Int. J. Adv. Res. Sci. Technol. Volume 15, Issue 7, 2026, pp.2169-2173.

IJARST

International Journal of Advanced Research in Science and Technology

https://doi.org/10.62226/ijarst20262730

ISSN 2319 – 1783 (Print)

ISSN 2320 – 1126 (Online)

Feature Importance-Based Machine Learning Classification of Distribution Transformer Health Conditions Using Oil Test Data

Omar Ahmed Mohammed Hamood Al-Shaikh*[1], Arfah Ahmad[1], Muhammad Sharil Yahaya[1] and Ahmad Jazlan Haja Mohideen[2]

[1]Faculty of Electrical Technology and Engineering (FTKE), Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia
[2]Faculty of Engineering, International Islamic University Malaysia (IIUM), Kuala Lumpur, Malaysia

omar.utm18@gmail.com

ARTICLE INFO

Article history:

Received	30 June 2026
Accepted	10 July 2026
Available online	16 July 2026

Keywords:

Transformer Health Index, Distribution Transformer, Machine Learning, Oil Test Data, Feature Importance, Condition Assessment.

Indexed in:

doi Crossref

INDEX COPERNICUS INTERNATIONAL

Google scholar

and in major libraries

ABSTRACT

Transformer health condition assessment is essential for improving the reliability of electric power distribution networks. This paper presents a feature-importance-based machine learning framework for classifying the health condition of oil-immersed distribution transformers using routine oil-test data. The dataset contained 3,373 cleaned diagnostic records after removing exact duplicates and incomplete key identifiers. Thirteen diagnostic inputs were used, including dissolved gas analysis variables, dielectric breakdown strength, moisture, neutralization value, interfacial tension, color and 2-furfural. The continuous transformer health index was converted into five practical classes: Good, Acceptable, Need Caution, Poor and Critical. Seven classifiers were compared using an 80:20 stratified train-test split, namely logistic regression, k-nearest neighbours, support vector machine, multilayer perceptron, random forest, extra trees and gradient boosting. The best model was Gradient Boosting, which achieved 85.04% accuracy and 0.851 weighted F1-score on the independent test set. Feature importance analysis indicated that 2-furfural was the most influential diagnostic variable, followed by dissolved gas indicators such as C_2H_2, C_2H_4, C_2H_6 and CH_4, as well as oil-quality variables including dielectric breakdown strength, interfacial tension, neutralization value and moisture. The results show that machine learning can provide a practical decision-support layer for rapid transformer health screening and maintenance prioritization.

Introduction:

Oil-immersed distribution transformers are critical assets in electric power networks because their failure can interrupt supply, increase maintenance expenditure, and reduce service reliability. Diagnostic oil testing is widely used because the insulation system retains chemical and electrical signatures of thermal, electrical, and ageing processes. Dissolved gas analysis (DGA), oil quality analysis (OQA) and furan analysis (FA) are therefore central to condition-based maintenance programmes for transformer fleets [1-7].

Conventional transformer health index (THI) methods combine different diagnostic indicators into a single condition score using expert-defined limits, weighting factors, and condition classes. These approaches are useful because they are transparent and easy for asset managers to apply; however, fixed thresholds may not fully capture nonlinear relations between gases, oil quality, and paper-ageing indicators [8-12]. In large fleets, a data-driven classifier can support faster screening by converting routine oil-test measurements directly into health categories.

Machine learning has been increasingly applied to power transformer diagnostics, especially for DGA interpretation, fault classification, and condition assessment [13-17]. Compared with rule-based diagnosis, machine learning methods can learn nonlinear relationships from historical data, identify influential variables, and provide repeatable decisions when many samples need to be screened. Nevertheless, model interpretability remains important in engineering applications because maintenance decisions must be justified to asset owners [18-21].

This paper develops a concise machine learning classification framework suitable for practical implementation. The main contributions are: (i) a cleaned oil-test dataset is converted into five THI-based health-condition classes; (ii) seven classical and neural machine learning classifiers are compared under the same train-test protocol; (iii) model accuracy, precision, recall, F1-score and confusion-matrix behaviour are reported; and (iv) feature-importance analysis is used to identify diagnostic variables that most influence the health-condition decision.

Related Work:

DGA standards such as IEC 60599 and IEEE C57.104 provide interpretation guidance for gases generated in mineral-oil-immersed transformers [1], [2]. OQA guides such as IEEE C57.106 and IEC 60422 describe the significance of dielectric breakdown strength, moisture, acidity, interfacial tension, and other oil properties for equipment maintenance [3], [4]. Paper ageing is commonly assessed through furanic compounds, particularly 2-furfural, because paper degradation products migrate into the oil and reflect insulation ageing [5], [22], [23].

