Results in Engineering 30 (2026) 110933

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

Machine learning framework for multi-fault diagnosis within three-phase induction motor in Electric Vehicle applications

Mohamed Sharawy [a], Adel El-Nahas [a,*], M.A. Alahmar [a], Fahd A. Banakhr [b], Mohamed I. Mosaad [b], Mohamed Selmy [a,c]

[a] Electrical Engineering Department Faculty of Engineering at Shoubra, Benha University, Egypt
[b] Royal, Commission Yanbu Colleges Institutes, Yanbu Industrial College, Yanbu 46452, Saudi Arabia
[c] Innovation University, Cairo, Egypt

ARTICLE INFO

Keywords:
Electric vehicles
Induction motors
Machine learning
Fault diagnosis
Artificial intelligence

ABSTRACT

Induction Motors (IMs), which are known for their reliable performance, Low capital cost, and minimal operational expenditures, are a key component of Electric Vehicle (EV) powertrains. Nevertheless, they are susceptible to different electrical and mechanical faults that severely impact vehicle safety and performance. This necessitates the need for robust early fault detection systems. Although Machine Learning (ML) and Artificial Intelligence (AI) outperform traditional methods in terms of fault diagnostic features and accuracy, many existing studies remain limited by using simplistic models, focusing on single-fault investigations, a small number of input features, and a lack of comprehensive validations. To address these limitations, this paper presented a robust and extensively validated ML framework for IM fault diagnosis in EV settings. In this work, a high-resolution dataset was generated using MATLAB/Simulink at sampling rate of 150 kHz. The date set comprises many fault types, including Broken Rotor Bar (BRB), Inter-Turn Short Circuit (ITSC), Over/Under Voltage (OV/ UV), and Single Phasing (SPH) faults, under different load conditions. Five ML models, comprising Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Ensemble Model, and Decision Tree (DT), were optimized via Grid Search. These models were evaluated with strict sequential data partitioning, which in turn reduced data leakage. The five ML techniques achieved near-optimal accuracy (≈100%) in classifying IM faults. As explicitly clarified in this study, this exceptional performance establishes a robust theoretical baseline; it is primarily attributed to the high separability of features extracted from the clean, noise-free, and high-fidelity Simulink environment. Consequently, these results provide a strong proof-of-concept for the proposed diagnostic framework under idealized conditions.

1. Introduction

The rapid expansion of electric vehicles (EVs) has increased the demand for reliable, efficient, and fault-tolerant electric drive systems [1]. The appropriate electric motor selection is a key design choice that affects EVs performance, cost, and reliability. While various types are used, including brushless DC motors (BLDCs), permanent magnet synchronous motors (PMSMs), switched reluctance motors (SRMs), and induction motors (IMs), PMSMs and IMs are the most common. PMSMs are very popular because they have a very high-power density and peak efficiency due to the use of strong rare-earth magnets. However, this dependence on costly and geopolitically sensitive materials like neodymium and dysprosium creates significant risks in the supply chain, and costs can change significantly. Also, PMSMs could be dangerous and ineffective because they can generate power without control in faulty conditions, and because it is hard to manage demagnetization at high temperatures. Conversely, the IMs are a strong and affordable option as depicted in the comparison between other motor types shown in Fig. 1, especially the squirrel-cage version. Its simple, strong rotor design without permanent magnets, brushes, or slip rings makes it reliable, durable, and cheaper. Modern vector control methods have solved the old problem of precise speed control, which means that IMs can provide high starting torque and a wide speed range, even when running at high speeds for long periods. The IM is better for high-performance applications, especially in mainstream sedans and commercial vehicles, where scalability and value are most important. It is less expensive,

* Corresponding author.
E-mail address: adel.elnahas@feng.bu.edu.eg (A. El-Nahas).

https://doi.org/10.1016/j.rineng.2026.110933
Received 8 September 2025; Received in revised form 15 April 2026; Accepted 6 May 2026
Available online 7 May 2026

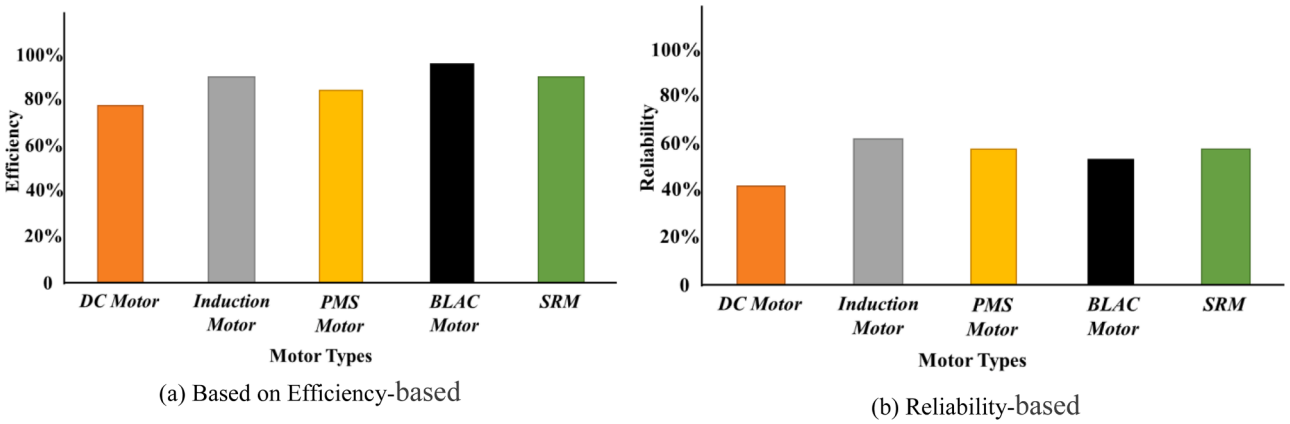


Fig. 1. Comparison between motor types used in EV, [3].

