energies

MDPI

Article

Autoencoder-Based Missing Data Imputation for Enhanced Power Transformer Health Index Assessment

Seung-Yun Lee, Jeong-Sik Oh, Jae-Deok Park, Dong-Ho Lee and Tae-Sik Park *

Department of Electrical Engineering, Mokpo National University, Muan 58554, Republic of Korea; seungyun8970@mokpo.ac.kr (S.-Y.L.); ferdinet105@naver.com (J.-S.O.); asfgd2000@naver.com (J.-D.P.)
* Correspondence: tspark@mokpo.ac.kr

Abstract

Data sparsity, particularly the partial loss of diagnostic data caused by sensor failures, transmission errors, or missed inspections, frequently occurs in practical power transformer operations and significantly degrades the accuracy and reliability of health index (HI) assessments. In this study, a machine learning-based HI evaluation framework is developed using key diagnostic input parameters systematically derived from failure mode and effect analysis (FMEA) and established transformer diagnostic practices. To compensate for missing data, an unsupervised autoencoder (AE)-based imputation method is introduced and benchmarked against conventional statistical supplementation techniques, namely mean and mode imputation. The experimental results, obtained using real inspection-based transformer diagnostic records, demonstrate that the AE-based approach effectively preserves inter-variable correlations and latent data structures by learning nonlinear feature relationships. As a result, the proposed method maintains robust and consistent HI classification performance under varying missing-data conditions. Furthermore, validation using confirmed transformer failure cases shows that the AE method more accurately reconstructs missing dissolved gas analysis indicators and improves the identification of high-risk equipment compared with statistical imputation. Overall, the proposed approach provides decision-consistent HI evaluations even when diagnostic data are incomplete, thereby reducing uncertainty in maintenance planning and minimizing the need for additional follow-up inspections solely to compensate for missing information.

Keywords: health index; asset management; power transformer; missing data

Check for updates

Academic Editor: Adolfo Dannier

Received: 14 November 2025
Revised: 18 December 2025
Accepted: 24 December 2025
Published: 1 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Power transformers (TRs) are critical components in power systems, as they convert voltage levels and reduce energy losses, thereby fundamentally ensuring the stability and efficiency of transmission and distribution networks. Their deterioration or failure causes issues, such as degraded power quality, increased losses, overloads, and protective-device malfunction. In severe cases, these conditions may escalate to major grid disturbances, including short circuits and large-scale blackouts. Such events extend far beyond equipment-level failures, potentially triggering cascading effects that result in industrial disruptions, infrastructural damage, and significant economic losses. Thus, maintaining the reliability and integrity of power TRs remains a fundamental priority in power system operation [1,2].

The health index (HI) assessment provides a quantitative indicator of the overall condition of power TRs by integrating multiple diagnostic and operational factors, such as estimated lifetime, early anomaly signals, and potential failure explains. By condensing

heterogeneous condition information into a single index, HI-based assessment supports preventive maintenance planning and asset management decision-making. In practical utility environments, HI values are commonly used to optimize routine inspection schedules and maintenance prioritization, thereby improving operational efficiency and reducing costs associated with regular preventive maintenance activities rather than emergency fault-recovery actions [3–5]. However, inaccurate HI estimation may result in either unnecessary maintenance or delayed interventions, highlighting the importance of reliable assessment under practical data constraints.

Power-TR HI assessment utilizes different diagnostic methods and data analysis, with typical examples including dissolved gas analysis (DGA), partial discharge (PD) measurement, and sweep frequency response analysis (SFRA). The diagnostic results provide crucial information regarding insulation degradation and potential mechanical defects inside the TR. HI derived from such diagnostic data serve as useful indicators for maintaining TR reliability and supporting preventive maintenance. In many utility practices, such diagnostic data are collected as inspection-based snapshots rather than continuously sampled sensor data, reflecting the realistic data structure used in practical asset management.

Existing HIs are typically calculated using a weighted-sum approach, where predefined scores are assigned to each diagnostic item. Weights are generally applied based on importance, after which the results are summed. Although this approach is simple and intuitive, it may exhibit limitations, particularly under incomplete-data conditions Well-known examples of this approach include the Norwegian method (Figure 1; it is used to calculate TR HI based on inputs, such as load, temperature, DGA, oil condition, maintenance history, protective device information, and TR-nameplate data) and the Det Norske Veritas and Germanischer Lloyd (DNV GL) method [6]. Further, the weighted-sum method operates on a similar principle, where input data are multiplied by weights reflecting the importance of each factor, after which the product is summed to obtain the HI.

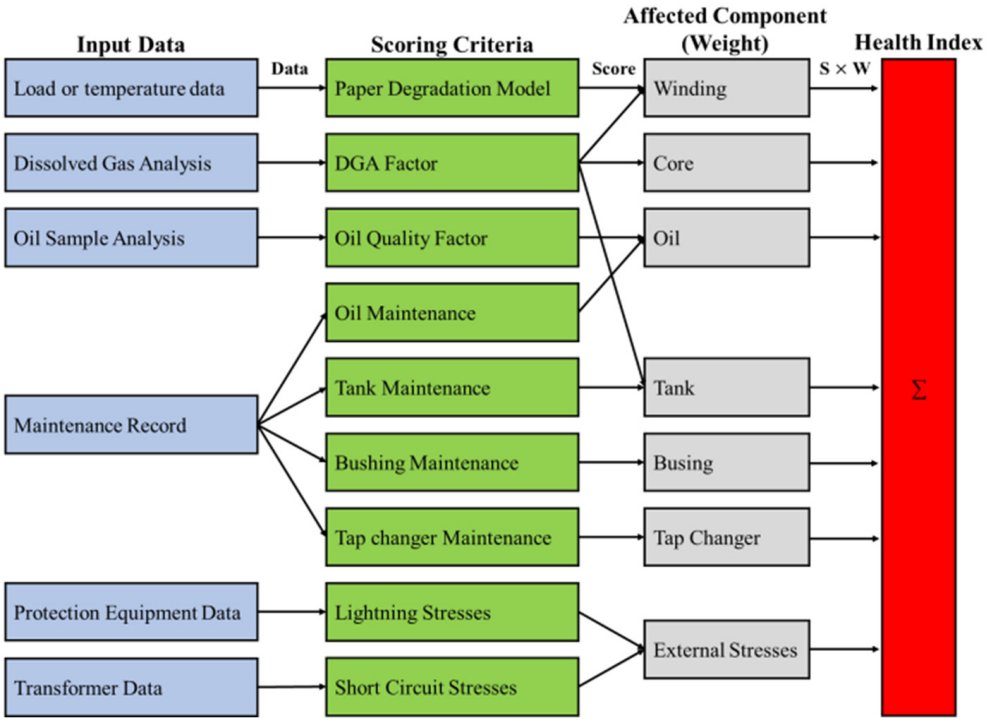


Figure 1. Norwegian Health Index (HI) Assessment Method.

The Norwegian HI assessment method exhibits inherent limitations, particularly under incomplete-data conditions. In this framework, diagnostic criteria and their associated weights are determined through expert knowledge and empirical experience, which ensures practical interpretability but also introduces a degree of subjectivity.

More importantly, when some diagnostic parameters are missing or unavailable, the weighted-sum aggregation structure may underestimate the influence of critical failure-related indicators. For instance, acetylene (C_2H_2) concentration in dissolved gas analysis (DGA) is closely associated with arcing and severe internal faults in transformer windings and cores. However, owing to the averaging nature of the weighted-sum calculation, if other available parameters receive relatively high scores, the overall HI may still be evaluated as healthy despite the presence of critical fault signals.

This effect highlights the difficulty of reliably handling missing diagnostic data within weighted-sum-based HI formulations, rather than a fundamental limitation of the HI concept itself.

Figure 2 shows the HI calculation method proposed by DNV GL. This method employs utilization, statistical, and condition data as inputs. Utilization data, typically obtained from operational records, describe the frequency and utilization mode of the asset. Statistical data, typically obtained from maintenance records, reflect general patterns from similar equipment types. Condition data reflect the physical state of the asset based on inspection and monitoring. In this model, the remaining life is estimated as the minimum value derived from the utilization- and condition-based results. By incorporating statistical data and taking the minimum value, the DNV GL method partly mitigates the critical parameter dilution inherent in the Norwegian method. However, as the underlying formulas and weighting principles remain largely expert-driven, some degree of subjectivity persists in practical applications. Additionally, this method may not fully capture the granularity of risk levels associated with different errors that can occur during health assessments.

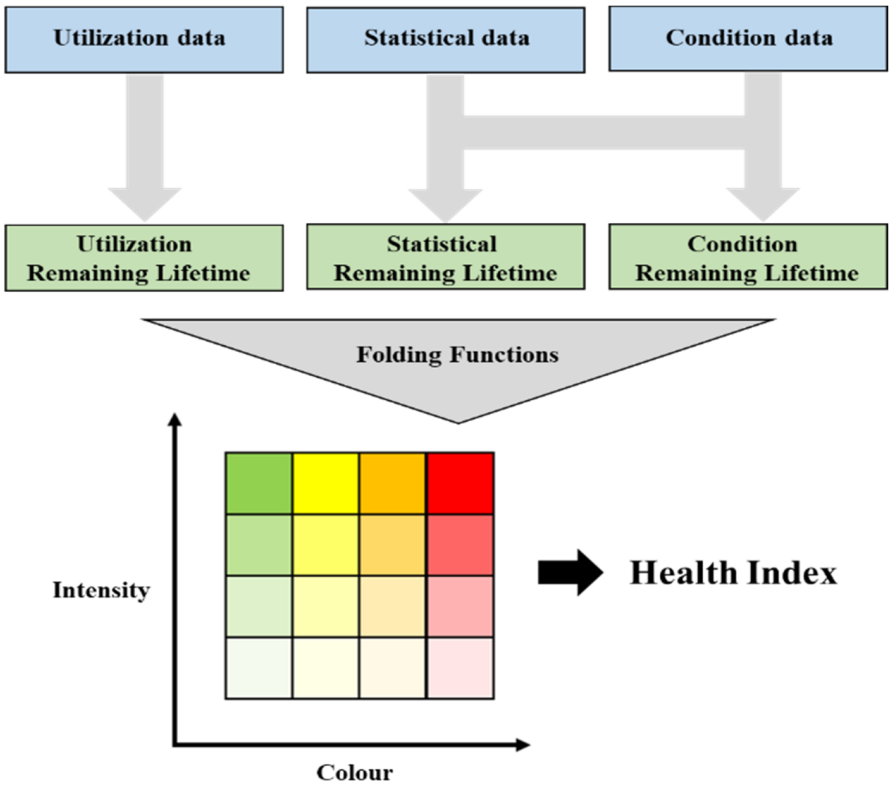


Figure 2. DNV GL HI Assessment Method.

To overcome these limitations, machine learning (ML)-based health assessment methods have been actively explored in recent years. These methods are well-suited for analyzing nonlinear relationships as they can learn complex TR conditions and abnormal patterns directly from large-scale data. Moreover, their performance continues to improve through learning, as more data become available, thereby enabling more accurate health assessments [7–9]. Nevertheless, many existing ML-based transformer HI studies implicitly assume that diagnostic datasets are complete and consistently available.

Alqudsi [10] developed an ML model to evaluate TR insulation condition using more than 1000 insulating-oil test records from two Middle Eastern power companies. The model utilized 10 oil test parameters (including H_2, C_2H_2, ethylene (C_2H_4), acidity, and moisture) as input variables to classify TR health into three categories: "good," "moderate," and "poor," by applying different pattern-recognition algorithms. A critical finding was that the model maintained high prediction accuracy even after the number of inputs was reduced by selecting only the most relevant features.

El-Rashidy [11] compared various ML models (e.g., decision trees and random forests) with deep learning (DL) models (e.g., long short-term memory (LSTM), gated recurrent unit (GRU), and hybrid LSTM–GRU) for predicting TR HI and remaining lifespan. Crucially, the author proposed a multi-task learning structure that simultaneously predicts TR HI and lifespan, achieving high prediction accuracy and improved model explainability.

These studies demonstrate the effectiveness of data-driven approaches; however, they primarily focus on scenarios with sufficiently large and complete datasets.

Existing TR-health-assessment algorithms demonstrate high performance only when all input data are complete and reliable. However, in practice, some diagnostic data are often missing or collected incompletely owing to factors, such as sensor faults, inconsistent measurement conditions, or changes in operating environments. Missing data prevent the full reflection of key health information, thereby increasing the risk of underestimating or overestimating the condition of the TR and reducing the overall accuracy and reliability of the assessment algorithm.

Particularly, machine learning and artificial intelligence-based evaluation models can achieve high predictive performance when trained with sufficiently large and complete datasets. However, as the proportion of missing data increases, their generalization performance degrades, leading to increased uncertainty in prediction results.

In practical transformer asset management, diagnostic information is often incomplete despite periodic condition monitoring. Missing data may occur due to reporting omissions, inconsistent measurement practices, unavailable test results, or variations in inspection conditions, rather than continuous sensor failures alone. Such data gaps complicate the determination of appropriate timing for health-based preventive maintenance and may result in unnecessary interventions or delayed corrective actions, ultimately increasing operational costs and risks.

To address these issues, we propose a deep neural network-based autoencoder (AE) method for compensating missing data. AEs effectively predict missing values by compressing the input data through unsupervised learning, followed by reconstructing them to a close-to-original level. Here, we develop an integrated algorithm that provides more accurate and reliable HI evaluation under missing-data conditions, specifically when combined with existing ML-based health-assessment models. The proposed AE-based approach addresses the limitations of existing assessment methods and reduces data uncertainty in practical operating environments. This is expected to improve the efficiency and reliability of TR maintenance, lower operating costs, and enable the planning of long-term asset management strategies.

2. Input Parameter Section for the Transformer Evaluation Algorithm

Input-parameter selection is crucial for predicting TR-failure probability and assessing overall TR health. Without suitable parameters, any evaluation model (ML or DL) may overlook crucial information or exhibit poor learning performance on account of missing key inputs. Selecting parameters that are strongly related to failure modes can mitigate overfitting from irrelevant data, enabling the model to focus on essential features for efficient learning. Additionally, selecting optimal input parameters lowers data processing and computational costs, ensuring improved TR-health-assessment reliability in different environments and under diverse operating conditions.

2.1. Power Transformer Failure Mode and Effect Analysis

Failure mode and effect analysis (FMEA) is a technique for assessing potential failure modes and their impacts in systems, equipment, or processes. FMEA has been applied across many industries, including product design, manufacturing processes, and medical systems; it is widely employed in the power sector. For power TRs, which are a critical component of power systems, numerous studies have detailed the application of FMEA therein.

Representative examples of FMEA for power TRs are documented in reports and standards from major organizations, such as the Institute of Electrical and Electronics Engineers (IEEE), Electric Power Research Institute (EPRI), and International Council on Large Electric Systems (CIGRE). Specifically, IEEE Std C57.140 provides a TR FMEA that categorizes subcomponents by type as well as details their possible causes, symptoms, test methods, and condition-assessment tools [12]. EPRI provides a detailed health framework, listing the effects of deterioration for each TR part, recommending detection and prevention strategies, and even including time codes for component aging to indicate the time to initial failure [13]. Conversely, CIGRE groups TR failures into three high-level categories: failure location, causes, and modes. Each category is only briefly described; the taxonomy lacks detailed classification and does not include specific diagnostic methods. Nevertheless, CIGRE identifies five major, distinct failure sites: windings, connections, mechanical structure, tap changer, and bushings [14].

