Preprints.org

Article

Enhancing Transformer Core Fault Diagnosis and Classification through Hilbert Transform Analysis of Electric Current Signals

Daryl Domingo , Akeem Bayo Kareem , Chibuzo Nwabufo Okwuosa , Paul Michael Custodio , Jang-Wook Hur

Article

Enhancing Transformer Core Fault Diagnosis and Classification through Hilbert Transform Analysis of Electric Current Signals

Daryl Domingo [1] ⓘ, **Akeem Bayo Kareem** [1] ⓘ, **Chibuzo Nwabufo Okwuosa** [1] ⓘ, **Paul Michael Custodio** [2] ⓘ, and **Jang-Wook Hur** [1,*] ⓘ

[1] Department of Mechanical Engineering (Department of Aeronautics, Mechanical and Electronic Convergence Engineering), Kumoh National Institute of Technology, South Korea

[2] ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea

* Correspondence: hhjw88@kumoh.ac.kr

Abstract: To enhance the overall reliability of the power system, engineers have redirected their focus toward health monitoring and early detection of faults in transformers. Among these faults, transformer core defects demand particular attention. While fault simulation using software has traditionally been the preferred approach, these methods suffer from data inaccuracies in real-world conditions. Consequently, conducting actual experimental setups with induced faults is imperative to investigate core issues. This study uses Hilbert Transform (HT) as a signal processing technique to extract crucial data characteristics, thereby enhancing the performance of the classifier model. The research involves analyzing electric current signals from a single-phase 1kVA transformer. A comparative assessment of our proposed model was conducted using raw data and Fast Fourier Transform (FFT), evaluating accuracy, precision, recall, F1-score, and computational time. The results demonstrate a significant improvement in all metrics for the classifier models, particularly the k-nearest neighbor (KNN) algorithm, which exhibited values of 83.89%, 84.39%, 83.89%, 83.79%, and 0.0156 seconds, respectively. Future work aims to extend this analysis under different loading conditions.

Keywords: core fault; electrical machines; fault classification; fault diagnosis; Hilbert transform; reliability; signal processing; transformer

1. Introduction

Predictive maintenance employs data-driven methodologies to anticipate potential failures in equipment or machinery, facilitating proactive maintenance measures and minimizing unforeseen downtime. It involves gathering data through strategically placed sensors in the equipment or by analyzing current and voltage levels for simplicity [1]. The ascendancy of data-driven techniques over model-based methods is attributed to the challenges inherent in developing and sustaining accurate physics-of-failure models. The significance of a data-driven approach in training artificial intelligence (AI) models must be considered in predictive maintenance. This approach is instrumental in harnessing the full potential of AI-based models, ensuring their efficacy in forecasting and preventing equipment failures [2].

The early stages of the core fault had minimal impact on the transformer. However, it can gradually lead to more severe damage. To obtain a high-quality energy supply, preventive evaluation for failure, mainly core fault, can effectively mitigate the risk of additional damage to the transformer, reducing outage duration and repair expenses [3]. Failure of the transformer is dangerous, especially to utility personnel, through explosion and fire. A highly challenging problem with transformers is the potential damage to their core due to increased exposure to the elements compared to other components. External conditions can activate numerous factors, namely short circuits, lightning strikes, and abrupt power surges. As a result, a process of identifying a fault is vital prior to the onset of malfunction [4].

Transformers play an essential role in power transmission and distribution networks. Any malfunction or fault in a transformer can significantly disrupt the reliable operation of the power system, resulting in considerable economic losses. Furthermore, given the high cost of transformers and the challenges associated with their maintenance, early fault detection is of paramount importance to facilitate timely repairs, ultimately reducing the risk of significant breakdowns [5]. The primary function of a transformer core is to transmit and convert electromagnetic energy. Any changes in the mechanical condition of the core represent a potential risk that can impact the safe and stable transformer operation. Statistical data reveals that core issues are the third most common problem among all transformer faults. Minor issues in the transformer's core might not immediately disrupt its operation and can go undetected. However, these issues can significantly reduce the transformer's operational lifespan and its ability to withstand faults in the future [6]. Consequently, it is imperative to regularly monitor the transformer's core to ensure its continued reliability. Usually, routine preventive maintenance has been undertaken concurrently with standard testing. However, the paradigm shift towards condition-based maintenance(CBM) has led to a decrease and, in some cases, the elimination of routine maintenance activities. This transformation is driven by the principle that maintenance interventions are executed only when the equipment's condition necessitates attention. Unlike the scheduled nature of routine maintenance, CBM relies on real-time data and predictive analytics to pinpoint optimal maintenance windows, thereby enhancing overall efficiency and reducing unnecessary operational disruptions [8,9].

In this study, we acquire electric current data from a 1kVA transformer in both its healthy state and faulty state. The distinctive behavior of current in these different states serves as a key indicator of fault patterns, particularly in advanced stages. Detecting faults in the raw data proves to be a challenging task, especially during the early stages of fault development. To preemptively address severe faults, we turn to signal processing as a pivotal technique in implementing condition monitoring. Signal processing excels in extracting valuable insights from raw data, enabling the identification and enhancement of features that might otherwise go unnoticed in unprocessed signals. The application of signal processing encompasses crucial functionalities such as data compression, noise reduction, and pattern recognition. These processes collectively contribute to improved decision-making and insights derived from complex datasets. Therefore, the significance of signal processing cannot be overstated in the context of modern data-driven applications. Utilizing various analyses, including time-domain, frequency-domain, and time-frequency analyses, allows for more comprehensive observation of a system's faults [10–14].

This research introduces a novel model for identifying core faults in transformers by leveraging electric current data to assess the health of the transformer core. These contributions collectively advance the comprehension of transformer health assessment, laying the foundation for more effective fault detection methodologies in power systems. The study offers the following significant contributions:

- We have designed a framework that acquires an electric current dataset for a comprehensive analysis of transformer health, particularly in identifying core faults.
- We employed time domain signal processing, specifically the Hilbert Transform, for extracting the magnitude envelope, which is an important process in enhancing the interpretation of signal analysis.
- We have established a comprehensive framework for robust feature engineering, focusing on the extraction of time domain statistical features and filter-based Pearson correlation feature selection. This process captures important features within the time-series data that are relevant before feeding into the classifier model.
- We have conducted a comparative analysis in terms of performance evaluation and determine the most effective model of existing machine learning models to validate the efficiency of the proposed framework.

The subsequent sections of the paper are organized as follows: Section 2 delves into a review of related works and provides insight into the motivation behind the study, while Section 3 explores

the theoretical background of the research. Section 4 presents the system model of the proposed fault diagnosis framework. The detailed experimental results of the study are discussed in Section 5, and the paper concludes in Section 6, where the findings are summarized, and future works are discussed.

2. Motivation and Review of Related Literature

The basic working principle of a transformer lies in the usage of alternating current. When an alternating current flows to a winding, it generates a magnetic field in the core thus inducing voltage in the secondary side. This is possible through the concept of electromagnetic induction. The core provides a path for the magnetic flux generated by the primary winding. It is commonly made of laminated thin sheets of electrical steel so that it is insulated from each other, which reduces eddy current losses, improving the transformer's efficiency. When a magnetic field changes within a conductor, circulating currents called eddy currents are induced within the core. These currents circulate in closed loops and result in the generation of heat due to the electrical resistance of the material [15,16]. As stated in [6], the core fault is one of the common failures in transformers, which is the main motivation of this study. Table 1 summarizes various core faults, including brief descriptions, causes, and effects studied in [7,19] namely, saturation and lamination fault, respectively. In this study, we focus on mechanical damage to the transformer core.

