Data-Driven Fault Diagnosis in Three-Phase IMs: Harnessing Current Signature Data with Machine Learning Algorithms

Uvesh Sipai[1][0000-0003-0253-9504], Hagos L.Shifare[2][0009-0003-1448-9425], and Nishant Kothari[3][0000-0003-2107-4466]

[1] Department of Electrical Engineering, Marwadi University, Rajkot, India
uvesh.sipai@gmail.com
[2] Department of Computer Engineering-DS, Marwadi University, Rajkot, India
hagosleshifare@gmail.com
[3] Department of Electrical Engineering, Marwadi University, Rajkot, India
nishantkothari77@gmail.com

Abstract. This paper presents a fault diagnosis technique in three-phase induction motors (IM) utilizing stator current signature as feature input to four different machine learning algorithms. The performance of k-Nearest Neighbour(kNN), Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF) classifiers has been analyzed for classifying external faults and abnormalities in 3-phase IM. A comparative evaluation of instantaneous and RMS currents has been carried as feature input for classifying Normal Load (NL), Overload (OL), Overvoltage (OV), Undervoltage (UV), Single Phasing (SP), and Voltage Unbalanced (VUB) conditions. Stator currents from both the experimental setup and simulation have been used to evaluate performance. The analysis suggests that the RMS current-based SVM classifier performed consistently and reliably with simulation and experimental datasets demonstrating better-generalized capabilities.

Keywords: Fault diagnosis· Three-phase induction motor· Stator current · Machine learning algorithms

1 Introduction

Among the different types of electric motors, IMs are the most widely used type, and this popularity is justified by their exceptional performance in different fields. IMs have consistently demonstrated their effectiveness due to several fundamental features that are desirable in a range of applications, including cost-effective operation, durability, high efficiency, and reliability, minimal maintenance requirements, and high starting torque[12]. The reliable operation and efficiency of three-phase IMs can be compromised by various internal and external faults [11]. Internal faults are faults that originate within the motor itself, such as short circuits and bearing faults, while external faults are faults that arise from external factors impacting the motor's performance.

One critical external fault is SP, resulting from the loss of one phase in the power supply due to blown fuses, damaged lines, or faulty connections. IMs facing SP conditions typically struggle to initiate, and if operational, experience a substantial torque reduction. This situation may lead to increased current imbalance, excessive vibration, decreased efficiency, and the imminent risk of motor burnout [2].

VUB, another significant external fault disturbance, arises from unequal voltages across the three phases due to faulty connections, overloaded lines, or unbalanced source voltage. Acceptable voltage unbalance is generally defined within a 5% difference between phases. Exceeding this threshold can result in significant performance degradation, manifesting as increased rotor current, excessive vibration, reduced efficiency, and potential bearing damage [7].

Voltage fluctuations beyond the rated range introduce OV and UV conditions. OV, exceeding 110%, induces increased iron losses, increased torque, and potential insulation breakdown. UV, falling below 90% of the rated voltage, poses risks of reduced torque, heightened current, and overheating. Both scenarios jeopardize motor performance, carrying implications for efficiency and insulation integrity [17].

The OL conditions, stemming from prolonged operation above 110% of the rated load, can result from excessive load demands, improper sizing, or jammed mechanisms. Adverse effects include increased rotor current, excessive heat generation, heightened vibration, and the looming risk of winding insulation breakdown [16].

In contrast, operating under NL allows the motor to function within its intended design parameters, ensuring optimal efficiency, performance, and lifespan. Recent research underscores the significance of accurately identifying and maintaining NL conditions. Real-time monitoring and data analysis play a crucial role in distinguishing NL from early-stage fault conditions, enabling preventive maintenance and mitigating potential damage [3].

Over the years, various methods have been employed for fault detection. While conventional techniques were somewhat effective, Artificial Intelligence-based approaches have emerged as superior, offering greater reliability and robustness. In this research, the primary objective is to assess and compare the efficacy of four distinct models in classifying external faults/abnormalities in three-phase IM utilizing stator current signatures. The outcomes of the proposed work are anticipated to refine fault identification strategies, ensuring a more balanced and efficient operation of three-phase IMs.

