

University of Windsor

## Scholarship at UWindor

---

Electrical and Computer Engineering  
Publications

Department of Electrical and Computer  
Engineering

---

2024

# Explainable Artificial Intelligence Approach for Diagnosing Faults in an Induction Furnace

Sajad Moosavi  
*University of Windsor*

Roozbeh Razavi-Far  
*University of New Brunswick*

Vasile Palade  
*Coventry University*

Mehrdad Saif  
*University of Windsor*

Follow this and additional works at: <https://scholar.uwindsor.ca/electricalengpub>



Part of the [Electrical and Computer Engineering Commons](#)

---

### Recommended Citation

Moosavi, Sajad; Razavi-Far, Roozbeh; Palade, Vasile; and Saif, Mehrdad. (2024). Explainable Artificial Intelligence Approach for Diagnosing Faults in an Induction Furnace. *Electronics*.  
<https://scholar.uwindsor.ca/electricalengpub/481>

This Article is brought to you for free and open access by the Department of Electrical and Computer Engineering at Scholarship at UWindor. It has been accepted for inclusion in Electrical and Computer Engineering Publications by an authorized administrator of Scholarship at UWindor. For more information, please contact [scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca).

## Article

# Explainable Artificial Intelligence Approach for Diagnosing Faults in an Induction Furnace

Sajad Moosavi <sup>1,\*</sup>, Roozbeh Razavi-Far <sup>2</sup>, Vasile Palade <sup>3,\*</sup> and Mehrdad Saif <sup>1</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada; msaif@uwindsor.ca

<sup>2</sup> Faculty of Computer Science, University of New Brunswick, Fredericton, NB E3B 5A3, Canada; roozbeh.razavi-far@unb.ca

<sup>3</sup> Centre for Computational Science and Mathematical Modelling, Coventry University, Coventry CV1 5FB, UK

\* Correspondence: moosavis@uwindsor.ca (S.M.); vasile.palade@coventry.ac.uk (V.P.)

**Abstract:** For over a century, induction furnaces have been used in the core of foundries for metal melting and heating. They provide high melting/heating rates with optimal efficiency. The occurrence of faults not only imposes safety risks but also reduces productivity due to unscheduled shutdowns. The problem of diagnosing faults in induction furnaces has not yet been studied, and this work is the first to propose a data-driven framework for diagnosing faults in this application. This paper presents a deep neural network framework for diagnosing electrical faults by measuring real-time electrical parameters at the supply side. Experimental and sensory measurements are collected from multiple energy analyzer devices installed in the foundry. Next, a semi-supervised learning approach, known as the local outlier factor, has been used to discriminate normal and faulty samples from each other and label the data samples. Then, a deep neural network is trained with the collected labeled samples. The performance of the developed model is compared with several state-of-the-art techniques in terms of various performance metrics. The results demonstrate the superior performance of the selected deep neural network model over other classifiers, with an average F-measure of 0.9187. Due to the black box nature of the constructed neural network, the model predictions are interpreted by Shapley additive explanations and local interpretable model-agnostic explanations. The interpretability analysis reveals that classified faults are closely linked to variations in odd voltage/current harmonics of order 3, 11, 13, and 17, highlighting the critical impact of these parameters on the model's prediction.

**Keywords:** induction furnace; fault diagnostic; deep neural network; explainable AI; XAI; LIME explanation; SHAP interpretation



**Citation:** Moosavi, S.; Razavi-Far, R.; Palade, V.; Saif, M. Explainable Artificial Intelligence Approach for Diagnosing Faults in an Induction Furnace. *Electronics* **2024**, *13*, 1721. <https://doi.org/10.3390/electronics13091721>

Academic Editor: Bingqiang Chen

Received: 26 March 2024

Revised: 24 April 2024

Accepted: 26 April 2024

Published: 29 April 2024

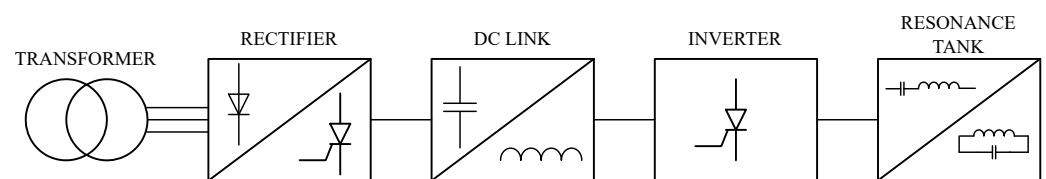


**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

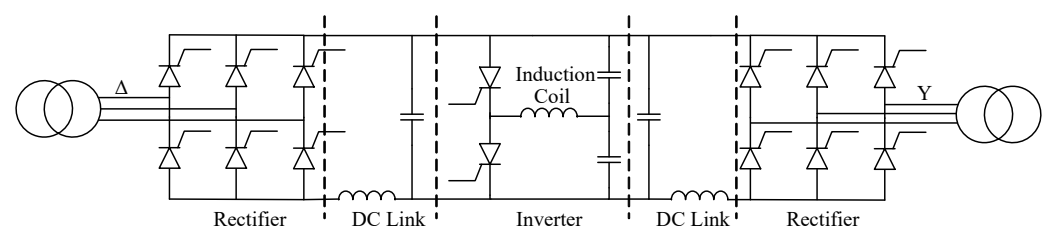
Induction heating furnaces operate based on eddy current losses in electrically conductive materials when exposed to varying magnetic fields. Induction furnaces (IFs) are characterized by efficiency, rapid heating, controllability, and cleanliness. Thus, they are widely used in industry for heating, melting, welding, or hardening metals. As shown in Figure 1, the power supply of an IF comprises four main parts: a rectifier, a DC link, an inverter, and a resonance tank [1]. The inverter topology could be a current source or voltage source feeding a parallel or series resonant circuit, respectively. In this work, the furnaces have voltage source-type series-resonant inverter configurations, as shown in Figure 2. The six-phase fully controlled rectifier consists of silicon-controlled rectifiers (SCRs) that facilitate the primary AC to DC conversion. The rectifier produces a fixed DC voltage and the SCRs are used for AC line interruption in the case of faults in circuits following the rectifier. The current limiting reactor is placed in series with the inverter to limit the inrush current during a short circuit and SCR faults. The DC capacitors filter the rectifier output voltage ripple and act as voltage sources for series resonance circuits. The furnace coil and

series AC capacitors form a resonance circuit with a variable resonant frequency as the inductance varies depending on the charge state inside the power coil. When driving the series resonant load, there is always a phase shift between the output voltage and current except at the resonance frequency in an ideal state. Therefore, there are anti-parallel diodes with SCRs to bypass the current in this state. In order to regulate the power, the switching frequency is varied above or below the resonance. The circuit impedance of the tank circuit increases, and, consequently, the amount of power delivered to the coil decreases. The power of a furnace is automatically limited to regulate these parameters: the AC capacitor voltage, inverter current, inverter frequency, and furnace voltage. If either of these values exceeds its preset value, the inverter's power is then automatically diminished by the inverter control board even if the power is set to maximum by the operator.



**Figure 1.** General block diagram of induction furnaces.

In high-power IFs, the system components must be cooled with a coolant, such as water, to dissipate heat generated in power semiconductors, busbars, capacitors, and induction coil. The coolant circulates through heatsinks, copper tubes, and an induction coil to transfer the heat to the cooling towers. Any interruption in water flow may cause serious damage to components, particularly power semiconductors. The operations of a foundry are characterized by their continuous nature, where the chemical and metallurgical processes are at an elevated temperature and the products have to remain hot [2]. Interruption and stoppage in their operations are very costly and must be minimized. Therefore, a data-driven diagnostic system can enhance the safe operation of systems and help to remain productive [3–5]. Furthermore, when a detected fault alerts the operator to a potential failure, providing an explanation will support the prediction and could even prevent an unnecessary shutdown. This clarification is essential for the integration of AI into critical applications like manufacturing, where decisions can have a significant impact. In this paper, a deep learning-based framework has been proposed for diagnosing IF faults. A post hoc explainable artificial intelligence (XAI) module has been devised in the proposed framework, which can further interpret the constructed black box deep learning model to provide understandable outcomes and establish trust.



**Figure 2.** Induction furnace with series resonance half-bridge inverter.

## 2. Related Works

Fault detection and diagnosis of machinery are among the primary applications of artificial intelligence (AI) in industry, as the health of machines and equipment plays a key role in enhancing the productivity and efficiency of systems [6]. In general, feature extraction and selection techniques have been widely used along with data-driven diagnostic systems [7,8] as they enable the identification of faulty patterns and features from vast amounts of data samples, facilitating the accurate diagnosis of faults and enhancing the system's ability to detect potential faults effectively. However, these techniques cannot provide any explanation of the diagnostic system's decision and the impact of features.

