electricity Distribution (CIRED) (Vol. 5, No. 1, pp. 5-15).

Hughes, D. et al. (2006) 'Condition based risk management (CBRM) — enabling asset condition information to be central to corporate decision making', Engineering Asset Management, pp. 1212–1217. doi:10.1007/978-1-84628-814-2_133.

Ibrahim, K. et al. (2016) 'Selection criteria for oil transformer measurements to calculate the health index', IEEE Transactions on Dielectrics and Electrical Insulation, 23(6), pp. 3397–3404. doi:10.1109/tdei.2016.006058.

IEC 60050-191 International electrotechnical vocabulary Chapter 191 -Dependability and quality of service, IEC Standard, 1990.

IEC 60076-5(2006) 'Ability to withstand short circuit'.

IEC 60076-7(2005) 'Loading guide for oil-immersed transformers'.

IEC 60076-7(2018), Power transformers - Part 7: Loading guide for mineral-oil-immersed power transformers, https://webstore.iec.ch/publication/34351

IEC 60475 (2011) 'method of sampling insulating liquids (no date a) iTeh Standards'. Available at: https://standards.iteh.ai/catalog/standards/iec/55db517a-b9e2-46dd-a28a-ac2172b4414d/iec-60475-2011 (Accessed: 01 August 2021).

IEC 60505 (2011) 'Evaluation and qualification of electrical insulation systems', Released:2011-07-11.

IEC 60505(2004) 'Evaluation and qualification of electrical insulation systems'. 3rd ed. Geneva: International Electrotechnical Commission.

IEC 61198 (1993) Mineral insulating oils - Methods for the determination of 2-furfural and related compounds, IEC [Online].

IEC 62535 (2008) 'Insulating liquids – Test method for detection of potentially corrosive sulphur in used and unused insulating oil'. Available at: https://webstore.iec.ch/publication/7167 (Accessed: 23/Feb/2020)

IEC Std. Publication 567 (2019) 'Guide for the Sampling of Gases and of Oil-Filled Electrical Equipment and for the Analysis of Free and Dissolved Gases'.

IEC White Paper, 'Strategic asset management of power networks, 2016.Available at: https://www.iec.ch/basecamp/strategic-asset-management-power-networks

IEC, 60076-1 (2011) Power transformers - Part 1: General.

IEC, 60076-14 (2013) 'Liquid-immersed power transformers using high-temperature insulation materials'.

IEC, 60076-2 (2011) 'Power transformers - Part 2: Temperature rise for liquid-immersed transformers'.

IEC, 60076-7 (2018) 'Loading guide for mineral-oil-immersed power transformers'.

IEC, 60099-4 (2014) ‘Surge arresters - Part 4: Metal-oxide surge arresters without gaps for a.c. Systems’.

IEC, 60099-5 (2013) ‘Surge arresters - Part 5: Selection and application recommendations’.

IEC, 60137(2008) - Insulated bushings for alternating voltages above 1000 V.

IEC, 60156 (2018) ‘Insulating liquids - Determination of the breakdown voltage at power frequency - Test method’.

IEC, 60214-1 (2014) ‘Tap-changer- Part 1: Performance requirements and test methods’.

IEC, 60214-2 (2018) ‘Tap-changer- Part 2: Application guide’.

IEC, 60247 (2004) ‘Insulating liquids - Measurement of relative permittivity, dielectric dissipation factor (tan d) and d.c. Resistivity’.

IEC, 60270 (2000) ‘High-voltage test techniques - Partial discharge measurements’.

IEC, 60296 (2012) ‘Fluids for electrotechnical applications — Unused mineral insulating oils for transformers and switchgear’, Edition 4.0 ed.

IEC, 60422 (2013) ‘Mineral insulating oils in electrical equipment – Supervision and maintenance guidance’, Edition 4.0 ed.

IEC, 60450 (2004) ‘Measurement of the average viscometric degree of polymerization of new and aged cellulosic electrically insulating materials’.

IEC, 60599(2007) ‘Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis’.

IEC, 60599(2015) ‘Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis’.

IEC, 60599(2022) ‘Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis’, released:2022-05-25.

IEC, 60814 (1997) ‘Insulating liquids - Oil-impregnated paper and pressboard - Determination of water by automatic coulometric Karl Fischer titration’.

IEC, 61099 (2010) Insulating liquids - Specifications for unused synthetic organic esters for electrical purposes, 2010.

IEC, 61203 (1992) Synthetic organic esters for electrical purposes - Guide for maintenance of transformer esters in equipment, Edition 1.0 ed., 1992-12.

IEC, 61649 (2008) ‘Weibull Analysis’.

IEC, 62770 (2013) ‘Fluids for electrotechnical applications – Unused natural esters for transformers and similar electrical equipment’, Edition 1.0 ed.

IEC, 62975 (2021) ‘Natural esters – Guidelines for maintenance and use in electrical equipment’.

IEC, TECHNICAL BROCHURES 662 (2016) Guidelines for partial discharge detection using conventional (IEC 60270) and unconventional methods.

IEC, Technical Report 61464(1998) "Guide for the interpretation of dissolved gas analysis (DGA) in bushings where oil is the impregnating medium of the main insulation, (generally paper)",” ISBN: 2- 8318- 4459-2.

IEEE C57.117-1986: IEEE guide on reporting failure data for power transformers and shunt reactors on electric utility power systems, IEEE, “ 1986.

IEEE Guide, C57.104 (2019) ‘Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers, released:01.11.2019, ISBN:978-1-5044-5973-0.

IEEE Guide, C57.106 (2015) ‘Guide for Acceptance and Maintenance of Insulating Mineral Oil in Electrical Equipment, released:23.03.2016, ISBN:978-1-5044-0097-8.

IEEE Std 62-Part 1 (1995). Guide for Diagnostic Field Testing of Electric Power Apparatus - Part 1: Oil-Filled Power Transformers, Regulators, and Reactors.

[online] Available at: https://standards.ieee.org/standard/62-1995.html [Accessed 27 Jul. 2023].

IEEE Std C57.139 (2015) 'IEEE Guide for Dissolved Gas Analysis in Transformer Load Tap Changers'.

IEEE Std, C57.104 (2008) 'Guide for Interpretation of Dissolved Gases Generated in Insulating Oil Immersed Power Transformer Insulating Oil'.

IEEE Std, C57.106 (2006) 'Guide for Acceptance and Maintenance of Insulating Oil in Equipment'.

IEEE Std, C57.106 (2015) 'Guide for Acceptance and Maintenance of Insulating Oil in Equipment'.

IEEE Std, C57.12.00 (2015) 'General Requirements for Liquid-Immersed Distribution', Power and Regulating Transformers.

IEEE Std, C57.121 (1998) 'Guide for Acceptance and Maintenance of Less Flammable Hydrocarbon Fluid in Transformers'. Available at: https://ieeexplore.ieee.org/document/7862345/ (Accessed: 01 August 2023).

IEEE Std, C57.140 (2017) 'Guide for Evaluation and Reconditioning of Liquid Immersed Power Transformers'.

IEEE Std, C57.143 (2012) 'Guide for Application for Monitoring Equipment to Liquid-Immersed Transformers and Components'.

IEEE Std, C57.146 (2018) 'Guide for the Interpretation of Gases Generated in Silicone-Immersed Transformers'.

IEEE Std, C57.147 (2008) 'Guide for Acceptance and Maintenance of Natural Ester Fluids in Transformers'.

IEEE Std, C57.149 (2012) 'Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers'.

IEEE Std, C57.154 (2012) 'Design, Testing, and Application of Liquid-Immersed Distribution, Power, and Regulating Transformers Using High-Temperature Insulation Systems and Operating at Elevated Temperatures'.

IEEE Std, C57.155 (2014) 'Guide for Interpretation of Gases Generated in Natural Ester and Synthetic Ester-Immersed Transformers'.

IEEE Std, C57.19.00 (2018) 'General Requirements and Test Procedure for Power Apparatus Bushings'.

IEEE Std, C57.19.01(2017) 'Performance Characteristics and Dimensions for Power Transformer and Reactor Bushings'.

IEEE Std, C57.19.100 (2000)' IEEE Guide for Application of Power Apparatus Bushings'. New York: Institute of Electrical and Electronics Engineers.

IEEE Std, C57.637(2007) 'Guide for the Reclamation of Insulating Oil and Criteria for its use'.

IEEE Std, C57.91 (2011) 'Guide for Loading Mineral-Oil-Immersed Transformers'.

IEEE Std, IEEE C57.152 (2013) 'Guide for Diagnostic Field Testing of Fluid-Filled Power Transformers, Regulators, and Reactors'.

IEEE Std, IEEE C62.22 (2009) 'IEEE Guide for the Application of Metal-Oxide Surge Arresters for Alternating-Current Systems'.

Institute of Asset Management. (2020). Home. https://theiam.org/

Islam, M. et al. (2018) "Calculating a health index for Power Transformers using a subsystem-based GRNN approach," IEEE Transactions on Power Delivery, 33(4), pp. 1903–1912. Available at: https://doi.org/10.1109/tpwrd.2017.2770166.

Islam, M.M., Lee, G. and Hettiwatte, S.N. (2016) 'A nearest neighbour clustering approach for incipient fault diagnosis of Power Transformers', Electrical Engineering, 99(3), pp. 1109–1119. doi:10.1007/s00202-016-0481-3.

Islam, M.M., Lee, G. and Hettiwatte, S.N. (2017a) 'A review of condition monitoring techniques and diagnostic tests for lifetime estimation of Power Transformers', Electrical Engineering, 100(2), pp. 581–605. doi:10.1007/s00202-017-0532-4.

Islam, M.M., Lee, G. and Hettiwatte, S.N. (2017b) 'Application of a general regression neural network for health index calculation of Power Transformers', *International Journal of Electrical Power; Energy Systems*, 93, pp. 308–315. doi:10.1016/j.ijepes.2017.06.008.

ISO 18095 (2018), 'Condition monitoring and diagnostics of power transformers'. Available at: https://www.iso.org/obp/ui/#!iso:std:61423:en.

ISO 55000 series, Asset Management, ISO 55000:2014, BS ISO 55001:2014, and ISO 55002:2014, First edition 2014-01-15, Reference number ISO 55000:2014(E), Corrected version 2014-03-15.

Istad, M., Solvang, E., Catrinu-Renström, M., Johnsen, J. (2018). Risk monitoring function in Statnett: SAMBA WP6 report. SINTEF Energy Research1

Jagers et al., (2009) 'Evaluation of transformer reliability data based on national and utility statistics', 16th Int. Symp. High Volt, Eng Cape Town, South Africa, pp. 1–9.

Jahromi, A. et al. (2009) 'An approach to power transformer asset management using health index', IEEE Electrical Insulation Magazine, 25(2), pp. 20–34. doi:10.1109/mei.2009.4802595.

Jakob, F. et al. (2003) 'Use of gas concentration ratios to interpret LTC and OCB Dissolved Gas Data', Proceedings: Electrical Insulation Conference and Electrical Manufacturing and Coil Winding Technology Conference (Cat. No.03CH37480) [Preprint]. doi:10.1109/eicemc.2003.1247901.

Jalbert, J. and Lessard, M.-C. (2015) 'Cellulose chemical markers relationship with insulating paper post-mortem investigations', IEEE Transactions on Dielectrics and Electrical Insulation, 22(6), pp. 3550–3554. doi:10.1109/tdei.2015.005295.

Jalbert, J. et al. (2007) 'Identification of a chemical indicator of the rupture of 1,4-β-glycosidic bonds of cellulose in an oil-impregnated insulating paper system', Cellulose, 14(4), pp. 295–309. doi:10.1007/s10570-007-9124-1.

Jalbert, J. et al. (2012) 'Methanol: A novel approach to power transformer asset management', IEEE Transactions on Power Delivery, 27(2), pp. 514–520. doi:10.1109/tpwrd.2012.2185957.

Jalbert, J. et al. (2014) 'Kinetics of the production of chain-end groups and methanol from the depolymerization of cellulose during the ageing of paper/oil systems. Part 3: Extension of the study under temperature conditions over 120 °C', Cellulose, 22(1), pp. 829–848. doi:10.1007/s10570-014-0516-8.

Jalbert, J., Duchesne, S., et al. (2012) 'Robust and sensitive analysis of methanol and ethanol from cellulose degradation in mineral oils', Journal of Chromatography A, 1256, pp. 240–245. doi:10.1016/j.chroma.2012.07.069.

Jalbert, J., Lessard, M.-C. and Ryadi, M. (2013) 'Cellulose chemical markers in transformer oil insulation part 1: Temperature Correction factors', IEEE Transactions on Dielectrics and Electrical Insulation, 20(6), pp. 2287–2291. doi:10.1109/tdei.2013.6678881.

Jan et al. (2015) 'Transformer failures, causes and impact' (2015) International Conference Data Mining, Civil and Mechanical Engineering (ICDMCME'2015) Feb. 1-2, 2015, Bali (Indonesia) [Preprint]. doi:10.15242/iie.e0215039.

Jie, H. and Zheng, G. (2019) 'Calibration of torque error of permanent magnet synchronous motor base on polynomial linear regression model', IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society [Preprint]. doi:10.1109/iecon.2019.8927537.

Jin, L. et al. (2022) 'Oil-Immersed Power Transformer Condition Monitoring Methodologies': A Review', Energies, 15(9), p. 3379. doi:10.3390/en15093379.

Jucker, B. and Helmach, M. (2013) 'ABB Review. Zurich', Suiza: ABB Technology.

Jürgensen, J.H. et al. (2017) 'Health index as condition estimator for power system equipment: A critical discussion and case study', CIRED - Open Access Proceedings Journal, 2017(1), pp. 202–205. doi:10.1049/oap-cired.2017.1174.

Jürgensen, J.H.et al. (2016) 'Individual failure rates for Transformers within a population based on diagnostic measures', Electric Power Systems Research, 141, pp. 354–362. doi:10.1016/j.epsr.2016.08.015.

Kadim, E. *et al.* (2018) 'Transformers health index assessment based on neural-fuzzy network', *Energies*, 11(4), p. 710. Available at: https://doi.org/10.3390/en11040710.

Karthik, R., Sree Renga Raja, T. and Sudhakar, T. (2013) 'Deterioration of solid insulation for thermal degradation of transformer oil', Open Engineering, 3(2), pp. 226–232. doi:10.2478/s13531-012-0051-z.

Kavikondala, A. et al. (2019) 'Automated retraining of Machine Learning Models', International Journal of Innovative Technology and Exploring Engineering, 8(12), pp. 445–452. doi:10.35940/ijitee.l3322.1081219.

Khamis Al-Romaimi (2015) A053 Conference Paper 'Strategic Thinking Towards Best Selection and Implementation of Enterprise Asset Management (EAM) - Asset Management Solution for GCC Electrical Utilities', In CIGRE GCC Power 2015, 2-5 November 2015, Saudi Arabia.

Khamis Al-Romaimi (2015) A060 Conference Paper 'Risk-Based Investment Planning for GCC Electrical Utilities'. In CIGRE GCC Power 2015, 2-5 November 2015, Saudi Arabia.

Khamis Al-Romaimi et al., (2014) 'Critical Role of Failure Investigations in Effective Asset Management', CIGRE Paris, France.

Khamis Salim (2011) 'Adopting Best Practice Asset Management and Achieving the PAS 55 Standard', In GCC CIGRE Conference.

Khamis Salim (2011) ‘Adopting Best Practice Asset Management and Achieving the PAS 55 Standard’, In Government and Enterprise Asset Management Congress Conference, Dubai, UAE.

Khamis Salim (2022) ’Power Transformer Asset Management Health Index in Electrical Utilities maintenance system’, 2nd Annual Postgraduate Research Conference, Sunderland University, 15 September 2022.

Khamis Salim (2023) ‘Power Transformer Asset Management Health Index in Electrical Utilities maintenance system’, 3rd Annual Postgraduate Research Conference, Sunderland University, 14 September 2023.

Khamis Salim Al-Romaimi (2015) ‘Ensuring Capability Programmes Are Firmly Institutionalised in Your Organisation from Start to Finish’, In Government and Enterprise Asset Management Congress, Dubai, UAE, 18-21 May 2015.

Khamis Salim Al-Romaimi et al., (2014) ‘Adopting Best Practice Asset Management and Achieving the PAS 55 Standard’, In Government and Enterprise Asset Management Congress Conference, Dubai, UAE, 19-22 May 2014.

Khamis Salim and Jean-Louis Habert (2012) ‘Gas-Insulated Substation Service Continuity’, GCC Conference Paper, Muscat, Oman.

Khamis Salim et al. (2012) ‘Critical Role of Failure Investigations in Effective Asset Management’, CIGRE GCC, Muscat, Oman.

Khamis Salim, David Baglee and Derek Dixon (2023) ‘Health Index Assessment for Power Transformer Strategic Asset Management in Electrical Utilities', Int. J. of Strategic Engineering Asset Management. DOI: 10.1504/IJSEAM.2023.10059373

Khamis Salim, David Baglee and Derek Dixon (2023) ‘Power Transformer Failure Rate Investigation Using Historical Maintenance Inspection Data’, 12th IMA International Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) 4-6 July 2023 University of Nottingham, Jubilee Campus, UK.

Khamis Salim, David Baglee and Derek Dixon (2024) ‘Principles of Asset Management Health Index Development in Driving Maintenance Strategy for Electrical Utilities’, MDPI

Khamis Salim, David Baglee and Derek Dixon (2024) ‘Risk-Stratify Power Transformers Based on Health Index Condition Utilizing Asset Management’, Int. J. Risk Assessment and Management.

Khamis Salim, David Baglee and Derek Dixon (2024) ‘Strategic Asset Management Health Index for Predicting Power Transformer Health Conditions', Int. J. of Strategic Engineering Asset Management.

Khamis Salim, David Baglee and Derek Dixon (2024) ‘The Critical Role of Data Asset Management Power Transformer Health Index Decision Making for Electrical Utilities Strategic Asset Planning Lifecycle', Int. J. of Strategic Engineering Asset Management.

Khan, M., and Hussain, A. (2019) ‘Condition Assessment of Power Transformers in Service Using PD Monitoring’, International Journal of Electrical and Computer Engineering (IJECE), 9(6), 5163-5170. https://doi.org/10.11591/ijece.v9i6.pp5163-5170

Kherif, O. et al. (2021) ‘Accuracy improvement of power transformer faults diagnostic using KNN classifier with decision tree principle’, IEEE Access, 9, pp. 81693–81701. doi:10.1109/access.2021.3086135.

Kinnunen, S.-K. et al. (2016) ‘Decision making situations define data requirements in Fleet Asset Management’, Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015), pp. 357–364. doi:10.1007/978-3-319-27064-7_33.

Krishnamoorthi, R. et al. (2022) ‘A novel diabetes healthcare disease prediction framework using machine learning techniques’, Journal of Healthcare Engineering, 2022, pp. 1–10. doi:10.1155/2022/1684017.

Krüger, M., Koch, M., Kraetke, A., Rätzke, S. and Schnitzler, J. (2012) 'New Tools for the Assessment of High Voltage Bushings', in High Voltage Symposium. Stuttgart University, Germany.

Kumar Bhowmik, T. (2015) 'Naive Bayes vs logistic regression: Theory, implementation and experimental validation', INTELIGENCIA ARTIFICIAL, 18(56), pp. 14–30. Available at: https://doi.org/10.4114/ia.v18i56.1113.

Kurtz, C. (2005) 'Managing Aged Transformer, Utility develops repair/refurbish/replace strategy using innovative risk-based methodologies', Kansas City Power & Light; Gary Ford and Mark Vainberg, PowerNex Associates Inc.; Mike Lebow, Coplanar Consulting Inc.; and Barry Ward, EPRI. Available at: https://studylib.net/doc/18526899/managing-aged-transformers

Lapworth, J. and McGrail, T. (1998) 'Transformer failure modes and planned replacement', IEE Colloquium Transformer Life Management [Preprint]. doi:10.1049/ic:19981013.

Leauprasert, K. et al. (2020) 'Intelligent machine learning techniques for condition assessment of Power Transformers', 2020 International Conference on Power, Energy and Innovations (ICPEI) [Preprint]. doi:10.1109/icpei49860.2020.9431460.

Lelekakis, N., Martin, D. and Wijaya, J. (2012) 'Ageing rate of paper insulation used in power transformers part 1: Oil/paper system with low oxygen concentration', IEEE Transactions on Dielectrics and Electrical Insulation, 19(6), pp. 1999–2008. doi:10.1109/tdei.2012.6396958.

Li, S. and Dong, H. (2019) 'Transformer apparent age estimation based on Probabilistic Health index', 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia) [Preprint]. Available at: https://doi.org/10.1109/isgt-asia.2019.8881798.

Li, S. et al. (2017) 'Bayesian information fusion for Probabilistic Health index of power transformer', IET Generation, Transmission; Distribution, 12(2), pp. 279–287. doi:10.1049/iet-gtd.2017.0582.

Li, S. et al. (2018) ‘A probabilistic framework for Transformer Health Condition Assessment’, 12th International Conference on the Properties and Applications of Dielectric Materials (ICPADM) [Preprint]. Available at: https://doi.org/10.1109/icpadm.2018.8401221.

Li, S. et al. (2020) ‘Probabilistic health index-based apparent age estimation for power transformers’, IEEE Access, 8, pp. 9692–9701. Available at: https://doi.org/10.1109/access.2020.2963963.

Linhell, D. and Lundgaard, L.E. (2021) ‘A setup for measuring bubbling from paper in transformer oil under simulated sudden overload’, 2021 IEEE Electrical Insulation Conference (EIC) [Preprint]. doi:10.1109/eic49891.2021.9612373.

Lorin, P. et al. (2019) ‘Transformer Health index and probability of failure based on failure mode effects analysis (FMEA) of a Reliability Centered Maintenance (RCM) program’, e-cigre. Available at: https://e-cigre.org/publication/A2-110_2016 (Accessed: 01 August 2023).

Lundgaard et al. (2015) ‘Transformer windings ageing, diagnosis and asset management’. version 4, SINTEF. Available at: https://www.sintef.no/en/publications/publication/1060702/ (Accessed: 26 Mar 2019).

Lundgaard L. et al. (2014) ‘Managing an ageing transformer fleet; Norwegian situation, research and initiatives’, CIGRE A2 Committee.

Luo, Y. et al. (2021) ‘Deep learning-based anomaly detection in cyber-physical systems’, ACM Computing Surveys, 54(5), pp. 1–36. doi:10.1145/3453155.

Lutz. (2013) ‘Learning Python, 5th Edition Powerful Object-Oriented Programming’. O’Reilly Media, Inc.

M, H. and M.N, S. (2015) ‘A review on evaluation metrics for Data Classification Evaluations’, International Journal of Data Mining and Knowledge Management Process, 5(2), pp. 01–11. Available at: https://doi.org/10.5121/ijdkp.2015.5201.

Machado et al.(2016) “A fuzzy logic application in virtual education”, Procedia Computer Science, Vol 91, pages 19-26, doi: 10.1016/j.procs.2016.07.037

Madavan, R. and Balaraman, S. (2016) ‘Failure analysis of transformer liquid — solid insulation system under selective environmental conditions using Weibull statistics method’, Engineering Failure Analysis, 65, pp. 26–38. doi:10.1016/j.engfailanal.2016.03.017.

Madshaven, I. et al. (2022) ‘On-line direct clamping pressure monitoring of power transformer windings’, 2022 7th International Advanced Research Workshop on Transformers (ARWtr) [Preprint]. doi:10.23919/arwtr54586.2022.9959909.

Malik, H.et al. (2012) ‘Application research based on modern-technology for Transformer Health index estimation’, International Multi-Conference on Systems, Signals; Devices [Preprint]. doi:10.1109/ssd.2012.6198012.

Martin, D. et al. (2017) ‘On the development of power transformer failure models: An Australian case study’, 2017 IEEE Power; Energy Society General Meeting [Preprint]. doi:10.1109/pesgm.2017.8274571.

Martins, M.A. (2014) ‘Condition and risk assessment of Power Transformers: A general approach to calculate a health index’, Ciência; Tecnologia dos Materiais, 26(1), pp. 9–16. doi:10.1016/j.ctmat.2014.09.002.

Maulud, D. and Abdulazeez, A.M. (2020) ‘A review on linear regression comprehensive in machine learning’, Journal of Applied Science and Technology Trends, 1(2), pp. 140–147. doi:10.38094/jastt1457.

McGrail, T. (2020) ‘An Intelligent System for Condition Assessment of Power Transformers’, [online] Available at: https://www.doble.com/wp-content/uploads/transformers-magazine-think-of-a-number-mcgrail-nov-2020.pdf [Accessed 31 Jul. 2023].

Methods, International Journal of Engineering Research & Technology (IJERT), ACMEE - 2016 Conference Proceedings, ISSN: 2278-0181.

Metwally, I. (2011) 'Failures, monitoring and new trends of Power Transformers', IEEE Potentials, 30(3), pp. 36–43. doi:10.1109/mpot.2011.940233.

Meyer, D. and Hornik, K. (2009) 'Generalized and customizable sets inr', Journal of Statistical Software, 31(2). doi:10.18637/jss.v031.i02.

MGI-The-Age-of-Analytics-Full-report. (2016). Competing in a data-driven world. McKinsey Global Institute. Available at: https://www.mckinsey.com/~/media/McKinsey/Industries/Public%20and%20Social%20Sector/Our%20Insights/The%20age%20of%20analytics%20Competing%20in%20a%20data%20driven%20world/MGI-The-Age-of-Analytics-Full-report.pdf

Miller et al. (2010) 'A practical approach to condition and risk-based power', IEEE explore. Available at: https://ieeexplore.ieee.org/document/5549580 (Accessed: 01 Feb 2021).

Milosavljevic, S. and Janjic, A. (2020) 'Integrated Transformer Health Estimation methodology based on Markov chains and evidential reasoning', Mathematical Problems in Engineering, 2020, pp. 1–12. doi:10.1155/2020/7291749.

Mitchell, T.M. (2015) 'Generative and Discriminative Classifiers: Naive Bayes And Logistic Regression', In Machine Learning; Mitchell, T.M., Ed.; McGraw Hill: New York, NY, USA.

Mohamadeen, K.I., Sharkawy, R.M. and Salama, M.M. (2014) 'Binary cat swarm optimization versus binary particle swarm optimization for Transformer Health index determination', 2014 International Conference on Engineering and Technology (ICET) [Preprint]. doi:10.1109/icengtechnol.2014.7016812.

Mohammed Nagib Omara, Khamis Salim Al-Romaimi, and Rashid Said Al-Abri, (2007) 'Transmission Pricing in the Restructured Electricity Market', International Conference on Communication, Computer and Power (ICCCP'07), Muscat, Oman.

Murugan, R. and Ramasamy, R. (2019) 'Understanding the power transformer component failures for health index-based maintenance planning in electric utilities', Engineering Failure Analysis, 96, pp. 274–288. doi:10.1016/j.engfailanal.2018.10.011.

Naderian, A. et al. (2008) 'An approach to determine the health index of power transformers', Conference Record of the 2008 IEEE International Symposium on Electrical Insulation [Preprint]. doi:10.1109/elinsl.2008.4570308.

Nasrat, L.S., Kassem, N. and Shukry, N. (2013) 'Aging effect on characteristics of oil-impregnated insulation paper for Power Transformers', *Engineering*, 05(01), pp. 1–7. doi:10.4236/eng.2013.51001.

National Grid (2019) 'Annex NGET_A14.04_ITOMS', December 2015.Available at https://www.nationalgrid.com/electricity-transmission/document/132776/download (Accessed: 23 July 2021)

Nemeth, B., Voros, C. and Csepes, G. (2014) 'Health index as one of the best practices for condition assessment of transformers and substation equipments - Hungarian experience'. CIGRE.

Nhu, V.-H. et al. (2020) 'Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve Bayes tree, artificial neural network, and support vector machine algorithms', International Journal of Environmental Research and Public Health, 17(8), p. 2749. doi:10.3390/ijerph17082749.

Nitu, M.-C. et al. (2017) 'Power transformer fault diagnosis using fuzzy logic technique based on dissolved gas analysis and furan analysis', 2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM); 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP) [Preprint]. doi:10.1109/optim.2017.7974968.

Niu, H., Lu, Q. and Wang, C. (2018) 'Color correction based on histogram matching and polynomial regression for image stitching', 2018 IEEE 3rd International

Conference on Image, Vision and Computing (ICIVC) [Preprint]. doi:10.1109/icivc.2018.8492895.

Notingher, P.V. et al. (2008) 'The use of conductivity factors for estimating the degradation state of insulation systems of medium Power Electrical Machines', 2008 International Conference on Condition Monitoring and Diagnosis [Preprint]. doi:10.1109/cmd.2008.4580246.

Oommen, T.V. (1992) 'Further Experimentation on Bubble Generation During Transformer Overload', EPRI Report EL-7391. March 1992.

Oommen, T.V. and Lindgren, S.R. (2001) 'Bubble evolution from Transformer overload', 2001 IEEE/PES Transmission and Distribution Conference and Exposition. Developing New Perspectives (Cat. No.01CH37294) [Preprint]. doi:10.1109/tdc.2001.971223.

Ortiz, F. et al. (2016) 'Health Indexes for Power Transformers: A case study', IEEE Electrical Insulation Magazine, 32(5), pp. 7–17. doi:10.1109/mei.2016.7552372.

Ozgonenel, O. and Karagol, S. (2014) 'Power Transformer protection based on decision tree approach', IET Electric Power Applications, 8(7), pp. 251–256. doi:10.1049/iet-epa.2013.0407.

Padmanaban et al.(2023) 'Determination of power transformers health index using parameters affecting the transformer's life', IETE Journal of Research, 69(11), pp.8467-8488.

Patel, M. et al. (2021) 'Simulation and analysis of DGA Analysis for power transformer using advanced control methods', ASIAN JOURNAL OF CONVERGENCE IN TECHNOLOGY, 7(1), pp. 102–109. doi:10.33130/ajct.2021v07i01.022.

Pereira, F., Mitchell, T. and Botvinick, M. (2009) 'Machine Learning Classifiers and fmri: A tutorial overview', NeuroImage, 45(1). doi:10.1016/j.neuroimage.2008.11.007.

Perez, J. (2012) 'Fundamental principles of transformer thermal loading and protection', 11th IET International Conference on Developments in Power Systems Protection (DPSP 2012) [Preprint]. doi:10.1049/cp.2012.0103.

Petchrompo, S. and Parlikad, A.K., 2019. A review of asset management literature on multi-asset systems. Reliability Engineering & System Safety, 181, pp.181-201.

Petersen, A. and Austin, P. (2004) 'Impact of recent transformer failures and fires: Australian and New Zealand experiences', CIGRE Conference.

Picher, P et al. (2012) 'New technologies for monitoring transformer tap-changers and bushings and their integration into a modern IT infrastructure', CIGRE 2012. Available at: https://www.researchgate.net/profile/Francois-Leonard-2/publication/281101468_New_technologies_for_monitoring_transformer_tap-changers_and_bushings_and_their_integration_into_a_modern_IT_infrastructure/links/55d4e61408ae1e651663716b/New-technologies-for-monitoring-transformer-tap-changers-and-bushings-and-their-integration-into-a-modern-IT-infrastructure.pdf

Picher, P. et al. (2013) 'Experience with frequency response analysis (FRA) for the mechanical condition assessment of transformer windings', 2013 IEEE Electrical Insulation Conference (EIC) [Preprint]. doi:10.1109/eic.2013.6554237.

Picher, P. et al. (2014) 'Use of health index and reliability data for transformer condition assessment and fleet ranking'. Proceedings of the CIGRÉ Biennial Session A2-101, Paris, France, pp.24-29.

Piotrowski, T. et al. (2020) 'Using the analysis of the gases dissolved in oil in diagnosis of transformer bushings with paper-oil insulation—A case study', Energies, 13(24), p. 6713. doi:10.3390/en13246713.

plotmf (Plot membership functions for input or output variable) (no date) Plot membership functions for input or output variable - MATLAB. Available at: https://www.mathworks.com/help/fuzzy/plotmf.html (Accessed: 04 June 2020).

Pompili, M. and Scatiggio, F. (2015) 'Classification in iso-ATTENTION CLASSES OF HV transformer fleets', IEEE Transactions on Dielectrics and Electrical Insulation, 22(5), pp. 2676–2683. doi:10.1109/tdei.2015.005252.

Prasojo, R.A. et al. (2021) 'Dealing with data uncertainty for Transformer Insulation System Health index', IEEE Access, 9, pp. 74703–74712. doi:10.1109/access.2021.3081699.

Psarommatis, F., Prouvost, S., May, G. and Kiritsis, D. (2020) 'Product quality improvement policies in Industry 4.0: characteristics, enabling factors, barriers, and evolution toward zero defect manufacturing', Frontiers in Computer Science, 2, p.26.

Publicly Available Specification PAS 55-1 and 2 (2008), Asset Management, ICS code: 03.100.01: Specification for optimized management of physical assets, doi:10.3403/30171836.

Puluhem et al. (2011) 'Improvement of an Health Index from ageing studies and field outcome', 8th International Conference on Insulated Power Cables, Jicable'11, 19-23 June, Versailles, France. EDF R&D, Moret-sur-Loing, France. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.jicable.org/2011/abstracts/JJicable11_0096.pdf

Qiu, J. et al. (2015) 'Nonparametric regression-based failure rate model for electric power equipment using lifecycle data', IEEE Transactions on Smart Grid, 6(2), pp. 955–964. doi:10.1109/tsg.2015.2388784.

R studio version 4.1.3 library (caret), function (confusionMatrix) help: (1. Kuhn, M. (2008), "Building predictive models in R using the caret package, " Journal of Statistical Software, (doi: 10.18637/jss.v028.i05),

Raj, J.T. (2019) What to do when your classification dataset is imbalanced., Medium. Towards Data Science. Available at: https://towardsdatascience.com/what-to-do-when-your-classification-dataset-is-imbalanced-6af031b12a36 (Accessed: January 16, 2023).

Rao et al. (2019) 'Condition monitoring of in-service oil-filled transformers: Case studies and experience', IEEE Electrical Insulation Magazine, 35(6), pp.33-42.

Rausand, M. et al. (2020) 'System reliability theory', Wiley Series in Probability and Statistics [Preprint]. doi:10.1002/9781119373940.

Rediansyah, D. and Prasojo, R.A. (2021) 'Study on artificial intelligence approaches for power transformer health index assessment', In 2021 International Conference on Electrical Engineering and Informatics (ICEEI) (pp. 1-4). IEEE.

Rediansyah, D. et al. (2021) 'Artificial Intelligence-based power transformer health index for handling data uncertainty', IEEE Access, 9, pp. 150637–150648. doi:10.1109/access.2021.3125379.

Rishit Mishra (2019) 'Using agile software development model for IT projects', International Journal of Engineering Research and, V8 (06). doi:10.17577/ijertv8is060300.

Rodriguez-Celis, E.M. et al. (2015) 'Understanding ethanol versus methanol formation from insulating paper in Power transformers', Cellulose, 22(5), pp. 3225–3236. doi:10.1007/s10570-015-0693-0.

Roncero-Clemente, C. and Roanes-Lozano, E. (2018) 'A multi-criteria computer package for power transformer fault detection and diagnosis', Applied Mathematics and Computation, 319, pp. 153–164. doi:10.1016/j.amc.2017.02.024.

Rubio-serrano, J. et al. (2012) 'Electro-acoustic detection, identification and location of partial discharge sources in oil-paper insulation systems', IEEE Transactions on Dielectrics and Electrical Insulation, 19(5), pp. 1569–1578. doi:10.1109/tdei.2012.6311502.

S. Russell, P. Norvig, (2003). 'Artificial Intelligence: A Modern Approach', (2nd ed.). Prentice Hall. ISBN 978-0137903955, 2003.

Saha, T.K. (2003) 'Review of modern diagnostic techniques for assessing insulation condition in aged transformers', IEEE Transactions on Dielectrics and Electrical Insulation, 10(5), pp. 903–917. doi:10.1109/tdei.2003.1237337.

Saha, T.K. and Purkait, P. (2004) 'Investigation of an expert system for the condition assessment of transformer insulation based on dielectric response measurements', IEEE Transactions on Power Delivery, 19(3), pp. 1127–1134. doi:10.1109/tpwrd.2004.829100.

Saha, T.K. et al. (2014) 'Pattern recognition techniques for determining the health index of oil-paper insulation of in-service power transformers', CIGRE. Available at: https://cigreindia.org/CIGRE%20Lib/CIGRE-Paris-Session-2014-papers-Report/A2/A2_105_2014.pdf

Samanta, B., Al-Balushi, K.R. and Al-Araimi, S.A. (2003) 'Artificial neural networks and support vector machines with genetic algorithm for Bearing Fault Detection', Engineering Applications of Artificial Intelligence, 16(7–8), pp. 657–665. doi:10.1016/j.engappai.2003.09.006.

San, H. (2013) 'Addressing the Data Management Challenge in Asset Management', In: Pinedo, M., Walter, I. (eds) Global Asset Management. Palgrave Macmillan, London. https://doi.org/10.1057/9781137328878_24 BS 55001 series, Asset Management:2014, and BS ISO 55002:2014, First edition 2014-01-15, Reference number ISO 55000:2014(E), Corrected version 2014-03-15.

Scatiggio, F. and Pompili, M. (2013) 'Health index: The Terna's practical approach for Transformers Fleet Management', 2013 IEEE Electrical Insulation Conference (EIC) [Preprint]. doi:10.1109/eic.2013.6554228.

Scatiggio, F. et al. (2016) 'Risk prevention for HV transformers: Beyond the health index', 2016 IEEE Electrical Insulation Conference (EIC) [Preprint]. doi:10.1109/eic.2016.7548690.

Schaut, A., Autru, S. and Eeckhoudt, S. (2011) 'Applicability of methanol as a New Marker for Paper Degradation in Power Transformers', IEEE Transactions on

Dielectrics and Electrical Insulation, 18(2), pp. 533–540. doi:10.1109/tdei.2011.5739459.

Sefidgaran, M., Mirzaie, M. and Ebrahimzadeh, A. (2011) 'Reliability Model of the Power Transformer with ONAF cooling' International Journal of Electrical Power; Energy Systems [Preprint]. doi:10.1016/j.ijepes.2011.10.002.

Senoussaoui, M.E., Brahami, M. and Fofana, I. (2021) 'Transformer Oil Quality Assessment Using Random Forest with Feature Engineering', Energies, 14(7), p. 1809. doi:10.3390/en14071809.

Sermsukroongsakul, S. et al. (2019) 'A study of remaining lifetime assessment of generator step-up transformer using degree of polymerization', 2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia) [Preprint]. doi:10.1109/gtdasia.2019.8715949.

Setayeshmehr, A. et al. (2008) 'Dielectric spectroscopic measurements on transformer oil-paper insulation under controlled laboratory conditions', IEEE Transactions on Dielectrics and Electrical Insulation, 15(4), pp. 1100–1111. doi:10.1109/tdei.2008.4591233.

Sharma, J.P. (2020) 'Regression approach to power Transformer Health Assessment Using Health index', Lecture Notes in Mechanical Engineering, pp. 603–616. doi:10.1007/978-981-15-5463-6_54.

Shintemirov et al. (2010) 'Transformer winding condition assessment using frequency response analysis and evidential reasoning', IET Electric Power Applications, 4(3), p. 198. doi:10.1049/iet-epa.2009.0102.

Sima, W. et al. (2016) 'Impact of time parameters of lightning impulse on the breakdown characteristics of oil paper insulation', High Voltage, 1(1), pp. 18–24. doi:10.1049/hve.2016.0009.

Simple Text Classification with Python and TextBlob. On-line. Available at: http://stevenloria.com/how-to-build-a-text-classification-system-with-python-and-textblob/ [Accessed 27 Jul. 2023].

Singh et al. (2013). “Real-life applications of fuzzy logic. Advances in Fuzzy Systems”, 2013, pp.3-3.

Singh, A. and Swanson, A.G. (2018) ‘Development of a plant health and risk index for Distribution Power Transformers in South Africa’, SAIEE Africa Research Journal, 109(3), pp. 159–170. doi:10.23919/saiee.2018.8532192.

Singh, J. and Singh, S. (2016) ‘Transformer Failure Analysis: Reasons and

Soni, R. and Mehta, B. (2022) ‘Evaluation of Power Transformer Health Analysis by Internal Fault Criticalities to Prevent Premature Failure Using Statistical Data Analytics Approach’, *Engineering Failure Analysis*, 136, p. 106213. doi:10.1016/j.engfailanal.2022.106213.

SreePrakashNarayana et al. (2019) 4 data collection techniques: Which One is right for you? - Atlan: Humans of data, Atlan. Available at: https://humansofdata.atlan.com/2017/08/4-data-collection-techniques-ones-right/ (Accessed: 20 July 2020).

Sukrislismono, Y. and El Nur, A.K. (2015) ‘Case study on asset health indices implementation on distribution transformer asset at PLN Bali Indonesia’, Asset Management Conference 2015 [Preprint], doi:10.1049/cp.2015.1747.

Sun, L. et al. (2016) ‘Research on multi-attribute decision-making in condition evaluation for power transformer using fuzzy AHP and modified weighted averaging combination’, IET Generation, Transmission; Distribution, 10(15), pp. 3855–3864. doi:10.1049/iet-gtd.2016.0381.

Susa et al., (2009) ‘Power transformer on-line multi-gas monitoring system’, Chalmers University of Technology, ISBN: 978-91-633-4880-8, Page(s):271 – 274, In Proceedings of the 21th Nordic Insulation Symposium: Nord-IS 09: June 15-17, 2009, Gothenburg, Sweden.

Suwanasri, T., Phadungthin, R. and Suwanasri, C. (2013) ‘Risk-based maintenance for asset management of Power Transformer: Practical experience in Thailand’,

International Transactions on Electrical Energy Systems, 24(8), pp. 1103–1119. doi:10.1002/etep.1764.

Suwnansri, T. (2014) 'Asset management of power transformer: Optimization of operation and maintenance costs', *2014 International Electrical Engineering Congress (iEECON)* [Preprint]. doi:10.1109/ieecon.2014.6925972.

T. Oommen, "Moisture Equilibrium curves - use and misuse," in Doble Client Conference, Boston, 2003.

Taengko, K. and Damrongkulkamjorn, P. (2013) 'Risk assessment for power transformers in pea substations using health index', 2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology [Preprint]. doi:10.1109/ecticon.2013.6559602.

Tamma, W.R., Azis Prasojo, R. and Suwarno, S. (2020) 'Assessment of high voltage power transformer aging condition based on health index value considering its apparent and actual age', 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE) [Preprint]. Available at: https://doi.org/10.1109/icitee49829.2020.9271778.

Tanasescu, G. et al. (2016) 'Monitoring and diagnosis of electrical equipment using a web software. Health index and remaining lifetime estimation', 2016 Conference on Diagnostics in Electrical Engineering (Diagnostika) [Preprint]. doi:10.1109/diagnostika.2016.7736495.

Tang, S., Hale, C. and Thaker, H. (2014) 'Reliability Modelling of Power Transformers with maintenance outage', Systems Science; Control Engineering, 2(1), pp. 316–324. doi:10.1080/21642583.2014.901930.

Tang, W.H. and Wu, Q.H. (2011) 'Condition monitoring and assessment of Power Transformers using computational intelligence', Power Systems [Preprint]. doi:10.1007/978-0-85729-052-6.

Tao, F. et al. (2018) 'Data-driven smart manufacturing', Journal of Manufacturing Systems, 48, pp. 157–169. doi: 10.1016/j.jmsy.2018.01.006.

Tee, S. et al. (2018) 'Practice of IEC 60422 in the ageing assessment of in-service transformers', Available at: https://www.researchgate.net/publication/318109817_Practice_of_IEC_60422_in_ageing_assessment_of_in-service_transformers (Accessed: 31 July 2023).

Tenbohlen, S. et al. (2016) 'Diagnostic measurements for Power Transformers', Energies, 9(5), p. 347. doi:10.3390/en9050347.

teNyenhuis, E.G. et al. (2002) 'Calculation of core hot-spot temperature in power and distribution transformers', IEEE Power Engineering Society Summer Meeting, [Preprint]. doi:10.1109/pess.2002.1043282.

The data science puzzle revisited (2017) KDnuggets. Available at: https://www.kdnuggets.com/2017/01/data-science-puzzle-revisited.html (Accessed: 13 June 2020).

The Institute of Asset Management. (2015) 'Asset Management, an anatomy, version 3', December 2015. The Institute of Asset Management. Available at: https://www.theiam.org/knowledge-library/asset-management-an-anatomy/ (Accessed: 23 July 2023).

Thompson, P. (2004) 'Researching Family and Social Mobility with Two Eyes: Some Experiences of the Interaction between Qualitative and Quantitative Data', International Journal of Social Research Methodology, 7, 237-257. https://doi.org/10.1080/1364557021000024785

Three Rivers DC Data Quality Strategy - Three Rivers District Council Data Quality Strategy March (no date) Studocu. Available at: https://www.studocu.com/ph/document/far-eastern-university/auditing-in-a-cis-environment/three-rivers-dc-data-quality-strategy/36165938 (Accessed: 13 July 2020).

Tran et al.(2020) 'A review of health assessment techniques for distribution transformers in smart distribution grids', Applied Sciences, 10(22), p.8115.

Trillas, E. and Eciolaza, L. (2015) "Fuzzy logic", Springer International Publishing. DOI, 10, pp.978-3.

U.S. Department of the Interior (2000) 'Facilities instructions, standards, and techniques - volume 3-30', In Transformer Maintenance; United States Bureau of Reclamation: Denver, CO, USA

U.S. Department of the Interior (2020) 'Facilities instructions, standards, and techniques - volume 4-1a', www.usbr.gov. Available at: https://www.usbr.gov/power/data/fist/FIST_4-1A_(7-2020).pdf (Accessed: 25 July 2023).