Table 1. Different types of transformer core fault.

Type	Description	Cause	Effect
Saturation	Occurs when the magnetic flux reaches its limit	Overloading, sudden changes in load, or system faults	Overheating, and distorted output waveform
Lamination Fault	Deterioration of insulation between laminations	Aging, excessive moisture, or manufacturing defects	potential short circuit and reduced insulation resistance
Mechanical Damage	Physical damage to the core, such as bending or cracking	Mechanical stress, operational stress, or overloading	Altered magnetic properties, increased core losses

Research on data-driven fault detection on transformer's health created a new perspective for researchers as a new approach. However, only a few focused on using ML algorithms to develop predictive models. In [17], the authors presented a method for analyzing core looseness fault by acquiring vibration signals and used different time-frequency analysis methods to compare which method works well. Their work did not focus on predictive maintenance. Another development in transformer fault analysis is the usage of the feature extraction method on vibration signals based on variational mode decomposition, and Hilbert Transform is applied to obtain the Hilbert Spectrum of the signal [18]. While in [19], detection and classification of lamination fault, namely, edge burr and lamination fault, have been analyzed. They extracted Average, Fundamental, Total Harmonic Distortion (THD), and Standard Deviation (STD) features from the collected current signal and fed them to SVM, KNN, and DT algorithms.

In [21], authors tackled the problem of obtaining data on vibration and acoustic, so they came up with a simulation method of transformer core fault based on multi-field coupling. They verify their simulation through physical experiments. However, they did not perform a technique to remove noise in the data since it is prone to collecting unwanted signals. In [22], they were able to use signal processing and machine learning techniques to develop a prognosis model of a transformer based on vibration signals. In [23], they introduce an online technique called the On-load Current Method (OLCM); this method utilizes vibration measurements to distinguish the condition of the transformer

core. The usage of current as a way to analyze faults in moving machines such as motors was also used by [10,24].

3. Theoretical Background

3.1. Fast Fourier Transform

Fast Fourier Transform (FFT) is another way of processing signals designed for the computation of the Discrete Fourier Transform (DFT). To comprehend the intricacies of DFT, it is essential to first delve into the concept of Fourier Transform (FT). FT analyzes a signal in the time domain, breaking it down into a representation that exposes its frequency components. It explains the extent to which each frequency contributes to the original signal. Furthermore, FT within a discrete time domain is referred to as DFT, and FFT is recognized as an algorithm specifically tailored for the rapid computation of a large number of DFTs. The FT of a function, denoted as f(t), is shown below [27,28]:

$$\mathcal{F}\{f(t)\} = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\, dt \tag{1}$$

The FFT employs complex exponentials or sinusoids of varying frequencies as its basis functions, effecting a transformation into a distinct set of basis functions. Originally devised as an enhancement to the conventional DFT, the FFT significantly diminishes computational complexity from (N^2) to $(N \log N)$, rendering it especially beneficial for the efficient processing of extensive datasets and real-time applications. Mathematically, the FFT can be succinctly expressed as [29]:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-\frac{2\pi i}{N}kn}, \quad k = 0, 1, 2, \ldots, N-1 \tag{2}$$

3.2. Hilbert Transform

The derivation of an analytic signal from a real-valued signal entails the utilization of the Hilbert transform. The resultant analytic signal finds widespread application in signal processing and communication systems, serving diverse purposes such as analyzing frequency content, extracting envelope information, and facilitating phase-sensitive operations [30,31]. The Hilbert transform of a real-valued signal $f(t)$ is given by:

$$H[f(t)](t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{t - \tau}\, d\tau \tag{3}$$

or in terms of the Cauchy principal value:

$$H[f(t)](t) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \left[\int_{-\infty}^{t-\epsilon} \frac{f(\tau)}{t - \tau}\, d\tau + \int_{t+\epsilon}^{\infty} \frac{f(\tau)}{t - \tau}\, d\tau \right] \tag{4}$$

The analytic signal $F_H(t)$, combining the original signal $f(t)$ and its HT, is given by:

$$F_H(t) = f(t) + j \cdot H[f(t)](t) \tag{5}$$

The properties of analytical signal include:

- Complex Representation: The analytic signal is complex, with both real and imaginary components. The actual component signifies the original signal, while the imaginary component represents the Hilbert transform of the signal.
- 90-Degree Phase Shift: The positive frequency shifts to a negative 90-degree angle while the negative frequency shifts to a positive 90-degree.The HT introduces a phase shift of 90 degrees between the original signal and its HT. This phase shift is crucial in applications such as demodulation and phase-sensitive analysis.

- Frequency Content: The analytic signal provides a representation of the original signal that separates positive and negative frequency components. This property is useful in analyzing the frequency content of a signal.
- Enveloping: The envelope of the original signal can be extracted from the magnitude of the analytic signal. The envelope represents the slowly varying magnitude of the signal and is useful in applications such as amplitude modulation.

3.3. Analysis using Time Domain and Frequency Domain

Analyzing signals is of paramount importance in various scientific and engineering applications, and understanding the time and frequency domains is fundamental to this analysis. In the time domain(TD), examining signals provides insights into their temporal behavior, allowing researchers and engineers to understand how the signal varies with respect to time. This is crucial for assessing the transient response of systems, studying the duration and shape of pulses, and investigating dynamic behaviors such as switching events. TD analysis is essential for evaluating the stability, response time, and overall performance of circuits and systems [32]. In the frequency domain(FD), analysis reveals the frequency components present in the signal. This is particularly valuable for characterizing harmonic content, identifying resonances, and assessing the spectral distribution of the current. FD analysis is indispensable in the design and optimization of power distribution systems, as well as in identifying and mitigating issues related to electromagnetic interference and power quality [33].

3.4. Overview of Selected Machine Learning Algorithms

The AdaBoost classifier (ABC) is a smart way to group things, often used for figuring out what category something belongs to. It combines many simple decision-makers to create a strong classifier. ABC pays more attention to things it got wrong before and keeps improving with each try, making it less chance of overfitting [34]. The roots of the K-Nearest Neighbors(KNN) algorithm can be traced back to the field of pattern recognition. The idea of using the nearest neighbors for classification purposes started gaining attention. The KNN algorithm is a fundamental and versatile supervised machine learning algorithm used for classification and regression tasks [35,36]. It classifies a new data point by evaluating the class labels of its 'k' closest neighbors. For classification, the algorithm calculates distances to all points in the training set, identifies the 'k' nearest neighbors, and assigns the most frequent class label among them to the new data point [37]. Logistic Regression (LR) is a straightforward method used a lot in picking between two options. It looks at the chance that something belongs to a certain group. LR is liked for being easy to understand and efficient, especially when the connection between features and the target group is simple [38]. The Multi-Layer Perceptron (MLP) is a kind of computer program designed for learning tasks. It has different layers, and during learning, it adjusts itself to get better at predicting. The non-convex nature of its loss function, along with the presence of local minima, means that initializing weights randomly may lead to fluctuations in validation accuracy [39]. Stochastic Gradient Descent (SGD) is a helpful tool for training learning models. It works by getting better after looking at each example, making it fast and useful for big sets of data. SGD is popular because it balances between speed and accuracy in machine learning [40]. The Support vector machine (SVM) is a strong method for grouping and predicting. It's good at finding the best way to separate groups, especially in complicated situations [41].

