Graphical User Interface Aided Intelligent Diagnosis of Stator Faults in Induction Motors

Jitendra Kumar Dewangan
IIIT Naya Raipur
Chhattisgarh, India
jitendra21101@iiitnr.edu.in

Harsh Patel
IIIT Naya Raipur
Chhattisgarh, India
harsh21100@iiitnr.edu.in

Aparna Sinha
IIIT Naya Raipur
Chhattisgarh, India
aparna.sinha@ieee.org

Debanjan Das
IIT Kharagpur
West Bengal, India
debanjan@ieee.org

***Abstract*—Induction motors form the backbone of many modern industries; therefore, the performance and safety of these industries rely heavily on motor health. The stator faults account for more than one-third of all motor faults. The existing literature primarily focuses on current/vibration-based models for stator fault diagnosis, without any attention to providing a suitable Graphical User Interface as well. To alleviate these drawbacks, an end-to-end intelligent solution using vibration signal has been proposed that allows the end-users to upload the acquired motor data in the application and get step-by-step fault diagnosis results without requiring any coding knowledge. The solution has been made flexible for use on a wide variety of datasets by applying AutoKeras model with the capacity to select a suitable deep-learning model and adjust its parameters in accordance with the specific characteristics of the dataset. The proposed method performs stator fault detection and severity prediction with 99.81% accuracy. The prediction outcome is interpreted by Explainable AI to improve model transparency and provide trustworthiness. The user-friendly Graphical User Interface using Streamlit enables easy dataset upload, EDA visualization, and interaction with predictions, bridging ML methodologies with practical applications. This research will prove very beneficial in facilitating Industry 4.0.**

***Index Terms*—Stator fault diagnosis, Graphical User Interface, AutoKeras, Explainable AI , Streamlit**

I. Introduction

When talking about modern industries, one of the fundamental and commonly used machines are the induction motors because of their durability, affordability, and low maintenance [1]. The reliability and efficiency of induction motors significantly impact industrial processes, making them essential for powering various equipment and ensuring seamless operations in manufacturing and other industries [2]. Fault diagnosis of induction motors at an early stage is essential, as it can cause unexpected downtime, loss of yield, and severe injuries for factory workers and engineers [3]. The operation of induction motors can be affected by various electrical, environmental, or physical factors. Challenges like extreme temperatures, moisture-induced corrosion, and dust accumulation causing motor wear are often addressed using traditional manual techniques. Figure 1 shows the distribution of several common faults occurring in the induction motors, of which stator faults cover nearly one-third of all faults.

Over the years, several researchers have proposed various stator fault diagnosis techniques for induction motors. Existing techniques use current signature analysis [4], vibration data [5], acoustic signal [1], or infrared thermography [6] for fault diagnosis. Although these existing solutions are accurate, their main drawback is the constrained practical implementation. Many of these solutions are confined to software-based simulations only and lack a suitable user interface for easy implementation without requiring much knowledge of coding. Further, most of the proposed methods include rigid models that can't adjust their parameters in accordance with the specific characteristics of the dataset. Hence, these models are not generalized to be used for induction motors with varying specifications.

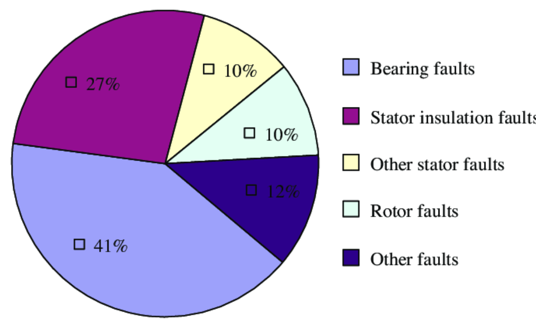

Fig. 1. Categories of induction motor faults

Therefore, to mitigate the above-mentioned problems, this paper proposes a stator fault detection and severity estimation technique for induction motors. A graphical user interface (GUI) has been designed to enable the end user to upload the collected vibration data and visualize the step-by-step fault diagnosis results, including feature extraction and selection, model training, validation, and explanation, without any knowledge of the back-end coding. An AutoKeras model is deployed for fault identification that automatically chooses the optimal model from a specified range of hyperparameters. As a result, the generated model is lightweight and capable of providing accurate results rapidly. Explainable AI (XAI) [7] has been utilized for in-depth interpretation of the 'black-box' model outcome to improve trustworthiness in case of field deployment.