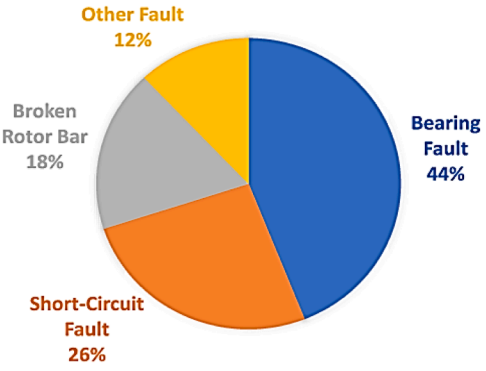


Fig. 2. Percentages of IM faults [4].

fault-tolerant, and can handle harsh operating conditions [2,3].

Common faults in Ims as depicted in Fig. 2 include inter-turn short circuits (ITSC), broken rotor bars (BRB), overvoltage (OV), undervoltage (UV), and single phasing (SPH). These faults negatively impact efficiency, torque stability, and operational reliability, making accurate and early fault detection a critical requirement for ensuring safe and continuous operation [4,5].

As EV adoption has increased dramatically, accurate fault detection of the electric motor has become essential. EV motors are prone to failures that could affect their performance, efficiency, and safety because they require intricate electrical systems, unlike combustion engines. Undetected faults can lead to increased energy consumption, costly failures, and power loss. Diagnostic technologies are a significant emphasis in recent vehicle engineering.

Traditional fault detection approaches, including signal-based and model-based methods, suffer from several limitations such as dependency on expert knowledge, sensitivity to noise, and limited real-time applicability [6,7]. In contrast, machine learning (ML) techniques have emerged as powerful tools for data-driven fault diagnosis, offering improved accuracy, adaptability, and automation [8,9].

Despite the growing adoption of ML in fault detection, several research gaps still exist. Many previous studies rely on a single ML algorithm, which limits model generalization and robustness. In addition, fault diagnosis is often restricted to a limited number of fault types, failing to represent realistic operating conditions. Another limitation is the dependence on a single type of feature, such as current or vibration signals, which reduces diagnostic effectiveness. Furthermore, most studies evaluate models under fixed operating conditions, without considering the impact of varying load levels. From a performance perspective, achieving very high classification accuracy remains a challenge, particularly when dealing with complex fault scenarios. Moreover, limited dataset sizes and insufficient validation strategies in prior work reduce the reliability and practical applicability of the developed models.

To address these limitations, this study proposes a comprehensive and robust ML-based framework for fault detection in induction motors. Multiple machine learning algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees (DT), and Ensemble methods, are implemented and optimized using hyperparameter tuning based on grid search. A wide range of fault conditions (ITSC, BRB, OV, UV, and SPH) is investigated under different loading scenarios, including no load (NL), half load (HL), and full load (FL). A large-scale dataset is generated through independent simulation, ensuring sufficient data diversity and reliability. In addition, a multi-source feature set consisting of stator and rotor currents, torque, and speed is utilized to enhance fault discrimination capability. To ensure robust evaluation, the dataset is divided into training and testing sets (80/20 split), and 5-fold cross-validation is applied. Furthermore, repeated experiments are conducted, and statistical performance metrics, including mean accuracy, standard deviation, confidence intervals, and extreme values, are reported.

The principal contributions of this study can be summarized as follows:

- **Comprehensive Comparative Framework:** Developing a rigorous benchmarking framework that evaluates and compares five distinct machine learning architectures (ANN, KNN, SVM, DT, and Ensemble) for the simultaneous diagnosis of multiple electrical and mechanical faults in induction motors.
- **Comprehensive Multi-Fault Diagnostic Scope:** Developing a robust framework capable of identifying and classifying a wide range of both electrical and mechanical faults (including BRB, ITSC, OV, UV, and SPH), providing a more holistic diagnostic solution for EV induction motors compared to single-fault studies.
- **Large-Scale High-Resolution Dataset:** Generating an exceptionally large dataset consisting of over 2.25 million samples recorded at a high sampling frequency of 150 kHz. This massive data scale ensures that the ML models capture subtle transient behaviors and high-frequency fault signatures that are often missed in smaller datasets.
- **Diverse Operational Loading Scenarios:** Evaluating the diagnostic system under various load conditions specifically Full-Load (FL), Half-Load (HL), and No-Load (NL). Unlike studies restricted to constant loads, this approach ensures the model's reliability across the dynamic power demands typical of real-world electric vehicle operation.
- **Extensive Comparative Benchmarking:** Conducting a rigorous performance evaluation of five distinct machine learning architectures (ANN, KNN, SVM, DT, and Ensemble). This large-scale

Table 1
Summary of research gaps and proposed solutions.

No.	Ref.	Identified Research Gap	Limitation in Previous Studies	Proposed Solution in This Study
1	[10,11, 31–40]	**Use of a single ML model**	Limits generalization and robustness across different datasets and operating conditions	Multiple ML models are implemented and compared (ANN, SVM, KNN, DT, Ensemble) to enhance robustness and performance
2	[12,13, 41–48]	**Limited fault coverage**	Most studies focus on one or two faults only (e.g., BRB or ITSC), not reflecting real-world conditions	A wide range of electrical and mechanical faults is considered (ITSC, BRB, OV, UV, SPH)
3	[14,15, 49,50]	**Dependence on single feature type**	Using only current or vibration signals reduces diagnostic accuracy	Multi-source feature set is used (Isa, Isb, Isc, Ira, Irb, Irc, Speed, Torque) for improved fault discrimination
4	[16,17, 51–53]	**Low or moderate accuracy**	Existing models struggle to achieve very high classification accuracy	Hyperparameter optimization (Grid Search) is applied to maximize model performance
5	[18–21]	**Fixed operating conditions**	Models are tested under limited or constant load conditions	Fault detection is evaluated under multiple loading conditions (NL, HL, FL)
6	[17,22, 23,54]	**Lack of statistical validation**	Many studies report only single-run accuracy results	Repeated experiments (6 runs) with statistical metrics (mean, std, confidence interval, min, max) are conducted
7	[19–21]	**Limited Dataset Size**	Small datasets reduce model reliability and generalization	Large-scale dataset generated (up to 150,000 samples per fault per load condition)

comparison identifies the most effective algorithms for high-dimensional fault data and establishes a performance baseline for future research.