United States Department of The Interior Bureau Of Reclamation (2003) 'Transformer Diagnostics, Facilities Instructions, Standards, And Techniques', Volume 3-31 Ed.

US Department of Commerce, N. (2021) 'Dew Point vs humidity, National Weather Service', Available at: https://www.weather.gov/arx/why_dewpoint_vs_humidity (Accessed: 25 Jan 2022).

Varoquaux, G. (2018) 'Cross-validation failure: Small sample sizes lead to large error bars', NeuroImage, 180, pp. 68–77. doi:10.1016/j.neuroimage.2017.06.061.

Vasović et al.(2021) 'Comprehensive power transformers evaluation based on multi-criteria decision-making approach', Tehnički vjesnik, 28(5), pp.1448-1455.

Velasquez, R.M. and Lara, J.V. (2016) 'Electrical assessment by lightning phenomenon in power lines of double circuit', IEEE Latin America Transactions, 14(5), pp. 2217–2225. doi:10.1109/tla.2016.7530416.

Velasquez, R.M. and Lara, J.V. (2017) 'Expert system for power transformer diagnosis', 2017 IEEE XXIV International Conference on Electronics, Electrical

Engineering and Computing (INTERCON) [Preprint]. doi:10.1109/intercon.2017.8079640.

Velásquez, R.M. and Lara, J.V. (2017) 'Principal Components Analysis and adaptive decision system based on fuzzy logic for power transformer', Fuzzy Information and Engineering, 9(4), pp. 493–514. doi: 10.1016/j.fiae.2017.12.005.

Velasquez, R.M. and Lara, J.V. (2018) 'Methodology for overhead line conductor remaining life aging infrastructure and Asset Management', 2018 IEEE PES Transmission and Distribution Conference and Exhibition - Latin America (T and D-LA) [Preprint]. Available at: https://doi.org/10.1109/tdc-la.2018.8511752.

Vermeer, M. and Wetzer, J. (2014) 'Transformer health and risk indexing, including data quality management'. CIGRE, SC A2 Transformers.

Vermeer, M. et al. (2013) 'Asset management decision support modelling, using a health index, for maintenance and replacement planning'. In CIGRE SC A2 and C4 Joint Colloquium.

Vermeer, M. et al. (2015) 'Asset-management decision-support modelling, using a health and risk model', 2015 IEEE Eindhoven PowerTech [Preprint]. doi:10.1109/ptc.2015.7232556.

Vermeer, M.E. et al. (2016) 'The application of health- and risk indices as a decision-support tool for utilities', Asset Management Conference (AM 2016) [Preprint]. doi:10.1049/cp.2016.1431.

Wahab, M.A.A. et al. (1999) 'Novel modeling for the prediction of aged transformer oil characteristics', Electric Power Systems Research, 51(1), pp. 61–70. doi:10.1016/s0378-7796(98)00156-4.

Wallace, D.A., Kluss, J.V. and Ahmed, Z. (2018) 'Voltage dependence of dissipation factor measurements on high voltage instrument transformers', 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE) [Preprint]. doi:10.1109/ichve.2018.8642007.

Wang et al., (2022) 'Review on evolution of intelligent algorithms for transformer condition assessment', Frontiers in Energy Research, 10, p.904109.

Wang, J. et al. (2018) 'Deep Learning for Smart Manufacturing: Methods and Applications', Journal of Manufacturing Systems, 48, pp. 144–156. doi: 10.1016/j.jmsy.2018.01.003.

Wang, J. et al. (2021) 'Anomaly detection method of packet loss node location in heterogeneous hash networks', Computer Communications, 178, pp. 201–211. doi: 10.1016/j.comcom.2021.07.003.

Wang, M., Vandermaar, A.J. and Srivastava, K.D. (2002) 'Review of condition assessment of Power Transformers in service', IEEE Electrical Insulation Magazine, 18(6), pp. 12–25. doi:10.1109/mei.2002.1161455.

Wang, Z. et al. (2015). 'Ageing assessment of transformers through oil test database analyses and alternative diagnostic techniques' In CIGRE SC A2 Colloquium (p. 0663).

Wani and Ranga (2023) 'A review of health index-based condition assessment techniques for power and distribution transformers', Insight-Non-Destructive Testing and Condition Monitoring, 65(11), pp.625-630.

Wattakapaiboon, W. and Pattanadech, N. (2016) 'The new developed health index for Transformer Condition Assessment', 2016 International Conference on Condition Monitoring and Diagnosis (CMD) [Preprint]. doi:10.1109/cmd.2016.7757760.

Wouters, P.A., van Schijndel, A. and Wetzer, J.M. (2011) 'Remaining lifetime modeling of Power Transformers: Individual Assets and Fleets', IEEE Electrical Insulation Magazine, 27(3), pp. 45–51. doi:10.1109/mei.2011.5871368.

Xian, Y. et al. (2018) 'Research of influence of Garton effect on the measurement values of dielectric loss of capacitor voltage transformer', 2018 International Conference on Power System Technology (POWERCON) [Preprint]. doi:10.1109/powercon.2018.8601602.

Yi Cui, Ma, H. and Saha, T.K. (2013) 'Power transformer condition assessment using support vector machine with heuristic optimization', 2013 Australasian Universities Power Engineering Conference (AUPEC) [Preprint]. doi:10.1109/aupec.2013.6725452.

Yuan, Y. et al. (2020) 'Condition assessment on the 500kV OIP bushings of power transformer', 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE) [Preprint]. doi:10.1109/ichve49031.2020.9279671.

Yuan, Y. et al. (2021) 'Analysis of Vacuum Tap Changer failures in power and converter transformers and maintenance suggestions', The 16th IET International Conference on AC and DC Power Transmission (ACDC 2020) [Preprint]. doi:10.1049/icp.2020.0083.

Zach (2021) 'How to interpret diagnostic plots in R, Statology', Available at: https://www.statology.org/diagnostic-plots-in-r/ (Accessed: February 9, 2023).

Zaengl, W. (2003) 'Applications of Dielectric Spectroscopy in Time and Frequency Domain for HV Power Equipment, Part II, DEIS', IEEE, 19(6)

Zaengl, W.S. (2003) 'Dielectric spectroscopy in time and frequency domain for HV power equipment. I. Theoretical Considerations', IEEE Electrical Insulation Magazine, 19(5), pp. 5–19. doi:10.1109/mei.2003.1238713.

Zeinoddini-Meymand and Vahidi (2016) 'Health index calculation for power transformers using technical and economical parameters', IET Science, Measurement; Technology, 10(7), pp. 823–830. doi:10.1049/iet-smt.2016.0184.

Zeinoddini-Meymand, H., Saeid, M. and Vahidi, B. (2024) 'Calculation of power transformer health index using the combination of principal component analysis and mass-spring-damper model', Electrical Engineering [Preprint]. doi:10.1007/s00202-024-02459-9.

Zhang, H. (2004) 'The Optimality of Naive Bayes'. In Proceedings of the 17th International FLAIRS Conference 2004, Menlo Park, CA., AAAI Press.

Zuo, W. et al. (2016) ‘Calculation of a health index of oil-paper transformers insulation with binary logistic regression’, Mathematical Problems in Engineering, 2016, pp. 1–9. doi:10.1155/2016/6069784.

Appendix A: Publications
Appendix B: Research Milestones
Appendix C: Recent Development in Power Transformer HI
Appendix D: Data Management Processing and Issues
Appendix E: Health Index Formulation
Appendix F: Failure Modes and Degradation Mechanisms
Appendix G: Scoring Assessment Matrix's
Appendix H: Modelling and Stimulation Results

Appendix A | Publications

INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

Int. J. Strategic Engineering Asset Management, Vol. 4, No. 1, 2023 81

Health index assessment for power transformer strategic asset management in electrical utilities

Khamis Salim Al-Romaimi*, David Baglee and Derek Dixon

Faculty of Technology,
University of Sunderland,
Sunderland, UK
Email: bh34bl@research.sunderland.ac.uk
Email: researcher8@outlook.com
Email: david.baglee@sunderland.ac.uk
Email: derek.dixon@sunderland.ac.uk
*Corresponding author

Abstract: The health index (HI) has helped in driving electrical utilities' decision-making for strategic investment planning. This approach increases business ability to implement robust investment decision objectives that makes physical assets safe, efficient, and cost effective. Asset management (AM) ensures delivering the best quality of services by operating electrical network at the required performance cost while managing risks. This paper examines HI as an important section in AM. The HI is a building block in the AM process. It discusses transformer condition assessment using conventional method, data interpretation and international standards. Data management has improved, but a robust master dataset is either limited or not available due to a range of factors, including huge costs for capturing live data and data uncertainty challenges. A small number of electrical utilities are using recent data technology best practices for a large fleet of power transformers (PTs). Preliminary analysis for data management using Python Programming language (PPL) is considered.

Keywords: power transformer; asset management; health index; electrical utility; decision making; strategic investment planning; Python programming language; PPL.

Reference to this paper should be made as follows: Al-Romaimi, K.S., Baglee, D. and Dixon, D. (2023) 'Health index assessment for power transformer strategic asset management in electrical utilities', *Int. J. Strategic Engineering Asset Management*, Vol. 4, No. 1, pp.81–99.

Biographical notes: Khamis Salim Al-Romaimi has over 15 years employment experience and moved to various roles in his career as started as electrical engineer in electricity and water industries. During his work experience in this field, he is involved in operation and maintenance practices, data analysis, strategic planning, program management, asset planning, economics, system studies, design, projects, load dispatch, risk management, business continuity and international best practices for asset management development for PAS55 and ISO55001 certifications. He received his MSc in Electrical Engineering from Curtin University of Technology, Australia and his BEng in Automatic Control System Engineering, University of Sheffield. His research interests focus on asset management, electrical engineering, power transformer condition assessment, artificial intelligence, and transformer aging.

Copyright © 2023 Inderscience Enterprises Ltd.

INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

Int. J. Strategic Engineering Asset Management Vol. X, No. Y, 2023 1

Health Index Assessment for Power Transformer Strategic Asset Management in Electrical Utilities

Khamis Salim Al-Romaimi*
Faculty of Technology, University of Sunderland,
Email: bh34bl@research.sunderland.ac.uk, researcher8@outlook.com
* Corresponding author

Professor David Baglee
Faculty of Technology, University of Sunderland,
Email: david.baglee@sunderland.ac.uk

Dr Derek Dixon
Faculty of Technology, University of Sunderland,
Email: derek.dixon@sunderland.ac.uk

Abstract:

The use of Health Index (HI) has helped in driving electrical utilities decision making for strategic investment planning including operation, and maintenance (O&M) programmes. This approach increases the ability of the business to implement robust investment decision objectives that makes physical assets safe, productive, efficient, and cost effective. Asset Management (AM) framework is defined by ISO55001 to ensure electrical network operators are delivering best quality of services by operating their network at high performance, low cost while managing unexpected risks.

Data monitoring and recording have improved but still robust master dataset is either limited or not available due to a range of factors including huge costs for capturing live data, the lack of monitoring tools, limited or no data collection and, data uncertainty challenges. According to (Jahromi, 2009), there are a small number of electrical utilities around the globe who capture data using recent technologies that work in line with information best practice which serve large fleet of Power Transformers (PTs). The implementation of strategic AM standard ISO55001 framework is considered for the HI tool in this research. Recently, innovative systems became an alternative approach in structuring big data to support condition assessment and condition monitoring tools which are reliant on various data sources.

This paper is developed to examine HI as an important section in AM, it is considered the HI as a building block in AM process. It discusses HI scoring for transformer condition assessment using conventional methods that can add value to the AM practice. This includes defining HI model requirements. Power transformer health index data interpretation analysis will be considered using international standard: Institute of Electrical Electronics & Engineers (IEEE) C57.104 transactions on industrial informatics. Preliminary analysis for data management using Python Programming Language (PPL) is considered.

Copyright © 2023 Inderscience Enterprises Ltd.

DOI: 10.1504/IJSEAM.2023.10059373

INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

Int. J. Strategic Engineering Asset Management, Vol. X, No. Y, xxxx 1

Strategic Asset Management Health Index for Predicting Power Transformer Health Conditions

[Do not insert author name or contact details now - only to be completed after acceptance]

Abstract: Asset Management assists in operating electrical utilities at high performance and low cost. The Power Transformer Health Index (PTHI) is considered a good health condition evaluation and decision-making tool. PTHI is used to prioritize maintenance decisions, drive maintenance strategy, manage failure impact before it occurs, asset lifecycle planning, deferral big capitals, manage spare parts plan, and extend power transformer life. This paper presents the PTHI models' investigation which was conducted on 4324 transformer records using various Artificial Intelligent Machine Learning (ML) algorithms: Artificial Neural Network (ANN), Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT), Random Forest (RF) and k-Nearest Neighbours (KNN) in R programming language. Several evaluation metrics present comparable analyses using accuracy, sensitivity, specificity, and F1 score. According to the results, the SVM model was found applicable to local electrical utility transformers' health condition assessment. The paper addressed integrating international best practices and AM into the HI model.

Keywords: Power Transformer; Asset Management; Health Index; Electrical Utility; Decision Making; Strategic Investment Planning; R Programming Language; Machine Learning

1 Introduction

Asset management is described in (ISO55001, 2014) as "systematic and coordinated activities that linked together to achieve the common business strategy, vision, and mission". Power transformers are considered as one of the important, reliable, and expensive assets in electricity networks, this asset is assumed to remain in service for at least 30-40 years (Hillary et al., 2017).

Over the past decade, the Power Transformer Health Index (PTHI) has been considered a reliable health condition assessment tool to classify the power transformer health condition based on the Health Index (HI) model results. PTHI is considered a building block in the overall AM process (Deloitte, 2014). The best performer among the electrical utilities is the ones that operate their power system network on high service and low cost (ITOMS, 2019; ISO55001, 2014). PTHI objectives must be defined in the asset maintenance strategy (Jardine and Tsang, 2021). In this paper the power transformer health condition categorised into the following five classes: very good, good, moderate, bad, and very bad.

Artificial Intelligence (AI) is becoming an innovative solution to improve actual data accuracy for PTHI calculations. Over the last few years, AI has become the best tool for

Copyright © 202x Inderscience Enterprises Ltd.

INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

Int. J. Risk Assessment and Management, *Vol. X, No. Y, xxxx* 1

Risk-Stratify Power Transformers Based on Health Index Condition Utilizing Asset Management

[Do not insert author name or contact details now - only to be completed after acceptance]

Abstract: Recently, a cumulative records of power transformer are aging. The power transformer ageing argues necessity of a massive replacement budget. The money has become the main driver towards implementing the capital plan and reducing the asset management replacement cost. According to the statistical analysis in Python programming language, the age factor was the most correlated factor with Health Index (HI) value. This is counted as a reliable risk investigator that examine health condition of the large fleet of power transformer. HI utilized to define faulty transformers. Regression model developed in R programming language. Probability failure analysis is mapped to HI condition score and the expected degradation outcomes. The risk assessment is based on HI method applied to 1177 power transformers dataset for electrical utility. The results prove the practicality of this approach to identifying the risk condition of the transformer to ensure fast concrete maintenance decision are made.

Keywords: Asset Management, Failure Rate, Power Transformer Health Index (PTHI), Apparent Age, Risk Matrix, R programming language, Python Programming language.

1 Introduction

The paper is promoting the integration of Power Transformer Health Index (PTHI) into an overall asset management framework as a strategic solution for the electrical utility. The contribution of improving business knowledge and introducing AM culture is challenged to support effective decision making especially within diverse situations for the categorized assets based on a good understanding of AM Standard ISO55001 framework as strategic solutions. Both the technical and strategic solutions are considered in this research to enable good maintenance decision for the power transformers through a good condition assessment.

PTHI works by aggregating and processing available real data from the electrical utility at different locations and with different functions of the asset performance system into an overall condition evaluation (Velásquez,2018). Each electrical utility has it is own scoring method that is different from other electrical utilities due to various reasons but not limited to the cost investment on the software, assets size, type, age, location, climate…etc. The electrical utility must use the real data in the scoring calculation of the PTHI. This to gain a true results and value of implementing PTHI. The electrical utilities must calibrate annually the

Copyright © 202x Inderscience Enterprises Ltd.

University of Sunderland
CERTIFICATE
OF PARTICIPATION
This certificate is proudly given to
Khamis Al-Romaimi
For participating in the University of Sunderland
Postgraduate Researchers' Conference 2023
held on the 14th. September 2023
CERTIFICATE PROUDLY AWARDED ON BEHALF OF
The Conference Organisation Team

Appendix B | Research Milestones

Table B-1: The research stages over the four years.

Task	Description	Timing
1	Analyse the literature review to understand the research topic area deeply, including the importance of the Asset Management Power Transformer Health Index.	Year 1
1.1	Analyse the published literature, articles, and findings about the latest Asset Management (AM) maintenance Health Index (HI) development and submit Form RC/REV1 14/02/2019. Subgroup notes sent to Student and Director of Studies 19/02/19	Four months
1.2	Relationship between the Health Index and Asset Management standard ISO 55001, British Standard.	1/2 month
1.3	Define problem statement and research questions	One month
1.4	Importance of health check and gap analysis.	1/2 month
1.5	Health Index calculation scores	1/2 month
1.6	Understand Analytics phase	1/2 month
1.7	Understand measuring the critical factors	1/2 month
1.8	Understand determining the relative importance of scoring and weighting factors	1/2 month
1.9	Study the relationship between technical regulators, electrical utility, researchers, and manufacturing challenges.	1/2 month
	Chapter one completion	
2	Analysis of Power Transformer Requirements from Asset Management Health Index.	>>Year 1
2.1	Power transformer knowledge and speciality	1/2 month
2.2	Failure modes and ageing mechanism	One month
2.3	Power Transformer Health Index	One month

Task	Description	Timing
2.4	Detail discussion: on power transformer components	1/2 month
2.5	Transformer outer strain	1/2 month
2.6	Isolating fluid role in transformer condition analysis	1/2 month
	Chapter Two completion: Submit chapter two for AMR 25/11/2019. The feedback from the Annual Monitoring Review Assessment panel (from ANNMON form A). 28/01/2020	
3	Analysis of Health Index Condition Assessment Process: Failure mode analysis	Year 2
3.1	Understand fleet screening assessment.	
3.2	Understand detailed assessment.	
3.3	Learn intelligent algorithms assessment.	
4.	Analysis of Available Asset Health Index Models by Lead Industrial Utilities	>>year 2
4.1	Model 1 - DNV KEMA	½ month
	Model 2 - Hydro-Québec	½ month
	Model 3 – Kinetrics	½ month
	Model 4 - EDF method	½ month
	Model 5 - ABB method, Analysis, and evaluation	½ month
	Model 6 – Deloitte	½ month
	Model 7 – Electrical Utility UK	16 months
5	Dealing with data uncertainty	>>Year 3
5.1	Understanding data issues, including data collection, analysis, and preparation in electrical utilities	Two months

Task	Description	Timing
5.2	Understanding the mathematical modelling background of the Health Index	Two months
6	Health Index Model Development – Fuzzy Logic approach	2months
6.1	Health Index Model Data Analysis Based on Literature Review	½ month
6.2	Study and build the HI model based on the literature review sample data using the Fuzzy logic approach	1 ½ month
7	Health Index Model Development – Conventional method approach	Two months
7.1	Study and Build the HI model using the conventional weight average method.	Two months
8	Health Index Model Development – Artificial Intelligence Machine Learning approach	<< Year 4
8.1	Study the literature about the use of AI ML models in building Health index condition assessment.	One month
8.2	Study and build a binary logistic regression model for Power transformer HI assessment for sample data from the literature.	One month
8.3	Study and build an artificial neural network model for Power transformer HI assessment for sample data from the literature.	One month
8.4	Study and build an artificial support vector machine model for Power transformer HI assessment for sample data from the literature.	One month
8.5	Study and build the Naïve Bayes model for Power transformer HI assessment for sample data from the literature.	One month
8.6	Study and build a decision tree model for power transformer HI assessment using sample data from the literature.	One month
8.7	Study and build a random forest model for power transformer HI assessment using sample data from the literature.	One month

Task	Description	Timing
8.8	Experiment with the artificial intelligence machine learning models on actual data from the electrical utility	Three months
8.8	Evaluate the model performance and perform a comparative analysis	One month
9	Build a web app HI model.	
9.1	Choose one of the HI models that proved high-reliability results, then develop the web app	½ month
9.2	Choose the best interface to build the model	½ month
9.3	Study and understand how to deploy the model in the app	½ month
9.4	Build the tool using suitable technology software for the model's environment.	One month
9.5	Experiment and check the tool's functionality	½ month
10	Manage transformer ageing using calculated health index.	
10.1	Statistical analysis of the data	½ month
10.2	Study how the calculated HI result can be integrated into asset management strategy and help prioritise decision-making for power transformer maintenance action, whether replacement, refurbishment, or inspection.	Two months
10.3	Conduct financial impact analysis	Two months
11	PhD Thesis Completion - Final review and updates	Six months

Appendix C | Recent Development in Power Transformer HI

Table C-1: Recent developments in power transformer HI.

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
1.	Islam et al. (2017)	✓ This paper combines PT condition methods using quantitative approach for developing HI. ✓ The paper consider loading and inspection records to check HI accuracy.	✓ Measurements Of 345 Power Transformers
2.	Heywood and McGrail (2016)	✓ This paper describes the author knowledge and practical experience for developing HI using Delphic format (question and answer). ✓ This paper uses condition monitoring data and inspection records to study the failure causes and issues that causes PT failure.	DGA, Maintenance Assessment records.
3.	Jürgensen et al. (2017)	✓ This paper presents a critical discussion by listing advantages and disadvantages of HI ✓ The paper also highlights the weights based on different indicators and issues of missing data. ✓ The paper considers linguistic approach for critical observation of power transformers health state. ✓ The paper discussed oil solid paper ageing for transformer winding based on the DP test results. ✓ The HI present PT condition.	Diagnostic measurements, Weights Condition ratings, Paper age, Load factor

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
4.	Li et al. (2018)	✓ This article considers Bayesian Belief Network to study transformer oil immersed. ✓ The paper consider quality of predicting health condition in developing HI and estimate apparent age. ✓ The article utilises case study approach by benefiting from literature review information that includes DGA, AQA, records of 192 in china but the data is not shared.	Field datasets including DGA, oil quality test, age, failure statistics, some data fusion based intelligent methods,
5.	Abu-Elanien et al. (2012)	✓ This paper demonstrates how to use fuzzy logic to calculate power transformer health index for voltages 69kV and under. ✓ The results show it is a reliable method.	Dissolved gas analysis, oil and furan analysis were considered.

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
6.	Jahromi et al. (2009)	✓ This paper demonstrated an approach to calculated health index using weighing and factor scores of multiple diagnostic tests. ✓ It shared a practical approach of giving scores and weights for each element of these diagnostic test based to derive health index value that an indication of health condition. ✓ It also shared the interpretation of these values to assist in prioritize maintenance decisions.	
7.	Velásquez and Lara (2018)	✓ This research describes for studying the overhead lines health condition, it is insufficient to consider only the age and visual inspection, neither to keep the study only based on diagnostic tests according to ASTM. Therefore, these circumstances, require to consider all available parameters including overhead lines, underlines cables, age, and diagnostic test parameters for developing a good health index.	Mechanical tests, Service Record, Physical condition, Endurance test, Electric tests and maintenance records.

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
8.	The Canadian Electricity Association, Deloitte, 2015	✓ The investigation considers interview, transformer data, trends to come up with comprehensive Asset Health Index (AHI). ✓ The AHI is examined for 15 Canadian utility companies using all main data sources. ✓ The analysis suggested it is vital to improve data source within the utility for better decision making.	The report classified the AHI into five main data sources. Asset data is organised into five main elements.
9.	Vermeer, M. et al. (2013)	✓ The article highlighted how to prioritise maintenance needs considering age and occurring incidents. ✓ The study used MCS simulation and risk colour codes for risk approach which reflect to the probability failure impact into in operation assets.	
10.	Algudsi and El-Hag (2019)	✓ This paper demonstrated the convenience of using ML for Health Index problems using varies ML algorithms. The ML models in the experiments achieved high accuracy scores on both all features and reduced features experiments with three training/testing scenarios, NB (94.4%), ANN (95.1%), SVM (92.1%), KNN (95.6%), and RF (96.6%).	

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
11.	Benhmed et al. (2018)	✓ This paper examined asset selection rely in it to PT complexities using asset classes and feature selection. ✓ This shows a reduction on choosing the most important key factors for developing transformer assessment. ✓ The investigation stated the importance of the feature selection process in producing good assessment. ✓ It was found BDV, oil acidity, 2FL, water content important features to estimate the transformer HI. ✓ Accuracy achieved was 91.81% using RF and 91.81% with KNN with the K value equal to three.	
12.	Foros and Istad (2020)	✓ The article presented HI of PT by considering its remaining life, and its breakdown probability using three basic models. ✓ The three models are statistics data, monitoring data and winding deterioration analysis. ✓ Results of the combined method is used for maintenance and replacement decisions. This supports the utility to identify power transformer in bad condition and by that it helps in prioritise maintenance decisions.	

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
13.	Kadim et al. (2018)	✓ The research used NF model for total of 73 transformers but shared only 15 records. ✓ Data divided actual and simulated data using MCS. ✓ The FN consider two training sets. ✓ The author claimed to use ML for developing HI. ✓ The HI outcomes is similar to scoring conventional method. ✓ The study assessed 62% of transformers similarly using scoring method and FN trained data by MCS.	DGA, BDV, 2FL, Oil acidity and colour, DF, and Age.
14.	Rediansyah and Prasojo (2021)	✓ This paper presented an PT approach assessment considering AI. ✓ The study conducted for 504 PT records utilised 150kV to develop HI studying seven ML performance KNN, SVM, NB, ANN, DT and RF, ✓ Random Forest was the best with a 97.3% accuracy score.	oil quality, DGA, and paper condition
15.	Taengko and Damrongkulkamjorn (2013)	✓ The article introduces risk using HI approach is calculated using two parts: transformer history score and condition weighted scores. ✓ 237 power transformers records are considered to analyse risk and scoring factors. ✓ Weighted scores from transformer history score and condition score	

	Publication	Literature Analysis:	Inputs Data: (Tests or Site Inspection)
		parts are calculated to get the HI for the power transformer. ✓ The study demonstrated excellently by finding risks for related PT. ✓ Demonstrated results shows that this approach can identify the risk condition of power transformer	
16.	Varoquaux (2018)	✓ This paper is a review on cross-validation procedures for decoding in neuroimaging. ✓ The study results demonstrate that sample size negatively impacts classification accuracy.	
17.	Pereira et al. (2009).	✓ This paper study vital features in ML to enable a good training fMRI data sets classifier to interpret mental health, stimuli and human body behaviour. ✓ It was found from the literature that a small sample size can result in overfitting of the model, resulting in poor generalisation	
18.	Arbabshirani et al. (2017)	✓ This paper discussed the common downsides from machine learning perspective. ✓ It has been also found that sample size negatively impacts classification accuracy.	

Table C- 2: Power Transformer Active Components.

No.	Active Component	Description
1	Core	The core controls the magnetic flux from one winding to the next with low losses (ABB Review 1 Transformer Challenges, 2008; Jucker and Helmach, 2013; Diaby and Stiegemeier, 2017). This control uses a low magnetic reluctance to control the magnetic flux and keep a high resistance to stabilise losses of the eddy current. The core comprises a few steel sheets with a 0.2-0.3 thickness per sheet (ABB Review 1 Transformer Challenges, 2008). Core failure caused by overheating will cause mechanical distortion (AJ et al., 2018).
2	Winding	Generally, windings use various electrically isolated copper conductors to minimise eddy currents (Diaby and Stiegemeier, 2017). Manufacturers cover each conductor with a varnished sheet and paper isolation. These smaller conductors form a bundle spun by solid paper to form one winding conductor set. The last solid isolated sheet protects the winding conductor from the other two-phase conductor, nearby winding rolls or the grounded core (Lundgaard et al., 2015; ABB Review 1 Transformer Challenges, 2008). Winding failure caused by overvoltage, mechanical break, or thermal stress will cause open circuits, flashovers, or short circuits (Jan et al., 2015). A study by Arumugam (2021) investigated the causes of failure in core and winding assemblies, explicitly focusing on the detrimental effects of electrical stressors and partial discharges on PT performance.

No.	Active Component	Description
3	On Load Tap Changer (OLTC)	OLTC installation is to correct transformer voltage variance. This correction is essential for several causes, including improving voltage variations or giving adaptability derived from the overall voltage of the power network (Diaby and Stiegemeier, 2017). OLTC in the transmission network installation is part of the transformer system components of the industry transformers system since it relies on controlling transformer voltage output (ABB Review 1 Transformer Challenges, 2008). Power station step-up transformers do not require tap changers based on voltage improvement adopted using generator magnetisation. The tap changer tapping position has an automatic operation. It mechanically uses the necessary low voltage position number for the related winding turns by altering the turn and ratio of transformer voltage. OLTC is considered one of the active parts and critical transformer subsystems (Jucker & Helmach, 2013). Tap changer failure can cause the transformer to be out of service for a costly repair (ABB Review 1 Transformer Challenges, 2008; Yuan et al., 2021).
4	Bushings	The bushings are crucial mechanisms for PT operation. This operation is functional in controlling the current flow in the power grid to the PT energetic section. The current shall go through the transformer's main tank. The bushings tolerate this without negotiating the tank strength (Foros and Istad, 2020). The bushings are inexpensive mechanisms that substitute during the failure. The bushing is a vital subsystem to the transformer's health, and it is restricted even though serious bushing failures might lead to severe damage to other costly transformer mechanisms. Therefore, failure modes and deterioration considerations which disturb

No.	Active Component	Description
		transformer bushing are significant (ABB Review 1 Transformer Challenges, 2008). Bushing failure causes an oil-fuelled fire, heavy damage to the transformer and nearby equipment, labour, and harms the neighbourhood's environment (AJ et al., 2018; Jan et al., 2015).
5	Oil	Generally, oil is the liquid that isolates the tank and winding. Oil has two primary beneficial characteristics: firstly, it forms an insulating system, and secondly, it is valuable because of its fluidity. Research conducted by Jin et al. (2022) specifically examined the deterioration of insulation materials in transformers, including solid (paper), liquid (oil), and other materials. The practical consequences formed by a regular converter on a transformer will make the oil act as a cooling operator, passing the transformer's active section (Diaby and Stiegemeier, 2017). Moreover, scholars have identified thermal ageing and moisture ingress as significant elements that greatly influence the overall condition of the transformers. Losing oil quality is a substantial cause of generating a combination of gases and chemicals, which shows what circumstances can occur inside the transformer. Oil test occurs over laboratory examination. These examinations examine failure modes on significant mechanisms, which then can go unnoticed. (ABB Review 1 (Transformer Challenges, 2008; Lundgaard et al., 2014; Martin et al., 2017). Discharges and overheating areas produce gases that can occur in transformers' oils and might yield different hydrocarbon developments. Oil failure can cause flashovers, short circuits, electrical arcing, harm to labour at

No.	Active Component	Description
		work, severe damage to a transformer and nearby equipment, and an energy supply outage (Singh and Singh, 2016).
6	Tank and Auxiliary Equipment	The transformer tank is another active component for keeping the oil inside the tank for isolation purposes arrangement from peripheral aspects. This isolation includes mechanical destruction and moisture access. However, the tank permits an increase in oil level and prevents overflowing. It also lowers the temperature of its structure. Diaby and Stiegemeier (2017) stated that tank function fails due to outer coating mechanical damage and seals or gaskets failure due to ageing. The transformer tank is protected from corrosion and rust using paint coating. Any scratch or damage on the tank can cause oil leaks over the long run or short term. However, through regular maintenance, the importance of the main tank for having good seals and gaskets requires a frequent check and replacements to avoid any degradation that can occur (Lundgaard et al., 2015). Tank failure can cause an increase of extra force in the tank that leads to internal arc due to oil contamination, transformer parts degradation, and electrical storm. These problems could lead to the blowing up of the whole transformer, and the fire would harm nearby equipment and labour within the station (Jan et al., 2015).
7	Transformer Outer Strains/failures	According to Jin et al. (2022), considering failure reasons due to transformer subsystems as vital parts of the transformer's health condition, external fault reasons can lead to transformer failures, too. Outer faults can be due to thermal, chemical, mechanical, and electrical failure, described as fault reasons outside the transformer tank as

No.	Active Component	Description
		events of thermal, chemical, and mechanical faults. Also, it can occur outside bushings due to electrical fault reasons. Hence, outer fault reasons are non-electrical or electrical fault reasons. This failure, based on AJ et al. (2018), causes the effect to be reliant upon the transformer's environmental location and energy supply source area within the electrical network. Furthermore, it differs from transformer to transformer, whether it is a strategic or critical asset or not, and the type of customers in the area (industry, commercial, or diplomatic). According to IEC 60076-7(2005), the impact of outer failure observation is not immediate. Still, the effect will significantly impact future failure rates and simultaneously lead to catastrophic conditions and cost impacts (Jürgensen et al., 2016).

Appendix D | Data Management Processing and Issues

Part 1: Data Issues

Part 2: Data Processing (Interpretation Diagnosis Parameters and Analysis)

- 2.1 DGA
- 2.2 OQA
- 2.3 Furfural Analysis (Furans Content)

Part 3: Power transformer literature reviews Datasets

Part 4: Data Interpretation

Part 1: Data Issues

Table D- 1: Main Data Issues

No.	Issue	Description
1	Inaccurate Data Entry	During manual data entry, spelling, typing, and recording inaccuracies can occur which impact on the data calculation output. Since these results provide wrong response from HI condition assessment, hence data check and validation are must. Table D-2 presents various types of data validation methods based on international best practices and international standards.
2	Importance of Data Validation to Incorrect Data	Data validation is important for all cases, not limited to certain cases. The validation reflects on the quality and correctness of the data as well expresses the incorrect data during the absence of correct data. Data augmentation consideration is ideal before processing HI assessment. But this requires care and sufficient investigation to ensure the required level of data correctness and quality. The incorrect data are considered part of the missing data to enable a wise decision with respect to use solutions given in table D3 below. Accordingly, the antiquated data to a certain degree is certainly correct and not necessarily to be incorrect. This data can be applied to replace the incorrect/missing data. Therefore, these data are utilised in aggregation with amount of insecurity in comparative with the data age and the average deterioration rate.

Table D- 2: Data validation methodologies (CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761, 2019)

Validation Method	Degree of Difficulty	Objective	Advantage	Disadvantage
Ensure Data Dependability	Medium	- Units under attention like transformer, HI data attributes must only let transformer reported data	- Merely related data is recognised	- No content validation
Using Ranges /Boundaries	Medium	- Use conditional format to limit defined range to reflect with the predefined list, or standard values are specific range. - Once data is outside the defined range which can be for example year of installation (elderly priority), brand name spelling to prioritise transformers use commercial, domestic or industry based on it is criticality, condition, and others.	- Quality validation based on mistyping, numerical values, anything outside the predefined ranges.	- No validation within the predefined ranges. Predefined ranges including outliers must be designed and managed precisely.
Validation of numbers of digits	Low	- Validate digits number including actions prioritised.	- Friendly installation	- Restricted validation resources
Validation of Data Consistency	Medium	- Validate data consistency using various data attribute including transformer load condition values within the categorised voltage values.	- Very effective validation range.	- The source data attributes must be correctly accessible.
Verification of Date Format	Low	- Verify data date format within register as yyyy-mm-dd.	- Keep standard format of data register.	- Validate only the type of data format

Table D- 3 shows various solutions for controlling data unobtainability during PTHI development(CIGRE Technical Brochure 576, 2014; CIGRE Technical Brochure 761, 2019).

	Solution	Degree of Difficulty	Objective	Advantage	Disadvantage
1	Algorithms Machine learning	- Highest	- Algorithms machine learning to calculate the missing input parameters.	- PTHI scores are adaptable and can give talented scores and possibly creates innovative relationships and visions.	- Involves big data and costly to apply.
2	Apply statistical assumption on numerous input parameters	- High	- A lot of PTs with missing data. Insecurity rate is given for numerous inputs. Insecurity rate is obtained statistically using simulations tool and given with MATLAB scores.	- Pros: Insecurity rate is given and scale too which fit with local PTs situations. - Better in managing various missing input parameters.	- Capability lest to fine tune single PT state. - Statistical data is needed.
3	Use statistical assumption on partial figure of input parameter	- High	- Assessment of units more than 100, with partial missing data. Rate of insecurity is given as 1/2 input parameters. Rate of insecurity is calculated and revealed within PTHI scores.	- Insecurity rate is given and scale too which fit with local PTs situations.	- Valid in missing 1/2 inputs. - Assets less for PT state, - statistical data is needed.

	Solution	Degree of Difficulty	Objective	Advantage	Disadvantage
4	Apply a default scale	- Average	- A default rate is given with a viable scale rate. PTHI scores are with a scale of viable scores: 1) min, 2) max, or 3) moderate.	- PTHI is given, a scale w.r.t degree of confidence is obtained.	- Very intricate to develop. - The scale cannot be adjusted to local situations.
5	Apply a default rate	- Low	- Categorising PT units if less than 100 between low to medium range, then apply default rate: 1) Good, 2) Moderate, or 3) Bad, which always permits PTHI score even it can be erroneous.	- Straightforward to develop. - cost less.	- Avoid review, hard finding risky PT and signs for missing data.
6	Attribution by outer and inner situations	- High	- A lot of PTs with missing data and different with local situations. Identified basic affecting failure modes are judged in the assessment, accepting a statistical assessment attributing the expected rate of missing input parameters, involving insecurity rate.	- Insecurity rate is given and scale too which fit with local PTs situations.	- Effect of local affecting factors and statistical data is needed.
7	Prevent review, or estimate and calculate the missing data manually	- Very low	- For smaller amount of PT units like less than 10. PTHI assessment using algorithms won't be wholly computerised, so, the PTHI designer should calculate PT health state manually.	- Straightforward to develop. - cost less. - Straightforward to find the missing data.	- Avoid review, hard finding risky PT and missing data calculated manually.

Part 2: Data Processing (Interpretation Diagnosis Parameters and Analysis)

2.1 DGA

Table D- 4: Gas limit recommendations [PPM] based on various standards (Leauprasert et al., 2020).

Gas	IEC	IEEE	Dornenburg	Bureau of Reclamation
H2	100	100	200	500
CH4	75	120	50	125
C2H6	75	65	35	75
C2H4	75	50	80	175
C2H2	3	35	5	7
CO	700	350	500	750
CO2	7000	2500	6000	10000

Table D- 5: IEC/IEEE codes for DGA analysis techniques (Notingher et al.,2008)

fault code	fault type	gas ratios		
		C_2H_2 / C_2H_4	CH_4 / H_2	C_2H_4 / C_2H_6
0	no fault	< 0.1	0.1-1.0	< 0.1
1	<150^{0}C thermal fault	< 0.1	0.1-1.0	1.0-3.0
2	150^{0}C-300^{0}C thermal fault	< 0.1	> 1.0	< 1.0
3	300^{0}C-700^{0}C thermal fault	< 0.1	> 1.0	1.0-3.0
4	>700^{0}C thermal fault	< 0.1	> 1.0	> 3.0
5	low energy partial discharges	< 0.1	< 0.1	< 1.0
6	high energy partial discharges	0.1-3.0	< 0.1	< 1.0
7	low energy discharges	> 0.1	0.1-1.0	> 1.0
8	high energy discharges	0.1-3.0	0.1-1.0	> 3.0

Table D- 6: Transformer rating based on DGA factor analysis (Jahromi et al., 2009).

Rating Code	Condition	Description
A	Good	DGAF < 1.2
B	Acceptable	1.2 < DGAF < 1.5
C	Need Caution	1.5 < DGAF < 2
D	Poor	2 < DGAF < 3
E	Very poor	DGAF > 3

Table D- 7: scoring and weight factors for gas levels [PPM] (Jahromi et al., 2009).

	Score (S;)						W_i
Gas	1	2	3	4	5	6	
H2	≤100	100-200	200-300	300-500	500-700	>700	2
CH4	≤75	75-125	125-200	200-400	400-600	>600	3
C2H6	≤65	65-80	80-100	100-120	120-150	>150	3
C2H4	≤50	50-80	80-100	100-150	150-200	>200	3
C2H2	≤3	3-7	7-35	35-50	50-80	>80	5
CO	≤350	350-700	700-900	900-1100	1100-1400	>1400	1
CO2	≤2500	≤3000	≤4000	≤5000	≤7000	>7000	1

2.2 OQA

Table D- 8: Gas limit references assessment (Jahromi et al., 2009; IEEE Std, C57.104, 2008; IEEE Std, C57.106, 2006; CIGRE Working Group 05 1983).

Parameter	ASTM recommended by IEEE (IEEE Std, C57.106, 2006; IEEE Std, C57.104, 2008)	IEC recommended by CIGRE (IEEE Std, C57.106, 2006; CIGRE Working Group 05 1983)
Dielectric Breakdown	D877 D1816	IEC60156
Water content	D1533	IEC 60814
Power Factor	D924	IEC247
IFT	D971	ISO 6295
Acidity	D644 D974	IEC62021
Colour	D1500	ISO 2049

Table D- 9: Gading technique for oil analysis parameters (Jahromi et al., 2009).

	u ≤ 69 kV	69 kV <U< 230 kV	230kV≤U	Score	Wi
Dielectric Strength kV (2 mm gap)	≥45	≥52	≥60	1	3
	35-45	45-52	50-60	2	
	30-35	35-45	40-50	3	
	≤30	≤35	≤40	4	
IFT dyne/cm	≥25	≥30	≥32	1	2
	20-25	23-30	25-32	2	
	15-20	18-23	20-25	3	
	≤15	≤ 18	≤20	4	
Acid Number	≤0.05	≤0.04	≤0.03	1	1
	0.05-0.1	0.04-1.0	0.03-.07	2	
	0.1-0.2	1.0-0.15	0.07-.10	3	
	≥0.2	≥0.15	≥0.10	4	

	u ≤ 69 kV	69 kV <U< 230 kV	230kV≤U	Score	Wi
Moisture (ppm)	≤20			1	4
	20-30			2	
	30-40			3	
	>40			4	
Colour	≤ 1.5			1	2
	1.5-2.0			2	
	2.0-2.5			3	
	≥2.5			4	

2.3 Furfural Analysis (Furans Content)

Table D- 10: Furfural test or age ratings (Jahromi et al., 2009).

Rating Code	Furaldehyde [ppm]	Age Years	Condition
A	0-0.1	<20	Good
B	0.1-0.5	20-40	Acceptable
C	0.5-1	40-60	Need Caution
D	1-5	>60	Poor
E	>5	-	Very Poor

Table D- 11: Transformer components failure tests (Raj, 2019; Wattakapaiboon and Pattanadech, 2016; Hernanda et al., 2014; Naderian et al., 2008, Khamis et al., 2023).