The primary contributions of this paper include:

- This study introduces an intelligent GUI-aided approach for stator fault detection and severity estimation in in-

duction motors. The GUI empowers users of diverse technical backgrounds to effortlessly upload datasets, visualize EDA results, and interact with predictions.
- The proposed fault diagnosis technique ensures model flexibility by employing AutoKeras, which trains numerous models simultaneously and facilitates identifying the most suitable hyperparameter configuration.
- The integration of Explainable AI (XAI) techniques demystifies the decision-making process of complex models, providing stakeholders with transparent insights into the factors driving each prediction.

The remaining sections of this proposed paper are arranged as follows: Section II outlines the research gaps from the existing research, while Section III presents the Proposed Fault Detection Technique, followed by Section IV, which delves into the Results and Discussion. Finally, Section V encapsulates the conclusion and prospects for future research endeavors.

II. Related Prior Research and Research Gaps

Fault diagnosis techniques for induction motors are broadly categorized as model-based and data-driven. Model-based approaches, such as [8] and [9], require accurate modeling of system dynamics, which can be challenging and prone to errors. These methods also rely on theoretical assumptions, which may not fully capture real-world complexities. These drawbacks can be addressed by data-driven solutions that are becoming more common due to the advancements in Machine Learning (ML) and Artificial Intelligence (AI). These techniques operate without necessitating any pre-existing understanding of the particular system under consideration. Machine learning models are used to predict when electrical equipment, such as induction motor, transformers and circuit breakers, will fail.

Information fusion of current and vibration signals has been utilized in [10], employing Wavelet packet decomposition for time-frequency characteristics extraction, and Support Vector Machines for motor fault diagnosis. Similarly, the stator fault diagnosis framework for BLDC motors based on vibration and current signals has been proposed in [11]. Merging vibration and current data into a unified model poses significant challenges and requires considerable time due to the disparate sources of data: the accelerometer and the control system. Moreover, acquiring current data involves invasive procedures. Consequently, opting solely for vibration signals emerged as a more pragmatic choice, as they can be effortlessly gathered through non-intrusive means. Since they are already used widely for bearing fault detection [12], the same setup can be utilized for stator fault detection as well. The usability of only vibration signals for stator fault diagnosis in induction motors has been established and validated in [5] and [13]. However, these solutions face challenges in generalizing the models to different motor specifications, as they include rigid models that are incapable of adjusting their parameters according to the varying datasets. This problem can be mitigated using Automated Machine Learning (AutoML) [14] instead of a fixed-model method, as AutoML can change its architecture and hyper-parameters according to data fed into it. A significant drawback of data-driven algorithms is their *black box* nature, which leads engineers and field personnel to view them as unreliable and untrustworthy. Consequently, efforts must be made to enhance the transparency of the proposed model, allowing for the interpretation of prediction results through Explainable AI (XAI) techniques. However, none of the existing studies provide a suitable GUI that enables the end-users to upload the collected data and give step-by-step results for stator fault diagnosis, severity prediction, and model interpretation without requiring knowledge of coding.

The primary research deficiencies tackled in this study are:
- Lack of user-friendly fault diagnosis systems exclusively utilizing vibration data for stator fault detection in induction motors.
- Limited exploration of the application of vibration signals, commonly used for bearing fault detection, in diagnosing stator faults through a user interface-based system.
- Lack of robust model that can perform well in the variation of data available in the real-life scenario
- Scarcity of research on the integration of Explainable AI techniques in fault diagnosis systems tailored for industrial applications, particularly in the context of induction motor fault diagnosis.

The research gaps are consolidated in Table I.