Health-index methods were introduced to convert multiple diagnostic observations into practical asset-condition indicators. Jahromi et al. used health index concepts for transformer asset management, while Naderian et al. proposed a health index approach for power transformer condition assessment [8], [9]. Later work explored fuzzy logic, multi-criteria scoring and utility-oriented condition assessment for transformer ranking and maintenance prioritization [10-12]. These methods are valuable but still depend heavily on threshold selection and manually assigned weighting factors.

Artificial intelligence methods have also been used for transformer diagnosis. Wang et al. reviewed early applications of artificial intelligence to incipient transformer faults, and later studies applied neural networks, support vector machines, ensemble learning, and fuzzy systems for fault or condition classification [13-17]. In the broader machine learning literature, random forests, extremely randomised trees, support vector machines, k-nearest neighbours, gradient boosting, and multilayer perceptrons provide complementary modelling assumptions for classification problems [24-29]. Evaluation should include multiple metrics because accuracy alone can be misleading when health classes are imbalanced [30-33].

Existing machine learning studies on transformer diagnostics have mainly focused on dissolved gas analysis-based fault classification, conventional fault interpretation, and direct transformer health index prediction [13-17]. However, most previous studies give less attention to multi-class transformer health-condition classification using combined routine oil-test data from dissolved gas analysis, oil quality analysis, and furan analysis [1-5], [8], [9]. In addition, feature-importance-based interpretation is often limited, although it is important for explaining model decisions in maintenance planning. Therefore, this study addresses this gap by comparing seven machine learning classifiers for transformer health-condition classification and by applying feature-importance analysis to identify the diagnostic variables that most influence the final health-condition decision.

Methodology:

The proposed workflow consists of data cleaning, health-class labelling, feature preparation, model training, performance evaluation, and feature-importance analysis. The input data were obtained from routine transformer oil tests. Exact duplicate rows were removed, and records without a sample date or serial number were excluded. The final modelling dataset contained 3,373 records, 564 transformer serial numbers, and 189 sites. The sampling dates ranged from 2004 to 2015.

Thirteen diagnostic features were used as model inputs: H2, CH4, CO, CO2, C2H4, C2H6, C2H2, dielectric breakdown strength, moisture, neutralization value (NV), interfacial tension (IFT), color and 2-furfural (2FAL). These variables represent combustible gases, oil dielectric/chemical quality, and cellulose ageing. Identifier variables such as site name, transformer number, serial number, and sampling date were excluded from model inputs to avoid identity-based learning.

The health-condition label was derived from the calculated THI percentage. Five classes were assigned as follows: Good for THI greater than 80%, Acceptable for 60%-80%, Need Caution for 40%-60%, Poor for 20%-40%, and Critical for 20% or below. This structure converts a continuous maintenance indicator into operational classes suitable for practical asset prioritization.

The dataset was split into training and testing subsets using an 80:20 stratified split with a fixed random seed. Stratification preserved the class proportions in both subsets. Seven classifiers were evaluated: logistic regression, k-nearest neighbors (KNN), radial-basis-function support vector machine (SVM-RBF), multilayer perceptron (MLP), random forest, extra trees, and gradient boosting. Standardization was applied to distance-based and gradient-based models through a pipeline, while tree-based models were trained on the original feature scale.

Model Parameter Selection:

No exhaustive hyperparameter optimization was performed in this study. All models were implemented using baseline scikit-learn parameter settings to provide a transparent and reproducible comparison of commonly used machine learning classifiers. This choice was made because the main objective was to evaluate baseline classification performance for rapid transformer health screening rather than to optimize a single model extensively. Logistic regression used L2 regularization with C = 1.0; KNN used five neighbours with uniform weighting; SVM used the radial-basis-function kernel with C = 1.0 and gamma = scale; MLP used one hidden layer with 100 neurons, ReLU activation and the Adam optimizer; random forest and extra trees used 100 estimators; and Gradient Boosting used 100 estimators with a learning rate of 0.1. A fixed random seed was used where applicable to improve reproducibility.

Performance was measured using accuracy, weighted precision, weighted recall, weighted F1-score and macro F1-score. Weighted metrics account for class frequency, while macro F1 gives equal importance to each class. The best model was selected based on weighted F1-score, and its confusion matrix and feature-importance profile were analyzed.