No.	Test	Description
1	Turn ratio (TR) test	The turn ratio detects short circuit between turns or open circuits between turns in winding and tap changer. It measures the ratio between primary and secondary windings of the transformer for all tap positions of all three phases. The results are compared with the name plate details.
2	Winding resistance (WR) test	The winding resistance can detect short circuit between strands and open circuit contacts in winding conductor and tap changer. It measures the resistance in the range of 20 Ω down to mΩ. WR varies with oil temperature. During the test, the temperature must be recorded. For future evaluations, the WR must be converted to a reference temperature. The test requires an outage and isolation of the transformer.
3	Core resistance (CR) test	The core-to-ground resistance test can detect the failure modes of ungrounded core, and multiple core grounding. During the test, the core grounding can be disconnected outside the tank to measure the insulation resistance of the core to ground. This test is performed by applying a voltage in the range of 2.5 kV using Megger instrument between the core test terminal and earth for one minute. Prior to the test, ground between the grounded tank and the core must be removed. If the core ground is intact, the resultant resistance should be very low. Experience may help to locate the source of the problem.
4	Sweep frequency response analysis (SFRA)	The SFRA technique is a very sensitive method that is used to detect all mechanical, and core failure modes includes winding conductor tilting/bending, axial instability, bulk movement of winding, buckling, loose clamping structure, winding lead deformation and core deformation etc. The test is carried out by injecting a sinusoidal low voltage (about 2 to 15 V) with variable frequency (several

No.	Test	Description
		Hz to several MHz), to each phase of the HV and the LV windings with all other windings open circuited. The results of SFRA plots are compared with frequency sub band limits i.e. < 2 kHz -Core deformation, open circuits, and shorted turns, 2 kHz to 400 kHz-Bulk winding movement between windings and clamping structure and 400 kHz to 1 MHz - Deformation within the main or tap windings.
5	Short circuit impedance (SCI)	Short circuit impedance can identify the winding and core deformations. An AC voltage 415 V is applied to each phase of the HV winding with the related LV winding short circuited. The voltage and current across the HV winding are measured. The interpretation of the results is mainly based on comparing the whole impedance with the nameplate or factory values.
6	CO2/CO ratio test	The ratio CO2/CO test is used to perform as a part of DGA. It indicates aging or cellulose degradation failure modes. The normal ratio should be 3 to 10 for healthy cellulose insulation. A ratio below 3 indicates degrading paper insulation by arcing while above 10, it indicates cellulose aging from thermal heating.
7	Break down voltage test (BDV)	This test is performed to determine the dielectric breakdown voltage of service-aging oil and water content (such as water, dirt, or other conducting particles in the oil) failure modes. The oil samples are collected from the bottom of the transformer tank, then kept in BDV test chamber with 2.5 mm gap. Apply an AC voltage from 0 to until it breaks down. For each sample, the corresponding results were noted down. The average of six samples indicates the overall condition of the oil.
8	Interfacial tension (IFT)	This test is used to measure degradation of oil due to water and aging of oil failure modes. In an IFT test, the interface tension between oil and water is measured. There are several methods of

No.	Test	Description
		interfacial tension measurements. Ring or drop volume method is commonly used for testing and its units are given in dynes/cm.
9	Frequency domain spectroscopy (FDS)	Frequency Domain Spectroscopy (FDS) test indicates the DF-WCP and DF-WCO failure modes of the oil-paper insulation system. This test provides information about the high moisture in the oil-paper insulation. The procedure is similar to performing a power factor test. Using DIRANA test kit, AC voltage about 100 V with variable frequency range typically from 0.001 Hz to 1000 Hz is applied between the shorted test object of HV or LV and obtain the test results through the shorted HV or LV terminals.
10	Conductivity factor	The health index is strongly influenced by the dielectric losses which occur inside electro-insulating components volumes, but also by the conductivity factor and the polarization index despite both the wo factors are expensive for electrical utilities. Conductivity factor is calculated based on absorption/resorption currents, using (Gorgan, B. at al. (2010): $$K_c = \frac{i_a(30) - i_r(30)}{i_a(60) - i_r(60)} \qquad (D-1)$$ where $i_{a,r}$ (30, 60) represent the values of absorption/resorption currents measured at 30 and 60 s from the applied voltage start (Gorgan, B. at al. 2010; Notingher et al.,2008). Conductivity factor k_c

No.	Test	Description
		values allow the estimation of the oil's total electrical conductivity, respectively of its degradation state.
11	Degree of Polymerisation	The DP is calculated using the furan analysis from the oil sample. A score of 1-4 is used where 4 indicates the end of life of a transformer and has a weight of 0.3 in the PHI. The expected life of a transformer in the Eskom Distribution network is 40 years and it is not expected that the paper will degrade to catastrophic proportions in the first 10 to 15 years, even if highly loaded. The DP is a long-term assessment of the state of the transformer. A DP score of 4 would move the transformer into the next highest PHI category (Singh and Swanson, 2018).
12	Moisture in Paper	There are 14,000 possible PHI scores between 0 and 7.4 (Singh and Swanson, 2018). The DP, DGA and moisture score have low sensitivity to an input error. An error would give a maximum score of 0.9 which would push it into a higher category, transformers scored in a higher category would require maintenance or intervention and further oil samples would confirm the error. The oil score, however, has a high sensitivity, an error of 3 for poor oil quality would push it into the unhealthy category. Transformer oil can be replaced or conditioned, thus it requires immediate intervention if it is poor. Further oil samples would confirm the error.

Part 3: Power transformer literature reviews Datasets

Table D- 12: 30 Transformer's Medium Dataset (Abu-Elanien et al.,2012; Zuo, W. et al.,2016).

Tx No.	Year of installation	Voltage ratio (kV), and Power rating (kVA)	Acidity	BDV	DCG	Furfuraldehyde	Fuzzy logic Calculation (Abu-Elanien et al.,2012)		HI (Zuo, W. et al.,2016)	
TX01	1980	22/6.9; 5000	0.024	32.5	483	0.86	0.36	G	0	Healthy
TX02	1980	22/6.6; 3.15	0.098	4.05	254	0.65	0.3	G	1	Unhealthy
TX03	1980	22/6.6; 8	0.033	58	78	0.26	0.3	G	0	Healthy
TX04	1980	22/6.9; 8	0.226	48.7	215	5.53	0.78	B	1	Unhealthy
TX05	2000	22/6.6; 5000	0.01	75	126	0.06	0.2	VG	0	Healthy
TX06	1981	22/6.6; 4	0.075	71	38	0.53	0.3	G	0	Healthy
TX07	1981	22/6.6; 10	0.167	70.1	149	0.78	0.3	G	0	Healthy
TX08	1980	22/7.2; 15	0.092	67.8	28	0.69	0.3	G	0	Healthy
TX09	1980	22/6.9; 8000	0.035	62.7	9	0.21	0.22	VG	0	Healthy
TX10	1981	22/3.3; 3.15	0.088	37.6	197	0.31	0.3	G	0	Healthy
TX11	1970	22/6.9; 8000	0.181	25.5	35	8.76	0.94	VB	1	Unhealthy
TX12	1980	23.5/11; 25600	0.155	37.5	53	7.29	0.93	VB	1	Unhealthy
TX13	1980	11/23.5; 25600	0.115	25.6	78	9.6	0.94	VB	1	Unhealthy
TX14	1980	11/23.5; 25600	0.21	57.2	53	6.69	0.83	B	1	Unhealthy
TX15	1980	22/6.6; 8	0.089	30.4	336	5.12	0.78	B	1	Unhealthy

Tx No.	Year of installation	Voltage ratio (kV), and Power rating (kVA)	Acidity	BDV	DCG	Furfuraldehyde	Fuzzy logic Calculation (Abu-Elanien et al.,2012)		HI (Zuo, W. et al.,2016)	
TX16	1980	21.45/6.976; 5000	0.106	45.8	30	0.24	0.3	G	0	Healthy
TX17	1980	21.45/6.9; 8000	0.012	29.9	504	1.68	0.53	M	0	Healthy
TX18	1980	21.45/7.2; 15000	0.07	39.2	22	0.5	0.3	G	0	Healthy
TX19	1980	21.45/6.9; 8000	0.073	28.7	30	15.7	0.94	VB	1	Unhealthy
TX20	1980	21.45/6.9; 8000	0.063	64.5	69	0.15	0.3	G	0	Healthy
TX21	1980	22/0.4; 1600	0.019	28.4	144	0.02	0.15	VG	0	Healthy
TX22	1981	22/0.4; 630	0.01	66.9	71	0.05	0.11	VG	0	Healthy
TX23	1980	23.5/11; 25600	0.216	41.1	48	7.54	0.94	VB	1	Unhealthy
TX24	2000	22/6.98; 5000	0.01	67.6	427	0.08	0.3	G	0	Healthy
TX25	1981	22/6.6; 10	0.152	37.2	81	1.14	0.51	M	0	Healthy
TX26	1997	6.6/0.4; 1600	0.032	67.2	119	0.04	0.11	VG	0	Healthy
TX27	1980	22/6.6; 3.15	0.147	60.8	168	0.92	0.42	M	0	Healthy
TX28	1986	22/6.6; 4	0.16	70.1	10	0.06	0.3	G	0	Healthy
TX29	1981	0.38/19.2; 100	0.091	51.6	8	0.32	0.3	G	0	Healthy
TX30	1981	0.38/19.2; 100	0.229	41.4	24	1.07	0.48	M	0	Healthy

Table D- 13: 15 Transformer's small Dataset (Kadim et al.,2018)

Transformer	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
Furans (ppm)	3.34	1.68	1.43	1.4	0.6	0.1	0.1	0.13	0.32	0.401	0	0.12	0.08	0.08	0.05
H_2 (ppm)	22	98	70	8	12	8	88	180	503	277	11	13	12	14	134
CH_4 (ppm)	103	122	6	4	1	2	168	126	7	20	4	2	2	2	32
CO (ppm)	783	781	578	631	17	172	356	473	305	335	395	206	146	127	245
CO_2 (ppm)	3501	4784	8551	6535	286	3382	6080	7821	6561	2416	2700	2764	2123	1766	3086
C_2H_4 (ppm)	18	46	2	116	0	15	24	16	2	1	4	14	6	4	3
C_2H_6 (ppm)	93	105	3	1	0	0	615	373	15	35	1	0	0	0	44
C_2H_2 (ppm)	0	5	0	0	2	0	0	0	0	0	0	0	0	0	0
Moisture (ppm)	35	23	9	10	12	12	4	4	7	7	3	11	10	8	3
BDV (kV)	27	25	57	48	48	26	49	49	55	69	97	57	27	31	94
DF (%)	0.26	0.034	0.008	0.12	0.03	0.004	0.03	0.03	0.02	0.033	0.002	0.008	0.003	0.003	0.003
Acidity (mg KOH/g)	0.43	0.23	0.15	0.085	0.086	0.03	0.02	0.01	0.03	0.045	0.02	0.031	0.017	0.015	0.014
IFT (mN/m)	16	15	17	22	18	32	25	25	24	23	30	30	32	35	30
Color	7	4	3	5	3	2	1.5	1.5	1	2	1	2	1.5	1	0.5
Age (years)	23	21	11	29	18	16	7	7	4	8	14	16	16	16	4
Scoring method HI	VB	B	M	B	G	G	M	M	G	G	VG	VG	VG	VG	VG
	28	34.6	55	41.6	76.6	70	66.6	61.6	73	80	100	88.3	86.6	93.3	91.6
NF-IS HI	VB	VB	B	M	VG	G	G	G	G	G	G	VG	VG	VG	VG
	19.8	18	30	65	94	70	71	73	73	80	78	90	88	87	87
NF-MCS HI	VB	VB	B	B	M	G	G	G	G	G	VG	VG	VG	VG	VG
	20.4	22.6	35.5	45	57	80	78	81	71	73	88	85	85.6	91	86

Where: VG = Very Good; G = Good; M = Moderate; B = bad; VB = Very Bad.

Table D- 14: 3 Transformer's small Dataset (Singh and Swanson, 2018)

Parameter	T1	T2	T3
Age	12	37	12
H_2 (ppm)	0	37	9
CH_4 (ppm)	11	28	277
C_2H_6 (ppm)	2	3	2259
C_2H_4 (ppm)	12	59	36
C_2H_2 (ppm)	9	115	0
CO (ppm)	152	84	208
CO_2 (ppm)	2260	626	2571
Breakdown (kV)	15	65	72
Acid number	0.04	0.02	0.01
Water in oil (ppm)	215	10	7
DP	910	910	1300
Water in paper (ppm)	4	2.78	1.18

Part 4: Data Interpretation

Table D- 15: Health Index scoring calculation and required decision (Jahromi et al., 2009)

Category	Condition	Score (S)	Lifetime	Decision
A	Healthy	S < 1.01	>10 years	Normal maintenance
B	Moderately healthy	1.01 ≤ S < 2.01	3-10 years	Increased diagnostic testing, possible remedial work or replacement need depending on criticality
C	Moderately unhealthy	2.01 ≤ S < 3.01	<3 years	Start planning process to replace or rebuild considering risk and consequence of failure
D	Unhealthy	3.01 ≤ S	Near to the end of life	Immediately assess risk; replace or rebuild based on assessment

Table D- 16: Health Index scoring interpretation (Naderian et al.,2008).

HI score	Condition	Lifetime	Decision
0-30	Very Poor	Near to the end of life	Immediately assess risk; replace or rebuild based on assessment
30-50	Poor	<3 years	Start planning process to replace or rebuild considering risk and consequence of failure
50-70	Fair	3-10 years	Increased diagnostic testing, possible remedial work or replacement need depending on criticality
70-85	Good	>10 years	Normal maintenance
85-100	Very Good	>15 years	Normal maintenance

Table D- 17: Equations for calculating the final health index score.

Sr. No.	Year	Authors	Parameters	Equation/ Methodology
1	2009	Jahromi et.al	DGAF	$DGAF = \frac{\sum_{i=1}^{7} S_i \times W_i}{\sum_{i=1}^{7} W_i}$ S_i and W_i are scoring and weighting factors respectively.
			Load Factor	$LF = \frac{\sum_{j=0}^{4} (4-i) \times N_i}{\sum_{j=0}^{4} N_i}$
			HI	$HI = 60\% \times \frac{\sum_{j=1}^{14} K_j \times HIF_j}{\sum_{j=1}^{14} 4K_j} + 40\% \times \frac{\sum_{j=15}^{17} K_j \times HIF_j}{\sum_{j=15}^{17} 4K_j}$ Kj and HIFj are weighting and health index factors. 60% for main tank weightage and 40% for LTC weightage.
2	2013	Haema and Phadungthin	Power Factor	$(\%)PF = \frac{P}{VI} \times 100$ P = power, V = voltage, I = current
			DGA of main tank oil	$(\%)DGAF = \frac{\sum_{i=H_2}^{i=CO_2} (S_i \times W_i)}{\sum_{i=H_2}^{i=CO_2} (S_{maxi} \times W_i)} \times 100$
			DGA in OLTC	$(\%)DGATF = \frac{\sum_{i=1}^{i=n} (S_i \times W_i)}{\sum_{i=1}^{i=n} (S_{maxi} \times W_i)} \times 100$
			Oil quality in the main tank	$(\%)OQF = \frac{\sum_{i=BD}^{i=Colour} (S_i \times W_i)}{\sum_{i=BD}^{i=Colour} (S_{maxi} \times W_i)} \times 100$
			Oil quality in OLTC	$(\%)OQF = \frac{\sum_{i=BD}^{i=Colour} (S_i \times W_i)}{\sum_{i=BD}^{i=Colour} (S_{maxi} \times W_i)} \times 100$
			Final HI	$(\%)HI = [\, 0.6 \times \frac{\sum_{j=1}^{17} K_j \times HIF_j}{\sum_{j=1}^{17} HIF_{max} \times K_j} + 0.4 \times \frac{\sum_{j=15}^{17} K_j \times HIF_j}{\sum_{j=15}^{17} 4 \times K_j}]$
3	2014	Hernada et al.	HI each parameter	$HIeachparameter = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} W_i}$
			HI final	$HIfinal = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} 4 * W_i}$
4	2019	Murugan and Ramasamy		$HI = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} S_{max} \times W_i} \times 100\%$
5	2023	Khamis et al.		$DGAF = \frac{\sum_{i=1}^{7} S_i \times W_i}{\sum_{i=1}^{7} W_i}$ S_i and W_i are scoring and weighting factors respectively. $OQF = \frac{\sum_{i=1}^{7} S_i \times W_i}{\sum_{i=1}^{7} W_i}$ S_i and W_i are scoring and weighting factors respectively. $HI = \frac{\sum_{i=1}^{n} S_i \times W_i}{\sum_{i=1}^{n} S_{max} \times W_i} \times 100\%$
6	2011	Puluhem et al.		HI = Σ [Wfi x Ifi]

Table D- 18: Scoring of DP assessment

	Score
End of Life (0<DP<200)	4
Extensive deterioration (200<DP<350)	3
Moderate deterioration (350<DP<899)	2
Healthy (DP>900)	1

Table D- 19: Scoring of Age Assessment.

	Score
Healthy (Age<5)	1
Moderate deterioration(5<Age<50)	2
Extensive deterioration(20<Age<35)	3
End of Life (Age>35)	4

Table D- 20: Scoring of TDCG assessment.

TDCG	Score
Low TDCG <100	1
TDCG < 720 ppm	2
1920 > TDCG ≥ 720 ppm	3
4630 > TDCG ≥ 1920 ppm and CO2 less 50%	4
TDCG ≥ 4630 ppm	8

Table D- 21: Scoring of Basic Gas Ratio Assessment

Basic Gas Ratio	Score
None	0
D1: Discharges of low energy	2
D2: Discharges of high energy	5
T1: Thermal fault of ≥ 150 ◦C and < 300 ◦C	2
T2: Thermal fault of ≥ 300 ◦C and < 700 ◦C	3.5
T3: Thermal fault of ≥ 700 ◦C	5
PD: Partial discharge	3

Table D- 22: C2H2 Rating Scores

	Min	Max	Score
C2H2 <5 ppm	0	5	0
5 ppm ≤ C2H2 <15 ppm	5	15	4
15 ppm ≤ C2H2 <35 ppm	15	35	6
35 ppm ≤ C2H2	35	999999	8

Table D- 23: H2 Rating Scores

	Score
H2<50 ppm	0
50 ppm ≤H2<150 ppm	4
150 ppm ≤H2<250 ppm	6
250 ppm ≤H2	8

Table D- 24: C2H6 Rating Scores

	Score
C2H6<50 ppm	0
50 ppm ≤C2H6<100 ppm	4
100 ppm ≤C2H6<150 ppm	6
150 ppm ≤C2H6	8

Table D- 25: C2H4 Rating Scores

	Score
C2H4<50 ppm	0
50 ppm ≤C2H4<100 ppm	4
100 ppm ≤C2H4<150 ppm	6
150 ppm ≤C2H4	8

Table D- 26: CH4 Rating Scores

	Score
CH4<75 ppm	0
75 ppm ≤CH4<150 ppm	4
150 ppm ≤CH4<250 ppm	6
250 ppm ≤CH4	8

Table D- 27: CO Rating Scores

	Score
CO<500 ppm	0
500 ppm ≤CO<750 ppm	1
750 ppm ≤CO<1000 ppm	2.5
1000 ppm ≤CO	4

Table D- 28: Scoring of moisture in paper

Moisture in paper	Score
Low moisture	1
3% > % Moisture per dry-weight ≥ 2%	2
5% > % Moisture per dry-weight ≥ 3%	3
% Moisture per dry-weight ≥ 5%	4

Table D- 29: Scoring of Moisture in Oil Assessment

Moisture in Oil	Score
Moisture ≤ 10 ppm	0
10 ppm < Moisture ≤ 20 ppm	2
20 ppm < Moisture ≤ 30 ppm	4
30 ppm < Moisture ≤ 40 ppm	10
Moisture > 40 ppm	30

Table D- 30: Scoring of Acidity Assessment

Acidity	Score
Acidity< 0.1 mg KOH/g	0
0.1 mg KOH/g ≤ Acidity < 0.25 mg KOH/g	5
0.25 mg KOH/g ≤ Acidity < 0.4 mg KOH/g	12
Acidity > 0.4 mg KOH/g	30

Appendix E | Health Index Formulation

- HI Design Factors.
- Types of Health Indexs (HIs)
- Health Index Calculation Score

- **HI Design Factors.**

Table E- 1: HI design principles

No.	Factors	Description
1	Influencing Factors (IF)	Finding the factors which affect links performances is the first step in HI development. This means, when designing HI scores, users should determine the purpose of the HI assessment score and identify the failure modes and mechanisms which are relevant to be included in the HI scores. The test data and the information which are available for the assessment whether to include failure mode or mechanisms. The failure modes and mechanisms are underlying conditions which may cause failure including partial discharge and diagnostic results indications that are used to assess the underlying conditions that are not likely to impede operating functionality. The diagnostic outcomes are implemented to show the improvement of a specific failure mechanism. The level of a dissolved gas in the dielectric fluid, and how scores for every failure mechanism integrated to unlike indices and determining how each failure mode will be assessed.

No.	Factors	Description
	Grading Factors (GF)	The second step involves the translation, for every influencing factor of the link characteristics into a mark IF, a high value being representative of a poor condition regarding the concerned influencing factor. It is tough to properly understand and quantitatively weight the signs without seeing the specified failure mechanisms. It is critical that calculations are steady, so users are stimulated to develop a scoring matrix, so steadily, the scores are assigned respectively to every failure mechanism of every transformer. As soon as the pointers are involved to a failure mechanism and applied to find the probability of failure in a selected period. Then will be better to join the improvement of unlike failure mechanisms into a complete transformer HI assessment score.
	Weighting Factors (WF)	This is the third step; power transformer experts with the HI developers must work together in developing the weighting factors considering the facts of important subsystem from the cost point of view. The bushings and OLTCs are considered as a key player for the transformer replacement decision especially when both multiple subsystems require quick response. Therefore, the replacement decision of the whole transformer comes first as the cheapest decision. Although, both the subsystems are considered throughout the refurbishment or repair activity. The relative importance of the influencing factors is estimated through a WF. It requires developing a calibrated classification for classifying failure modes or mechanisms (scoring matrix) and then calculate HI score for each transformer in the fleet. The HI calculation scores are essential to remain joined to permit a transformer with a single failure mechanism close to failure to remain emphasised. Although, those transformers with various minor difficulties to be emphasised above the ones with less minor

No.	Factors	Description
		difficulties. This is a complicated process due to various levels of data and is supposed to be vital for unlike signs, and several arrangements are applied or have been projected. It requires in the technique of grouping consideration of the probability of failure from each failure mechanism and then multiply the survival rates from every failure mechanism simultaneously to reach the complete survival rate. This is equivalent to considering the failure of a chain knowing the failure rate of every link. Unluckily, the condition of information of the failure mechanisms and the value of the signs is not enough to use this technique with sureness, nevertheless the overall method is operative in satisfying the grouping of the above measures. Weighted average systems is implemented to join scores, nevertheless the optimal of weightings is hard and qualitative weightings might present important partiality. Using the averaging method, the extra failure mechanisms which are involved, the tougher it is to permit specific corrupted element to have the importance that earns the total. AM team shall know that various numbering scoring calculations offer a score that is suitable only for weighting transformers. The absolute weight of the score might not visibly show the total state of the transformer or the firmness to report complications with the transformer. Though, as soon as a HI is implemented to indicate the worst transformers in the fleet, the evaluation applied to obtain the calculation score will be studied to decide the right decisions and the time to respond.

- **Types of Health Indexs (HIs)**

Table E- 2: Types of Health Indexs (HIs)

No.	HI Type	Description
1	**Reliability Health Index (RHI)**	This index is implemented to rate transformer reliant on probability of failure or projected time to failure or the transformer that has degraded from its novel state. Although, the index rates all transformers that need any kind of decision reliant on its health condition urgency. AM technical team must find the most suitable decision for transformers that are requiring care. Furthermore, the Reliability or Health Index cannot usually affect in a score which straight links to the reliability or health of the transformer. When the index is well constructed, it can support in rating the transformers of large fleet, hence, the transformers which are less fit or less reliable are emphasised.

No.	HI Type	Description
2	**Replacement Index (RI)**	To identify the transformers in the fleet which should be considered and prioritised for replacement. This may require either Fleet Screening or Detailed Assessment for practical scenarios. This index is implemented to rate transformers that most require to be replaced. The scores for each transformer are reliant on: • Irreparable failure modes or • Single or multiple adjustable failure modes that are grouped, will not economically be repairable. This will only be accomplished by thought of economic aspects through the construction of this index. Adjustable failure modes refer to failure modes which can be fine-tuned or somewhat fine-tuned by repair or replacement of a component or sub-component. For instance, failure modes connected with a bushing may be adjustable as the bushing can be replaced. Failure modes that are fine-tuned using repair index are predominating and not completely adjusted. AM team can approve during comprising apparatuses such as bushings, tap changers and insulating fluids to be part of the Replacement index. These apparatuses will be repairable or replaceable. Also, the AM team might decide its more economical to replace the whole transformer, once more than one apparatus in bad state.
3	**Repair Index (RI)**	To identify the transformers that should be considered for repairs or major maintenance activities. This index is not designed to identify transformers that are due for routine maintenance including maintenance of the On-Load Tap Changer. Routine maintenance should generally be carried out in accordance with the organisation's maintenance plan.

No.	HI Type	Description
		This index is implemented to rate transformers that most require repairs or need normal maintenance. The transformer scoring method is reliant on adjustable failure modes that may be adjusted by repair or by carrying out normal maintenance. AM team can conduct further comprehensive financial and health condition investigation on transformers scores using repair index prior starting repairs to confirm that the transformers are effectively and economically possible to be repaired. A Replacement Index might also be implemented in combination with a Repair Index to confirm that transformers are identified reliant on repairs requirements that also do not have permanent failure modes or not known as expensive to repair. AM team might implement the Reliability Index in combination with a Replacement Index to find transformers which require only repair. The transformer, which is found defective and cannot be replaced, might be repaired.
4	**Refurbishment Lifetime Index (RLI)**	To identify transformers that should be considered for major refurbishment work including work that is more complex and costly than the repair work that might be carried out to the transformers highlighted by the repair index. This index is implemented to rate transformers requirement reliant on the technical refurbishment need. This index concentrates naturally on various adjustable failure modes or failure modes that includes unnecessary costs and lifetime extension. The transformers found are uneconomical to repair. This requires a detailed analysis that may cause uneconomic replace decision rather than repair decision. A good constructed RLI can confirm these transformers that developed permanent failure modes are not found appropriate for refurbishment decision. Therefore,

No.	HI Type	Description
		refurbishment is arranged for transformers emphasised by a "Refurbishment Lifetime Index" despite the fact of consideration the criticality and other matters reliant on the facts of other asset health issues. This subsection describes that Refurbishment Index uses the transformer main tank which is non repairable subsystem and the corrosive Sulphur scores. The transformers cannot be refurbished at the "Red" or "Pink" score based on adequate residual life expectancy. The scores of the other 8 subsystems are utilised to define the transformers that urgently required to be refurbished. Those 8 subsystems are scored into five groups. The nonlinear score generates much bigger scores (CIGRE Technical Brochure 323, 2007). Hence, "Red" scores are multiplied by 84 based on Non-Linear Score method 3.
5	**Composite Index (CI)**	This index groups the results of the replacement and refurbishment or repair need. It comprises a technical evaluation and common financial thoughts to permit for the best result seeing transformer maintainability failure modes. This index groups the unlike types of indices in one index. For instance, mentioning maintainable indices and non-maintainable indices in one composite index. The index finds the adjustability of a failure mode and proposes a replacement or maintenance decision.
6	**Mitigation Index (MI)**	This index is implemented to improve the technical practicability of transformer refurbishment or repair health condition is needed. When the score $RS_{Original}$ (Refurbishment Score) shows which transformer can be repaired or refurbished, then AM team will calculate $RS_{Mitigated}$ (Mitigated Refurbishment Score) via equivalent index and scoring method to calculate the score reliant on the estimated transformer health condition, then the decision will be

No.	HI Type	Description
		effectively finalised. AM team will calculate $RS_{Mitigated}$ to find whether the refurbishment or repair decision is required. For instance, refurbishment or repair decision is not made without confirming that $RS_{Mitigated}$ falls lower prearranged threshold, then the mitigated score is calculated using equation (E-1). $$MS = RS_{Original} - RS_{Mitigated} \quad \text{(E-1)}$$ A higher value for MS shows which the refurbishment or repair is ended with further impact change on transformer state. Nevertheless, the value of a mitigated score can be reliant on the implemented scoring method. Several numeric scoring methods result in a value which agrees transformers to be graded; but the score is subjective. AM team shall be familiar with the scoring method to ensure the weight and value via mitigated scores implementation is achieved successfully.

- **Health Index Calculation Score**

Table E- 3: The two decisions are managed in these circumstances as follows:

No.	Index Type	Explanation
1	Develop Replacement Index with reversible failure modes/mechanisms	This has important impact on the assessment score: ▪ Highlighting the multiple reversible failure modes/mechanisms or ▪ Uneconomic to repair a reversible failure mode/mechanism or ▪ Uneconomic reversible failure modes or mechanisms of a poor transformer state (less remaining life). The transformer specialist manages manually and simply these assessments because it's more difficult using machine learning. Weightings of scores can be doubtlessly needed, as this can assist in finding other problems.

No.	Index Type	Explanation
2	Develop Replacement Index without reversible failure modes/mechanisms;	This is through finding the transformers that have active reversible failure modes/mechanisms which need decision and ensure economical correct decision. Transformers that need actions which are uneconomically cannot be justified can be added to the list of transformers that are prioritised for replacement. Though this case manages the problems allied with a Replacement Index, including the problems that can happen when designing other HIs. In overall, this will be fixed by designing a more complex HI consist of a composite HI or using one or more simple HIs to design initial outcome which can lead to more investigation.

Table E- 4: HI scoring methods (CIGRE Technical Brochure 761, 2019).

	Method	Calculation
1.	Average	$$HI = \frac{\sum_{i=1}^{N} W_{FM_i} \times S_{FM_i}}{\sum_{i=1}^{N} W_{FM_i}}$$ $$HI = \frac{(10 + 0 + 2 + 4)}{(4)} = 4$$ $$HI_{max} = 10$$ Where, **S_{FM}** is the score of a specific failure mode, **W_{FM}** is the weighting factor per failure mode, **N** is the total number of failure modes. ***Note:*** A simple set of linear or non-linear scores can be added to the scoring matrix. Weighting factor are also can be added as per the AM team concern.
2.	Summation	$$HI = \sum_{i=1}^{N} S_{FM_i}$$ *HI* = FM1 + FM2 + FM3 + FM4 TAI = (1 × 10) + (1 × 4) + (1 × 2) + (0 × 1) + (1 × 0) = 16, HI_{max} = 40. S_{FM} is the score of a specific failure mode score, ***Note:*** A simple set of linear or non-linear scores can be added to the scoring matrix. Weighting factors are also can be added as per method 2 and the AM team concern.
3.	Worst case	$$HI = worst(S_{FM})$$ S_{FM} is the score of a specific failure mode. This is another way to show the number of failure modes which are calculated using the worst-case score. **Score** = **Red 3** present that transformer has 3 failure modes which is scored Red. HI_{max} = **Red or E**
4.	Hybrid Score	HI = 1E or **1 Red** The worst-case score is applied in the calculation defined above. For instance, a simple added score is joined with a Worst-Case score. These scores are in the below formula: TX1 = **64 Red** TX2 = **64 Orange** Its visibly that TX1 require immediate action, the number score also is like TX2.

	Method	Calculation
5.	Numerical	HI = 1 – ((1-est. PoF_{FM1}) x (1-est. PoF_{FM2}) x (1-est. PoF_{FM3}) x… (1-est. PoF_{FMn})) Benefiting from the projected Probability of Failure, reliant on test results and other calculated data, which is projected for each failure mode. Hence, the score is calculated. The projected Probability of Failures for this case are: FM1 – Red = 0.1 (10%) FM4 – Pink = 0.03 (3%) FM3 – Orange = 0.01 (1%) FM2 – Green = 0.002 (0.2%) HI=1 – (1-0.1) x (1-0.03) x (1-0.01) x (1-0.002), TAI = 0.1375
6.	Non-linear	$$HI = \sum_{n=0}^{k=1} x_n i^n \qquad (E-2)$$ Where: • i is the number base, can be like or more than the failure modes number that counted in the HI. • x_n is the number of failure modes per class. • k is the number of classes indicated in the calculations of the failure mode $$HI = \sum_{n=0}^{k=1} x_n i^n$$ $$HI = 1 \times 4^0 + 0 \times 4^1 + 1 \times 4^2 + 1 \times 4^3 + 1 \times 4^4$$ HI = 1 + 0 + 16 + 64 + 256 = 337, HI_{max} = $4x4^4$ = 1024 Where, **i** is the number base or radix, which is equal to or greater than the number of failure modes included in the HI, $\mathbf{x_n}$ is the number of failure modes per category, **k** is the number of categories included in the failure mode assessment, **n** is the counter in this equation. *Radix is a Latin word for "root". Root can be considered a synonym for base, in the arithmetical sense.
7.	Count per category	The HI is exposed as a set of numbers, rather than a specific quantity number. The failure modes number is weighed as shown in each class, for instance, applying the risk matrix of five-colour as per Table B1, a score for a transformer with 12 failure modes weighed as shown below: 3 5 3 1 0 1E 1D 1C 0B 1A

	Method	Calculation
8.	Machine learning	The score index ensures not to apply a limited method to find the required score, then the score used as a replacement that apply smart data analytic like algorithms including neural networks, to investigate both states of the data and the failure.

Appendix F | Failure Modes and Degradation Mechanisms

- List of failure modes and ageing mechanisms.
- Condition Assessment (CIGRE Technical Brochure 761, 2019)
- Images of Transformer Failure Modes
- Transformer active parts component's function, failure causes and failure sources.

- **List of failure modes and ageing mechanisms.**

Table F- 1: List of failure modes (CIGRE Technical Brochure 761, 2019)

S. No	Component	Failure mode
1.	Core	Short circuit laminations (E)
2.		Broken ground condition (E)
3.		Unintentional ground (E)
4.	Winding	Deformation of winding (M)
5.		Breakdown of insulation (E)
6.		High moisture (C)
7.	OLTC	Coking of contacts (E/C/T)
8.		Burnt resistor (E)
9.		Jammed mechanism (M)
10.	DETC	Coking of contacts (E/C/T)
11.		Jammed mechanism (M)
12.	Bushings	Corona and internal discharges (E)
13.		Loose field distributor (E/M)
14.		High resistance in connections (E/T)
15.		Puncture of capacitive insulation (E/M)
16.		Cracks in outer coating (M)
17.		Leaks (M)
18.		Pollution of outer surface (C)
19.	Oil	Formation of sludge (C)
20.		Corrosivity (C)
21.		Contamination of particles (C)
22.		Water in oil (C)
23.	Cooling	Loss of gasket sealing (M)
24.		Failure of fans or pumps (M/E)
25.	Tank	Overheating from stray flux or circulating currents (T)
26.		Leaks (M)

Table F- 2: Ageing mechanisms list (CIGRE Technical Brochure 761, 2019).

S. No	Component	Aging mechanisms
1.	Core	Loss of lamination pressure (M)
2.	Winding	Aging of paper insulation (M/C)
3.		Loss of winding pressure (M)
4.		Carbonisation of paper (T)
5.		Overheating of contacts and joints (E/T)
6.		Copper corrosion (C)
7.		Discharges or partial discharges (E)
8.	OLTC	Aging of oil (C)
9.		Aging of insulation (C)
10.		Endurance of mechanical parts (M)
11.		Silver corrosion (C)
12.	DETC	Aging of insulation (C)
13.	Bushings	Partial discharges (E)
14.		Degradation of paper insulation (C)
15.		Moisture ingress (C)
16.	Cooling	Corrosion (C)
17.		Clogging of heat exchangers (C)
18.		Clogging of water coolers (M)
19.	Oil	Oxidation of oil (C)
20.		High moisture (C)
21.	Tank	Aging of gaskets (C)
22.		Corrosion (C)
23.		Loss of sealing pressure (M)

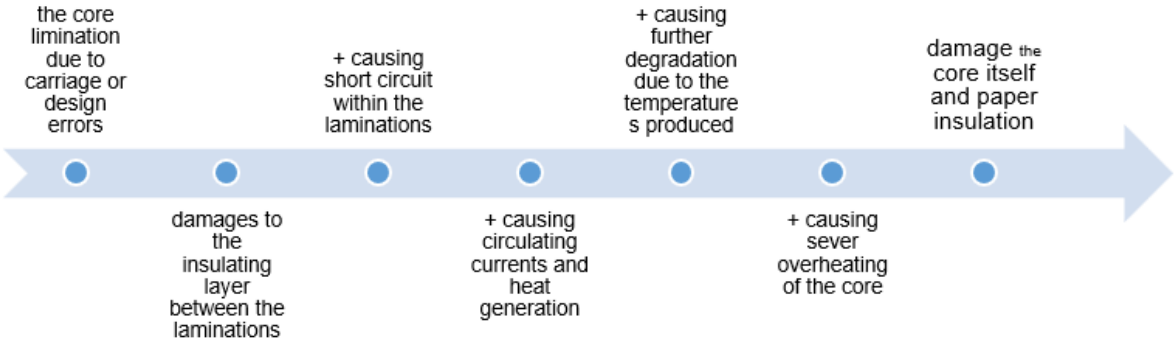


Figure F- 1: The process of core and paper insulation damage.

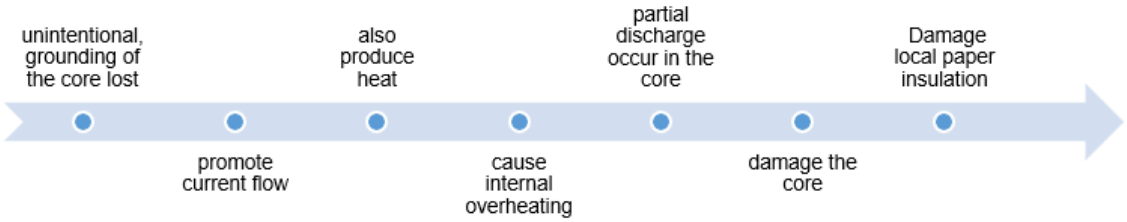


Figure F- 2: The process occurring on the damage of paper insulation.

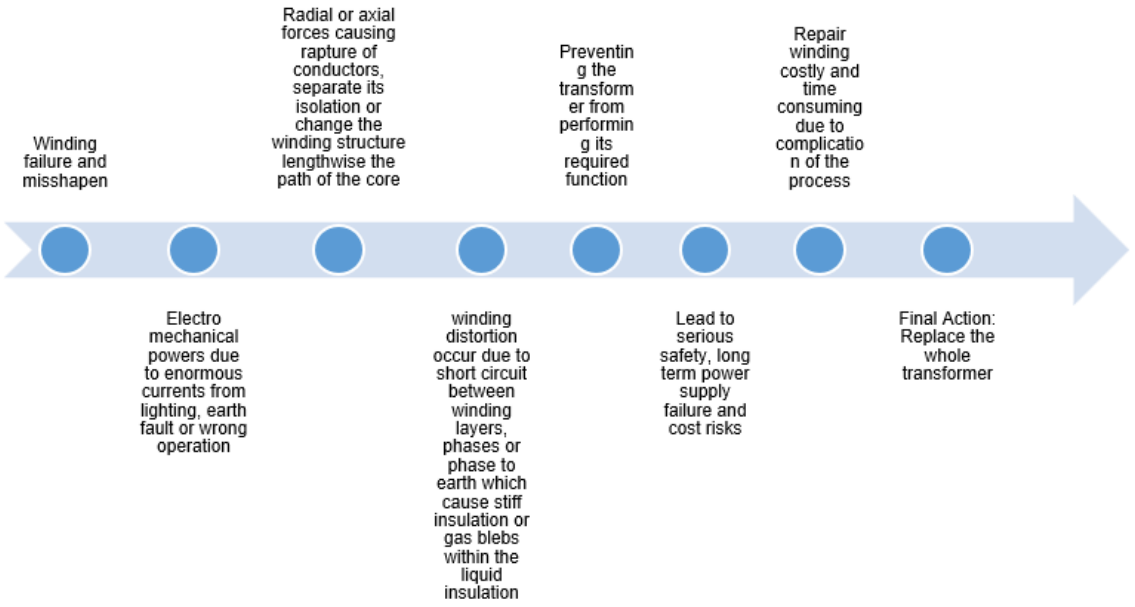


Figure F- 3: The process of winding failure that leads to replacing the whole PT.

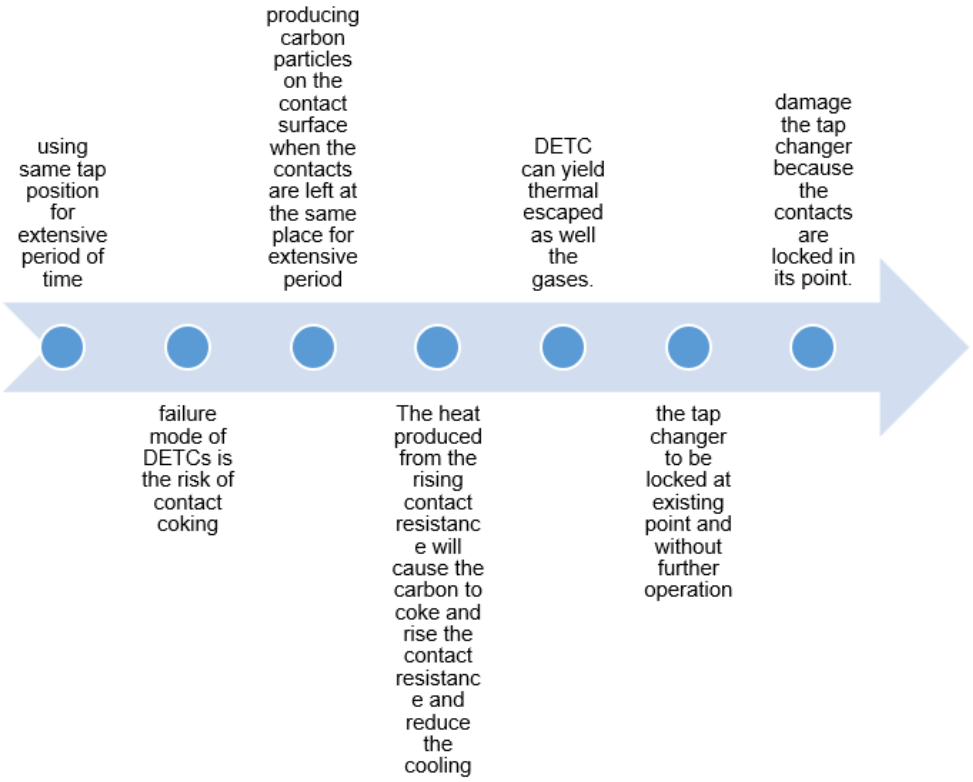


Figure F- 4: The damage process due to the tap changer.

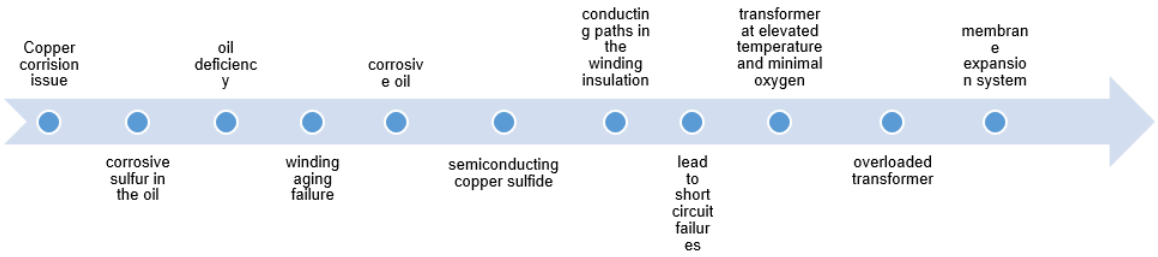


Figure F- 5: Copper corrosion process that led to membrane expansion.

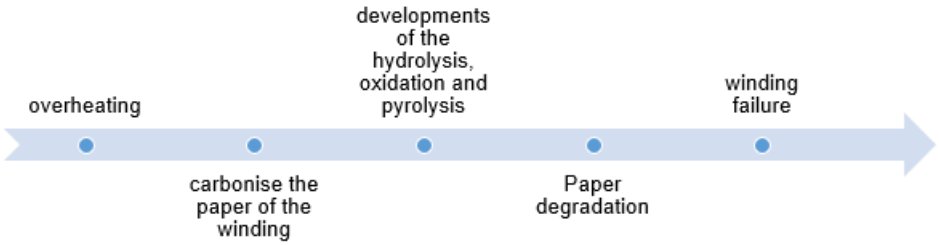


Figure F- 6: The overheating process that led to winding failure.

- **Condition Assessment**

Table F- 3: State evaluation methods for transformer based on failure mode (CIGRE Technical Brochure 761, 2019).

	Method	Explanations
1.	Advanced state evaluation	Advanced state evaluation methods are implemented to improve information on the definite state of a subsystem of a transformer. Normally, more advanced methods deliver more correct information than simple evaluations.
2.	Basic state evaluation	Basic evaluations are implemented to improve information on the definite state of a subsystem of a transformer. It is projected that basic methods deliver a practical sign.
3.	Visual Inspections	Visual inspections are implemented to improve information on the definite state of a subsystem of the transformer. The finding is restricted to degradation that is visible, audible, which can be smelt or felt. The inspections are independent to human elucidations.
4.	Rational reasoning	When evaluation or inspection is not offered, rational reasoning is used. This classified as indirect factor or information formed due to loading and age, to improve information about the most expected state of a subsystem of the transformer, • Implementing related practices of product / types / materials • Utilising environmental data and facts of physical methods • Utilising numerical interpretation of identified parameters in other similar transformers.

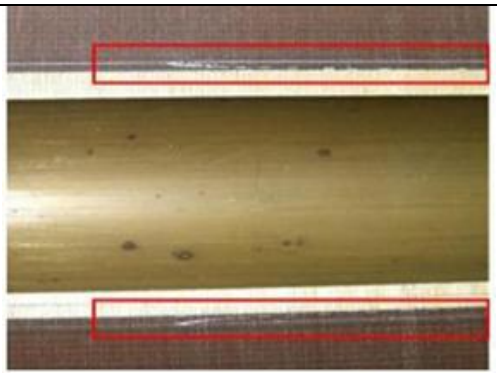
Figure F- 7: Phase to Earth Fault in fluid-filled 11kV Cable Box.

Figure F- 8: Degradation and Crystallisation of insulating liquid in 11kV and 33kV Cable Boxes (CIGRE Technical Brochure 625, 2015).

Figure F- 9: 33kV Dry Type Cable Box Failure due to PD Activity (CIGRE Technical Brochure 625, 2015).

– **Images of Transformers Failure Modes**

The images shown below represent ideal faults and failure modes of HV transformer bushings. Bushing catastrophes regularly result in combustion or blast.

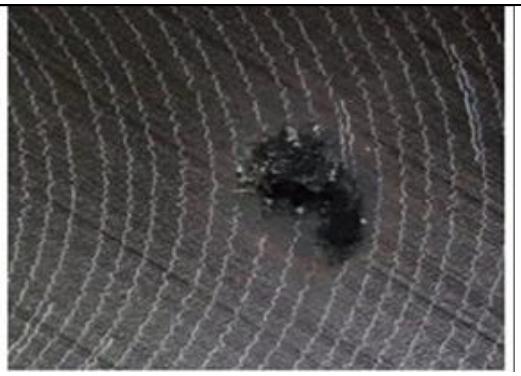


There are connection issues between the extreme scoring stratum and the electrode tube.

Access to moisture into the resin exterior on incorrect storage.

Connection catastrophe at the gauging tap because of water access and deterioration.

The connection catastrophe occurs between the internal gauging tap joint and the extreme scoring stratum due to the 'cold' fusion joint.



Fast transient over-voltages cause partial breakdown between grading layers or at the overlapping of a grading layer.