TABLE I
Comparison between existing solutions and proposed method

Papers	Data-driven	Non-invasive	Self-learning	XAI	GUI
[8]	✗	✗	✗	✗	✗
[9]	✗	✗	✗	✗	✗
[10]	✓	✗	✗	✗	✗
[11]	✓	✗	✗	✗	✗
[5]	✓	✓	✗	✓	✗
[13]	✓	✓	✗	✓	✗
Current work	✓	✓	✓	✓	✓

III. The Proposed Fault Detection Technique

The proposed process flow for identifying stator faults is presented in Figure 2. The proposed technique is concisely outlined in Algorithm 1. Initially, vibration signals from the stator motor are collected to form a dataset, which will be used for fault diagnosis. Moreover, the data will undergo preprocessing, which involves the removal of invalid sample points and normalization using a standard scaler. Subsequently, an AutoKeras model will be constructed with a predetermined number of trials and trained on the processed data to yield an optimal model for fault diagnosis. A detailed description of the methodology is presented in the following sections.

Algorithm 1 Optimize Acquisition Function

Input: $\mathcal{A}$, r, S_{low}
$S \leftarrow 1$, $P \leftarrow PriorityQueue()$
$c_{min} \leftarrow$ lowest c in $\mathcal{A}$
for $(f, \boldsymbol{\theta_f}, c) \in \mathcal{H}$ **do**
    $P.Push(f)$
end for
while $P \neq \varnothing$ **and** $S > S_{low}$ **do**
    $S \leftarrow S \times r$, $f \leftarrow P.Pop()$
    for $o \in \Omega(f)$ **do**
        $f' \leftarrow \mathcal{M}(f, \{o\})$
        if $e^{\frac{c_{min} - \alpha(f')}{T}} > Rand()$ **then**
            P.Push(f')
        end if
        if $c_{min} > \alpha(f')$ **then**
            $c_{min} \leftarrow \alpha(f')$, $f_{min} \leftarrow f'$
        end if
    end for
end while
Return: The ancestor which is nearest of f_{min} in $\mathcal{A}$, the operation sequence to reach f_{min}

A. Dataset Description

The motor vibration data for this study is drawn from practical experiments performed in [5] using a power system hardware setup and Brüel & Kjær (B&K) accelerometer.The data is time series data containing vibrational signal. The data is collected via B&K data acquisition system (DAQ) at 800 Hz frequency. A rheostat has introduced stator inter-turn short circuit faults of 25%, 50% and 75% severity levels.

B. Feature Extraction

Machine-learning models exhibit limited accuracy when trained on raw data alone. Feature extraction emerges as a crucial process for identifying the most distinguishing characteristics within a signal. This facilitates data compatibility for ML model inputs, leading to reduced complexity of the model and improved performance. In this study, various time-domain features including max, min, min-max difference, median, mean, skewness, standard deviation, and kurtosis, along with frequency-domain features like peak frequency and spectral entropy, were extracted. To extract these features, a window which will slide over samples with a size of 100 and an overlap length of 50 was utilized. Assume that the data obtained from the accelerometer is denoted by an array x_i. Equations (1)-(5) provide the formulas for extracting time domain features:

$$\text{mean}(x) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{1}$$

$$\text{median}(x) = \frac{x\left[\frac{n}{2}\right] + x\left[\frac{n}{2}+1\right]}{2}, \tag{2}$$

$$\text{std}(x) = \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}, \tag{3}$$

$$\text{skew}(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)\cdot\sigma^3}, \tag{4}$$

$$\text{kurtosis}(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n\sigma^4}, \tag{5}$$

where n is the number of samples.

Equations (6) and (7) provide the formula for extracting frequency domain features, including peak frequency and spectral entropy, respectively:

$$\text{Peak Frequency} = f_{\text{peak}} = \frac{k_{\text{max}} \cdot f_{\text{sample}}}{N}, \tag{6}$$

where k_{max} is the index of the maximum magnitude in the FFT output, f_{sample} is the sampling frequency, N is the number of points in the FFT.

$$\text{Spectral Entropy} = H_{\text{spec}} = -\sum_{k=1}^{N} P(k)\log_2 P(k), \tag{7}$$

where P(k) is the normalized magnitude of the k_{th} frequency bin, given as:

$$P(k) = \frac{|X(k)|}{\sum_{k=1}^{N}|X(k)|}, \tag{8}$$

where $X(k)$ represents frequency magnitude.