2 Literature review

In recent studies addressing fault diagnosis and condition assessment of three-phase IMs, various methodologies leveraging machine learning techniques have emerged. The authors in the work developed by Sandeep Sharma et al. [18] proposed a Multi-Class Extreme Learning Machine (ELM) for external fault classification. This approach utilized RMS values of 3-phase voltages and currents, achieving robust identification of six fault types. Comparatively, another recent work by Vanga et al. [20] introduced a Bidirectional Long Short Term Memory (Bi-LSTM) network, demonstrating its efficiency in fault classification. Leveraging line voltages and currents, the Bi-LSTM outperformed traditional LSTM networks, showcasing advantages in convergence and accuracy.

In parallel, Chudasama et al. [6] presented a Subtractive Clustering-Based Sugeno Fuzzy Inference System for external fault detection in low-voltage three-phase IMs. This fuzzy logic-based method surpassed conventional thermal relays in accuracy, providing a robust solution for fault identification. In [7] the authors addressed fault monitoring and diagnosis using an Artificial Neural Network (ANN) integrated into batch simulation. The simulated results demonstrated that well-trained neural networks can precisely detect and diagnose early faults in three-phase IMs, presenting a reliable and effective protection scheme.

A practical approach for external fault identification in three-phase IMs using a Proximal Support Vector Machine (PSVM) was explored in the research of [15] where the PSVM algorithm showcased faster investigations, leading to a reduction in computational load. Furthermore, the study outlined in [12] presented an ANN-based technique for identifying various faults in three-phase IMs. Their approach, utilizing three-phase currents and voltages, demonstrated effectiveness in fault identification during simulated and online testing.

Adding to this discourse, a novel approach for external fault identification in three-phase IMs using Motor Electrical Signature Analysis (MESA) was investigated in the study conducted by the authors of [5]. The research delves into the exploration of multilabel classification techniques, employing Ensemble Bagged Tree and Support Vector Machine , demonstrating their effectiveness even in the presence of noise in the dataset.

Nonetheless, the existing body of research, as outlined in the reviewed papers, has not collectively implemented diverse models for both hardware and simulation data, including features like RMS and instantaneous current, to offer a comprehensive comparative analysis. The proposed work utilizes stator current signature for fault diagnosis in three-phase IMs.

3 Methodology

In the course of its operational lifespan, a three-phase IM experiences fluctuations in loads and exposure to various environmental conditions. These factors play a significant role in influencing the motor's overall performance and longevity. In the context of this research, the focus lies on the analysis of external faults in IMs, specifically those arising from variations in load, voltage conditions, and SP incidents. To achieve this, various models were implemented to discern distinct fault signatures present in the instantaneous and RMS current data. In the proposed work KNN, SVM, LR, and RF were selected for a comprehensive comparison and in-depth study. Key theoretical principles and parameters employed in the analysis of fault classification are discussed for each technique.

3.1 K-Nearest Neighbors (KNN)

KNN stands out as a straightforward yet powerful algorithm for classification and regression[9]. Its principle is based on assigning a data point's class by considering the majority class among its k nearest neighbors. The versatility and non-parametric nature of KNN contribute to its suitability across diverse applications. To understand KNN for classification better, refer to [1].

3.2 Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm utilized for both classification and regression tasks, with Support Vector Classifier (SVC) [4] specifically employed for binary and multiclass classification tasks [13]. The fundamental concept of SVC is to find the optimal hyperplane that maximally separates different classes in the feature space. The algorithm works by identifying support vectors, which are data points that lie closest to the decision boundary. For more insight into the mechanisms of the SVM algorithm reference [4] is recommended.

3.3 Logistic Regression (LR)

LR is a widely used classification algorithm[10] that models the probability of a data point belonging to a specific class. It utilizes the logistic function to map the linear combination of input features to a probability score. In this study, Multinomial Logistic Regression is applied as it is used to model the relationship involving a categorical dependent variable with more than two categories. Reference [19] is advised for a better understanding of the LR model.

3.4 Random Forest (RF)

RF is an ensemble learning algorithm [8] that builds a multitude of decision trees during training. For classification tasks, each tree independently predicts the class, and the final output is determined by voting [14]. The ensemble nature of RF enhances its resilience to noisy data and contributes to improved generalization performance.