Unlike time–frequency analysis-based diagnostic approaches, which typically require high-resolution signal acquisition, extensive preprocessing, and considerable computational resources, the proposed intelligent algorithm-based framework operates on inspection-derived diagnostic indicators. These indicators represent summarized outcomes of standard diagnostic tests, such as dissolved gas analysis, insulation measurements, and frequency response assessments. As a result, the computational burden during both data preparation and health index evaluation is significantly reduced, enabling scalable and repeatable assessment suitable for inspection-based transformer asset management.

Table 1 resents the FMEA for power TRs, developed by integrating insights from the IEEE, EPRI, and CIGRE, defining failure sites (winding, core, steel, oil, bushing, OLTC), along with their causes and phenomena.

Table 1. Power Transformer (TR) Failure Mode and Effect Analysis (FMEA).

Component	Sub-Component	Failure Mode	Defect Cause	Effect
Winding	Conductors Lead	Turn-Turn Fault Ground Fault Lead Fault Magnetic Insulation Fault, etc.	Through Fault Overvoltage Design Flaw, etc.	Overheating, Gassing, Short-Circuit Current, Partial Discharge

Table 1. *Cont.*

Component	Sub-Component	Failure Mode	Defect Cause	Effect
Winding	Insulation	Magnetic Insulation Fault, Short Circuit	Material, Deterioration, Static Electrification, etc.	Overheating, Gassing, Short-Circuit Current, PD, DC Discharge
Core	Steel Insulation	Magnetic Insulation Fault, Short Circuit	Eddy Current Loss Stray-Load Loss Leakage Flux, etc.	Overheating, Gassing, Short-Circuit Current, PD
Steel	Casing	Explosions, Flashes and Oil Leaks	Bolt Loosening	Overheating Marks
Oil		Oil Leaks	Lack of Antioxidant	Gassing
Bushing	Bushing	Magnetic Insulation Fault, Breakdown, Lightning, etc.	Bolt Loosening, Lack of Maintenance, etc.	Discharge, Short-Circuit, Carbonization
OLTC	Diverter	Interphase Contact Discharge, Bonding Discharge, etc.	Bolt Loosening, Wear	Crack, Flashover, Overheating, Discharge, etc.

Table 2 is a power-TR diagnostic technique that summarizes major diagnostic techniques, associating each technique with specific fault types and characteristic symptoms. Moreover, a more accurate fault diagnosis can be achieved by integrating multiple diagnostic methods.

Table 2. Power-TR Diagnostic Methods and Checklist.

Classification	Inspection Method
Basic Electrical	Winding Ratio
	Winding Resistance
	Magnetization Current
	Capacitance and Dissipation Factor/Power Factor
	Leakage Reactance
	Core Ground Test
	Frequency Response Analysis
	Polarization/Depolarization
Advanced Electrical	Frequency-Domain Spectroscopy
	Recovery Voltage Method
	Electrical Detection of PO
	Acoustical Detection of PD
	Ultra-High Frequency Detection of PD
	Dissolved Gas Analysis (DGA)

A winding-resistance test is conducted to assess TR windings for short circuits, open circuits in parallel windings, and high contact resistance at connection points. During this test, the line-to-line resistance is measured, after which the phase resistance is calculated using the appropriate phase-conversion formula. The insulation resistance (IR) test measures the insulation between windings and ground by applying direct current (DC) between the primary and secondary windings (or across all windings) of a three-phase system and the

grounded tank. The polarization index (PI) test is performed in conjunction with the IR test; the PI test is used to evaluate the contamination and moisture levels in the insulation system. In this test (PI test), the leakage current is measured by comparing the currents after 1 and 10 min to calculate the PI ratio. Furthermore, localized insulation defects can be identified by observing the so-called "kick phenomenon."

The dissipation factor (DF) test evaluates the condition and deterioration degree of a TR's insulation system by applying alternating current (AC) voltage. The DF test, often employing a Schering bridge, measures the dielectric-loss angle with the increasing AC voltage, thereby enabling the detection of impurities in the dielectric as well as the average deterioration of the insulation system. Finally, DGA detects faults by analyzing the concentrations of hydrocarbons and other gases dissolved in TR-insulating oil [15,16].

Although the diagnostic methods listed in Table 1 are traditionally analyzed using threshold-based or time–frequency techniques, such approaches often require manual interpretation and may struggle to capture complex interactions among multiple diagnostic indicators. In contrast, AI-based methods can integrate heterogeneous diagnostic information and implicitly learn nonlinear relationships across failure modes, without explicitly increasing computational complexity during routine assessment.

2.2. Power Transformer Cross Matrix

A cross-matrix is a table that provides a quick overview of diagnostic possibilities for TRs by matching failure modes with diagnostic techniques. Figure 3 illustrates a cross-matrix that combines the FMEA (summarized in Table 1) with the list of diagnostic methods (summarized in Table 2). By systematically organizing the diagnostic options for each failure mode, the cross-matrix identifies the applicability and limitations of each technique. For instance, the winding-resistance test is only effective for specific failures, such as interlayer shorts or poor connections, whereas structural issues, such as mechanical defects or iron-core damage, are detected via SFRA. Notably, DGA is considered the most effective of the diagnostic techniques; it detects most major faults and estimates fault severity by leveraging gas ratios to evaluate the occurrence temperature. Thus, DGA is a recommended core method for TR condition diagnosis by international standards (e.g., IEEE C57.104 and IEC 60599) and is widely deployed as a basic diagnostic tool in power-facility management worldwide [17,18].

It should be noted that the proposed cross matrix is constructed during the model-design stage and does not impose additional computational burden during online HI evaluation, as it serves as a structural mapping rather than a real-time calculation process.

2.3. Input Parameter Selection for the Proposed Transformer Health Index

Table 3 presents the selected input parameters for the TR HI. These parameters are rigorously derived from FMEA, key diagnostic techniques, and cross-matrix analysis. The feature set (parameters) is structured to cover the five critical areas of potential failure: insulation condition, mechanical deformation, gaseous decomposition products, electrical characteristics, and external defects, thereby reflecting diverse TR-failure mechanisms.

Particularly, the DGA variable is the first to respond to a failure, serving as a key indicator for the early detection of thermal and electrical abnormalities within the TR. Further, gaseous components, such as H_2, methane (CH_4), C_2H_2, C_2H_4, and ethane (C_2H_6), correlate closely with failure modes, including insulation deterioration, discharge, and arcing. These species are consistently classified as parameters exhibiting high prediction performance in TR-health models because of their relatively reliable measurement via DGA.

However, in this study, we augmented these sensor-based parameters with test-based inspection items, such as IR, SFRA, and short-circuit-current tests, as key input variables.

Although these items are measured less frequently and are more often missing from datasets, they provide highly reliable diagnostic data that directly reflect the physical characteristics and causes of failures when available.

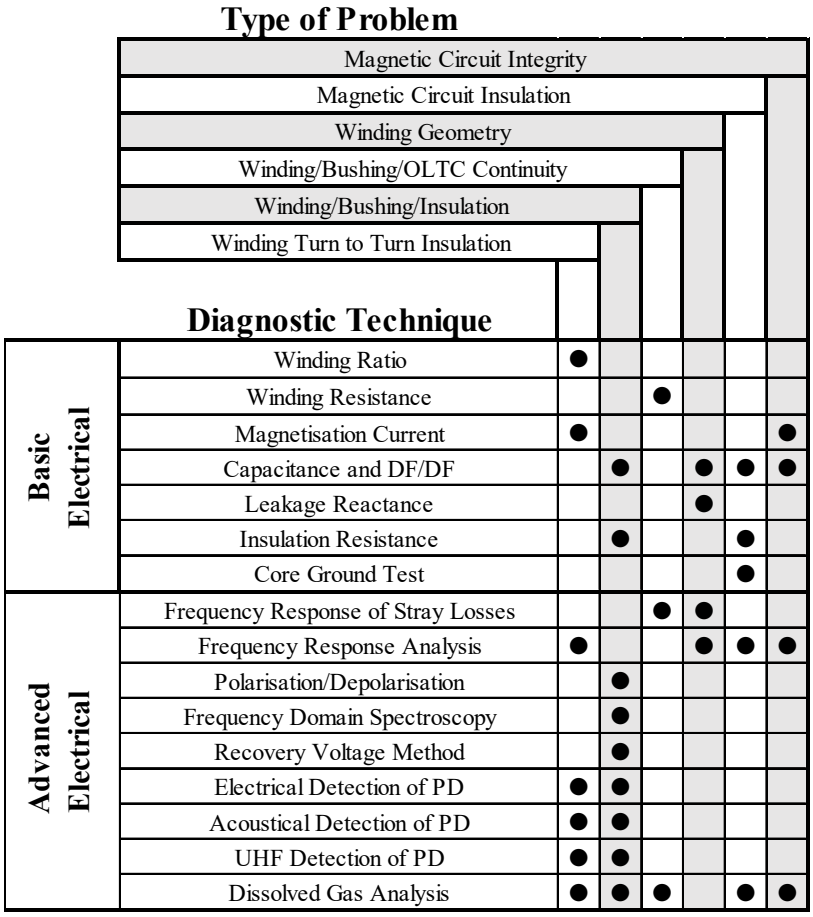


Figure 3. Power-TR Cross-Matrix.

Table 3. Proposed Power-TR HI Input Parameters.

No.	Input Parameter	No.	Input Parameter
1	Coil Displacement	12	C_2H_4
2	Winding Insulation	13	CH_4
3	IR	14	C_2H_6
4	Oil Breakdown Voltage Test	15	Total Combustible Gas (TCG)
5	SFRA	16	Gas Increase Rate
6	Double Test	17	PD Test
7	Magnetization Current & Short-Circuit Test	18	Damage Position (Insulator)
8	Angular Displacement	19	Degradation
9	Turn-Ratio Test	20	Age
10	H_2	21	Water Content
11	C_2H_2	22	IR test on a Bushing Current TR(BCR)

For example, SFRA identifies mechanical structural defects by detecting changes in the natural frequency response. Similarly, IR tests quantitatively measure the accumulated

condition of insulation deterioration. Both assessments facilitate the accurate classification of failure types. These test-based items are crucial for improving the explainability and diagnostic reliability of predictive models, revealing early failure signs or providing a clear basis for interpreting abnormal conditions.

Therefore, we developed a more precise HI-assessment algorithm by integrating sensor-based items (for the sensitive detection of early failure signs) and test-based items (for the structural evidence of failure mechanisms), thereby ensuring a balance between measurement frequency and diagnostic effectiveness.

The input parameters listed in Table 3 are derived from standard transformer inspection reports and diagnostic records, reflecting the availability of test results rather than continuously measured real-time sensor signals. Therefore, the proposed framework does not require additional sensing infrastructure beyond conventional inspection practices commonly adopted by utilities.

As the proposed input parameters are obtained from existing inspection workflows, the practical instrumentation cost impact is minimal, making the proposed approach suitable for real-world utility applications.

3. Power Transformer Health Index Assessment Algorithm Based on Existing Machine Learning Models

However, in this study, we augmented these sensor-based parameters with test-based inspection items, such as IR, SFRA, and short-circuit-current tests, as key input variables. Although these items are measured less frequently and are more often missing from datasets, they provide highly reliable diagnostic data that directly reflect the physical characteristics and causes of failures when available.

For example, SFRA identifies mechanical structural defects by detecting changes in the natural frequency response. Similarly, IR tests quantitatively measure the accumulated condition of insulation deterioration. Both assessments facilitate the accurate classification of failure types. These test-based items are crucial for improving the explainability and diagnostic reliability of predictive models, revealing early failure signs or providing a clear basis for interpreting abnormal conditions.

Therefore, we developed a more precise HI-assessment algorithm by integrating sensor-based items (for the sensitive detection of early failure signs) and test-based items (for the structural evidence of failure mechanisms), thereby ensuring a balance between measurement frequency and diagnostic effectiveness.

3.1. Machine Learning Model

Recently, machine learning (ML)-based algorithms for transformer health index (HI) assessment have attracted significant attention in condition diagnosis and maintenance applications. Unlike traditional rule-based diagnostic methods, ML algorithms can learn complex nonlinear relationships from heterogeneous diagnostic data, enabling more accurate and consistent condition assessment. Transformer faults are characterized by diverse failure modes, subtle early-stage symptoms, and strong nonlinearity, which makes data-driven approaches particularly suitable for health assessment and prognosis.

In practical transformer asset management, however, diagnostic data are typically inspection-based, limited in sample size, and collected at irregular intervals from multiple test methods. Considering these data characteristics, this study adopts classical machine learning models rather than high-capacity neural network-based classifiers. Margin-based and instance-based classifiers have been widely reported to provide stable generalization performance and robustness under small-sample and heterogeneous data conditions, without requiring large-scale training datasets.

Accordingly, support vector machines (SVMs), k-nearest neighbors (kNNs), and ensemble-based classifiers were selected, as they offer complementary strengths in handling high-dimensional diagnostic features, local similarity structures, and model variance reduction, respectively. These models are well suited for inspection-driven transformer health assessment, where interpretability, robustness, and reliability are prioritized over high-capacity model representation. Artificial neural networks, which typically require sufficiently large and diverse datasets to achieve stable generalization, were therefore not considered as the primary assessment models in this study. Figure 4 illustrates the structure of the SVM-ML algorithm.

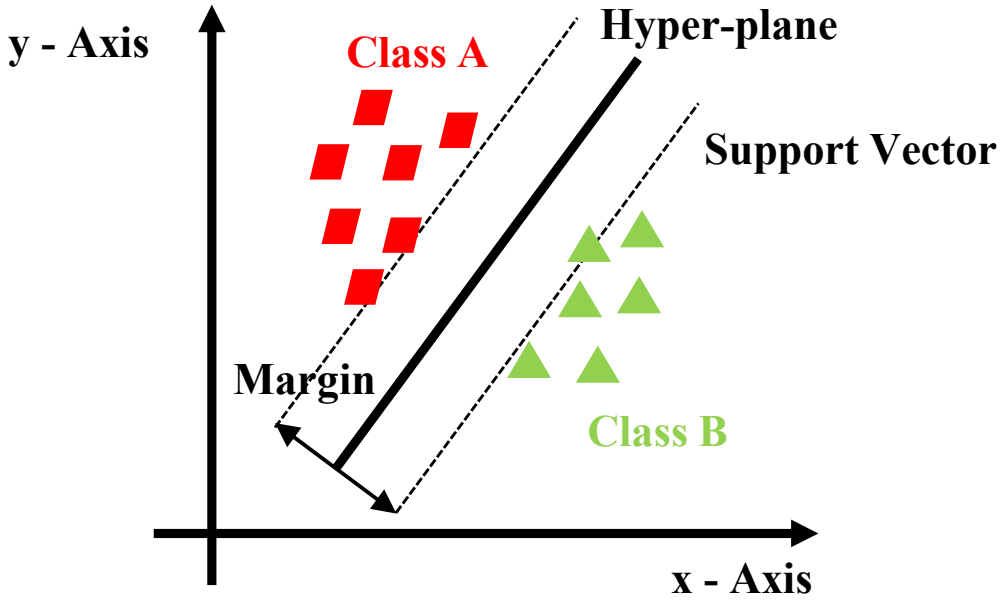


Figure 4. Support Vector Machine (SVM) Structure.