With the success of using intelligent diagnostic tools, there is a great demand for their explainability to establish trust. XAI helps in discovering intelligent predictive models and explaining their outcomes for the operators. In the literature, there is no research work available to address diagnosing faults in IFs. However, various research works have been accomplished to address the problem of diagnosing faults in other systems by XAI tools. In the following, different approaches are discussed and compared in terms of incentives for using XAI. For fault detection in high-speed rail systems, four knowledge extraction methods are compared in [9], including case-based reasoning (CBR), association rule learning (ARL), the Bayesian network (BN), and a neuro-fuzzy system with respect to five characteristics. The Bayesian network is chosen for the work because of the explainability of the constructed model integrated with correlation-based feature selection (CFS). The health of water pumps is monitored for preventive maintenance in [10]. The proposed method makes use of a Type-2 fuzzy logic system (FLS) to make the model interpretable. With 100 rules and four antecedents per rule, the system provides high explainability and trust compared to other opaque box models. In [11], bearing faults have been diagnosed using Shapley additive explanations (SHAP) followed by a K-nearest neighbor (KNN) classifier. The proposed approach could be adapted to work on different datasets with different configurations, as a generalized model. Anomaly detection of rotating machinery was performed in an unsupervised manner in [12]. The work compares 11 different methods and chooses the isolation forest method due to its good overall performance. Then, using a model-agnostic SHAP and a model-specific local depth-based isolation forest feature importance (DIFFI), the feature importance is extracted. Finally, the root cause analysis outputs the most important specific features. The linear motion guide faults are discovered in [13]. This work employs a 1D convolutional neural network (CNN) for time-domain training, and, then, uses a frequency-domain-based gradient-weighted class activation map (FG-CAM) to visualize the classification criteria. The notion of Grad-CAM is to interpret the model by using the process of learning in reverse. The authors expect that the proposed method can be applied to various complex physical models. Similar approaches are presented in [14–16]. For anomaly detection and prognosis of gas turbines, Bayesian long short-term memory (LSTM) is employed and the outputs are explained using SHAP [17]. Alongside the prediction, two output layers, the AU layer and the EU layer, are added to generate data and parameter uncertainty. The uncertainty mirrors the model's confidence in predictions. The root-mean-square error (RMSE) and early prediction score are calculated for the performance evaluation. Moreover, for the evaluation of the SHAP explanation, two metrics—local accuracy and consistency—are examined. Artificial neural networks (ANNs) and support vector machines (SVMs) are used in [18] for heat recovery failure detection in the air handling unit (AHU). Then, the model's decisions are justified using the explanation provided by local interpretable model-agnostic explanations (LIME). The author states that the trustworthiness of the model is the aim of explanations provided to the user. For fault diagnosis of gearboxes [19], a Deep CNN (DCNN) has been developed based on the layer-wise relevance propagation (LRP) technique. The LRP is an explanation method for the DCNN classifier, which quantifies the contribution of the individual inputs to the output. In other words, LRP brings more transparency to the decision made by DCNN and opens up the opportunity for widespread DCNN usage for machine fault diagnosis. In another work, Grezmak et al. [20] utilize CNN and LRP for motor fault diagnosis. They investigate the performance of the CNN by time-frequency spectra images of vibration signals measured on an induction motor. Amarasinghe et al. [21] propose a framework for deep neural network (DNN)-based anomaly detection and a post hoc explanation generated by LRP. Similarly, condition monitoring of hydraulic systems has been proposed using a framework that combines a DNN with DeepSHAP for interpretability [22]. An XAI-based chiller fault detection and diagnosis is presented in [23]. This work uses the extreme gradient boosting (XGBoost) model, an ensemble of classification and regression trees. The explanation introduced by LIME helps to detect preliminary faults by possibly resorting to human operators. In addition, this information improves accuracy,

reduces fault detection time, and establishes trust with the field personnel. The remaining useful life (RUL) is an estimate of the remaining time for the repair or replacement of a machine. To predict the RUL of turbofan engines, [24] implemented a DNN model coupled with LIME. The outcomes indicate that the model effectively understands the underlying physics with simpler data sets, and LIME provides clear explanations. However, both the model and explainer struggle to deliver satisfactory results with more complex data. In the same application, three algorithms are employed in [25]: CNN, LSTM, and Bidirectional LSTM. The effect of each input variable is confirmed through SHAP. The SHAP provides an intuitive visualization of the effect of each feature on the predicted results. Conventional CNN filters are not transparent and they include noisy and undesirable spectral shapes. This issue is addressed in [26] with the utilization of SincNet. SincNet encourages the first layer of the CNN to generate interpretable filter kernels. The method, referred to as Deep-SincNet, demonstrated better performance, greater interpretability, enhanced immunity to noise, and lower implementation costs compared to conventional CNNs. Resembling the F. B. Abid et al. method, T. Li et al. [27] replaced the first layer of the CNN with a wavelet convolutional layer (CWConv). This wavelet-driven deep neural network is called WaveletKernelNet (WKN) and represents more meaningful kernels. The accuracy of WKN is 10% higher than the standard CNN, plus the CWConv layer is interpretable.

Apart from explainability, two data-driven studies were conducted by Yolim Choi et al. [28,29] to predict the RUL of an IF refractory. They prove that multilayer perceptron (MLP) performs better than recurrent neural networks (RNNs) and LSTM methods in predicting RUL. The study employs basic electrical parameters such as input/output power, voltage and current, converter DC voltage, and frequency. In the second work, a novel s-convolutional LSTM method is proposed to predict RUL. In terms of RMSE and Pearson correlation coefficient (PCC) metrics, it is shown that s-ConvLSTM outperforms MLP and LSTM.

In our study, the time and frequency domain parameters are extracted from the IF power supply. Then a framework is introduced consisting of, first, a semi-supervised anomaly detector to identify faulty samples from normal data samples; second, a DNN to classify fault type in a supervised manner; and third, the use of XAI methods to interpret the predictions. To the best of the authors' knowledge, this is the first study that introduces XAI-based fault diagnosis in IF, utilizing empirical data samples from the industry.

### 3. Problem Statement

Fault detection and diagnosis help take recovery actions before a critical failure happens, causing subsequent damages to other parts of equipment, or, in extreme cases, exposing personnel to safety hazards. IFs play a key role in heating and melting metals in industry. In foundries with a continuous casting machine, the uninterrupted operation of the furnaces is essential, as any stoppage could trigger a breakout in the production process. Additionally, a prolonged shutdown of a furnace during operation can lead to melt blocking in the crucible if the charge is not dumped. Early detection and the diagnosis of faults in IFs provide the following benefits:

- Prevents further damage to components, especially power semiconductors.
- Reduces repair time and overall downtime.
- Reduces the repair cost.
- Enhances furnace system health and performance.
- Boosts productivity.
- Mitigates potential safety hazards.

Data-driven fault detection and diagnostic models employ deep learning solutions and are built by means of sensor data. There are no previous studies on fault diagnosis of IFs in the literature. In this study, we provide a deep learning solution to detect and diagnose faults using data samples collected from an energy analyzer device. There are several faults related to the electrical components. In the high-power furnace cabinet, there are several AC and DC oil-filled water-cooled capacitors. Capacitors are very sensitive to