4. Proposed Methodology

This section outlines the framework for the core fault detection of a transformer, as shown in Figure 1. The overall stages are as follows: collecting data on the electric current in both healthy and faulty conditions of the transformer; employing signal processing techniques, specifically using Hilbert Transform; conducting time-domain statistical feature extraction to derive relevant features; employing Pearson correlation filter-based approach to pinpoint highly correlated features; feeding

the selected features for model training and testing; and finally, conducting performance evaluation to validate the effectiveness of the model.
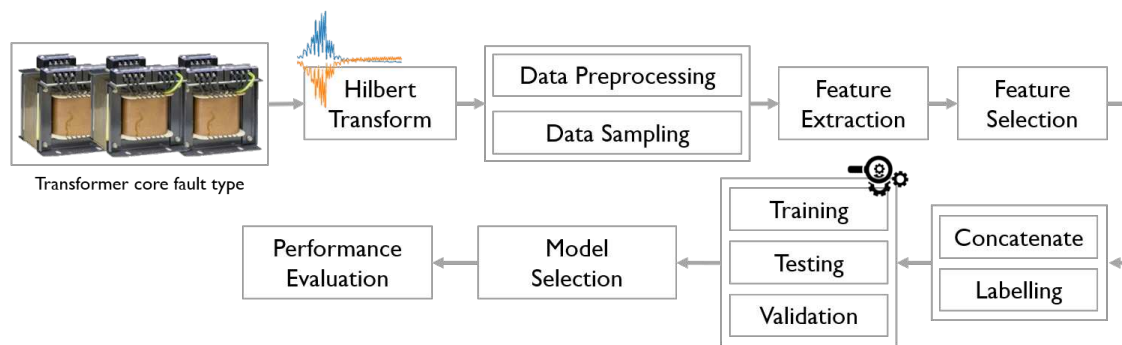

Figure 1. Proposed fault diagnosis framework.

4.1. Experimental Setup and Data Collection

Figure 2 depicts the experimental test setup utilized to acquire the current signals. This setup was designed and executed at the Defense and Reliability Laboratory, Kumoh National Institute of Technology, South Korea. The supply voltage is from a standard convenience outlet, featuring a rating of 220V at a frequency of 60Hz. Current measurements were conducted using the national instruments NI 9246 current module, seamlessly interfaced with LABVIEW software through the national instruments NI cDAQ-9174. The acquisition of current data occurred on the primary side of the circuit. Concurrently, on the secondary side, an electric fan was connected to function as a motor load for the transformer. The comprehensive circuit diagram is presented in Figure 3. The NI 9246 specifications are listed as follows:

- Three isolated analog input channels were employed, each operating at a simultaneous sample rate of 50 kS/s, ensuring comprehensive data collection.
- The system offers a broad input range of 22 continuous arms, with a ±30 Apeak input range and 24-bit resolution exclusively for AC signals.
- Specifically designed to accommodate 1 A/5 A nominal CTs, ensuring compatibility and accuracy during measurements.
- Channel-to-Earth isolation of up to 300 Vrms and Channel-to-Channel CAT III isolation of 480 Vrms guarantees safety and accuracy during experimentation.
- It has ring lug connectors tailored for up to 10 AWG cables, ensuring secure and reliable connections.
- It operates within a wide temperature range from -40 °C to 70 °C and is engineered to withstand 5 g vibrations and 50 g shocks, ensuring stability and functionality across varying environmental conditions.
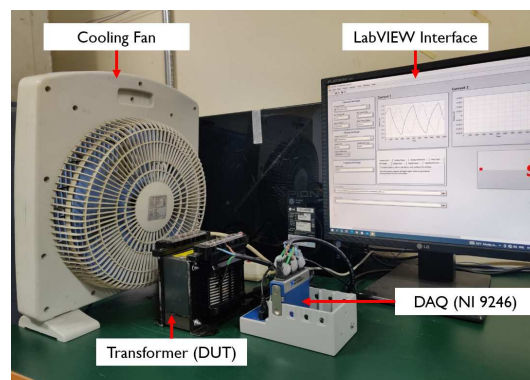

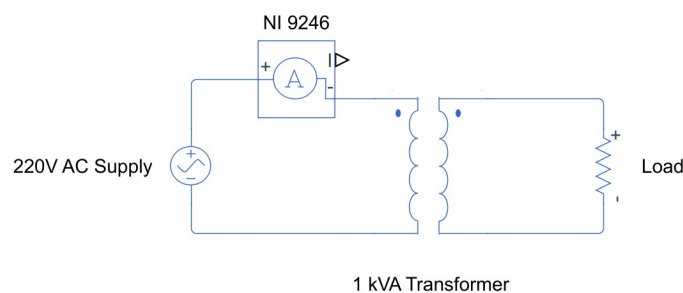**Figure 2.** Experimental Testbed Setup for Transformer Analysis.

Figure 3. Circuit Diagram

In this study, we obtained three datasets representing different conditions of transformers: a healthy state (labeled as HLTY), a state with one hole in the core (labeled as 1HCF), and a state with two holes in the core (labeled as 2HCF). To simulate 1HCF, a 5mm hole was drilled diagonally through the edge of the core. This was done to replicate damage focused on the edge of the transformer. In the 2HCF, an additional 5mm hole was drilled straight through the core from top to bottom, simulating core damage away from the edge of the transformer. Figure 4 illustrates the actual replication of these faults conducted during our experiment in the laboratory.

(**a**)   (**b**)

Figure 4. Actual faults induced: (**a**) 1HCF and (**b**) 2HCF.

4.2. Applying Signal Processing Technique

In this study, we employed signal processing techniques to unveil crucial details within the signals obscured in raw data. To assess the efficacy of our proposed model utilizing Hilbert Transform (HT) on electric current data, we conducted a comparative analysis using Fast Fourier Transform (FFT) without employing any signal processing technique. Following the signal processing step, we applied a window size of 25 samples to the data before proceeding with statistical feature extraction. Figure 5 a-c displays the electric current data obtained from the modules under three working conditions: HLTY, 1HCF, and 2HCF, represented by green, black, and orange. The data values range from -0.5 to 0.5 in all working conditions. During HLTY, the plot reveals that the maximum current in the circuit can reach -0.5A to 0.5A, with a notably cleaner waveform compared to other operating conditions. In the case of 1HCF, the current ranges from -0.4 to 0.4A, which is lower than the HLTY condition. The plot exhibits a random pattern with distortions in every cycle. Transitioning to 2HCF, the range of values is relatively similar to the 1HCF condition, varying from -0.4A to 0.4A. However, compared to HLTY, the waveform pattern differs. Upon close examination of each plot, it seems that using raw data could potentially aid in identifying core faults in transformers. However, upon closer examination in Figure 5 d, which presents the plots for all working conditions, it becomes evident that there is no significant difference when comparing HLTY with the faulty conditions (1HCF and 2HCF). Figure 6 illustrates the FFT plots under various operating conditions, revealing the limitation of FFT in capturing essential changes across all scenarios. Upon observation, the plots in all conditions exhibit minimal variation, indicating that when features are extracted, there is a lack of discriminative information. In