To derive the margin length, we define the classification boundary as $\omega^T x + b = 0$. The vector, ω, is the normal vector that is perpendicular to this boundary. Assuming ω is expressed as a two-dimensional (2D) vector, $(\omega_1, \omega_2)^T$, the equation of the line will be given by $\omega^T x + b = \omega_1 x_1 + \omega_2 x_2 + b = 0$, with the slopes of this line and ω being $-\omega_1/\omega_2$ and ω_2/ω_1, respectively, indicating their perpendicularity. The relationship between the vectors, x^+ and x^-, located on the positive plane can be defined using Equations (1)–(3), as illustrated in Figure 5. Figure 5 illustrates how an SVM constructs an optimal hyperplane to separate data points. This hyperplane acts as an $(N - 1)$-dimensional boundary that divides the N-dimensional feature space. The margin is defined as the distance between the two dotted lines (parallel planes), i.e., plus- and minus-planes, that are closest to the boundary. As shown in the figure, SVM determines $\omega^T x + b = 0$ by maximizing this margin [19,20].

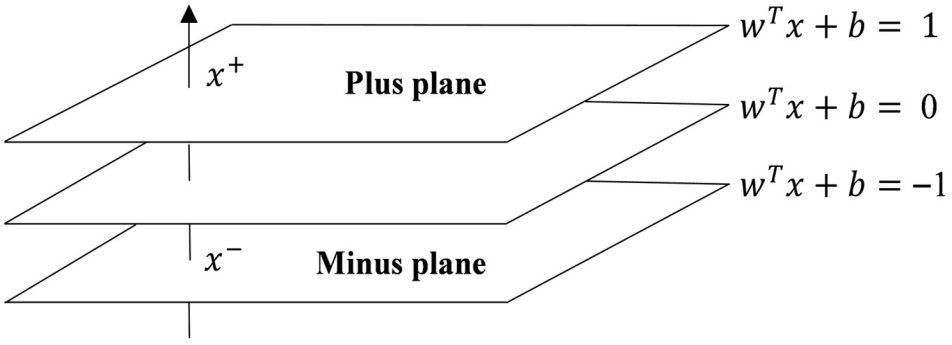


Figure 5. SVM Classification Boundary ($\omega^T x + b = 0$).

$$\omega^T x^+ + b = 1 \tag{1}$$

$$\omega^T x^- + b = -1 \tag{2}$$

$$x^+ = x^- + \lambda\omega. \tag{3}$$

Using the three conditions above, the regularization parameter λ was determined empirically through cross-validation, with the objective of balancing margin maximization and classification error under limited data availability. This data-driven selection approach avoids overfitting while maintaining stable generalization performance in inspection-based transformer diagnostic datasets. Although alternative parameter selection strategies, such as exhaustive grid search or Bayesian optimization, could be applied, these methods mainly affect fine-grained performance tuning and do not alter the overall behavior of the proposed HI assessment framework.

$$\omega^T\left(x^- + \lambda\omega\right) + b = 1$$

$$\omega^T x^- + b + \lambda\omega^T\omega = 1$$

$$-1 + \lambda\omega^T\omega = 1$$

$$\boldsymbol{\lambda = \frac{2}{\omega^T\omega}}$$

The margin can be defined by rearranging the above equation based on the relationship between vectors $\boldsymbol{x}^+$ and $\boldsymbol{x}^-$ and the value of λ:

$$\omega^T x^+ + b = 1$$

$$\begin{aligned}\text{Distance } \omega^T\left(x^+, x^-\right) &= ||x^+ - x^-||_2 \\ &= ||x^+ + \lambda\omega - x^-||_2 \\ &= ||\lambda\omega||_2 \\ &= \lambda\sqrt{\omega^T\omega} \\ &= \frac{2}{\omega^T\omega}\sqrt{\omega^T\omega} \\ &= \frac{2}{\omega^T\omega} \\ &= \frac{2}{||\omega||_2}\end{aligned}$$

Since SVM is designed to identify the boundary that maximizes the margin, the distance between $\boldsymbol{x}^+$ and $\boldsymbol{x}^-$ is defined in such a manner that it is assigned the highest possible value:

$$max\frac{2}{||\omega||_2} \rightarrow min\frac{1}{2}||\omega||_2^2. \tag{4}$$

The SVM method is well-suited to handle high-dimensional data and can operate efficiently with large datasets. In supervised learning, models are often susceptible to overfitting, becoming too closely fitted to the training data and performing poorly on real-world cases. SVM addresses this drawback by applying kernel techniques to map the training data into a higher-dimensional space, which enables more complex, yet generalized, classification boundaries and effectively mitigates overfitting.

Figure 6 shows the structure of the kNN ML, a classification algorithm that assigns a label to new data based on the k-closest neighbors. Although a simple model, kNN has been effectively deployed in fault diagnosis, demonstrating effectiveness in handling

diverse data types, such as sensor readings and image data. In practice, this algorithm identifies the NNs of the to-be-classified data point and determines its label by referring to the labels of those neighbors (e.g., using majority vote).

Selecting an optimal *k*-value for the specific dataset is crucial since it significantly affects prediction accuracy [21].

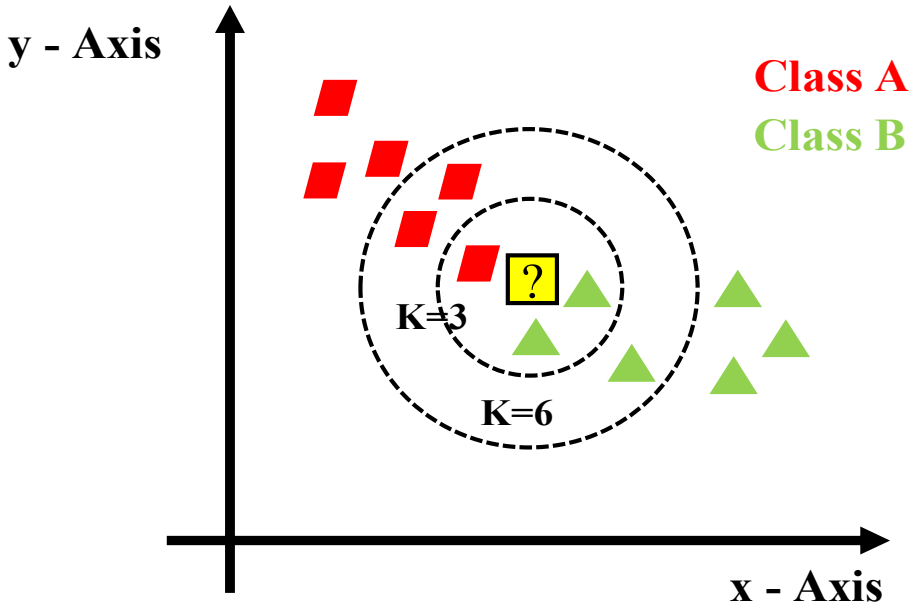


Figure 6. Kernel Nearest Neighbor (NN), kNN, Structure.

First, NNs are identified by calculating the distance between the to-be-calculated data point and its surrounding points using conventional approaches, such as the Euclidean or Manhattan distance, with the former being the most widely deployed approach. Notably, the distance between two data points, x and y, can be expressed as the Euclidean distance, as expressed in Equation (5):

$$d(x,y) = \sqrt{(x_1 - y_1)^2} + (x_2 - y_2)^2 + \cdots + (\omega_n - y_n)^2 \quad (5)$$

Neighbor selection involves choosing the kNNs of the to-be-classified data point, where the *k*-value must be set by the user. A smaller k increases the sensitivity of the model to noise and local patterns (potentially overfitting), whereas a larger one yields a simpler model with a smoother decision boundary (potentially underfitting). Thereafter, classification proceeds via majority voting, which assigns the most often appearing label among the *k*-selected neighbors. Formally, for a data point, x, and its neighbors, $N(x)$, the majority-voting rule is applied to predict the label, y, of x. For example, at k = 3, the classification result will be Class A if two of the neighbors belong to Class A and one belongs to Class B.

$$y = argmax\ (class_i) \sum I(y_i = class_i), \quad (6)$$

where $class_i$ represents a possible class. The function, $I(x)$, returns 1 and 0 if the condition is true and otherwise, respectively, corresponding to the selection of the class with the largest count among the kNNs of x. In certain cases, weights are applied based on the distance to each neighbor, ensuring that NNs exert greater influences. Thus, prediction is performed using a weighted average, as shown in Equation (7).

$$y = argmax(class_i) \sum \omega_i^* I(y_i = class_i) \quad (7)$$

In this equation, weight ω_i is defined as the inverse of the distance between the *i*-th neighbor and *x*; therefore, NNs are assigned larger weights. Although kNN is a simple and intuitive classification model, its computational cost increases rapidly with the increasing dataset, and its prediction performance can decrease if an optimal *k*-value is not selected. Additionally, class imbalance can decrease prediction performance. Therefore, the choice of *k* and the class-imbalance effect must be carefully considered when applying kNN.

In kNN-based classification, the choice of distance metric and neighborhood structure directly affects the local decision boundaries and sensitivity to noise. Smaller neighborhood sizes or overly complex distance metrics may emphasize local variations and increase the risk of overfitting, whereas larger neighborhoods tend to produce smoother decision boundaries at the cost of reduced sensitivity to incipient fault patterns.

To investigate this effect, multiple kNN variants were evaluated, including dense (fine), medium, sparse (coarse), cosine-distance–based, cubic-distance–based, and weighted kNN models. Among these variants, the weighted kNN model consistently achieved the highest classification accuracy by assigning greater influence to closer neighbors, thereby balancing local sensitivity and generalization. Consequently, weighted kNN was selected as the representative kNN classifier in this study.

Furthermore, ensemble is an ML algorithm that integrates multiple ML models to offset their weaknesses and improve overall predictive accuracy and stability. The conceptual structure of the ensemble-based health index assessment model adopted in this study is illustrated in Figure 7. In this model, the final decision is typically reached by aggregating the predictions of several models through majority voting or weighted average. Ensemble methods are generally divided into two: bagging and boosting. Bagging generates multiple datasets in three steps: resampling the training data with replacement, training each learner independently, and integrating the results by weighted averaging or majority voting. A representative example of this method is the random forest method.

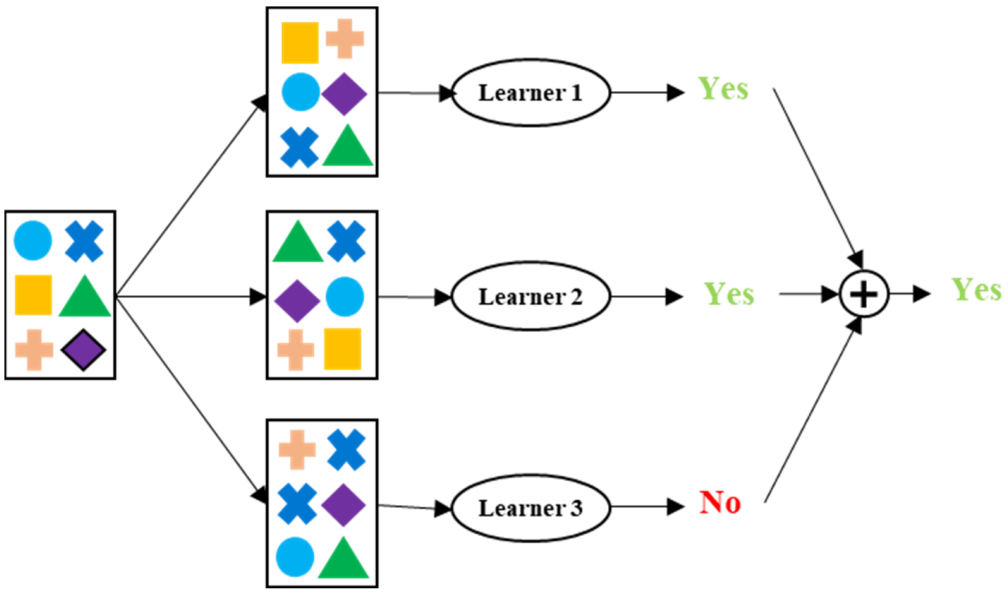


Figure 7. Ensemble Machine Learning Structure.

Conversely, boosting is a sequential learning technique, where each new model is designed to correct the errors of its predecessor(s). In this method, misclassified samples are assigned larger weights so that the subsequent learner pays more attention to them. This process reduces bias, gradually enhancing overall prediction accuracy. Dissimilar to bagging, which reduces variance by averaging independent learners, boosting constructs models sequentially to correct residual errors.

Another well-known boosting method is adaptive boosting (AdaBoost), which iteratively trains a sequence of weak learners, often decision trees, and updates weights of misclassified training examples based on their classification errors. In AdaBoost, each learner's output is combined through weighted voting, where the contribution of each learner to the final decision is proportional to its performance, and the ensemble is formed by weighted voting. However, the reliance of AdaBoost on accurately weighting errors renders it sensitive to noisy data and outliers.

Additionally, gradient boosting improves prediction by fitting subsequent learners to the negative gradient of a chosen loss function. It is widely implemented with decision trees as base learners and can be finely tuned using parameters, such as learning rate, tree depth, and number of boosting rounds. To control overfitting, regularization methods, such as subsampling or limiting tree depth, are applied [22–24].

To mitigate class overfitting under limited data availability, leave-one-out cross-validation (LOOCV) was employed, enabling effective use of all samples while providing an unbiased estimate of generalization performance.

3.2. Machine Learning-Based Power Transformer Health Assessment Algorithm

In this study, we implemented training using data from 43 TRs (TR1–TR43), encompassing normal and abnormal units, with 22 input parameters. The TR condition was classified into five distinct grades: Grade 1 (normal status), Grade 2 (normal condition but requiring imminent inspection, e.g., oil replacement), Grade 3 (requiring minor inspections, e.g., tightening of loose bolts), Grade 4 (major-component replacement), and Grade 5 (irreparable failures).

Additionally, To reliably evaluate the performance of the trained machine learning models under limited data availability, a leave-p-out cross-validation (LpOCV) strategy was employed. In this approach, p samples are randomly excluded from the dataset during each iteration and used for validation, while the remaining samples are used for training.

Compared with conventional k-fold cross-validation, which may yield unstable estimates when the dataset is small, LpOCV provides improved robustness by reducing performance variance caused by data partitioning. In contrast to leave-one-out cross-validation (LOOCV), LpOCV achieves lower computational burden by excluding multiple samples per iteration, while still mitigating class overfitting and maintaining reliable generalization performance.

Accordingly, the LpOCV method was selected as an effective validation strategy for the inspection-based transformer diagnostic dataset used in this study. The results of this evaluation are shown in Figures 8–10.