- **Methodological Integrity and Transparency:** Implementing a strict hierarchical data isolation protocol and advanced multi-domain feature engineering (Time-Frequency). By utilizing Chi-Square analysis for feature ranking, the study provides a transparent justification for the models' success and establishes a "theoretical upper-bound benchmark" for induction motor diagnostics.

The identified research gaps and the corresponding contributions of this work are summarized in Table 1.

2. Induction motor modeling

The purpose of this section is to develop a high-fidelity mathematical representation of the IM that serves as the foundation for data generation. By utilizing a dynamic D-q axis model, we translate the electrical input variables (voltages) into mechanical outputs (torque and speed) through a set of differential equations. Crucially, this modeling framework is designed not only to simulate healthy motor operation but also to allow for the precise 'injection' of specific electrical and mechanical faults such as ITSC and BRB by modifying the internal resistance and inductance parameters. This approach ensures that the dataset captures the subtle physical deviations caused by each fault, providing the ML models with realistic and discriminative features for training.

The Modeling phase involved creating a comprehensive simulation under both healthy and faulty conditions. The motor specifications are summarized in the Appendix. The model was executed by using a fixed-step solver with a step size of (1/150,000) *sec*, ensuring precise numerical integration and stable simulation results.

2.1. Modeling of healthy state

The dynamic model of an IM as depicted in Fig. 3 can be derived by examining the equivalence of three-phase and two-phase machines [24]. The basis for the equivalency is the MMF generated in two-phase and three-phase windings, together with identical current magnitudes. The two-phase winding will have $3N_s/2$ turns per phase for MMF equality, if each of the three-phase windings has N_s turns per phase and equal current magnitudes. Resolving the MMF of the three-phase along the d and q axes yields the MMF of the d and q axes. The existing equalities remain once the common term, the number of turns, is cancelled on both sides of the equations. Flux links can be used as variables to represent the dynamic model of an IM in any reference frame. This entails reducing several variables in dynamic equations, which makes it much easier to

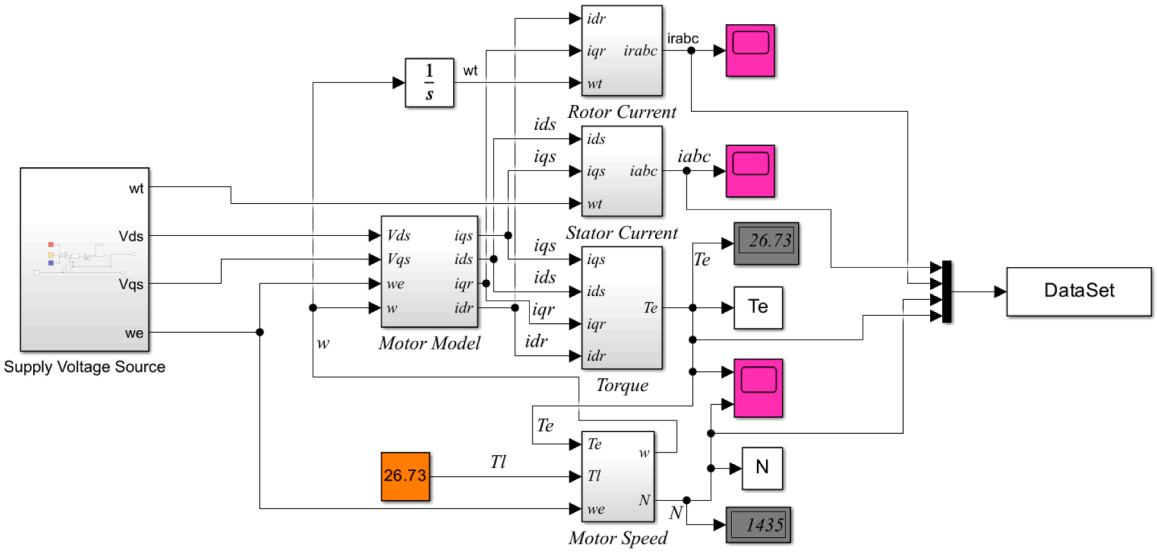


Fig. 3. Implementation of mathematical modeling of IM.

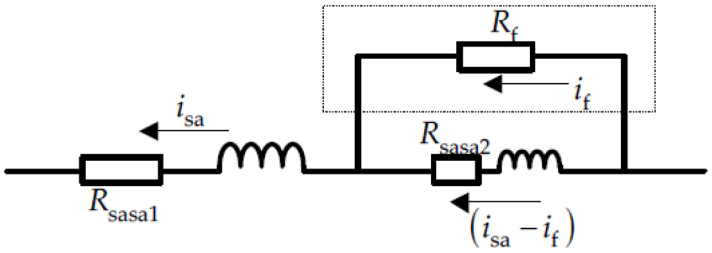


Fig. 4. Representation of ITSC fault at phase (A) [26].