4 System Modelling

4.1 Simulation Model

The simulation model of the proposed method is shown in Fig.1. The model includes a three-phase, 4kW/5.4 HP, 400 V, 50 Hz, 4-pole squirrel-cage IM developed in MATLAB/Simulink software. The rated load and speed of the motor are 26.72 N-m and 1430 RPM. Three-phase instantaneous and RMS currents have been utilized as feature input to the classifiers for classifying external faults as shown in Fig 1.

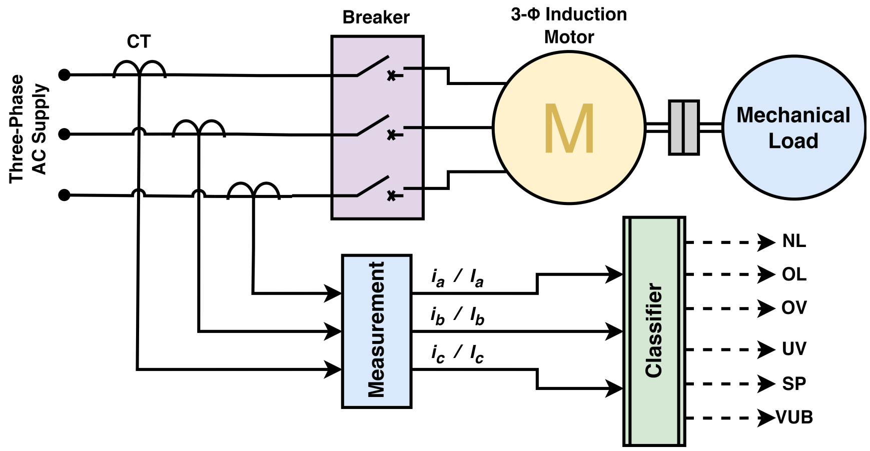

Fig. 1. Proposed fault diagnosis schem

Various external faults/abnormalities such as OL, OV, UV, VUB and SP have been simulated at different load conditions. Figure 2 shows instantaneous currents for above-mentioned conditions.

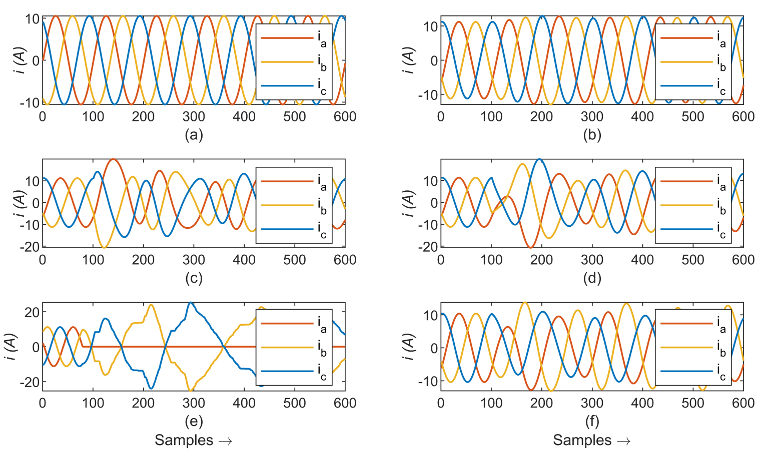

Fig. 2. Instantaneous currents for (a) NL, (b) OL, (c) OV, (d) UV, (e) SP, and (f) VUB conditions

Table 1 shows values selected for different operating conditions of three-phase IM. Three different load conditions have been considered for simulating various faults and abnormalities.

Table 1. Parameters for various fault/abnormal conditions in IM.

Operating Condition	**Values**	**Number of Cases**
Overload	111% to 120% of Rated Load	$10 \times 3^{*} = 30$
OV	112% to 121% of Rated Voltage	$10 \times 3^{*} = 30$
UV	89% to 80% of Rated Voltage	$10 \times 3^{*} = 30$
Voltage Unbalance	V_R=100%, V_Y=94%, V_B=92% V_R=100%, V_Y=92%, V_B=94% V_R=94%, V_Y=100%, V_B=92% V_R=92%, V_Y=100%, V_B=94% V_R=94%, V_Y=92%, V_B=100% V_R=92%, V_Y=94%, V_B=100% (Similarly for 100%, 93%, and 91% of rated voltage)	$12 \times 3^{*} = 36$
Single Phasing	Instances of SP w.r.t VR (0°, 30°, 60°, 90°, 115°)	$5 \times 3^{\#} \times 3^{*} = 45$
Normal Condition	75% to 100% of Rated Load	26
Total No. of Cases		197

***** Three different load conditions: 100%, 90%, and 80% of rated load
\# Single-phasing condition in each phase: R, Y, and B

Further, out of total cases simulated, cases for training and testing of the algorithms have been separated out as mentioned below in Table 2.