The SVM model effectively captured complex nonlinear relationships in the data and correctly classified most TRs (TRs) across normal and faulty states. It demonstrated high sensitivity and accurately predicted faulty TRs. Although there were some overestimation instances, such as with TR3 being overestimated to a higher grade than its actual status (Grade 3). Nevertheless, SVM realized the highest overall classification accuracy across the dataset and provided stable evaluation results.

Conversely, the kNN algorithm, which predicted solely based on data similarity, misclassified TR4 and TR8. Particularly, it severely overestimated TR8, an actual Grade 3 TR requiring only simple inspection, as an irreparable Grade 5 TR, resulting in unnecessary maintenance. This error occurred because of the strong effect of neighboring data on the classification performance of kNN, causing it to overreact to certain local patterns.

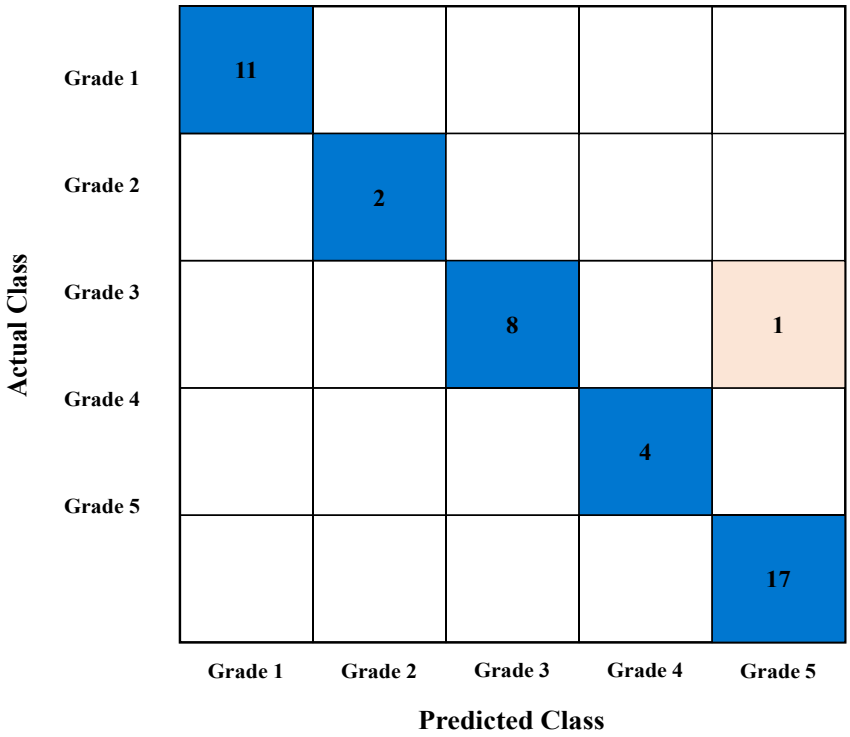


Figure 8. Conventional Method (SVM).

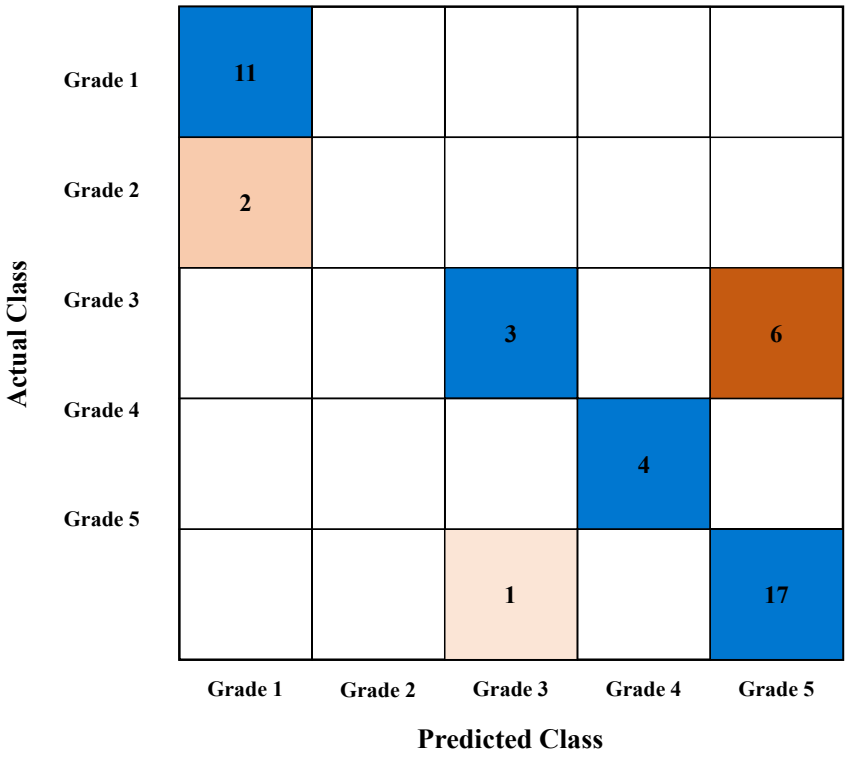


Figure 9. Conventional Method (kNN).

Ensemble ML is designed to improve accuracy by integrating multiple models. However, the accumulation of prediction errors from individual models results in the misclassification of some TRs (e.g., TR12). Particularly, the limited amount of real-world data for Grade 2 TRs restricted the models from learning sufficient features during training, thus reducing classification accuracy for that class. Consequently, although the ensemble

demonstrated high overall stability, its accuracy was lower for specific TR grades with insufficient data.

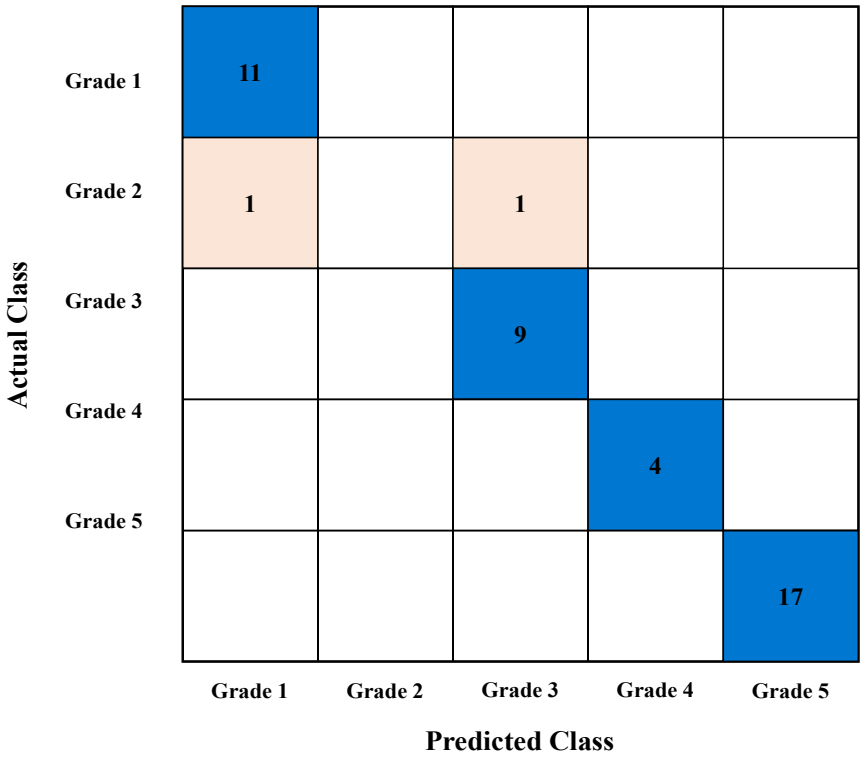


Figure 10. Conventional Method (Ensemble).

4. Effect of Missing Data in Machine Learning-Based Evaluations

4.1. Missing Data Type

Accurate and reliable data are essential for the assessment of power–TR HIs. However, various factors, such as sensor errors, transmission failures, and missed maintenance schedules, can result in missing-data occurrences. Such missing data can be generally classified into three categories: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) [25]. The conceptual differences among these missing-data mechanisms are illustrated in Figure 11. Although missing data can be categorized into MCAR, MAR, and MNAR, this study adopts the MCAR assumption for experimental evaluation in order to avoid introducing unverifiable dependencies among diagnostic parameters under limited data availability.

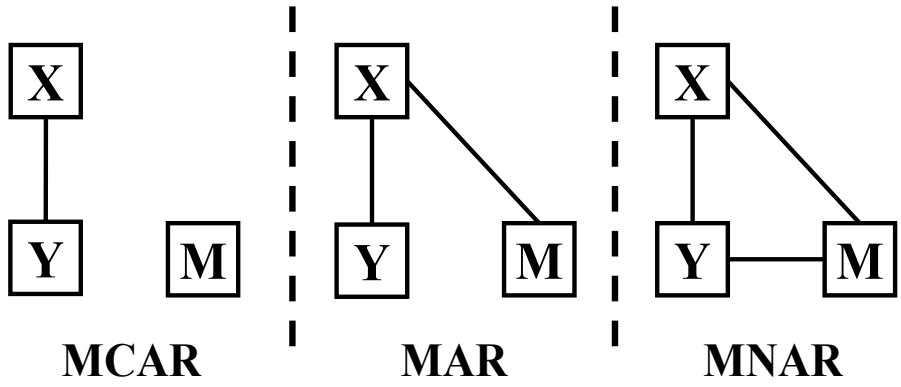


Figure 11. Missing Data Mechanism: MCAR, MAR, MNAR.

4.1.1. Missing Completely at Random

MCAR refers to cases where data are missing in a purely random manner, independent of other variables (observed and unobserved). In this case, the missingness probability does not depend on the value of any variable; thus, no systematic bias is introduced into the analysis. In inspection-based transformer datasets, MCAR-type missingness commonly arises from reporting omissions, temporary measurement failures, or incomplete maintenance records.

4.1.2. Missing at Random

MAR-type missing data refers to when the occurrence is related to other observed variables but is independent of the value of the missing data itself. Put differently, missingness depends on the relationship with other observed variables but not on the actual unobserved value. For example, some data may be missing because a particular manufacturer consistently omits specific inspections or because exciting-current and short-circuit-current tests are omitted during low-load periods. In the case of continuously operating industrial TRs, limited inspection schedules can also result in the missingness of crucial measurements, such as IR. While MAR-type missingness may occur in practical environments, explicitly modeling such dependency structures requires additional assumptions regarding inter-variable relationships, which are difficult to verify in inspection-based datasets.

4.1.3. Missing Not at Random

MNAR refers to a critical missing-data occurrence that is directly related to the value of the variable itself (the unobserved value). This happens when data are not recorded under specific conditions or when outliers are automatically removed. Further, MNAR refers to cases where missing value occurrence is due to specific, nonrandom reasons, and the unobserved values themselves influence their probability of missing. This intrinsic dependency complicates data correction and imputation as well as strongly affects subsequent data analysis. For example, if TR insulation is severely degraded, SFRA or voltage testing may be interrupted and their results not saved. Similarly, in an overloaded TR, some sensor data may fail to record owing to abnormal operating conditions.

Such missing data greatly reduce the reliability and predictive performance of health-assessment algorithms. Notably, actual failures may go undetected if the missing values correspond to key diagnostic indicators. Conversely, they may trigger unnecessary TR maintenance. Moreover, substantial missing data reduce representativeness, further degrading model accuracy. In practice, the incorrect assessment of the true condition of TRs complicates maintenance scheduling, increases unnecessary costs, and hinders effective risk assessment.

Based on these missing-data mechanisms, representative missingness scenarios were conceptually analyzed for each diagnostic parameter to provide contextual understanding of practical data-loss patterns.

In this study, MCAR examples include age, insulator damage, angular-displacement tests, bushing current TR (BCT) measurements, and control-circuit IR tests. Such missing data can be due to incomplete equipment records, missing maintenance reports, or temporary measuring-instrument malfunctions.

In this study, examples of MAR-type missing data include coil distortion, IR, excitation-current and short-circuit-current tests, and insulating-oil-withstand-voltage tests. Such missing-data occurrences may be observed in TRs from certain manufacturers that omit data collection or analysis, or when measurements cannot be conducted under specific load or operating conditions.

In this study, MNAR examples include DGA parameters, such as H_2, C_2H_2, C_2H_4, and C_2H_6. Such missingness occurs when TR insulation is too degraded to measure, or when the condition is so poor, e.g., under severe discharge, that data cannot be collected.

Here, we developed an HI calculation model using 22 diagnostic input parameters for TRs. These variables captured diverse physical and chemical conditions, including insulation status, mechanical displacement, DGA, electrical characteristics, and external fault indicators. Table 3 presents the complete list of variables. Notably, each variable was further categorized as a core parameter or an auxiliary one, depending on its diagnostic contribution and sensitivity as a model input.

Particularly, gas analysis items from the DGA series, such as CH_4, C_2H_2, C_2H_4, C_2H_6, H_2, the total combustible gas (TCG), the gas-increase rate, and bushing-oil DGA, are directly linked to major failure mechanisms, including deterioration, arc discharge, and insulation damage, within the TR. Thus, these items were defined as key variables. They exerted the greatest influence on prediction accuracy when missing, thus playing a critical role in ensuring fault-detection sensitivity and reliability in the model.

Conversely, items, such as IR, turn-ratio test, coil displacement, SFRA, water content, insulator-damage location, BCT IR, and ambient or operating conditions, were classified as auxiliary variables, as they contribute to improving predictive performance and compensatory analysis. They also enhance the interpretability of failure causes, facilitate a more precise distinction between normal and abnormal states, and strengthen the predictive stability and interpretability of the model when combined with the core variables.

Next, to evaluate algorithm performance under missing-data conditions, variables, such as usage years and DGA items (H_2, C_2H_2, C_2H_4, C_2H_6, and CH_4), were intentionally processed as missing values, reflecting real-world possibilities. The data for usage years may become missing because of the following three factors: the installation date of a TR not clearly recorded, the operating age of an older TR unit not being determined, or obscurity of management records owing to relocation or a change in operating entity.

The missingness of DGA data arises from errors or defects during the sampling and analysis of transformer oil, including sensor faults, equipment malfunctions, or sample contamination. Additionally, if a technical error occurs during measurement, or if the value of a gas item is abnormally high or low, the corresponding data may be excluded and treated as missing.

In the experimental evaluation presented in this study, missing data were generated under the MCAR assumption by randomly removing selected diagnostic values across variables. This design choice enables controlled performance comparison while avoiding unverified assumptions regarding inter-variable dependency structures.

A missing rate of 20% was selected to reflect realistic inspection-data loss levels commonly observed in utility maintenance records, while avoiding extreme sparsity that would obscure comparative model behavior.

To simulate real-world missing-data conditions, 20% of the training data were intentionally processed as missing (a 20% missing rate), after which the performance of the algorithm was compared. Figure 12 illustrates the resulting HI prediction outcomes obtained using the SVM model under this missing-data condition. The results revealed that the accuracy of the SVM model decreased to 83.7% after introducing the missing-data rate. This decrease indicated that the absence of certain critical variables can impair the final decision of the model, resulting in the inaccurate assessments (overestimation and underestimation) of TR conditions (true health grades).