Partial discharge in an OIP insulation

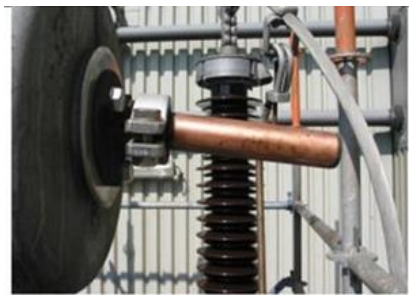
 Seepage in the enlargement container	 Damage of Oil compression
 The distorted linking screw.	 Arcing occurs among phases due to tree branches.
 The sealing gasket at the lid of the gauging tap is faulty or lost.	 The local failure of the paper insulation between the high-voltage electrode and the electrode tube occurs during fast transient overvoltage faults.

Figure F- 10: Forms of bushing failure modes (CIGRE Technical Brochure 625, 2015).

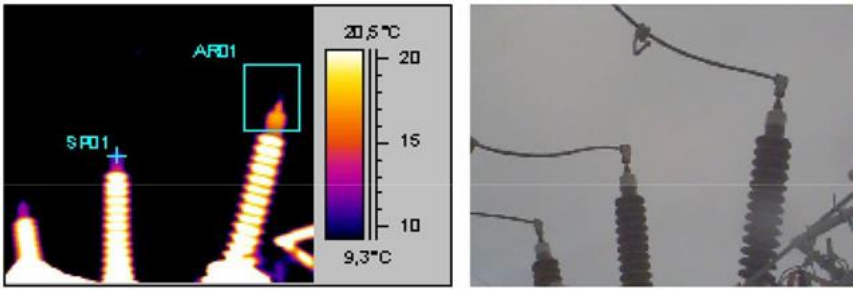


Figure F- 11: Overheating of the above station via application of IR/sensitive camera (CIGRE Technical Brochure 625, 2015).

Figure F- 12: Silver Sulphide specimen's development on OLTC acquaintances (CIGRE Technical Brochure 625, 2015).

- Blocked Radiator Valves

Figure F- 13: Particle and sledge development in water cooler (CIGRE Technical Brochure 625, 2015).

- Thermographic Inspection

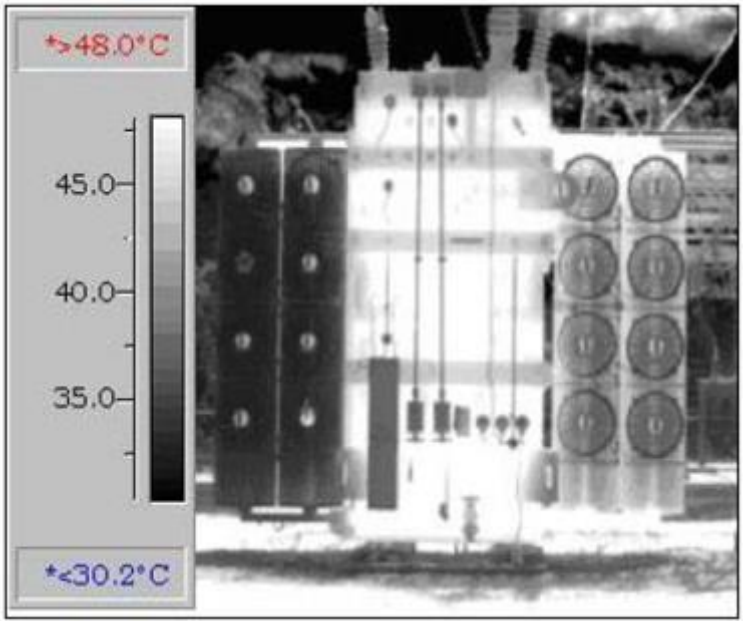


Figure F- 14: IR image demonstrating obstructed radiators (CIGRE Technical Brochure 625, 2015).

- Infrared survey

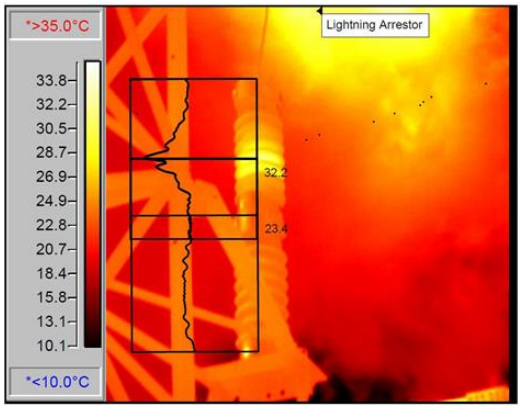


Figure F- 15: Infrared imaging displaying hot spot (U.S. Department of the Interior, 2000)

Table F- 4: The various processes for power transformer data validation (CIGRE Technical Brochure 761, 2019).

Outcomes	Validator
Some of the diagnosis tests were conducted before 36 months based on the routine tests and the outcomes were not found or doubtful.	0
Some of the diagnosis tests were conducted between 36 to 24 months based on the routine tests and the outcomes were uncertain or unsatisfactory.	4
Some of the diagnosis tests were conducted within 24 to 6 months old, based on the routine tests and the outcomes were acceptable.	7
The whole tests linked to the key validator were conducted with the routine tests and the outcomes collected were acceptable.	10

– **Transformer active parts component's function, failure causes and failure sources.**

Table F- 5: Power transformer component primary and sub-failure modes.

Transformer Active Parts	Diagnostics Test	Root cause of the failure modes	Failure modes Type
Core	Power factor (PF)	Short-circuited core laminations	Electrical
		Multiple core grounding	
		Ungrounded core	
		Core deformation	Mechanical
Winding	Turn ration (TR)	Short-circuit between turns	Electrical
	Winding resistance (WR)	Short-circuit between strands	
	Core insulation (CI) resistance	Short-circuit to ground	
		Overloading of the transformer, cooling system failure, Arcing, low-energy sparking, partial discharge, local over heating of the winding insulating material, and hot spots,	Thermal
	Sweep Frequency Response Analysis (SFRA)	Conductor tilting	Mechanical
	Short circuit impedance (SCI)	Conductor bending	
		Bulk movement of winding	
		Axial instability	
		Loose clamping structure	
		Winding buckling (hoop/compression types)	
		Winding internal lead deformation	
Tap changer	Power factor (PF)	Open circuit contacts	Electrical
		Coking of contacts	Thermal
		Drive mechanism faults	Mechanical
Bushing		Short-circuit between layers	Electrical
	DGA, CO2/CO ratio test	Network overloading, Arcing, low-energy sparking, partial discharge, local over heating	Thermal

Transformer Active Parts	Diagnostics Test	Root cause of the failure modes	Failure modes Type
		of the bushing insulating material, and hot spots.	
		Porcelain damage due to electrical flashover, lightning, external short circuit force, and cracks in outer coatings	Mechanical
Oil-Paper insulation		Water in oil	Insulation Degradation (chemical)
		Water in paper	
		High temperature of oil and paper	
	Frequency domain spectroscopy (FDS) Breakdown voltage (BDV) Interfacial tension (IFT)	Aging of oil and paper due to thermal decomposition of oil and paper insulation and contamination from moistures.	
Tank		Leakage	Mechanical/ Thermal
		Rupture	
Cooling system		Pump	Electrical and Mechanical
		Fan	
		Radiator	
Others		Operational errors	Unknown causes
		Lack of maintenance	
		Protection system failure	

Table F- 6: Transformer active parts component's function, failure causes, and failure sources (Murugan and Ramasamy, 2019).

	Power Transformers Components				
S. No	Name	Ingredients	Function	Five Failures Categories	Failures caused by
1.	Core	The laminated cores electrically isolated and magnetically coupled from each other by thin coatings of insulating material which are used to reduce eddy-current losses. The core is insulated from the grounded mechanical assemblies then intentionally grounded to a single point.	To carry magnetic flux	1. Electrical, 2. Thermal, 3. Mechanical, 4. Insulation (oil and paper) degradation and 5. Others. If one or more of these stresses exceed the design and operational limits, the service life of the transformers may be drastically reduced, or failure occurs.	1. Various manufacturing defects, 2. inadequate operational conditions and 3. deteriorated processes 4. combination of operational stresses
2.	Winding	Paper insulated, current carrying conductors wound around the core.			
3.	Tap Change	an on-load tap changer (OLTC) and a deenergised tap changer (DETC). diverter switch (used for higher voltages) and a selector switch (used for lower voltages).	To regulate the output voltage of a transformer		

	Power Transformers Components				
S. No	Name	Ingredients	Function	Five Failures Categories	Failures caused by
4.	Bushing	Solid and capacitance graded. In solid type bushings, mineral oil is used between the conductor and the insulator. In capacitance, graded bushings, oil impregnated paper (OIP), resin impregnated paper (RIP), and resin bonded paper (RBP) is using a typical insulation.			
5.	Insulation	solid (paper/pressboard) and oil	1. Paper to provide dielectric and mechanical isolation to the windings. 2. Oil common cooling systems were Oil Natural Air Natural (ONAN) and Oil Natural Air Forced (ONAF). Insulation provides sufficient cooling, dielectric strength, and minimise contact with oxygen and other materials that are at risk of oxidation.		

	Power Transformers Components				
S. No	Name	Ingredients	Function	Five Failures Categories	Failures caused by
6.	cooling system	Cooling fans, oil pumps, and water-cooled heat exchangers.	Reduces the heat produced in transformers due to copper and iron losses.		
7.	Tank		To be a container for the oil used in it. Moreover, the tank is also used to provide a support structure for all power transformer accessories and control equipment.		
8.	Others	Operational errors, lack of maintenance and protection system trips and failures are the major factors contributing to the other failure factors of power transformers.			difficult to locate and diagnose due to the complexity of internal structure of the transformers

Table F- 7: List power transformer sub-systems affected by failure modes and ageing mechanisms.

S. No	Sub System	Prevent from	Failure mode*	Aging Mechanism*	Tools to Measure Failure	Type of Application used	Solution for the problem
1.	Core	- Partial Discharge activity or - Overheating problems	Short circuited laminations (E)	Loss of lamination pressure (M)	laboratory DGA	Continuous monitoring	
			Broken ground condition (E)				
			Unintentional ground (E)				
2.	Winding		Deformation of winding (M)	Aging of paper insulation (M/C)		Site visit inspection	
			Breakdown of insulation (E)	Loss of winding pressure (M)			
			High moisture (C)	Carbonisation of paper (T)			
				Overheating of contacts and joints (E/T)			
				Copper corrosion (C)			
				Discharges or partial discharges (E)			
				Aging of paper insulation (M/C)			
3.	On Load Tap Changer		Coking of contacts (E/C/T)	Aging of oil (C)			
			Burnt resistor (E)	Aging of insulation (C)			
			Jammed mechanism (M)	Endurance of mechanical parts (M)			

S. No	Sub System	Prevent from	Failure mode*	Aging Mechanism*	Tools to Measure Failure	Type of Application used	Solution for the problem
				Silver corrosion (C)			
4.	De Energised Tap Changer		Coking of contacts (E/C/T)	Aging of insulation (C)			
			Jammed mechanism (M)				
5.	Bushings		Corona and internal discharges (E)	Partial discharges (E)			
			Loose field distributor (E/M)	Degradation of paper insulation (C)			
			High resistance in connections (E/T)	Moisture ingress (C)			
			Puncture of capacitive insulation (E/M)				
			Cracks in outer coating (M)				
			Leaks (M)				
			Pollution of outer surface (C)				
6.	Oil		Formation of sludge (C)	Oxidation of oil (C)			
			Corrosivity (C)	High moisture (C)			
			Contamination of particles (C)				
			Water in oil (C)				
7.	Tank		Overheating from stray flux or circulating currents (T)	Aging of gaskets (C)			

S. No	Sub System	Prevent from	Failure mode*	Aging Mechanism*	Tools to Measure Failure	Type of Application used	Solution for the problem
			Leaks (M)	Corrosion (C)			
				Loss of sealing pressure (M)			
8.	Bushings		Corona and internal discharges (E)	Partial discharges (E)			
			Loose field distributor (E/M)	Degradation of paper insulation (C)			
			High resistance in connections (E/T)	Moisture ingress (C)			
			Puncture of capacitive insulation (E/M)				
			Cracks in outer coating (M)				
			Leaks (M)				
			Pollution of outer surface (C)				
9.	Cooling		Loss of gasket sealing (M)	Corrosion (C)			
			Failure of fans or pumps (M/E)	Clogging of heat exchangers (C)			
				Clogging of water coolers (M)			

Appendix G | Scoring Assessment Matrix's

Table G- 1: Categorised failure modes of the scoring matrix (CIGRE Technical Brochure 761, 2019)

A	– Asset healthy and is in new condition. – Asset aging or deterioration signs are at minimum. – Asset projected for long term reliability service.
B	– Asset healthy and is in a good condition. – Asset aging or deterioration signs are obvious. – Asset projected for long term reliability service.
C	– Asset condition is okay. – Asset aging or deterioration signs are serious. – Asset for medium term operation reliability projected. – Asset is in condition-based maintenance schedule.
D	– Asset in a poor condition. – Asset shortly require planned repair/replacement. – Asset reliability service can be impaired or compromised. – Asset operation can lead to deleterious impacts. – Asset in need for rating and health condition review.
E	– Asset is near the end-of-life cycle. – Asset is about to fail. – Asset is in very poor condition. – Asset immediately require repair or replacement. – Asset is in limited service and must be taken out of service. – Asset replacement is highly required because it's operation not possible.
F	– Asset is at the end-of-life cycle. – Asset immediately must be taken out of service. – Asset to be repaired prior back to service.

Table G- 2: Thermal Assessment Scoring Matrix (CIGRE Technical Brochure 761, 2019).

	Dielectric Evaluation Scoring Matrix
A	Wholly assessment outcomes stable: no effect of acetylene; stable rate of further gases and no sign of issues from electrical assessments.
B	Signal display slight deflection from innovative, still use regular deterioration performance. Ordinary manoeuvre is projected for long period.
C	Signs display inactive or irregular arcing/sparking/tracking or breakdown defect including core insulation failure.
D	Signs display stable arcing/sparking/tracking or dielectric breakdown defect but not becoming bad with period, showing possible breakdown on medium period.
E	Signs display arcing/sparking/tracking or dielectric breakdown defect of the most vital insulation is obtaining bad and displaying breakdown in short period.
F	Dangerous partial discharge defect in transformer oil tank, expected to affect impending breakdown. Electrical assessments show definite failure of insulation.

Table G- 3: Isolating Fluids Scoring Matrix (CIGRE Technical Brochure 761, 2019).

	Dielectric Fluids Risk Scoring Matrix
A	New state: no indication for deterioration, the asset is working fine.
B	Good state. The indication for deterioration is obvious. Diminish the specimen period to control the deterioration amount.
C	Satisfactory state: by serious indications of deterioration. Starting by the assessment of the grounds led to deterioration. Later, refurbishing, repairing, replacing the issue consider overhauling, repossessing, or replacing the dielectric fluid.
D	Bad state: The grounds for the dielectric fluid deterioration must be refurbished, replaced, or repaired including oil seepage. Overhauling, repossessing, or replacement of the dielectric fluid is agreed to be arranged shortly. Damaging impacts on the active part can occur including speeding up of deterioration degree of insulating paper because of acids or sludge growth.
E	Extremely bad state: Extreme probability of failure. Can be anticipated losses on the active part or other mechanisms because of the poor state of the dielectric fluids which will cause transformer failure. The grounds for the dielectric fluid deterioration must be actioned and repaired including oil seepage. Overhauling, repossessing, or replacement of the dielectric fluid as rapidly as required is advised.
F	Shutdown as early as conceivable: System normalisation is arranged only once the repair, or the issue is fixed.

Appendix H | Modelling and Stimulation Results

Part 1 Conventional Method Health Index Scoring Rules in Spreadsheet

Part 2 Fuzzy Logic Rules

Part 3 Machine Learning Methods

Part 4 Mobile, Website Application and Flex Dashboard

Part 5 Power Transformer Age Study and Failure Rate Experiments

Part 6 Polynomial Regression Models

Part 1 Conventional Method Health Index Scoring Rules in Spreadsheet

1. To calculate the DGA Factor

Table H- 1: Grading Method for DGA Factor based FOR U<=69 kV

Individual Gases	Measurement	Min	Max	Score	Weight
H_2	<=100	0	101	1	2
	100-200	101	200	2	
	200-300	200	300	3	
	300-500	300	500	4	
	500-700	500	700	5	
	>700	700	9999999	6	
CH_4	<=75	0	76	1	3
	75-125	76	125	2	
	125-200	125	200	3	
	200-400	200	400	4	
	400-600	400	600	5	
	>600	600	999999	6	
C_2H_6	<=65	0	66	1	3
	65-80	66	81	2	
	80-100	81	101	3	
	100-120	101	121	4	
	120-150	121	151	5	
	>150	151	999999	6	
C_2H_4	<=50	0	51	1	3
	50-80	51	80	2	
	80-100	80	100	3	
	100-120	100	120	4	
	120-150	120	150	5	
	>150	150	999999	6	
C_2H_2	<=3	0	4	1	5
	3-7	4	7	2	
	7-35	7	35	3	
	35-50	35	50	4	
	50-80	50	80	5	
	>80	80	999999999	6	
CO	<=350	0	351	1	1
	350-700	351	700	2	
	700-900	700	900	3	
	900-1100	900	1100	4	
	1100-1400	1100	1400	5	

Individual Gases	Measurement	Min	Max	Score	Weight
	>1400	1400	999999	6	
CO_2	<=2500	0	2501	1	1
	<=3000	2501	3001	2	
	<=4000	3001	4001	3	
	<=5000	4001	5001	4	
	<=7000	5001	7001	5	
	>7000	7001	999999	6	

2. **To calculate the Oil Quality Factor**

Table H- 2: Grading Method for Oil Test Parameters based FOR U<=69 kV

Individual Oil Parameters	U<=69kV	Rating	Min	Max	Score (Si)	Weight (Wi)
Dielectric Strength kV (2 mm gap)	<=30	D	0	31	4	
	30-35	C	31	35	3	
	35-45	B	35	45	2	
	>=45	A	45	999999	1	3
IFT dyne/cm	<=15	D	0	15	4	
	15-20	C	15	20	3	
	20-25	B	20	25	2	
	>=25	A	25	999999	1	2
Acid Number	<=0.05	A	0	0.05	1	
	0.05-0.1	B	0.05	0.1	2	
	0.1-0.2	C	0.1	0.2	3	
	>=0.2	D	0.2	999999	4	1
Moisture (ppm)	<=20	A	0	21	1	
	20-30	B	21	30	2	
	30-40	C	30	40	3	
	>40	D	40	999999	4	4
Colour	<=1.5	A	0	1.6	1	
	1.5-2.0	B	1.6	2	2	
	2.0-2.5	C	2	2.5	3	
	>=2.5	D	2.5	999999	4	2
Dissipation Factor	<=0.1	A	0	0.09	1	
	0.1-0.5	B	0.1	0.4	2	
	0.5-1.0	C	0.5	0.9	3	
	>=1.0	D	1	999999	4	3

Table H- 3: Transformer Rating based on DGA Factor

Rating	Condition	Description	Min	Max	**Score**	Weight
A	Good	DGAF<1.2	0	1.2	4	10
B	Acceptable	1.2<=DGAF<1.5	1.2	1.5	3	
C	Need Caution	1.5<=GAF<2	1.5	2	2	
D	Poor	2<=DGAF<3	2	3	1	
E	Very Poor	DGAF>=3	3	999999	0	

An oil quality factor (OQF) similar to DGAF is developed using the score and weighting factor.

Table H- 4: Transformer Rating based on OQF Factor.

Rating	Condition	Description	Min	Max	**Score**	Weight
A	Good	OQF<1.2	0	1.2	4	10
B	Acceptable	1.2<=OQF<1.5	1.2	1.5	3	
C	Need Caution	1.5<=OQF<2	1.5	2	2	
D	Poor	2<=OQF<3	2	3	1	
E	Very Poor	OQF>=3	3	999999	0	

Table H- 5: Age Rating

	Description	Rating Code	Min	Max	**Score**	Weight
Age	Less than 20	A	0	20	4	3
	20-40	B	20	40	3	
	40-60	C	40	60	2	
	More than 60	D	60	999999	1	

Table H- 6: Furan Rating

	Description	Rating Code	Min	Max	**Score**	Weight
Furan	Good	A	0	0.1	4	8
	Acceptable	B	0.11	0.5	3	
	Need caution	C	0.51	1	2	
	Poor	D	1.01	5	1	
	Very Poor	E	5.01	999999	0	

Table H- 7: Interpretation of Output

Health Index	Min	Max	Condition	Maintenance Action
0-29	0	29	Very Bad	Replacement
30-49	30	49	Bad	Replacement
50-69	50	69	Moderate	Repair and Refurbishment
70-84	70	84	Good	Inspection
85-100	85	100	Very Good	Inspection

Table H- 8: The electrical utility Interpretation of HI value

Min	Max	Condition	Maintenance Action
0	0.5	Very Good	Inspection
0.5	5.4	Good	Inspection
5.5	9	Fair	Repair and Refurbishment
10	100	Poor	Replacement
100	99999	Very Poor	Replacement

– Part 2 Fuzzy Logic Rules

Table H- 9: Fuzzy-Logic simple rules for the PTHI model.

Rule No.	Rule
1	If (2-Furfuraldehyde is very_bad) then (Health_index is very_bad)
2	If (2-Furfuraldehyde is bad) and (DCG is not bad) then (Health_index is bad)
3	If (2-Furfuraldehyde is bad) and (DCG is bad) then (Health_index is very_bad)
4	If (2-Furfuraldehyde is high_moderate) and (DCG is bad) then (Health_index is bad)
5	If (2-Furfuraldehyde is high_moderate) and (BDV is bad) and (DCG is not bad) and (Water_content is bad) and (Acidity is bad) then (Health_index is bad)
6	If (2-Furfuraldehyde is high_moderate) and (DCG is not bad) and (Water_content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is bad)
7	If (2-Furfuraldehyde is high_moderate) and (BDV is bad) and (DCG is not bad) and (Water_content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is bad)
8	If (2-Furfuraldehyde is high_moderate) and (BDV is bad) and (DCG is not bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is bad)
9	If (2-Furfuraldehyde is high_moderate) and (DCG is not bad) and (Water_content is not bad) and (Acidity is not bad) then (Health_index is moderate)
10	If (2-Furfuraldehyde is high_moderate) and (BDV is not bad) and (DCG is not bad) and (Water_content is not bad) then (Health_index is moderate)
11	If (2-Furfuraldehyde is high_moderate) and (BDV is not bad) and (DCG is not bad) and (Dissipation_factor is not bad) then (Health_index is moderate)
12	If (2-Furfuraldehyde is high_moderate) and (BDV is not bad) and (DCG is not bad) and (Acidity is not bad) then (Health_index is moderate)
13	If (2-Furfuraldehyde is high_moderate) and (DCG is not bad) and (Acidity is not bad) and (Dissipation_factor is not bad) then (Health_index is moderate)

Rule No.	Rule
14	If (2-Furfuraldehyde is low moderate) and (DCG is bad) and (Water_content is not bad) and (Acidity is not bad) then (Health_index is moderate)
15	If (2-Furfuraldehyde is good) and (DCG is bad) and (Water content is not bad) and (Acidity is not bad) then (Health index is moderate)
16	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is bad) and (Water_content is not bad) then (Health_index is moderate)
17	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is bad) and (Water content is not bad) then (Health index is moderate)
18	If (2-Furfuraldehyde is low moderate) and (DCG is bad) and (Water_content is not bad) and (Dissipation_factor is not bad) then (Health index is moderate)
19	If (2-Furfuraldehyde is good) and (DCG is bad) and (Water content is not bad) and (Dissipation_factor is not bad) then (Health_index is moderate)
20	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is bad) and (Acidity is not bad) then (Health index is moderate)
21	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is bad) and (Acidity is not bad) then (Health_index is moderate)
22	If (2-Furfuraldehyde is low moderate) and (DCG is bad) and (Acidity is not bad) and (Dissipation_factor is not bad) then (Health_index is moderate)
23	If (2-Furfuraldehyde is good) and (DCG is bad) and (Acidity is not bad) and (Dissipation factor is not bad) then (Health index is moderate)
24	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is bad) and (Dissipation_factor is not bad) then (Health_index is moderate)
25	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is bad) and (Dissipation factor is not bad) then (Health index is moderate)
26	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is bad) and (Water_content is bad) and (Acidity is bad) then (Health_index is bad)
27	If (2-Furfuraldehyde is low moderate) and (DCG is bad) and (Water_content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health index is bad)
28	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is bad) and (Water_content is bad) and (Acidity is bad) then (Health_index is bad)
29	If (2-Furfuraldehyde is good) and (DCG is bad) and (Water content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health index is bad)
30	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is bad) and (Water_content is bad) and (Dissipation_factor is bad) then (Health index is bad)
31	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is bad)
32	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is bad) and (Water_content is bad) and (Dissipation_factor is bad) then (Health_index is bad)

Rule No.	Rule
33	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is bad)
34	If (2-Furfuraldehyde is low moderate) and (DCG is moderate) and (Water_content is bad) then (Health_index is moderate)
35	If (2-Furfuraldehyde is low moderate) and (DCG is moderate) and (Acidity is bad) then (Health_index is moderate)
36	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is moderate) then (Health_index is moderate)
37	If (2-Furfuraldehyde is low moderate) and (DCG is moderate) and (Dissipation_factor is bad) then (Health_index is moderate)
38	If (2-Furfuraldehyde is good) and (DCG is moderate) and (Water_content is bad) then (Health_index is moderate)
39	If (2-Furfuraldehyde is good) and (DCG is moderate) and (Water_content is bad) and (Acidity is bad) then (Health_index is moderate)
40	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is moderate) then (Health_index is moderate)
41	If (2-Furfuraldehyde is good) and (DCG is moderate) and (Dissipation_factor is bad) then (Health_index is moderate)
42	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is moderate) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation_factor is not bad) then (Health_index is good)
43	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is moderate) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation_factor is not bad) then (Health_index is good)
44	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Water_content is bad) and (Acidity is bad) then (Health_index is moderate)
45	If (2-Furfuraldehyde is low moderate) and (DCG is good) and (Water_content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is moderate)
46	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Water_content is bad) and (Acidity is bad) then (Health_index is moderate)
47	If (2-Furfuraldehyde is good) and (DCG is good) and (Water_content is bad) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is moderate)
48	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Water_content is bad) and (Dissipation_factor is bad) then (Health_index is moderate)
49	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is moderate)
50	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Water_content is bad) and (Dissipation_factor is bad) then (Health_index is moderate)
51	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Acidity is bad) and (Dissipation_factor is bad) then (Health_index is moderate)

Rule No.	Rule
52	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation factor is not bad) then (Health index is good)
53	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation_factor is not bad) then (Health index is good)
54	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation factor is not good) then (Health index is good)
55	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation_factor is not good) then (Health index is good)
56	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not bad) and (Dissipation factor is not good) then (Health index is good)
57	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is not good) and (Dissipation factor is not bad) then (Health index is good)
58	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not good) and (Acidity is not bad) and (Dissipation factor is not bad) then (Health index is good)
59	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not good) and (Acidity is not bad) and (Dissipation factor is not bad) then (Health index is good)
60	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Water_content is bad) then (Health_index is moderate)
61	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Acidity is bad) then (Health index is moderate)
62	If (2-Furfuraldehyde is low moderate) and (BDV is bad) and (DCG is good) and (Dissipation_factor is bad) then (Health_index is moderate)
63	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Water content is bad) then (Health index is moderate)
64	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Acidity is bad) then (Health_index is moderate)
65	If (2-Furfuraldehyde is good) and (BDV is bad) and (DCG is good) and (Dissipation factor is bad) then (Health index is moderate)
66	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is bad) and (Dissipation_factor is not bad) then (Health index is good)
67	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Acidity is bad) and (Dissipation_factor is not bad) then (Health index is good)
68	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is bad) and (Dissipation_factor is not bad) then (Health index is good)
69	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Acidity is bad) and (Dissipation_factor is not bad) then (Health_index is good)

Rule No.	Rule
70	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is bad) and (Acidity is not bad) then (Health index is good)
71	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Acidity is not bad) and (Dissipation_factor is bad) then (Health index is good)
72	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is bad) and (Acidity is not bad) then (Health_index is good)
73	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Acidity is not bad) and (Dissipation_factor is bad) then (Health_index is good)
74	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is bad) then (Health index is good)
75	If (2-Furfuraldehyde is low moderate) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Dissipation_factor is bad) then (Health index is good)
76	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Acidity is bad) then (Health_index is good)
77	If (2-Furfuraldehyde is good) and (BDV is not bad) and (DCG is good) and (Water_content is not bad) and (Dissipation_factor is bad) then (Health index is good)
78	If (2-Furfuraldehyde is low moderate) and (BDV is good) and (DCG is good) and (Water_content is good) and (Acidity is good) and (Dissipation factor is good) then (Health index is very good)
79	If (2-Furfuraldehyde is good) and (BDV is good) and (DCG is good) and (Water_content is good) and (Acidity is good) and (Dissipation_factor is good) then (Health index is very good)

Part 3 Machine Learning Methods

1. Screenshot of Binary Logistic regression Implementation on the 30 transformers

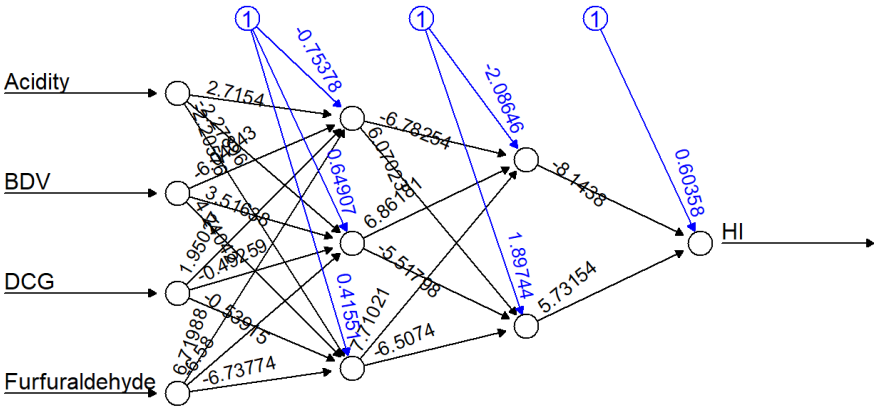
```
Call:
glm(formula = HI ~ Acidity + BDV + DCG + Furfuraldehyde, family = "binomial",
    data = training)

Deviance Residuals:
          Min             1Q         Median
-0.0000114883  -0.0000000211  -0.0000000211
           3Q             Max
-0.0000000211   0.0000130127

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -42.764 491134.903       0        1
Acidity           -2.026 416672.174       0        1
BDV               11.636 446989.352       0        1
DCG               -6.624 356595.048       0        1
Furfuraldehyde   205.209 520487.353       0        1


(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 23.58223512551355  on 21  degrees of freedom
Residual deviance:  0.00000000040254  on 17  degrees of freedom
AIC: 10

Number of Fisher Scoring iterations: 25

> |
```

```
Confusion Matrix and Statistics

          predicted
            0 1
          0 4 0
          1 1 3


               Accuracy : 0.875
                 95% CI : (0.4735, 0.9968)
    No Information Rate : 0.625
    P-Value [Acc > NIR] : 0.135

                  Kappa : 0.75

 Mcnemar's Test P-Value : 1.000

            Sensitivity : 0.800
            Specificity : 1.000
         Pos Pred Value : 1.000
         Neg Pred Value : 0.750
             Prevalence : 0.625
         Detection Rate : 0.500
   Detection Prevalence : 0.500
      Balanced Accuracy : 0.900

       'Positive' Class : 0
```

Figure H- 1: BLR Models Result

2. Screenshot of ANN Implementation on the 30 transformers

```
                       Length Class      Mode
call                    10    -none-     call
response                22    -none-     numeric
covariate               88    -none-     numeric
model.list               2    -none-     list
err.fct                  1    -none-     function
act.fct                  1    -none-     function
linear.output            1    -none-     logical
data                     5    data.frame list
exclude                  0    -none-     NULL
net.result               5    -none-     list
weights                  5    -none-     list
generalized.weights      5    -none-     list
startweights             5    -none-     list
result.matrix          145    -none-     numeric
>
```



Error: 0.018555 Steps: 66

```
Confusion Matrix and Statistics

          predicted
            0 1
          0 4 0
          1 0 4
```

```
               Accuracy : 1
                 95% CI : (0.6306, 1)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : 0.003906

                  Kappa : 1

 Mcnemar's Test P-Value : NA

            Sensitivity : 1.0
            Specificity : 1.0
         Pos Pred Value : 1.0
         Neg Pred Value : 1.0
             Prevalence : 0.5
         Detection Rate : 0.5
   Detection Prevalence : 0.5
      Balanced Accuracy : 1.0

       'Positive' Class : 0
```

Figure H- 2: ANN Model Results

3. Screenshot of SVM Implementation on the 30 transformers

```
Call:
svm(formula = as.factor(ytrain) ~ ., data = xtrain[,
    2:5], cross = 10, probabilty = TRUE, scale = TRUE)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  13

 ( 8 5 )


Number of Classes:  2

Levels:
 0 1

10-fold cross-validation on training data:
```

```
          Confusion Matrix and Statistics

             y_pred
               0 1
            0 5 0
            1 1 3


                  Accuracy : 0.8889
                    95% CI : (0.5175, 0.9972)
       No Information Rate : 0.6667
       P-Value [Acc > NIR] : 0.1431

                     Kappa : 0.7692

    Mcnemar's Test P-Value : 1.0000

               Sensitivity : 0.8333
               Specificity : 1.0000
            Pos Pred Value : 1.0000
            Neg Pred Value : 0.7500
                Prevalence : 0.6667
            Detection Rate : 0.5556
      Detection Prevalence : 0.5556
         Balanced Accuracy : 0.9167

          'Positive' Class : 0
```

Figure H- 3: SVM Model Result.

4. Screenshot of NB Implementation on the 30 transformers

```
Confusion Matrix and Statistics

   y_pred
     0 1
  0 8 1
  1 0 3
```

```
               Accuracy : 0.9167
                 95% CI : (0.6152, 0.9979)
    No Information Rate : 0.6667
    P-Value [Acc > NIR] : 0.05395

                  Kappa : 0.8

 Mcnemar's Test P-Value : 1.00000

            Sensitivity : 1.0000
            Specificity : 0.7500
         Pos Pred Value : 0.8889
         Neg Pred Value : 1.0000
             Prevalence : 0.6667
         Detection Rate : 0.6667
   Detection Prevalence : 0.7500
      Balanced Accuracy : 0.8750

       'Positive' Class : 0
```

Figure H- 4: NB Model Result

5. Implementation of Multiple ML models for condition Modelling using the Health Index on Electrical Utility Data

This appendix contains all the tables used for the HI scoring method for chapter 8. The tables are explained more thoroughly in chapter 8.

- **SVM Model Building Experiments and Testing**

Due to the dataset Imbalance problem, an oversampling technique is used for the minority classes by stimulating more data. So that the predictive model can predict the cases of the minority class correctly, the count of the real dataset for under 69kV transformers includes both records for 66 And 33kV transformers since they apply the same rules of IEC and IEEE standards.

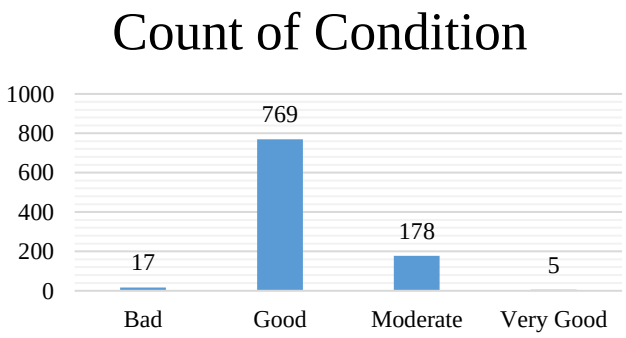


Figure H- 5: Original Count NPG records based on condition.

- **Experiment 1**

In this experiment, it was to use fuzzy rules to specify the ranges of input parameters and the output class of the stimulated data for Very Bad, Bad, Moderate/Fair and Very Good classes.

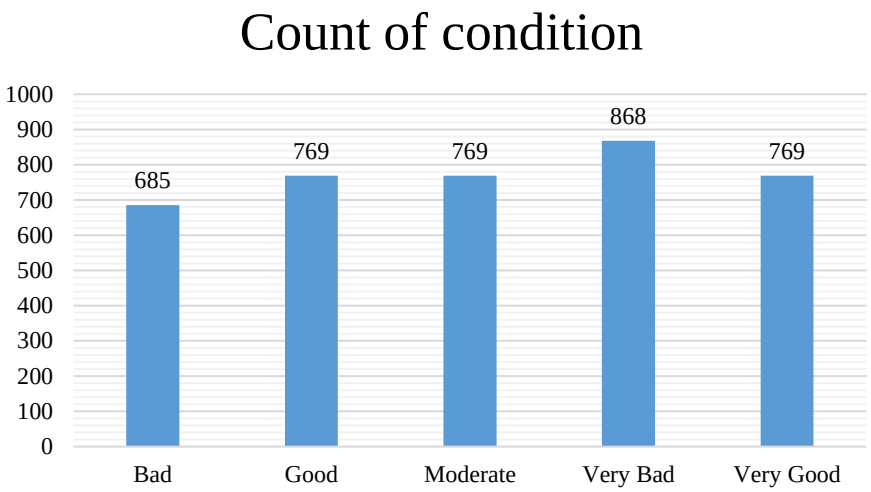


Figure H- 6: Number of electrical utility records stimulated-based fuzzy rules after adding stimulated data.

The stimulated data was saved as CSV and then used as input to an SVM predictive model built in R, which resulted in a low accuracy of 24%, which is not accepted. The below figure shows a summary of the SVM predictor model.

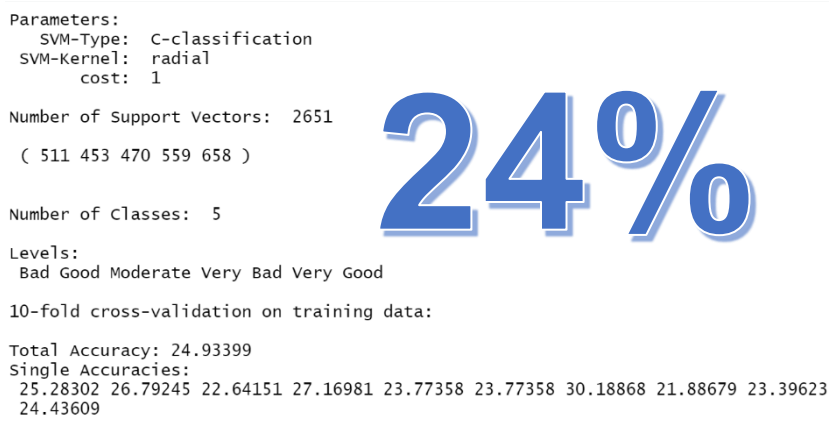


Figure H- 7: Output of the SVM model showing accuracy.

- **Experiment 2**

As a result of the low accuracy of experiment 1, it requires double-checking the output columns using the model built in Excel earlier, which resulted in a different output. The output parameters were considered from the Excel ranking model as these proved reliable earlier by comparing them with the Electrical Utility Health Index results and the findings with other journal papers. Then, the newly updated dataset was merged with the actual dataset using the file named DataSetUnder69kVUpdate.csv.

```
#Loading the data set

TCAdata<-read.csv("DataSetUnder69kVUpdate.csv",sep=",",header=T)
```

Figure H- 8: Using the DataSetUnder69kVUpdate.csv file, reading through the CSV in Rstudio

```
> str(TCAdata)
'data.frame':   3560 obs. of  16 variables:
 $ Health_Index      : num  79 71 66.1 66.1 76.6 ...
 $ condition         : chr  "Good" "Good" "Moderate" "Moderate" ...
 $ Maintenance_Action: chr  "Inspection" "Inspection" "Repair & Refurbis
hment" "Repair & Refurbishment" ...
 $ H2                : chr  "7" "30.1" "6.7" "6.3" ...
 $ CH4               : num  1.38 2.21 1.59 1.59 3.44 3.99 1.43 1.55 1.86
 2.36 ...
 $ CO                : num  37 166 107 86 165 216 67 64 80 86 ...
 $ CO2               : num  569 1271 1070 1361 2356 ...
 $ C2H4              : num  1.58 2.13 1.35 1.69 6.12 7.25 1.24 1.26 0.57
 0.96 ...
 $ C2H6              : num  0.67 1.42 0.4 0.5 4.61 4.07 0.57 0.56 1.97
 5.94 ...
 $ C2H2              : num  0.74 0.59 0.01 0.23 0.18 0.01 0.01 0.01 0.01
 0.01 ...
 $ Moisture          : num  14 17 17 15 19 18 17 20 11 16 ...
 $ BDV               : num  73 74 73 74 52 57 58 54 71 62 ...
 $ DF                : num  0.039 0.113 0.212 0.255 0.151 0.19 0.047 0.0
57 0.033 0.045 ...
 $ Acidity           : num  0.06 0.07 0.2 0.21 0.11 0.14 0.1 0.11 0.01
```

Figure H- 9: Understanding the structure of the data frame using the str() function

```
> head(TCAdata)
  Health_Index condition     Maintenance_Action   H2  CH4  CO  CO2
1     79.03226      Good             Inspection    7 1.38  37  569
2     70.96774      Good             Inspection 30.1 2.21 166 1271
3     66.12903  Moderate Repair & Refurbishment  6.7 1.59 107 1070
4     66.12903  Moderate Repair & Refurbishment  6.3 1.59  86 1361
5     76.61290      Good             Inspection 10.8 3.44 165 2356
6     76.61290      Good             Inspection  9.7 3.99 216 2199
  C2H4 C2H6 C2H2 Moisture BDV    DF Acidity Age Furan
1 1.58 0.67 0.74       14  73 0.039    0.06  57    NA
2 2.13 1.42 0.59       17  74 0.113    0.07  57    NA
3 1.35 0.40 0.01       17  73 0.212    0.20  45   0.6
4 1.69 0.50 0.23       15  74 0.255    0.21  45   0.8
5 6.12 4.61 0.18       19  52 0.151    0.11  60    NA
6 7.25 4.07 0.01       18  57 0.190    0.14  60    NA
```

Figure H- 10: The above figure shows the first six records of the data frame using the head() function

Understanding the structure of the data frame is essential so we can remove the unwanted columns and prepare the data for modelling. Creating a data frame utilising only the columns we need for the prediction model and saving it as TCAdf.

```
TCAdf<-data.frame(condition=TCAdata$condition,
                  Moisture=TCAdata$Moisture,
                  Acidity=TCAdata$Acidity,
                  BDV=TCAdata$BDV,
                  DF=TCAdata$DF,
                  H2=TCAdata$H2,
                  CH4=TCAdata$CH4,
                  C2H6=TCAdata$ C2H6,
                  C2H4=TCAdata$C2H4,
                  C2H2=TCAdata$C2H2,
                  CO=TCAdata$CO,
                  CO2=TCAdata$CO2,
                  Furan=TCAdata$Furan,
                  Age=TCAdata$Age)
```

Figure H- 11: Creating a data frame

```
> str(TCAdf)
'data.frame':   3560 obs. of  14 variables:
 $ condition: chr  "Good" "Good" "Moderate" "Moderate" ...
 $ Moisture : num  14 17 17 15 19 18 17 20 11 16 ...
 $ Acidity  : num  0.06 0.07 0.2 0.21 0.11 0.14 0.1 0.11 0.01 0.01 ...
 $ BDV      : num  73 74 73 74 52 57 58 54 71 62 ...
 $ DF       : num  0.039 0.113 0.212 0.255 0.151 0.19 0.047 0.057 0.033
0.045 ...
 $ H2       : chr  "7" "30.1" "6.7" "6.3" ...
 $ CH4      : num  1.38 2.21 1.59 1.59 3.44 3.99 1.43 1.55 1.86 2.36 ...
 $ C2H6     : num  0.67 1.42 0.4 0.5 4.61 4.07 0.57 0.56 1.97 5.94 ...
 $ C2H4     : num  1.58 2.13 1.35 1.69 6.12 7.25 1.24 1.26 0.57 0.96 ...
 $ C2H2     : num  0.74 0.59 0.01 0.23 0.18 0.01 0.01 0.01 0.01 0.01 ...
 $ CO       : num  37 166 107 86 165 216 67 64 80 86 ...
 $ CO2      : num  569 1271 1070 1361 2356 ...
 $ Furan    : num  NA NA 0.6 0.8 NA NA NA NA 0.04 0.04 ...
 $ Age      : num  57 57 45 45 60 60 57 57 15 15 ...
```

Figure H- 12: Checking the data frame structure and data types.

```
TCAdataC <- as.data.frame(unclass(TCAdf),
                          stringsAsFactors = TRUE)
```

Figure H- 13: Convert the columns to factor so the SVM classifier can handle them.

