**Deep Learning and Explainable AI Techniques for detection and diagnosis of faults in control systems**

**Abstract**

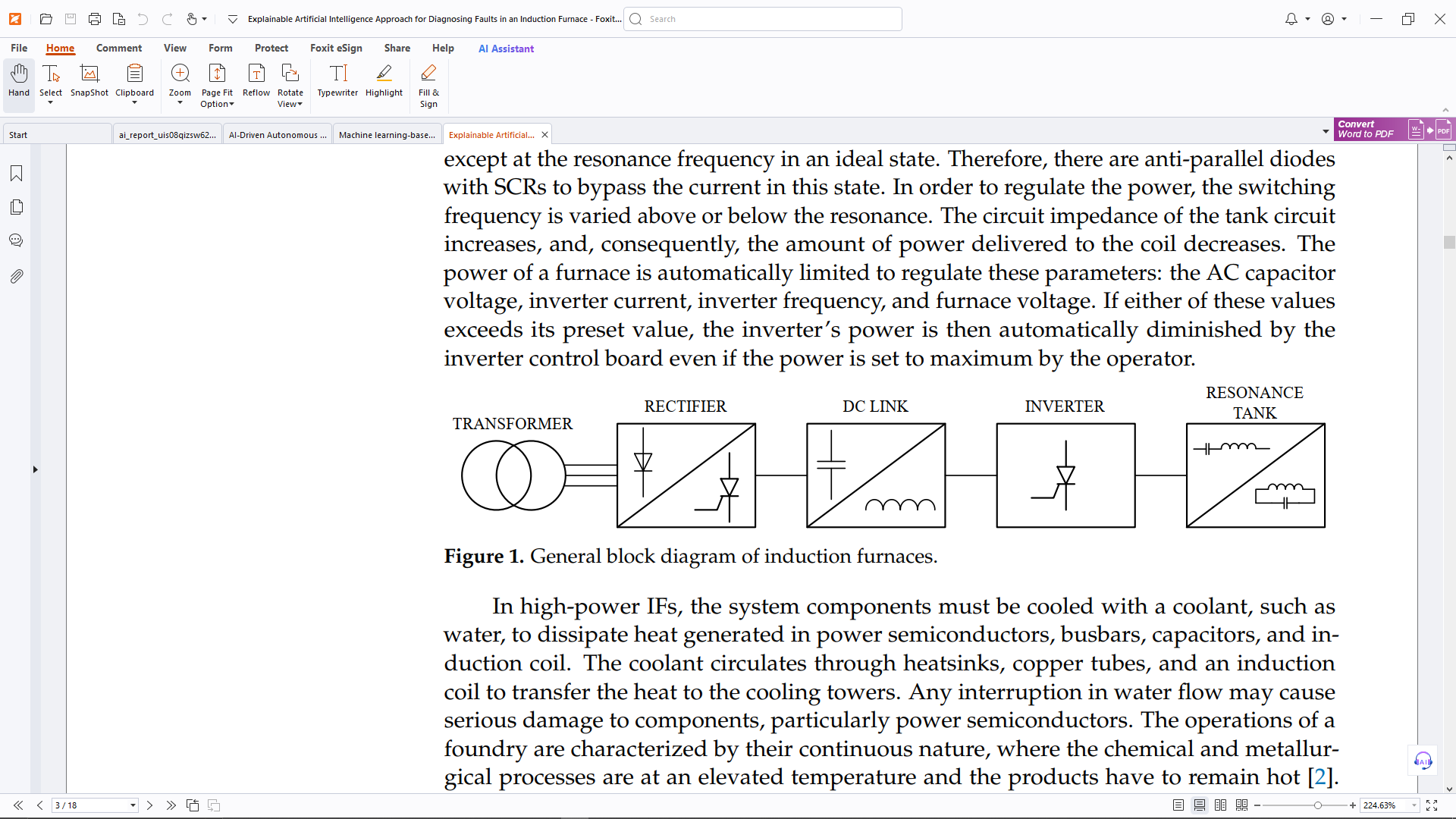
This article introduces a previously unexplored data control approach to diagnosing inductive stove problems. Induction stoves have been used in foundries for over a century to heat and melt metal. This allows for high melting and heating speeds with optimal efficiency. However, unplanned shutdowns and errors can interfere with production and pose security risks. The proposed architecture of deep neural networks continuously monitors supply-side electrical parameters to identify electrical errors in real-time. To collect sensory and experimental data, Foundry uses a variety of devices for its energy analysers. The data samples are then marked using half-surveillance learning technology known as the local outlier factor to distinguish between normal and defective authorities. The marked data is used to train deep neural networks. The performance of the developed model is evaluated using several metrics in several advanced techniques. The results show that the deep neural network model exceeds the other classifiers and achieves an average F-measure of 0.9187. Considering the fact that neural networks act as black boxes, predictions are interpreted by Shapley Additive's explanation and locally interpretable models-logical explanations. Interpretability analysis shows that odd voltage/electric harmonic anomalies in orders 3, 11, 13, and 17 are strongly related to the identified errors, highlighting the important role of these parameters in model prediction.

**Keywords**: Artificial intelligence (AI), control systems, diagnose faults, Deep Learning

**Introduction**

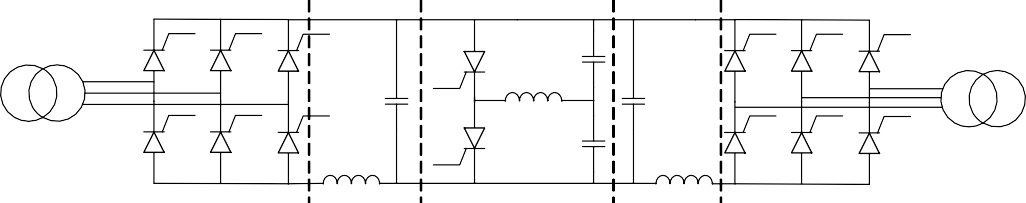
Electrically conductive materials experience vertebral current losses when exposed to fluctuating magnetic fields, which is the principle behind the operation of induction stoves. Induction stoves (IFs) are known for their high efficiency, cleanliness, accurate control and high-speed heat generation capabilities. As a result, they are often used to heat and melt, weld or harden metals in industrial processes. As shown in Figure 1 [1], the power supply consists of four main components, it is a rectifier, DC member, inverter, and resonant tank. The inverter can be designed to deliver resonant circuits continuously or in parallel and act as a voltage source or power source. As shown in Figure 2, the voltage source types investigated in this study use a series resonance inverter design. The 6-phase fully controlled rectifier consists of a silicon-controlled rectifier (SCR) that facilitates AC-to-DC conversion. The SCR is used to isolate alternating current lines when circuit questions arise, according to a rectifier. The rectifier creates a specific DC voltage. The current limit reactor is connected to the inverter to mitigate lash flow during short circuits or SCR errors (Kato et al., 2016). Recent research on direct power smoothing control methods still focuses on pitch angle control and rotor kinetic energy control (Zhu et al., 2021; Zhu et al., 2022), then the utilisation of DC-link capacitor deserves to be researched comprehensively to improve the power smoothing effect. The initial voltage wave from the rectifier is smoothed by a DC capacitor, which also serves as a voltage source for series resonances. With series changes, the stove coil creates a resonant group with variable resonance frequencies that vary based on the load state of the competing coil according to the inductance (Waluyo et al., 2022).

In a series of resonant load operations, the phase difference between the output voltage and electricity is usually observed, except under ideal conditions of resonant frequencies. In this regard, parallel diodes with silicon-controlled rectifiers (SCRs) are paired to manage current flow. Power control is achieved by adjusting the switching frequency above or below the resonance point. Increased circuit resistance in the tank circuit reduces the current supplied to the coil. The electricity supplied to the oven is automatically limited to maintain control of alternating current capacitor voltages, inverter, inverter frequency, and oven voltage. The inverter control plate should reduce power if any of these parameters exceed the specified limit, regardless of the user's maximum performance settings.



**Fig.1: General block diagram of induction furnaces**

In high-performance intermediate frequency systems (IFs), it is important to use half of the electric waste, bath hair, capacitors, coolant, and usually water to direct the heat generated by the induction coil. This coolant circulates copper pipes, heat sinks and induction coils to transfer heat to the cooling tower. Interference in water flow can be caused by substantial damage to the components, especially the power semiconductors. The continuous nature of casting operations is characterized by the high temperatures involved in chemical and metallurgical processes and the need to maintain the heating of the product. Given the high costs associated with these operations, it is important that they are maintained continuously. Implementing a data-controlled diagnostic system can improve both productivity and security within the system. By being vigilant about potential failures and statement submissions, such a system can verify and possibly avoid unnecessary shutdowns if the defect is recognized. This clarity is extremely important for integrating artificial intelligence into key sectors such as production where decisions can have great consequences. This study introduced a deep learning-based method for diagnosing errors in IF systems. To make the results understandable and improve trustworthiness, the Artificial Intelligence Post-Hook Module (XAI) has been developed to further clarify the integrated black-box deep learning model.



Δ

Induction Coil

Y

Rectifier

DC Link

Inverter

DC Link

Rectifier

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

**Literature Review**

One of the most important uses of artificial intelligence (AI) in an industrial environment is the detection and diagnosis of defects as the state of machines and devices is extremely important to improve production and operational efficiency. Data-oriented diagnostic systems often use distinctive extraction and selection methods to facilitate problematic patterns and properties from a wide range of data records, improving the accuracy of error diagnosis and the ability to effectively identify potential problems. However, these methods do not provide an explanation of the effect of the properties or the process of selecting the diagnostic system.

With the growing demand for intelligent diagnostic tools, these tools need to be explained to ensure reliability. Explanatory AI (XAI) provides operator insight into the outcomes and helps in developing intelligent prediction algorithms. In particular, there was no research into the diagnosis of defects in existing literature. A lot of research has been conducted to tackle the challenges of Xai Tools in identifying defects in various systems. This review evaluates and compares several incentive systems for Xai applications. The study analyzes four extraction methods, case-based discussion (CBR), Association Rul Learning (ARL), Bayesian Networks (BN), and Neuro-Fuzzy systems, in relation to five criteria for error detection in high-speed rail systems. Bayesian networks were selected for this project based on the description of the developed model and the use of correlation-based feature selection (CFS) (Bounemeur et al. 2019; Bounemeur et al. 2021).

The condition of the water pump is evaluated to promote preventive maintenance [10]. The proposed approach uses a type 2 fuzzy logic system (FLS) to improve the interpretability of the model. With a total of 100 rules, including four anti-Zedesia, the system offers superior explanation and reliability compared to existing opaque models. [11] found a problem related to warehouses using the KNN classifier (the most neighbouring Knearest Neighbor). This methodology can be tailored and function as a generalised model for use with a variety of data records with different configurations. In [12], 11 different methods were evaluated 12], and unmanned anomaly detection of the rotating machine was performed, resulting in overall performance, leading to the selection of isolated forest technology. The importance of the properties was then determined using both model tag shaping and model-specific local depth-based forest properties (diff). Finally, the causal analysis identified the most important specific features. Defects in linear motion instructions were identified in [13]. After training of 1D folding networks (CNNs) over time ranges, the classification criteria for this study are presented by the frequency-based gradient-weighted class activation card (FG-CAM). Grad-CAM methodology aims to improve model understanding by tracking the learning process. The authors argue that the proposed technique can be used in a wide range of complex physical models. An equivalent method is explained in [14–16]. Gas turbine anomaly detection and prediction were performed using Bayesian Long Short-Term Memory (LSTM), and results are explained in Form [17]. To take into account data and parameter uncertainty, the prediction is extended with two starting layers:

The AU and EU layers reflect the model's confidence in that prediction.