Section 4, we introduced our proposed method that leverages HT to enhance crucial characteristics of the current. Figure 7 demonstrates the substantial differences revealed after applying HT to the transformer core dataset, particularly distinguishing between healthy and faulty conditions. This observation underscores the effectiveness of our proposed signal processing technique's usefulness in analyzing transformers' core health based on current data. Identifying relevant characteristics and patterns in the raw signal proves to be pivotal in the initial stages of our methodology, as these factors significantly impact the overall performance of the ML model.
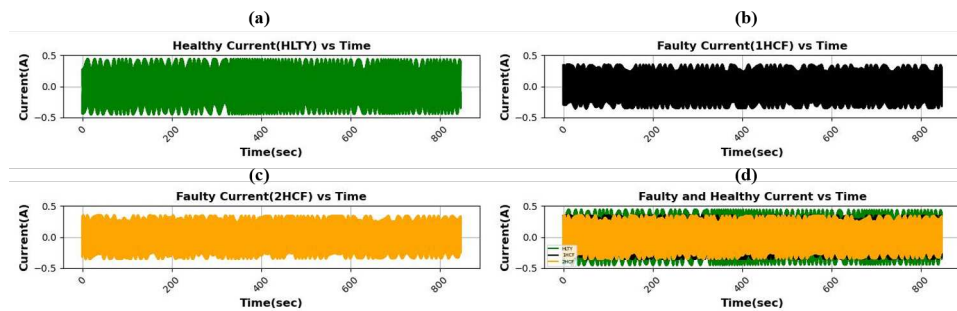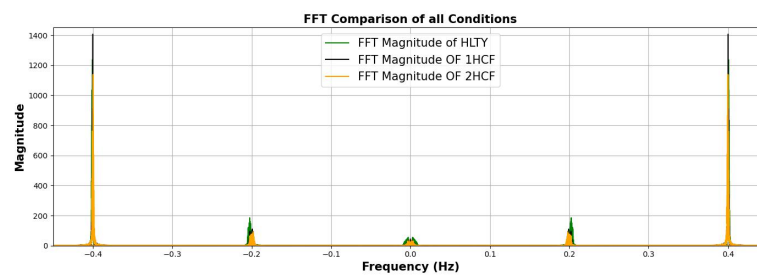

Figure 5. Plot of raw current signal.
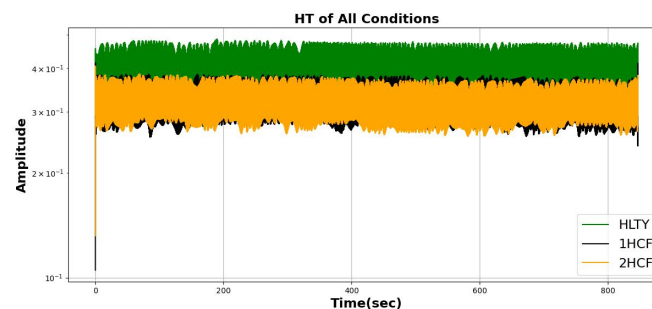

Figure 6. FFT of all working conditions.

Figure 7. HT of all working conditions.

4.3. Statistical Feature Extraction and Selection

It is imperative to represent data in a simplified manner, emphasizing only essential features before inputting it into the model to enhance the speed and accuracy of an ML model. Time-domain statistical features are extracted to capture the relevant aspects of the data. The primary objective of this process is dimensionality reduction while retaining crucial properties or features. Moreover, transforming the raw data into a more concise representation yields several advantages, including improved overall model performance by reducing complexity, decreasing computational time, and mitigating the risk of overfitting. In this paper, we employ feature engineering through statistical feature extraction, and the 14 extracted features in time-based and frequency-based domains are

illustrated in Table 2. The linear relationship between two variables can be statistically measured on a scale ranging from -1 to 1. If the value is close to 1, they are highly correlated, and the sign indicates a positive or negative correlation. Negative values indicate a negative correlation, signifying that as one variable increases, the other decreases. Conversely, positive values suggest a positive correlation, meaning that both variables increase or decrease simultaneously. A value of 0 denotes no correlation between the variables. This approach is widely employed in data analysis, statistics, and machine learning for feature selection and comprehending the relationships between variables [25,26]. The formula is presented below:

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} \tag{6}$$

The numerator represents the covariance of variables X and Y. X and Y are the features extracted from the HT values to simplify data representation, highlighting only the essential features. Covariance quantifies the extent to which the two variables change. The product of the standard deviations of X and Y is calculated in the denominator. Standard deviation gauges the degree of variation or dispersion from their respective averages. This study selected features exhibiting a Pearson correlation coefficient exceeding 95% to ensure distinctiveness among the remaining features. Subsequently, the retained features undergo concatenation and labeling before being employed in the model training and testing phases.

4.4. Correlation Matrix of Extracted and Selected Time Domain Statistical Feature

Figure 8 (a) illustrates the correlation plot of the features extracted from the HT. The red intensity in the plot indicates the strength of the correlation among features, with a gradient from more red showing a stronger correlation to less red and eventually blue indicating a weaker correlation. This matrix visually represents the relationships between each feature, providing valuable insights for analysis and utilization. A notable correlation between the mean and other features, namely max, peak-to-peak, RMS, variance, and standard deviation, with correlation coefficients of 0.98, 0.94, 1, 0.97, and 0.97, respectively. Recognizing such high correlations is crucial, as incorporating highly correlated features into the model can significantly and negatively impact its performance.
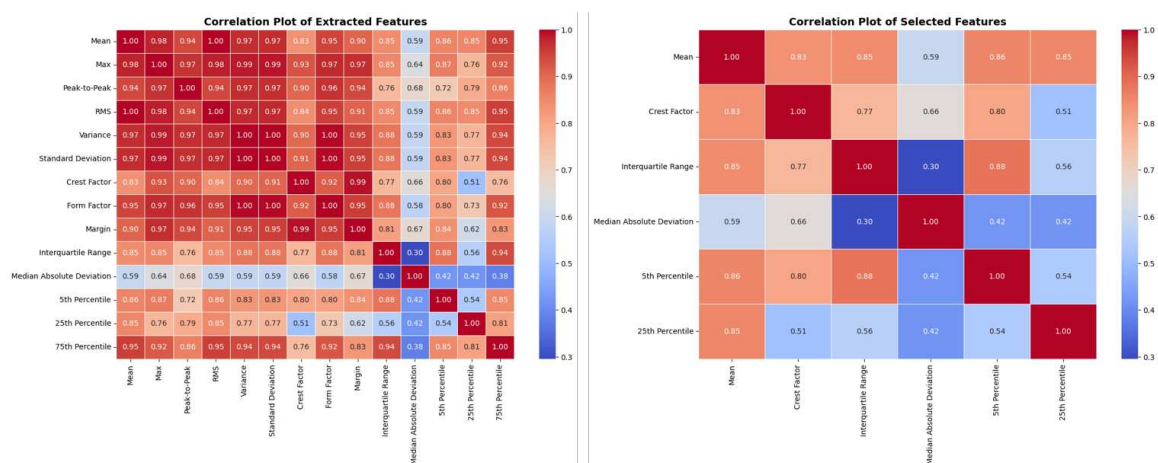

Figure 8. Statistical Correlation matrix (**a**) Feature extraction and (**b**) Feature Selection.

Upon extracting features and generating the correlation matrix, it becomes evident that the features are highly correlated and could impact the model's performance. To address this, we employ filter-based statistical feature selection. As illustrated in Figure 8 (b), out of the initially extracted 14 features, only six were retained after eliminating those with high correlations. This step further refined the dataset before feeding it into the ML model, enhancing its ability to capture relevant patterns and relationships in the data.

Table 2. Time domain statistical features and formulas.