A part of understanding the data frame is using the summary function to find the minimum and maximum mean for each column and how much data is missing from each column, as shown in the figure below. The furan has the most columns with missing data of 265. Whereas other columns like DF have 23 NA's, Acidity and BDV have 9 NA's, filled by the column's mean value.

```
> summary(TCAdata)
  Health_Index    condition         Maintenance_Action
 Min.   :15.32   Length:3560        Length:3560
 1st Qu.:38.71   Class :character   Class :character
 Median :62.10   Mode  :character   Mode  :character
 Mean   :51.57
 3rd Qu.:68.55
 Max.   :91.94

      H2                 CH4                 CO
 Length:3560        Min.   :   0.0427   Min.   :   1.353
 Class :character   1st Qu.:   4.7417   1st Qu.: 113.846
 Mode  :character   Median :  62.3886   Median : 313.547
                    Mean   : 213.8219   Mean   :1496.253
                    3rd Qu.: 198.5339   3rd Qu.: 892.526
                    Max.   :2207.0000   Max.   :9991.405

      CO2                C2H4               C2H6
 Min.   :    0.01   Min.   :   0.010   Min.   :  0.000
 1st Qu.: 1005.69   1st Qu.:   6.645   1st Qu.:  3.803
 Median : 2129.80   Median :  43.122   Median :  50.596
 Mean   : 3150.03   Mean   : 164.216   Mean   :163.889
 3rd Qu.: 3995.86   3rd Qu.:  99.549   3rd Qu.:100.341
 Max.   :12210.00   Max.   :4230.000   Max.   :999.946
```

```
      C2H2              Moisture             BDV
 Min.   :  0.0094   Min.   :    0.25   Min.   :  0.026
 1st Qu.:  0.4317   1st Qu.:   24.00   1st Qu.: 44.056
 Median :  3.5342   Median :   51.68   Median : 70.000
 Mean   :136.9799   Mean   :   60.96   Mean   : 85.726
 3rd Qu.: 34.3417   3rd Qu.:   76.07   3rd Qu.:121.692
 Max.   :994.9239   Max.   :17366.00   Max.   :610.000
                                       NA's   :9

       DF                Acidity              Age
 Min.   : 0.00004   Min.   :  0.00002   Min.   :  0.01362
 1st Qu.: 0.12100   1st Qu.:  0.14000   1st Qu.: 13.13593
 Median : 6.90453   Median :  3.50882   Median : 45.68978
 Mean   :26.83645   Mean   : 25.65603   Mean   : 43.03884
 3rd Qu.:53.11472   3rd Qu.: 50.52656   3rd Qu.: 59.00000
 Max.   :99.97525   Max.   : 99.99730   Max.   :119.89017
 NA's   :23         NA's   :9
     Furan
 Min.   :0.00024
 1st Qu.:0.04000
 Median :0.09099
 Mean   :0.34678
 3rd Qu.:0.69374
 Max.   :6.20000
 NA's   :265
```

Figure H- 14: Checking the data summary

```
TCAdataC$Moisture[is.na(TCAdataC$Moisture)] <- mean(TCAdataC$Moisture, na.rm = TRUE)
TCAdataC$Acidity[is.na(TCAdataC$Acidity)] <- mean(TCAdataC$Acidity, na.rm = TRUE)
TCAdataC$BDV[is.na(TCAdataC$BDV)] <- mean(TCAdataC$BDV, na.rm = TRUE)
TCAdataC$DF[is.na(TCAdataC$DF)] <- mean(TCAdataC$DF, na.rm = TRUE)
```

Figure H- 15: Replacing the null value with the column's average

The dataset forms into a training set and a testing set. The figure shows those dimensions, which apply 60% of the data as a training set.

```
> dim(xtrain)
[1] 2136   14
> dim(xtest)
[1] 1424   14
```

Figure H- 16: Checking the data dimensions

Building an SVM Classifier to predict the condition of the transformer involved categorizing the data into five classes: Bad, Moderate, Good and Very Good. It has 414 support vectors and obtained an accuracy of 79%. The below figure shows the model summary.

```
Call:
svm(formula = as.factor(ytrain) ~ ., data = xtrain[, 2:14],
    cross = 10, probability = TRUE, scale = FALSE)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1
```

```
Number of Support Vectors:  1672

 ( 243 355 473 582 19 )


Number of Classes:  5

Levels:
 Bad Good Moderate Very Bad Very Good

10-fold cross-validation on training data:

Total Accuracy: 79.43118
Single Accuracies:
 81.12245 77.15736 74.61929 80.71066 78.17259 81.72589 79.18782 83.24873
 80.71066 77.66497
```

79%

Figure H- 17: Output of the SVM model showing the accuracy after experiment 2.

Total Accuracy is 79%, as shown in the above figure. In order to improve the accuracy, an investigation that experiments the class “Very Good” had only 26 records, which is a low number compared to other classes.

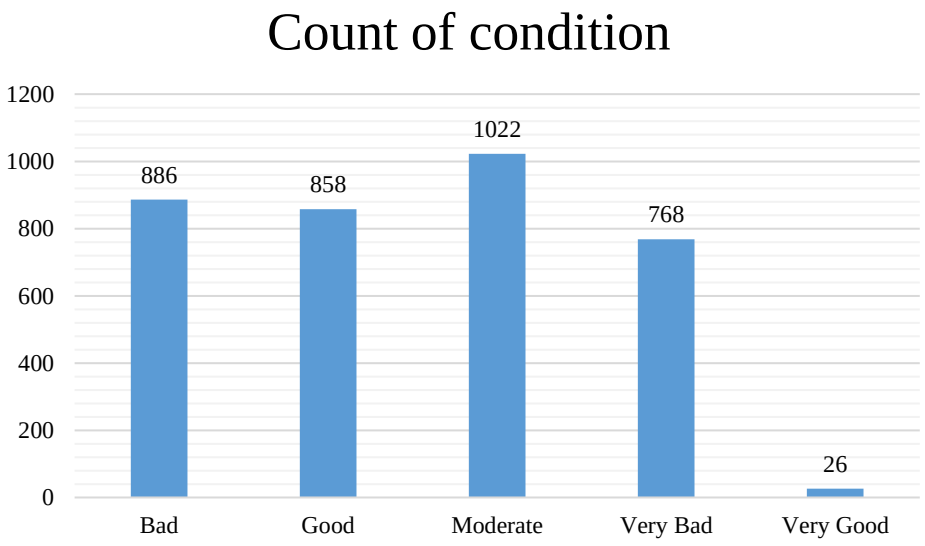


Figure H- 18: Number of the records per class/condition.

- **Experiment 3**

Add more records to the “Very Good” Class to balance the dataset, then save the model as svm_ConditionClassifier.

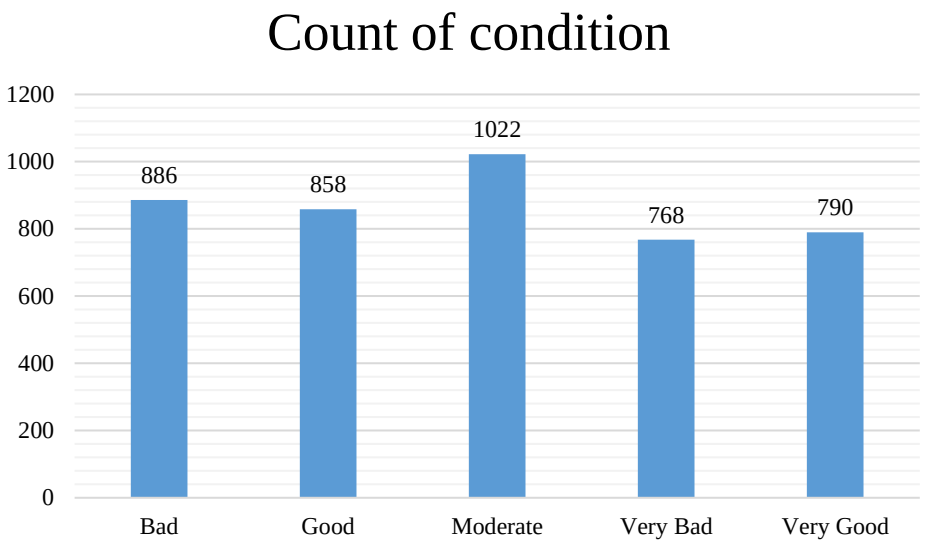


Figure H- 19: Numer of electrical utility records with the stimulated.

```
> str(TCAdataC)
'data.frame':   4324 obs. of  14 variables:
 $ condition: Factor w/ 5 levels "Bad","Good","Moderate",..: 2 2 3 3 2 2
 2 2 2 2 ...
 $ Moisture : num  14 17 17 15 19 18 17 20 11 16 ...
 $ Acidity  : num  0.06 0.07 0.2 0.21 0.11 0.14 0.1 0.11 0.01 0.01 ...
 $ BDV      : num  73 74 73 74 52 57 58 54 71 62 ...
 $ DF       : num  0.039 0.113 0.212 0.255 0.151 0.19 0.047 0.057 0.033
 0.045 ...
 $ H2       : Factor w/ 3782 levels "0.001984648",..: 2858 1620 2564 255
1 86 3494 3482 116 3478 2177 ...
 $ CH4      : num  1.38 2.21 1.59 1.59 3.44 3.99 1.43 1.55 1.86 2.36 ...
 $ C2H6     : num  0.67 1.42 0.4 0.5 4.61 4.07 0.57 0.56 1.97 5.94 ...
 $ C2H4     : num  1.58 2.13 1.35 1.69 6.12 7.25 1.24 1.26 0.57 0.96 ...
 $ C2H2     : num  0.74 0.59 0.01 0.23 0.18 0.01 0.01 0.01 0.01 0.01 ...
 $ CO       : num  37 166 107 86 165 216 67 64 80 86 ...
 $ CO2      : num  569 1271 1070 1361 2356 ...
 $ Furan    : num  NA NA 0.6 0.8 NA NA NA NA 0.04 0.04 ...
 $ Age      : num  57 57 45 45 60 60 57 57 15 15 ...
```

Figure H- 20: Structure of the data frame.

```
> summary(TCAdataC)
    condition       Moisture          Acidity
 Bad      : 886   Min.   :    0.25   Min.   :  0.00001
 Good     : 858   1st Qu.:   39.58   1st Qu.:  0.05333
 Moderate :1022   Median :   56.61   Median :  0.66931
 Very Bad : 768   Mean   :   61.03   Mean   : 21.11877
 Very Good: 790   3rd Qu.:   70.39   3rd Qu.: 40.39552
                  Max.   :17366.00   Max.   : 99.99730
                                     NA's   :9
      BDV                DF                 H2
 Min.   :  0.026   Min.   : 0.00002   10.8   :  12
 1st Qu.: 32.877   1st Qu.: 0.05200   11.6   :  11
 Median : 60.000   Median : 0.16600   8.9    :  11
 Mean   : 72.231   Mean   :22.07274   10     :  10
 3rd Qu.:104.178   3rd Qu.:43.55768   11.8   :  10
 Max.   :610.000   Max.   :99.97525   7.9    :   9
 NA's   :9         NA's   :23         (Other):4261
```

```
      CH4                C2H6              C2H4
 Min.   :   0.0427   Min.   :  0.000   Min.   :   0.01
 1st Qu.:   6.1660   1st Qu.:  4.953   1st Qu.:  11.26
 Median :  38.2150   Median : 29.278   Median :  29.12
 Mean   : 179.6049   Mean   :137.552   Mean   : 138.74
 3rd Qu.: 173.5358   3rd Qu.: 94.192   3rd Qu.:  93.88
 Max.   :2207.0000   Max.   :999.946   Max.   :4230.00

      C2H2                CO                CO2              Furan
 Min.   :  0.0055   Min.   :   0.024   Min.   :    0.01   Min.   :0.00024
 1st Qu.:  0.4579   1st Qu.:  70.000   1st Qu.:  708.65   1st Qu.:0.04000
 Median :  2.0775   Median : 211.953   Median : 1544.92   Median :0.07698
 Mean   :112.9518   Mean   :1240.908   Mean   : 2678.88   Mean   :0.29132
 3rd Qu.: 24.7794   3rd Qu.: 827.321   3rd Qu.: 3700.48   3rd Qu.:0.61982
 Max.   :994.9239   Max.   :9991.405   Max.   :12210.00   Max.   :6.20000
                                                          NA's   :265
      Age
 Min.   :  0.01362
 1st Qu.:  7.87570
 Median : 21.00000
 Mean   : 36.32171
 3rd Qu.: 56.08985
 Max.   :119.89017
```

Figure H- 21: Summary of the data frame columns

```
> dim(xtrain)
[1] 2594   14
> dim(xtest)
[1] 1730   14
```

Figure H- 22: Dimension of training and testing set.

Summary of svm_ConditionClassifier, which, as shown, has 1755 support vectors and is used to predict five classes, achieved an accuracy of 83%, which is quite good and acceptable.

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  1755

 ( 137 239 327 582 470 )


Number of Classes:  5

Levels:
 Bad Good Moderate Very Bad Very Good

10-fold cross-validation on training data:

Total Accuracy: 83.95568
Single Accuracies:
 84.36214 84.01639 84.42623 84.77366 83.60656 86.47541 81.06996 83.60656 81.9
6721 85.2459
```

83%

Figure H- 23: Output of the SVM second model showing the accuracy after experiment 3.

- **Experiment 4**

Following the previous experiment, another predictor classifier considers the available dataset for the same machine-learning algorithm. This classifier intends to output the maintenance action of 3 classes: Inspection, Repair and Refurbishment, and Replacement, based on the specified condition. This experiment was successful, with an accuracy of 86%.

86%

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  1721

 ( 343 796 582 )


Number of Classes:  3

Levels:
 Inspection Repair & Refurbishment Replacement

10-fold cross-validation on training data:

Total Accuracy: 86.78703
Single Accuracies:
 86.00823 86.47541 84.83607 87.65432 85.65574 87.29508 87.65432 87.70492 84.83607
 89.7541
```

Figure H- 24: Output of the SVM first model for the accuracy during experiment 4.

After using scaling to True and specifying 20 in the cross, the model accuracy improved to 93%

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  470

 ( 140 162 168 )


Number of Classes:  3

Levels:
 Inspection Repair & Refurbishment Replacement

20-fold cross-validation on training data:

Total Accuracy: 93.43186
Single Accuracies:
 96.69421 93.44262 95.90164 90.98361 91.80328 91.73554 94.2623 96.72
131 90.98361 93.44262 90.08264 95.08197 90.98361 95.08197 94.2623 9
5.04132 91.80328 91.80328 95.90164 92.62295
```

Figure H- 25: Output of the SVM first model showing the final accuracy during experiment 4.

Later on, when executing the SVM condition classifier model again, the error below was shown.

```
>
> ## Train SVM on data using e1071 package
> ConditionClassifier <- e1071::svm(as.factor(ytrain)~ ., data=xtrain[,2:14],
+                                   scale=TRUE,cross=20,probability=TRUE)
Error in if (any(co)) { : missing value where TRUE/FALSE needed
```

Figure H- 26: Error screen during experiment 4.

A solution was to ensure that the pre-processing of the data was appropriately done by applying the normalisation function to each column and replacing empty cells with

0, which solved the above problem. The model accuracy now becomes **93%,** as shown below.

```
Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  1

Number of Support Vectors:  549

 ( 48 132 146 165 58 )


Number of Classes:  5

Levels:
 Bad Good Moderate Very Bad Very Good

20-fold cross-validation on training data:

Total Accuracy: 93.22938
Single Accuracies:
 94.21488 93.44262 90.16393 92.62295 95.08197 90.98361 92.56198
 93.44262 92.62295 94.2623 95.90164 95.90164 90.98361 93.38843
 92.62295 90.16393 91.80328 96.72131 92.62295 95.08197
```

93%

Figure H- 27: Final SVM model accuracy after experiment 4.

6. ML Modelling Experiments:

This section focuses on performance evaluation using quantitative measures. It evaluates machine learning models using statistics measures in R studio. The following describes the evaluation measures and their calculation, including a 2x2 confusion matrix table with the following representation.

		Prediction	
		Transformer Faulty	Transformer Healthy
Actual	Transformer Faulty	TP	FN
	Transformer Healthy	FP	TN

Figure H- 28: Illustration of confusion Matrix

Table H- 10: The formulas for the quantitative measures (Caret: Classification and regression training - the comprehensive R, 2023).

Quantitative Measure	Formula
Sensitivity	$\frac{TP}{(TP + FP)}$
Specificity	$\frac{TN}{FN + TN}$
Prevalence	$\frac{(TP + FP)}{(TP + FP + TN + FN)}$
Pos Pred Value	$\frac{(Senstivity * Prevalence)}{((Sensitivty * Prevalence) + ((1 - Specifity) * (1 - Prevalence)))}$
Neg Pred Value	$\frac{(Specifity * (1 - Prevalence))}{(((1 - Sensitivty) * Prevalence) + ((Specifity) * (1 - Prevalence)))}$
Detection Rate	$\frac{TP}{(TP + FP + TN + F)}$
Detection Prevalence	$\frac{(TP + FN)}{(TP + P + TN + FN)}$

Quantitative Measure	Formula
Balanced Accuracy	$\frac{(Sensitivity + Specifity)}{2}$
Precision	$\frac{TP}{(TP + FN)}$
Recall	$\frac{TP}{(TP + FP)}$
F1	$\frac{(1 + beta^2) * Precision * Recall}{((beta^2 * Precision) + Recall)}$

For class Bad, prevalence, detection rate and detection prevalence can be described as follows.

1. Prevalence is the rate of all actual bad transformers in the whole population (training dataset).
2. Detection Rate is the rate of actual bad transformers correctly predicted as bad transformers.
3. Detection Prevalence: the rate of all predicted bad transformers in the whole population (training dataset).

The prevalence of the ANN model is 0.1859, which means that 0.1859 of the training datasets are actual transformers in “Bad” conditions. On the other hand, 0.1809 is detected as bad, which is the detection rate. Also, the detection prevalence is 0.1834, a rate for all predicted “Bad” in the training dataset. Therefore, since all rates are similar, very few errors occurred, which means the model is acceptable.

Since the model is multi-classified and the transformer has multiple condition classes, the results of evaluating each factor level to the remaining levels utilise the one-versus-all approach.

7. Machine Learning Models Implementation in R Programming Language

Binary Logistic Regression (BLR)

The model in R statistical using the glm () function language defines the family as binominal, specifying the regression as binary logistic regression.

MAcFadden metric is to evaluate the fit of the binary logistic regression model. The metric value ranges from 0 to just under 1. Values over 0.40 suggest that a model fits the data very well. On the other hand, values near 0 show that the model has no predictive power (Atmathew, 2015). As seen from the figure, the value of McFadden 0.9 is close to 1, and this indicates that the HI model fits the data very well and has high predictive power the model fits the data very well.