solve those using analogue and hybrid computers. The flux connections are continuous even with discontinuous voltages and currents. This offers the benefit of numerical stability in differentiating these variables. Furthermore, motor drives use flux linkages to illustrate how the torque and flux channels in synchronous and induction machines decouple. The various flux linkage equations are expressed as,

$$\frac{d\Psi_{qs}}{dt} = \omega_b\left[V_{qs} - \left(\frac{\omega_e}{\omega_b}\right)\Psi_{ds} - \left(\frac{R_s}{X_{ls}}\right)(\Psi_{mq} - \Psi_{qs})\right] \tag{1}$$

$$\frac{d\Psi_{ds}}{dt} = \omega_b\left[V_{ds} - \left(\frac{\omega_e}{\omega_b}\right)\Psi_{qs} - \left(\frac{R_s}{X_{ls}}\right)(\Psi_{md} - \Psi_{ds})\right] \tag{2}$$

$$\frac{d\Psi_{qr}}{dt} = \omega_b\left[V_{qr} - \left(\frac{\omega_e - \omega_r}{\omega_b}\right)\Psi_{dr} - \left(\frac{R_r}{X_{lr}}\right)(\Psi_{mq} - \Psi_{qr})\right] \tag{3}$$

$$\frac{d\Psi_{dr}}{dt} = \omega_b\left[V_{dr} - \left(\frac{\omega_e - \omega_r}{\omega_b}\right)\Psi_{qr} - \left(\frac{R_r}{X_{lr}}\right)(\Psi_{md} - \Psi_{dr})\right] \tag{4}$$

Where,

$$\Psi_{mq} = X_{ml}\left[\frac{\Psi_{qs}}{X_{ls}} + \frac{\Psi_{qr}}{X_{lr}}\right],\ and\ \Psi_{md} = X_{ml}\left[\frac{\Psi_{ds}}{X_{ls}} + \frac{\Psi_{dr}}{X_{lr}}\right] \tag{5}$$

Then substituting the values of the flux linkages to find the currents,

$$I_{qs} = 1\Big/\left[\frac{1}{X_m} + \frac{1}{X_{ls}} + \frac{1}{X_{lr}}\right] \tag{6}$$

$$I_{ds} = \frac{1}{X_{ls}}(\Psi_{ds} - \Psi_{md}) \tag{7}$$

$$I_{qr} = \frac{1}{X_{lr}}(\Psi_{qr} - \Psi_{mq}) \tag{8}$$

$$I_{dr} = \frac{1}{X_{lr}}(\Psi_{dr} - \Psi_{md}) \tag{9}$$

Based on the above equations, the Electromagnetic torque and rotor speed can be determined as follows,

$$T_e = \left(\frac{3}{2}\right)\left(\frac{P}{2}\right)\left(\frac{1}{\omega_b}\right)(\Psi_{ds}I_{qs} - \Psi_{qs}I_{ds}) \tag{10}$$

$$\omega_b = \int\left(\frac{P}{2J}\right)(T_e - T_l) \tag{11}$$

2.2. Mathematical model of induction motor with BRB fault

The air gap between the rotor and stator of IMs has been symmetrically distributed. This symmetry could be broken by any malfunction that arises in those motors. When a BRB fault occurs in an IM, the rotor phase resistance changes and the rotor phases become unbalanced, which disrupts the electromagnetic field's symmetry. Consequently, the motor currents exhibit additional frequency components due to this imbalanced situation. Thus, the extra resistance that will be added to the rotor phase resistance in the mathematical model of the IM makes it simple to simulate the BRB fault. Within this modeling study, the effect of the rotor end-ring has been ignored; and it was accepted that stator resistance and inductance didn't change [25]. In that case, the rotor resistance matrix modified (R_r^*) on three-phase domain can be written as,

$$R_r^* = \begin{bmatrix} (R_r + \Delta R_{ra}) & 0 & 0 \\ 0 & (R_r + \Delta R_{rb}) & 0 \\ 0 & 0 & (R_r + \Delta R_{rc}) \end{bmatrix} \tag{12}$$

Where: ΔR_{ra},ΔR_{rb}, ΔR_{rc} represent the rotor resistances change that occur because of the BRB fault in the phases.

$$\Delta R_{ra,b,c} = R_r\left[\frac{3n_{bb}}{N_b - 3n_{bb}}\right] \tag{13}$$

Where: n_{bb}, N_b, R_r represent the number of BRBs, total number of the rotor bars and rotor resistance equivalent to one phase.

When ignoring the effect of the end-ring, rotor resistance equivalent to one phase in a healthy IM can be expressed as,

$$R_r \approx R_b\left[\frac{(2N_s)^2}{\frac{N_b}{3}}\right] \tag{14}$$

Where: R_b represents rotor bar resistance.

In that case, a rotor phase equivalent resistance that has n_{bb} number of BRB faults is presented below.

$$R_r^* \approx R_b\left[\frac{(2N_s)^2}{\frac{N_b}{3} - 3n_{bb}}\right] \tag{15}$$

Where: N_s represents the number of equivalent stator winding turns.

In that case, the change in the rotor phase resistances because of the BRB fault is defined with ΔR,

$$\Delta R_r = R_r^* - R_r = R_r\left[\frac{3n_{bb}}{N_b - 3n_{bb}}\right] \tag{16}$$

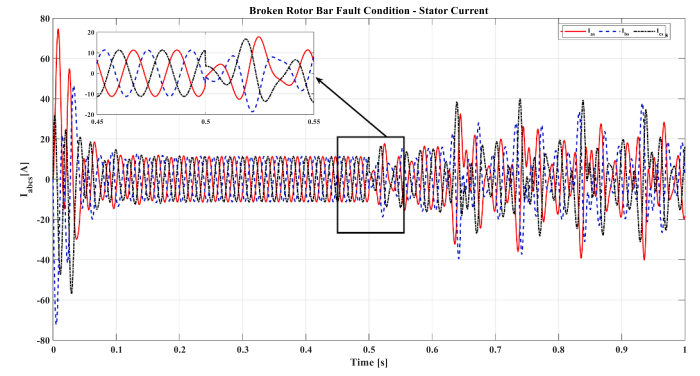
As clearly seen from the equations an additional resistance having a value convenient to a rotor phase resistance is added in an IM, a desired number of BRB faults can easily be modelled.

In this research, we performed IM modeling of the BRB Fault by using $N_b = 32$ bars and $n_{bb} = 5$.