Root Mean Square Error (RMSE) and initial predictions are calculated for performance evaluation. Additionally, two criteria (precise accuracy and consistency) are examined to evaluate the form description. [18] Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) can be used to recognise issues with heat recovery in airspace (AHUs).

The results of the model are provided by the findings provided by the Regional Interpretable Model Aggregation Declaration (LIME). The authors argue that user-supplied explanations should improve the reliability of the model. A deep folding neural network (DCNN) using LSRP technology (layer-white association propagation) has been developed for the diagnosis of gear errors. The explanatory method used by the DCNN classifier, LRP evaluates the effect of each input on the final result. This means that LRP improves the transparency of DCNN's decision-making process and encourages wider use in machine error diagnosis. In another study, Grezmak et al. CNN and LRP were used to record motion irregularities and assess the effectiveness of CNNs by time-frequency spectral images of vibration signals recorded by the induction engine. Amarasinghe et al. We proposed a framework for the recognition of deep neural networks (DNNS) and post-hoc explanations generated via the use of LRP. A comparable architecture has been proposed for deep ship integration of DNNs for interpretability to monitor the condition of hydraulic systems. Another study presents a model for the diagnosis and identification of artificial intelligence (XAI) cold secession (XAI) for the diagnosis and identification of chiller errors, using an extreme gradient boost (XGBoost) model connecting regression trees. The explanation provided by LIME is likely to aid in the early detection of errors, with the participation of human operators. Additionally, this information improves accuracy, reduces the time required to identify errors, and increases field personnel's confidence.

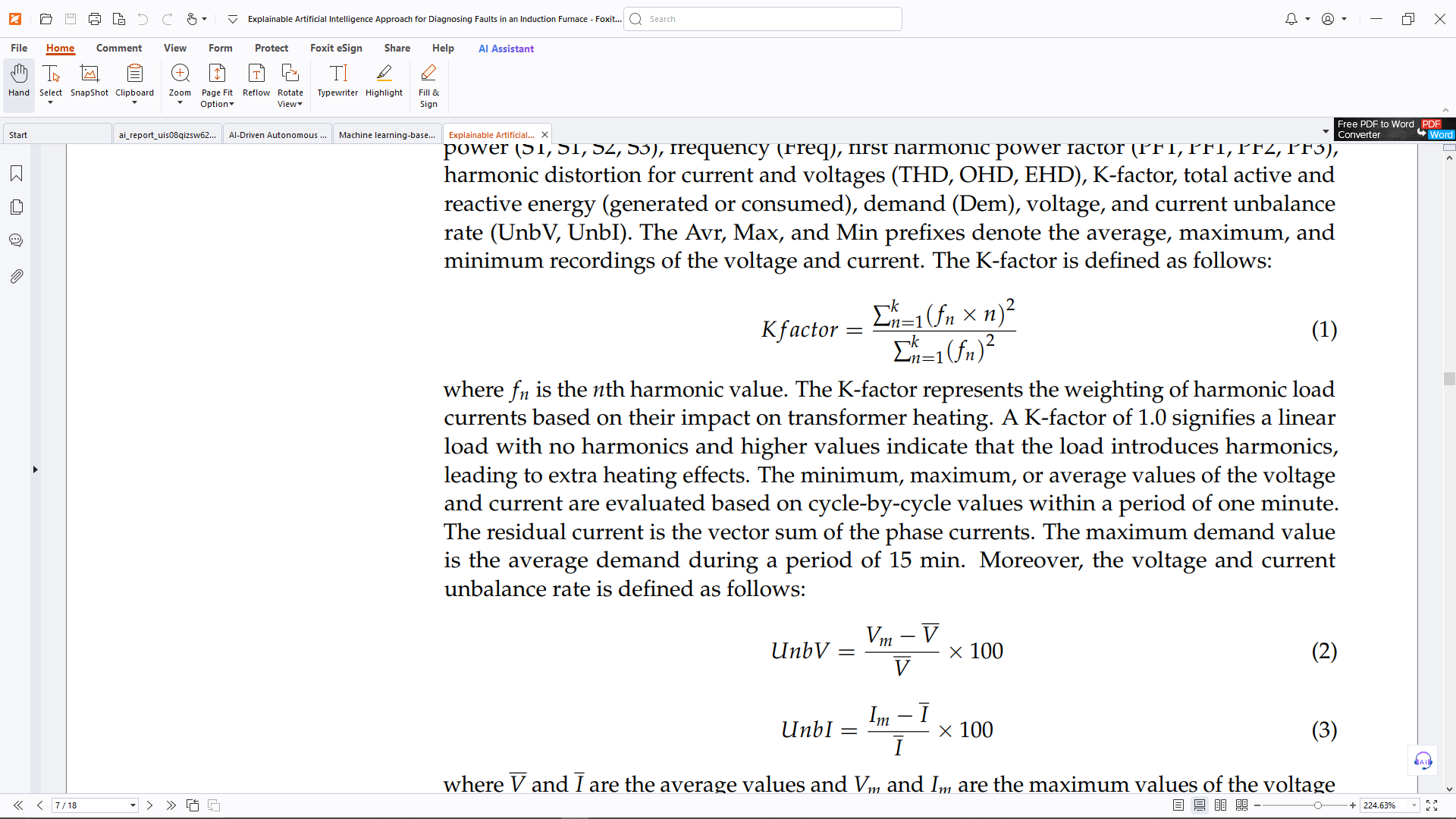
The estimated period required to repair or replace a machine is called the remaining service life (rules). [24] used a model with deep neural networks (DNNs) in addition to Lime to predict the rules of turbofan engines. The results show that the model effectively captures basic physics with fewer data records and provides a clear explanation of lime. However, when both the model and its explanations relate to increasingly complex data, we face challenges in providing appropriate results. [25] uses three algorithms - CNN, LSTM, and bidirectional LSTM- in the same context. SHAP examines the effects of all input variables and provides a clear, visual representation of how all attributes affect the expected outcome. Typically, traditional CNN filters are opaque and can contain unwanted, crazy spectral patterns. To meet this challenge, Sincnet was introduced in [26] to improve the interpretability of filter codes in the first layer of CNN. Compared to standard CNNs, the low-sink approach showed superior performance, improved interpretability, improved intoxication, and reduced implementation costs. Following the methodology of F.B Abid et al., T. Li et al. [27] The initial layer of CNN was replaced with a wavelet folding layer (CWCONV). This wavelet-controlled deep neural network, called WaveletKernelNet (WKN), contains important cores, and provides a 10% improvement in accuracy compared to traditional CNNs, allowing you to interpret the CWCONV layer. Furthermore, Yolim Choi et al. [28, 29] conducted two data-driven studies to assess instructors in IF-resistant rice. This demonstrates that multilayered perceptrons (MLPs) surpass Rulevage's recurrent neuron networks (RNNS) and LSTM technologies. The research uses important electrical characteristics such as input/output power, voltage, electricity, DC voltage, and frequency. In the second study, a new SSTM method for RUL prediction is proposed. S-ConvlStm MLP and LSTM have been shown to rise with regard to Pearson correlation coefficients (PCC) and RMSE metrics.

Investigate the time and frequency domain characteristics of the IF power supply. The system is then proven in three parts of the

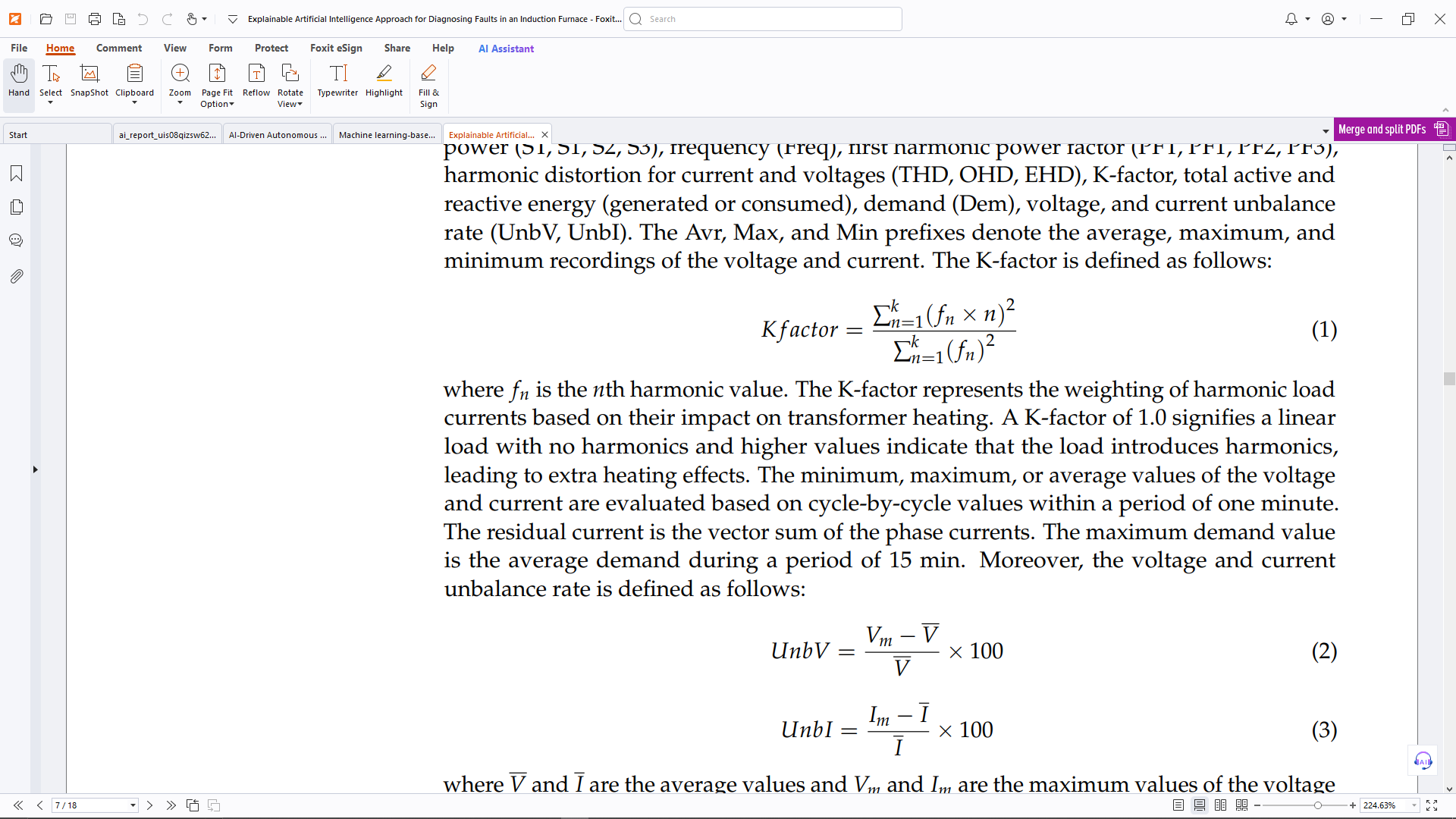
A DNN for classification of error types. A semi-salx anomaly detector that distinguishes false rehearsals of normal data samples. Using the XAI method. According to the author's best knowledge, this is the first study to represent Xai-based error detection in the use of real data samples from the industry.