temperature increases; therefore, the water outlet temperature typically must not exceed 45 °C. The most common failure in these capacitors is dielectric damage due to excessive current, voltage, or temperature. A damaged capacitor often results in a phase-to-phase or phase-to-ground short circuit and draws significant inrush current. IFs rely on electronic circuits for rectifier and inverter control while providing protection. The control boards issue trigger signals to power circuit switching elements. Issues in these electronic circuits have diverse effects on the operation of the furnace. For instance, a faulty inverter control card may trigger thyristors on opposite legs at the same time, thus causing a short circuit across the DC rail. Improper firing might lead to a short circuit or harmonic generation. Conducted/radiated interference, power supply issues, loose connection, and component overheating are common reasons for the control system malfunctioning. The changeover switch (COS) is used for transferring power from one crucible to another in case of crucible stops for relining or patching. Water flows through the switch to stabilize its temperature. These high-current offload switches are made up of high-conductivity copper with silver contacts to minimize resistance and loss. A common failure in COS is the corrosion of switchblades due to misalignment or overheating. The misalignment or gap between moving and stationary contacts leads to small arcs and gradually melts the contact's surface. Additionally, a drop in water pressure or a rise in inlet water temperature can lead to erosion of the silver contacts due to overheating. Earth leakage is a prevalent fault among others and occurs when there is a current leakage to the ground at the inverter output side. The most severe accident is when molten metal penetrates through the furnace lining and touches the coil. The ground leak detector (GLD) circuit protects the furnace from such accidents. Since the power circuit is isolated from the ground, the GLD senses ground leaks by applying a DC voltage between the inverter output and the ground. A malfunctioning GLD either fails to detect earth leakage current or may trigger false alarms. The other group of faults is related to semiconductors including thyristors, flywheel diodes, or components in the snubber circuit. These types of faults are due to over-voltage, over-current, short circuits, earth leakage, overheating, or loose connection. In rare cases, the  $di/dt$  reactor fails with an inductance change or insulation breakdown. There are plenty of fixed temperature switches on the water supply manifolds to shut off the inverter when the temperature surpasses a threshold level. The temperature of the components generally rises due to cooling system problems such as faulty pumps, clogged water hoses or pipes, and water pressure drops. Based on the authors' expertise in IFs, faults often originate from a single source within the furnace and can escalate to affect other furnace components. The main goal of this framework is to detect faults early to prevent subsequent damage.

The deep learning approach attempts to predict the IF faults with a high degree of precision. However, only diagnosing a potential fault would not be enough to justify the system status and subsequent actions, particularly when there is a substantial risk of unintended bias during the learning phase, and a false shutdown can lead to serious consequences. Thus, it is imperative that the system's alarms are supported by a comprehensive set of justifications, which provide the control room operators with confidence. The XAI techniques can analyze and clarify the predictions of the trained models, such as neural networks, to pinpoint the features that contribute most to the decision. This ensures more reliable predictive models, thereby securing the trust of end-users.

#### 4. Proposed Approach

This work proposes a transparent and efficient method for diagnosing faults in IFs that are solely based on real-time monitoring of electrical parameters collected by a power quality and energy analyzer device. The system setup is shown in Figure 3. There are 218 parameters recorded by each energy analyzer in two sets: general and harmonic parameters. All measurements are sent to a central data server and can be visualized or exported for monitoring and further analysis. General parameters include line voltage (VL12, VL23, VL31), phase voltage (VT, V1, V2, V3), phase current (IT, I1, I2, I3), residual current (Inull), active power (PT, P1, P2, P3), reactive power (QT, Q1, Q2, Q3), apparent



power (ST, S1, S2, S3), frequency (Freq), first harmonic power factor (PFT, PF1, PF2, PF3), harmonic distortion for current and voltages (THD, OHD, EHD), K-factor, total active and reactive energy (generated or consumed), demand (Dem), voltage, and current unbalance rate (UnbV, UnbI). The Avr, Max, and Min prefixes denote the average, maximum, and minimum recordings of the voltage and current. The K-factor is defined as follows:

$$K_{factor} = \frac{\sum_{n=1}^k (f_n \times n)^2}{\sum_{n=1}^k (f_n)^2} \quad (1)$$

where  $f_n$  is the  $n$ th harmonic value. The K-factor represents the weighting of harmonic load currents based on their impact on transformer heating. A K-factor of 1.0 signifies a linear load with no harmonics and higher values indicate that the load introduces harmonics, leading to extra heating effects. The minimum, maximum, or average values of the voltage and current are evaluated based on cycle-by-cycle values within a period of one minute. The residual current is the vector sum of the phase currents. The maximum demand value is the average demand during a period of 15 min. Moreover, the voltage and current unbalance rate is defined as follows:

$$UnbV = \frac{V_m - \bar{V}}{\bar{V}} \times 100 \quad (2)$$

$$UnbI = \frac{I_m - \bar{I}}{\bar{I}} \times 100 \quad (3)$$

where  $\bar{V}$  and  $\bar{I}$  are the average values and  $V_m$  and  $I_m$  are the maximum values of the voltage and current. Moreover, harmonic parameters consist of harmonic data up to the 22nd order for the phase voltage and current. Assuming  $m_n$  as the RMS value of the  $n$ th harmonic, the total harmonic distortion (THD) is determined using the following equation:

$$THD = \frac{\sqrt{\sum_{n=2}^{max} m_n^2}}{m_1} \quad (4)$$

Similarly, even harmonic distortion (EHD) and odd harmonic distortion (OHD) are defined as follows:

$$EHD = \frac{\sqrt{m_2^2 + m_4^2 + m_6^2 + \dots}}{\sqrt{m_1^2 + m_2^2 + m_4^2 + m_6^2 + \dots}} \quad (5)$$

$$OHD = \frac{\sqrt{m_3^2 + m_5^2 + m_7^2 + \dots}}{\sqrt{m_1^2 + m_3^2 + m_5^2 + m_7^2 + \dots}} \quad (6)$$

In total, there are 92 general parameters and 126 harmonic parameters. From these, 15 features with low correlation were removed from the dataset. The recording intervals for the general and harmonic parameters differ. General parameters are recorded every minute, while harmonic parameters are recorded every 10 min. Therefore, the date and time stamp of the samples are used to downsample the general parameters to properly match the harmonic parameters.

Most of the faults cause subsequent damage to other components such as semiconductors, capacitors, snubbers, and control systems. Therefore, the initial cause of a failure is identified in the maintenance reports and flagged as a fault case. Considering these criteria, a total of 290 faults are considered in this study. Then, for each fault case, the energy analyzer data samples were cropped from the beginning of the furnace cycle to the time of the fault occurrence. A furnace cycle is considered from the time of filling the furnace with cold scrap and starting the inverter to the moment of inverter shutdown by the operator and discharging the molten metal. For the above furnace, this cycle lasts for roughly two hours and normally repeats ten times a day according to the furnace production reports.



**Figure 3.** The general diagram of the IF system.

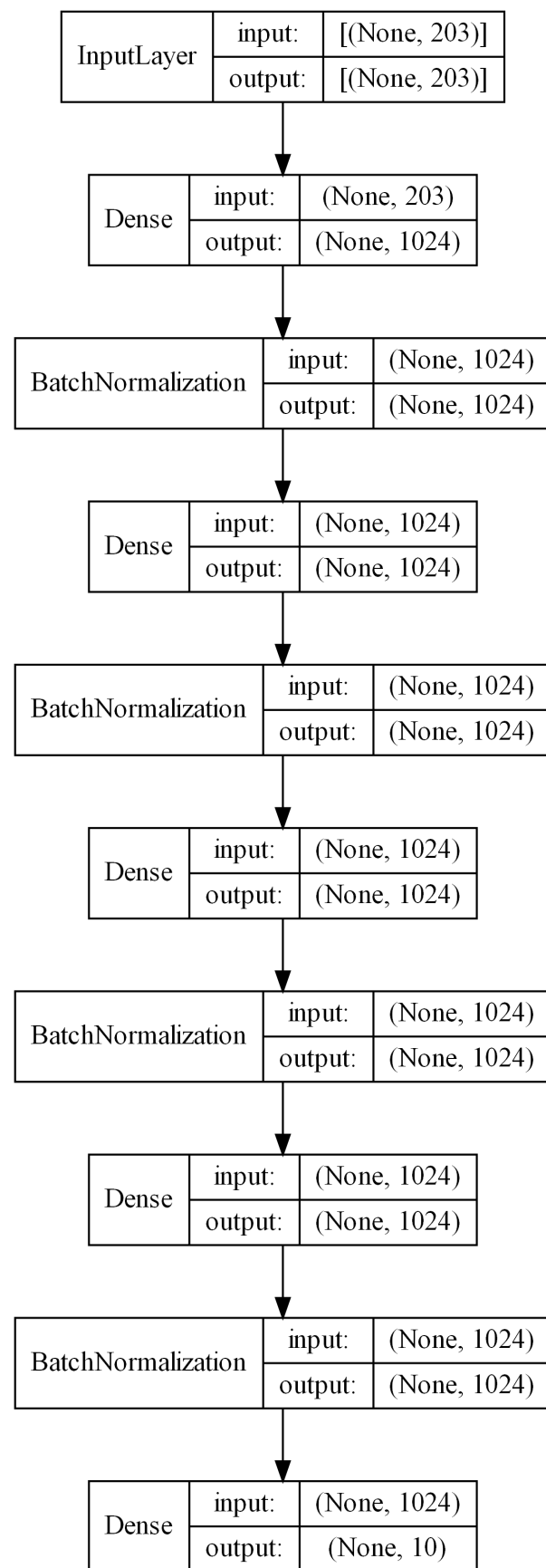
In order to label the samples, we employ the local outlier factor (LOF) algorithm, which is an unsupervised anomaly detection method. LOF calculates the local density deviation of a given data point in relation to its neighbors. As there are many normal samples available in our database, we use LOF in a semi-supervised manner and then try to find novelty in the samples. As a result, we can find outliers among collected samples. Detected outliers are assigned a label according to the corresponding fault case.