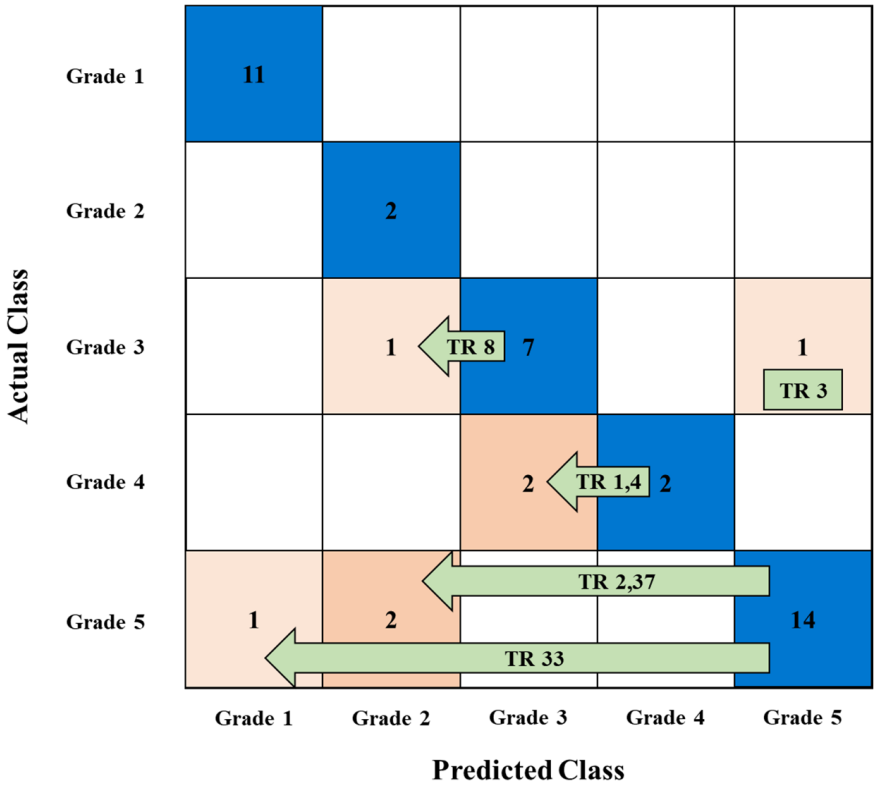


Figure 12. Conventional Method for Missing-Data Analysis (SVM).

The lack of TCG values for TR1 and TR8 caused significant misclassification errors, with their predicted ratings decreasing from the actual Grades 4 and 3 to Grades 3 and 2, respectively. This discrepancy strongly illustrates that the TCG content of a TR is a key indicator of internal abnormalities. In TR37, the absence of TCG data resulted in particularly severe misclassification, with the predicted rating decreasing from the actual status, Grade 5, to Grade 2, highlighting the critical influence of missing TCG data on HI assessment.

For TR2, the absence of C_2H_2 data caused a significant discrepancy between the actual status, Grade 5, and predicted status, Grade 2, indicating that C_2H_2 is a critical variable for identifying severe inherent electrical abnormalities.

The usage-year data of TR22, TR29, TR34, TR35, and TR40 were missing, although these cases did not significantly alter the predicted grades. However, misclassifications may arise in the event of long-term or more complex missingness. For TR3, the absence of usage-year data forced the SVM ML model to misclassify the actual Grade 3 as Grade 5.

In TR4, the absence of C_2H_4 data caused the misclassification of actual Grade 4 TR status as Grade 3. Notably, C_2H_4 represents a key indicator for assessing the degree of internal thermal stress in the TR, and its absence forced the model to misclassify the TR unit as being healthier than it actually was.

Further, in cases characterized by the missing of only H_2 data (TR10, TR18, and TR30), the predicted and actual grades were almost identical, thus exerting minimal impact. However, misclassification occurred when H_2 was missing, alongside other variables, as observed in TR22.

A more severe instance was observed in TR33, where the simultaneous absence of CH_4, C_2H_4, and C_2H_6 resulted in drastic underestimation: the actual Grade 5 (irreparable failure) was misclassified as Grade 1 (normal status), demonstrating that the ability of the model to accurately assess TR condition can be significantly compromised by the simultaneous absence of major gas variables.

4.2. Existing Missing Data Compensation Methods and Performance Evaluation

In practical utility environments, simple statistical imputation techniques, such as mean and mode imputation, are commonly employed to handle missing diagnostic data. These simple methods replace missing values by first calculating the mean (average) or mode (most frequent value) of the corresponding variable across the dataset, followed by substituting the results into the missing location. For example, assuming C_2H_2 data were missing, the mean or mode of the C_2H_2 values measured from other TRs is calculated and substituted for the missing entry. For this reason, these methods are considered representative baselines for comparison in this study. Accordingly, the performance of learning-based imputation methods should be evaluated relative to these practically adopted baselines.

This approach is characterized by its simple implementation, computational efficiency, and applicability without requiring further training. Therefore, it has been widely adopted as a fast, repeatable processing method, particularly in power-facility operations requiring large-scale asset management. However, such statistical imputation methods are limited by three inherent structural drawbacks.

First, statistical methods, such as mean and mode imputations, cannot account for the relationships, correlations, or nonlinearities among variables. Consequently, the inherent patterns and semantic characteristics of the input variables may be distorted, potentially compromising the structural consistency of the dataset.

Second, averaging-based methods (mean imputation) tend to neutralize outliers, potentially obscuring anomalous values that serve as early equipment-failure indicators. This can produce critical errors, such as the inability to detect incipient faults, ultimately undermining the reliability of the health assessment.

Third, despite the effectiveness of mode imputation for categorical variables, it delivers inadequate performance for continuous data, potentially introducing distributional distortion. For instance, assuming the data of C_2H_2, an indicator of internal TR discharge, are missing and have been replaced via a statistical imputation method, the model may incorrectly assume normal operation, even under severe fault conditions. Such misclassification of faulty equipment can delay necessary maintenance and increase operational risks.

To address these issues, we experimented with actual TR diagnostic data to quantitatively evaluate the performance of the proposed missing-data-compensation technique. We utilized 43 TR diagnostic datasets, randomly excluding approximately 20% of the input variables, including key condition indicators (including C_2H_2, CH_4, and TCG, to simulate the missingness in the experimental data).

Existing imputation methods, i.e., mean and mode imputations, were applied when missing-data occurrence was encountered. The imputed datasets were subsequently used as inputs to the same machine learning-based health assessment models, and the resulting health index classifications were compared with the actual health grades.

Conventional pointwise reconstruction error metrics, such as RMSE or MAE, were not adopted in this study because transformer health index assessment is fundamentally an interval-based classification problem. In such frameworks, even a small numerical deviation may shift an imputed value across a health-grade boundary, resulting in a completely different maintenance decision, whereas a larger numerical error that remains within the same grade interval may have little practical impact. Therefore, pointwise reconstruction accuracy may be misleading in health index–driven asset management, and classification consistency was considered a more appropriate evaluation criterion.

As summarized in Table 4, mean and mode imputations deviated substantially from the actual HI values, particularly in cases, such as TR2 and TR33, characterized by the absence of multiple critical-failure indicators. Mean imputation generally underestimated HI, potentially causing the misclassification of hazardous equipment as low-risk ones.

Similarly, mode imputation decreased the sensitivity of anomaly detection, thereby limiting the ability of the model to capture early failure signs.

Table 4. SVM-Based HI Comparison for TRs with Missing Parameters (Mean vs. Mode).

TR	Actual HI	Mean	Mode	Missing Parameter
1	4	4	5	C_2H_2
2	5	2	2	CH_4
3	3	5	5	Age
4	4	3	3	C_2H_4
8	3	5	4	TCG
33	5	1	2	C_2H_2, C_2H_4 C_2H_6
37	5	4	3	TCG

Notably, the performances of the ML prediction models varied with the applied imputation method, as illustrated in Figures 13–18. Mean- and mode-imputed datasets failed to preserve the original data distribution, forcing the SVM model to learn distorted decision boundaries and reducing its prediction accuracy. This effect is clearly observed in Figures 13–18, where mean imputation, in particular, caused the model to misclassify equipment as normal in the absence of key diagnostic parameters, such as TCG and C_2H_2 data.

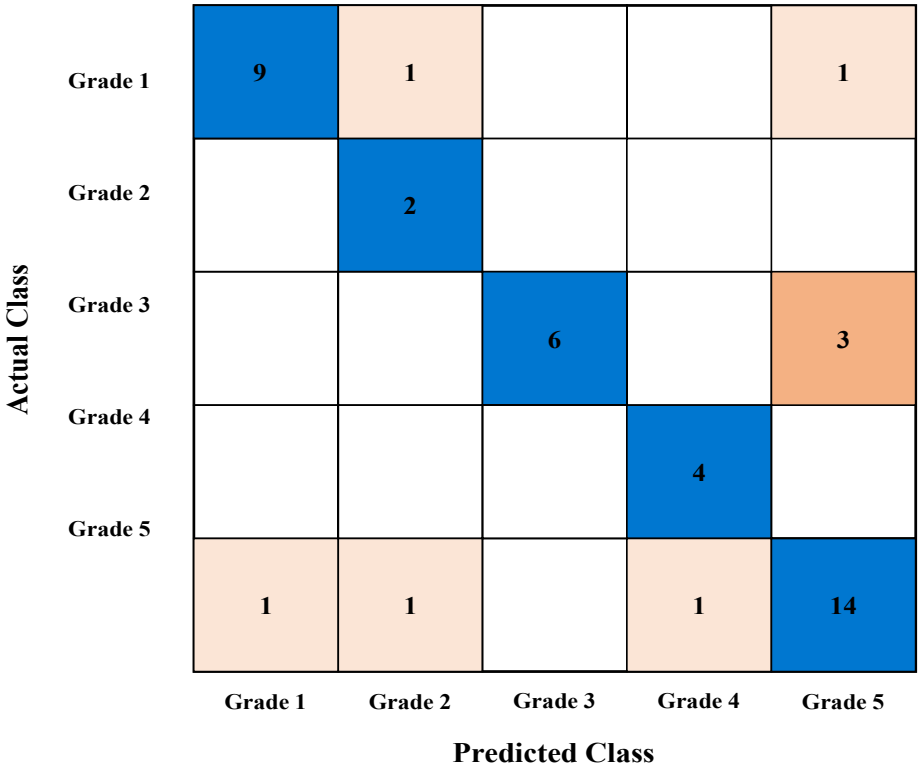


Figure 13. SVM Method (Mean Imputation).

In this experiment, we quantitatively demonstrate how simple imputation methods fail to capture fundamental data structure or the condition characteristics of equipment during imputation. Moreover, these simple imputation methods introduced larger errors with the increasing number of missing key condition indicators, resulting in the underestimation and overestimation of asset conditions. Such inaccuracies directly affect critical asset-management decisions, including maintenance prioritization and equipment-replacement

scheduling. These findings underscore the necessity of adopting more sophisticated, learning-based imputation techniques.

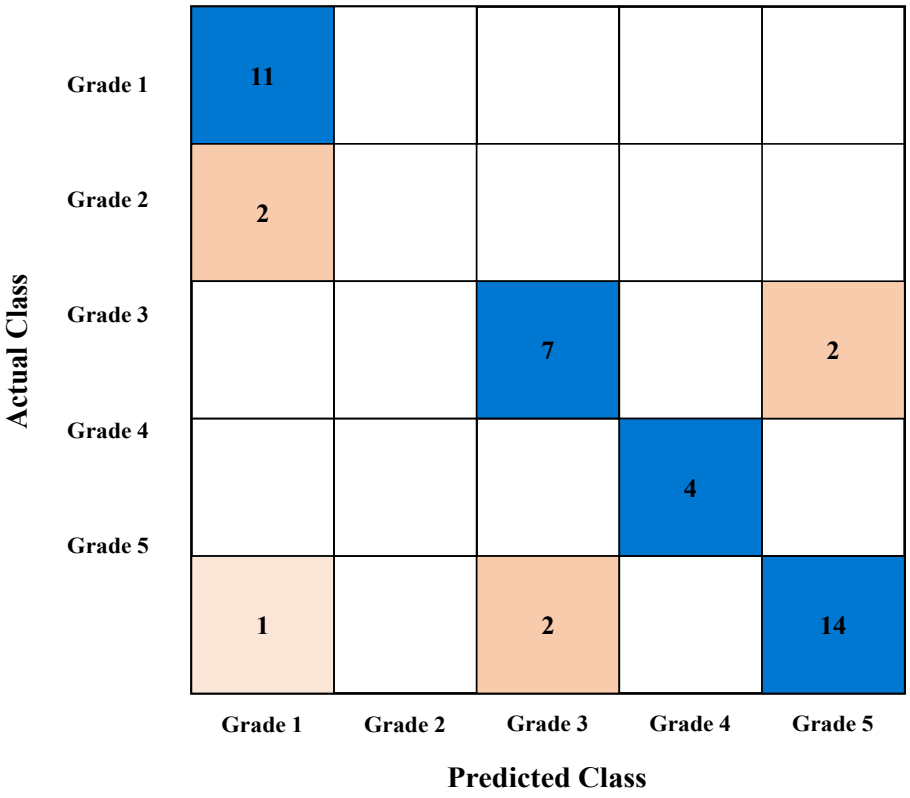


Figure 14. kNN Method (Mean Imputation).

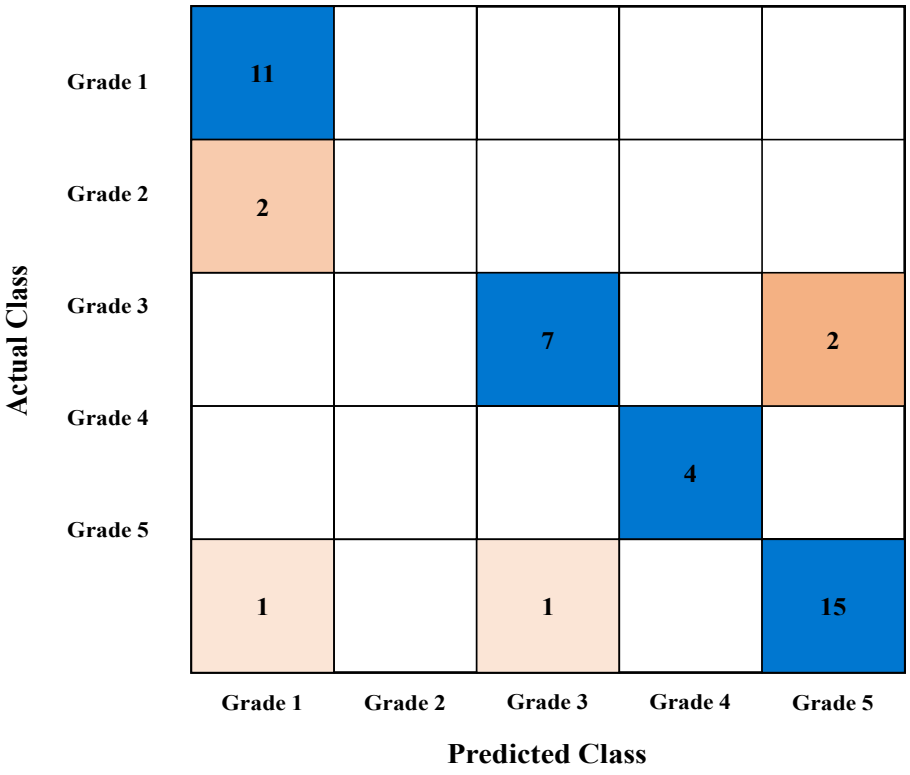


Figure 15. Ensemble Method (Mean Imputation).