**Methodology**

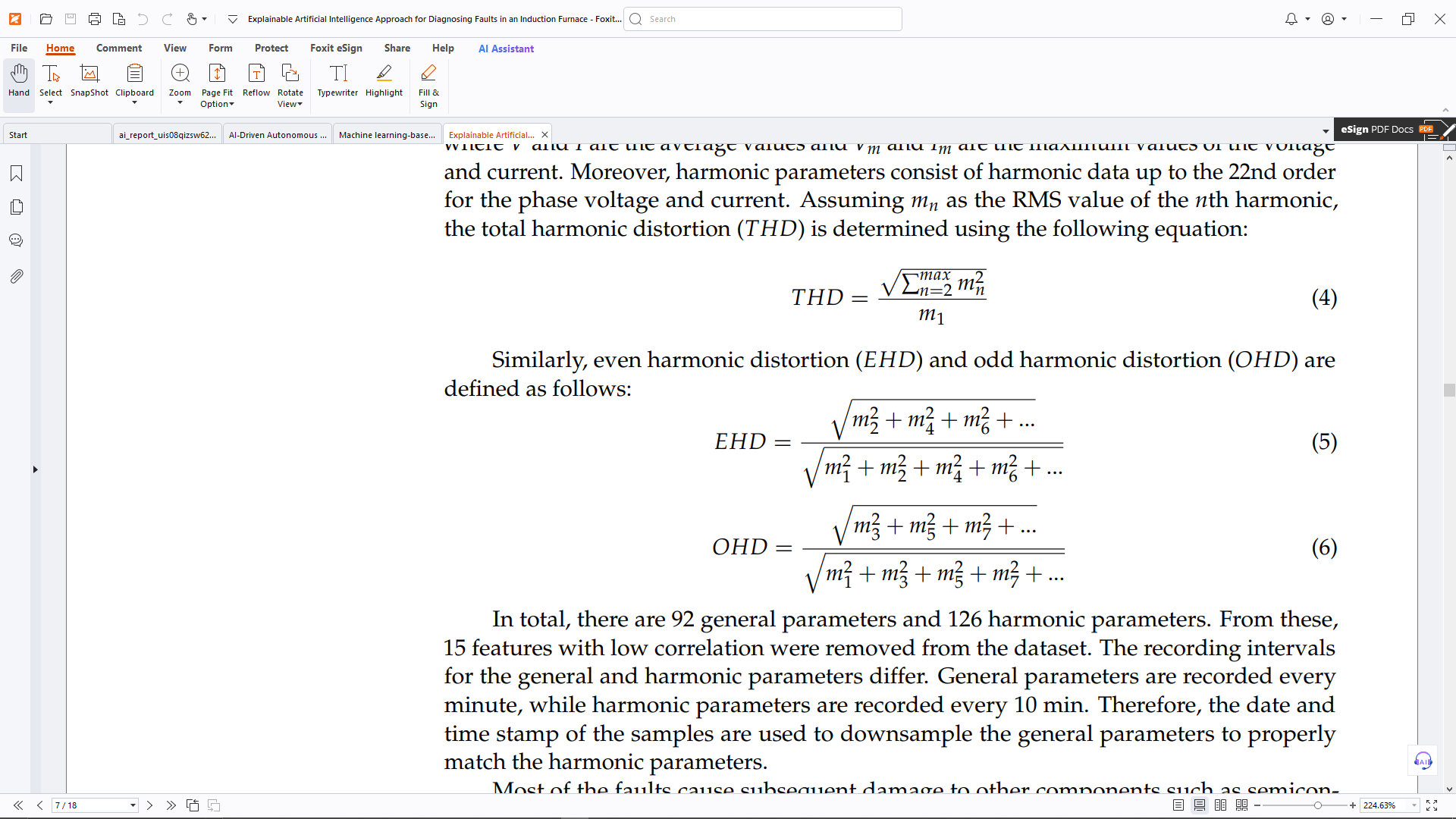
This work proposes an open and efficient method for diagnosing IF faults based only on real-time monitoring of electrical parameters collected by a power quality and energy analyzer equipment. The system setup is shown in Figure 3. Two sets of 218 parameters—harmonic and general parameters—are recorded by each energy analyzer. Each measurement is sent to a central data server for tracking and further analysis, where it may be accessed or exported. 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) are examples of general parameters. The voltage and current records' average, maximum, and minimum are indicated by the prefixes Avr, Max, and Min. The following is the definition of the K-factor:



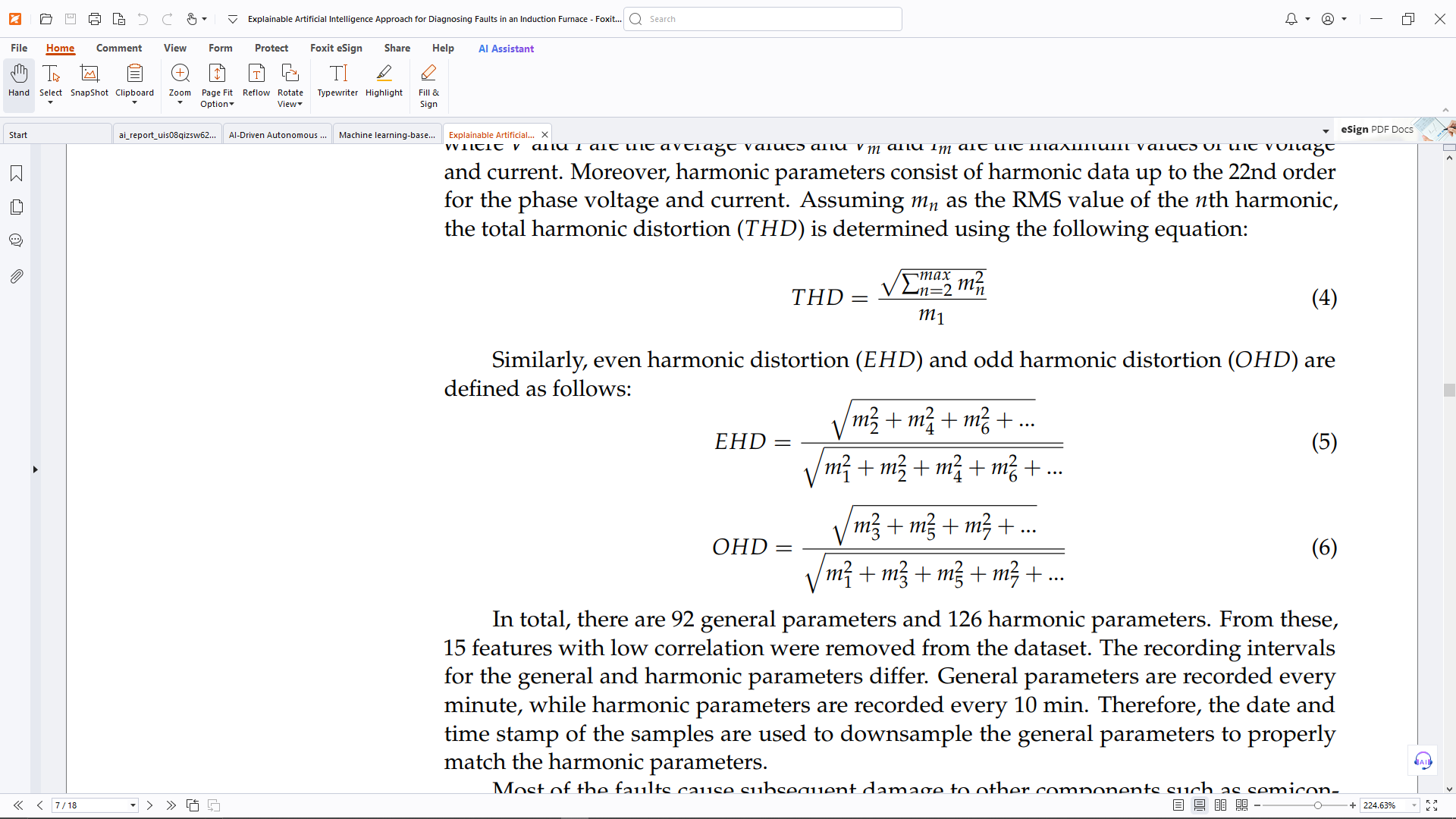
Where *fn* is the *n*th harmonic value. The K-factor represents the weighting of harmonic load currents based on their impact on transformer heating. A K-factor of 1.0 signifies a linear load with no harmonics and higher values indicate that the load introduces harmonics, leading to extra heating effects. 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:



Where *V* and *I* are the average values, and *Vm* and *Im* 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 *mn* as the RMS value of the *n*th harmonic, the total harmonic distortion (*THD*) is determined using the following equation:



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



It contains a total of 92 general parameters and 126 harmonious parameters. Fifteen properties with little correlation from the sample were removed. There are differences in recording intervals for general and harmonious properties. Harmonised parameters are recorded every 10 minutes, and general parameters are recorded every minute. The sample dates and timestamps are used to rehearse common parameters to satisfy the harmonious parameters accordingly.

control systems, capacitors, snubbers and semi-bearers are some of the other components that have been further damaged by most abnormalities. As a result, the main reason for the error is recorded as an error in the maintenance report. These criteria are used in this study to evaluate a total of 290 errors. A sample of the energy analyser data was then presented at each error event from the start of the oven cycle to the time of the error. When the inverter is turned on and the oven is equipped with cold scrap, the operator is turned off, and the melted material is removed, the oven cycle is taken into account. Data on oven production shows that this cycle of ovens is typically repeated 10 times a day for about 2 hours.