C. AutoKeras-based Fault Detection Module

The early detection and severity estimation of stator faults in induction motors is accomplished utilizing the *StructuredDataClassifier* function of AutoKeras, incorporating a parameter referred to as *max_trials*. The *max_trials* parameter determines the quantity of distinct combinations of hyperparameter values to be tested [11]. Following this, all produced models are assessed, and the most effective model is chosen.

D. Model Architecture

The AutoKeras framework utilized a Structured Data Classifier to explore 20 different combinations of hyperparameters and layers. The best-performing model identified through this process is a 10-layer multilayer perceptron. The 10 layers comprised of dense layer, normalization layer, Relu layer, dropout layer and final classification layer. The ReLU and

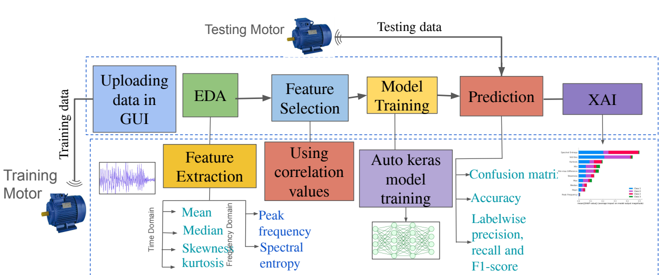

Fig. 2. Framework for the proposed methodology

Softmax activation functions are defined in Equations (9) and (10), respectively:

$$ReLU(x) = max(0, x), \tag{9}$$

$$S(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{N} e^{x_j}}. \tag{10}$$

E. Fault Detection Module

In this research, we are aiming to detect the phase-to-phase faults in an induction motor using an algorithm that can adjust its whole architecture, including hyperparameters according to the provided dataset. This flexible nature of the algorithm will help to boost the accuracy of the model significantly and make the algorithm robust. For this purpose, one of the popular tools named AutoKeras has been used for fault detection and severity prediction of phase-to-phase faults. AutoKeras is an open-source machine-learning library that automates the process of building and optimizing machine-learning models. It utilizes neural architecture search to automatically discover the most suitable neural network architecture for a given task, making it user-friendly for individuals with limited machine learning expertise.

F. Explainable AI

Explainable AI (XAI) is a framework that aims to make the decision-making processes of AI models understandable and interpretable for humans. It ensures transparency by providing prediction outcomes using various graphs and plots, which gives insights into how AI systems arrive at specific conclusions. In this paper, the SHapley Additive exPlanations (SHAP) technique, one of the many eXplainable AI (XAI) approaches, has been applied. Shapley values played a pivotal role in elucidating the decision-making process of the model. By computing the average impact of each feature on predictions across the entire dataset, Shap values provided a transparent and interpretable representation of variable importance [15]. This not only facilitated a comprehensive understanding of the factors influencing model outcomes but also empowered stakeholders to make informed decisions based on the nuanced insights gleaned from Shapley value analysis.

IV. Results and Discussion

A. Fault Identification

In instances involving sensor data, there is typically a substantial imbalance between normal and fault data, with the latter being significantly less prevalent. When dealing with imbalanced datasets, the confusion matrix proves more informative in evaluating the performance of the fault prediction module than relying solely on Accuracy. It offers a detailed perspective on the accuracy of predictions within each class and the specific types of errors occurring. Figure 3 presents the confusion matrix for the proposed fault detection model.

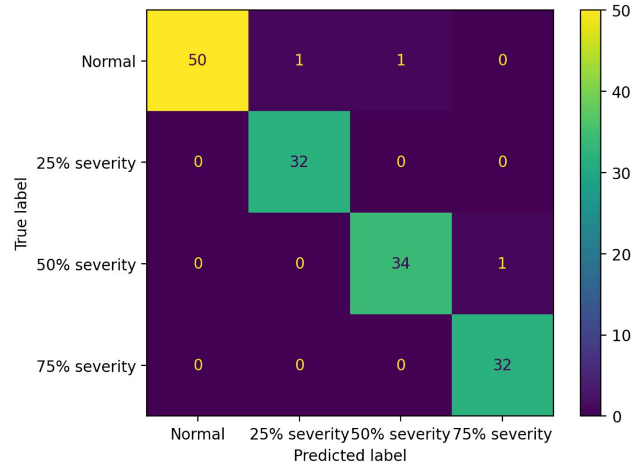

Fig. 3. Confusion matrix of the AutoKeras-based fault identification

The precision, recall, $f1-score$, and accuracy metrics are derived from the confusion matrix using the equations (11) through (14):