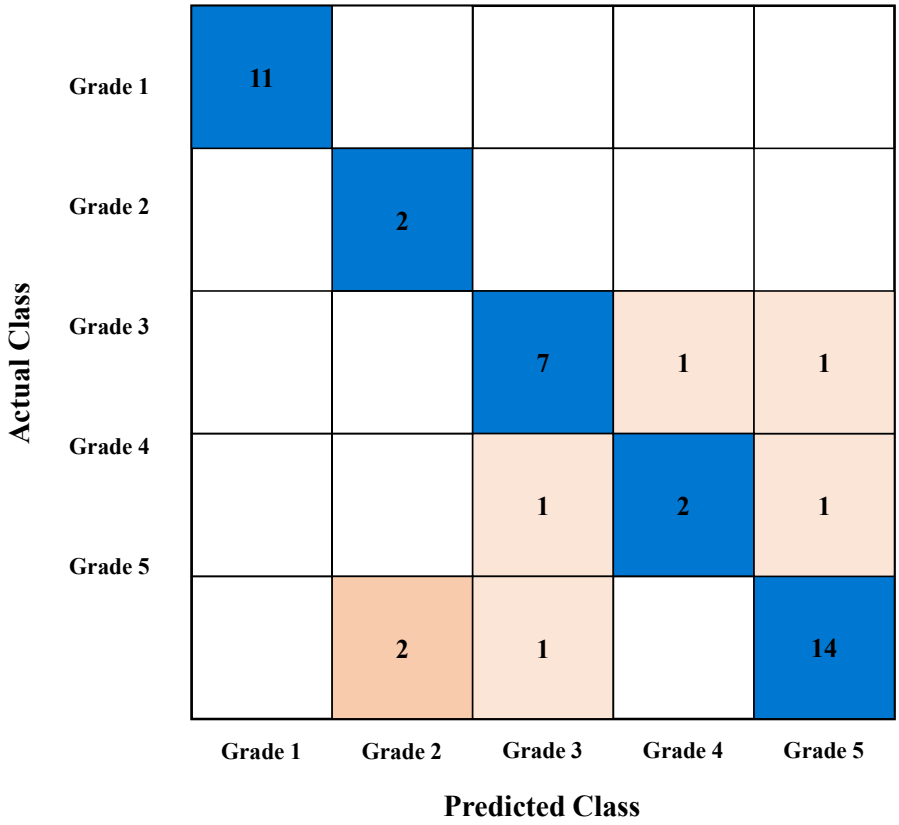


Figure 16. SVM Method (Mode Imputation).

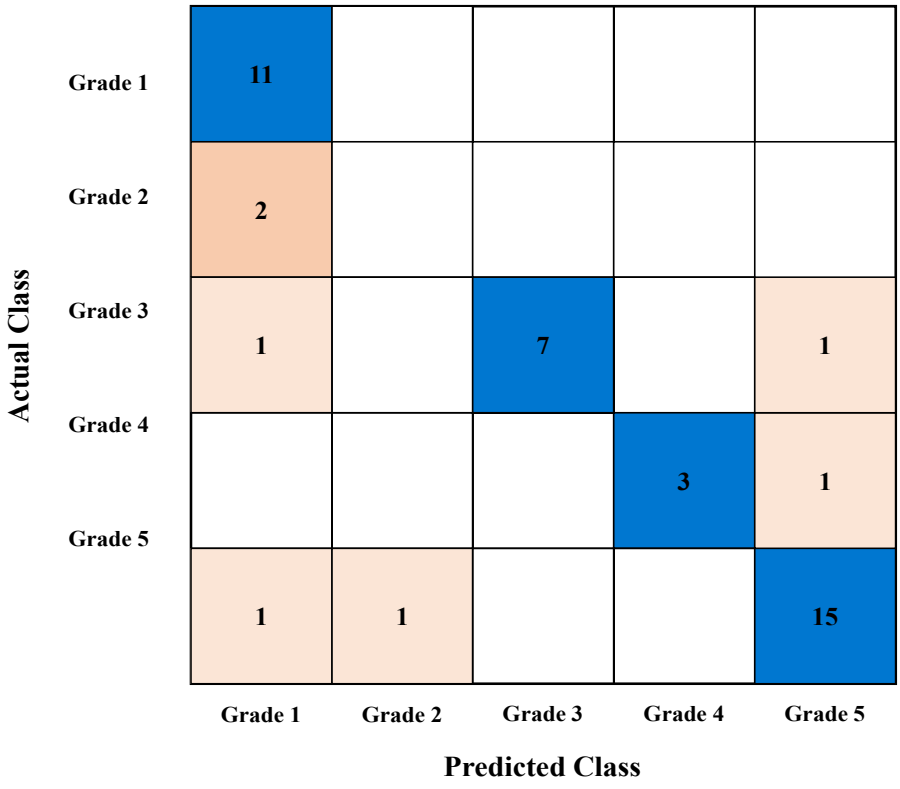


Figure 17. kNN Method (Mode Imputation).

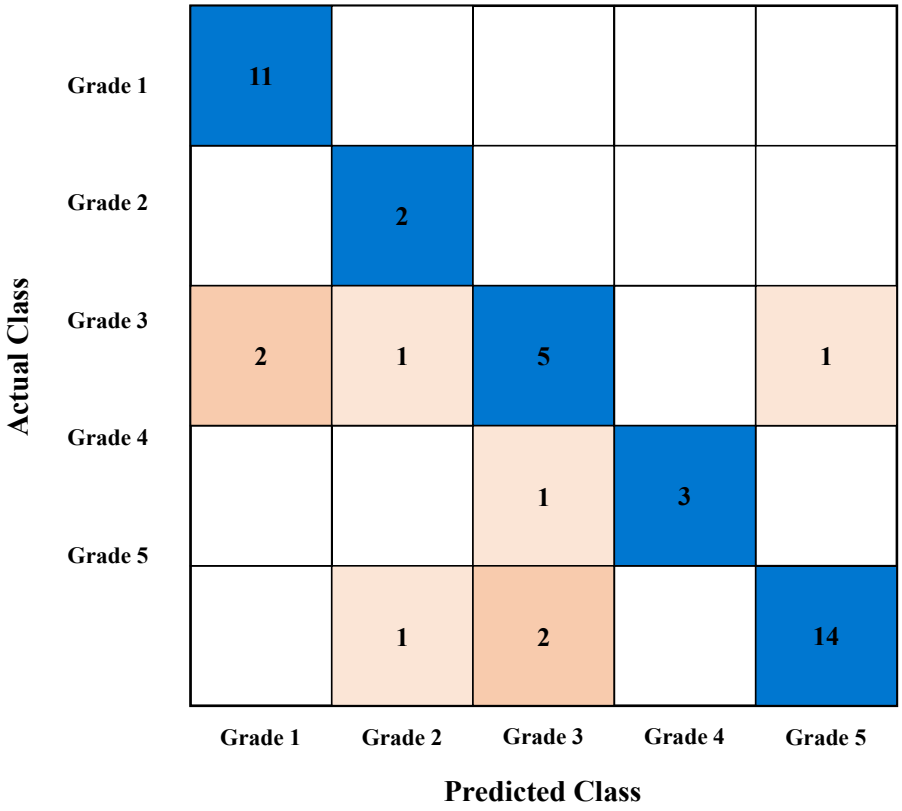


Figure 18. Ensemble Method (Mode Imputation).

5. Proposed Missing Data Compensation Method in This Study

In this study, we propose an AE-based imputation technique for the effective compensation for missing values in power-TR diagnostic data. Further, we present the theoretical background, model architecture, implementation process, and anticipated benefits of the approach. As discussed in the previous section, conventional imputation methods, such as mean or mode imputations, cannot capture complex interdependencies among variables, potentially yielding significant prediction errors. To address these limitations, we introduce an AE-based methodology that accurately compensates for missing data, thereby enabling a more reliable HI calculation. The methodology is based on an unsupervised DL model that learns correlations among multivariate diagnostic features.

Unlike conventional AE-based imputation studies that primarily focus on minimizing numerical reconstruction error, the proposed approach explicitly integrates the imputation process into a downstream transformer health index (HI) classification framework. The objective of the proposed AE is not exact signal reconstruction, but the preservation of diagnostic consistency required for reliable health-grade determination under missing-data conditions. This task-oriented imputation perspective differentiates the proposed method from existing AE-based imputation models.

Missing Data Imputation Method and Procedures Using Our Autoencoder Methodology

We enhanced the TR-health-assessment algorithm by applying the proposed AE-based missing-data-imputation method, thereby addressing the performance degradation caused by missing values [26–28].

An AE is an unsupervised learning algorithm designed to reconstruct its input at the output layer. As illustrated in Figure 19, the AE model comprises two main components: an encoder and a decoder. The encoder (consisting of the input and hidden layers) compresses

the input into a latent representation, and the decoder (consisting of hidden layers and the output layer) reconstructs the original input from this latent space.

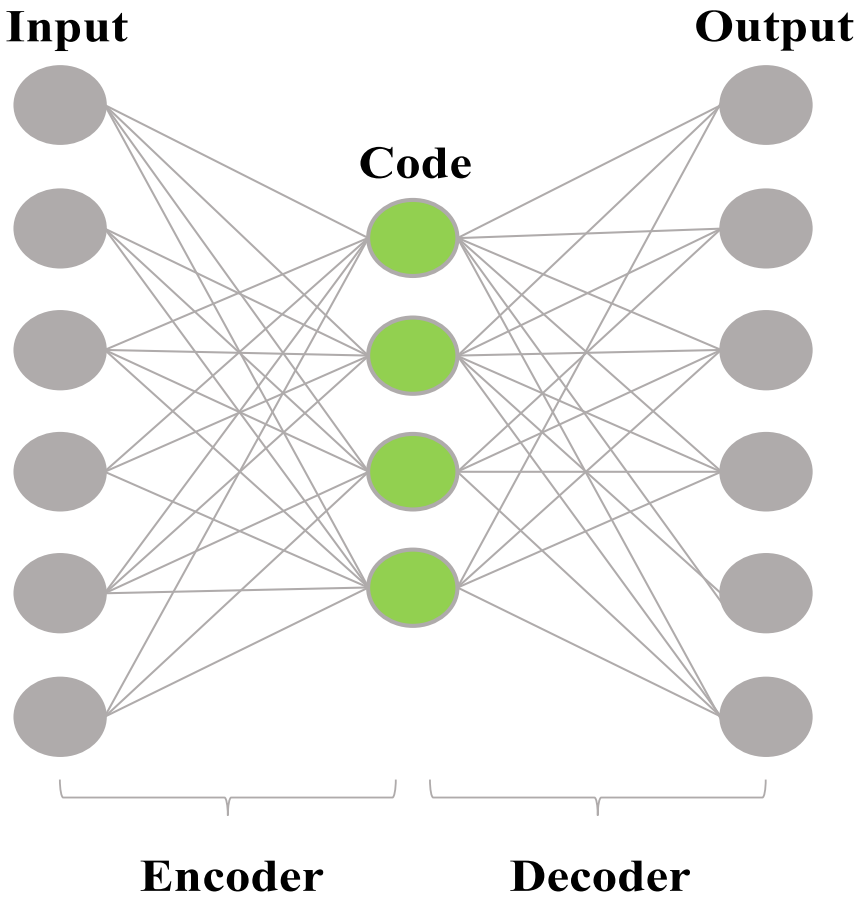


Figure 19. AE-Based Deep Learning Structure.

A basic AE comprises a single hidden layer, with equal numbers of nodes in the input and output layers. Assuming the number of nodes in each layer is nnn, mmm, and nnn, the encoder maps the high-dimensional input data, $x = \{x_1, x_2, \ldots, x_n\}$, into a low-dimensional hidden representation, $h = \{h_1, h_2, \ldots, h_m\}$, through the encoding function, f, following Equation (8):

$$h = f(x) = s_f(W_x + b). \tag{8}$$

where s_f denotes the activation function. Further, the encoder is parameterized by a weight matrix, $W \in R^{m \times n}$, and a bias vector, $b \in R^{m \times n}$, as follows:

$$x^* = g(h) = s_g(W^* h + b^*). \tag{9}$$

The decoder reconstructs the hidden representation, h, into $x^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$ via the function, g. Similarly, s_g is an activation function, and the decoder is parameterized by an n × m weight matrix, W^*, and a bias vector, b^*.

Activation functions typically leverage nonlinear functions, such as the hyperbolic tangent function or the sigmoid function.

AEs are trained to minimize the reconstruction error, $E(x, x^*)$, between x and x^*. Notably, the mean square error and cross entropy are conventionally used. Equation (10) calculates $E(x, x^*)$ based on cross-entropy (Equation (10)):

$$E(x, x^*) = -\sum_{i=1}(x_i log x_i^* + (1 - x_i) log(1 - x_i^*)) \tag{10}$$

Next, the loss function, $L(x, x^*)$, of the AE can be calculated using Equation (11) below, where λ is a regularization parameter, and W is a vector that flattens the network weights.

$$L(x, x^*) = -\sum_{x \in R^n} E(x, x^*) + \lambda ||W||^2 \quad (11)$$

The regularization term prevents overfitting during reconstruction learning and ensures stable latent representations under limited data availability. Since the AE is trained offline and only applied once for imputation, the additional computational burden introduced by the proposed method remains modest and does not affect real-time applicability of the subsequent HI assessment.

Although deep learning models are often associated with large-scale datasets, the proposed autoencoder is employed in an unsupervised manner and trained solely to capture latent correlations among diagnostic variables, rather than to perform direct classification. This unsupervised feature-learning characteristic enables stable training even with limited inspection-based datasets, as commonly encountered in transformer asset management.

The key advantage of AEs lies in their ability to iteratively extract essential features and filter out redundant information during training. During data compression and reconstruction, AEs capture inherent correlations among variables. They can restore and compensate for some missing data by leveraging the overall structure of the learned data. Dissimilar to conventional statistical methods, such as mean or median imputation, AEs more accurately compensate for missing data by reflecting the complex structure and nonlinearity of the dataset. Consequently, we adopted an AE-based technique for missing-data compensation, and the overall flow is illustrated in Figure 20.