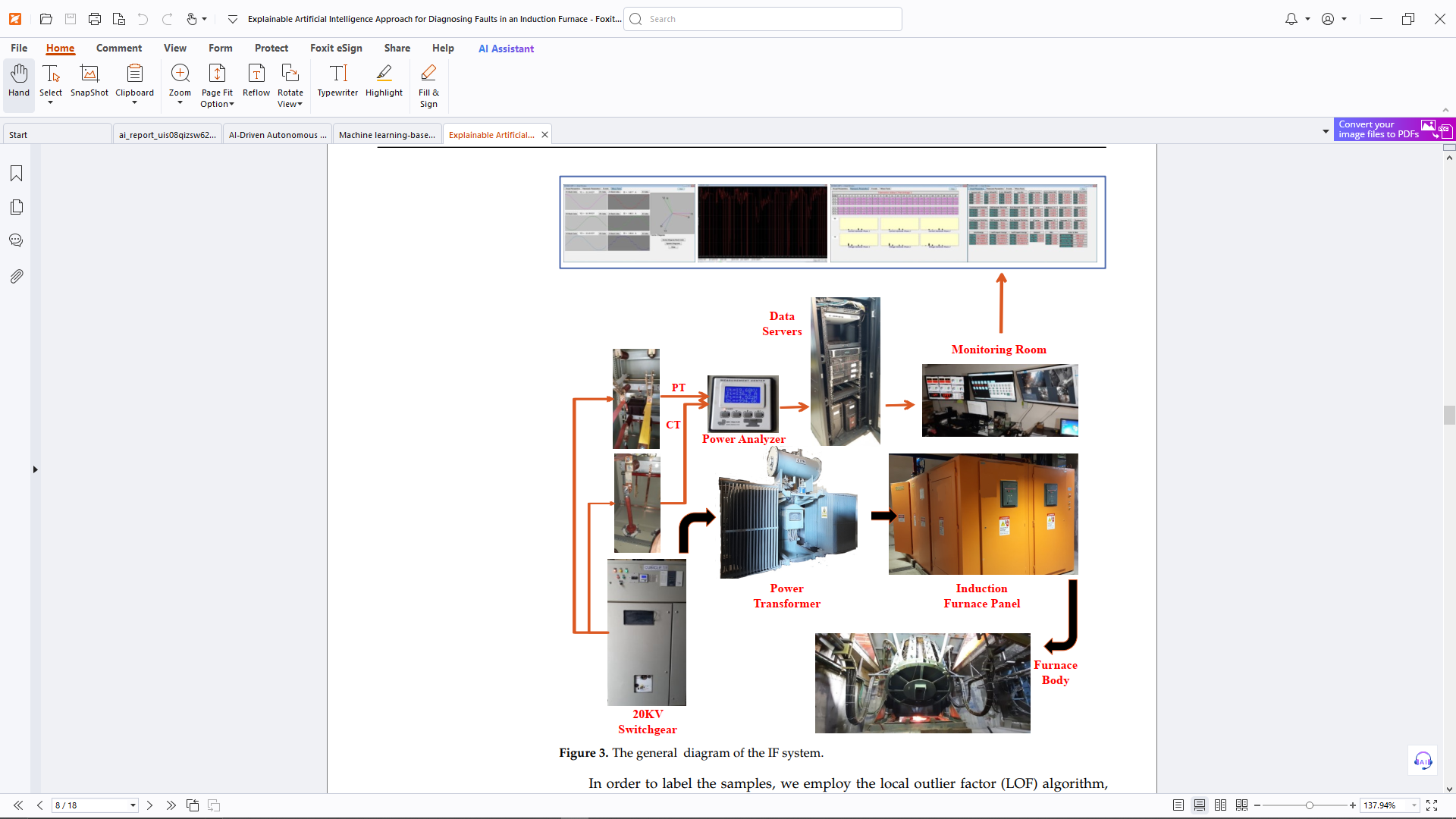
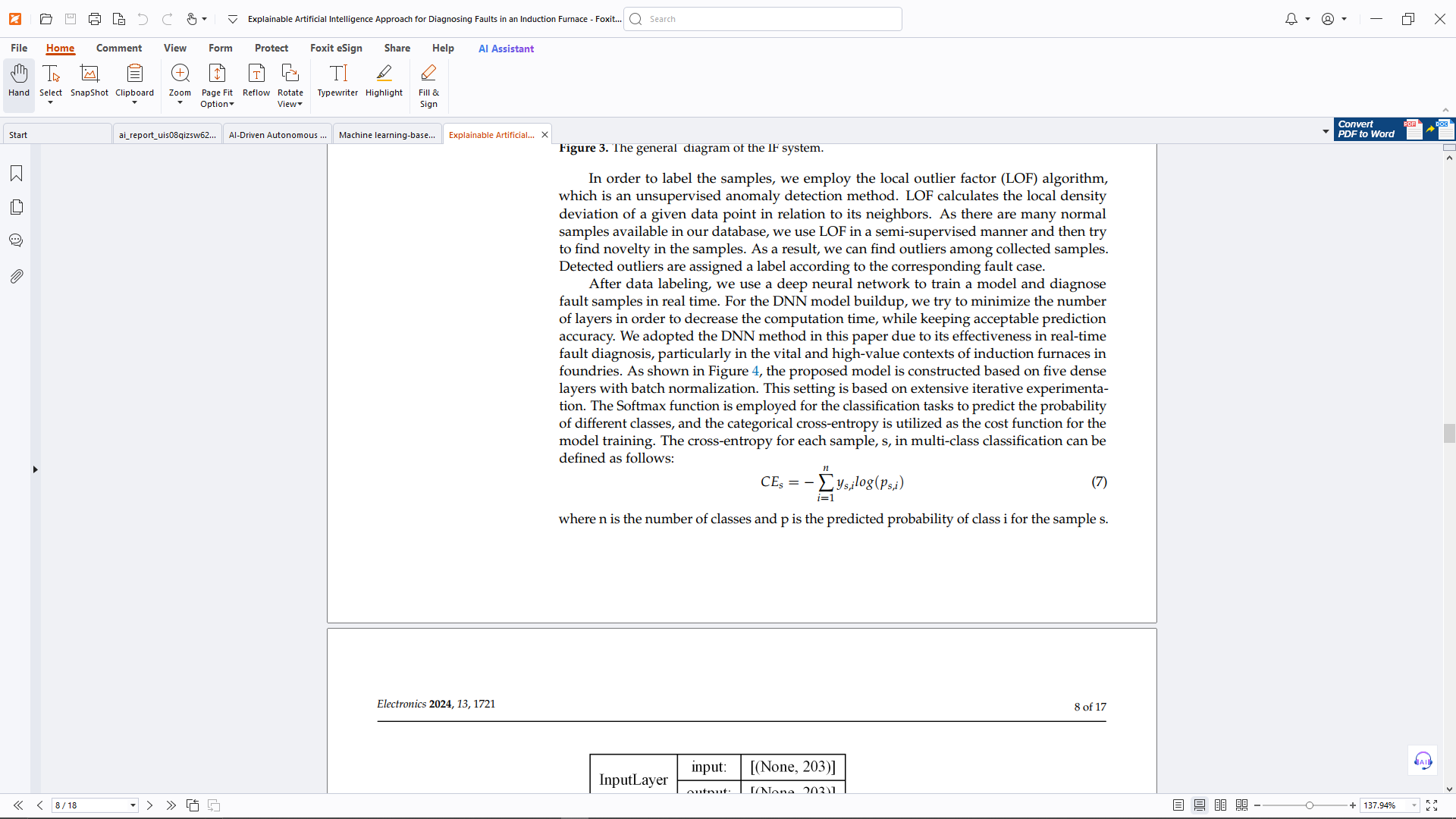


Fig.3 : The general diagram of IF system

To classify the data, we use an unmanned anomaly recognition method called the Local Outlier Coefficient (LOF) algorithm. LOF calculates the local density deviation of a data point compared to a neighbour. The database contains many regular samples, so use LOF in a half-trimmed way before trying to find uniqueness in the rehearsal. As a result, we can recognise abnormalities in the collected samples. The encountered exception is identified by the corresponding error. After classifying the data, the model is trained using deep neural networks to recognize failed samples in real-time. To save computational time and at the same time maintain a reasonable degree of predictability, try to reduce the number of layers in the DNN model development. This study used DNN technology in this study for its effectiveness in identifying real-time problems, particularly in the critical and critical environments of casting furnaces. Figure 4 shows how the proposed model is constructed using five thick layers with stacking. This attitude is based on extensive and repetitive exams. The SoftMax function is used for classification tasks to predict the possibilities of different classes, while the cross-entropy of categories is used as a cost function for model training. All rehearsal crosspieces and classifications containing some classes, can be defined as follows:



Where n is the number of classes and p is the predicted probability of class i for the samples.

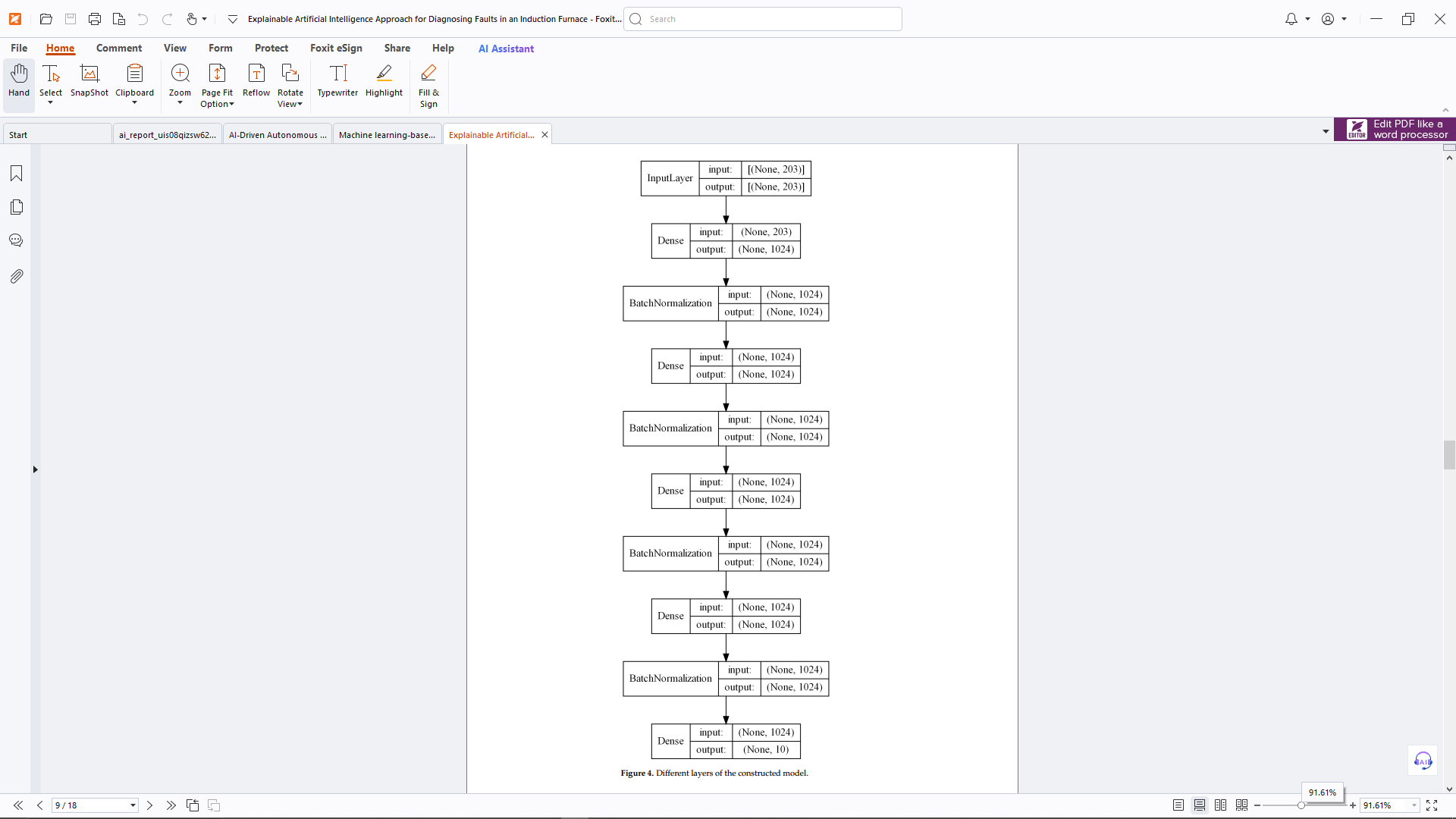


Fig. 4: Different layers of the constructed model

Interpretation of predictions in complex DNN models is of paramount importance. Descriptions allow users to understand and trust the predictions generated by the model. CAL Algorithm[30] is used to interpret the results of the proposed diagnostic model. Lime is a local approach that can explain the conditional interactions between characteristics and classes of a single sample. The algorithm follows the following steps at a high level:

* 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.