Class	THI (%)	Maintenance meaning
Good	> 80	Normal monitoring
Acceptable	60–80	Planned monitoring
Need Caution	40–60	Closer inspection
Poor	20–40	Maintenance planning
Critical	≤ 20	Urgent attention

Table 1. Transformer health-condition categories used for classification.

Item	Value
Records before cleaning	3,392
Final modelling records	3,373
Unique serial numbers	564
Diagnostic features	13
Sampling period	2004–2015

Table 2. Dataset summary after cleaning.

Class	Samples	Percent (%)
Good	716	21.23
Acceptable	879	26.06
Need Caution	993	29.44
Poor	659	19.54
Critical	126	3.74

Table 3. Distribution of THI-based health-condition classes.

Results and Discussion:

The final data were converted into five health-condition classes. The distribution was moderately imbalanced: Need Caution was the largest class with 993 samples, followed by Acceptable with 879, Good with 716, Poor with 659, and Critical with 126. The Critical class represented only 3.74% of the data, making macro-level evaluation important.

Gradient Boosting achieved the strongest overall performance, with 85.04% accuracy, 0.852 weighted precision, 0.850 weighted recall, and 0.851 weighted F1-score. Random forest and MLP also produced competitive performance, while KNN and logistic regression were weaker, indicating that the decision boundaries between THI classes are nonlinear.

The confusion matrix indicates that most classification errors occurred between adjacent health categories. This is expected because the class labels are derived from continuous THI thresholds; samples close to a threshold may share similar oil-test profiles but belong to different classes. From a maintenance perspective, adjacent-class errors are less severe than large jumps between Good and Critical conditions, but they still require careful interpretation when decisions affect inspection intervals.

Feature importance analysis shows that 2FAL was the dominant variable in the Gradient Boosting model. This is physically meaningful because 2FAL is a key indicator of cellulose paper insulation degradation, and paper ageing is closely related to the long-term remaining life of oil-immersed transformers. Dissolved gas variables such as C_2H_2, C_2H_4, C_2H_6 and CH_4 also appeared among the influential features because they are associated with thermal and electrical stresses inside the transformer. Oil-quality variables, including dielectric breakdown strength, interfacial tension, neutralization value and moisture, further contributed to the classification because they reflect the insulating strength, ageing condition, acidity and water contamination of transformer oil. Therefore, the feature ranking agrees with established transformer diagnostic knowledge and indicates that maintenance prioritization should not rely on one test only, but should consider the combined condition of paper insulation, dissolved gases and oil quality.

The proposed framework is not intended to replace expert judgement or standard diagnostic rules. Instead, it provides a repeatable screening tool that can rank many transformer oil samples quickly and highlight the variables most responsible for the predicted health class. The model should be retrained and validated when applied to new utilities, transformer designs, or sampling practices because the present label is based on a conventional THI calculation and the available historical data.

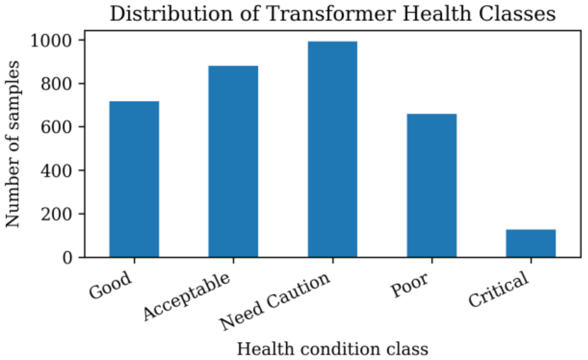

Fig 1. Distribution of transformer health-condition classes in the final dataset (N = 3,373)

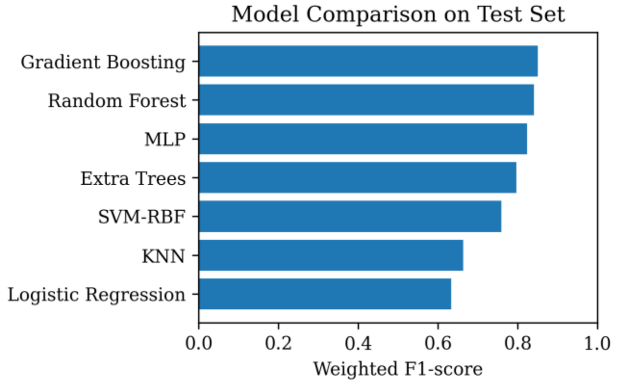

Fig 2. Comparison of machine learning classifiers based on weighted F1-score on the independent test set.