Domain	Features	Formulas
Time-based	Crest Factor	$\frac{\text{Peak Value}}{\text{RMS Value}}$
	Form Factor	$\frac{\text{RMS Value}}{\text{Mean Value}}$
	Interquartile Range	$Q_3 - Q_1$
	Margin	$\frac{\text{Peak Value}}{\text{Mean Value}}$
	Max	$\frac{\text{Peak Value}}{\text{Mean Value}}$
	Mean	$\frac{\sum_{i=1}^{N} x_i}{N}$
	Median Absolute Deviation	$median(|x_i - median(x)|)$
	nth Percentile(5, 25, 75)	$100 \frac{x - 0.5}{n}$
	Peak-to-peak	$x_{max} - x_{min}$
	Root Mean Square (RMS)	$\sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$
	Standard Deviation	$\sqrt{\frac{\sum_{i=1}^{N} (x_i - \text{Mean})^2}{N}}$
	Standard Error Mean	$\frac{standard deviation}{\sqrt{n}}$
	Variance	$\frac{\sum_{i=1}^{N} (x_i - \text{Mean})^2}{N}$
Frequency-based	Mean Frequency	$\frac{\sum_{k=1}^{K} f_k \cdot P(f_k)}{\sum_{k=1}^{K} P(f_k)}$
	Median Frequency	$(P_{\text{cum}} > 0.5)$ where P_{cum} is the cumulative power spectral density
	Spectral Entropy	$-\sum_{k=1}^{K} P(f_k) \cdot \log_2(P(f_k))$
	Spectral Centroid	$\frac{\sum_{k=1}^{K} f_k \cdot P(f_k)}{\sum_{k=1}^{K} P(f_k)}$
	Spectral Spread	$\sqrt{\frac{\sum_{k=1}^{K} (f_k - \text{SC})^2 \cdot P(f_k)}{\sum_{k=1}^{K} P(f_k)}}$
	Spectral Skewness	$\frac{\sum_{k=1}^{K} \left(\frac{f_k - \text{SC}}{\text{SS}}\right)^3 \cdot P(f_k)}{K}$
	Spectral Kurtosis	$\frac{\sum_{k=1}^{K} \left(\frac{f_k - \text{SC}}{\text{SS}}\right)^4 \cdot P(f_k)}{K}$
	Total Power	$\sum_{k=1}^{K} P(f_k)$
	Spectral Flatness	$\frac{\exp\left(\frac{1}{K} \sum_{k=1}^{K} \log(P(f_k))\right)}{\frac{1}{K} \sum_{k=1}^{K} P(f_k)}$
	Peak Frequency & Frequency	$(P(f_k) = \max(P(f_k)))$
	Peak Amplitude	$\max(P(f_k))$
	Dominant Frequency & Frequency	$(P(f_k) = \max(P(f_k)))$
	Spectral Roll-Off (n = 80%, 90%)	$\text{Frequency}(P_{\text{cum}} > n \cdot \text{Total Power})$ where P_{cum} is the cumulative power spectral density.

4.5. Model Training and Testing

The dataset must be split to train and test our model effectively. In our study, 80% of the dataset is allocated for training the model, while 20% is for testing. We employ six established machine learning models: ABC, KNN, LR, MLP, SGD, and SVC, as discussed in Section 3. The summary of parameters

of different models is summarized in Table 3. The objective is to assess and compare the performance of these models in accurately classifying conditions such as HLTY, 1HCF, and 2HCF.

Table 3. Machine learning model and parameter value.

ML Model	Parameter	Value
AdaBoost(ABC)	n estimators	50
k-Nearest Neighbor(KNN)	k	9
Logistic Regression(LR)	regularization	L2
Multi-layer Perceptron(MLP)	learning rate, n layers	constant, 100
Stochastic Gradient Descent(SGD)	loss function	perceptron
Support Vector Machine(SVC)	C, gamma	90, scale

4.6. ML-Model Metrics

Our study focuses on detecting and classifying faults in transformers; the significance lies in thoroughly evaluating the performance of our proposed model. Ensuring the effectiveness of our method through evaluation not only validates its reliability but shows its capability to proactively identify and classify faults, ultimately contributing to the overall reliability and efficiency of the transformer. The metrics used, along with the formula and brief descriptions, are presented below:

Accuracy: Measures the overall correctness of the model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{7}$$

Precision: Indicates the accuracy of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

Recall: Emphasizes the model's ability to capture all positive instances.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

F1 Score: Provides a harmonic mean by balancing precision and recall. It is particularly valuable in scenarios with uneven class distribution.

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

5. Results and Discussion

This section presents a comprehensive breakdown of the experimental results and a detailed discussion of our dataset description, conducted using the framework introduced in the preceding section.

5.1. Model Performance on Raw Transformer Core Dataset

Table 4 presents the metrics, providing a comprehensive evaluation of the ML model. LR emerges as the top-performing model with an accuracy of 65.23%, showcasing the highest metrics across all aspects except for computational time, where KNN proves to be the fastest with 0.0142s. The lowest-performing model is SGD, with an accuracy of 29.08%. A plot of performance metrics alongside corresponding ML models is presented in Figure 9 for a more intuitive representation. In Figure 10, the confusion matrix depicts the performance of different machine learning models without employing any signal processing techniques. Notably, numerous false positive (FP) instances exist, with SVC

exhibiting 164. On the other hand, the highest true positives (TP) were achieved by ABC and LR, as evident from the values along the diagonal. In general, the ML model's performance appears unsatisfactory when analyzing raw data, rendering it ineffective for detecting and classifying core faults.

Table 4. Performance evaluation for raw data.

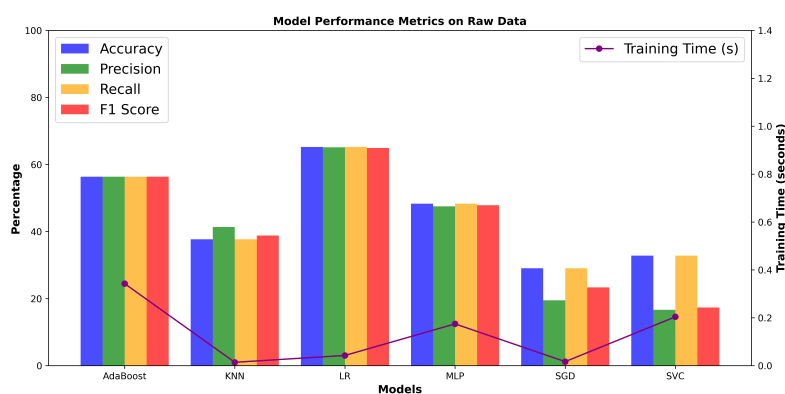| ML Model	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Time-cost(sec)
ABC	59.39	56.38	56.39	56.37	0.3425
KNN	37.72	41.35	37.72	38.84	0.1421
LR	**65.23**	**65.13**	**65.23**	**64.94**	0.0427
MLP	48.33	47.53	48.33	47.85	0.1745
SGD	29.08	19.51	29.08	23.34	0.0170
SVC	32.81	16.68	32.81	17.36	0.2040