$$precision(p) = \frac{tp}{tp + fp}, \tag{11}$$

$$recall(r) = \frac{tp}{tp + fn}, \tag{12}$$

$$f1 - score = \frac{2 \times p \times r}{p + r}, \tag{13}$$

$$accuracy = \frac{tp + tn}{tp + fn + tn + fp}, \tag{14}$$

where tp, fn, tn, and fp denote true positive, false negative, true negative, and false positive, respectively. Precision represents the proportion of correctly identified positives among all predicted positives. Recall indicates the proportion of predicted positives out of all actual positives. The $f1-score$ represents the harmonic mean of Precision and Recall. These metrics provide insights into model performance even in imbalanced data. High precision and recall values are indicative of a well-performing model. AutoKeras generated 20 models through permutations of various hyperparameters, and the best-performing model was selected based on these metrics.

TABLE II
THE MOST OPTIMAL MODEL ACQUIRED

Type of Layer	Output Shape	No. of Parameters
Input	(None, 10)	0
Multi Category	(None, 10)	0
Normalization	(None, 10)	21
Dense	(None, 32)	352
ReLU	(None, 32)	0
Dense	(None, 32)	1056
ReLU	(None, 32)	0
Dropout	(None, 32)	0
Dense	(None, 4)	132
Classification	(None, 4)	0

The model summary for the chosen model is presented in Table II. A grand sum of 1561 parameters were acquired, with 1540 of them being trainable. Following this, the model underwent training using 2204 training samples. After 100 epochs, the validation accuracy reached 99.81% with a loss of 0.006, as depicted in Figure 4. After training, the best hyper-tuned model was picked by AutoKeras, whose performance metrics have been provided in Table III. It is clear from this table that the F1-Score for overall classes is 0.9801, which is a good result shown by the chosen model.

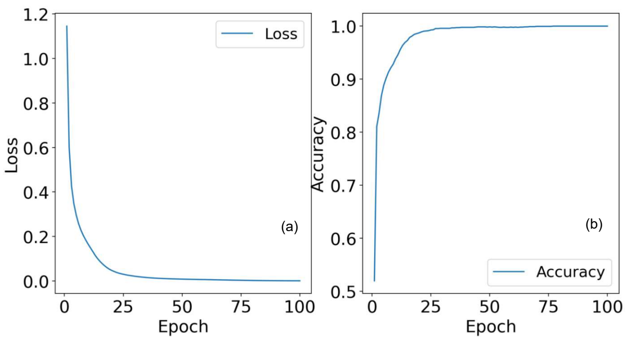

Fig. 4. Performance of the model (a) loss with epochs, (b) accuracy with epochs

TABLE III
METRICS FOR THE TOP-PERFORMING HYPERPARAMETER-TUNED MODEL ACHIEVED THROUGH AUTOKERAS

Class	Precision	Recall	F1-Score
Normal (0)	1.0000	0.9615	0.9804
25% severity (1)	0.9697	1.0000	0.9846
50% severity (2)	0.9714	0.9714	0.9714
75% severity (3)	0.9697	1.0000	0.9846
Overall	0.9805	0.9801	0.9801

Computational Efficiency: The final model requires less than two seconds for result prediction, and it operates efficiently on a computing device with only 1 GB of RAM and a clock speed of 900 MHz.

B. Interpretation using Explainable AI

1) Summary plot: The plot displays a summary of Shap values with variables along the y-axis and their corresponding Shap values along the x-axis. Each label is represented by distinct colors in Figure 5, providing a visual representation of the impact of different variables on the model's output. This visualization helps interpret the contribution of each feature to the model's predictions, enhancing understanding of variable importance and relationships within the dataset.

2) Force plot: A force plot is a local explanation plot explaining the model's prediction on a single data point. It shows the contribution of different variables on the output as shown in Figure 6.