Table 2. Parameters for various fault/abnormal conditions in IM.

Operating Condition	**Training Parameters**	**Training Cases**	**Testing Parameters**	**Testing Cases**
Overload	112%, 114%, 116%, 118%, and 120% of rated load	15	111%, 113%,115%, 117%, and 119% of rated load	15
Over Voltage	112%, 114%, 116%, 118%, 120% of rated voltage	15	113%, 115%, 117%, 119%, 121% of rated voltage	15
Under Voltage	88%, 86%, 84%, 82%, 80% of rated voltage	15	89%, 87%, 85%, 83%, 81% of rated voltage	15
Voltage Unbalance	100%, 94%, 92% of rated voltage	18	100%, 93%, 91% of rated voltage	18
Single Phasing	0°, 60°, 115°	27	30°, 90°	18
Normal Conditions	75% to 87% Load	13	88% to 100% Load	13
Total No. of Cases		103	94	

5 Experimental Set up

The experimental set up in the laboratory environment consists a set of a three-phase, 3.7 kW/5 HP, 415 V, 7.5A, 50 Hz, 1425 RPM, 4-pole IM coupled with a 3kW, 230V, 10.1A, 1500 RPM DC shunt generator. Figure 3 shows the experimental setup of three-phase IM. The DC generator is loaded with a lamp-load. Three-phase instantaneous currents were recorded using current probes and YOKOGAWA DLM 2024 digital storage oscilloscope (DSO). Subsequently, RMS currents were derived from instantaneous signals. From the experimental set-up various conditions of OL (9 cases), UV (16 cases), SP (16 cases), and VUB (16 cases) have been recorded. Due to supply voltage limitations, the OV condition was not considered for the experimental setup.

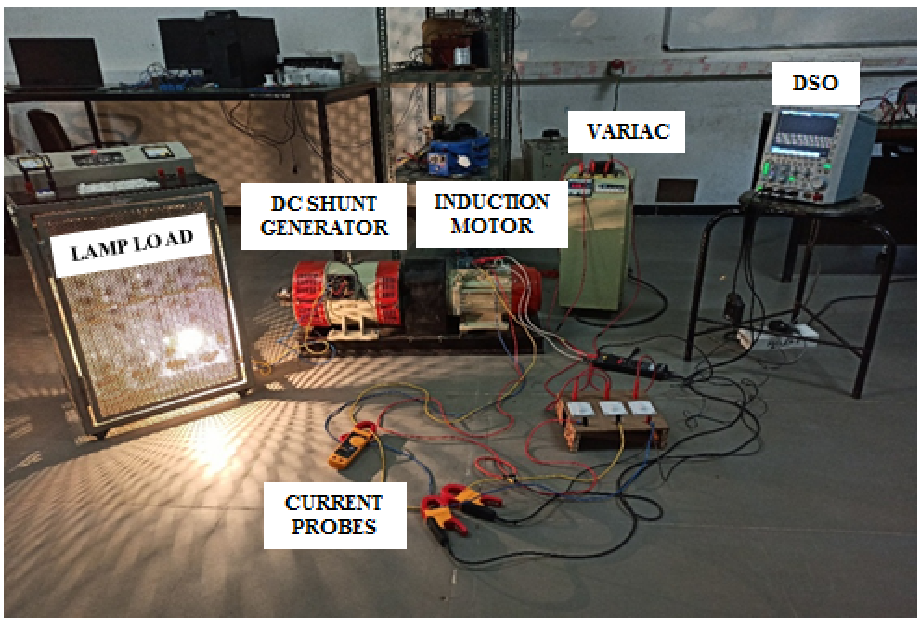

Fig. 3. Experimental setup of three phase IM coupled to DC generator.