After data labeling, we use a deep neural network to train a model and diagnose fault samples in real time. For the DNN model buildup, we try to minimize the number of layers in order to decrease the computation time, while keeping acceptable prediction accuracy. We adopted the DNN method in this paper due to its effectiveness in real-time fault diagnosis, particularly in the vital and high-value contexts of induction furnaces in foundries. As shown in Figure 4, the proposed model is constructed based on five dense layers with batch normalization. This setting is based on extensive iterative experimentation. The Softmax function is employed for the classification tasks to predict the probability of different classes, and the categorical cross-entropy is utilized as the cost function for the model training. The cross-entropy for each sample,  $s$ , in multi-class classification can be defined as follows:

$$CE_s = - \sum_{i=1}^n y_{s,i} \log(p_{s,i}) \quad (7)$$

where  $n$  is the number of classes and  $p$  is the predicted probability of class  $i$  for the sample  $s$ .





**Figure 4.** Different layers of the constructed model.

The interpretation of the predictions of the complex DNN model is of paramount importance. The explanations will enable users to understand and trust the predictions generated by the model. The LIME algorithm [30] is used to interpret the outcome of the proposed diagnostic model. LIME is a local approach that can explain the conditional interactions between features and classes for a single sample. At a high level, the algorithm follows these steps:

- It perturbs the desired prediction to create replicated feature data that has slight value modifications.
- It calculates the distance between each perturbed data point and the original sample.
- It obtains the outcomes of perturbed data using our black box model.
- It selects features that have the most contribution to the model outcome.
- It approximates a linear model using the perturbed data and selected features.
- An explanation is created based on the feature weights of the approximated linear model.

The explanation for the sample,  $s$ , is represented as follows:

$$\text{explanation}(s) = \operatorname{argmin}\{L(f, g, \pi_s) + \Omega(s)\} \quad (8)$$

where  $g$  is the linear approximation of the original model  $f$  around sample  $s$  with locality  $\pi_s$ .  $L$  is a function that measures the distance between  $f$  and  $g$  in the given neighborhood. It represents how closely the explanation matches the original model's prediction.  $\Omega(s)$  is a regularization term that measures the complexity of the explanations, i.e., the number of non-zero coefficients.

The LIME results are verified by applying SHAP [31]. SHAP is a unified framework based on the Shapley value. SHAP can produce local and global explanations for a model. In presenting a global explanation, we explore the importance of a particular feature for predictions made by the model.

## 5. Experimental Results and Discussion

The data were collected from a foundry operating three identical 15-ton, 5 MW series-resonance half-bridge inverters. The configuration of the furnaces is shown in Table 1. All furnaces are fed through a common 20 KV switchgear. However, dedicated voltage and current transformers are connected to the input of each furnace power transformer for energy analyzer measurements. The data collection began on the date of commissioning for each furnace.

**Table 1.** Induction furnace specification.

Specification	Value
Rated Power	5000 KW
Input Voltage	550–690 V
Output Frequency	200–400 Hz
Output Voltage	3–4 Kv

### 5.1. Outlier Detection

The raw data for each fault are collected and labeled accordingly. The LOF algorithm marks each sample as normal or an outlier. The algorithm was initially trained using 21,603 normal samples collected from fault-free furnace operations on random days. We ensured no faults occurred within 24 h before and after the sampling period. Each sample identified as an outlier is given an appropriate fault label at the end of that period. Table 2 displays the outliers detected by the LOF algorithm.

**Table 2.** Fault statistics.

Fault Types	Label	#Fault Cases	Total Samples	#Outlier Detected
Capacitor	1	37	487	108
Control System	2	24	333	82
Change Over Switch (COS)	3	13	181	19
Earth Fault	4	88	1071	302
Flywheel Diode	5	6	67	9
di/dt Reactor	6	7	55	23
Thyristor (SCR)	7	29	370	116
Temperature	8	86	1007	192

### 5.2. Model Evaluation

Upon processing the raw data, the proposed model is constructed and then evaluated with test data. To evaluate the efficiency of the model, the accuracy of its predictions is compared with those obtained by the ensemble learners, e.g., extreme gradient boosting (XGB), light gradient boosting (LGB), and random forest (RF), as well as classic machine learning (ML) methods including logistic regression (LR), MLP, KNN, SVM, DT, and NB classifiers. Comparing various classifiers enables us to determine the best predictive model for this application. Our proposed deep learning model employs the Stochastic Gradient Descent optimizer to control the rate of gradient descent. In the data pre-processing stage, we normalized the data by applying the min-max scaling technique. The batch size is set to 128, and the number of epochs varies, i.e., the training automatically stops when validation loss reaches a minimum value. Validation data constitutes 10% of the training data. The performance of the proposed model is assessed using four metrics: accuracy, precision, recall, and F-measure. The F-measure, also known as the F1 score, integrates precision and recall metrics to strike a balance between the two. Since data are class imbalanced, we use weighted averaging for each class metric's computation. For instance, the precision for class  $i$  is determined as follows:

$$Pr_i = \frac{TP_i}{TP_i + FP_i} \quad (9)$$

TP stands for true positive and FP means the number of false positives. The weight precision is defined as follows:

$$Pr = \sum_{i=1}^n W_i \times Pr_i \quad (10)$$

where  $n$  is the number of classes and  $W_i$  is the weight assigned to class  $i$ :

$$W_i = \frac{TP_i + FP_i}{\sum_{i=1}^n (TP_i + FP_i)} \quad (11)$$

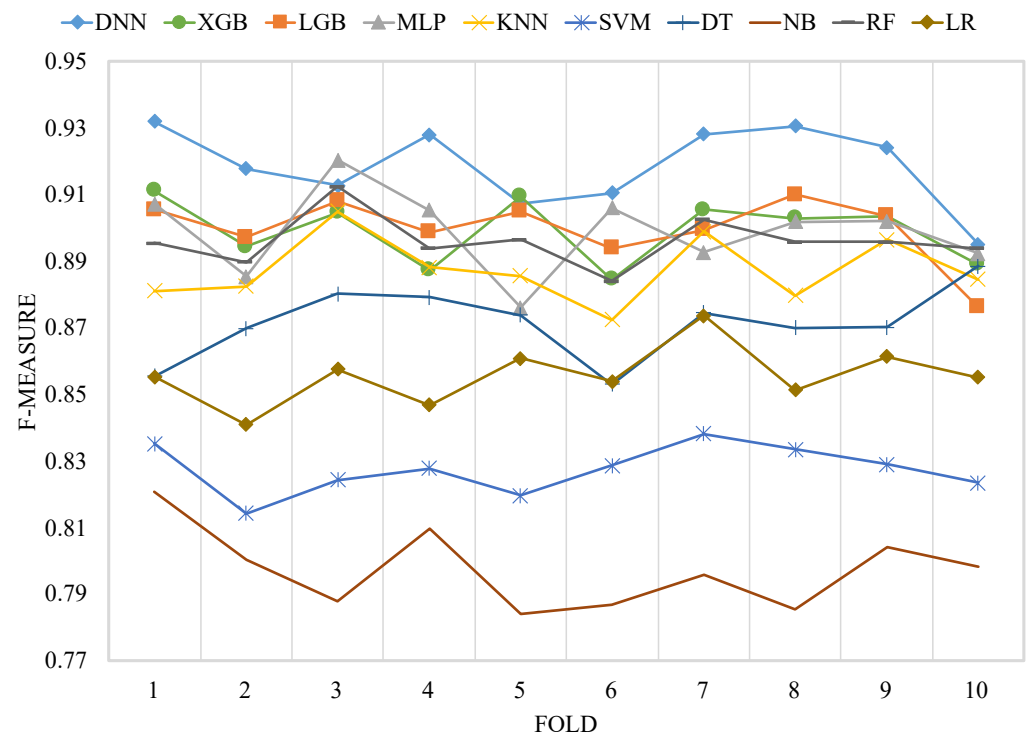
The model is evaluated through 10-fold cross-validation, which randomly partitions data into 90% for the training and 10% for the test. The class ratio is maintained throughout the 10 folds the same as the ratio in the original data. In each fold, 10% of the training set is kept for validation and tuning the model's hyper-parameters. In our analysis, various classifiers were compared in terms of accuracy, precision, recall, and F-measure.