```
fitting null model for pseudo-r2

                  McFadden
0.9999999999829305
```

Figure H- 29: McFadden value of the binary logistic regression model.

Artificial Neural Network (ANN)

The ANN model uses (a neural net) package in R programming, and it uses “RPROP+”, which is described as a direct adaptive method for faster backpropagation learning as it makes a local adaptation of the weight update according to the behaviour of the error function. This algorithm is for supervisory learning in a feedforward Artificial Neural Network.

The model had an error rate of 0.018, which is quite good and did 66 steps to perform. ANN is a network of equations that has four inputs in the KA-HI experimented model, which considers acidity, breakdown voltage, dissolved combustible gases, and furfuraldehyde, and has one output that presents the HI score of 0 and 1, as shown in the figure below. All the weighting factors assist the model adjustments, and the process frequently repeats as many times. The ANN learns by training the algorithm from the available dataset. Once the ANN model has sufficient training, the network will apply to any unsolved problems to identify the HI output.

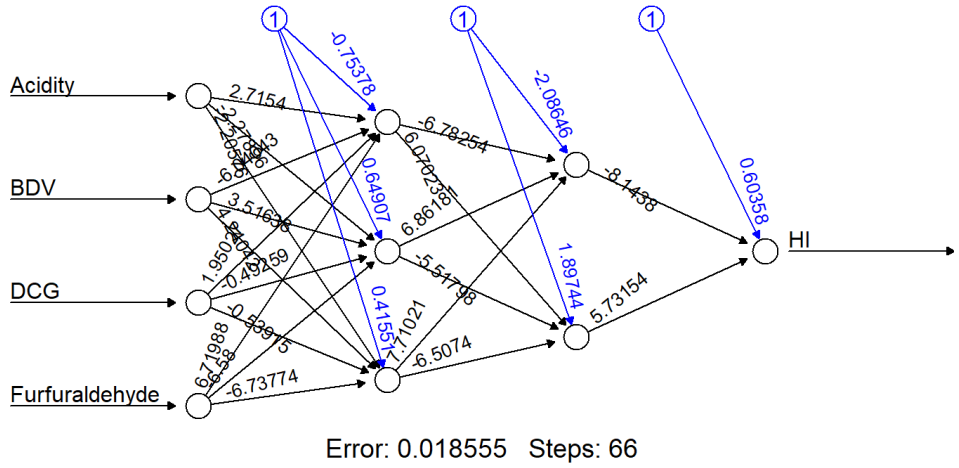


Figure H- 30: Artificial Neural Network model.

- **Result of Machine Learning Modelling on the Electrical Electricity Data**

ANN

Table H- 11: ANN Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	220	0	0	0	3
	Good	0	148	13	0	23
	Moderate	1	13	287	0	13
	Very Bad	0	3	2	226	6
	Very Good	5	0	9	0	244

Table H- 12: ANN Statistics By Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.9735	0.9024	0.9228	1	0.8443
Specificity	0.997	0.9658	0.9702	0.9889	0.9849
Pos Pred Value	0.9865	0.8043	0.914	0.9536	0.9457
Neg Pred Value	0.994	0.9845	0.9734	1	0.953
Precision	0.9865	0.8043	0.914	0.9536	0.9457
Recall	0.9735	0.9024	0.9228	1	0.8443
F1	0.98	0.8506	0.9184	0.9762	0.8921
Prevalence	0.1859	0.1349	0.2558	0.1859	0.2377
Detection Rate	0.1809	0.1217	0.236	0.1859	0.2007
Detection Prevalence	0.1834	0.1513	0.2582	0.1949	0.2122
Balanced Accuracy	0.9852	0.9341	0.9465	0.9944	0.9146

- **ANN model Output**

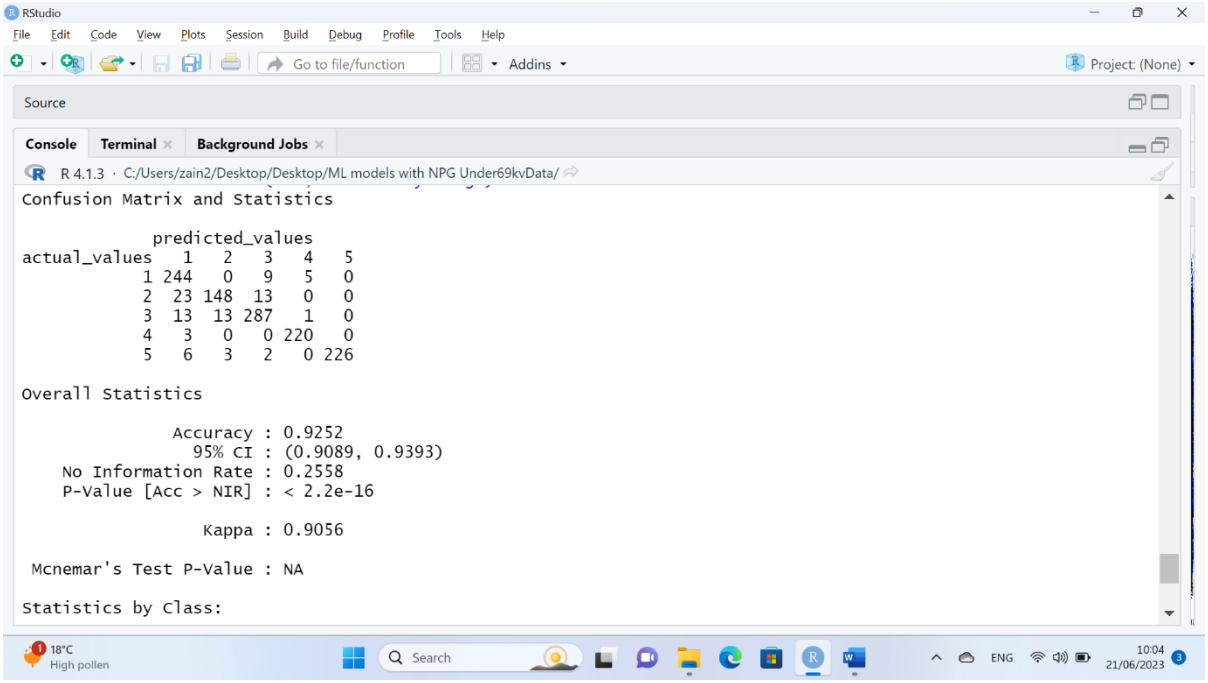


Figure H- 31: Screenshot from R studio showing ANN Model Output (1).

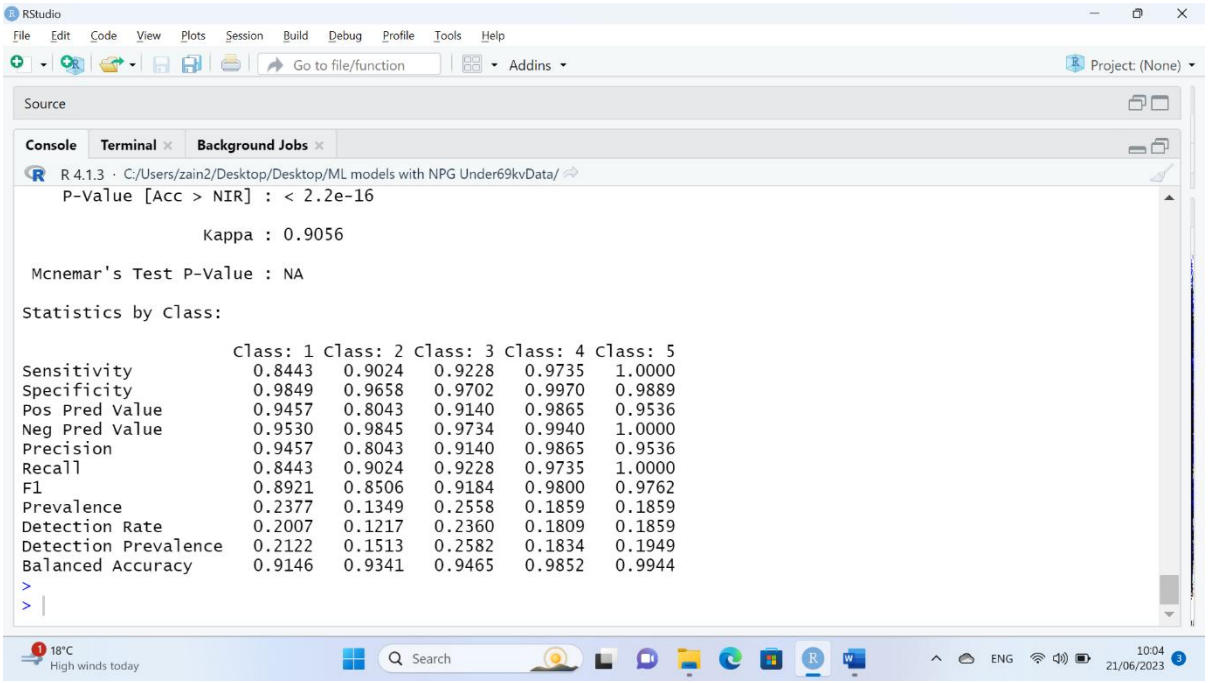


Figure H- 32: Screenshot from R studio showing ANN Model Output (2).

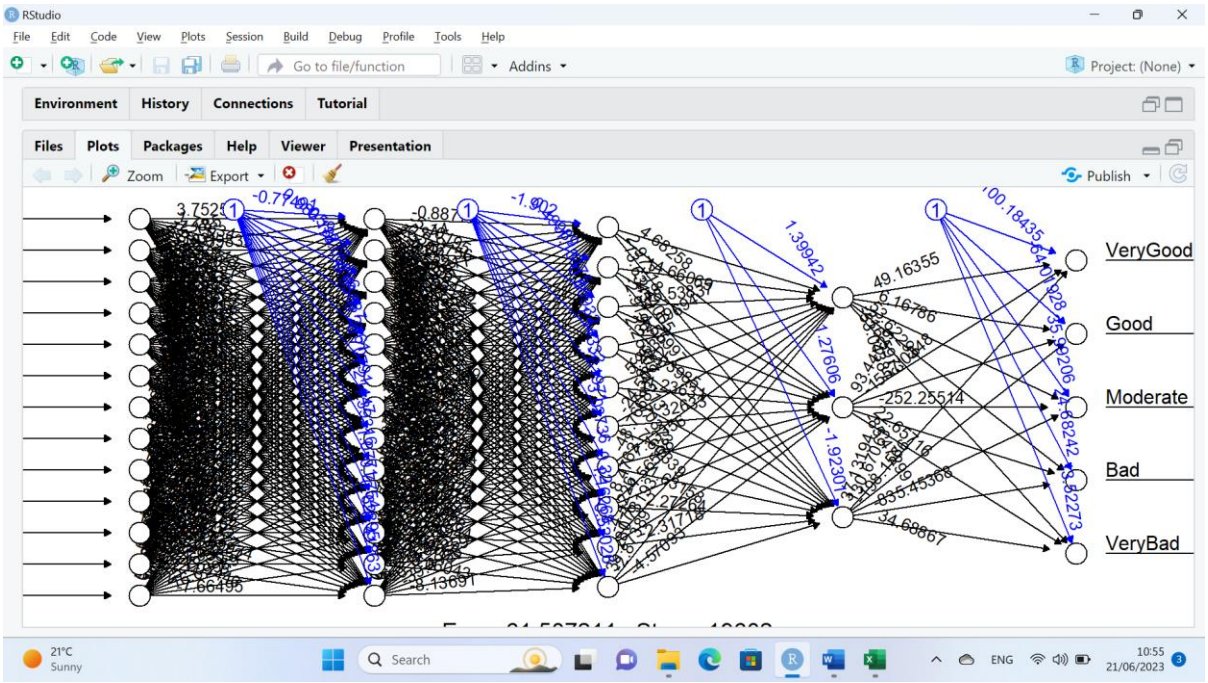


Figure H- 33: Screenshot from R studio showing ANN Model Output (3).

- **SVM Model**

Table H- 13: SVM Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	268	4	5	0	0
	Good	17	242	4	0	2
	Moderate	22	25	265	0	0
	Very Bad	0	0	0	224	0
	Very Good	5	1	0	0	213

Table H- 14: SVM Statistics By Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.859	0.8897	0.9672	1	0.9907
Specificity	0.9909	0.9776	0.9541	1	0.9945
Pos Pred Value	0.9675	0.9132	0.8494	1	0.9726
Neg Pred Value	0.9569	0.9709	0.9909	1	0.9981
Precision	0.9675	0.9132	0.8494	1	0.9726
Recall	0.859	0.8897	0.9672	1	0.9907
F1	0.91	0.9013	0.9044	1	0.9816
Prevalence	0.2406	0.2097	0.2113	0.1727	0.1658
Detection Rate	0.2066	0.1866	0.2043	0.1727	0.1642
Detection Prevalence	0.2136	0.2043	0.2406	0.1727	0.1689
Balanced Accuracy	0.9249	0.9336	0.9606	1	0.9926

- **SVM model Output**

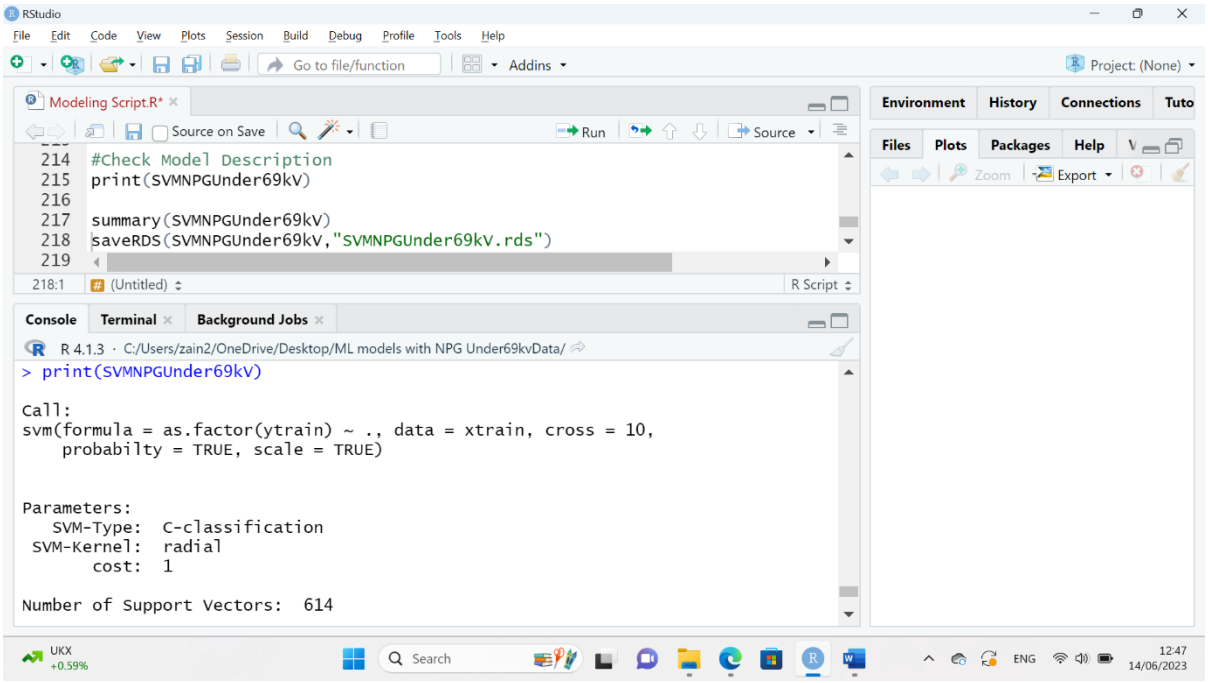


Figure H- 34: Screenshot from R studio showing SVM Model Output (1).

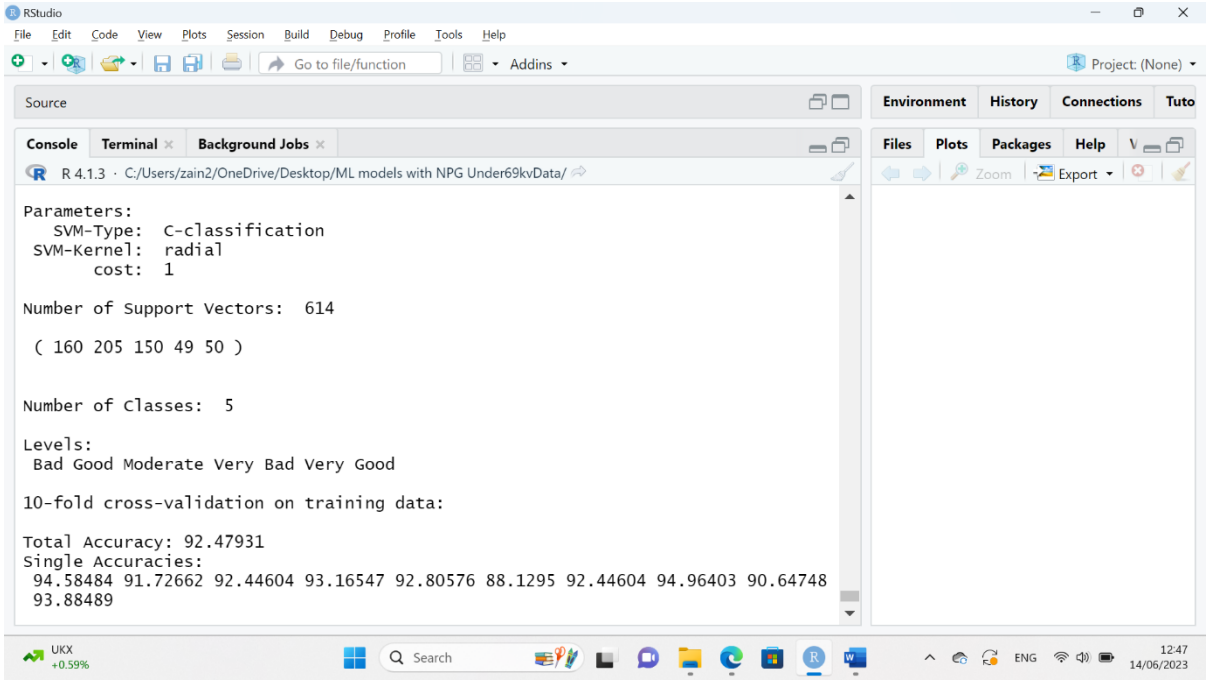


Figure H- 35: Screenshot from R studio showing SVM Model Output (2).

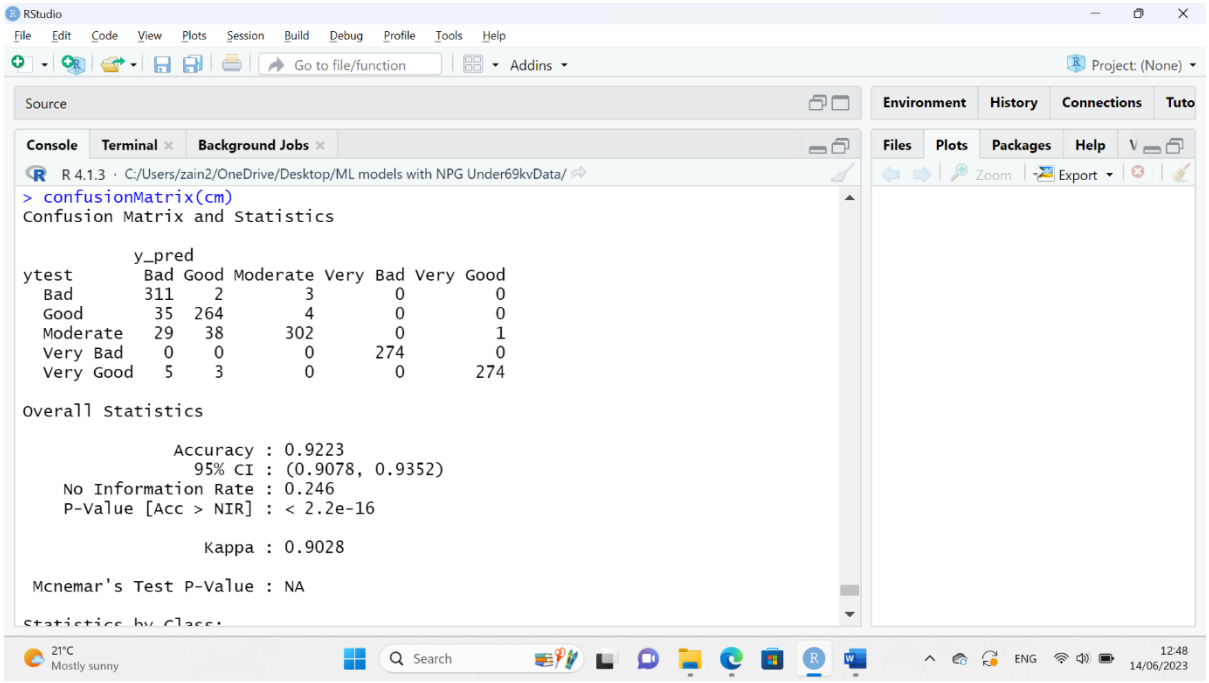


Figure H- 36: Screenshot from R studio showing SVM Model Output (3).

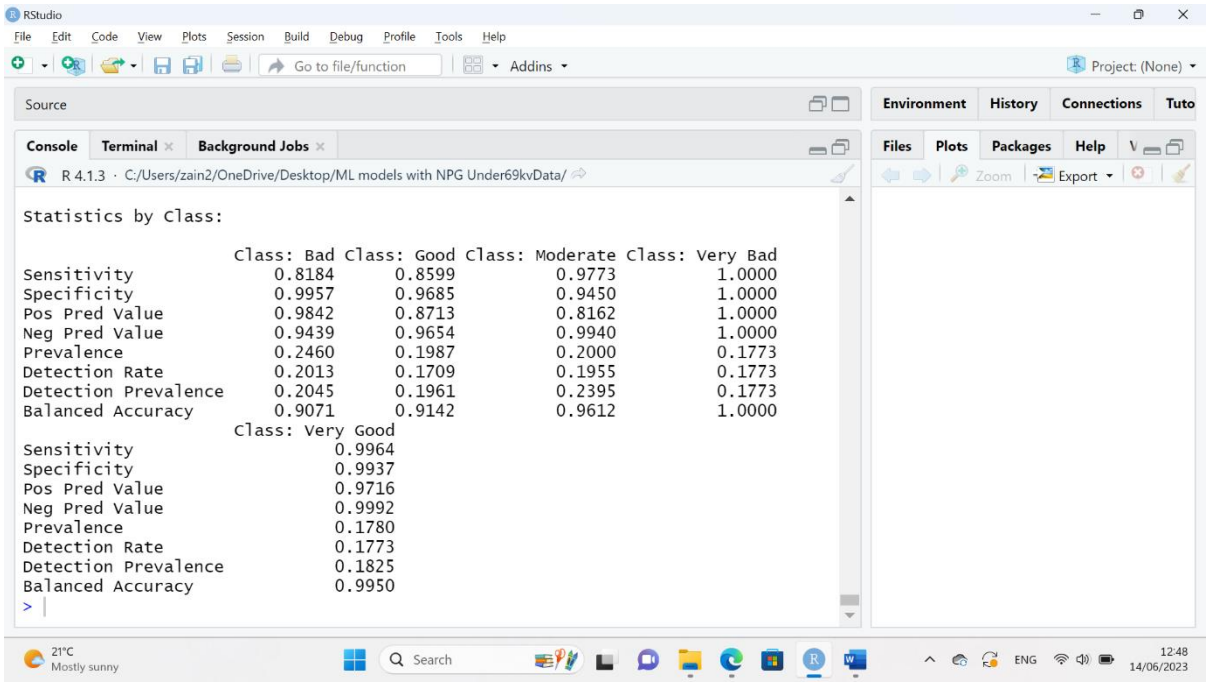


Figure H- 37: Screenshot from R studio showing SVM Model Output (4).

- **Random Forest (RF)**

Table H- 15: RF Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	299	6	11	0	0
	Good	28	263	12	0	0
	Moderate	12	17	338	0	3
	Very Bad	0	0	0	274	0
	Very Good	1	7	0	0	274

Table H- 16: RF Statistics by Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.8794	0.8976	0.9363	1	0.9892
Specificity	0.9859	0.9681	0.973	1	0.9937
Pos Pred Value	0.9462	0.868	0.9135	1	0.9716
Neg Pred Value	0.9666	0.9758	0.9804	1	0.9976
Precision	0.9462	0.868	0.9135	1	0.9716
Recall	0.8794	0.8976	0.9363	1	0.9892
F1	0.9116	0.8826	0.9248	1	0.9803
Prevalence	0.2201	0.1896	0.2337	0.1773	0.1793
Detection Rate	0.1935	0.1702	0.2188	0.1773	0.1773
Detection Prevalence	0.2045	0.1961	0.2395	0.1773	0.1825
Balanced Accuracy	0.9327	0.9328	0.9546	1	0.9914
Balanced Accuracy	0.9345	0.93	0.9558	1	0.9932

- **RF model Output**

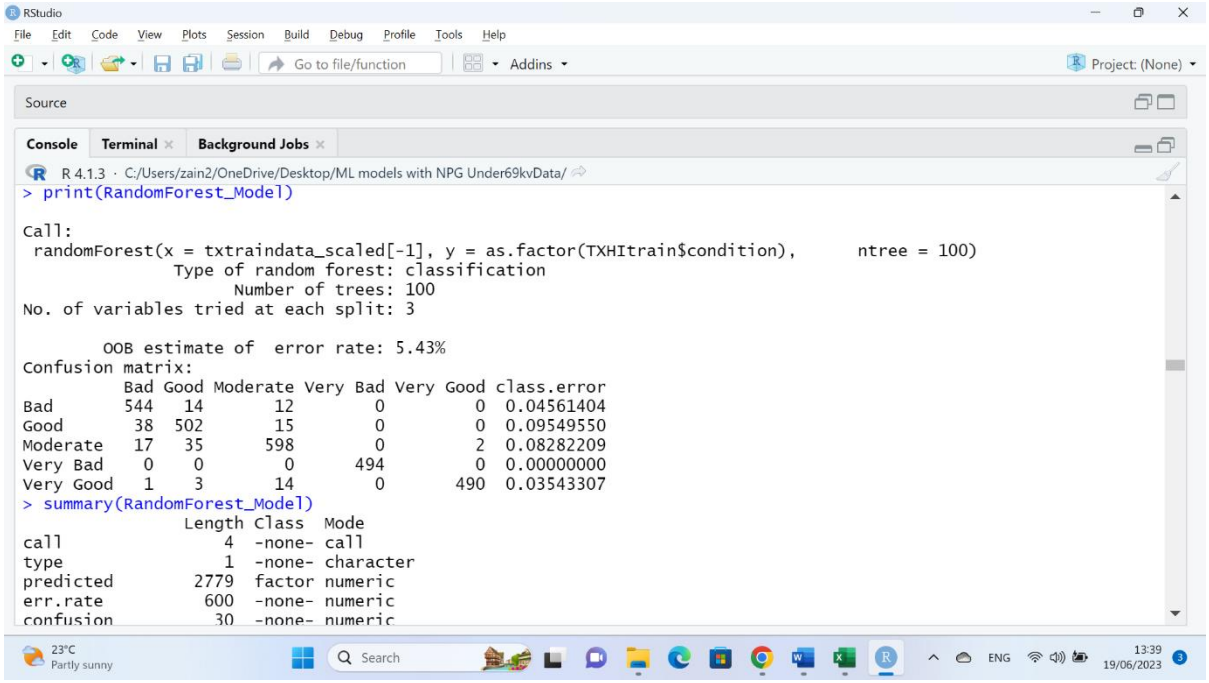


Figure H- 38: Screenshot from R studio showing RF Model Output (1).

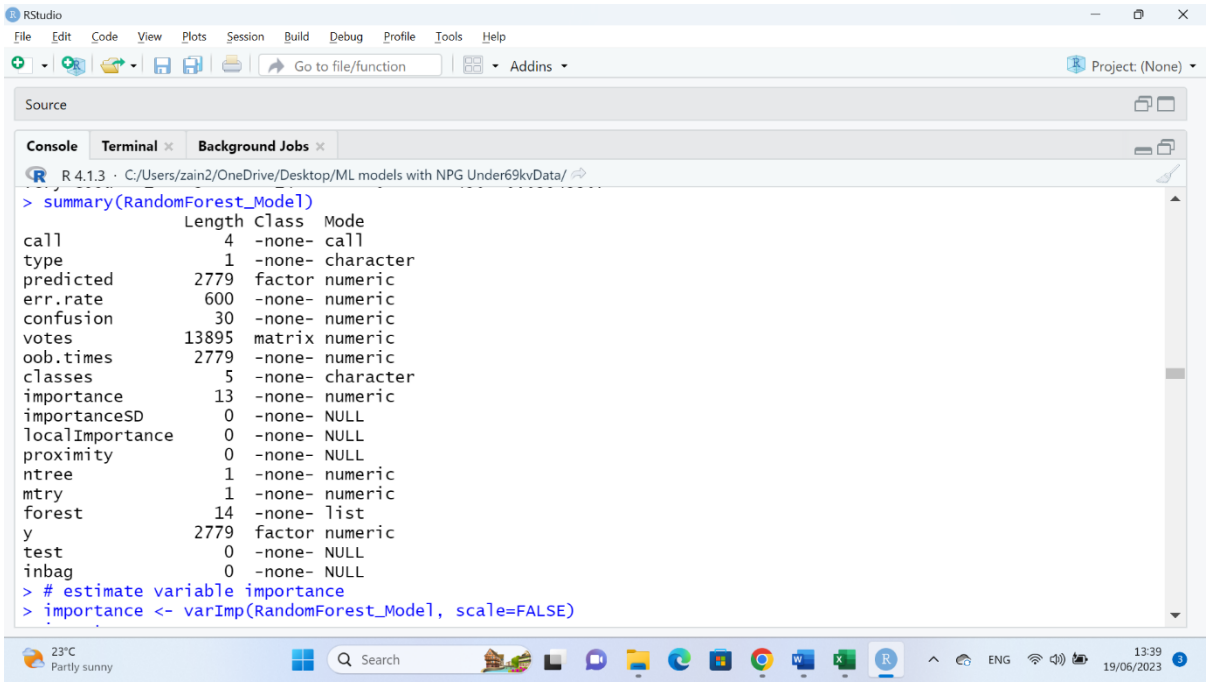


Figure H- 39: Screenshot from R studio showing RF Model Output (2).

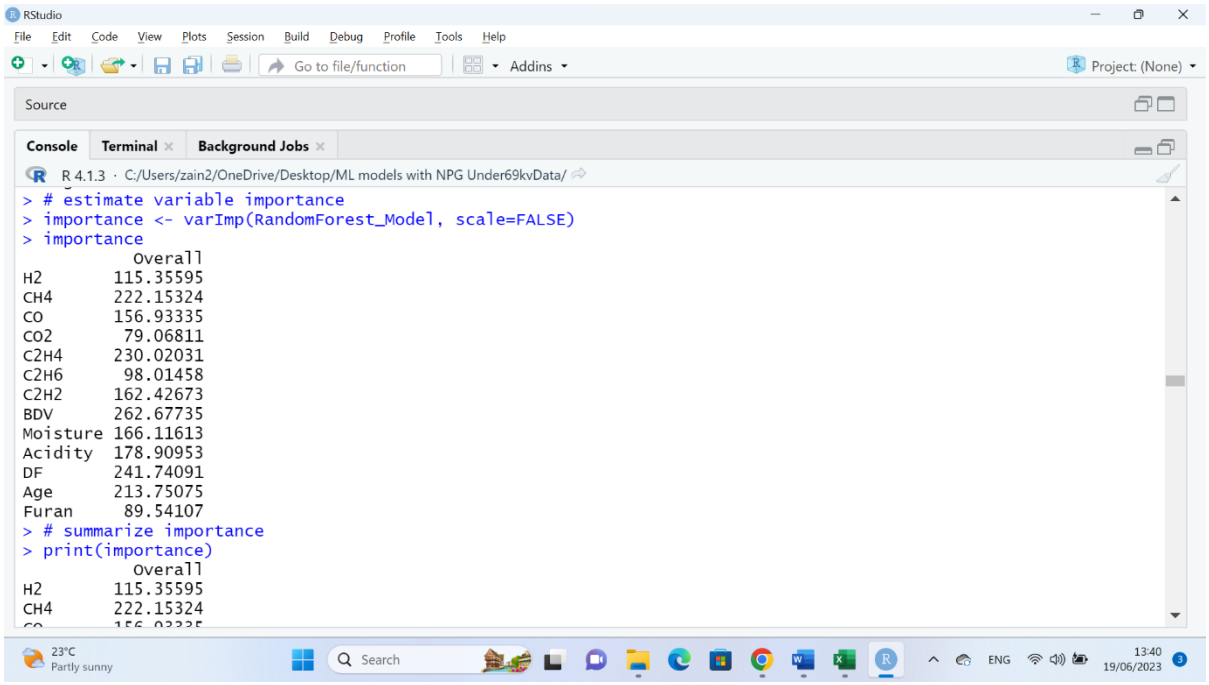


Figure H- 40: Screenshot from R studio showing RF Model Output (3).

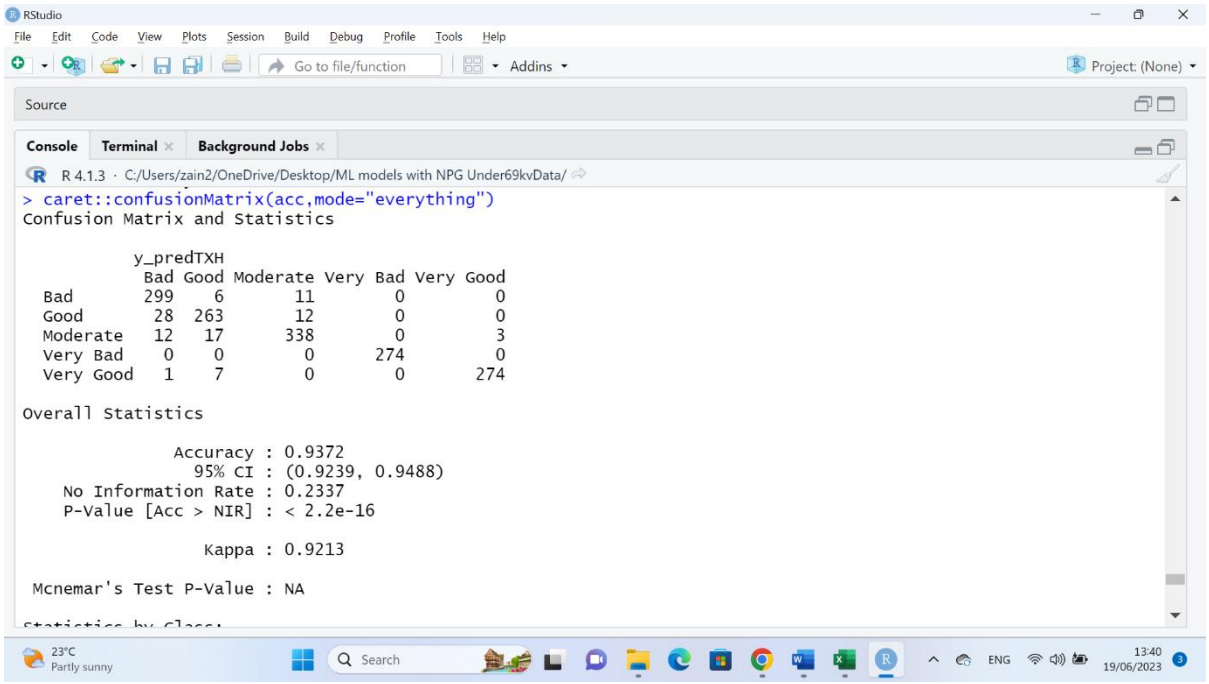


Figure H- 41: Screenshot from R studio showing RF Model Output (4).

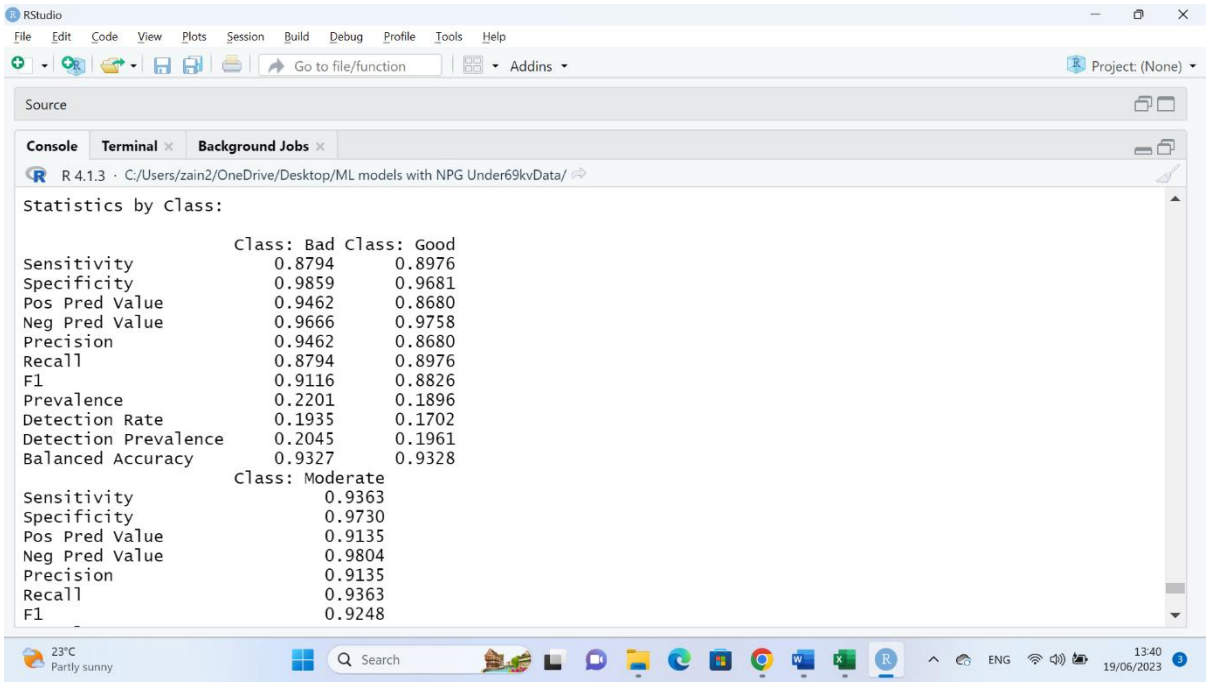


Figure H- 42: Screenshot from R studio showing RF Model Output (5).

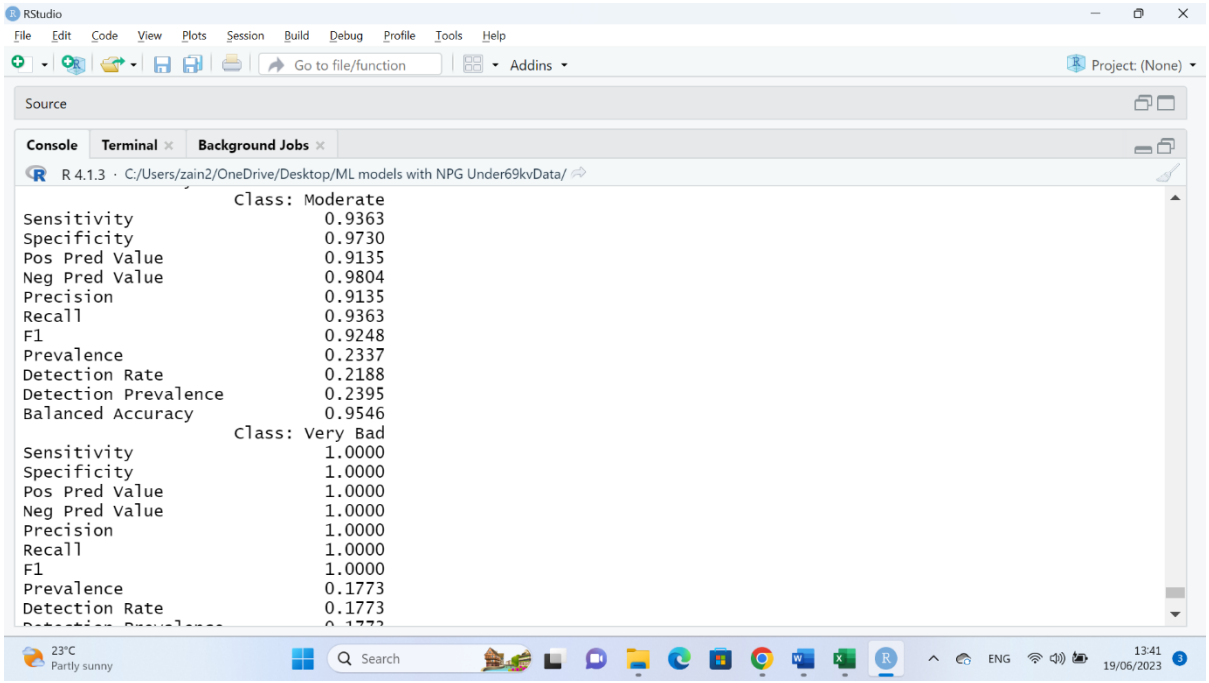


Figure H- 43: Screenshot from R studio showing RF Model Output (6).

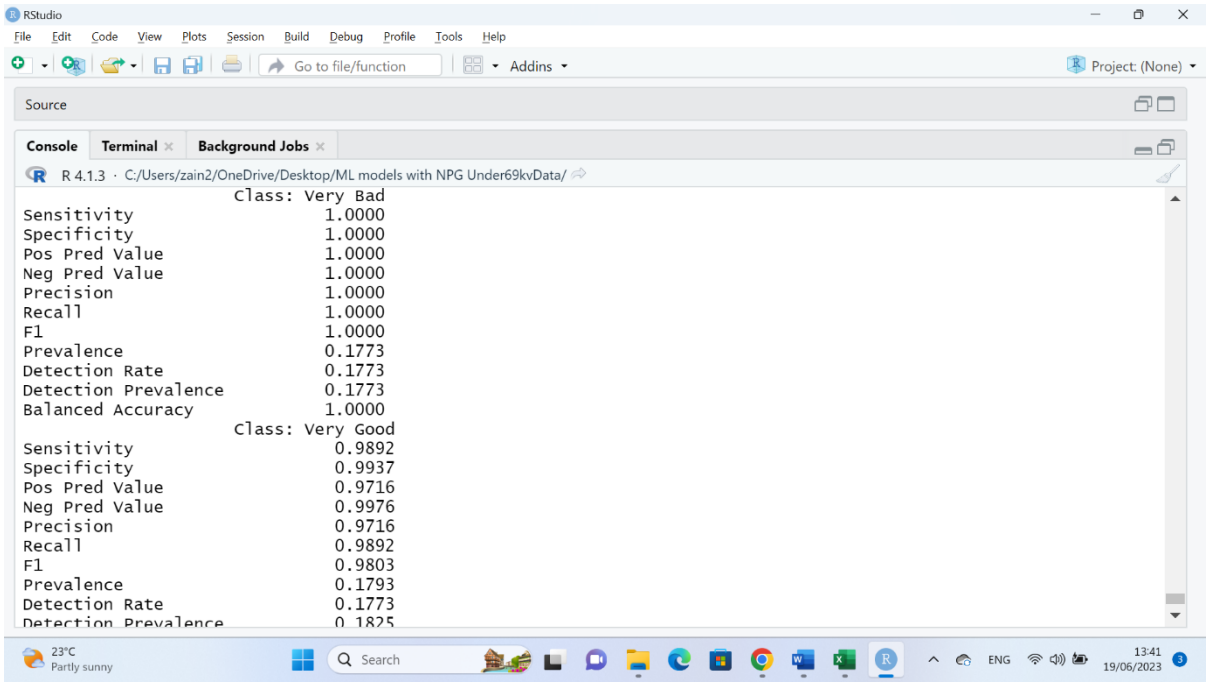


Figure H- 44: Screenshot from R studio showing RF Model Output (7).

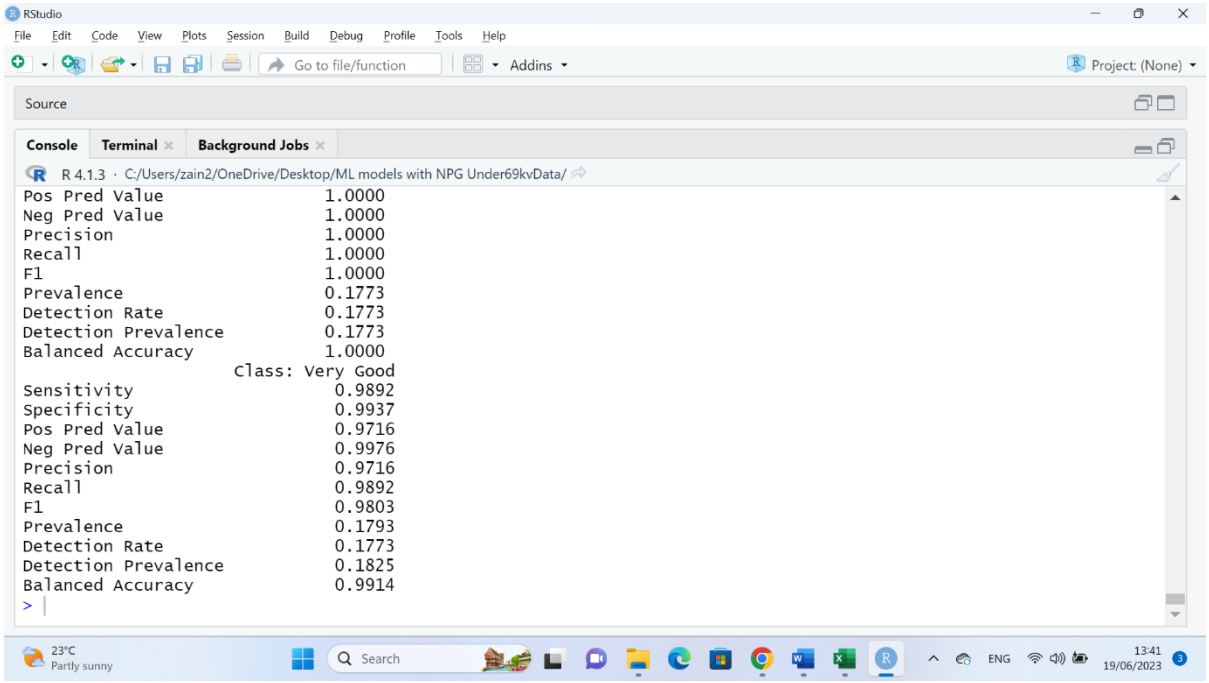


Figure H- 45: Screenshot from R studio showing RF Model Output (8).

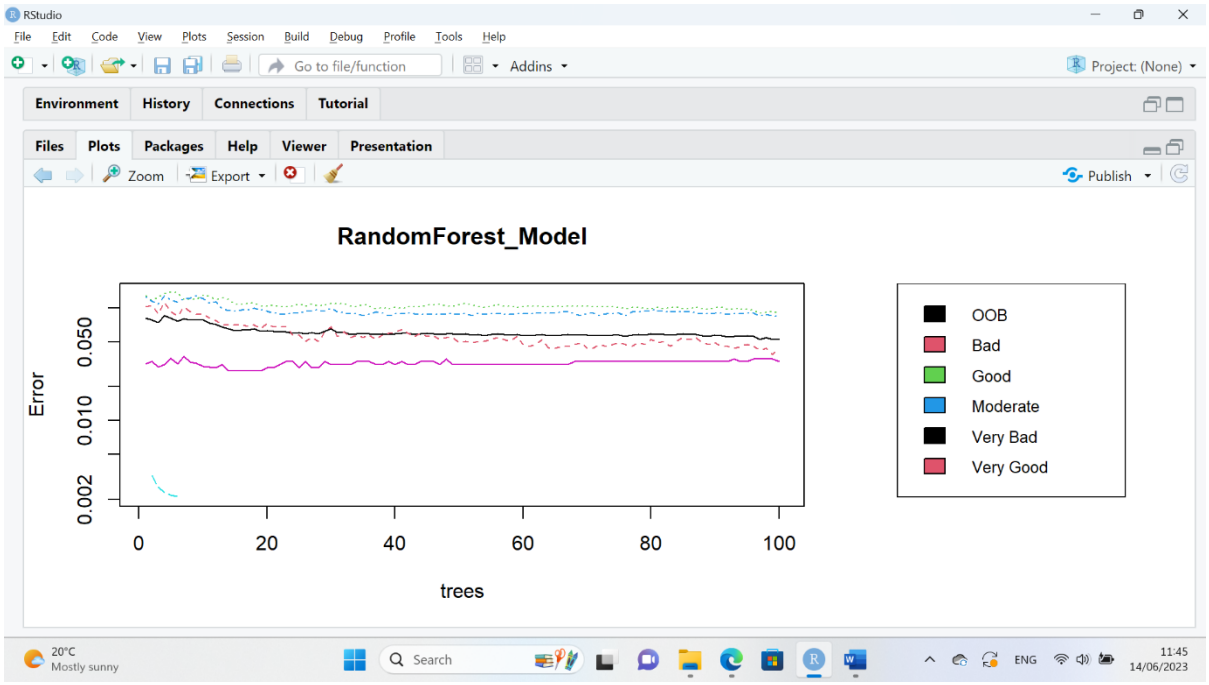


Figure H- 46: Screenshot from R studio showing RF Model Output (9).

- **KNN**

Table H- 17: KNN Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	253	16	27	0	6
	Good	2	228	26	0	5
	Moderate	9	12	253	0	0
	Very Bad	1	0	0	230	0
	Very Good	0	1	0	0	226

Table H- 18: KNN Statistics by Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.9547	0.8872	0.8268	1	0.9536
Specificity	0.9524	0.9682	0.9788	0.9991	0.9991
Pos Pred Value	0.8377	0.8736	0.9234	0.9957	0.9956
Neg Pred Value	0.9879	0.972	0.9481	1	0.9897
Precision	0.8377	0.8736	0.9234	0.9957	0.9956
Recall	0.9547	0.8872	0.8268	1	0.9536
F1	0.8924	0.8803	0.8724	0.9978	0.9741
Prevalence	0.2046	0.1985	0.2363	0.1776	0.183
Detection Rate	0.1954	0.1761	0.1954	0.1776	0.1745
Detection Prevalence	0.2332	0.2015	0.2116	0.1784	0.1753
Balanced Accuracy	0.9536	0.9277	0.9028	0.9995	0.9763

- **KNN model Output**

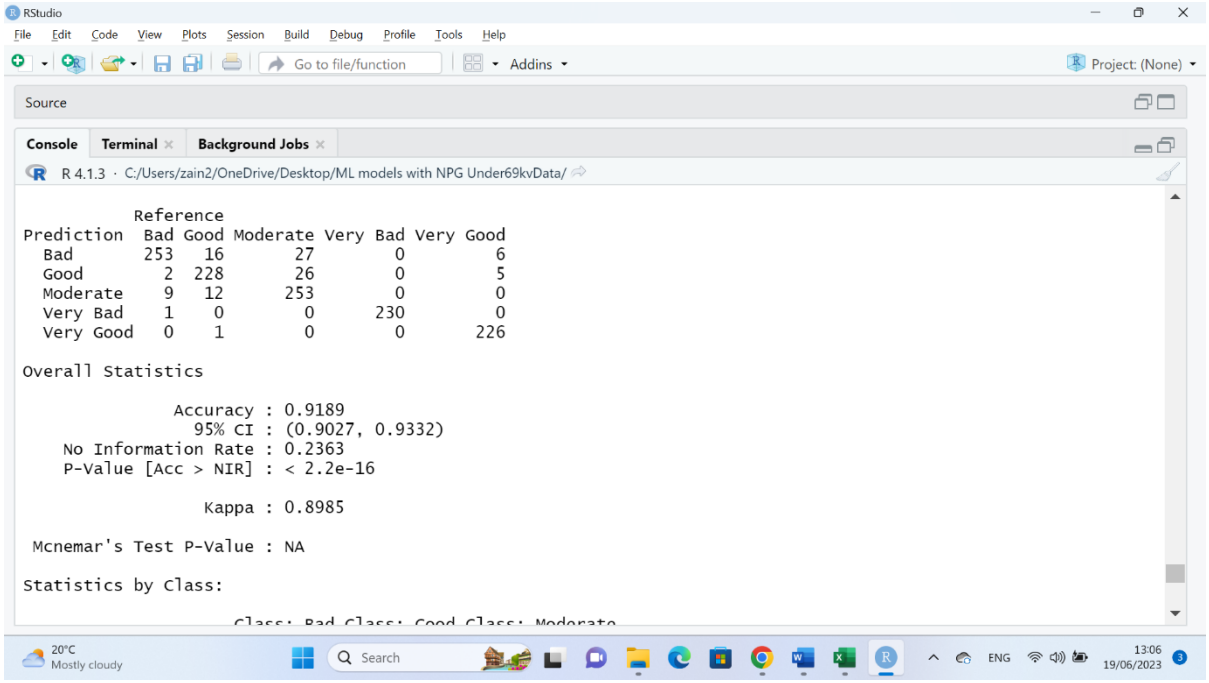


Figure H- 47: Screenshot from R studio showing KNN Model Output (1)

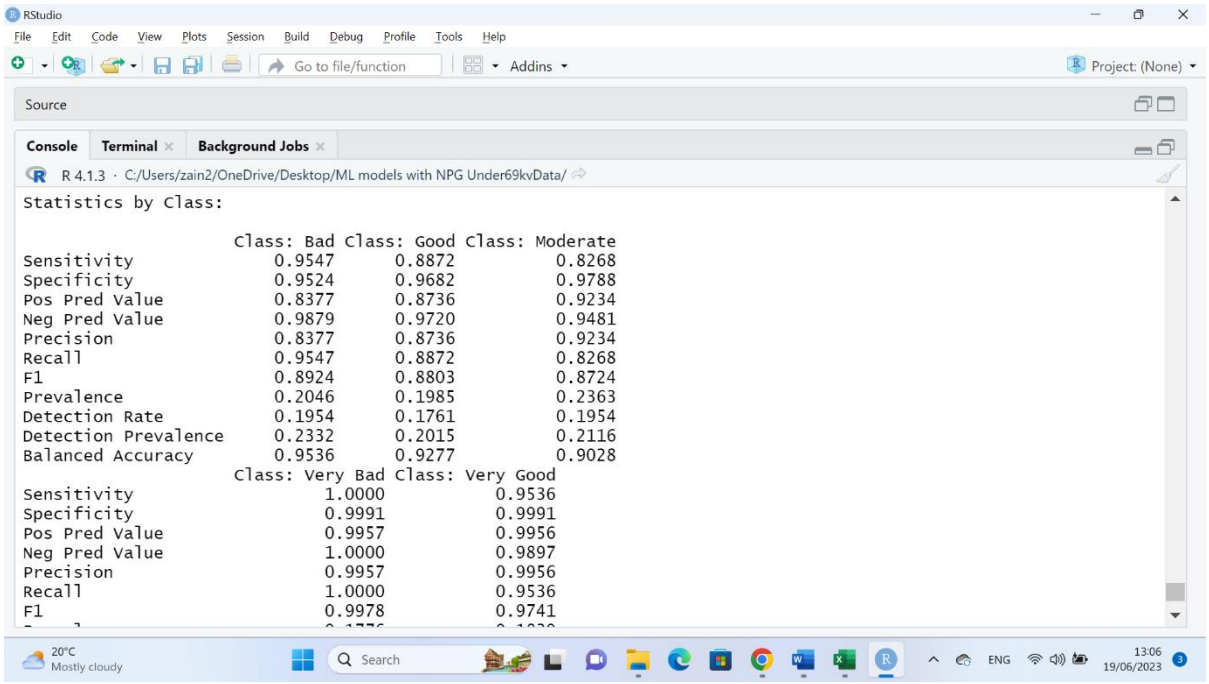


Figure H- 48: Screenshot from R studio showing KNN Model Output (2).

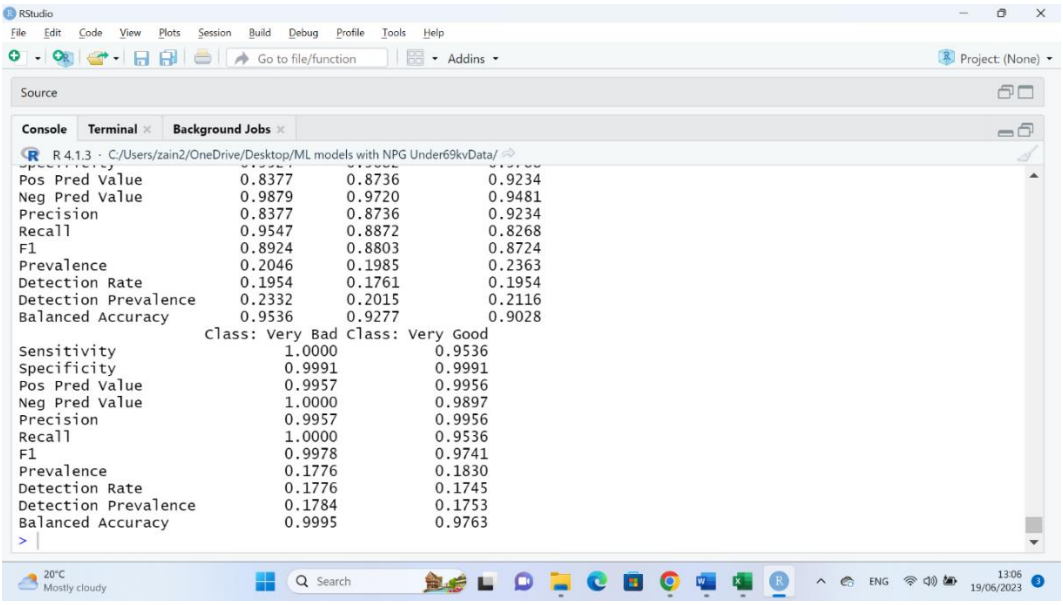


Figure H- 49: Screenshot from R studio showing KNN Model Output (3).

- **NB**

Table H- 19: NB Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	245	4	62	5	0
	Good	0	266	37	0	0
	Moderate	8	64	296	0	2
	Very Bad	0	0	0	274	0
	Very Good	0	3	5	0	274

Table H- 20: NB Statistics by Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.9684	0.7893	0.74	0.9821	0.9928
Specificity	0.945	0.9694	0.9354	1	0.9937
Pos Pred Value	0.7753	0.8779	0.8	1	0.9716
Neg Pred Value	0.9935	0.9428	0.9115	0.9961	0.9984
Precision	0.7753	0.8779	0.8	1	0.9716
Recall	0.9684	0.7893	0.74	0.9821	0.9928
F1	0.8612	0.8312	0.7688	0.991	0.9821
Prevalence	0.1638	0.2181	0.2589	0.1806	0.1786
Detection Rate	0.1586	0.1722	0.1916	0.1773	0.1773
Detection Prevalence	0.2045	0.1961	0.2395	0.1773	0.1825
Balanced Accuracy	0.9567	0.8793	0.8377	0.991	0.9932

- **NB model Output**

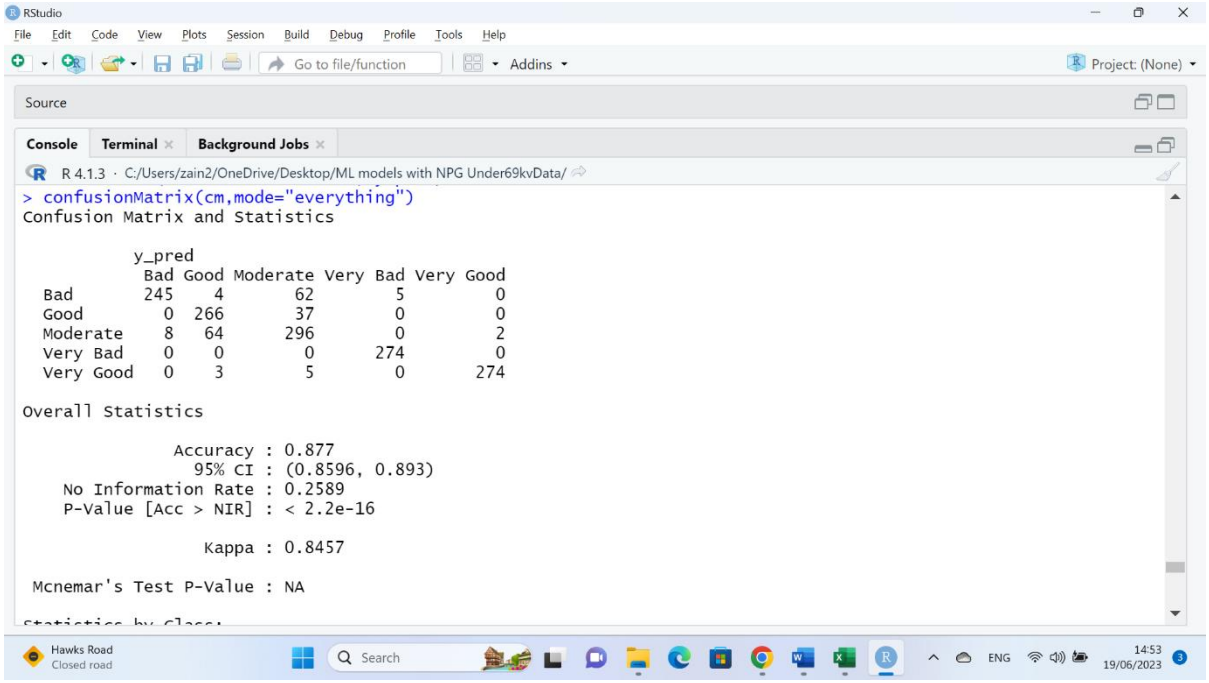


Figure H- 50: Screenshot from R studio showing NB Model Output (1).

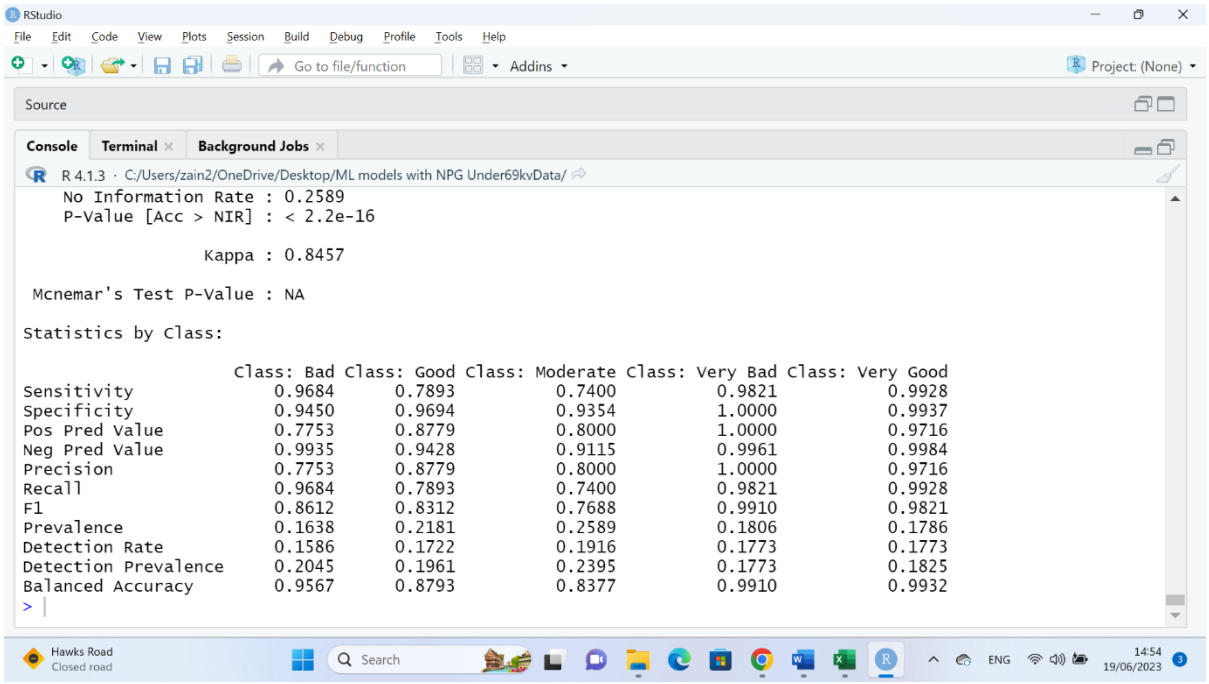


Figure H- 51: Screenshot from R studio showing NB Model Output (2).

- **DT**

Table H- 21: DT Confusion Matrix

		Prediction				
		Bad	Good	Moderate	Very Bad	Very Good
Actual	Bad	284	2	10	0	0
	Good	26	220	2	0	1
	Moderate	17	38	233	0	0
	Very Bad	0	0	0	237	0
	Very Good	3	1	1	0	222

Table H- 22: DT Statistics by Class/Condition

	Class: Bad	Class: Good	Class: Moderate	Class: Very Bad	Class: Very Good
Sensitivity	0.8606	0.8429	0.9472	1	0.9955
Specificity	0.9876	0.972	0.9477	1	0.9953
Pos Pred Value	0.9595	0.8835	0.809	1	0.978
Neg Pred Value	0.954	0.9609	0.9871	1	0.9991
Precision	0.9595	0.8835	0.809	1	0.978
Recall	0.8606	0.8429	0.9472	1	0.9955
F1	0.9073	0.8627	0.8727	1	0.9867
Prevalence	0.2544	0.2012	0.1897	0.1827	0.1719
Detection Rate	0.219	0.1696	0.1796	0.1827	0.1712
Detection Prevalence	0.2282	0.192	0.2221	0.1827	0.175
Balanced Accuracy	0.9241	0.9075	0.9474	1	0.9954

- **DT Model Output**

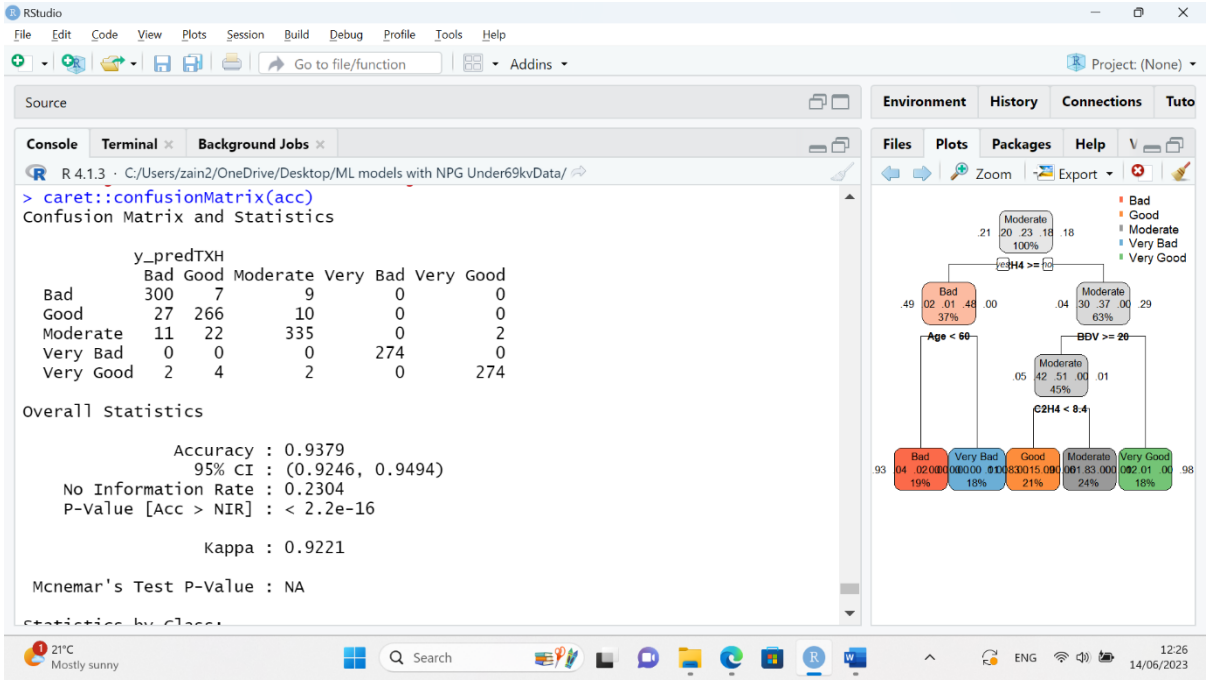


Figure H- 52: Screenshot from R studio showing DT Model Output (1).

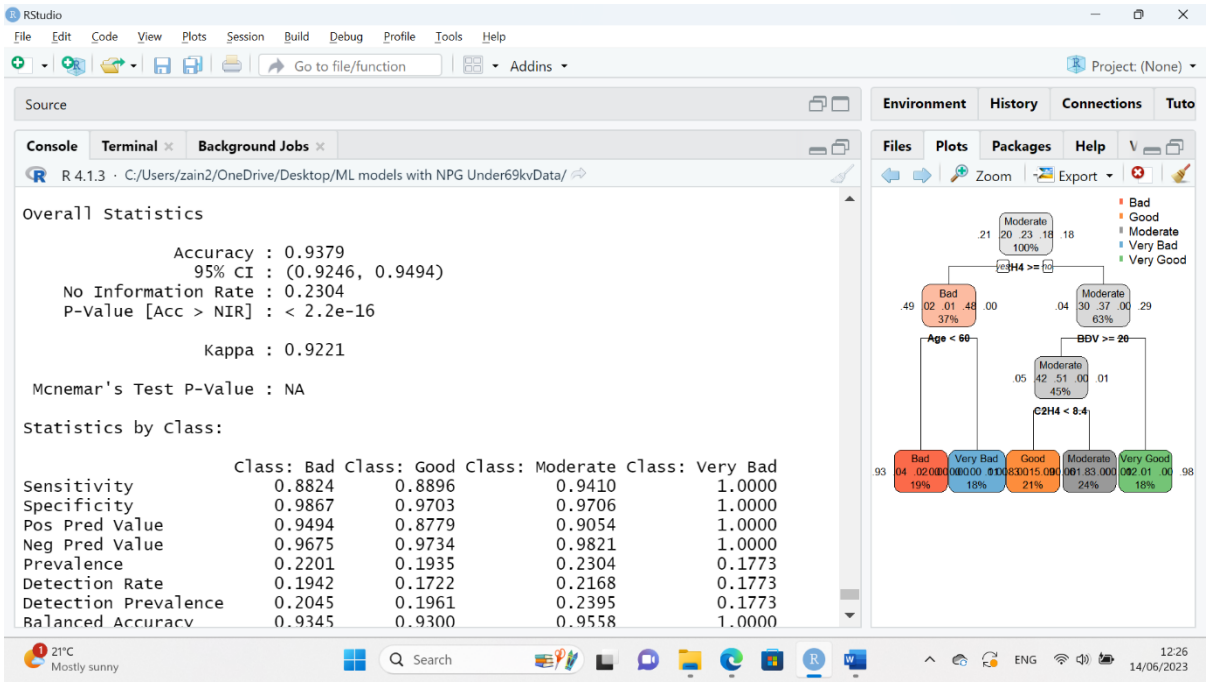


Figure H- 53: Screenshot from R studio showing DT Model Output (2).

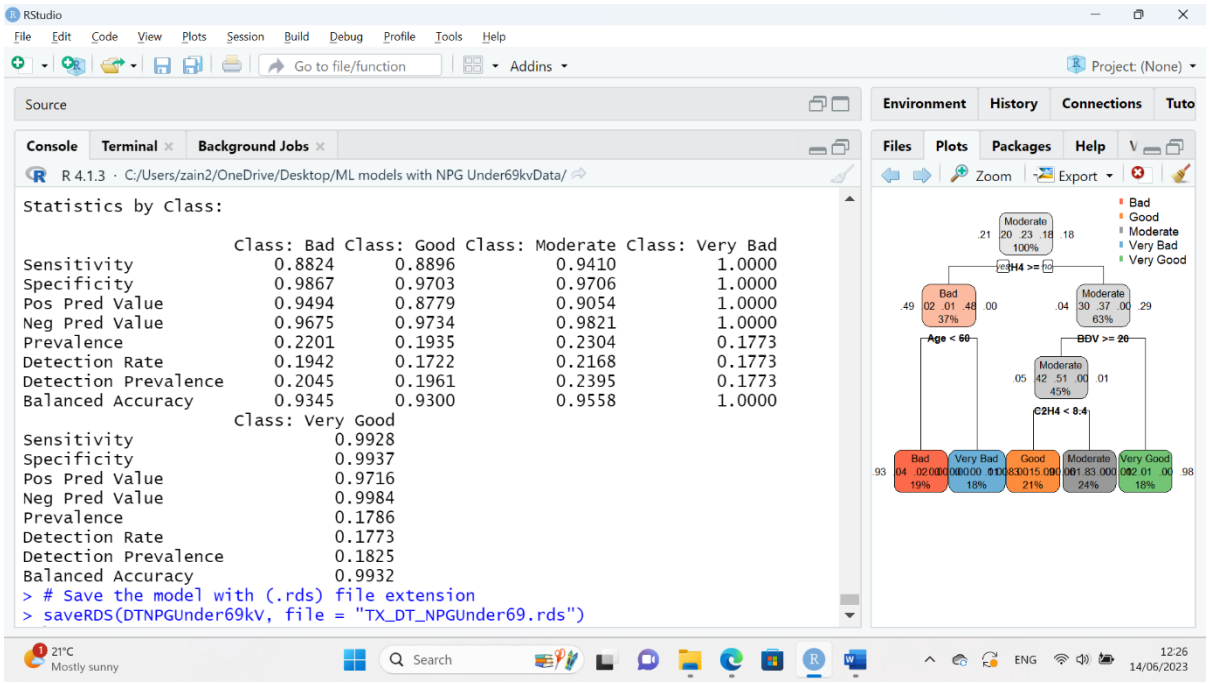


Figure H- 54: Screenshot from R studio showing DT Model Output (3).

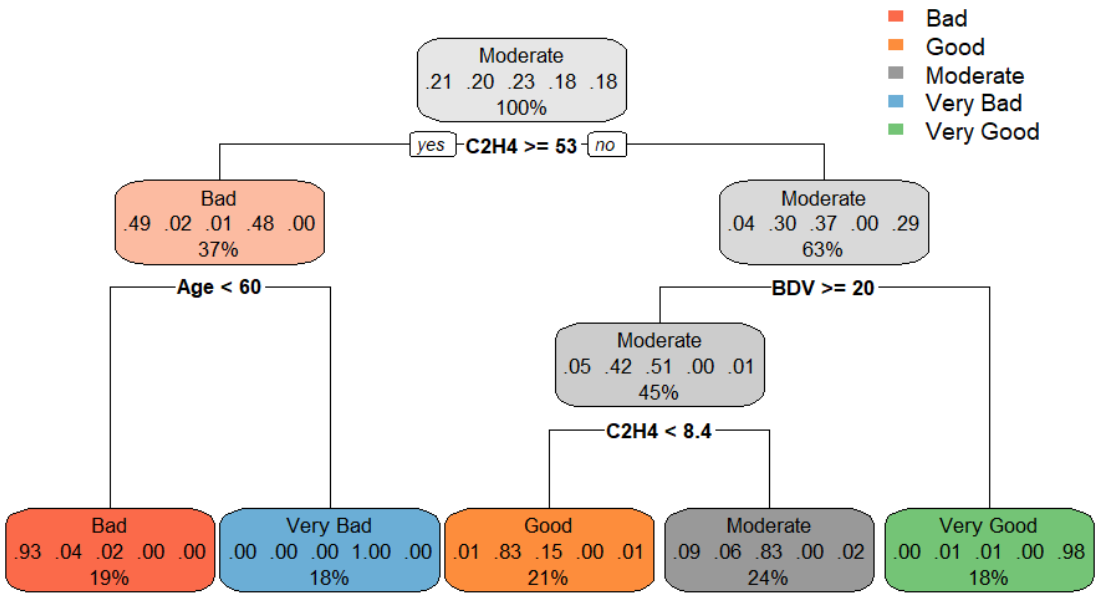


Figure H- 55: Decision Tree Model Plot

Part 4 Mobile and Website Application

Process for Developing PTHI Web App

1) Define the primary goal of the app
    a. Deploy the ML model to be a decision-making tool.
    b. Analysis of the dataset using
    c. Show information from the data using different visuals (Dashboard).
2) Sketch the web App design.
3) Write a short app description to show on one of the pages.
4) Write R code to perform data analysis on the dataset using R and present it in an R Markdown document as an EDA page.
5) Write R code to explain information from the dataset using visualisation (Data Dashboard)
6) Sketch the PTHI decision tool.

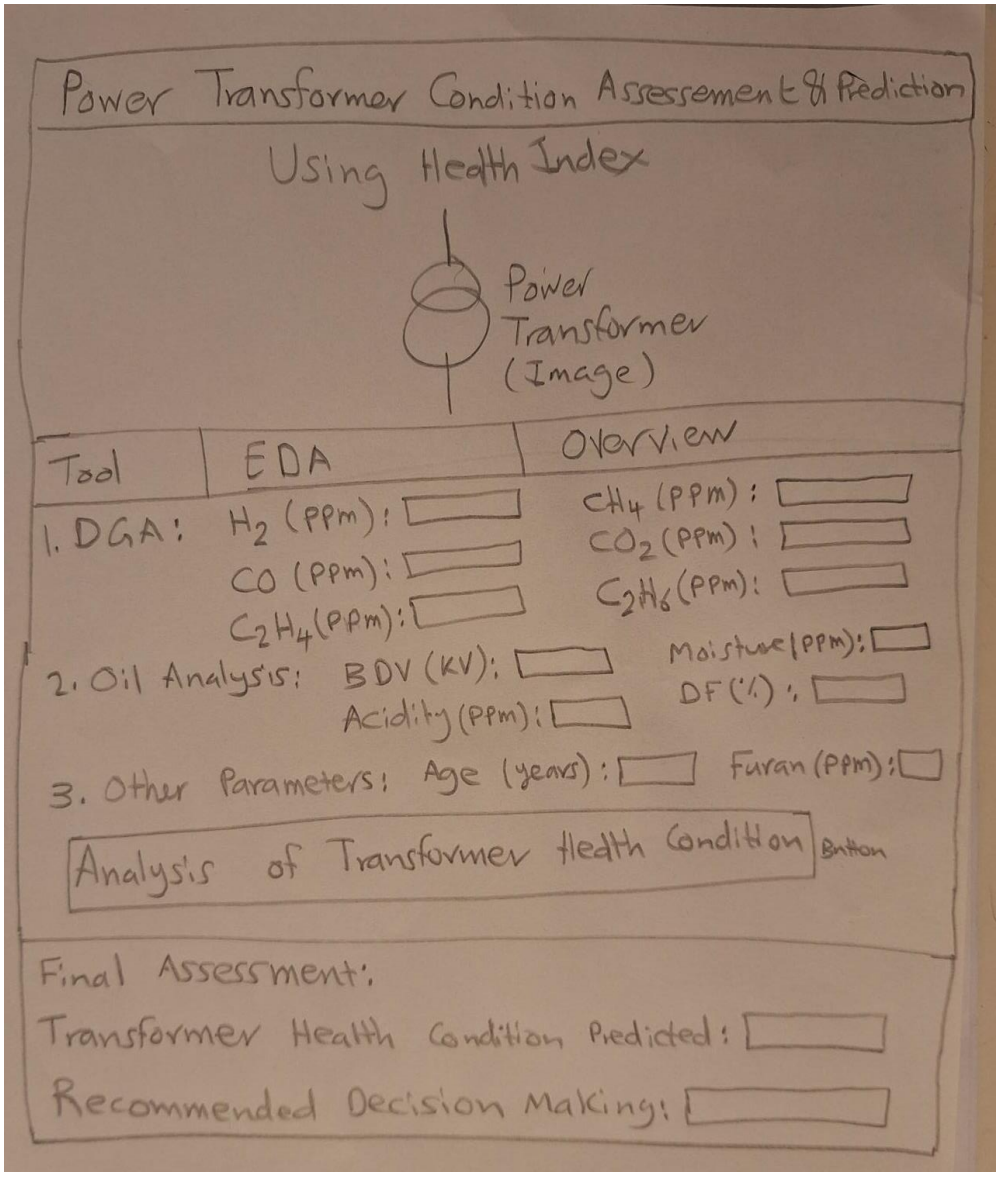


Figure H- 56: Web Application workflow.

Process of developing Dashboard Visualisation Idea

7) Step 5 codes were separated and developed using Flexdashboard in R Studio to visualise the result and deliver valuable information to AM decision-makers.