Model	Acc.	Prec.	Rec.	F1	Macro F1
GB	0.850	0.852	0.850	0.851	0.847
RF	0.843	0.844	0.843	0.840	0.803
MLP	0.824	0.825	0.824	0.823	0.810
ET	0.801	0.802	0.801	0.797	0.742
SVM	0.759	0.762	0.759	0.759	0.739
KNN	0.668	0.666	0.668	0.663	0.611
LR	0.631	0.643	0.631	0.633	0.587

Table 4. Test-set performance comparison of classification models.

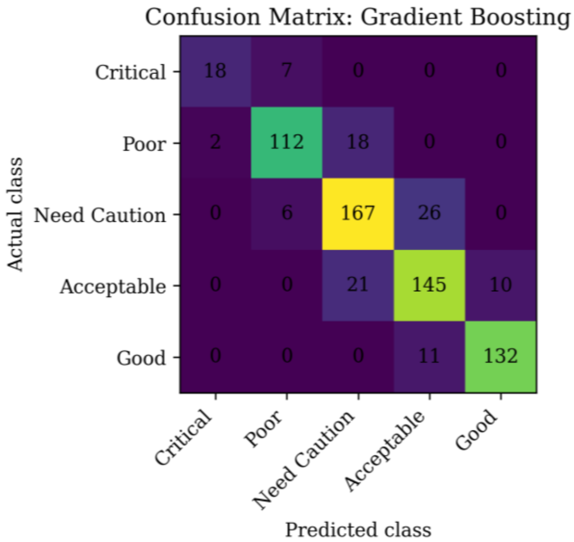

Fig 3. Confusion matrix of the Gradient Boosting model for five transformer health-condition classes.

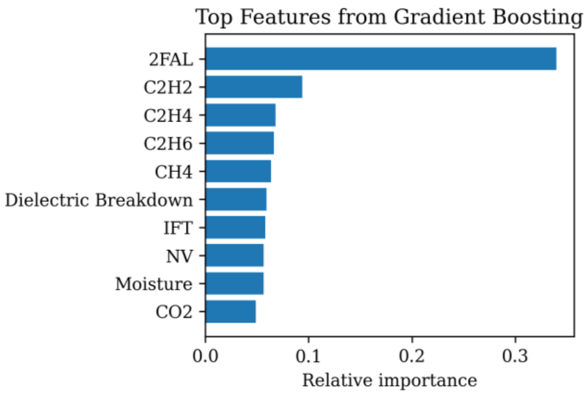

Fig 4. Top ten diagnostic variables ranked by feature importance from the Gradient Boosting model.

Rank	Feature	Importance
1	2FAL	0.340
2	C2H2	0.094
3	C2H4	0.068
4	C2H6	0.066
5	CH4	0.064
6	BDV	0.059
7	IFT	0.058
8	NV	0.057
9	Moisture	0.057
10	CO2	0.049

Table 5. Top ten diagnostic variables according to feature importance.

Conclusion:

This paper presented a feature-importance-based machine learning approach for classifying distribution transformer health conditions using routine oil-test data. Seven supervised learning models were compared, and Gradient Boosting achieved the best performance with 85.04% accuracy and 0.851 weighted F1-score on the independent test set. The most influential diagnostic variables were 2FAL, dissolved gas indicators such as C2H2, C2H4, C2H6, CH4, and oil-quality indicators such as dielectric breakdown strength, interfacial tension, neutralization value and moisture. From a practical utility perspective, the results suggest that 2FAL, DGA gases and oil-quality indicators should be prioritized for early warning, health screening, and maintenance planning. However, the model is based on historical data from a specific transformer population; therefore, retraining and validation are required before applying the framework to other utilities, transformer designs or sampling practices.

Acknowledgement:

The authors acknowledge the Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), for academic and research support.

Funding:

This work was supported by the Ministry of Higher Education (MOHE), Malaysia, under the Fundamental Research Grant Scheme (FRGS), Grant No. FRGS/1/2024/FTKE/F00579.

Conflict of Interest:

The authors declare no conflict of interest.

Data Availability:

The data used in this study were obtained from utility transformer diagnostic records and are not publicly available due to confidentiality restrictions.

References:

[1] I. IEC, "60599—Mineral Oil-Impregnated Electrical Equipment in Service–Guide to the Interpretation of Dissolved and Free Gases Analysis," *International Electrotechnical Commission: Geneva, Switzerland,* 2015.

[2] P. T. Committee, "IEEE guide for the interpretation of gases generated in mineral oil-immersed transformers," *IEEE Std C,* vol. 57, pp. 40–42.

[3] I. C. 106-, "IEEE guide for acceptance and maintenance of insulating mineral oil in electrical equipment," ed: Institution of Electrical and Electronics Engineers Piscataway, 2015.