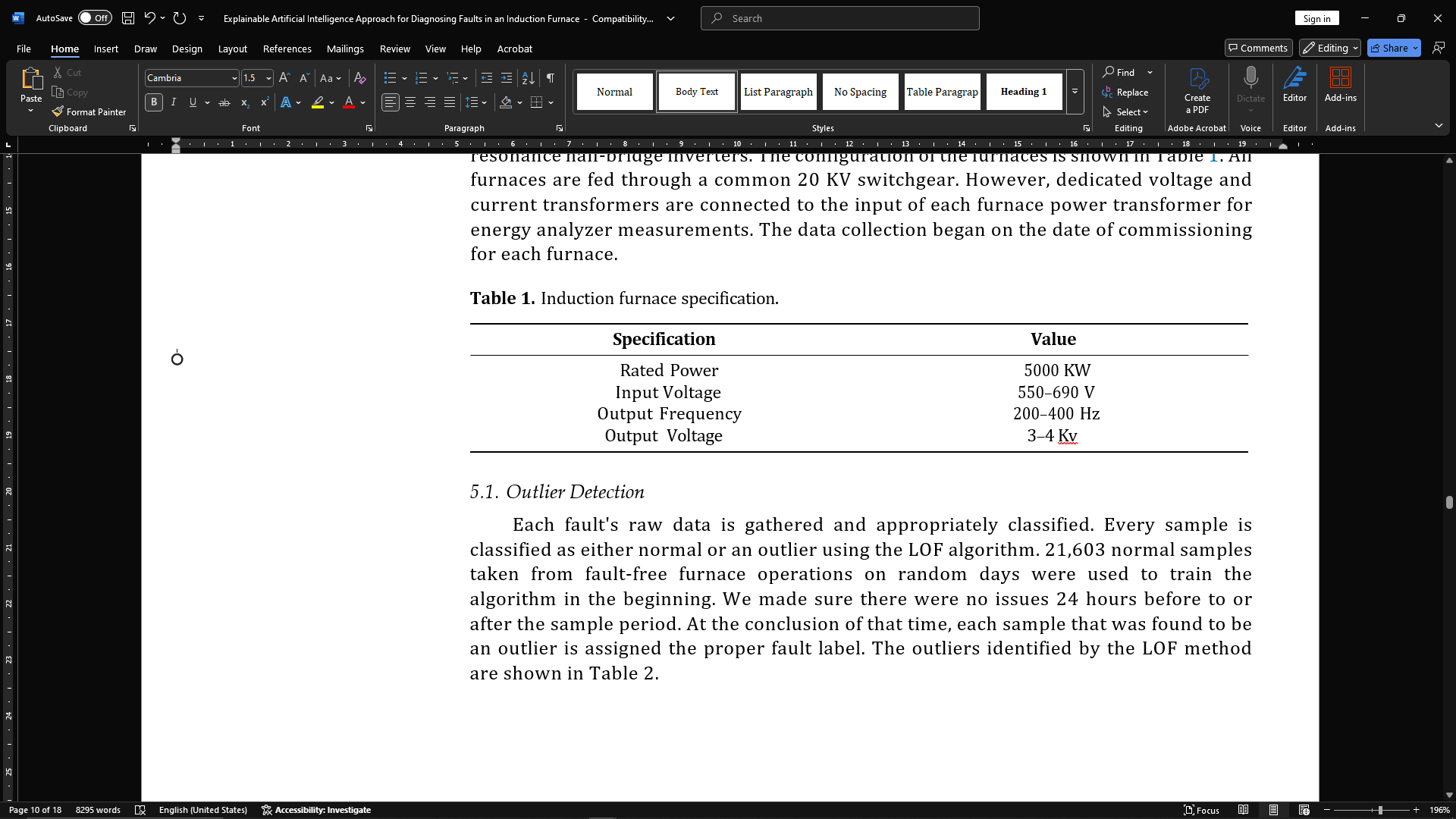
The explanation for the sample, s, is represented as follows:

*explanation*(*s*) = *argmin{L*( *f* , *g*, *πs*) + Ω(*s*)*}* (8)

Where g is a linear approximation of the original model f by sample S with local π. L is a function that measures the distance between F and G of a given neighbourhood. It shows how the prediction explanations of the original model correspond to. ω(s) is the concept of regularisation that measures the complexity of the explanation, namely the number of H. non-zero coefficients. The results of lime are validated by applying shapes [31]. Shap is a uniform framework based on Shapley values. SHAP can create local and global descriptions of models. In presenting a global explanation, we examine the importance of specific characteristics of model predictions.

**Results and Discussion**

A foundry running three identical 15-ton, 5 MW series-resonance half-bridge inverters provided the data. Table 1 displays the furnaces' arrangement. Every furnace receives its power from a single 20 KV switchgear. For energy analyser measurements, however, specific voltage and current transformers are linked to each furnace power transformer's input. On the day each furnace was commissioned, data collection got underway.



**Outlier Detection**

Raw data for each error is collected and categorised appropriately. Each sample is either normal or out of control using the LOF algorithm. 21,603 regular samples taken from error-free oven drainage on random days were first used to train the algorithm. We confirmed that there were no issues up to the sample time or sample time at 24 hours. At the end of this time, all samples found to be outliers are assigned the correct error label. The outliers identified according to the LOF method are listed in Table 2.