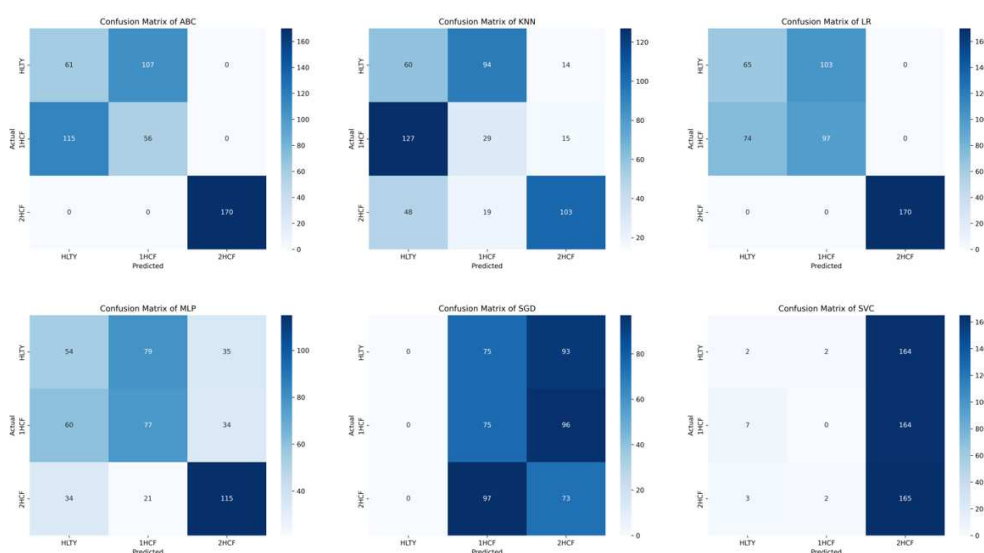**Figure 9.** Plot of ML model's performance evaluation when using raw data.

Figure 10. Confusion matrix of ML model when using raw data.

5.2. Model Performance using FFT on Transformer Core Dataset

To offer a comparative analysis, the incorporation of frequency domain analysis, mainly using FFT, provides a comprehensive and insightful comparative perspective. To thoroughly evaluate the model performance, metrics are tabulated in Table 5. ABC emerges as the leading model with an

accuracy of 61.49%, exhibiting the highest metrics across all aspects, and KNN delivers the fastest computational time. At the same time, SVC performed worse with an accuracy of 33.20%. A plot illustrating performance metrics alongside the corresponding machine learning models is presented in Figure 11 for enhanced visualization. Numerous FP instances are observed in SVC, with 163 cases, as shown in Figure 12. Conversely, the highest TP is recorded by ABC, which is evident from the values along the diagonal. Overall, the ML model exhibits low performance when conducting analyses in the frequency domain, thereby limiting its efficacy in detecting and classifying core faults.

Table 5. Performance evaluation for FFT.

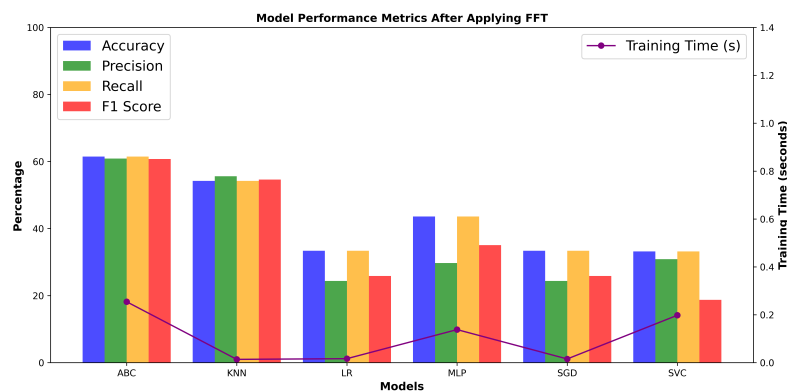| ML Model	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Time-cost(sec)
ABC	**61.49**	**60.88**	**61.49**	**60.74**	0.2545
KNN	54.22	55.63	54.22	54.63	0.0138
LR	33.40	24.37	33.40	25.88	0.0168
MLP	43.61	29.73	43.61	35.06	0.1381
SGD	33.40	24.37	33.40	25.88	0.0156
SVC	33.20	30.89	33.20	18.74	0.1980

Figure 11. Plot of ML model's performance evaluation when using FFT.
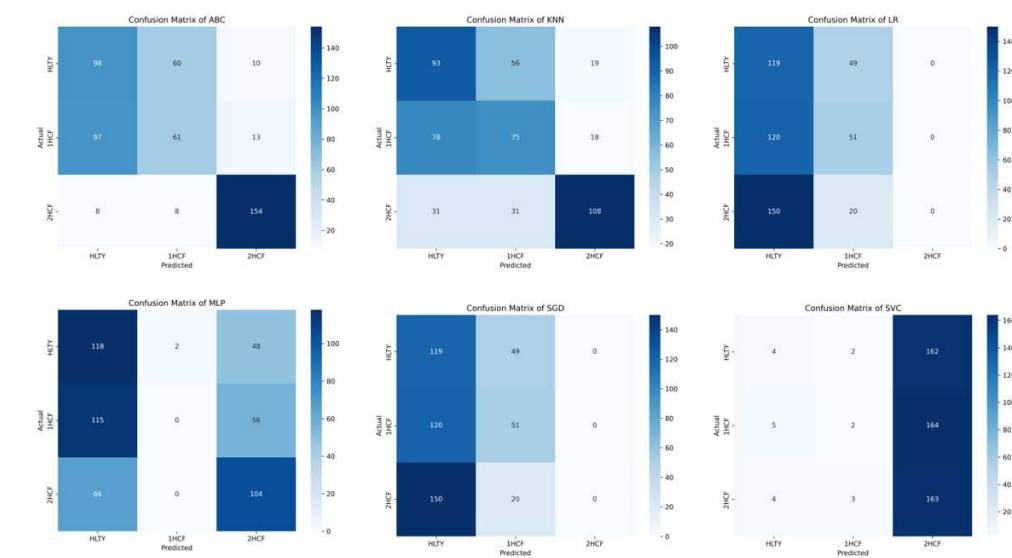

Figure 12. Confusion matrix of ML model when using Fast Fourier Transform.

5.3. Model Performance using HT on Transformer Core Dataset

Key performance metrics are presented in Table 6 to evaluate the ML models' performance comprehensively. Among these metrics, KNN stands out as the superior model with an accuracy of 83.89% and a computational time of 0.0156s, boasting the highest values across all parameters and delivering the fastest computational performance. Figure 13 is a graphical plot illustrating the performance metrics alongside the corresponding ML models. Figure 14 depicts the confusion matrix for all ML models during the execution of HT on raw signals. Notably, the ML models exhibited a consistent improvement in TP and a reduction in FP. Notably, the ABC and SGD models displayed the highest number of FP instances, erroneously predicting 121 and 171 cases, respectively. Conversely, other models demonstrated an increase in TP predictions, evident from the values along the diagonal. In essence, our model shows a notable enhancement in performance when employing the Hilbert Transform for analyses, particularly in diagnosing and classifying core faults.

Table 6. Performance evaluation for HT.