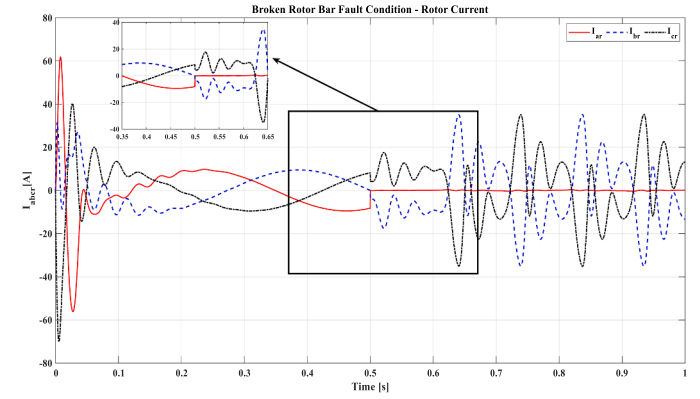
2.3. Mathematical model of induction motor with ITSC fault

As depicted in Fig. 4, if the motor exhibits a stator fault, an additional branch with short-circuit resistance R_f will be introduced to the stator winding's short section. Kirchhoff's current law states that the current of the fault section in the stator winding is $(I_{sa} - I_f)$ since the current I_f passes through the short circuit resistanceR_f [26].

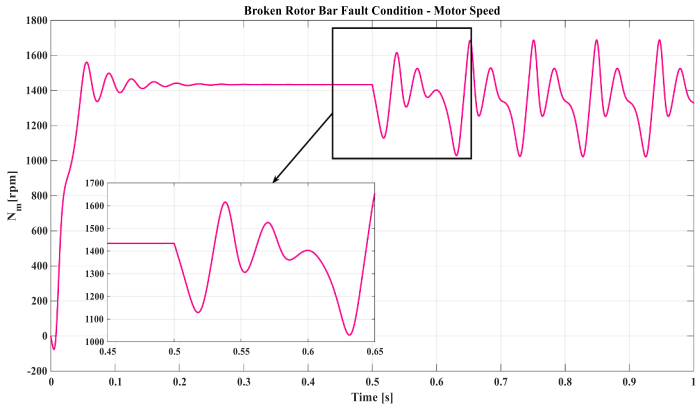
Also, it further shows that the stator fault is located at phase A. The parameters related to faulted turns are proportional to the number of shorted turns n_{sf}. For example, the stator resistance of the shorted part is



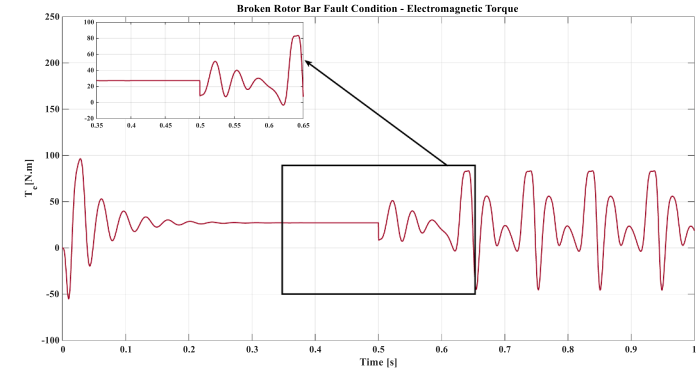
(a)



(b)

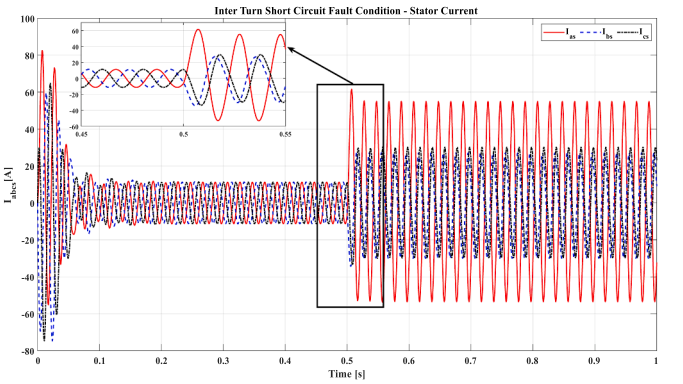


(c)

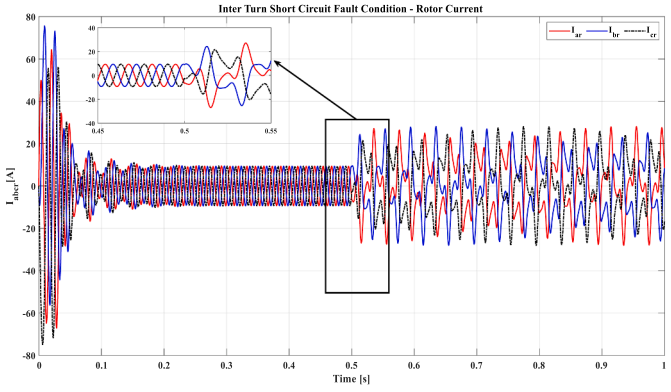


(d)

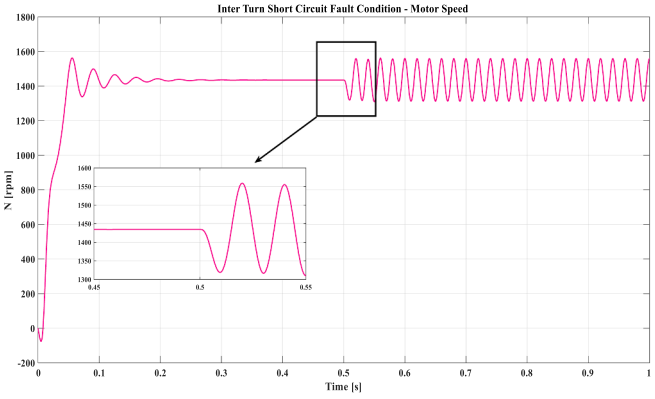
Fig. 5. Waveforms of IM during BRP Fault Condition, (a) Stator Current, (b) Rotor Current, (c) Motor Speed, and (d) Torque.