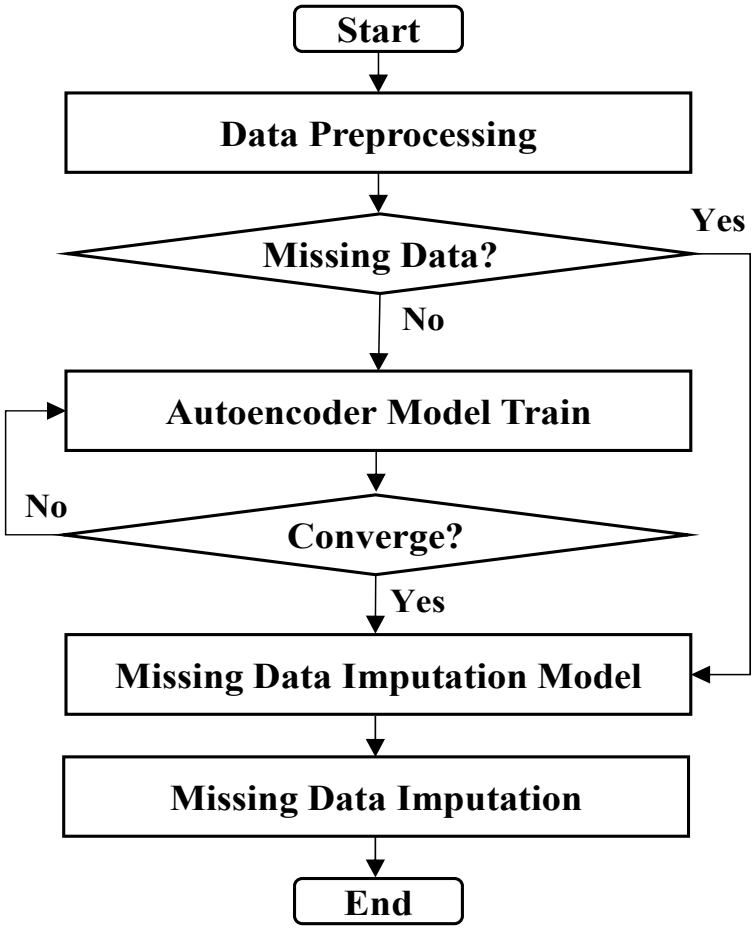


Figure 20. Proposed Power-TR HI Flowchart.

The data-preprocessing stage involves the normalization of power-TR diagnostic data, the removal of outliers, and the verification of missing-data occurrence. Thereafter, the data are categorized to reflect missing-data occurrence, and an AE model is trained using the complete subset without missing entries. This training is aimed at minimizing the reconstruction error between inputs and outputs.

Missing data detection is performed during the preprocessing stage by identifying undefined or null entries in inspection records. These detected missing positions are then masked and handled separately during the imputation stage, ensuring that only incomplete variables are reconstructed by the AE while observed values remain unchanged.

Afterward, the trained AE is applied for missing-data compensation. Specifically, the observed variables of the data are utilized as input, and the missing values are estimated using the reconstructed output. Finally, the reconstructed data with compensated values are merged into a complete dataset, which is subsequently employed as input for the power-TR health-assessment algorithm.

By compensating missing diagnostic data prior to health index assessment, the proposed AE-based framework reduces uncertainty propagation caused by incomplete inspections. This contributes to more reliable maintenance decision-making without altering existing diagnostic workflows or requiring additional sensing infrastructure.

6. Performance Evaluation of Missing Data Imputation Methods

Further, we analyze the impact of missing data on power-TR HI calculations and evaluate the performance of different imputation techniques. Particularly, we performed quantitative and visual analyses to compare conventional statistical methods (mean and mode imputations) with the proposed AE-based imputation technique.

Our experiment was conducted using real diagnostic data from 43 domestic TRs comprising 22 input parameters. The parameters were scaled to the [0, 1] range via min–max normalization, a preprocessing step that ensures consistent AE training and comparability across techniques.

Additionally, missing values were randomly assigned across all variables at missing rates of 5%, 10%, 20%, 30%, 40%, and 50%. For each missingness level, three imputation techniques (mean imputation, mode imputation, and AE reconstruction) were applied. Thereafter, the HI was predicted using the SVM, kNN, and ensemble ML models, followed by the evaluation of prediction errors by comparing the predicted and actual HIs.

Among the experiments, the 20% missing-rate case was selected as a representative example, and the restoration performance was deeply analyzed for key parameters exhibiting high missingness, particularly DGA items, such as CH_4, C_2H_6, and C_2H_2. Figure 21 shows the heatmaps of four data states, enabling an intuitive comparison of the restoration performance achieved by each method.

Unlike conventional missing-data studies that focus on pointwise numerical reconstruction accuracy, this study evaluates imputation performance based on its impact on downstream health index classification. Transformer HI assessment is inherently an interval- and grade-based decision problem, where even a small numerical deviation may shift an imputed value across a grade boundary and lead to a completely different maintenance decision. Conversely, a larger numerical error that remains within the same grade interval may have limited practical impact. Therefore, classification consistency and decision robustness were considered more appropriate evaluation criteria than conventional error metrics such as RMSE or MAE.

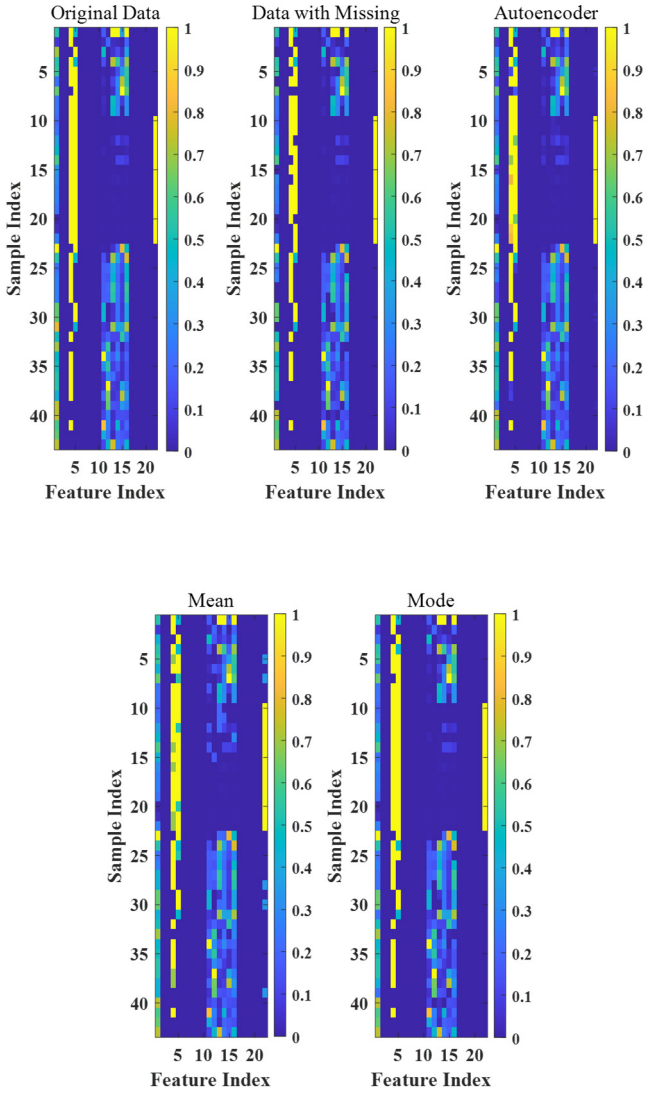


Figure 21. Power-TR Dataset Heatmap (Missing Rate of 20%).

6.1. Data States

6.1.1. Original Data

The first data state corresponds to the original data without missing values, and serves as the baseline for evaluating the similarity of the restored datasets with respect to the relative magnitudes and patterns across all variables.

6.1.2. Data with Missing Values

The second data state represents data with missing values randomly inserted at approximately 20% of the original entries. These missing values are processed as NaN (not a number), after which they are displayed visually. This visualization ensures an intuitive understanding of the extent and distribution of the missing data introduced for the experiment.

6.1.3. Autoencoder

This Figure 21 illustrates the data reconstructed using the proposed AE-based algorithm. Most of the original data patterns and structures were recovered. Particularly, the restoration achieved high fidelity in reproducing the patterns present before the missingness because the process preserves the inter-variable relationships learned through latent representations in the multidimensional space.

This indicates that the AE preserves inter-variable relationships through latent representations learned in a multidimensional space, rather than performing simple pointwise substitution.

6.1.4. Mean/Mode

The mean/mode-imputed datasets are visualized as heatmaps. As these methods replace missing entries with simple statistical values for individual variables, they do not consistently preserve the overall structural patterns of the data, and local distortions emerge. These visual analysis results intuitively demonstrate that the AE-based approach provides superior restoration performance compared with statistical imputation.

6.2. Results and Performance Analysis

Here, we quantitatively evaluate the impact of increasing missing rates on the predictive performances of different supplementation techniques, as summarized in Figure 22. To do this, the AE, mean, and mode imputation methods were tested at six missing rates: 5%, 10%, 20%, 30%, 40%, and 50%. Thereafter, the imputed datasets were applied to the same HI-prediction model to compare the overall accuracy. We employed three representative ML models: SVM, kNN, and Ensemble, for the analysis and assessed their performance trends across the supplementation methods at increasing missing rates.

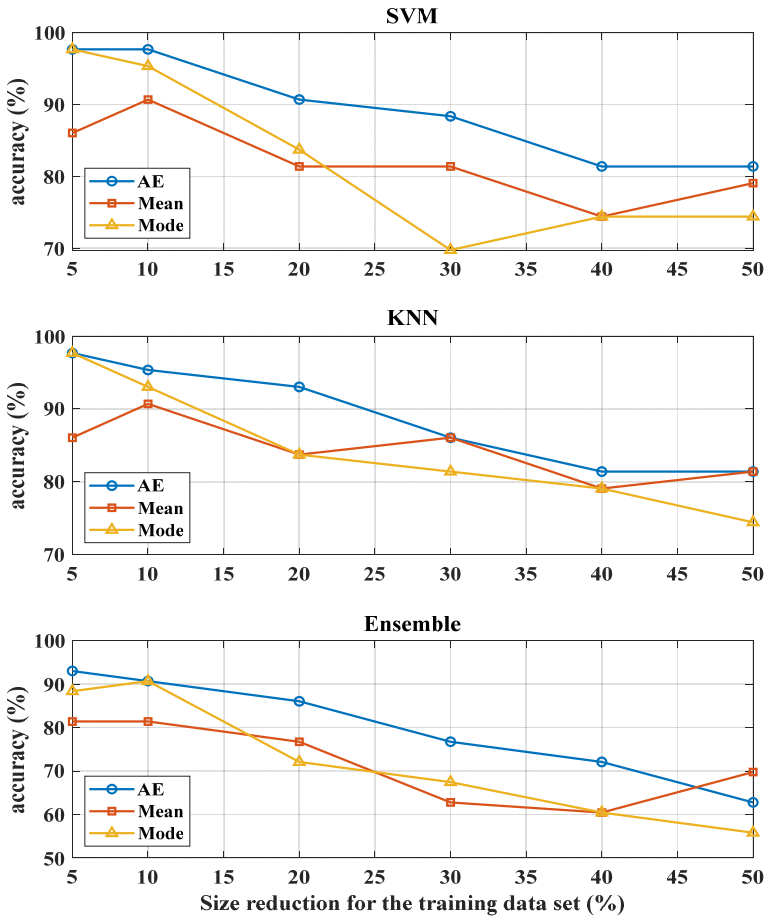


Figure 22. Accuracy comparison of AE, Mean, and Mode imputation methods with varying training-data reduction levels.

It should be emphasized that the reported accuracies reflect health index classification performance after imputation, rather than direct numerical reconstruction accuracy of missing values.

AE supplementation consistently achieved the highest accuracy across all missing rates in the SVM model. At 5% and 10% missingness, AE and mode imputations realized overall accuracies of 97.67% each, whereas mean recorded lower values of 86.05% and 90.70%, respectively. At 20% missingness, AE maintained a score of 90.69% compared with the mode (83.72%). Notably, when the missing rate exceeded 30%, mode suffered a sharp decrease to 69.77%, whereas AE maintained a high score, at 88.37% (~18.6% points higher). Even at 50% missingness, AE sustained a strong accuracy performance (81.39%), outperforming the mean (79.07%) and the mode (74.41%). These results confirm that AE supplementation effectively preserves correlations and nonlinear structures in the data, enabling stable performance, even at high missingness levels.

A similar trend was observed for the kNN model. Here, AE only exhibited a slight performance decrease, scoring 97.67%, 95.35%, and 93.02% at 5%, 10%, and 20% missingness, respectively. Conversely, mode imputation initially delivered similar performance to AE. However, the performance gap widened significantly once the missing rate exceeded 30%. At a 40% missingness level, AE maintained an accuracy of 81.39%, whereas that of the mode decreased to 79.07%, further decreasing to 74.42% at 50%. These results indicate that AE consistently sustains above-average performance across varying missing rates and provides higher predictive reliability than statistical mean and mode imputations.

The ensemble model exhibited a more pronounced performance degradation with increasing missingness, which magnified the differences among the three supplementation methods. Notably, AE maintained a high accuracy of 93.02% at 5% missingness, gradually decreasing to 90.70%, 86.05%, and 76.74% at 10%, 20%, and 30% missingness, respectively. Mode imputation demonstrated similar or slightly lower performance in this range, with 88.37%, 90.70%, and 72.09% at 10%, 20%, and 30% missingness, respectively. However, its performance decreased sharply at higher rates, plummeting to 60.47% and 55.81% at 40% and 50% missingness, respectively. Under the same high-missingness conditions, AE maintained significantly higher accuracies of 72.09% and 62.79%, demonstrating its superior robustness and predictive reliability, even under such conditions.

In summary, the AE method only suffered slight performance degradation with increasing missingness levels across all three models, consistently maintaining the highest overall accuracy. Conversely, the statistical methods (mean and mode methods) suffered significant, rapid declines once the missing rate exceeded 20%, highlighting their limited applicability in high-missing environments. Although the mode method initially performed similarly to AE in some cases, its predictive accuracy decreased significantly at higher missing rates owing to its inability to preserve the underlying data structure.

Overall, these results unequivocally demonstrate that AE-based supplementation restores missing information far more effectively than simple statistical substitution. It ensures stable HI prediction performance, even under high-missingness conditions, providing strong empirical evidence for the practicality and applicability of AE-based supplementation in real-world TR-condition diagnosis systems.

These trends confirm that the proposed AE-based imputation method provides superior robustness against increasing missingness, particularly in scenarios where multiple critical diagnostic variables are simultaneously unavailable.

6.3. Representative Case Analysis Using Support Vector Machine (20% Missing Rate)

In the previous section, we quantitatively assessed the stability and prediction accuracy of AE-based supplementation techniques by comparing their performances across varying missing rates. In this section, we explore the 20% missing-rate experimental condition as a representative case to deeply analyze the impact of the AE method on actual HI-prediction

results. Particularly, we highlight how the AE approach replaces missing values as well as enhances condition-classification accuracy.

In this experiment, missing data were generated by randomly removing approximately 20% of key diagnostic items from actual TR diagnostic records. Next, three supplementation techniques (mean, mode, and the proposed AE method) were applied, and their results were compared by the same SVM-based HI-prediction model. SVM, a representative ML algorithm that is well known for its ability to determine decision boundaries in high-dimensional spaces, was employed as the benchmark model to quantify prediction accuracy.