In prediction tasks, the F-score is a widely used evaluation metric that leads to a single score accounting for both precision and recall concerns. Table 3 illustrates the performance comparison results. The obtained metrics reported in this table are the average values over all folds. It can be seen that DNN reveals the best values across all measures compared to other classifiers. An average F-measure of 0.9187 is scored by DNN, followed by 0.8998 is reached by LGB. Tree ensemble methods stand in second place with nearly equal figures. MLP shows a higher precision of 0.9025 than tree ensemble algorithms but due to the increased false negative rate in MLP, the recall and F-measure values are alleviated. Despite achieving a high precision of 0.8736, NB performs poorly in other evaluation metrics.

The main reason for the poor performance of the Bayesian classifier is the unequal class distribution and the insufficient number of fault cases in our experimental data. Figure 5 shows the F-measure values obtained by each model at each fold of cross-validation. Using the F-score metric, it can be verified that the selected DNN model outperforms the lowest average score of the naïve Bayes classifier by 15.22%.

**Table 3.** Performance metrics obtained by each method.

Classifier	Precision	Accuracy	Recall	F-Measure
DNN	0.9127	0.9142	0.9287	0.9187
MLP	0.9025	0.8859	0.9007	0.8989
XGB	0.8894	0.8990	0.9188	0.8992
LGB	0.8914	0.8982	0.9175	0.8998
RF	0.8880	0.8972	0.9177	0.8960
KNN	0.8778	0.8834	0.9021	0.8874
SVM	0.8144	0.8573	0.8801	0.8274
DT	0.8746	0.8514	0.8697	0.8715
LR	0.8376	0.8647	0.8857	0.8557
NB	0.8736	0.7237	0.7440	0.7973

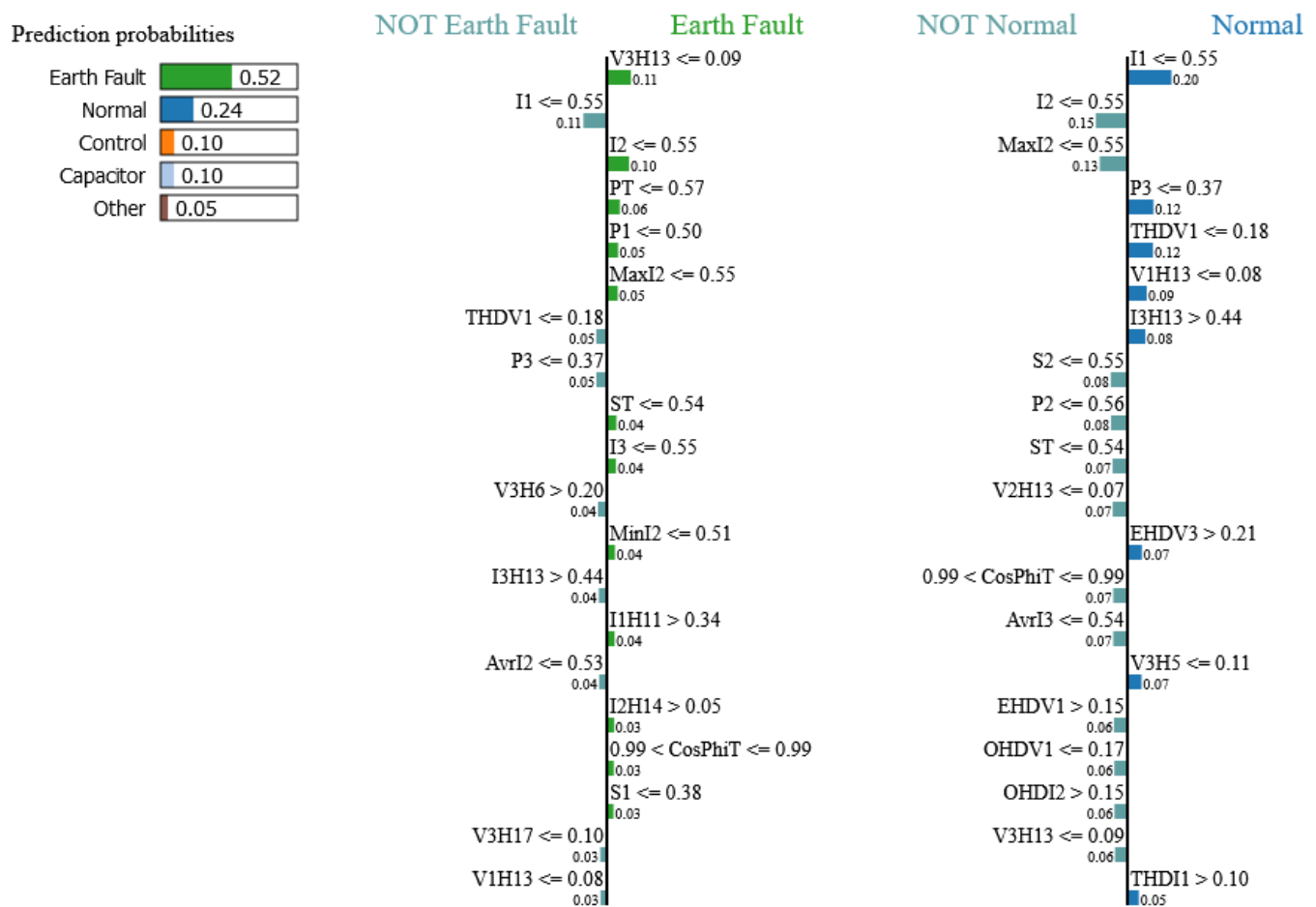


**Figure 5.** F-measure values obtained by each model at each fold of cross-validation.

## 6. Explanation for Predictions

The explanations provided by LIME and SHAP will enable users to understand and interpret the predictions generated by the model. We employ both techniques to explain the predictions of our model. A randomly chosen earth fault sample with a probability of 52% is used for further explanation. For faster and improved performance, the DeepSHAP [31] approach is used for our model to compute SHAP values. As shown in Figures 6 and 7, both algorithms are configured to show only the top 20 most important features. The features are sorted in descending order from top to bottom. In Figure 7, the top features are listed on the vertical axis while the horizontal axis indicates SHAP values. There are 11 common features identified by both methods. In both figures, it can be realized that the phase III voltage harmonic of order 13 (V3H13) has the most significant impact on the

predicted result. On the left side of the LIME plot, prediction probabilities of the different faults have been presented. There is a 52% probability of an earth fault while having a 24% chance of being in the normal state.



**Figure 6.** LIME feature weights for a single representative sample of the earth leak fault.

Therefore, the features with positive contributions are presented in green and placed on the right side of the vertical bar. On the contrary, the features with a negative effect on the decision are positioned on the left side of the bar. Likewise, in the SHAP plot, the red bars indicate factors with positive SHAP values, such as V3H13, I2, PT, etc., and negative ones are in blue, like I1 and THDV1, etc. This graph demonstrates how to transition from the mean of all predictions,  $E[f(x)]$ , to the predicted probability.

Furthermore, we can visualize SHAP values for a set of correct predictions of the same class to have a global explanation. A global summary plot is shown in Figure 8, combining feature importance with feature effects. The vertical axis illustrates the prominent features, whereas the horizontal axis corresponds to SHAP values. Each row represents a feature and each point illustrates a sample. The red dots show high feature values while the blue dots show low values. The features are ordered according to their importance. Some features have distinguishable red and blue dots distributed along the x-axis so that we can easily infer their impact. Apparently, the high values of the total power factor (CosPhiT) and 13th harmonic of phase I voltage (V1H13) together with low values of 13th harmonic of phase III voltage (V3H13), and current (I3H13) and 11th harmonic of phase I voltage (V1H11) elevate SHAP values and the probability of earth fault. Moreover, by comparing this plot and LIME local explanation for one of these samples in Figure 6, we can observe similar parameters like V3H13, I3H13, CosPhiT, and V1H13 contributing in the earth fault decision.