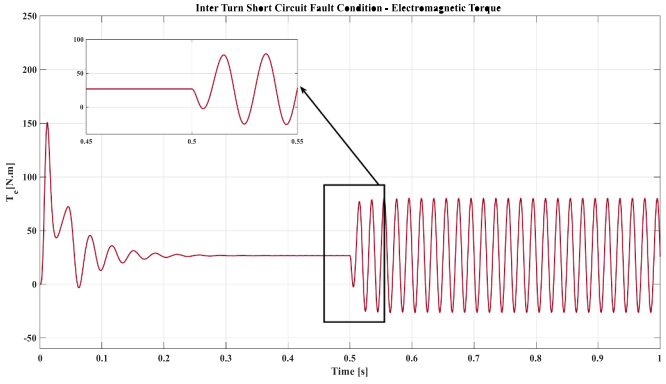
(a)



(b)



(c)



(d)

Fig. 6. Waveforms of IM during ITSC Fault Condition, (a) Stator Current, (b) Rotor Current, (c) Motor Speed, and (d) Torque.

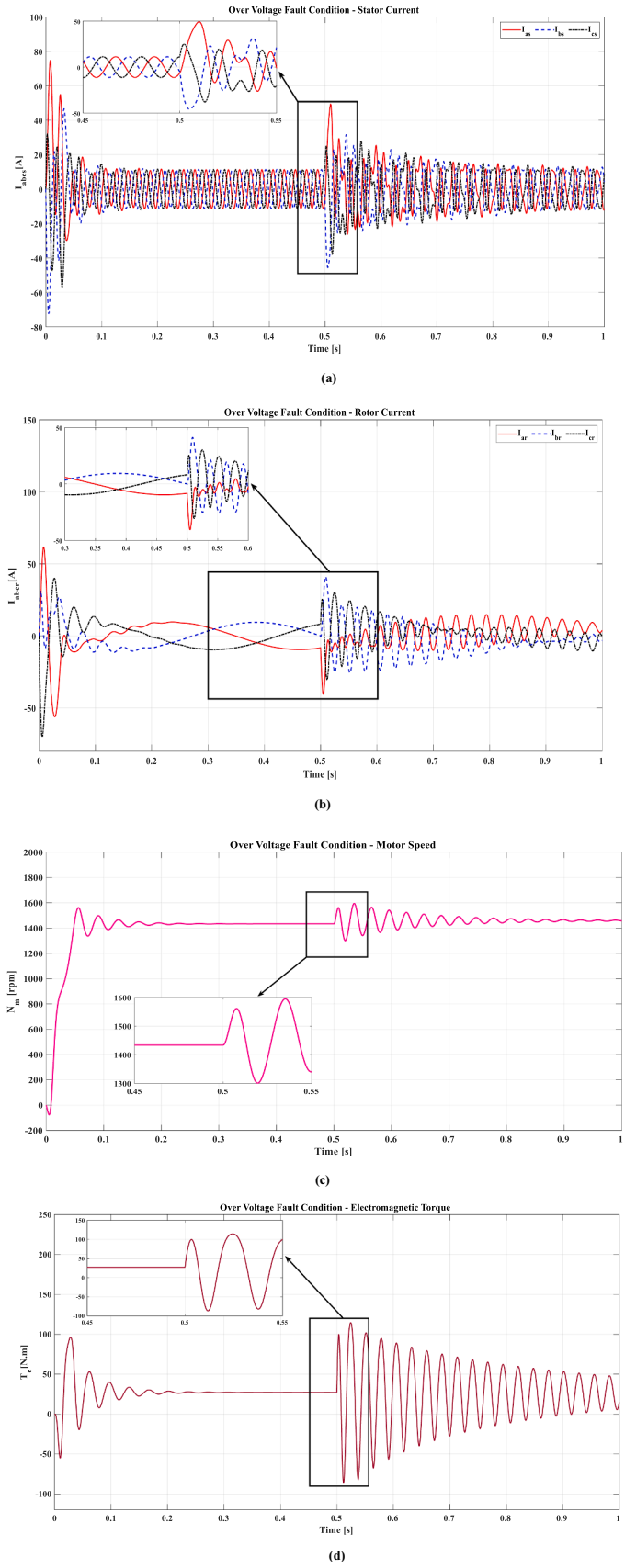


Fig. 7. Waveforms of IM during OV Fault Condition, (a) Stator Current, (b) Rotor Current, (c) Motor Speed, and (d) Torque.

$k_{sf}R_s$ ($k_{sf} = n_{sf}/z_s$, and z_s is the total number of stator slots), whereas the resistance of the healthy part is $(1 - k_{sf})\,R_s$. Hence, the voltage equations of the stator phase A and the shorted circuit are derived as,

$$V_{sa} = (1 - k_{sf})R_s i_{sa} + k_{sf}R_s(I_{sa} - I_f) + d\lambda_{sa}/dt \tag{17}$$

$$R_f I_f = k_{sf}R_s(I_{sa} - I_f) + d\lambda_f/dt \tag{18}$$

Notably, other equations related to the shorted phase should be updated accordingly. Consequently, the vector of the stator fault model can be simplified as,

$$\begin{bmatrix} V_{sabc} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} R_{sabc} & R_{sf} & pL_{sr} \\ R_{fs} & R_{ff} & pL_{fr} \\ pL_{rs} & pL_{rf} & R_r \end{bmatrix} \begin{bmatrix} i_{sabc} \\ i_f \\ i_r \end{bmatrix} + \begin{bmatrix} L_{sabc} & L_{sf} & L_{sr} \\ L_{fs} & L_{ff} & L_{fr} \\ L_{rs} & L_{rf} & L_r \end{bmatrix} p \begin{bmatrix} i_{sabc} \\ i_f \\ i_r \end{bmatrix} \tag{19}$$