8) As of the analysis date, the results show in the figures below that the electrical utility has a large, healthy fleet of transformers maintained well over the previous years. The study shows the requirement for 951 transformers for routine inspection and 198 transformers for annual maintenance. The tab set enables one to read through other essential indicators in the same way as maintenance. It shows transformers by condition, percentage of current transformers condition, moisture, acidity, and break-down voltage levels.

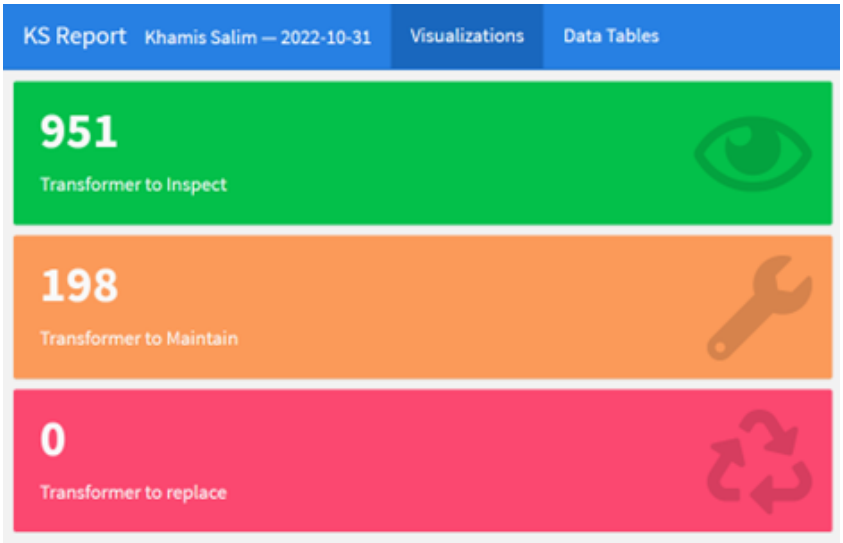


Figure H- 57: Screenshot for the PTHI analysis using flex dashboard.

Figure H- 58: Number of Transformers by Condition.

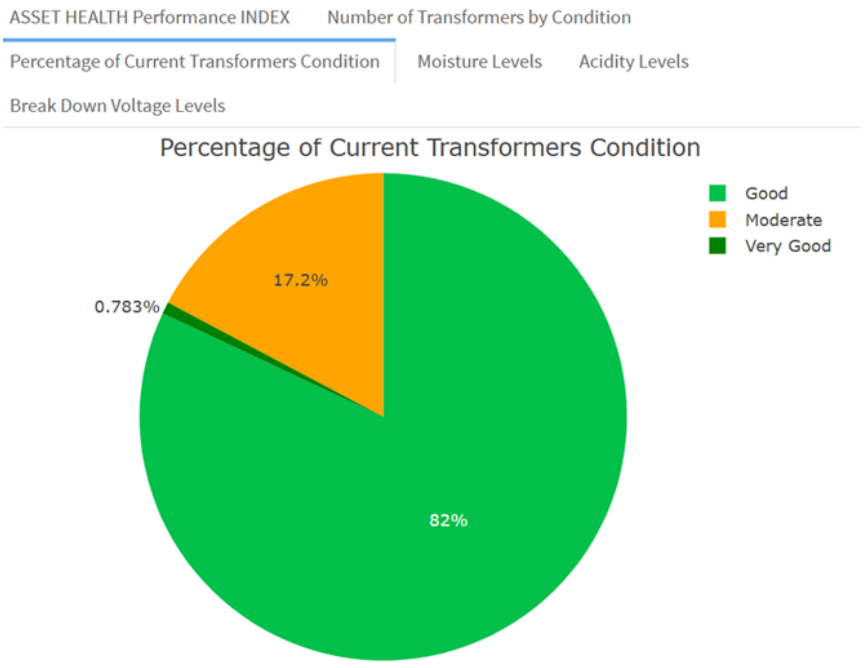


Figure H- 59: Percentage of Current transformers condition.

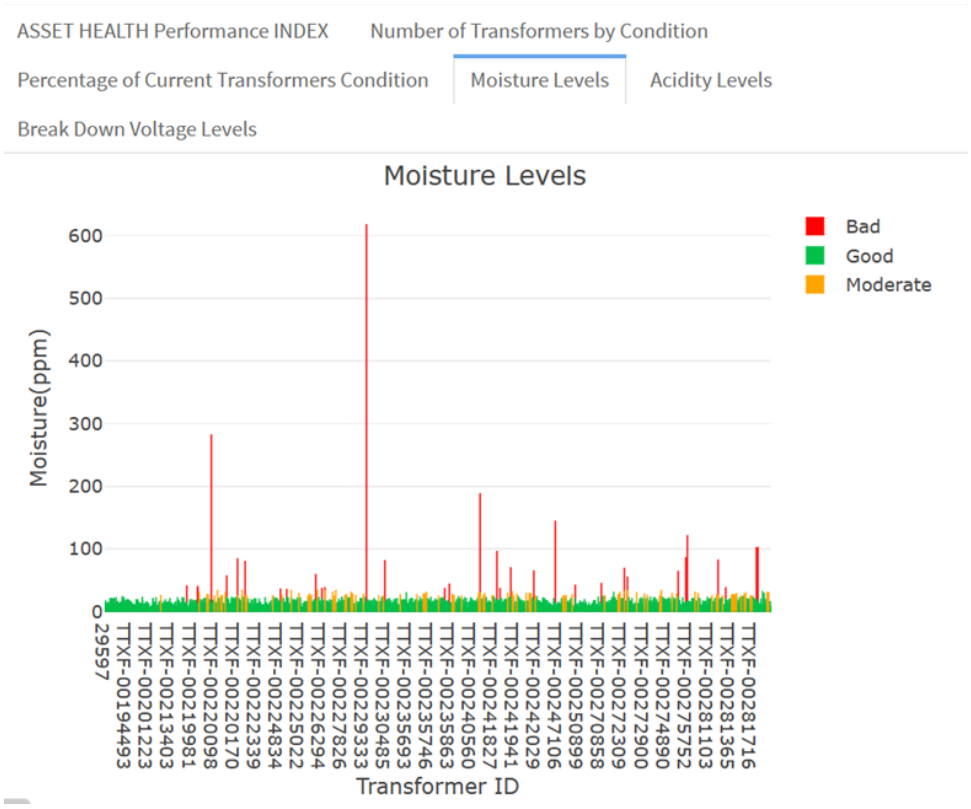


Figure H- 60: Transformers Moisture Levels.

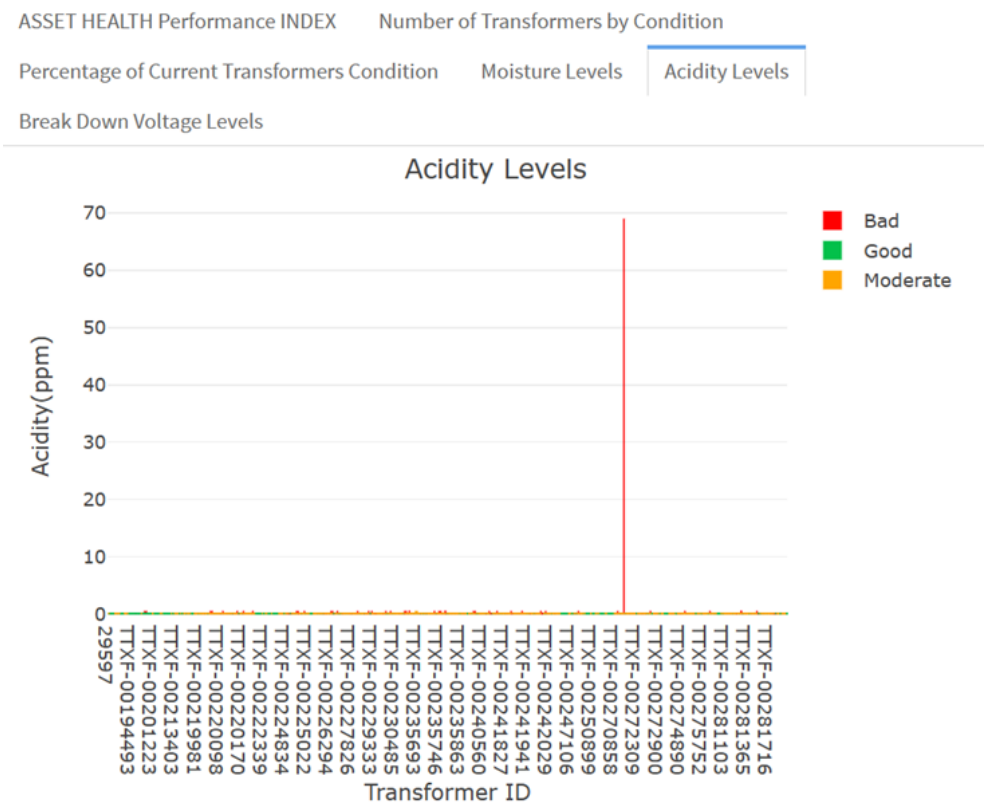


Figure H- 61: Transformer Acidity Levels.

ASSET HEALTH Performance INDEX | Number of Transformers by Condition | Percentage of Current Transformers Condition | Moisture Levels | Acidity Levels | Break Down Voltage Levels

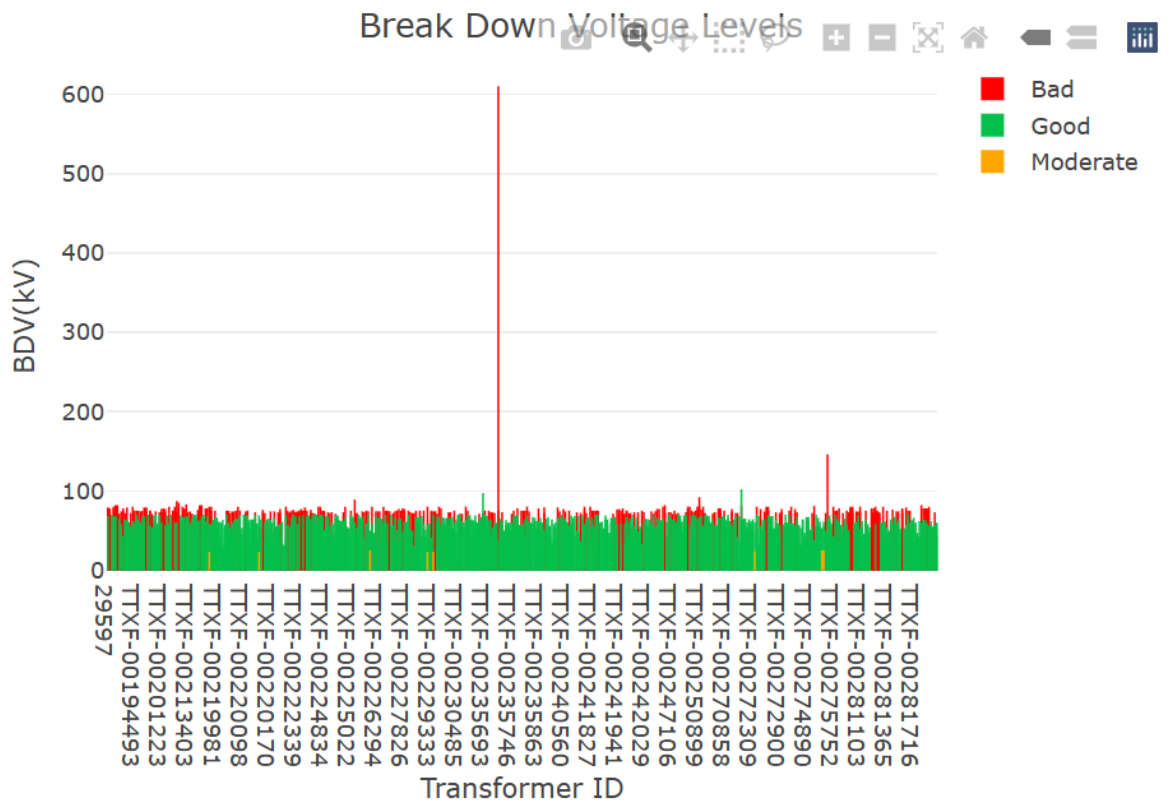


Figure H- 62: Breakdown Voltage Levels.

KS Report Khamis Salim — 2022-10-31 | Visualizations | Data Tables

132kV Transformers Data | 66kV Transformers Data | 33kV Transformers Data

Download Current Page | Download All Pages | Search:

	TX.Asset.ID	Substation.Name	Circuit.Name	Manufacturer...Specification	Serial.Number..Tx.....Tap.	Age	Sample_Date	Colour	Colour_Score
1	TTXF-00271692	ANNFIELD 132/66KV	GRID TRANSFORMER 1	BRUCE PEEBLES | BS 171	43512 | 1256	55	19/11/2019	Dark Yellow	4
2	TTXF-00271693	ANNFIELD 132/66KV	GRID TRANSFORMER 2	BRUCE PEEBLES | BS 171	43511 | 1255	55	19/11/2019	Dark Yellow	4
3	TTXF-00242025	ATTERCLIFFE 1365	Trans01	English Electric | 0	131C1718/53 | 1C1718.61	65	19/09/2018	Light Amber	1
4	TTXF-00241932	ATTERCLIFFE 1365	Trans02	English Electric | 0	131C1718/54 | 1C1718.62	65	19/09/2018	Amber	1
5	TTXF-00197803	BARRACK ROAD	GRID TRANSFORMER 1	ABB POWER T&D LTD | 0	1134114 | 1ZSC 8714 887	7	28/06/2019	Light Yellow	2

Showing 1 to 5 of 206 entries

Previous 1 2 3 4 5 ... 42 Next

Figure H- 63: Data tables showing the electrical utility records.

The web application starts with an authentication form to show its security, as only users with a username and password can access the app.

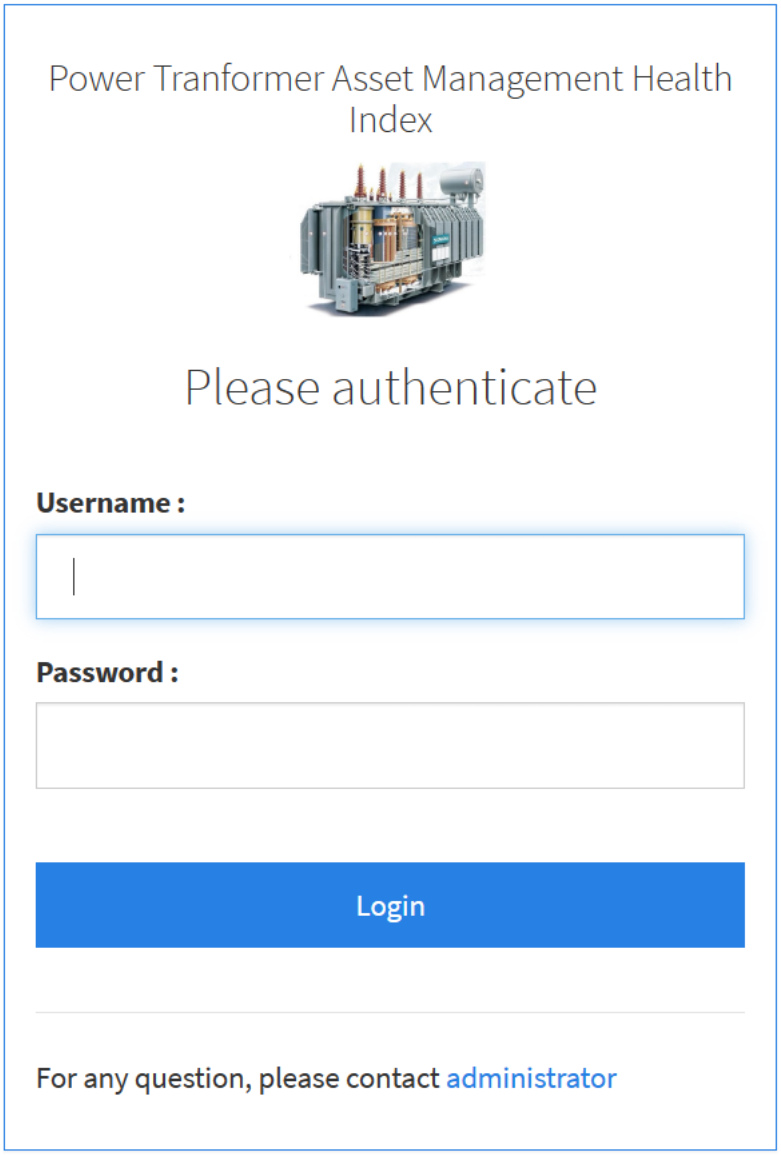


Figure H- 64: PTHI model in shiny web application login screen.

Figure H- 65: PTHI model in shiny web application tabs.

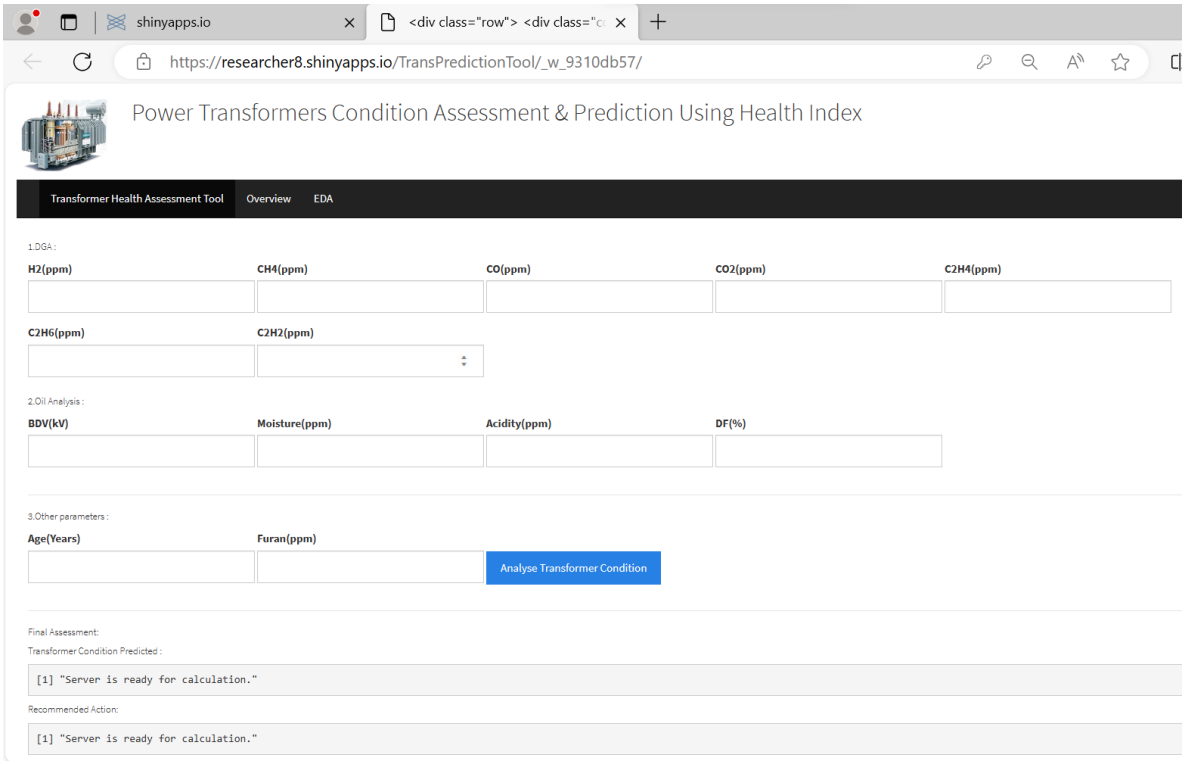


Figure H- 66: Health Index web application screenshot.

- EDA:

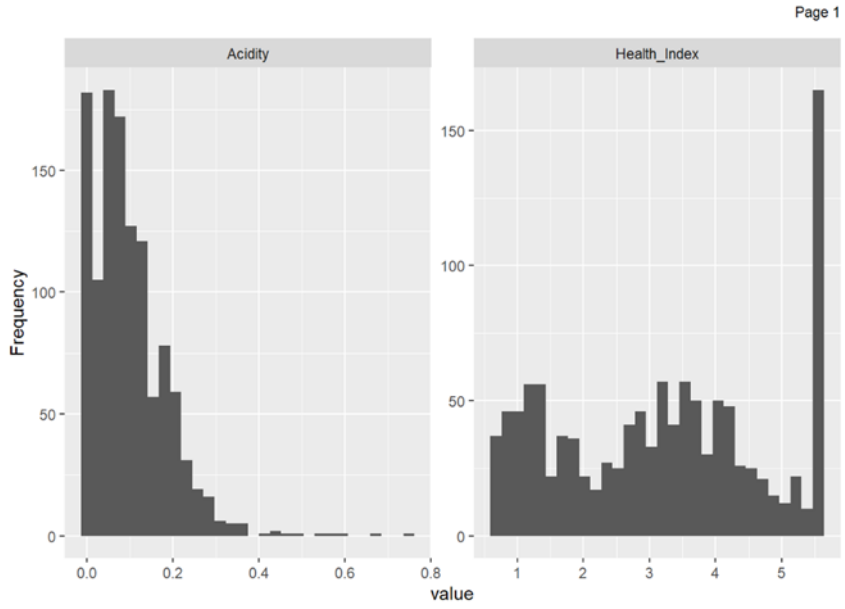


Figure H- 67: Health Index web application screenshot showing acidity and Health Index distribution.

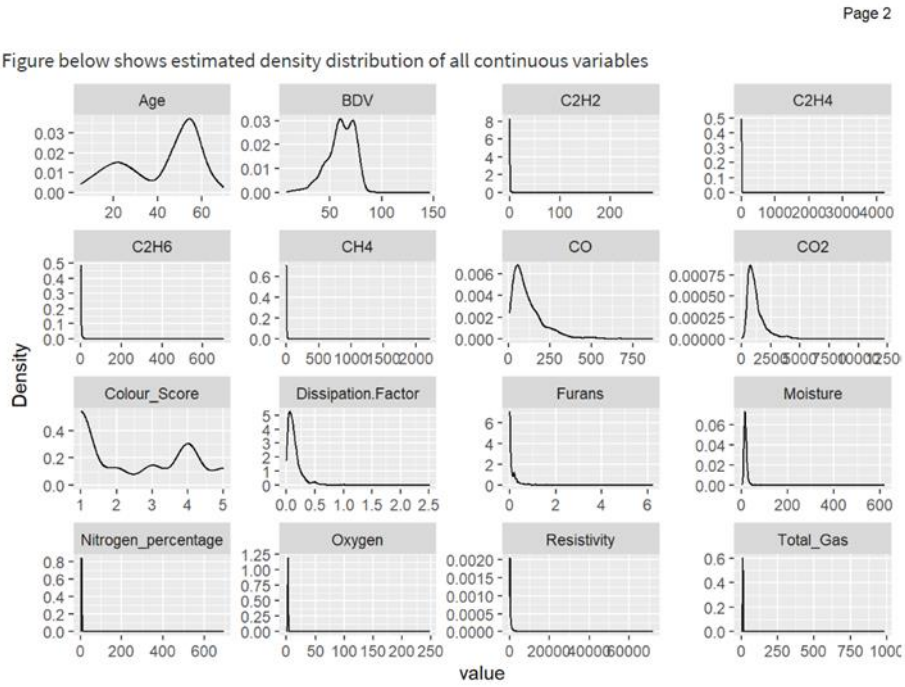


Figure H- 68: Health Index web application screenshot showing density distribution of all continuous variables.

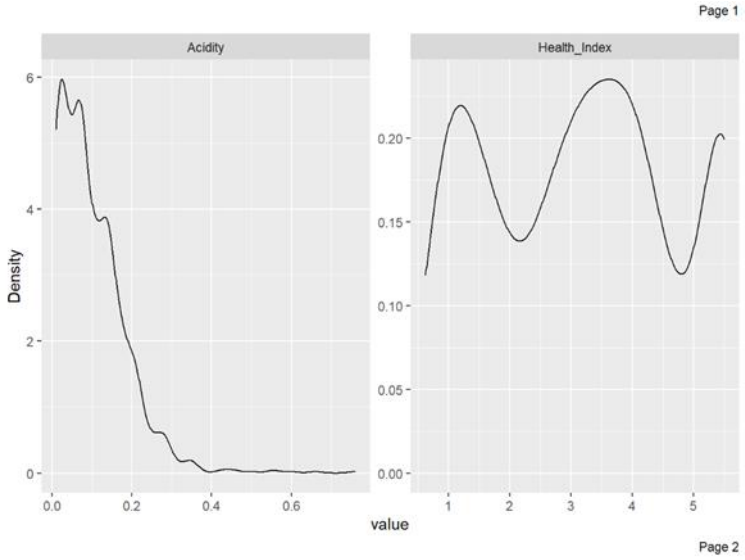


Figure H- 69: Health Index web application screenshot showing density plot of acidity and health Index.

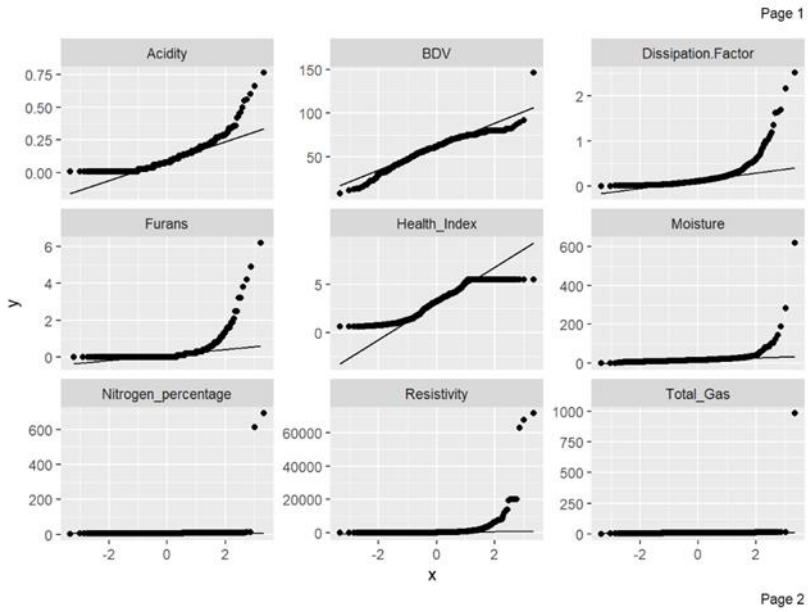


Figure H- 70: Health Index web application screenshot showing a quantile-quantile plot of all continuous variables (part 1).

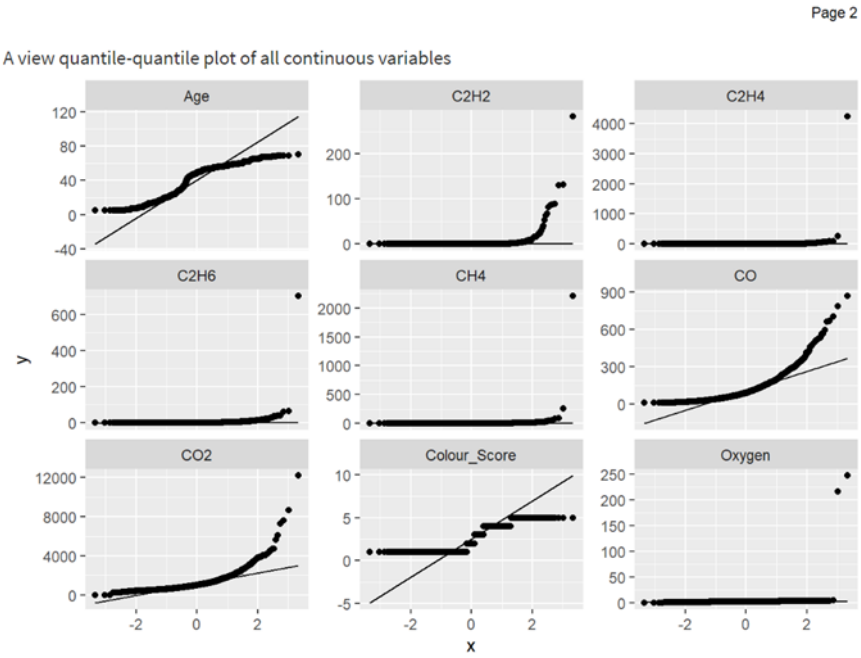


Figure H- 71: Health Index web application screenshot showing a quantile-quantile plot of all continuous variables (Part 2).

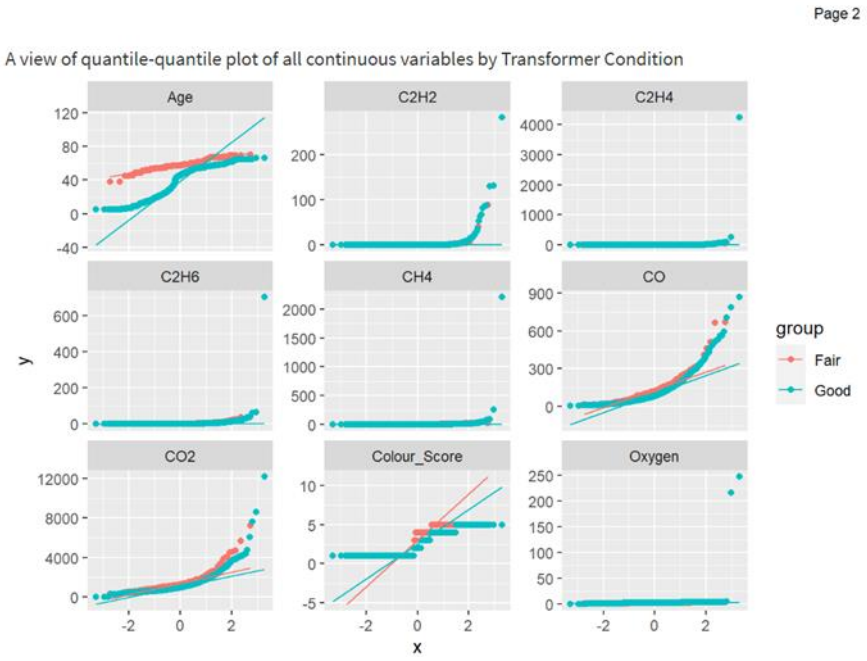


Figure H- 72: Health Index web application screenshot showing plots of continuous variables by transformer condition (Part 1).

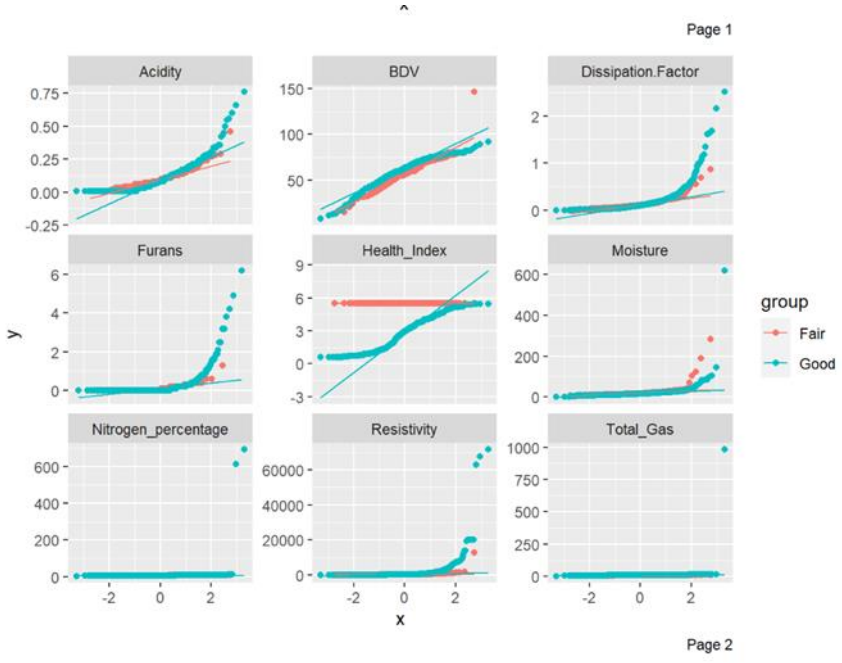


Figure H- 73: Health Index web application screenshot showing plots of continuous variables by transformer condition (Part 2).

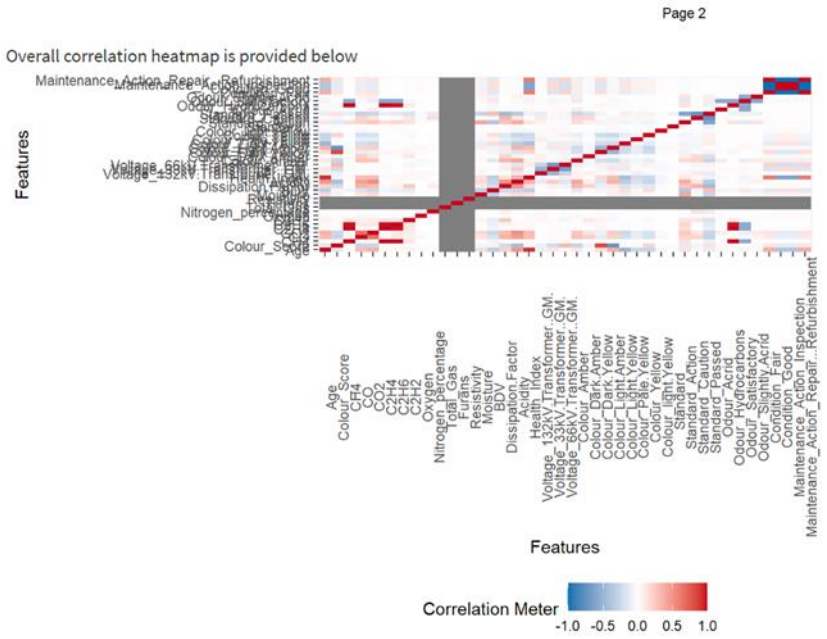


Figure H- 74: Health Index web application screenshot showing overall correlation heatmap.

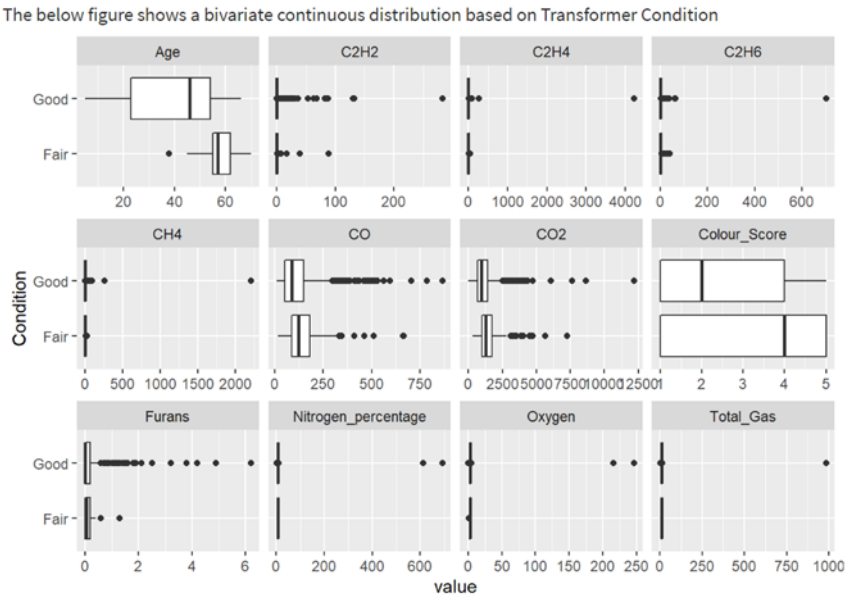


Figure H- 75: Health Index web application screenshot showing bivariate continuous distribution based on transformer condition (Part 1).

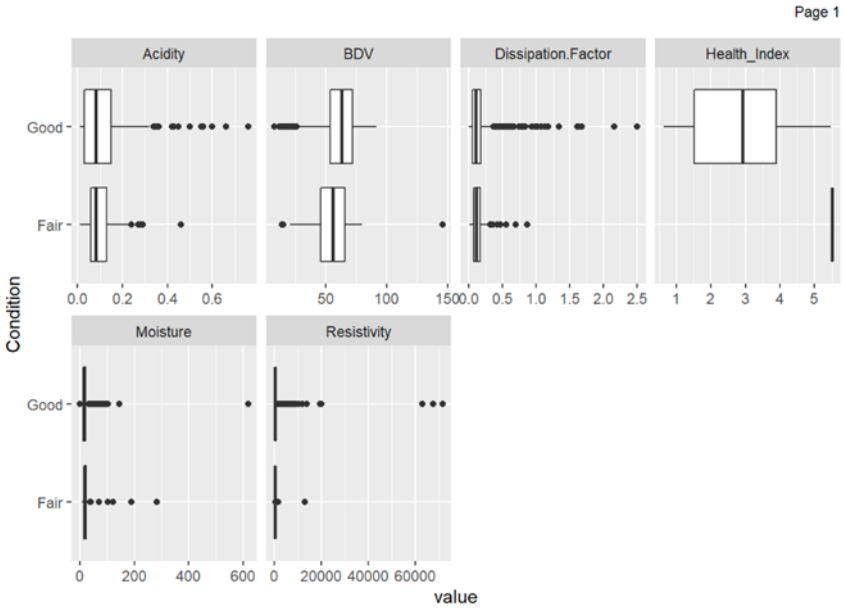


Figure H- 76: Health Index web application screenshot showing bivariate continuous distribution based on transformer condition (Part 2).

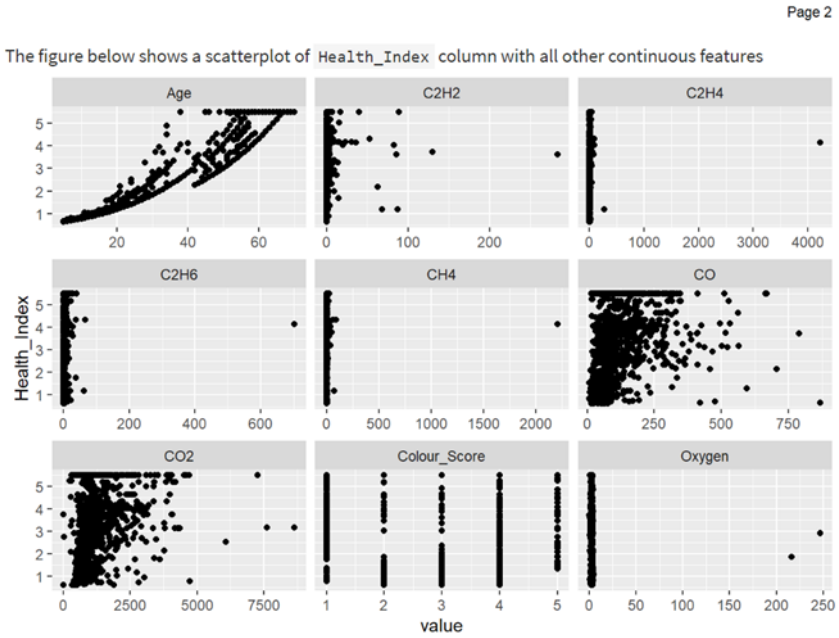


Figure H- 77: Health Index web application screenshot showing scatterplots of Health Index columns with other continuous variables (Part 1).

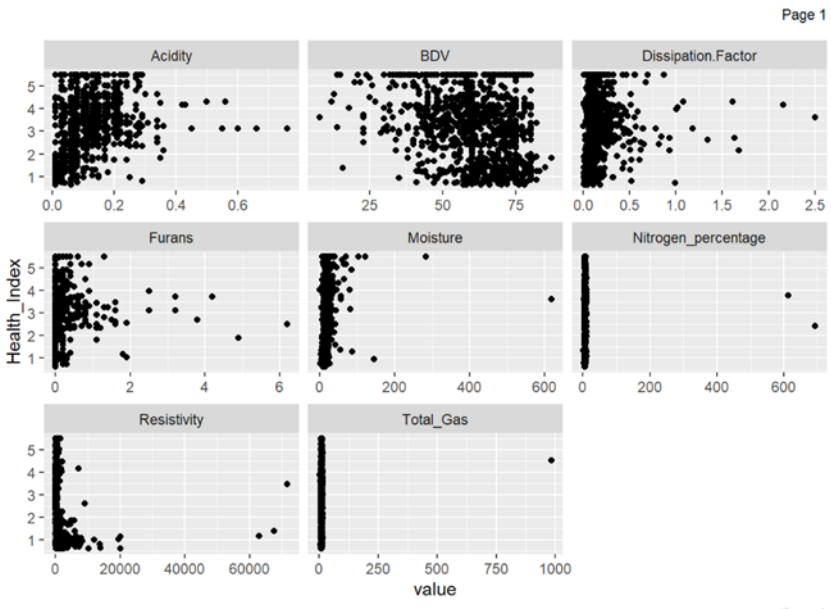


Figure H- 78: Health Index web application screenshot showing scatterplots of Health Index columns with other continuous variables (Part 2).

Part 5 Power Transformer Age Study and Failure Rate Experiments

The regression line is utilising the formula:

$$f(x) = a.x + b \qquad \text{(H-1)}$$

Where a is the line slope, x is the transformer's apparent age, b is the intercept, and F(x) is the calculated HI.

A process will assist in obtaining the distribution that fits well with the failure model by studying the PT failure data to determine the failure root cause, as shown in Figure H-79.

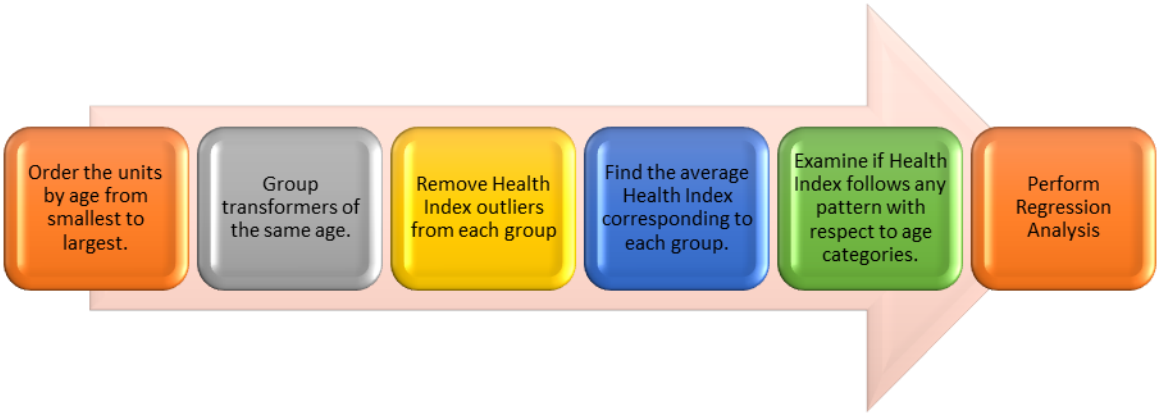


Figure H- 79: Studying PT failures using the Age and Health Index correlation.

Figure H-80 shows the PT hazard function based on international best practices, which shows that the failure rate is about 0.0002 failures/year per transformer through the constant rate phase. Then, the failure rate begins to climb at 20-25 years, reaching a failure rate of 0.1 at 40 years, 0.025 at 50 years, and over 0.5 at 60 years (Tenbohlen et al., 2016).

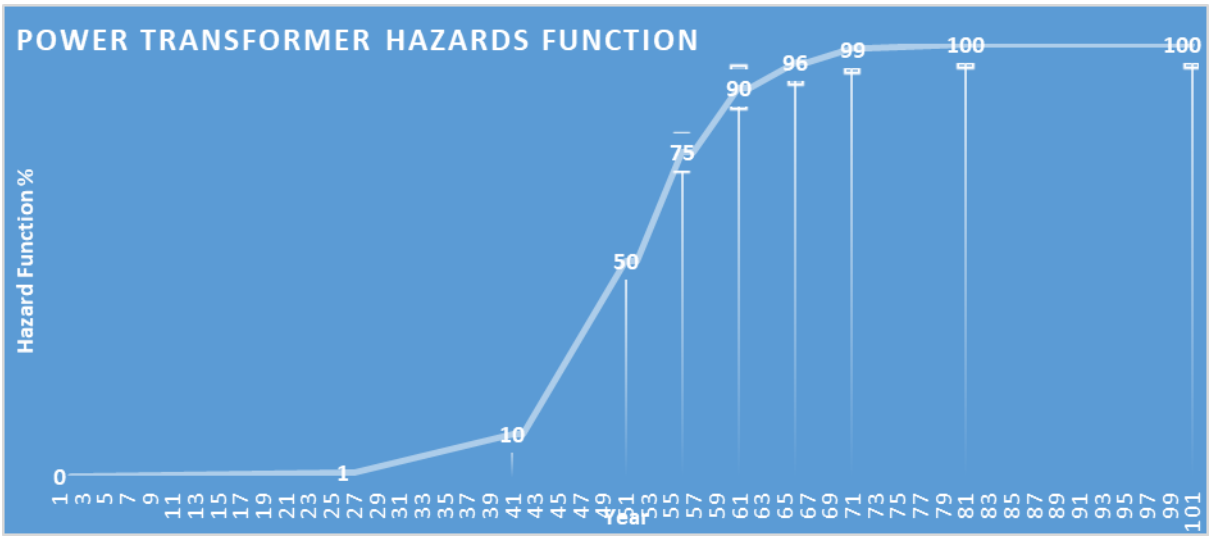


Figure H- 80: PT failure rate using Hazards Function.

The assessment basis of the probability of failure can be mapped to transformer age using hazard function as a transformer with <1% PoF considerably has very low failure probability, between 1% -25% is low failure probability, 25% -50% is medium, 50% - 75% is high failure probability, 75% at very high failure probability as shown in Table H-23.

Table H- 23: Probability assessment using age function.

Age Group	PoF Based on Age	PoF %
1-40 years	Very Low	<1%
40-50 years	Low	1% - 25%
50-60 years	Medium	25% - 50%
60-70 years	High	50% - 75%
>70 years	Very High	>75%

Table H-24 express PT's condition monitoring using HI to predict PT's remaining life.

Table H- 24: Transformer HI condition monitoring and the remaining life ranges.

Utility HI Range	Condition	Maintenance Action	Conventional HI Range	Remaining Life	Description
0- 0.5	Very Good	Inspection	85 - 100	>15 years	Some aging or minor deterioration of a limited number of a component
0.5 – 5.4	Good	Inspection	70 - 85	7-15years	Significant deterioration of some components
5.5 – 9	Fair	Repair & Refurbishment	50 - 70	3-7 years	Widespread significant deterioration or serious deterioration of specific components
10 -100	Poor	Replacement	30 - 50	2-3 years	Widespread serious deterioration
>100	Very Poor	Replacement	0 - 30	0-2 years	Extensive serious deterioration

Figure H-81 shows a regression line indicating the transformers fleet's average health condition to its age relationship, where transformer units above the line are healthy transformers. It classifies the transformers according to the age on the x-axis, whereas the HI on the y-axis. The HI range in PTHI is from 0 to 100, as 0 represents a “Very Bad” health condition, and 100 is a “Very Good” health condition according to the conventional HI Range in Table H-24.

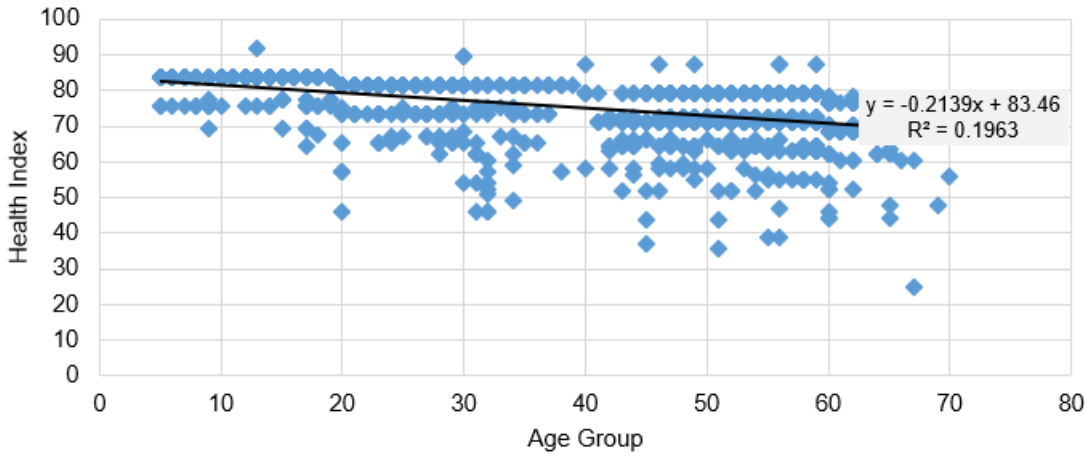


Figure H- 81: Age grouping of the transformers fleet by HI.

After plotting the transformer population using a scatter plot in Figure H-81, a regression and a linear regression line were investigated considerably. The regression line is the average condition of the large transformers fleet (Li and Dong, 2019). As

shown in Figure H-82, a clear negative correlation between HI value and transformer age is evident in the study.

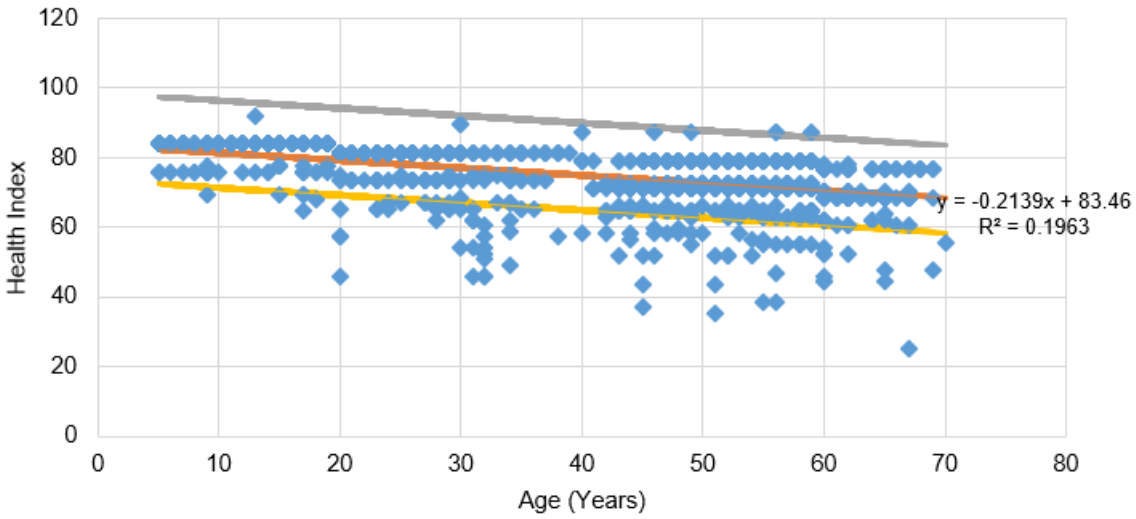


Figure H- 82: The regression line according to HI by age for the transformer population shows lower and upper confidence intervals.

The orange solid line in the middle is the regression line symbolising the transformer's population average. In this study, transformers with a significant differentiation between apparent age and actual age, which is the outliner, are not considered using the upper confidence interval value of more than fifteen years and lower confidence interval value of less than ten years. This period is another way of saying that the apparent age is modified to maximally fifteen years older (grey line) or ten years younger (yellow line) compared to the actual age. The outliners should not be counted in the extending asset life cycle and reducing the investment decision for new cost since the total capital is reduced based on increasing asset life span.

Another regression model after eliminating the outliners, and Figure H-83 shows the updated regression line.

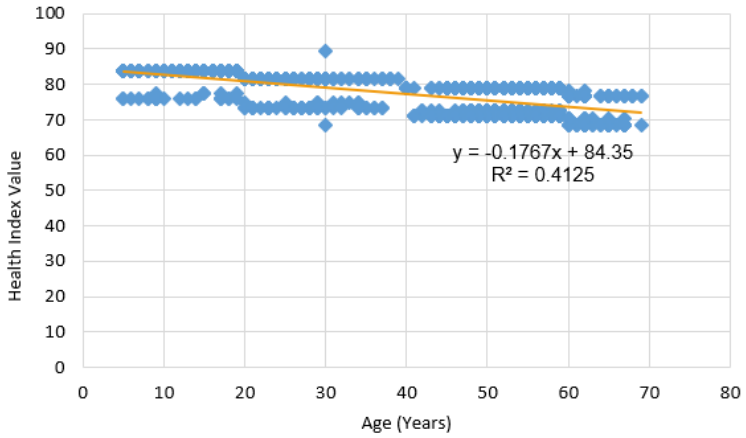


Figure H- 83: Regression line for calculating the apparent age.

Figure H-83 shows the updated regression performed after eliminating the outliners. It represents 984 transformers of the entire electrical utility transformer population. Also, the formula concluded from the above task represents the entire transformer population using the equation below:

$$\mathrm{HI} = -0.1767\mathrm{x} + 84.35 \qquad \text{(H-2)}$$

The above formula defines the transformer condition of the regression line and the transformer's apparent age. Therefore, apparent age investigation using the transformer's actual age and HI score utilising the regression line.

A statistical regression model was used by conducting several experiments to discover the best model that fits the age failure data. The first experiment, as Figure H-84 shows, is unreliable as the coefficient R^2 value is 0.1963, which means the model's accuracy is around 20%.