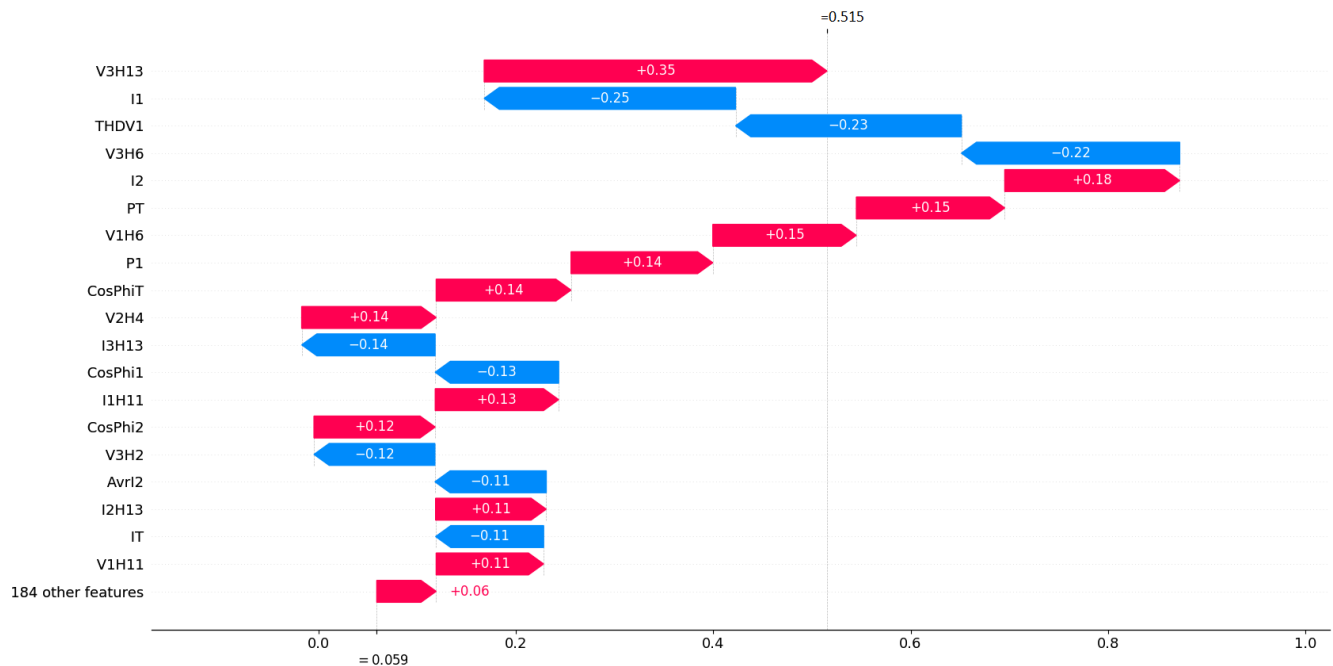


Figure 7. SHAP values for a single representative sample of the earth leak fault.

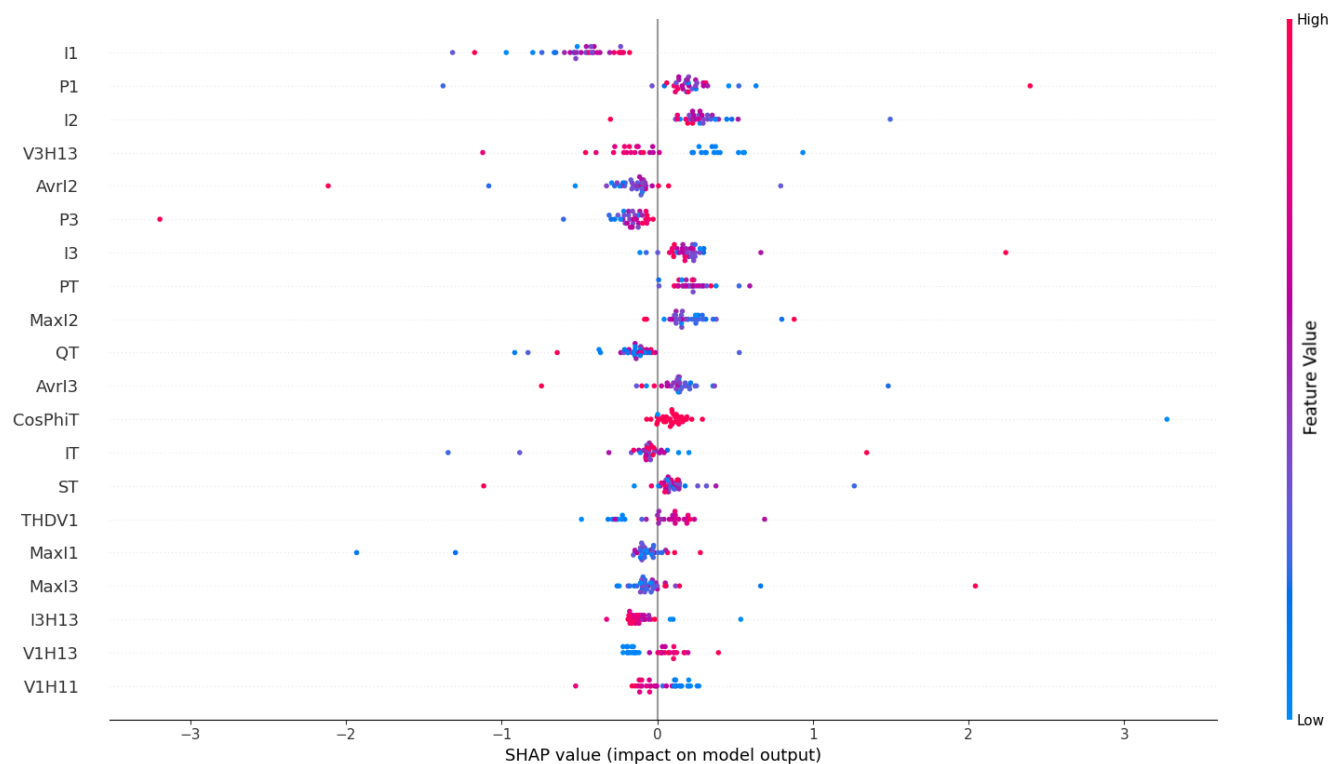


Figure 8. Global explanation using SHAP for the earth leak fault.

For further verification of explanations, we applied post hoc explainers to other constructed models. We have chosen a sample earth fault from the test set that all models correctly predict. As anticipated, each model presents varying probabilities for this particular case. Table 4 displays the top 20 features ranked by the score for each model, along with the associated probabilities. For instance, DNN identified the fault as an earth fault with 81% probability, while the XGB demonstrated higher confidence at a probability of 96%. We applied TreeSHAP for XGB and DT classifiers. Features that are reported by both



LIME and SHAP are marked in bold. It can be seen that in cases of deep learning model and SVM, both explainers uncover 55% of the same features. Additionally, it is evident that some features are highlighted by at least five models as reasons for the earth fault. These common features are V3H13, V3H21, P2, I2, and CosPhi3, where V3H13 is reported by all models. Therefore, we can conclude that a decrease in the 13th harmonic of phase III voltage could be a solid indication of the earth fault.