Where,

$$R_{sf} = \begin{bmatrix} -k_{sf}R_s & 0 & 0 \end{bmatrix}^T \tag{20}$$

$$R_{ff} = -(k_{sf}R_s + R_f) \tag{21}$$

$$L_{sf} = \begin{bmatrix} -k_{sf}(L_{ms} + L_{ls}) & 0.5k_{sf}L_{ms} & 0.5k_{sf}L_{ms} \end{bmatrix}^T \tag{22}$$

$$L_{ff} = -k_{sf}^2(L_{ms} + L_{ls}) \tag{23}$$

$$L_{fr} = k_{sf}L_{sar} \;\; (L_{sar} = [L_{sar1} \quad L_{sar2} \quad \ldots. \quad 0]) \tag{24}$$

$$R_{fs} = -R_{sf}^T \tag{25}$$

$$L_{fs} = -L_{sf}^T \tag{26}$$

$$L_{rf} = -L_{fr}^T \tag{27}$$

For the ITSC simulation, a fractional winding fault was introduced by setting the short-circuited turns (k_{sf}) to 5% of the total stator winding, representing an incipient fault. The fault was modeled with a short-circuit resistance (R_f) of 0.1 Ω to approximate a near-ideal low-impedance short circuit.

2.4. Modeling of over-voltage, under-voltage, and single-phasing faults

For OV, UV, and SPH faults, the simulation model is the same as the healthy model, but by controlling the supply voltages and connected phases across motor terminals. In OV fault, the input voltage was just increased by 25% above the motor's rated voltage, for UV fault, the input voltage was just reduced by 25% below the motor's rated voltage, and for SPH fault was modeled by setting one of the three input phases (phase c) that was disconnected and then simulating a single-phasing fault.

3. Simulation of IM and dataset extraction

This study specifically examines fault conditions in the three-phase IM EV powertrain operating at a constant speed of 1430 RPM. This operational mode is very important for applications that need to accurately track speed. Keeping a constant rotational speed gives you a stable baseline to find and study the signatures of different IM faults, like rotor bar defects and stator winding asymmetries, without the problems of transient operation.

3.1. Simulation of IM

Following simulations for both healthy and various fault conditions under different loading scenarios, the IM was driven during this phase. A total of 1 s was used on the condition. On the other hand, the motor was

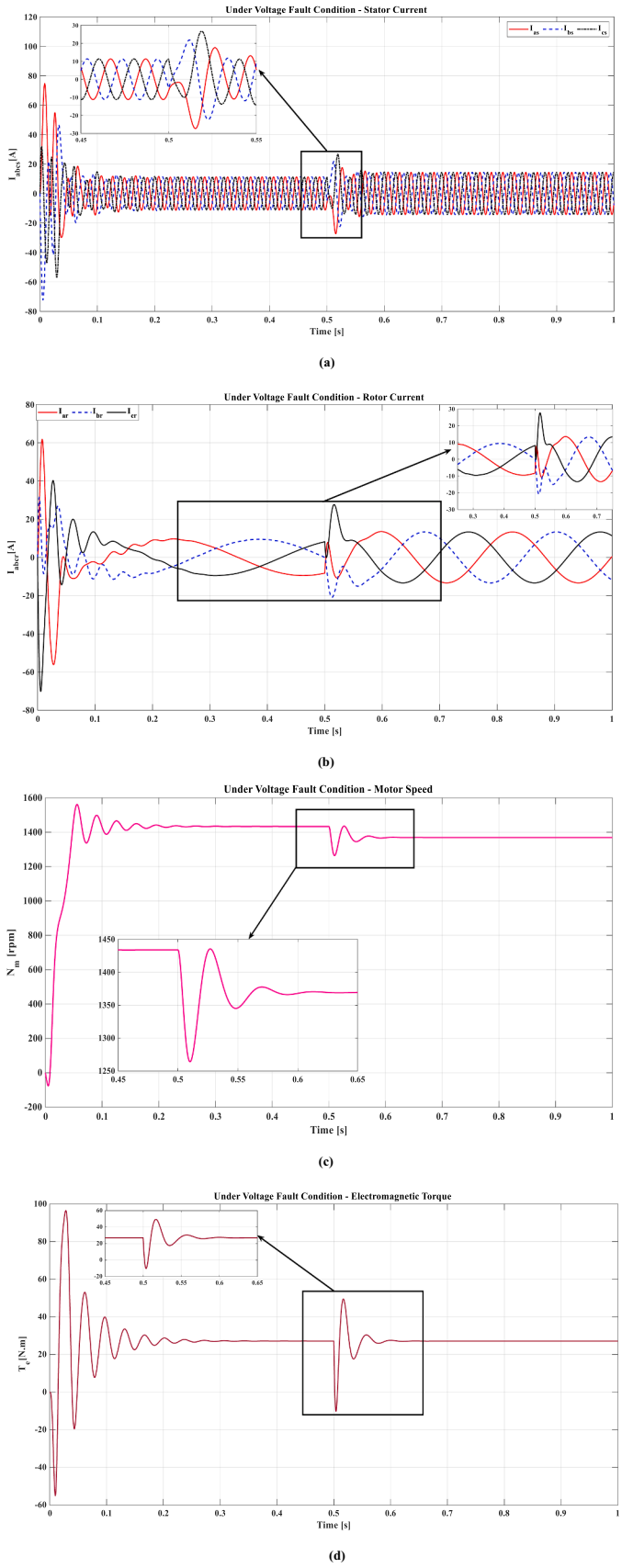


Fig. 8. Waveforms of IM during UV Fault Condition, (a) Stator Current, (b) Rotor Current, (c) Motor Speed, and (d) Torque.

simulated and modelled for the fault cases between 0 and 0.5 s while it was working normally and fault-free. The corresponding faults started at 0.5 s and persisted throughout the one-second simulation interval.