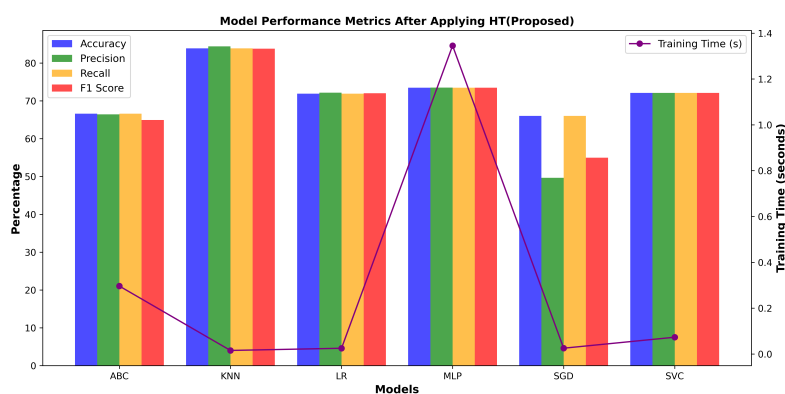| ML Model	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Time-cost(sec)
ABC	66.60	66.42	66.60	64.95	0.2963
KNN	**83.89**	**84.39**	**83.89**	**83.79**	0.0156
LR	71.91	72.16	71.91	72.01	0.0251
MLP	73.45	73.51	73.48	73.47	1.3454
SGD	66.01	46.66	66.01	54.99	0.0253
SVC	72.10	72.11	72.10	72.10	0.0732

Figure 13. Plot of ML models performance evaluation when using HT.
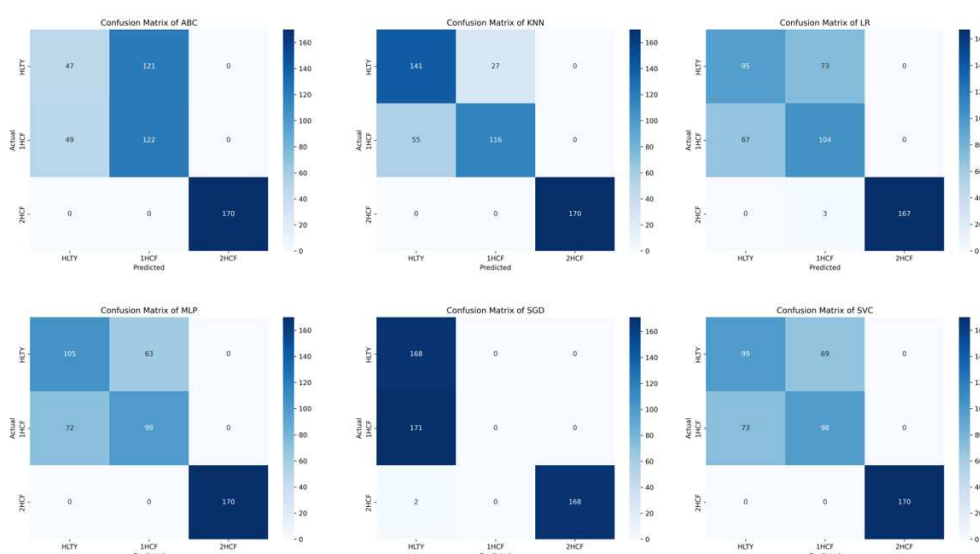
Figure 14. Confusion matrix of ML models when using Hilbert Transform.

6. Conclusion

This paper introduces HT as a signal processing technique and assesses the performance of various ML models in classifying the condition of a transformer's core. The evaluation is based on our electric current data collected during healthy and faulty conditions. The proposed method is compared under two scenarios: without any signal processing technique and when applying FFT. Our method demonstrates a notable enhancement in the overall performance of six well-established ML models, as evidenced by improvements in their performance metrics. Furthermore, the acquired current data serves as a baseline for future research focused on monitoring the transformer's core. Future work will involve comparing the proposed method under different load conditions and using it to monitor the health of the transformer winding.

Author Contributions: Conceptualization, D.D. and P.M.C; methodology, D.D. and C.N.O; software, D.D. and J.-W.H.; validation, D.D.; formal analysis, D.D.; investigation, D.D.; resources, A.B.K., C.N.O., and J.-W.H.; data curation, D.D.; writing—original draft preparation, D.D. and P.M.C.; writing—review and editing, D.D., P.M.C, C.N.O and A.B.K; visualization, D.D. and A.B.K.; supervision, D.D. and A.B.K.; project administration, A.B.K., C.N.O., and J.-W.H.; funding acquisition, J.-W.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study can be obtained upon request from the corresponding author. However, it is not accessible to the public as it is subject to laboratory regulations.

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

References

1. Kim, N.H.; An, D.; Choi, J.H. *Prognostics and Health Management of Engineering Systems: An Introduction*; Springer: Cham, Switzerland, **2017**; pp. 127–241. [CrossRef]
2. O. Fink, Q. Wang, M. Svensén, P. Dersin, W.-J. Lee, and M. Ducoffe. *Potential, challenges and future directions for deep learning in prognostics and health management applications*; Eng. Appl of Artif Intell, **2020**; vol. 92, Art. no. 103678. [CrossRef]

3.   L. Zhuang, B. K. Johnson, X. Chen and E. William, A topology-based model for two-winding, shell-type, single-phase transformer inter-turn faults. *2016 IEEE/PES Trans. and Dist. Conference and Exposition (T&D)* **2016**, pp.1-5. [CrossRef]

4.   S. S. Manohar, A. Subramaniam, M. Bagheri, S. Nadarajan, A. k. Gupta and S. K. Panda, Transformer Winding Fault Diagnosis by Vibration Monitoring. *2018 Condition Monitoring and Diagnosis (CMD)* **2018**, pp.1-6. [CrossRef]

5.   Islam, M.M., Lee, G. & Hettiwatte, S.N. *A nearest neighbour clustering approach for incipient fault diagnosis of power transformers*; S Electr Eng 99, 2017; vol. 99, no. 3, pp. 1109-1119. [CrossRef]

6.   M. Wang, A. J. Vandermaar and K. D. Srivastava, Review of condition assessment of power transformers in service. *IEEE Electrical Insulation Magazine* **2002**, vol. 18, no. 6, pp. 12-25. [CrossRef]

7.   Fritsch M.; Wolter M. Saturation of High-Frequency Current Transformers: Challenges and Solutions. *IEEE Transactions on Instrumentation and Measurement* **2023** vol. 72, pp. 1-10, 2023, Art no. 9004110. [CrossRef]

8.   Zhikai Xing, Yigang He, Multi-modal information analysis for fault diagnosis with time-series data from power transformer. *International Journal of Electrical Power & Energy Systems* **2023**, vol. 144, 108567, ISSN 0142-0615. [CrossRef]

9.   Hu, H.; Ma, X.; Shang, Y., A Novel Method for Transformer Fault Diagnosis Based on Refined Deep Residual Shrinkage Network. *IET Electr. Power Appl.* **2021**, 16(2), 206-223. [CrossRef]

10. Okwuosa, C.N., Hur, J.W. A Filter-Based Feature-Engineering-Assisted SVC Fault Classification for SCIM at Minor-Load Conditions. *Energies* **2022**, 15, 7597. [CrossRef]

11. Kareem, A.B., Hur, J.-W. Towards Data-Driven Fault Diagnostics Framework for SMPS-AEC Using Supervised Learning Algorithms. *Electronics* **2022**, 11, 2492. [CrossRef]

12. Shifat, T.A., Hur, J.W., ANN Assisted Multi Sensor Information Fusion for BLDC Motor Fault Diagnosis. *IEEE Access* **2021**, 9, 9429–9441. [CrossRef]

13. Lee, J.-H.; Okwuosa, C.N.; Hur, J.-W. Extruder Machine Gear Fault Detection Using Autoencoder LSTM via Sensor Fusion Approach. *Inventions* **2023**, 8, 140. [CrossRef]

14. Park, S.-H.; Kareem, A.B.; Joo, W.J.; Hur, J.-W. FEA Assessment of Contact Pressure and Von Mises Stress in Gasket Material Suitability for PEMFCs in Electric Vehicles. *Inventions* **2023**, 8, 116. [CrossRef]