An in-depth analysis was conducted utilizing simulation and experimental setups, where the currents obtained from both setups were employed as feature inputs for the classifiers. Recognizing the importance of parameter selection on the performance of machine learning classifiers, a meticulous approach was adopted in this research. Specifically, the parameters for the classifiers were carefully chosen through a systematic grid search method coupled with a rigorous five-fold cross-validation technique. The results of this parameter optimization are documented in Table 3, highlighting the best values identified through the aforementioned grid search and cross-validation techniques. The proposed work was carried out using Python 3.11.0 on a computer having an Intel Core i5-8250 CPU @ 1.65 GHz, 16 GB RAM, and 2 GB NVIDIA GeForce 940MX GPU.

Table 3. Model Parameters

Classifier	Best Hyperparameters
kNN	n_neighbors: 3, weights: distance
SVM	C: 100, gamma: 0.0001, kernel: rbf
LR	C: 100, penalty: l2, solver: liblinear
RF	max_depth: 10, max_features: log2, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 50

6 Results and Discussion

The performance evaluation of the classifiers with post-fault five cycles instantaneous and RMS currents for classifying external faults/abnormal conditions in three-phase IM is presented below. It is noteworthy that the classifiers were trained and tested using cases with distinct parameter values.

6.1 Performance with Simulation Data

The performance of the classifiers in terms of class-wise classification accuracy is depicted in Table 4 and 5 with instantaneous and RMS currents, respectively.

Table 4. Class-wise performance with instantaneous currents

Classifier	Class wise performance (% Accuracy)						Overall % Accuracy
	NL	OL	OV	UV	SP	VUB	
kNN	100	100	100	98.92	98.92	100	98.92
SVM	100	100	100	100	100	100	100
LR	100	100	100	100	98.92	98.92	98.92
RF	100	100	100	100	100	100	100

Table 5. Class-wise performance with RMS currents

Classifier	Class wise performance (% Accuracy)						Overall % Accuracy
	NL	OL	OV	UV	SP	VUB	
kNN	86.02	90.32	100	100	100	86.02	86.02
SVM	97.85	97.85	100	100	100	100	97.84
LR	100	100	100	100	100	100	100
RF	84.95	100	100	100	100	84.95	84.94

The overall performance in terms of Train and Test accuracies, F1 score, CV score, Recall and Precision is shown in Fig 4 with simulation data.

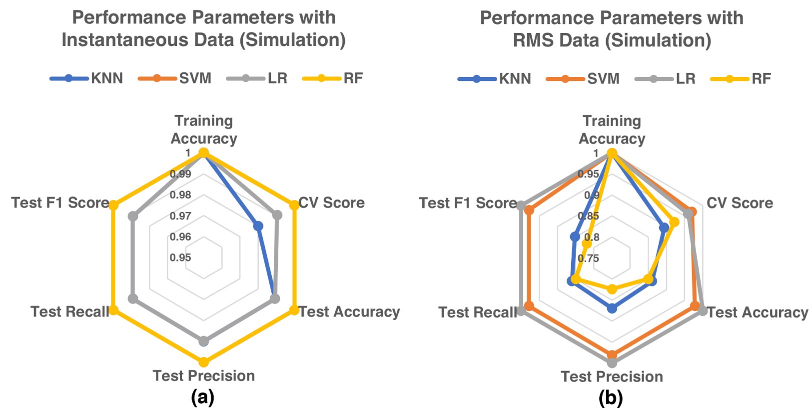

Fig. 4. Performance parameters with simulation data (a) Instantaneous, and (b) RMS currents

From above-mentioned results, it can be observed that the performance of SVM and LR is reliable with instantaneous as well as RMS currents achieving consistent score in all the performance parameters. The performance of RF is identical to SVM with instantaneous currents thus overlapping in Fig 4 (b). Nevertheless, with RMS currents RF classifier performance has declined significantly. Similarly, the performance of the kNN classifier also declined with RMS currents.

6.2 Performance with Experimental Data

Out of a total of 57 experimental cases recorded for four different operating conditions, 29 (OL-5, UV-8, SP-8, and VUB-8) cases were used for training and the performance of the trained model is evaluated with the remaining 28 cases. Therefore, the train-test ratio is nearly 50%. The class-wise performance with the experimental dataset is shown in Table 6 and Table 7 for instantaneous and RMS currents, respectively.