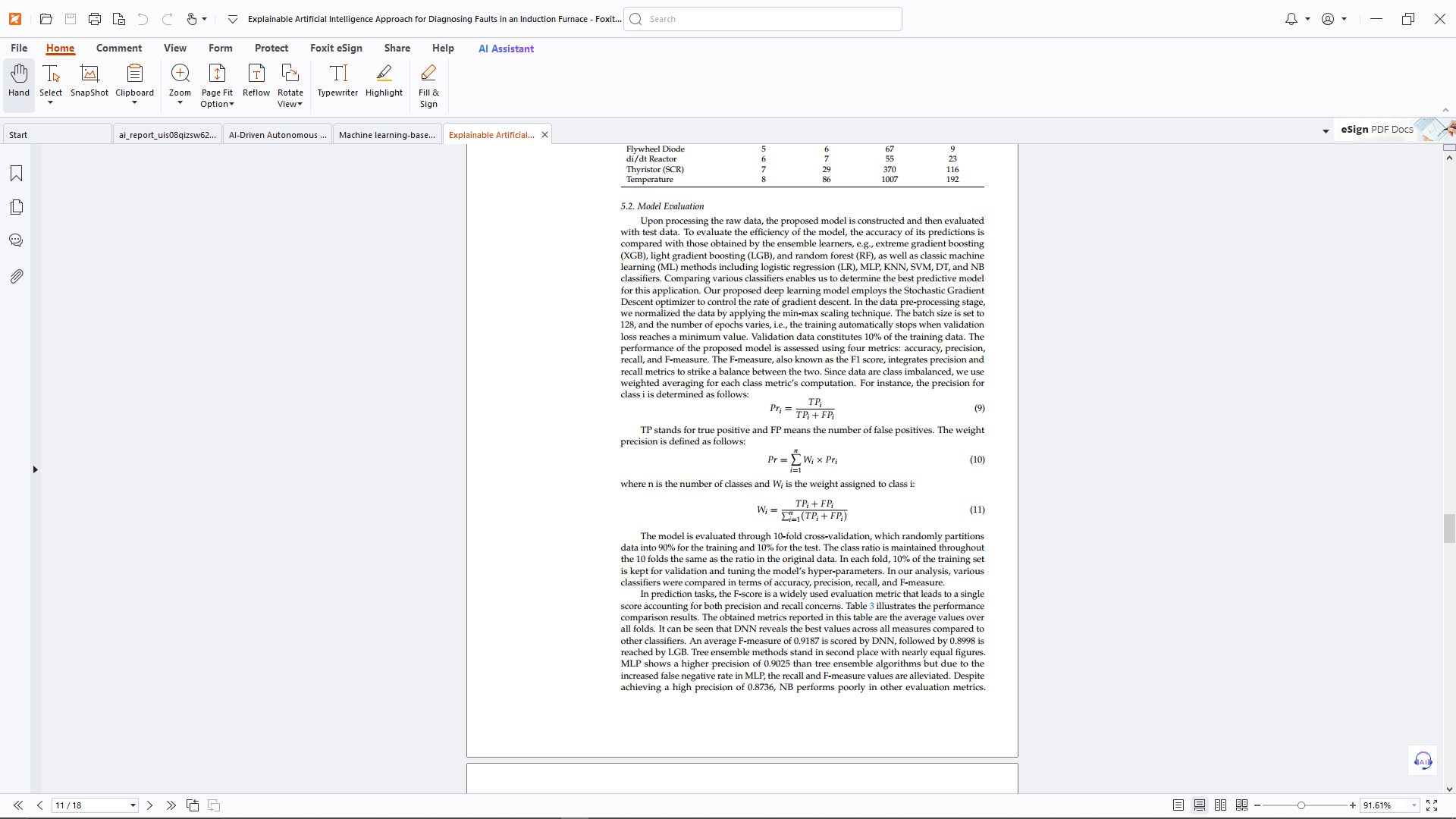
**Table 2. Fault statistics.**

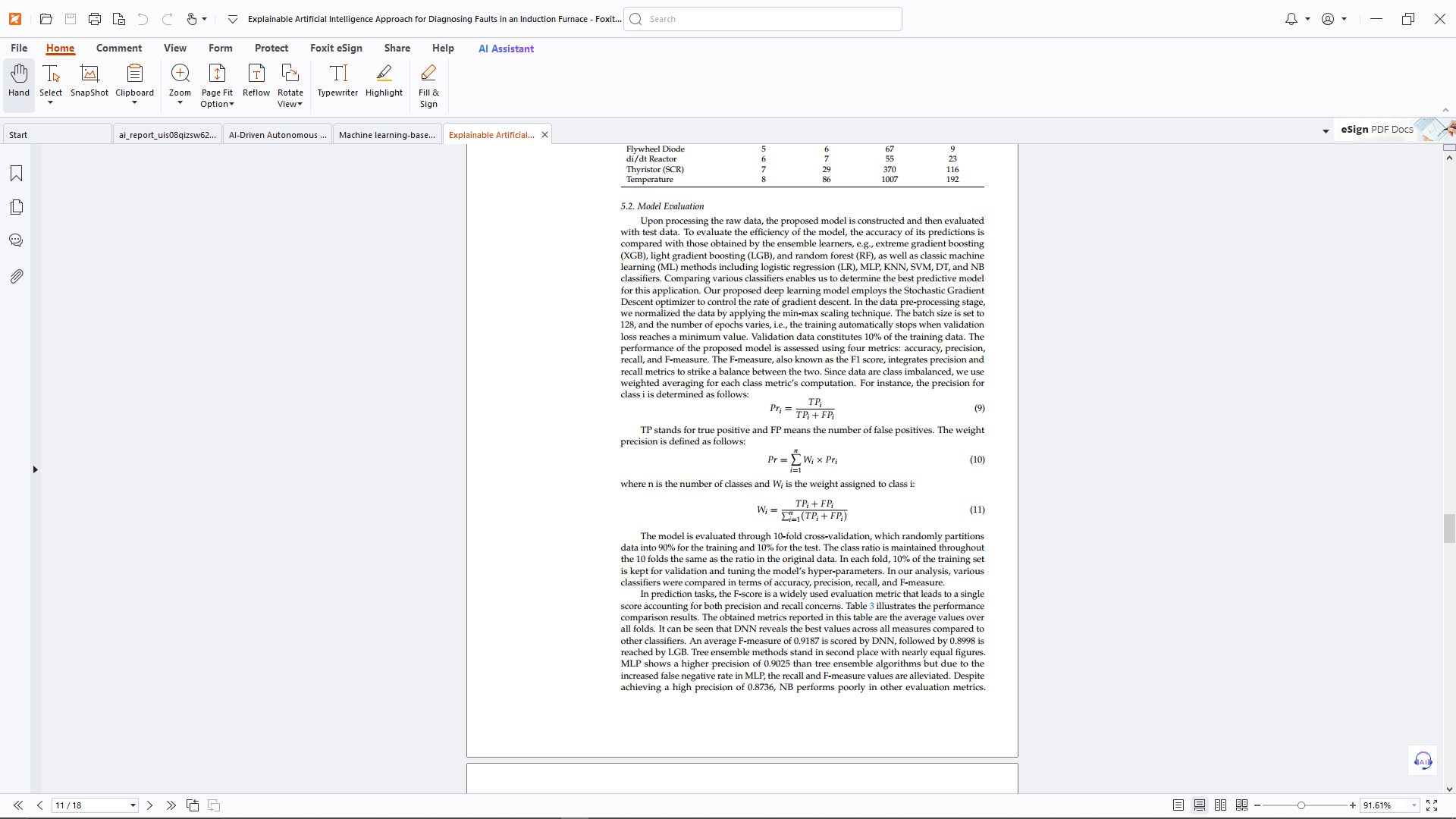
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fault Types Label** #**Fault** | | | **Total** | **#Outlier** |
|  |  | **Cases** | **Samples** | **Detected** |
| Capacitor | 1 | 37 | 487 | 108 |
| Control System | 2 | 24 | 333 | 82 |
| Change Over Switch (COS) | 3 | 13 | 181 | 19 |
| Earth Fault | 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 |

**Model Evaluation**

After processing the raw data and evaluating the test data, the proposed model is constructed. By examining the prediction accuracy of the model, we will introduce the prediction accuracy of the model (Ensemble (Random Forest (RF), Extreme Gradient Increase (XGB) and Optical Gradient Boost (LGB), and Predictive Accuracy (LR) of models such as traditional machine learning methods (ML), MLP, KNN, SVM, SVM, SVM, DT, NB -Klass -and nb Class-class popular MIN-MAX scaling technology. With a batch size and variability epoch of 128, training will automatically end when the loss of validation is reduced to a minimum. Validation data describes 10% of the training data. We use four metrics to evaluate the performance of the proposed model:

Accuracy, Accuracy, Recall, F Measurements. By integrating accuracy and memory measurements, F-Measurement-Aven as F1 points is known. The weighting of the data class processing allows you to calculate all class measurements using weighted averages. For example, the accuracy of class I is calculated by:





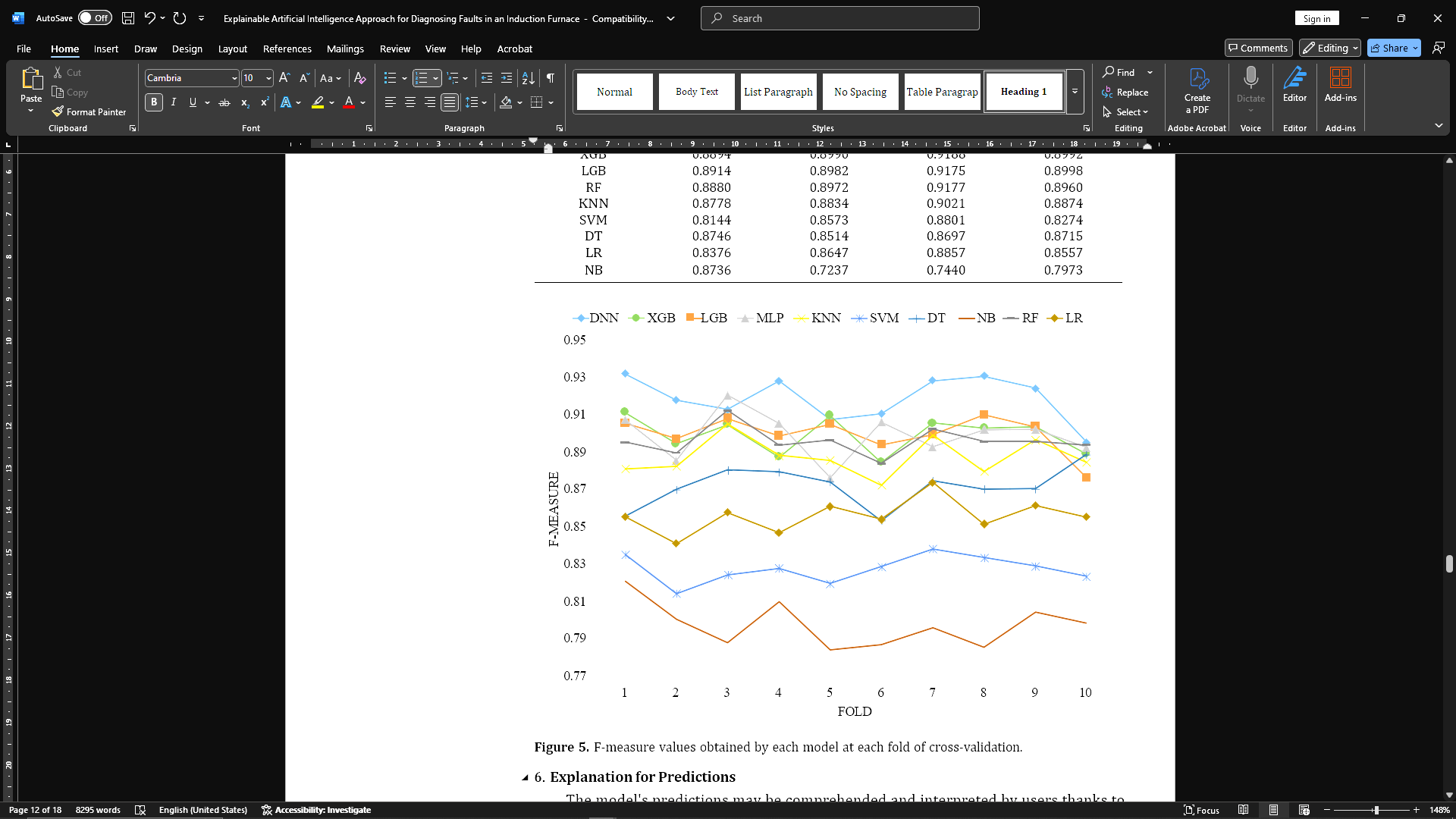
This model is evaluated using 10x cross-validation. This validation randomly divides the data into 90% for training and 10% for testing. The class ratio remains at the original data ratio of 10 times. Each fold holds 10% of the training set for validation and hyperparameter optimisation. In the study, many classifiers were evaluated according to F measurements, accuracy, accuracy, and recall.

F-scores are a broad evaluation metric for forecast tasks that generate a single point to be considered for both recall and accuracy concerns. The results of the performance comparison are shown in Table 3. The average values ​​for all surviving metrics are shown in this table. DNNS have the best performance over any scale compared to other classifiers. DNN receives a score of 0.9187 on average F measurements, while LGB receives a score of 0.8998. The second-placed tree ensemble technique had roughly the same numbers. The accuracy of the MLP is 0.9025 higher than that of the Tree Ensemble approach, but the recall and F measurements are lower due to the higher false negative rate. The NB accuracy is 0.8736, but it is insufficient compared to other evaluation criteria. The insufficient number of

error sediments and the unbalanced class distribution in the experimental data are the main reasons for the poor performance of the Bavarian classifier. The f-measurements for each model of each cross-validation fold are shown in Figure 5. The F-score metric confirms that the selected DNN model is at a minimum average of 15.22% above the Naive Bayes classifier.

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

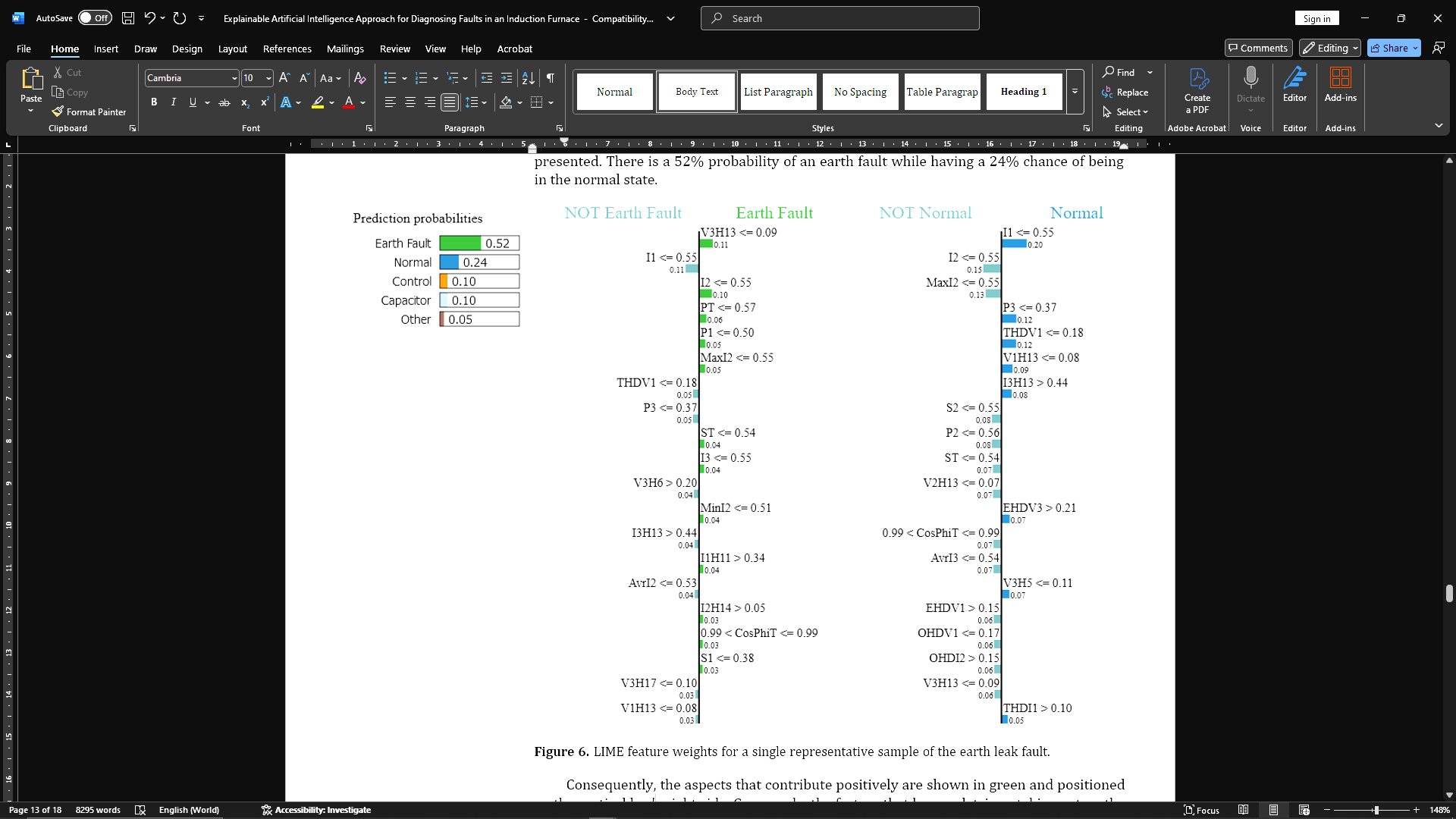
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Classifier** | **Precision** | **Accuracy** | **Recall** | **F-Measure** |
| DNN | 0.9127 | 0.9142 | 0.9287 | 0.9187 |
| MLP | 0.9025 | 0.8859 | 0.9007 | 0.8989 |
| XGB | 0.8894 | 0.8990 | 0.9188 | 0.8992 |
| LGB | 0.8914 | 0.8982 | 0.9175 | 0.8998 |
| 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 |