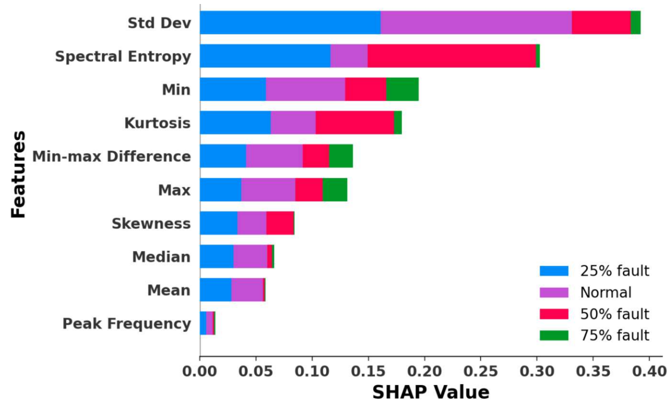

Fig. 5. Summary Plot depicting the impact of individual features on different target classes

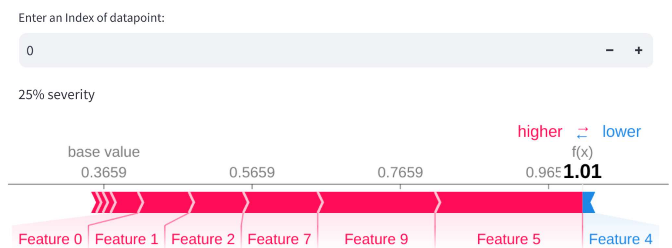

Fig. 6. Force Plot, where higher prediction is represented by features shown in red push, and the lower predictions are shown in blue push.

C. Model Sustainability

From the aforementioned outcomes, it's evident that the suggested solution facilitates swift stator motor fault detection. Because of its lightweight and adaptable design, the model can be applied even on edge devices with limited resources, where data collection takes place. This would lead to onboard execution, providing quicker fault detection outcomes. The entire process is automated, eliminating the need for human intervention in fault detection. Also, there will be no issues of trust with this model, as this paper proposes the method of implementing XAI on the model, which increases the model's interpretability as well.

D. Graphical User Interface

Our developed graphical user interface provides an intuitive platform for every step of the stator motor fault diagnosis system using AutoKeras and Explainable AI. Leveraging the *Streamlit* framework, users can seamlessly navigate through various stages, starting from dataset upload and exploratory data analysis (EDA) to feature selection, model training, prediction, and interpretation through Explainable AI techniques, within a few seconds. The interface offers an interactive experience, allowing users to visualize insights at each stage, facilitating efficient decision-making and enhancing transparency in the diagnostic process. Its user-friendly design fosters accessibility for both experts and non-experts, ensuring effective utilization of the diagnostic system. The user interface proposed by this research is provided in Figure 7. The dashboard on the left allows the user to select a particular functionality, and the corresponding output is displayed on the right. The app can be accessed via the following link: https://stator-motor-deploy-app-amybcmrxux5fiyg5pdsa9t.streamlit.app/.

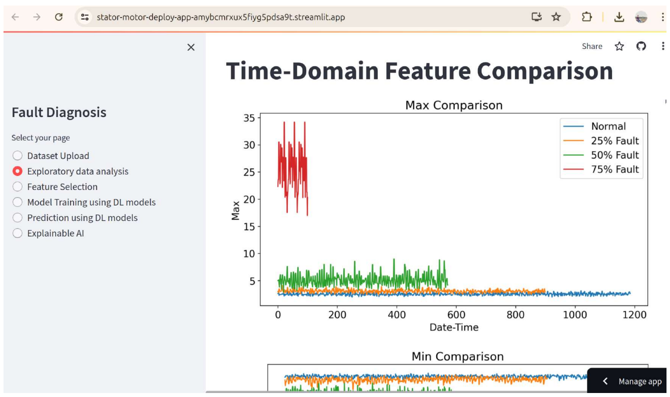

Fig. 7. User Interface showing the dashboard on the left and output on the right

V. Conclusions and Future Scope

Faults in induction motors such as Inter-turn short circuit can result in unexpected downtime, elevated manufacturing expenses, and safety risks in industrial settings. Many current studies rely on complex fault detection methods utilizing current data from motor control systems. Hence, this study presents a methodological framework utilizing data analysis for the detection of electrical faults, leveraging solely vibration data to simplify the data acquisition process. The major conclusions of this paper are summarized in the following sentences: (1) The AutoKeras model proposed achieves fault identification accuracy of 99.81%. (2) Explainable AI is successful in showing the importance of each feature for a particular prediction through summary plot and force plot. (3) The designed GUI provides a seamless and interactive experience for the end-users, regardless of their expertise in coding. For future scope, sound signals can be utilized instead of vibration signals originating from the induction motor for fault identification. Additionally, rather than focusing on a single component - as this paper does, considering only the induction motor - one could extend this project to encompass multiple industrial components, such as turbines and compressors.