Table 6. Class-wise performance with RMS currents

Classifier	**No. of misclassified cases**				**Total misclassified cases (out of 28)**	**Overall % Accuracy**
	OL (out of 4)	UV (out of 8)	SP (out of 8)	VUB (out of 8)		
kNN	0	0	1	0	1	96.42
SVM	0	0	1	0	1	96.42
LR	0	0	2	2	4	85.71
RF	0	0	0	6	6	78.57

Table 7. Class-wise performance with RMS currents

Classifier	**No. of misclassified cases**				**Total misclassified cases (out of 28)**	**Overall % Accuracy**
	OL (out of 4)	UV (out of 8)	SP (out of 8)	VUB (out of 8)		
kNN	0	0	0	0	0	100
SVM	0	0	0	0	0	100
LR	0	0	1	0	1	96.42
RF	0	0	0	0	0	100

The overall performance with the experimental dataset is shown in Fig 5.

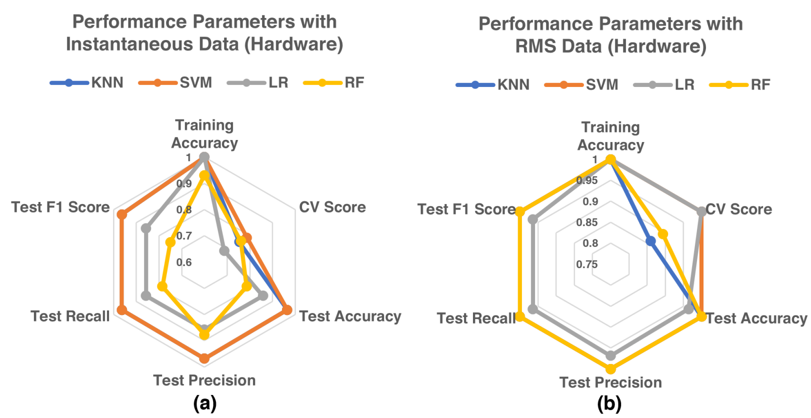

Fig. 5. Performance parameters with experimental data (a) Instantaneous, and (b) RMS currents

In comparison to simulation results, the performance of LR and RF classifiers declined significantly with experimental instantaneous signals. Although the test accuracy of kNN and SVM is consistent with experimental instantaneous signals, a poor CV score indicates chances of overfitting/underfitting of the models and may not perform well on new, unseen data. Considering the accuracy, the performance of SVM and LR classifiers is found reliable with experimental RMS current signals. However, considering overall performance analysis, the SVM classifier is found to be consistent and reliable for external faults and abnormalities classification using RMS current signals. The SVM classifier has achieved overall classification accuracies of 97.84% and 100% with RMS currents for simulation and experimental data, respectively.

7 Conclusion and Future Work

This study thoroughly assesses the effectiveness of instantaneous and RMS currents as feature inputs to the classifiers for classifying external faults and abnormal operating conditions in three-phase IM. Through simulations and experimental setups, kNN, SVM, LR, and RF classifiers were scrutinized using raw current signals.

Among the classifiers examined, SVM consistently outperforms the others, showcasing superior and consistent performance with raw current signals. Particularly, SVM exhibits better-generalized capabilities compared to the alternatives considered in this study. Moreover, the performance of SVM with RMS currents stands out, showing promise in comparison to using instantaneous currents.

Given the observed inconsistencies in classifier performance with instantaneous currents, future endeavors will explore employing diverse signal processing techniques to extract pertinent features from instantaneous currents. This avenue holds the potential for enhancing classifier performance and refining fault classification in three-phase IMs.