Table 5 summarizes the actual conditions, missing items, AE-supplemented values, and predicted results (using SVM) for five representative TRs: TR1, TR2, TR8, TR33, and TR37. These cases share missing data in key DGA indicators (C_2H_2, CH_4, C_2H_4, C_2H_6, and TCG). The results indicate that the conventional statistical supplementation methods tend to underestimate actual conditions.

Table 5. Proposed Power-TR HI grade evaluation.

TR	Real	Missing Data	AE Imputation	Missing Parameter
1	4	3	4	TCG
2	5	2	3	C_2H_2
3	3	5	5	Age
4	4	3	3	C_2H_4
8	3	2	5	TCG
33	5	1	5	CH_4, C_2H_4, C_2H_6
37	5	2	5	TCG

This case-based analysis complements the aggregated accuracy results (Figure 23) by illustrating how imputation quality directly affects individual maintenance decisions at the transformer level.

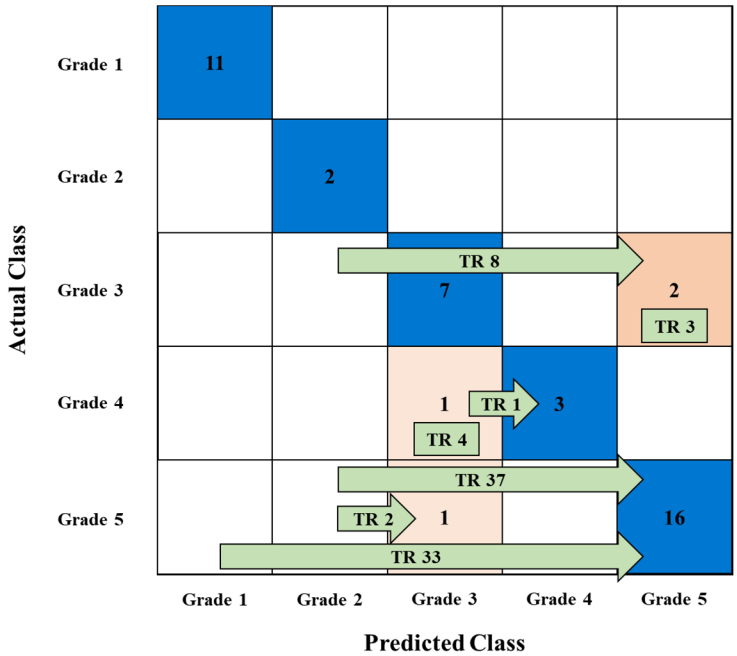


Figure 23. Proposed Method Using SVM with AE Imputation.

For instance, TR1 with an actual Grade 4 soundness level was underestimated as Grade 3 when the missing TCG data were supplemented via mean or mode imputation.

Conversely, the AE method correctly reclassified it as Grade 4 by restoring the missing TCG through learned correlations with related variables. Similarly, for TR33, where CH_4, C_2H_4, and C_2H_6 were simultaneously missing, the mean/mode methods resulted in a severe underestimation of the condition as Grade 1. However, AE accurately restored the state to Grade 5, which is consistent with the actual condition. These results demonstrate that AE transcends simple numerical substitution to provide structure-preserving restoration based on nonlinear inter-variable interactions.

These cases also highlight the criticality of missing DGA variables in HI-prediction models. Among these variables, C_2H_2, C_2H_4, CH_4, and C_2H_6 are strongly associated with internal discharge, insulation deterioration, and arcing and are consequently assigned high diagnostic weights. Thus, the model tends to underestimate and overestimate risk factors and health status, respectively, when these variables are missing, potentially delaying timely fault detection. However, the AE supplementation method resolves this issue by preserving variable correlations, thereby improving prediction sensitivity as well as enabling earlier detection of potential anomalies.

6.4. Verification of the Effectiveness of Missing Value Supplementation Methods Using Actual Failure Case

Here, we evaluated the effectiveness of various missing-value-compensation techniques using actual field failure histories. Our analysis revealed that the conventional imputation methods (mean and mode) tended to underestimate actual equipment condition in the absence of DGA items, a limitation that was effectively addressed by the proposed AE-based method, demonstrating that the AE approach transcends simple numerical substitution to enhance predictive performance by more accurately reflecting actual asset conditions.

In this subsection, the effectiveness of different missing value supplementation techniques is verified using representative transformer failure cases documented in internal inspection reports. These cases provide practical evidence of how missing diagnostic variables affect health index (HI) classification outcomes and how the proposed AE-based method mitigates such effects.

For TR1, internal inspection records indicated localized overheating and abnormal discharge phenomena in the on-load tap changer (OLTC) tap-selector terminals, which required a Grade 4 classification corresponding to major component replacement. However, when critical DGA indicators such as C2H2 and CH4 were unavailable, conventional mean and mode imputation methods substituted these values with global statistical estimates. As a result, the severity of the fault was underestimated and classified as Grade 3. In contrast, the proposed AE-based method reconstructed the missing gas indicators by exploiting correlations with related variables, leading to a Grade 4 classification that was consistent with the documented failure condition.

A similar trend was observed for TR2, where inspection records confirmed severe insulation degradation and arc discharge traces in the tap-lead region. The associated DGA data exhibited extremely high C2H2 concentrations, a well-known indicator of internal arcing. When this key variable was missing, statistical imputation diluted the fault signature and resulted in an underestimated HI. By contrast, the AE-based method restored the missing values based on learned multivariate relationships and produced a health grade closer to the actual critical condition.

In the case of TR8, inspection records reported severe contact wear in selector and diverter switches, which can induce thermal stress and insulation degradation during operation. Due to missing TCG data, conventional imputation methods underestimated the HI. The AE-based method, while generally improving classification accuracy, slightly overestimated the condition severity in this specific case. This observation indicates that

the proposed method, although more structure-preserving than statistical substitution, may still overestimate certain conditions when correlations are amplified. Nevertheless, the AE-based results remained significantly closer to the documented failure condition than those obtained using mean or mode imputation.

Overall, these case studies demonstrate that AE-based supplementation transcends simple numerical substitution by incorporating nonlinear inter-variable relationships, thereby improving the reliability of HI classification under missing-data conditions. While minor deviations may occur in individual cases, the proposed method consistently provides more realistic condition assessments aligned with actual failure histories

7. Conclusions

In this study, an autoencoder (AE)-based missing value supplementation technique was proposed to address the significant degradation in accuracy and reliability of health index (HI) assessment that occurs when key diagnostic variables, such as dissolved gas analysis (DGA), statistical indicators, and historical inspection information, are missing in power transformer condition evaluation.

Conventional mean and mode imputation methods rely on simple statistical substitution and therefore distort the intrinsic data structure, particularly when critical failure-related indicators are absent. As demonstrated in this study, such distortion often leads to the underestimation of equipment condition and delayed maintenance decisions. By contrast, the proposed AE-based approach supplements missing values by learning nonlinear inter-variable correlations in an unsupervised manner, enabling structure-preserving restoration rather than pointwise numerical substitution. As a result, the SVM-based HI classification accuracy improved significantly from 83.7% to 90.6% after applying AE-based supplementation.

Importantly, this study emphasizes that missing value compensation should not be treated as a mere preprocessing step, but as an integral component of asset-health assessment. Because transformer HI evaluation is fundamentally a grade-based decision problem, even small numerical deviations can cause a shift across health-grade boundaries and result in substantially different maintenance actions. Therefore, classification consistency and decision robustness were considered more appropriate evaluation criteria than conventional pointwise reconstruction metrics such as RMSE or MAE.

The practical effectiveness of the proposed method was further validated using representative real-world failure cases, including TR1, TR2, and TR8. In these cases, the AE-based method successfully restored missing critical DGA indicators and more accurately identified high-risk equipment compared with conventional statistical imputation techniques, demonstrating its applicability in realistic utility environments.

The main contributions of this study are threefold. First, the impact of missing diagnostic data on HI prediction performance was quantitatively analyzed using real inspection-based transformer datasets. Second, the superiority of the proposed AE-based supplementation method over conventional mean and mode imputation was systematically verified across multiple machine learning models. Third, the applicability of the proposed approach was validated through actual failure-case analysis, confirming its effectiveness in practical transformer condition assessment.

A limitation of this study is its dataset scale, as only 43 transformer inspection records were available. Future work will focus on validating the generalizability of the proposed approach using larger and more diverse datasets covering wider operating conditions and geographical regions. In addition, advanced generative restoration techniques, such as variational autoencoders, generative adversarial networks, and transformer-based models, may be explored to further enhance robustness under extreme missing data scenarios.

Overall, the proposed AE-based missing value supplementation technique enhances the reliability and robustness of transformer HI assessment under incomplete data conditions and provides practical support for asset management tasks, including maintenance planning, replacement decision-making, and risk-based operation of power facilities.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, and writing—original draft preparation, S.-Y.L.; writing—review and editing, J.-S.O. and J.-D.P.; resources and technical guidance, D.-H.L.; supervision, project administration, and writing—review and editing, T.-S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2025-25398164); and by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Climate, Energy, Environment (MCEE) of the Republic of Korea (No. RS-2025-07852969).

Data Availability Statement: The dataset used in this study is not publicly available due to institutional and project-related restrictions. As part of ongoing research, the reconstructed dataset cannot be shared. However, representative histograms illustrating the distribution characteristics of the dataset have been included in the manuscript. Data access requests may be sent to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AE	Autoencoder
CIGRE	International Council on Large Electric Systems
DGA	Dissolved Gas Analysis
DF	Dissipation Factor
DL	Deep Learning
DNV GL	Det Norske Veritas and Germanischer Lloyd
EPRI	Electric Power Research Institute
FMEA	Failure Mode and Effect Analysis
GRU	Gated Recurrent Unit
HI	Health Index
IEEE	Institute of Electrical and Electronics Engineers
IR	Insulation Resistance
kNN	k-Nearest Neighbors
LOOCV	Leave-One-Out Cross-Validation
LpOCV	Leave-p-Out Cross-Validation
LSTM	Long Short-Term Memory
MAR	Missing at Random
MCAR	Missing Completely at Random
ML	Machine Learning
MNAR	Missing Not at Random
OLTC	On-Load Tap Changer
PD	Partial Discharge
PI	Polarization Index
SFRA	Sweep Frequency Response Analysis
SVM	Support Vector Machine
TCG	Total Combustible Gas
TR	Power Transformer
UHF	Ultra-High Frequency

References

1. Ihendinihu, C.A.; Udofia, K.; Umana, T.I. The Impact of Transformer Failure on Electricity Distribution Network: A Case Study of Aba Area Network. *Int. Multiling. J. Sci. Technol.* **2023**, *8*, 6132–6141.
2. Bartley, W.H. Analysis of Transformer Failures, IMIA Working Group Paper WGP33-03. In Proceedings of the 36th Annual Conference of the International Association of Engineering Insurers (IMIA), Stockholm, Sweden, 15–17 September 2003.
3. Li, S.; Li, X.; Cui, Y.; Li, H. Review of Transformer Health Index from the Perspective of Survivability and Condition Assessment. *Electronics* **2023**, *12*, 2407. [CrossRef]
4. Al-Romaimi, K.; Baglee, D.; Dixon, D. Health Index Assessment for Power Transformer Strategic Asset Management in Electrical Utilities. *Int. J. Strateg. Eng. Asset Manag.* **2024**, *4*, 81–99. [CrossRef]
5. Jahromi, A.N.; Piercy, R.; Cress, S.; Service, J.R.R.; Wang, F. An Approach to Power Transformer Asset Management Using Health Index. *IEEE Electr. Insul. Mag.* **2009**, *25*, 20–34. [CrossRef]
6. Brandtzæg, G. Health Indexing of Norwegian Power Transformers. Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2015.
7. Taha, I.B.M. Power Transformers Health Index Enhancement Based on Convolutional Neural Network after Applying Imbalanced-Data Oversampling. *Electronics* **2023**, *12*, 2405. [CrossRef]
8. Zahra, S.T.; Imdad, S.K.; Khan, S.; Khalid, S.; Baig, N.A. Power Transformer Health Index and Life Span Assessment: A Comprehensive Review of Conventional and Machine Learning-Based Approaches. *Eng. Appl. Artif. Intell.* **2024**, *139*, 109474. [CrossRef]
9. Abdullah, A.M.; Ali, R.; Yaacob, S.B.; Ananda-Rao, K.; Uloom, N.A. Transformer Health Index by Prediction Artificial Neural Networks Diagnostic Techniques. *J. Phys. Conf. Ser.* **2022**, *2312*, 012002. [CrossRef]
10. Alqudsi, A.; El-Hag, A. Application of Machine Learning in Transformer Health Index Prediction. *Energies* **2019**, *12*, 2694. [CrossRef]
11. El-Rashidy, N.; Sultan, Y.A.; Ali, Z.H. Predicting Power Transformer Health Index and Life Expectation Based on Digital Twins and Multitask LSTM–GRU Model. *Sci. Rep.* **2025**, *15*, 1359. [CrossRef] [PubMed]
12. *IEEE Standard C57.140-2006*; IEEE Guide for the Evaluation and Reconditioning of Liquid Immersed Power Transformers. IEEE Power & Energy Society: New York, NY, USA, 2006.
13. Electric Power Research Institute (EPRI). *Equipment Failure Model and Data for Substation Transformer*; EPRI Report 1011989; EPRI: Palo Alto, CA, USA, 2005.
14. CIGRE Working Group A2.37. *Transformer Reliability Survey*; CIGRE Technical Brochure No. 642; CIGRE: Paris, France, 2015.
15. *IEEE Standard C57.152-2013*; IEEE Guide for Diagnostic Field Testing of Fluid-Filled Power Transformers, Regulators, and Reactors. IEEE Power & Energy Society: New York, NY, USA, 2013.
16. CIGRE Working Group A2.18. *Guide for Transformer Maintenance*; CIGRE Technical Brochure No. 445; CIGRE: Paris, France, 2010.
17. *IEEE Standard C57.104-2019*; IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers. IEEE Power & Energy Society: New York, NY, USA, 2019.
18. International Electrotechnical Commission (IEC). *IEC 60599: Mineral Oil-Impregnated Electrical Equipment in Service—Guide to the Interpretation of Dissolved and Free Gas Analysis*; IEC: Geneva, Switzerland, 2015.
19. Wang, L. (Ed.) *Support Vector Machines: Theory and Applications*; Springer: Boston, MA, USA, 2005.
20. Cervantes, J.; Garcia-Lamont, F.; Rodríguez-Mazahua, L.; Lopez, A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* **2020**, *408*, 189–215. [CrossRef]
21. Cover, T.M.; Hart, P.E. Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
22. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
23. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
24. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
25. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*; Wiley: Hoboken, NJ, USA, 2019.
26. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*; Rumelhart, D.E., McClelland, J.L., Eds.; MIT Press: Cambridge, MA, USA, 1986; pp. 318–362.
27. Baldi, P.; Hornik, K. Neural Networks and Principal Component Analysis: Learning from Examples without Local Minima. *Neural Netw.* **1989**, *2*, 53–58. [CrossRef]
28. Kramer, M.A. Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. *AIChE J.* **1991**, *37*, 233–243. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.