**Table 4.** Feature importance comparison of models W.R.T. individual fault samples.

DNN		XGB		KNN		SVM		DT		NB		LR	
p = 0.81		p = 0.96		p = 1.00		p = 0.56		p = 0.92		p = 1.00		p = 0.55	
SHAP	LIME	SHAP	LIME	SHAP	LIME	SHAP	LIME	SHAP	LIME	SHAP	LIME	SHAP	LIME
<b>I2</b>	I1	Freq	P2	<b>AvrI2</b>	V1H18	V3H3	I2	<b>P2</b>	S3	V1H14	V1H20	<b>Q2</b>	Q2
<b>MaxI2</b>	I2	<b>P2</b>	S2	<b>AvrI3</b>	I2H22	<b>P2</b>	P2	AvrI1	P2	V1H12	V3H13	<b>QT</b>	V3H13
<b>P2</b>	PT	PF1	UnbV	<b>V3H13</b>	V3H13	<b>Q2</b>	V3H17	UnbV	MinI2	V2H12	I1H13	<b>I1H13</b>	QT
AvrI3	V3H13	V3H21	V3H13	<b>AvrI1</b>	ST	<b>S2</b>	S2	I1H11	V3H13	<b>P2</b>	I3H5	<b>V3H13</b>	V3H21
<b>PT</b>	P2	CosPhi2	V1H18	P2	AvrI1	I1H13	V3H21	I2H8	I2H4	I3	P2	<b>I2</b>	P3
I3	MaxI2	UnbI	I2H8	I1	MinV3	IT	Q2	CosPhi3	I2H16	V2H14	PT	<b>V1H7</b>	I2
<b>V3H13</b>	CosPhi3	<b>S2</b>	I2H21	MaxI3	AvrI2	<b>Freq</b>	I1	I2H5	V2H20	<b>P1</b>	V3H21	<b>Freq</b>	Freq
<b>THDV1</b>	S1	<b>UnbV</b>	V1H15	MaxI2	AvrI3	<b>I2</b>	PT	Freq	Q3	AvrV2	ST	I2H19	V1H12
<b>VL23</b>	V3H21	V3H4	V2H12	Q2	MinI2	<b>I3</b>	ST	<b>S3</b>	MinV2	S2	V1H2	V2H13	AvrI1
<b>ST</b>	V3H17	AvrI1	V2H8	PT	P1	I2H15	S1	I2H11	KFactorI1	I3H15	KFactorI3	S2	I1H13
MinI2	ST	V1H14	I2H6	I3	I1H14	<b>V3H13</b>	V2H13	PFT	I2H19	<b>IT</b>	I3H20	I2H15	I3H6
CosPhiT	V2	CosPhi3	THDI3	<b>ST</b>	I3H20	<b>V2H13</b>	OHDV3	CosPhiT	AvrI3	MinI1	I3H9	I3	I2H2
<b>S1</b>	MinV2	<b>I2H15</b>	IT	MinI3	OHDI3	<b>I1</b>	I3	V3H19	OHDV1	UnbV	V1H6	MaxI1	V1H7
V1H12	THDV1	I2H16	I2H2	MinI1	S2	<b>PT</b>	S3	<b>V3H13</b>	V3H21	I2H19	V2H3	OHDV3	V1H13
<b>CosPhi3</b>	P3	<b>MinV1</b>	I3H15	<b>S2</b>	I2	QT	VL23	MinI3	EHDV1	EHDV2	I2H16	IT	KFactorV1
V1H19	VL12	<b>V3H13</b>	KFactorI3	<b>I2</b>	V2H20	CosPhi3	OHDV2	I2H10	I3H14	MaxI1	V1H18	THDV3	MinI2
V3H3	AvrI2	V2H14	I1H4	IT	I2H20	<b>OHDV3</b>	I2H7	I1H9	Q2	I2H6	KFactorV1I1		V2H10
VLT	MinI3	AvrI2	V3H15	<b>P1</b>	V1H12	<b>ST</b>	V3H21	I3H9	V3H5	I1H20	V1H22	V2H6	I3H14
I2H13	I2H19	V3H5	MinV1	S3	EHDI2	THDV3	V3H13	V2H3	EHDI2	Q3	IT	<b>V3H21</b>	CosPhi3
<b>V3H21</b>	VL23	I1H21	I2H15	MaxI1	VL23	CosPhi2	Freq	V3H7	V1H14	I2H14	CosPhi3	Q3	AvrI2

Note: Features identified by both LIME and SHAP are emphasized in bold.

Similarly, we could analyze other fault types for a global explanation using the SHAP explainer. Since LIME only provides a local explanation, we applied the algorithm to individual samples of each fault, repeatedly, then, gathered the recurring features. Hence, we can identify the features that contribute most to each fault. Table 5 lists the high-scored general and harmonic parameters. These parameters are sorted based on their importance. The general deduction from this table is that odd harmonics of order 3, 11, 13, and 17, as well as general parameters including phase currents, power factors, and harmonic distortion, have a significant impact on our model's predictions. In fact, rectifiers cause odd harmonics in the power line and a 12-pulse rectifier can extend these harmonics to a higher order while eliminating the 5th and 7th orders. Moreover, the primary Delta winding of the transformer remarkably attenuates triplen harmonics. Thus, we expect that any abnormalities may result in variations of other odd harmonics. This can be clearly noticed from the explanations gathered in Table 5 as voltage and current harmonics of orders 13 and 11 are mostly repeated in the results. Having discussed the global explanations, a domain expert could investigate the reasons behind DNN's decisions using SHAP/LIME local explanation and gain confidence by comparing them with the findings in Table 5.

**Table 5.** Global explanations for different fault types.

Class No.	Class Name	General Parameters	Harmonic Parameters
1	Capacitor	P2, I1, CosPhi1, CosPhi2, CosPhi3, CosPhiT	V2H13, V3H11, V1H6, V3H3, V1H3
2	Control System	I2, I1, P1, P2, KFactorV1, CosPhi2	V3H13, I2H15, V3H3, V1H3
3	COS	CosPhi3, CosPhi1, MaxI1, VL23, VL31, I2	V1H19, V3H13

**Table 5.** Global explanations for different fault types.

Class No.	Class Name	General Parameters	Harmonic Parameters
4	Earth	I1, I2, P1, PT, I3, QT, THDV1, CosPhiT	V3H13, I3H13, V1H13, V3H21, V1H11, I1H11
5	Flywheel Diode	CosPhiT, CosPhi1, I1, I2, CosPhi3, P3, P2, CosPhi3, CosPhi2, CosPhiT, I1, MaxI2, IT,	V3H3, I3H17, V3H17, V3H5, I1H13, I3H13
6	Reactor	P1, QT	V3H13, V1H13, V1H3, I2H13, V3H3,
7	Thyristor	CosPhi3, THDV1, I2, ST, I1, P1, S2, CosPhiT, KFactorV1	V3H13, V1H13, V2H17, I1H13, V2H13
8	Temperature	CosPhiT, P1, I1, OHDI2, EHDV3, THDV1, OHDI3, CosPhi3	V1H13, V3H3, V1H3, V3H13, V3H11, V1H7, V3H5

## 7. Conclusions

This work develops an efficient data-driven framework for diagnosing faults in induction melting furnaces. The framework uses real-time electrical parameters from power quality and energy analyzers on the supply side. We adopted a DNN classifier for early fault detection due to its effectiveness in real-time fault diagnosis, particularly in the critical and high-stakes environments of induction furnaces in foundries. The DNN was trained in a supervised manner, and its predictive capabilities were tested with experimental data. The model's performance was compared with other competitors, demonstrating that the proposed DNN model outperforms the ensemble learners, MLP, and classical ML methods in terms of precision, recall, accuracy, and F-measure. Finally, post hoc explanations of the model's predictions were presented to the user in terms of the relevance score. Model-agnostic LIME and SHAP techniques were employed to generate the explanations. The contribution of features to each fault class was extracted using global explanations obtained from the SHAP and LIME plots. Based on the explanation, one can find that classified faults are commonly associated with variations in odd voltage/current harmonics of orders 3, 11, 13, and 17. The main constraint of the work is the skewed class distribution, which forces models to be biased toward the majority class and can increase the error in the prediction of the minority class samples. A further constraint is the difficulty of interpreting explanations provided by the LIME algorithm, which requires advanced knowledge of the system.

**Author Contributions:** Conceptualization, S.M. and R.R.-F.; methodology, S.M. and R.R.-F.; software, S.M.; validation, S.M. and R.R.-F.; formal analysis, S.M. and R.R.-F.; investigation, S.M.; resources, S.M.; data curation, S.M.; writing—original draft preparation, S.M.; writing—review and editing, S.M., R.R.-F., and V.P.; visualization, S.M.; supervision, R.R.-F. and M.S.; project administration, R.R.-F., V.P., and M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are available upon reasonable request from the corresponding author, Sajad Moosavi. The data are not publicly available due to privacy considerations.