**Explanation for Predictions**

The descriptions of Lime and Shap help users understand and interpret the predictions of the model. Use both methods to explain the predictions of the model. It provides a different context using a randomly selected earth error test with a 52% chance. In our model, we use the deepshep [31] approach to calculate form values ​​to achieve faster and better performance. Both algorithms are configured to accurately deliver the 20 most important properties, as shown in Figures 6 and 7. Quality is listed in order of decreasing from top to bottom. In Figure 7, the horizontal axis shows the shape values and the vertical axis lists the features at the top. The two methods have 11 commonalities. Both figures show that the phase-III tension of order 13 (V3H13) has the most significant effect on the expected outcome.

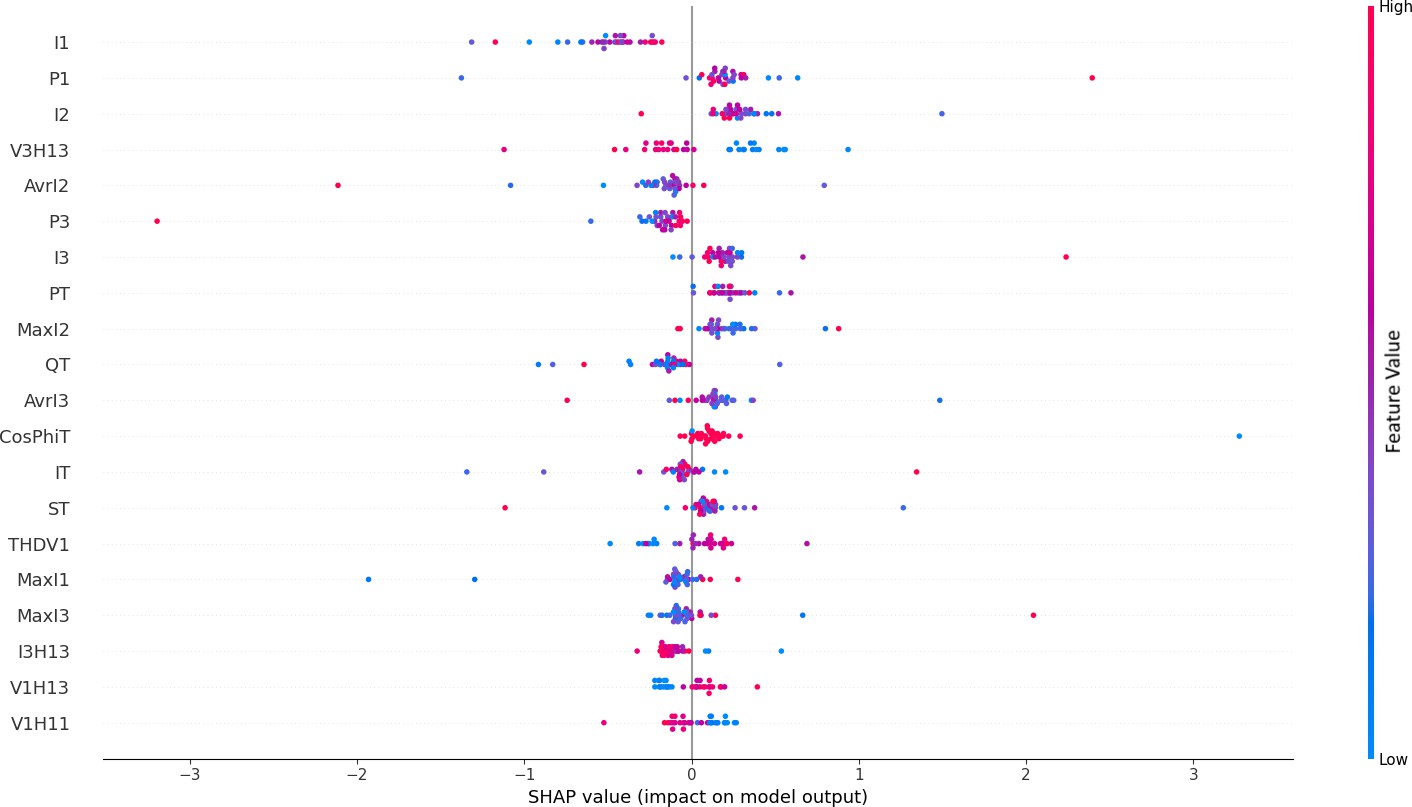
Predictive probabilities for various defects are displayed on the left side of the limestone diagram. Earth errors have a 52% chance of occurrence, and normal conditions can be 24%.



As a result, elements with positive effects are highlighted in green and placed on the right side of the vertical bar. On the other hand, the factors that negatively affect decisions are on the left side of the rod. The blue bars in the Shap diagram represent variables with negative form values ​​(such as I1 or Thdv1), while the red beam represent variables with positive form values ​​(such as V3H13, I2, PT). This diagram shows how to use the average of all predictions to obtain the predicted probability.

You can also provide a more thorough explanation by providing a set of accurate predictions from the same class. The combination of effects and characteristics relevant of the global summary graphics are shown in Figure 8. The vertical axis shows outstanding characteristics, while the horizontal axis shows shape values. Each row represents a property, and each point represents a sample. Red dots indicate high feature values, while blue dots indicate low. Properties are listed in order. By searching for various red and blue dots distributed along the X-axis, you can quickly determine the effect of a particular characteristic. The potential for form values ​​and earth error appears to be increased by the low values ​​of the 13th harmonic of the 11th harmonic (V1H11) of the current (I3H13) and phase-I voltage (V1H11), as well as the low values ​​of the overall performance factor (CoSPIT) and the 13th harmonic (CoSHP) and the high values ​​of the phase-I voltage (V1H13). For one of these samples in Figure 6, comparing this action with the LIME local declaration also reveals that similar features such as V3H13, I3H13, Cosphoit, and V1H13 contribute to the identification of earth errors.



**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.

To further confirm the description, we used post-hook explanators from additional models built. Sample errors were selected from the test set correctly predicted by each model. To be predictable, each model provides a different probability for this particular situation. Table 4 shows the top 20 attributes for each model evaluated with the associated probability after points. For example, DNN identified an error as an Earth error 81% of the time, while XGB showed more confidence 96% of the time. I used Treeshap for the XGB and DT classifiers. Features reported in both Shap and Lime are displayed by text printed in fat. It is clear that both explanations identify 55% of the same characteristics in both deep learning models and SVM scenarios.

It is also clear that at least five models emphasize certain elements rather than contributing to Earth's errors. All models report V3H13, and these frequent features include V3H13, V3H21, P2, I2, and Cosphi3. Therefore, it can be concluded that the 13th harmonic reduction of phase-III tension could be a reliable indicator of inter-earth mistakes.

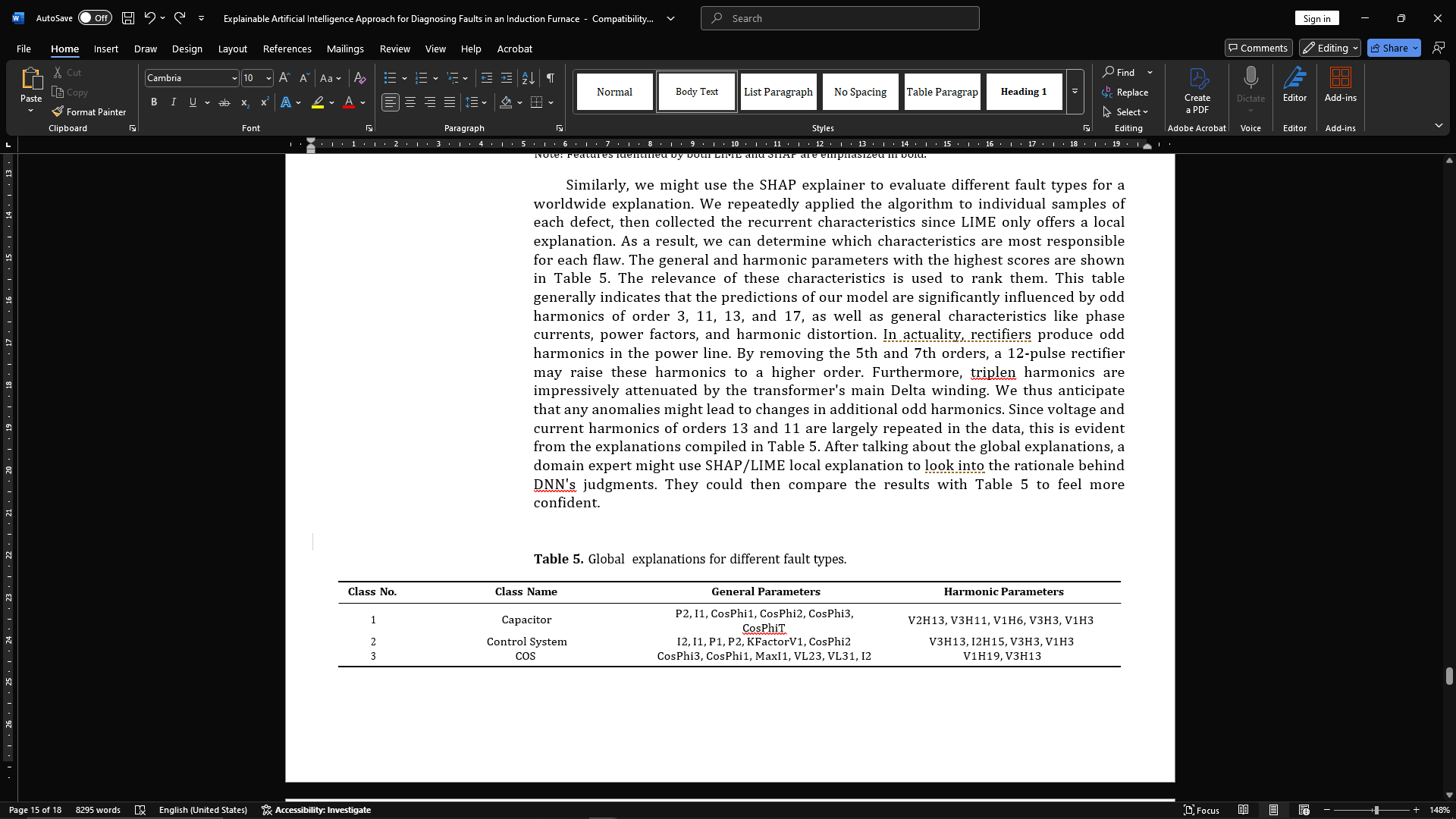
**Table 4.** Feature importance comparison of models W.R.T. individual fault samples.