15. Gao, B.; Yu, R.; Hu, G.; Liu, C.; Zhuang, X.; Zhou, P. Development Processes of Surface Trucking and Partial Discharge of Pressboards Immersed in Mineral Oil: Effect of Tip Curvatures. *Energies* **2019**, 12, 554. [CrossRef]

16. Liu, J.; Cao, Z.; Fan, X.; Zhang, H.; Geng, C.; Zhang, Y. Influence of Oil–Pressboard Mass Ratio on the Equilibrium Characteristics of Furfural under Oil Replacement Conditions. *Polymers* **2020**, , 12, 2760. [CrossRef]

17. Yuan, F.; Shang, Y.; Yang, D.; Gao, J.; Han, Y.; Wu, J. Comparison on multiple signal analysis method in transformer core looseness fault. In Proceedings of the IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, Dalian, China, 14–16 April 2021. [CrossRef]

18. Haoyang T.; Wei P. Min H.; Guogang Y.; Yuhui C. Feature extraction of the transformer core loosening based on variational mode decomposition. in: 2017 1st International Conference on Electrical Materials and Power Equipment (ICEMPE), Xi'an, China, 14–17 May 2017. [CrossRef]

19. Altayef E.; Anayi F.; Packianather M.; Benmahamed Y.; Kherif O. Detection and Classification of Lamination Faults in a 15 kVA Three-Phase Transformer Core Using SVM, KNN and DT Algorithms. *IEEE Access* **2022**, vol. 10, pp. 50925-50932. [CrossRef]

20. Pleite, J.; Gonzalez, C.; Vazquez, J.; Lázaro A. Feature extraction of the transformer core loosening based on variational mode decomposition. In Proceedings of the IEEE Mediterranean Electrotechnical Conference, Malaga, Spain, 16-19 May 2006. [CrossRef]

21. Yao, D.; Li, L.; Zhang, S.; Zhang, D.; Chen, D. The Vibroacoustic Characteristics Analysis of Transformer Core Faults Based on Multi-Physical Field Coupling. *Symmetry* **2022**, 14, 544. [CrossRef]

22. Bagheri M.; Zollanvari A.; Nezhivenko S. Transformer Fault Condition Prognosis Using Vibration Signals Over Cloud Environment. *IEEE Access* **2018**, vol. 6, pp. 9862-9874. [CrossRef]

23. Shengchang J.; Yongfen L.; Yanming L. Research on extraction technique of transformer core fundamental frequency vibration based on OLCM. *IEEE Transactions on Power Delivery* **2006**, vol. 21, no. 4, pp. 1981-1988. [CrossRef]

24. Okwuosa, C.N., Hur, J.W. An Intelligent Hybrid Feature Selection Approach for SCIM Inter-Turn Fault Classification at Minor Load Conditions Using Supervised Learning. *IEEE Access* **2023**, vol. 11, pp. 89907-89920. [CrossRef]

25. Metsämuuronen, J. Artificial systematic attenuation in eta squared and some related consequences: Attenuation-corrected eta and eta squared, negative values of eta, and their relation to Pearson correlation. *Behaviormetrika* **2023**, 50, 27–61. [CrossRef]

26. Denuit, M.; Trufin, J. Model selection with Pearson's correlation, concentration and Lorenz curves under autocalibration. *Eur. Actuar. J.* **2023**, 13, 1–8. [CrossRef]

27. Riza Alvy Syafi'i M.H.;Prasetyono E.; Khafidli M.K.;Anggriawan D.O.; Tjahjono A. Real Time Series DC Arc Fault Detection Based on Fast Fourier Transform. 2018 International Electronics Symposium on Engineering Technology and Applications (IES-ETA), Bali, Indonesia, 29-30 October 2018, pp. 25-30. [CrossRef]

28. Misra, S.; Kumar, S.; Sayyad, S.; Bongale, A.; Jadhav, P.; Kotecha, K.; Abraham, A.; Gabralla, L.A. Fault Detection in Induction Motor Using Time Domain and Spectral Imaging-Based Transfer Learning Approach on Vibration Data. *Sensors* **2022**, 22, 8210. [CrossRef]

29. Ewert, P.; Kowalski, C.T.; Jaworski, M. Comparison of the Effectiveness of Selected Vibration Signal Analysis Methods in the Rotor Unbalance Detection of PMSM Drive System *Electronics* **2022**, 11, 1748. [CrossRef]

30. El Idrissi, A.; Derouich, A.; Mahfoud, S.; El Ouanjli, N.; Chantoufi, A.; Al-Sumaiti, A.S.; Mossa, M.A. Bearing fault diagnosis for an induction motor controlled by an artificial neural network—Direct torque control using the Hilbert transform. *Mathematics* **2022**, 10, 4258. [CrossRef]

31. Dias, C.G.; Silva, L.C. Induction Motor Speed Estimation based on Airgap flux measurement using Hilbert transform and fast fourier transform *IEEE Sens. J.* **2022**, 22, 12690–12699. [CrossRef]

32. Badihi, H.; Zhang, Y.; Jiang, B.; Pillay, P.; Rakheja, S. A Comprehensive Review on Signal-Based and Model-Based Condition Monitoring of Wind Turbines: Fault Diagnosis and Lifetime Prognosis. *Proc. IEEE* **2022**, 110, 754–806. [CrossRef]

33. Ismail, A.; Saidi, L.; Sayadi, M.; Benbouzid, M. A New Data-Driven Approach for Power IGBT Remaining Useful Life Estimation Based On Feature Reduction Technique and Neural Network. *Electronics* **2018**, 9, 1571. [CrossRef]

34. Stavropoulos, G.; van Vorstenbosch, R.; van Schooten, F.; Smolinska, A.Random Forest and Ensemble Methods. *Chemom. Chem. Biochem. Data Anal.* **2020**, 2, 661–672. [CrossRef]

35. Yang, J.; Sun, Z.; Chen, Y. Fault Detection Using the Clustering-kNN Rule for Gas Sensor Arrays. *Sensors* **2016**, 16, 2069. [CrossRef]

36. Wang, X.; Jiang, Z.; Yu, D. An Improved KNN Algorithm Based on Kernel Methods and Attribute Reduction. In Proceedings of the 5th International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), Qinhuangdao, China, 18–20 September 2015; pp. 567–570.. [CrossRef]

37. Saadatfar, H.; Khosravi, S.; Joloudari, J.H.; Mosavi, A.; Shamshirband, S. A New K-Nearest Neighbors Classifier for Big Data Based on Efficient Data Pruning. *Mathematics* **2020**, 8, 286. [CrossRef]

38. Couronné, R.; Probst, P.; Boulesteix, A.-L. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinform.* **2018**, 19, 270. [CrossRef]

39. Carreras, J.; Kikuti, Y.Y.; Miyaoka, M.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; Nakamura, N.; Hamoudi, R. A Combination of Multilayer Perceptron, Radial Basis Function Artificial Neural Networks and Machine Learning Image Segmentation for the Dimension Reduction and the Prognosis Assessment of Diffuse Large B-Cell Lymphoma. *AI* **2021**, 2, 106–134. [CrossRef]

40. Huang, J.; Ling, S.; Wu, X.; Deng, R.GIS-Based Comparative Study of the Bayesian Network, Decision Table, Radial Basis Function Network and Stochastic Gradient Descent for the Spatial Prediction of Landslide Susceptibility. *Land* **2022**, 11, 436. [CrossRef]

41. Han, T.; Jiang, D.; Zhao, Q.;Wang, L.; Yin, K. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Trans. Inst. Meas. Control* **2018**, 40, 2681–2693. [CrossRef]