References

[1] A. Choudhary, R. K. Mishra, S. Fatima, and B. Panigrahi, "Multi-input CNN based vibro-acoustic fusion for accurate fault diagnosis of induction motor," *Engineering Applications of Artificial Intelligence*, vol. 120, p. 105872, 2023.

[2] G. Niu, X. Dong, and Y. Chen, "Motor Fault Diagnostics Based on Current Signatures: A Review," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–19, 2023.

[3] E. C.-Y. A. Garcia-Perez, R. d. J. Romero-Troncoso and R. A. Osornio-Rios, "The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, May 2011.

[4] F. Husari and J. Seshadrinath, "Stator Turn Fault Diagnosis and Severity Assessment in Converter-Fed Induction Motor Using Flat Diagnosis Structure Based on Deep Learning Approach," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 11, no. 6, pp. 5649–5657, 2023.

[5] A. Sinha and D. Das, "Machine Learning-based Explainable Stator Fault Diagnosis in Induction Motor using Vibration Signal," in *2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2023, pp. 1–6.

[6] A. Glowacz, "Thermographic fault diagnosis of electrical faults of commutator and induction motors," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 105962, 2023.

[7] E. Brusa, L. Cibrario, C. Delprete, and L. G. Di Maggio, "Explainable AI for Machine Fault Diagnosis: Understanding Features' Contribution in Machine Learning Models for Industrial Condition Monitoring," *Applied Sciences*, vol. 13, no. 4, p. 2038, 2023.

[8] S.-M. M. Moosavi, J. Faiz, M. B. Abadi, and S. M. Cruz, "Comparison of rotor electrical fault indices owing to inter-turn short circuit and unbalanced resistance in doubly-fed induction generator," *IET Electric Power Applications*, vol. 13, no. 2, pp. 235–242, 2019.

[9] M. H. Nazemi, F. Haghjoo, S. M. A. Cruz, and D. Gallehdar, "Stator and rotor turn-to-turn fault detection in wound rotor induction machines based on the air-gap magnetic field distortion," *IET Electric Power Applications*, vol. 14, pp. 2631–2639, 2020.

[10] M.-K. Liu, M.-Q. Tran, and P.-Y. Weng, "Fusion of Vibration and Current Signatures for the Fault Diagnosis of Induction Machines," *Shock and Vibration*, vol. 2019, p. 7176482, 2019.

[11] T. A. Shifat and J. W. Hur, "An Effective Stator Fault Diagnosis Framework of BLDC Motor Based on Vibration and Current Signals," *IEEE Access*, vol. 8, pp. 106 968–106 981, 2020.

[12] W. Rohouma, A. Zaitouny, M. F. Wahid, H. Ali, and S. S. Refaat, "A Machine Learning Framework for Bearing Fault Detection in Three-Phase Induction Motors," in *2024 4th International Conference on Smart Grid and Renewable Energy (SGRE)*, 2024, pp. 1–5.

[13] A. Sinha and D. Das, "An explainable deep learning approach for detection and isolation of sensor and machine faults in predictive maintenance paradigm," *Measurement Science and Technology*, vol. 35, no. 1, p. 015122, 2023.

[14] N. F. Fadzail, S. M. Zali, E. C. Mid, and R. Jailani, "Application of Automated Machine Learning (AutoML) Method in Wind Turbine Fault Detection," *Journal of Physics: Conference Series*, vol. 2312, p. 012074, 2022.

[15] D. Chowdhury, A. Sinha, and D. Das, "XAI-3DP: Diagnosis and Understanding Faults of 3-D Printer With Explainable Ensemble AI," *IEEE Sensors Letters*, vol. 7, no. 1, pp. 1–4, 2023.