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

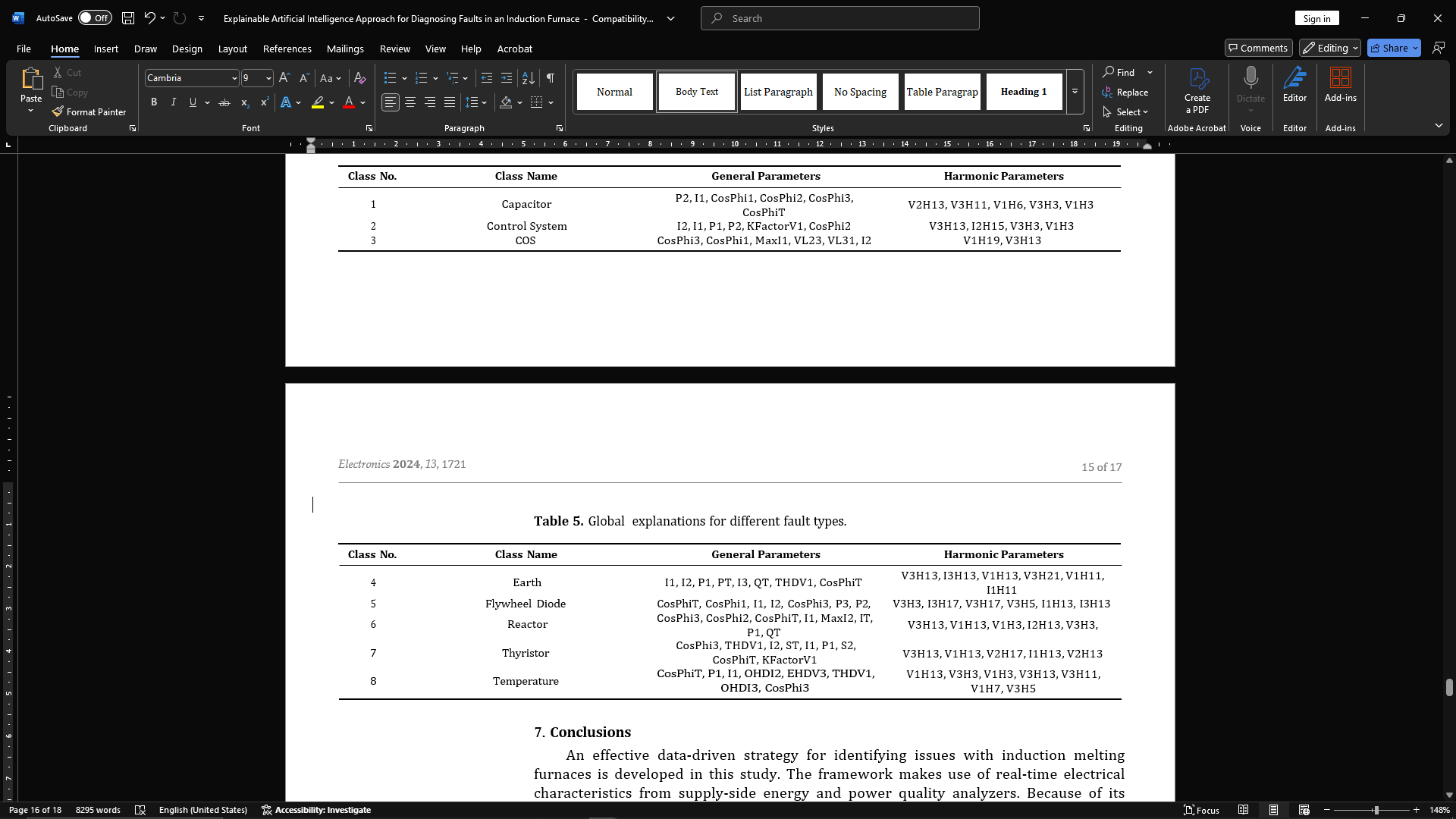
Note: Features identified by both LIME and SHAP are emphasized in bold.

You can also use Shap Declarer to evaluate different types of errors for global explanations. Since Lime only provides local descriptions, we repeatedly ran the algorithm for each error, before collecting repeated features. This allows you to identify the features created for all defects. Table 5 shows the general and harmonious parameters with the largest review. These attributes are classified according to their importance. This diagram shows that, as in the usual harmonious orders 3, 11, 13, and 17, general features such as phase flow, performance factors, and harmonized distortion have a major impact. In reality, the rectifier causes the power source to emit abnormal harmony. 12-pulse rectifiers can increase them to a higher order by eliminating the fifth and seventh orders. Furthermore, the main delta windings of the transformer significantly reduce the harmonious dray. Therefore, we expect anomalies could lead to other strange harmonious things. The explanations provided in Table 5 reveal that the data reiterate the harmoniousness of current instructions 13 and 11 primarily. Domain experts can use Shap/Lime's local description to explore the reasons behind the DNN conclusion after discussing the global description. Results can be compared to Table 5 to strengthen trust.

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



**Table 6.** Global explanations for different fault types.



**Conclusions**

This work develops an efficient data control approach to recognise the problem of inductive melted cubes. The real-time electrical properties of the analyser for energy and performance quality are used in the frame. In real-time error diagnosis, we selected the DNN classifier for early detection, especially in the critical and high operating environment of casting furnaces, for real-time disease diagnosis. After supervised training, the predictive ability of DNNs was assessed using experimental data. Comparisons of the proposed DNN model and the performance of the competitors show that they surpass ensemble learning, MLP, and traditional ML approaches in relation to accuracy, recall, accuracy, and F measurements. Finally, after relevance, a model prediction description is now available to users. Model tag brim and shaping techniques were used to generate explanations. The contribution of attributes for each error class was extracted using global descriptions obtained from the form and limestone diagrams. The explanation states that classified defects are often associated with abnormal voltage/current harmonic variations in the order of 3, 11, 13, 17. The main limitation of the study is the skewed class distribution, which can enforce prejudice in the majority class and increase prediction errors in the minority. Another drawback is the complexity of understanding the explanation of electronic gorisms, which requires extensive system knowledge.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

**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](http://doi.org/10.1007/s00202-013-0293-7)]
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](http://dx.doi.org/10.3182/20120829-3-MX-2028.00127)]
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](http://dx.doi.org/10.1016/j.est.2023.107575)]
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](http://dx.doi.org/10.1016/j.knosys.2023.110259)]
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](http://dx.doi.org/10.1016/j.neucom.2021.01.076)]
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](http://dx.doi.org/10.1109/MWSCAS.2018.8623823)]
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](http://dx.doi.org/10.1109/TDSC.2021.3118636)]
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](http://dx.doi.org/10.1109/ICPHM.2018.8448459)]
10. Upasane, S.J.; Hagras, 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](http://dx.doi.org/10.1109/FUZZ45933.2021.9494540)]
11. Hasan, M.J.; Sohaib, M.; Kim, J.M. An Explainable AI-Based Fault Diagnosis Model for Bearings. *Sensors* **2021**, *21*, 4070. [[CrossRef](http://dx.doi.org/10.3390/s21124070)] [[PubMed](http://www.ncbi.nlm.nih.gov/pubmed/34199163)]
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](http://dx.doi.org/10.1016/j.ymssp.2021.108105)]
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](http://dx.doi.org/10.1109/TII.2020.3012989)]
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](http://dx.doi.org/10.1109/ACCESS.2021.3095565)]
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](http://dx.doi.org/10.1109/ICPHM.2019.8819396)]
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](http://dx.doi.org/10.1109/ACCESS.2020.3006491)]
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](http://dx.doi.org/10.20944/preprints202109.0034.v1)]
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](http://dx.doi.org/10.1016/j.procir.2018.12.008)]
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](http://dx.doi.org/10.1109/JSEN.2019.2958787)]
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](http://dx.doi.org/10.1109/HSI.2018.8430788)]
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](http://dx.doi.org/10.1016/j.advengsoft.2022.103339)]
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](http://dx.doi.org/10.1109/MC.2021.3071551)]
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](http://dx.doi.org/10.1115/GT2022-80777)]
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](http://dx.doi.org/10.1109/ICKII50300.2020.9318912)]
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](http://dx.doi.org/10.1109/TIM.2019.2932162)]
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](http://dx.doi.org/10.1109/TSMC.2020.3048950)]
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](http://dx.doi.org/10.1109/BigComp48618.2020.00021)]
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](http://dx.doi.org/10.3390/pr9071121)]
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.
32. Bounemeur, A., chemachema, M., Zahaf, A., & Bououden, S. (2019, November). Adaptive fuzzy fault-tolerant control using nussbaum gain for a class of SISO nonlinear systems with unknown directions. In International conference on electrical engineering and control applications (pp. 493-510). Singapore: Springer Nature Singapore.
33. Bounemeur, A., & Chemachema, M. (2021). Adaptive fuzzy fault-tolerant control using Nussbaum-type function with state-dependent actuator failures. Neural Computing and Applications, 33, 191-208.
34. Waluyo, W., Ratna, S. U. S. A. N. A., & Robbie, K. M. (2022). Induction Heating Stove Prototype of 130 kHz using Arduino Uno. Electrotehnică, Electronică, Automatică (EEA), 70(1), 39-50
35. Kato, K., Gohda, T., & Iizuka, S. (2016). Production of Low Electron-Temperature and High Density Plasma by Using a Grid-Cage Electrode for RF Discharge. Journal of Scientific Research and Reports, 11(2), 1–9. <https://doi.org/10.9734/JSRR/2016/26362>
36. Zhu, Y., Guo, Y., Wang, Z., & Wei, Z. (2021). Kinetic energy based output power smoothing control and parameters design for PMSG-WECSs. International Journal of Electrical Power & Energy Systems, 131, 107077.
37. Zhu, Y., Ma, Z., & Wang, Z. (2022). An improved fuzzy logic based DC-link voltage control strategy for smoothing output power of the PMSG-WECS. Energy Reports, 8, 8413-8425.