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

## Abbreviations

The following abbreviations are used in this manuscript:

AHU	air handling unit
AI	artificial intelligence
ANN	artificial neural network
ARL	association rule learning
BN	Bayesian network
CBR	case-based reasoning
CFS	correlation-based feature selection
CNN	convolutional neural network

COS	changeover switch
CWConv	continuous wavelet convolutional
DCNN	deep convolutional neural network
DIFFI	depth-based isolation forest feature importance
DNN	deep neural network
DT	decision tree
EHD	even harmonic distortion
FG-CAM	frequency-domain-based gradient-weighted class activation map
FLS	fuzzy logic system
GLD	ground leak detector
IF	induction furnace
KNN	K-nearest neighbor
LGB	light gradient boosting
LIME	local interpretable model-agnostic explanations
LOF	local outlier factor
LR	linear regression
LRP	layer-wise relevance propagation
LSTM	long short-term memory
MLP	multilayer perceptron
NB	naïve Bayes
OHD	Odd harmonic distortion
PCC	Pearson correlation coefficient
RF	Random forest
RMSE	Root-mean-square error
RNN	Recurrent neural network
RUL	Remaining useful life
SHAP	Shapley additive explanations
SCR	Silicon-controlled rectifier
SVM	Support vector machine
THD	Total harmonic distortion
XAI	explainable artificial intelligence
XGB	extreme gradient boosting

## References

1. Tan, A.; Bayindir, K. Modeling and analysis of power quality problems caused by coreless induction melting furnace connected to distribution network. *Electr. Eng.* **2013**, *96*, 1–15. [\[CrossRef\]](#)
2. Barik, A. Good Maintenance Practice and Maintenance Hazards in a Foundry-Some Observations. December 2009. Available online: <https://www.researchgate.net/publication/322100996> (accessed on 24 April 2024).
3. Razavi-Far, R.; Kinnaert, M. Incremental Design of a Decision System for Residual Evaluation: A Wind Turbine Application. *IFAC Proc. Vol.* **2012**, *45*, 343–348. [\[CrossRef\]](#)
4. Zhang, C.; Zhao, S.; Yang, Z.; He, Y. A multi-fault diagnosis method for lithium-ion battery pack using curvilinear Manhattan distance evaluation and voltage difference analysis. *J. Energy Storage* **2023**, *67*, 107575. [\[CrossRef\]](#)
5. He, Z.; Zeng, Y.; Shao, H.; Hu, H.; Xu, X. Novel motor fault detection scheme based on one-class tensor hyperdisk. *Knowl.-Based Syst.* **2023**, *262*, 110259. [\[CrossRef\]](#)
6. Farajzadeh-Zanjani, M.; Hallaji, E.; Razavi-Far, R.; Saif, M. Generative adversarial dimensionality reduction for diagnosing faults and attacks in cyber-physical systems. *Neurocomputing* **2021**, *440*, 101–110. [\[CrossRef\]](#)
7. Farajzadeh-Zanjani, M.; Razavi-Far, R.; Saif, M. A Critical Study on the Importance of Feature Extraction and Selection for Diagnosing Bearing Defects. In Proceedings of the IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS), Windsor, ON, Canada, 5–8 August 2018; pp. 803–808. [\[CrossRef\]](#)
8. Farajzadeh-Zanjani, M.; Hallaji, E.; Razavi-Far, R.; Saif, M. Generative-Adversarial Class-Imbalance Learning for Classifying Cyber-Attacks and Faults—A Cyber-Physical Power System. *IEEE Trans. Dependable Secur. Comput.* **2022**, *19*, 4068–4081. [\[CrossRef\]](#)
9. Li, Y.F.; Liu, J. A Bayesian Network Approach for Imbalanced Fault Detection in High Speed Rail Systems. In Proceedings of the 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), Seattle, WA, USA, 11–13 June 2018; pp. 1–7. [\[CrossRef\]](#)

10. Upasane, S.J.; Hagrais, H.; Anisi, M.H.; Savill, S.; Taylor, I.; Manousakis, K. A Big Bang-Big Crunch Type-2 Fuzzy Logic System for Explainable Predictive Maintenance. In Proceedings of the 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Luxembourg, 11–14 July 2021; pp. 1–8. [\[CrossRef\]](#)
11. Hasan, M.J.; Sohaib, M.; Kim, J.M. An Explainable AI-Based Fault Diagnosis Model for Bearings. *Sensors* **2021**, *21*, 4070. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Brito, L.C.; Susto, G.A.; Brito, J.N.; Duarte, M.A. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. *Mech. Syst. Signal Process.* **2022**, *163*, 108105. [\[CrossRef\]](#)
13. Kim, M.S.; Yun, J.P.; Park, P. An Explainable Convolutional Neural Network for Fault Diagnosis in Linear Motion Guide. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4036–4045. [\[CrossRef\]](#)
14. Kim, M.S.; Yun, J.P.; Park, P. An Explainable Neural Network for Fault Diagnosis With a Frequency Activation Map. *IEEE Access* **2021**, *9*, 98962–98972. [\[CrossRef\]](#)
15. Saeki, M.; Ogata, J.; Murakawa, M.; Ogawa, T. Visual explanation of neural network based rotation machinery anomaly detection system. In Proceedings of the 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), San Francisco, CA, USA, 17–20 June 2019; pp. 1–4. [\[CrossRef\]](#)
16. Chen, H.Y.; Lee, C.H. Vibration Signals Analysis by Explainable Artificial Intelligence (XAI) Approach: Application on Bearing Faults Diagnosis. *IEEE Access* **2020**, *8*, 134246–134256. [\[CrossRef\]](#)
17. Nor, A.K.M.; Pedapati, S.R.; Muhammad, M. Application of Explainable AI (Xai) For Anomaly Detection and Prognostic of Gas Turbines with Uncertainty Quantification. *Preprints* **2021**, 2021090034. [\[CrossRef\]](#)
18. Madhikermi, M.; Malhi, A.K.; Främling, K. Explainable Artificial Intelligence Based Heat Recycler Fault Detection in Air Handling Unit. In *Explainable, Transparent Autonomous Agents and Multi-Agent Systems*; Calvaresi, D.; Najjar, A.; Schumacher, M.; Främling, K., Eds.; Springer: Cham, Switzerland, 2019; pp. 110–125.
19. Grezmak, J.; Wang, P.; Sun, C.; Gao, R.X. Explainable Convolutional Neural Network for Gearbox Fault Diagnosis. *Procedia CIRP* **2019**, *80*, 476–481. [\[CrossRef\]](#)
20. Grezmak, J.; Zhang, J.; Wang, P.; Loparo, K.A.; Gao, R.X. Interpretable Convolutional Neural Network Through Layer-wise Relevance Propagation for Machine Fault Diagnosis. *IEEE Sens. J.* **2020**, *20*, 3172–3181. [\[CrossRef\]](#)
21. Amarasinghe, K.; Kenney, K.; Manic, M. Toward Explainable Deep Neural Network Based Anomaly Detection. In Proceedings of the 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 4–6 July 2018; pp. 311–317. [\[CrossRef\]](#)
22. Keleko, A.T.; Kamsu-Foguem, B.; Ngouna, R.H.; Tongne, A. Health condition monitoring of a complex hydraulic system using Deep Neural Network and DeepSHAP explainable XAI. *Adv. Eng. Softw.* **2023**, *175*, 103339. [\[CrossRef\]](#)
23. Srinivasan, S.; Arjunan, P.; Jin, B.; Sangiovanni-Vincentelli, A.L.; Sultan, Z.; Poolla, K. Explainable AI for Chiller Fault-Detection Systems: Gaining Human Trust. *Computer* **2021**, *54*, 60–68. [\[CrossRef\]](#)
24. Protopapadakis, G.; Apostolidis, A.; Kalfas, A.I. Explainable and Interpretable AI-Assisted Remaining Useful Life Estimation for Aeroengines; Volume 2: Coal, Biomass, Hydrogen, and Alternative Fuels; Controls, Diagnostics, and Instrumentation; Steam Turbine. In Proceedings of the Turbo Expo: Power for Land, Sea, and Air 2022, Rotterdam, The Netherlands, 13–17 June 2022. [\[CrossRef\]](#)
25. Hong, C.W.; Lee, C.; Lee, K.; Ko, M.S.; Hur, K. Explainable Artificial Intelligence for the Remaining Useful Life Prognosis of the Turbofan Engines. In Proceedings of the 2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII), Kaohsiung, Taiwan, 21–23 August 2020; pp. 144–147. [\[CrossRef\]](#)
26. Abid, F.B.; Sallem, M.; Braham, A. Robust Interpretable Deep Learning for Intelligent Fault Diagnosis of Induction Motors. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 3506–3515. [\[CrossRef\]](#)
27. Li, T.; Zhao, Z.; Sun, C.; Cheng, L.; Chen, X.; Yan, R.; Gao, R.X. WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *52*, 2302–2312. [\[CrossRef\]](#)
28. Choi, Y.; Kwun, H.; Kim, D.; Lee, E.; Bae, H. Method of Predictive Maintenance for Induction Furnace Based on Neural Network. In Proceedings of the 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), Busan, Republic of Korea, 19–22 February 2020; pp. 609–612. [\[CrossRef\]](#)
29. Choi, Y.; Kwun, H.; Kim, D.; Lee, E.; Bae, H. Residual Life Prediction for Induction Furnace by Sequential Encoder with s-Convolutional LSTM. *Processes* **2021**, *9*, 1121. [\[CrossRef\]](#)
30. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *arXiv* **2016**, arXiv:1602.04938.
31. Lundberg, S.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**, arXiv:1705.07874.

**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.