```
> lm <-lm(Health_Index~ Age)
> summary(lm)

Call:
lm(formula = Health_Index ~ Age)

Residuals:
    Min      1Q  Median      3Q     Max
-44.127  -2.256   1.694   5.265  16.258

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 83.46011    0.57049  146.30   <2e-16 ***
Age         -0.21392    0.01263  -16.94   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.412 on 1175 degrees of freedom
Multiple R-squared:  0.1963,    Adjusted R-squared:  0.1956
F-statistic:   287 on 1 and 1175 DF,  p-value: < 2.2e-16
```

Figure H- 84: Summary of the initial regression model using R programming language.

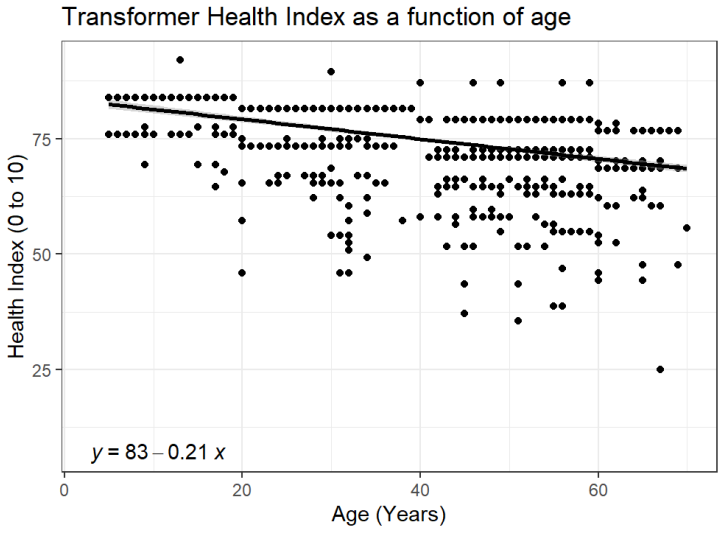


Figure H- 85: initial regression model plot.

The regression model summary in the second experiment uses the 132kV dataset, as presented in Figures H-86 and H-87. Even though the p-value is very small <2.2e-16, which can indicate that changes in HI are associated with transformer age, the R square is 0.3003, equal to 30%, which is inappropriate to use and shows a reliable regression model. The fitted regression equation is as follows:

$$HI = 85 - 0.21*(Age) \qquad \text{(H-3)}$$

R 4.1.3 · C:/Users/zain2/Desktop/Romaim/khamis NPG Data/Age HI Regression Analysis/Experiment 3/

```
Call:
lm(formula = Health_Index ~ Age)

Residuals:
    Min      1Q  Median      3Q     Max
-18.800  -2.944   1.304   3.448  14.309

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 85.33053    0.94985  89.836   <2e-16 ***
Age         -0.21259    0.02272  -9.357   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.6 on 204 degrees of freedom
Multiple R-squared:  0.3003,    Adjusted R-squared:  0.2969
F-statistic: 87.55 on 1 and 204 DF,  p-value: < 2.2e-16
```

Figure H- 86: Regression model built for 132kV transformers using R programming language.

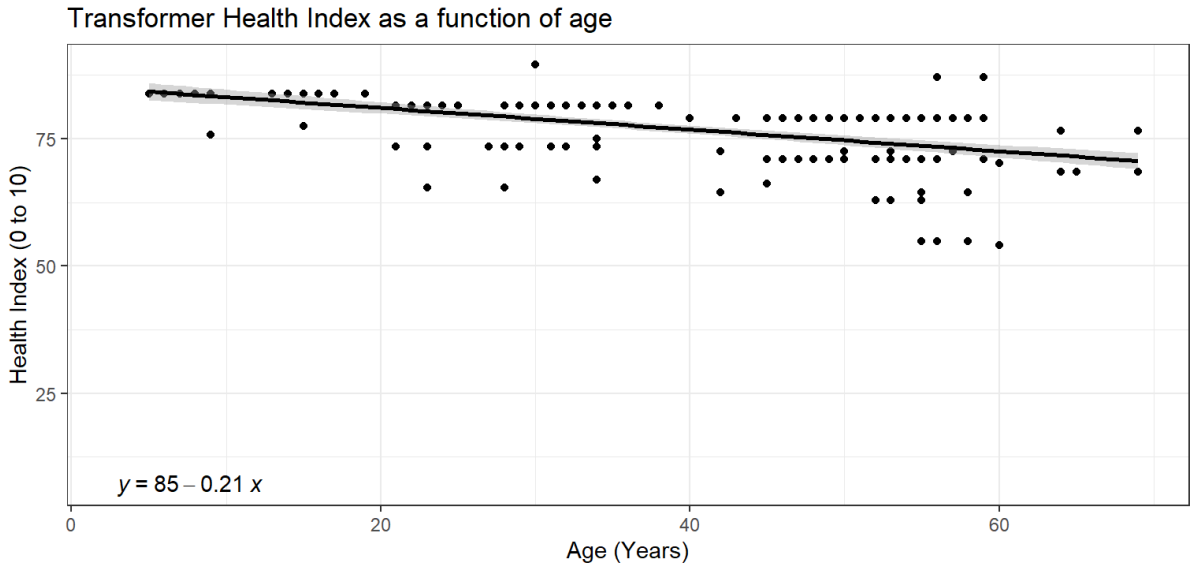


Figure H- 87: Regression model plot for 132kV transformers.

The third experiment aims to remove the outliners and perform two regression models, the first for 132kV data and the second for 66kV data. Figures H-88 and H-89 show that the coefficient R^2 value improved to 39% and 48%, respectively. However, it still cannot be considered reliable as the accuracy factor is deficient.

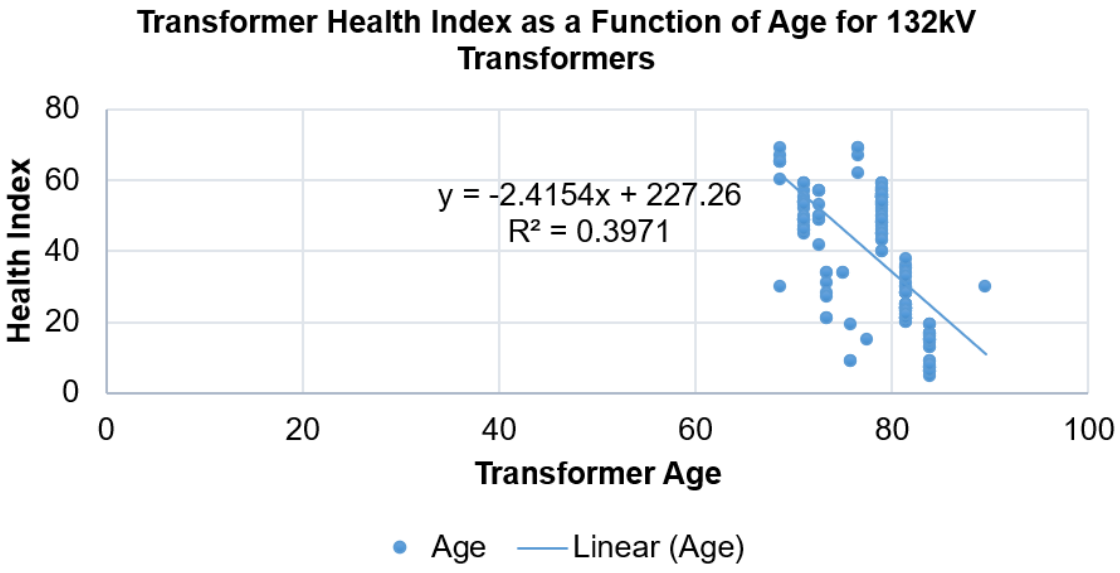


Figure H- 88: Regression model for 132kV transformers after removing the outliners.

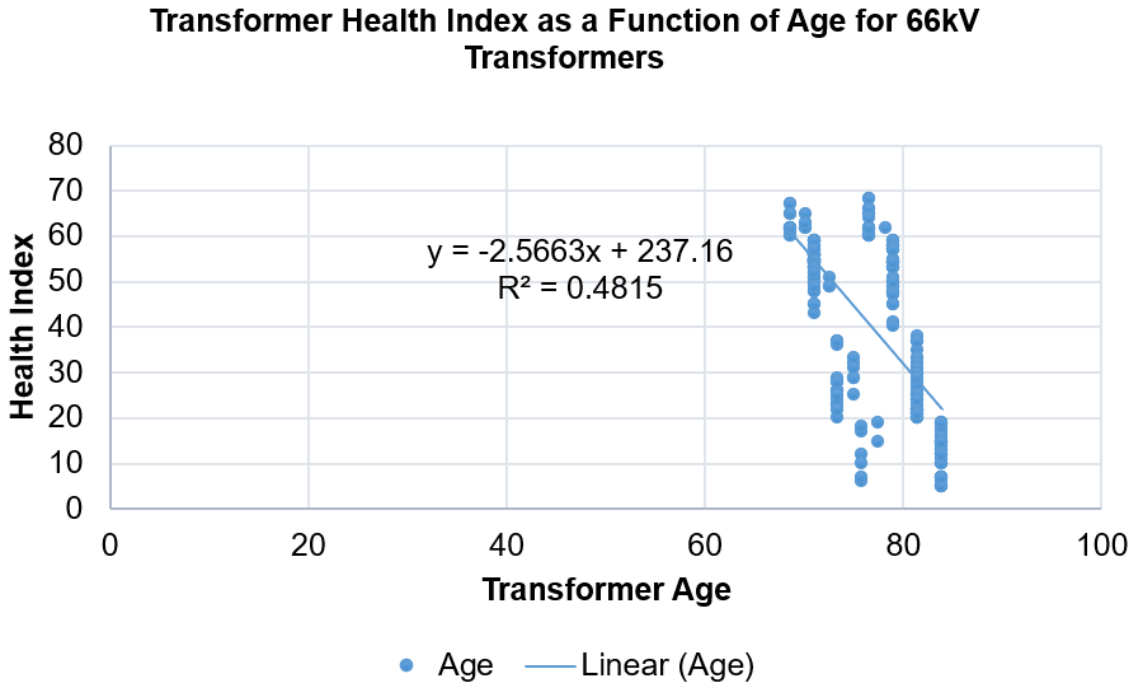


Figure H- 89: Regression model for 66kV transformers.

The fourth experiment applies the dataset received from the electrical utility employing the computed HI. The model shows high reliability and could state that it is the finest model that matches the age failure data as the coefficient R^2 value enhanced to be 0.8162, equivalent to 82% accuracy, as shown in Figure H-90. Improvement in

the accuracy of the regression model is evidence of the importance of getting all the 132kV datasets, which was not provided, especially for DGA.

Since the R^2 value is enhanced to be 82%, this model is suitable for calculating the transformer's apparent age since it is the best fit among the models built for the age HI data. As shown in Figure H-90, the p-value of the regression model is very significant (<2e-16), and the R-squared value shows a “Very Good” descriptive power (over 81% of the variability in HI explanation by age). Also, it shows that the residual standard error is 0.6575, meaning that the minor the residual typical error, the better the fit of the regression model to the calculated dataset and vice versa.

```
Call:
lm(formula = data$Health.Index ~ data$Age, data = df)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9211 -0.5161 -0.1198  0.3614  2.6380

Coefficients:
             Estimate Std. Error t value
(Intercept) -0.21562    0.05061  -4.261
data$Age     0.08099    0.00112  72.282
             Pr(>|t|)
(Intercept)  2.2e-05 ***
data$Age     < 2e-16 ***
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6575 on 1174 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8165,    Adjusted R-squared:  0.8164
F-statistic:  5225 on 1 and 1174 DF,  p-value: < 2.2e-16
```

Figure H- 90: Regression model utilising the electrical utility HI System.

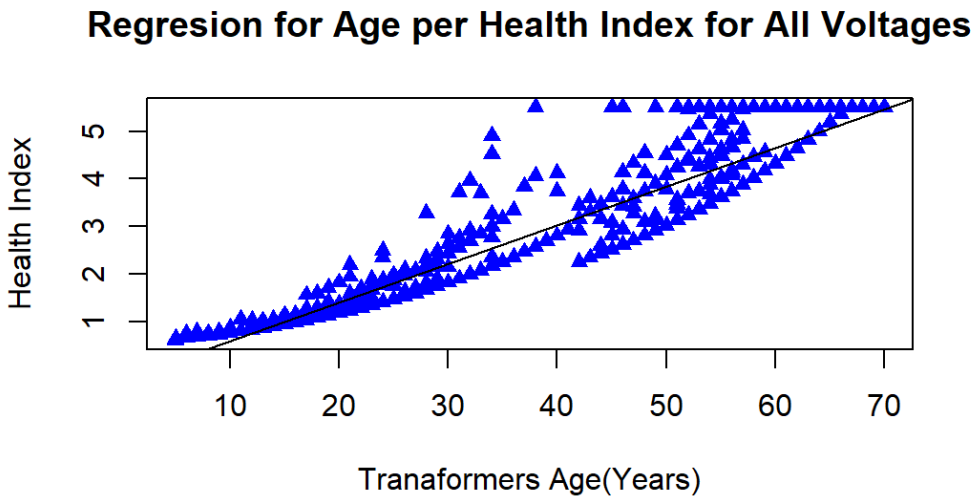


Figure H- 91: Enhanced scatterplot to better understand the model for all voltage.

Figure H-92 shows the data points according to voltage levels to understand the model further.

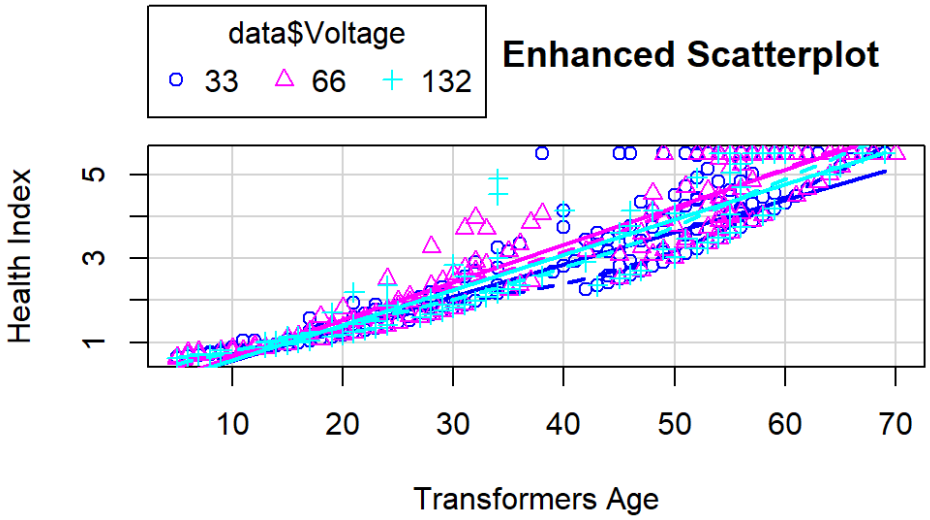


Figure H- 92: Enhanced scatterplot to better understand the model by voltage.

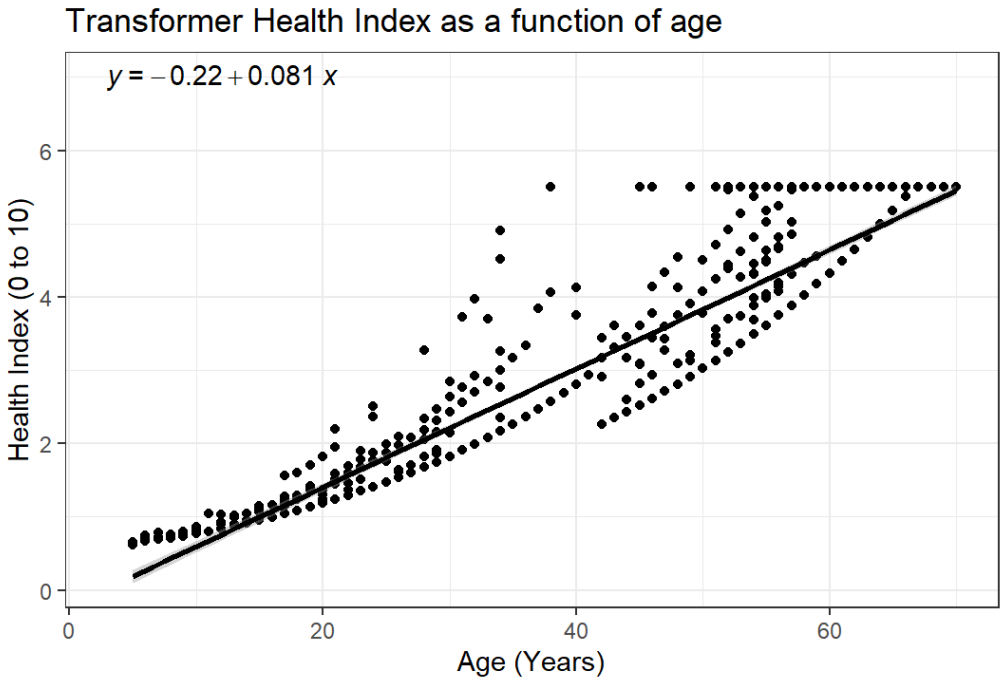


Figure H- 93: Transformer HI as a function of age.

So, the regression model after calculating intercept (a) and slope (b) as follows:

$$y = -0.22 + 0.081\,x \qquad \text{(H-4)}$$

Where y is the dependent variable (HI), x is the independent variable (Age), the intercept (value of Y when x is 0) is equal to -0.22, and the slope of the line is b and equal to 0.081.

The model plot shows age on the x-axis and HI on the y-axis. The HI value ranges from 0 to infinity, and data interpretation uses Table H-25.

Table H- 25: The electrical utility interpretation of HI value according to the HI model.

Range	Condition	Maintenance Action
0- 0.5	Very Good	Inspection
0.5 – 5.4	Good	Inspection
5.5 – 9	Fair	Repair & Refurbishment
10 -100	Poor	Replacement
>100	Very Poor	Replacement

The following Figure H-100 presents the diagnostic plots for the regression model.

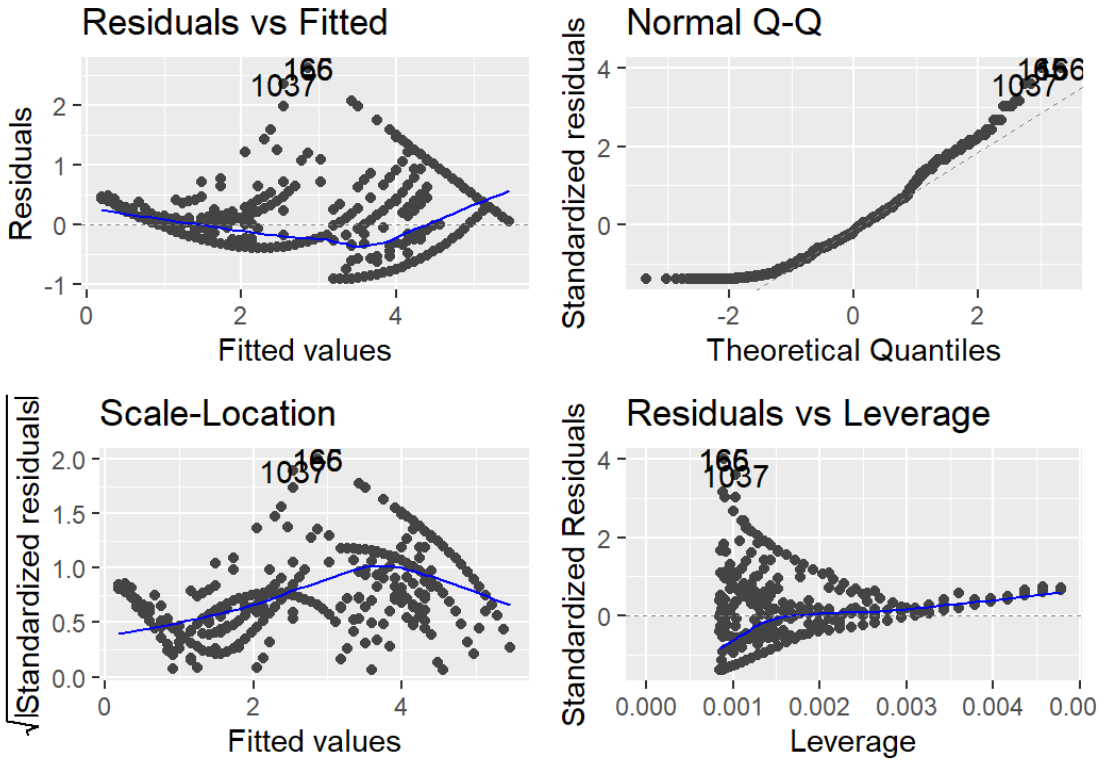


Figure H- 94: Diagnostic plots for the regression model.

Figure H-100 Residual vs Fitted, the plot in the upper left is Residual vs Fitted Diagnostic Plot for the dataset statistical analysis. This kind of plot is typically applied to verify the linear relationship hypotheses. As the plot shows, a roughly horizontal line states that a regression model suits the available transformer's data.

Figure H-100, the Q-Q plot in the upper right is applied to verify whether the regression model residuals are normal distribution. In this plot, it appears that the points fall nearly beside our regression line, which shows the normal distribution, and the spread of residual is increasing. In addition, points 1037 and 1656 are outliers, as the outliers have substantial residual rates.

Figure H-100, the plot at the bottom left is Scale-Location to confirm the hypothesis of homoscedasticity between the model residuals. A nearly horizontal line throughout the plot proves that the hypothesis of identical variance is probably confirmed. This plot shows heteroskedasticity as fitted values increase, and there is a more significant variance in the square root of standardised residuals.

Figure H-100, the plot of the right bottom is Residual vs Leverage. This plot is utilised to recognise influential observations by observing how the standardised residual spread increases as the leverage or the sensitivity of the fitted model y_i to a change in y_i. This spread means a point beyond Cook's distance conforms to an influential observation due to influential outliers such as points 1037 and 166, which negatively impact the calculated regression PTHI model. This plot shows that the standardised residuals are in decrease, demonstrating heteroskedasticity, which shows the error variance of variables over various values of periods.

Part 6 Polynomial Regression Models

The earlier selection of a linear regression model on the high positive correlation value between age and health index was (0.91). However, since it is visible in the plot, the points could be best presented using a curved line. Hence, an experiment will be necessary to solve and compare the performance of linear, second-order polynomial, and third-order polynomial regression on the same dataset. According to (Maulud & Abdulazeez, 2020), polynomial regression is a particular case of multiple regression (MLR) where the equation results in a curvilinear relationship between the dependent

and independent variables x, x^2 and perhaps even x^3,x^4 and x^5.Polynomial is simply many different terms of x raised to second order power, which is using the formula :

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_h x^h + \varepsilon \qquad \text{(H-5)}$$

where h is the polynomial degree (Jie & Zheng, 2019; Niu et al., 2018).

Comparison between the linear and second-order polynomial (quadratic) using Residual vs Fitted Plot

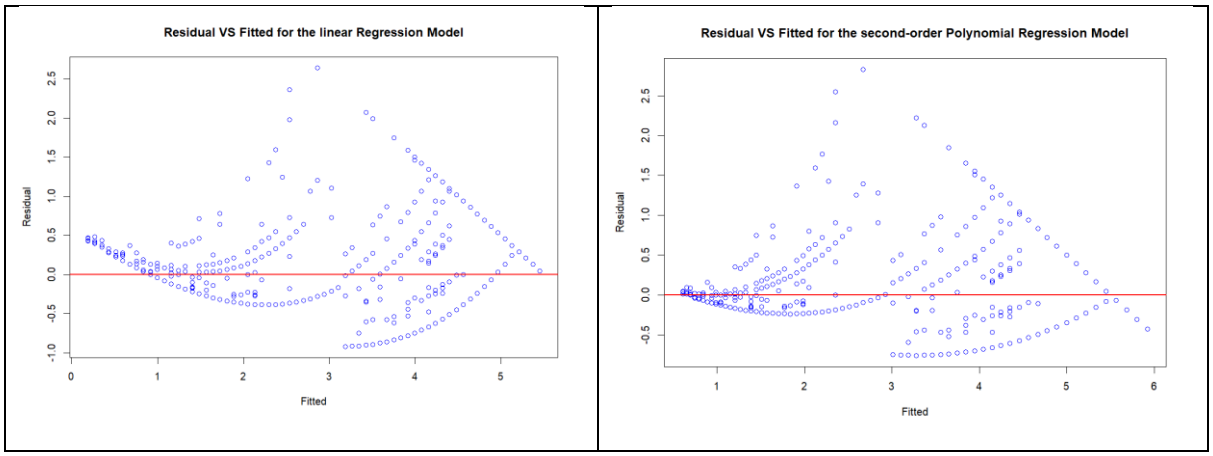


Figure H- 95: Residual vs Fitted for linear (left) and second-order polynomial (right).

Suppose the residuals had no clear pattern, were randomly spread, and were available on the horizontal axis. In that case, this is a sign that the model efficiently describes the relationship between the variables. As shown on both right and left plots, the residual shows a clear pattern with a curved shape, indicating that the linear model and even the second-order polynomial do not sufficiently describe the underlying relationship in the data. The curve pattern of the residual suggests that both models are too simple and that a complex model could fit the data better.

- 132kV Model Scoring Rules

1. To calculate DGA Factor (Jahromi et al.,2009)

Table H- 26: Grading Method for DGA Factor based FOR U<=69 kV

Individual Gases	Measurement	Min	Max	Score	Weight
H2	<=100	0	101	1	2
	100-200	101	200	2	
	200-300	200	300	3	
	300-500	300	500	4	
	500-700	500	700	5	
	>700	700	9999999	6	
CH4	<=75	0	76	1	3
	75-125	76	125	2	
	125-200	125	200	3	
	200-400	200	400	4	
	400-600	400	600	5	
	>600	600	999999	6	
C2H6	<=65	0	66	1	3
	65-80	66	81	2	
	80-100	81	101	3	
	100-120	101	121	4	
	120-150	121	151	5	
	>150	151	999999	6	
C2H4	<=50	0	51	1	3
	50-80	51	80	2	
	80-100	80	100	3	
	100-120	100	120	4	
	120-150	120	150	5	
	>150	150	999999	6	
C2H2	<=3	0	4	1	5
	3-7	4	7	2	
	7-35	7	35	3	
	35-50	35	50	4	
	50-80	50	80	5	
	>80	80	999999999	6	
CO	<=350	0	351	1	1
	350-700	351	700	2	
	700-900	700	900	3	
	900-1100	900	1100	4	
	1100-1400	1100	1400	5	
	>1400	1400	999999	6	
CO2	<=2500	0	2501	1	1
	<=3000	2501	3001	2	
	<=4000	3001	4001	3	

Individual Gases	Measurement	Min	Max	Score	Weight
	<=5000	4001	5001	4	
	<=7000	5001	7001	5	
	>7000	7001	999999	6	

2. To calculate OQA Factor (Jahromi et al.,2009)

Table H- 27: Grading Method for Oil Test Parameters based FOR 69 kV < U < 239 kV.

Individual Oil Parameters	69 kV < U < 239 kV	Rating	Min	Max	Score (Si)	Wi
Dielectric Strength kV (2 mm gap)	<=35	D	0	36	4	
	35-47	C	36	47	3	
	47-52	B	47	52	2	
	>=52	A	52	999999	1	3
IFT dyne/cm	<=18	D	0	19	4	
	18-23	C	19	23	3	
	23-30	B	23	30	2	
	>=30	A	30	999999	1	2
Acid Number	<=0.04	A	0	0.05	1	
	0.04-0.10	B	0.05	0.1	2	
	0.10-0.15	C	0.1	0.15	3	
	>=0.15	D	0.15	999999	4	1
Moisture (ppm)	<=20	A	0	21	1	
	20-30	B	21	30	2	
	30-40	C	30	40	3	
	>40	D	40	999999	4	4
Color	<=1.5	A	0	1.6	1	
	1.5-2.0	B	1.6	2	2	
	2.0-2.5	C	2	2.5	3	
	>=2.5	D	2.5	999999	4	2
Dissipation Factor	<=0.1	A	0	0.09	1	
	0.1-0.5	B	0.1	0.4	2	
	0.5-1.0	C	0.5	0.9	3	
	>=1.0	D	1	999999	1	3

3. Second Step using the second formula: 4 factors considered: Furan, Age, DGA and OQF

3.1 DGA Factor

Table H- 28: Transformer Rating based on DGA Factor

Rating	Condition	Description	Min	Max	Score	Weight
A	Good	DGAF<1.2	0	1.2	4	10
B	Acceptable	1.2<=DGAF<1.5	1.2	1.5	3	
C	Need Caution	1.5<=GAF<2	1.5	2	2	
D	Poor	2<=DGAF<3	2	3	1	
E	Very Poor	DGAF>=3	3	999999	0	

3.2 An oil quality factor (OQF) similar to DGAF is developed using the score and weighting factor.

Table H- 29: Transformer Rating based on OQF Factor

Rating	Condition	Description	Min	Max	Score	Weight
A	Good	OQF<1.2	0	1.2	4	10
B	Acceptable	1.2<=OQF<1.5	1.2	1.5	3	
C	Need Caution	1.5<=OQF<2	1.5	2	2	
D	Poor	2<=OQF<3	2	3	1	
E	Very Poor	OQF>=3	3	999999	0	

Table H- 30: Age Rating

	Range	Rating Code	Min	Max	Score	Weight
Age	Less than 20	A	0	20	4	3
	20-40	B	20	40	3	
	40-60	C	40	60	2	
	More than 60	D	60	999999	1	

Table H- 31: Furan Rating

	Description	Rating Code	Min	Max	Score	Weight
Furan	Good	A	0	0.1	4	8
	Acceptable	B	0.11	0.5	3	
	Need caution	C	0.51	1	2	
	Poor	D	1.01	5	1	
	Very Poor	E	5.01	999999	0	

Table H- 32: transformer oil colour scoring and weighting. (Jahromi et al., 2009).

	Range	Category	Score	Weight
Color	<=1.5	A	1	
	1.5-2.0	B	2	
	2.0-2.5	C	3	
	>=2.5	D	4	2

Table H- 33: Interpretation of the electrical utility data oil colour into nominal values.

Colour*	Nominal Value Range	Condition	Score
Light Amber	0 ≤ colour < 1.2	Very Good	1
Amber	1.2 ≤ colour < 1.5	Good	2
Pale Yellow	1.5 ≤ colour < 2.0	Moderate	3
Yellow	2.0 ≤ colour < 2.5	Bad	4
Dark Yellow	colour ≥ 3.0	Very Bad	5