References

1. Abu Alfeilat, H.A., Hassanat, A.B., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Eyal Salman, H.S., Prasath, V.S.: Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data **7**(4), 221–248 (2019). https://doi.org/10.1089/big.2018.0175
2. Al-Yoonus, M.A., Alyozbaky, O.S.A.d.: Detection of internal and external faults of single-phase induction motor using current signature. International Journal of Electrical and Computer Engineering **11**(4), 2830 (2021). https://doi.org/DOI:10.11591/IJECE.V11I4.PP2830-2841
3. Badr, B.E., Altawil, I., Almomani, M., Al-Saadi, M., Alkhurainej, M.: Fault diagnosis of three-phase induction motors using convolutional neural networks. Mathematical Modelling of Engineering Problems **10**(5) (2023)
4. Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., Lopez, A.: A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing **408**, 189–215 (2020). https://doi.org/https://doi.org/10.1016/j.neucom.2019.10.118
5. Chandran, L.R., Ilango, K., Nair, M.G., Kumar, A.A., Kumar, A.A., et al.: Multilabel external fault classification of induction motor using machine learning models. In: 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT). pp. 559–564. IEEE (2022). https://doi.org/10.1109/ICICICT54557.2022.9917882
6. Chudasama, K., Shah, V., Shah, S.: Induction motor relaying scheme for external faults detection and classification using subtractive clustering based sugeno fuzzy inference system. Electric Power Components and Systems **44**(10), 1149–1162 (2016). https://doi.org/10.1080/15325008.2016.1149255
7. Eldin, E.M.T., Emara, H.R., Aboul-Zahab, E.M., Refaat, S.S.: Monitoring and diagnosis of external faults in three phase induction motors using artificial neural network. In: 2007 IEEE Power Engineering Society General Meeting. pp. 1–7. IEEE (2007). https://doi.org/10.1109/PES.2007.385469

8. Gron, A.: Hands-on machine learning with scikit learn keras&tensorflow . 2019
9. Helmina, A., Lailani, F.K., Fatihaturahmi, F., Jalinus, N., Waskito, W.: Use of the k-nearest neighbor (knn) algorithm in new categories based on books. Jurnal Pendidikan dan Konseling (JPDK) **5**(1), 2570–2579 (2023)
10. Hilbe, J.M.: Logistic regression. International encyclopedia of statistical science **1**, 15–32 (2011)
11. Khanjani, M., Ezoji, M.: Electrical fault detection in three-phase induction motor using deep network-based features of thermograms. Measurement **173**, 108622 (2021)
12. Kolla, S., Varatharasa, L.: Identifying three-phase induction motor faults using artificial neural networks. ISA transactions **39**(4), 433–439 (2000). https://doi.org/10.1016/S0019-0578(00)00031-8
13. Kothari, N.H., Bhalja, B.R., Pandya, V., Tripathi, P.: A rate-of-change-of-current based fault classification technique for thyristor-controlled series-compensated transmission lines. International Journal of Emerging Electric Power Systems **23**(3), 289–304 (2021)
14. Kothari, N.H., Bhalja, B.R., Pandya, V., Tripathi, P., Jena, S.: A faulty section identification scheme in thyristor controlled series compensated transmission lines using superimposed currents. In: 2019 8th International Conference on Power Systems (ICPS). pp. 1–6. IEEE (2019)
15. Mittal, A., Malik, H., Rastogi, S., Talur, V.: External fault identification experienced by 3-phase induction motor using psvm. In: 2014 6th IEEE Power India International Conference (PIICON). pp. 1–6. IEEE (2014)
16. Reuben, J., Onah, C., JU, A.: Modeling and simulation of three phase induction motor electrical faults using matlab/simulink. International Journal of Modern Trends in Engineering and Research (IJMTER) **5**(5), 176–187 (2018)
17. Salau, A.O., Kanchana, K., Anoop, K.J., Markus, E.D., Braide, S.L.: Suppression of over voltage in sic-based inverter fed induction motor. Australian Journal of Electrical and Electronics Engineering pp. 1–18 (2023)
18. Sharma, S., Malik, H., Khatri, A.: External fault classification experienced by three-phase induction motor based on multi-class elm. Procedia Computer Science **70**, 814–820 (2015)
19. Starbuck, C.: Logistic regression. In: The Fundamentals of People Analytics: With Applications in R, pp. 223–238. Springer (2023)
20. Vanga, J., Ranimekhala, D.P., Jonnala, S., Jamalapuram, J., Gutta, B., Gampa, S.R., Alluri, A.: Fault classification of three phase induction motors using bi-lstm networks. Journal of Electrical Systems and Information Technology **10**(1), 1–15 (2023)