[4] B. EN, "60422: 2013 Mineral insulating oils in electrical equipment. Supervision and maintenance guidance," *International Electrotechnical Commission, Geneva, Switzerland,* 2013.

[5] M. I. Oils, "Methods for the Determination of 2-Furfural and Related Compounds," *Standard IEC,* vol. 61198, 1993.

[6] B. IEC, "60076-7: 2018," *Power Transformers, Part,* vol. 7, 2018.

[7] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE electrical Insulation magazine,* vol. 18, no. 3, pp. 8–17, 2002.

[8] A. Jahromi, R. Piercy, S. Cress, J. Service, and W. Fan, "An approach to power transformer asset management using health index," *IEEE Electrical Insulation Magazine,* vol. 25, no. 2, pp. 20–34, 2009.

[9] A. Naderian, S. Cress, R. Piercy, F. Wang, and J. Service, "An approach to determine the health

index of power transformers," in *Conference Record of the 2008 IEEE International Symposium on Electrical Insulation*, 2008: IEEE, pp. 192–196.

[10] A. Esmaeili Nezhad and M. H. Samimi, "A review of vibration-based techniques for the condition assessment and failure detection of transformers," *Journal of Vibration Engineering & Technologies,* vol. 13, no. 1, p. 95, 2025.

[11] M. Arshad and S. M. Islam, "Power transformer condition monitoring and assessment for strategic benefits," *Curtin University of Technology Department of Electrical & Computer Engineering, Perth,* 2003.

[12] A. Abu-Siada and S. Islam, "A new approach to identify power transformer criticality and asset management decision based on dissolved gas-in-oil analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 19, no. 3, pp. 1007–1012, 2012.

[13] M.-H. Wang, "A novel extension method for transformer fault diagnosis," *IEEE transactions on power delivery,* vol. 18, no. 1, pp. 164–169, 2003.

[14] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electrical insulation magazine,* vol. 18, no. 6, pp. 12–25, 2003.

[15] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *IEEE transactions on dielectrics and electrical insulation,* vol. 10, no. 5, pp. 903–917, 2003.

[16] J. Faiz and M. Soleimani, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 25, no. 5, pp. 1798–1806, 2018.

[17] I. B. Taha, S. S. Ghoneim, and H. G. Zaini, "Improvement of Rogers four ratios and IEC Code methods for transformer fault diagnosis based on Dissolved Gas Analysis," in *2015 North American Power Symposium (NAPS)*, 2015: IEEE, pp. 1–5.

[18] A. E. Abu-Elanien, M. Salama, and M. Ibrahim, "Calculation of a health index for oil-immersed transformers rated under 69 kV using fuzzy logic," *IEEE transactions on power delivery,* vol. 27, no. 4, pp. 2029–2036, 2012.

[19] W. Cigré, "A2. 37; Transformer Reliability Survey," *Technical Brochure,* vol. 642, 2015.

[20] L. A. Zadeh, "Fuzzy sets," *Information and control,* vol. 8, no. 3, pp. 338–353, 1965.

[21] T. Hastie, "The elements of statistical learning: data mining, inference, and prediction," ed: springer, 2009.

[22] A. Emsley and G. Stevens, "Review of chemical indicators of degradation of cellulosic electrical paper insulation in oil-filled transformers," *IEE Proceedings-Science, Measurement and Technology,* vol. 141, no. 5, pp. 324–334, 1994.

[23] A. de Pablo, "Furfural and ageing: How are they related," in *IEE Colloquium on Insulating Liquids (Ref. No. 1999/119)*, 1999: IET, pp. 5/1–5/4.

[24] L. Breiman, "Random forests," *Machine learning,* vol. 45, no. 1, pp. 5–32, 2001.

[25] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning,* vol. 63, no. 1, pp. 3–42, 2006.

[26] C. Cortes and V. Vapnik, "Support-vector networks Machine learning Vol. 20," ed: Kluwer Academic Publisher, Boston, MA, 1995.

[27] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory,* vol. 13, no. 1, pp. 21–27, 1967.

[28] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics,* pp. 1189–1232, 2001.

[29] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature,* vol. 323, no. 6088, pp. 533–536, 1986.

[30] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of machine learning research,* vol. 13, no. 2, 2012.

[31] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information processing & management,* vol. 45, no. 4, pp. 427–437, 2009.

[32] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering,* vol. 21, no. 9, pp. 1263–1284, 2009.

[33] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research,* vol. 3, no. Mar, pp. 1157–1182, 2003.