3.1.1. *Broken rotor bar fault condition*

Within the BRB fault condition, it's noticed that the stator and rotor currents operate abnormally, as illustrated in Fig. 5a and b. In contrast to a healthy condition, there are clearly disturbances in the electrical waveforms. These distortions show up as random oscillations and swings in current waveforms range from (- 40A to + 40A). The consequences of the BRB fault on speed and torque waveforms are seen in Fig. 5c and d. Due to the distribution of electromagnetic torque inside, the motor becomes disrupted in the BRB fault condition, there are visible abnormal oscillations and regular variations in both torque and speed waveforms, in which speed oscillations range from (1700 to 1050 rpm) and torque oscillations range from (- 45 to + 80 N.m).

3.1.2. *Inter-turn short circuit fault condition*

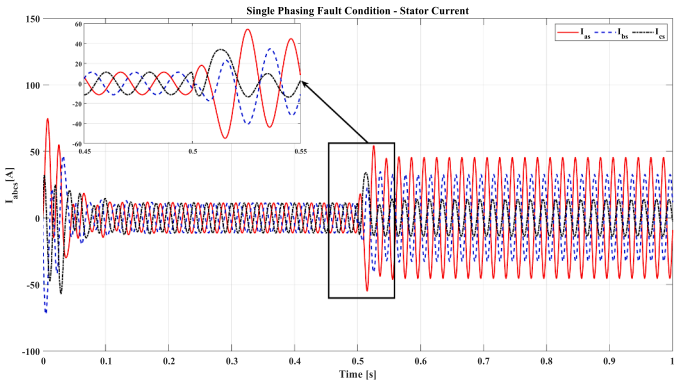
The impact of ITSC faults on the stator and rotor current waveforms is illustrated in Fig. 6a and b. Due to circulating currents in the shorted turns of the stator, the current waveforms exhibit unbalanced oscillations and distorted amplitude ranges, varying from (−50 to +50 A for stator currents and −25 to +25 A for rotor currents), in contrast to those observed under healthy conditions. An imbalance in the electromagnetic field arises from a significant increase in the impacted duration. The ITSC fault significantly influences both motor speed and electromagnetic torque waveforms, as demonstrated in Fig. 6c and d. Continuous torque fluctuations ranging from −20 to + 60 N·m, manifested as irregular ripples, are produced by the ITSC fault, which alters the electromagnetic distribution. The motor's uneven internal forces result in an unstable speed waveform, leading to unexpected swings in rotational speed ranging from 1320 to 1560 rpm.

3.1.3. *Over voltage fault condition*

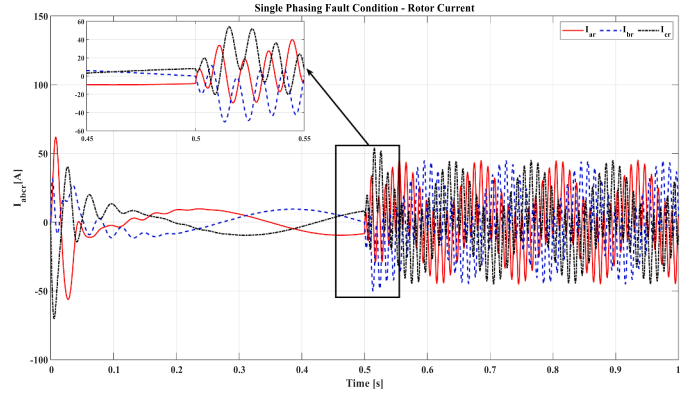
The stator and rotor currents provide an unusual behavior under the OV fault condition, as illustrated in Fig. 7 a and b. Partial saturation of the magnetic field in the core is caused by a significant rise in current amplitude, which commonly exceeds the rated values. Extreme peaks and distortion of the waveform result from this. When OV continues, current waveforms illustrate imbalances between phases that range from (−50 to +50 A for stator currents, and −40 to +40 A for rotor currents) and vary from their sinusoidal shape. This fault can also lead to quicker insulation degradation and excessive heating. An OV condition has a significant effect on motor's speed and torque. Fig. 7c shows how OV fault results in speed oscillations. This fault may lead motor speed to increase initially. Still, with time, it will begin to show indications of instability in its rotational speed, fluctuating about its normal operating point, which ranges from (1150 to 1750 rpm). Fig. 7d also shows how the increased voltage produces a slight increase and oscillations in instantaneous torque, ranges from (−80 to +120 N.m), which can initially appear advantageous. Yet, due to phase imbalance and magnetic saturation, this quickly results in periodic torque oscillations.

3.1.4. *Under voltage fault condition*

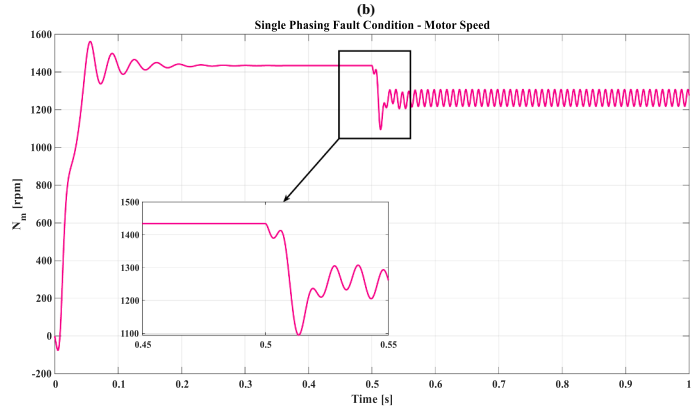
A uniform UV across all 3 phases directly affects the stator and rotor currents, as illustrated in Fig. 8a and b. Due to the lower input voltage, the overall current amplitude drops, but the balance of phase is maintained, ranging from (−16 to + 16 A). While there is a noticeable reduction in amplitude, they are kept almost sinusoidal and symmetrical. This reduction indicates that the electromagnetic force of the motor becomes weaker. The reduced input power reduces the motor performance regardless of waveform symmetry. UV fault reduces the electromagnetic torque produced by the motor to 25 N.m, as seen in Fig. 8d. The capability of the motor to produce torque is directly affected by a UV fault, even though it remains balanced. With average and instantaneous torque values under those of normal operation, the torque waveform



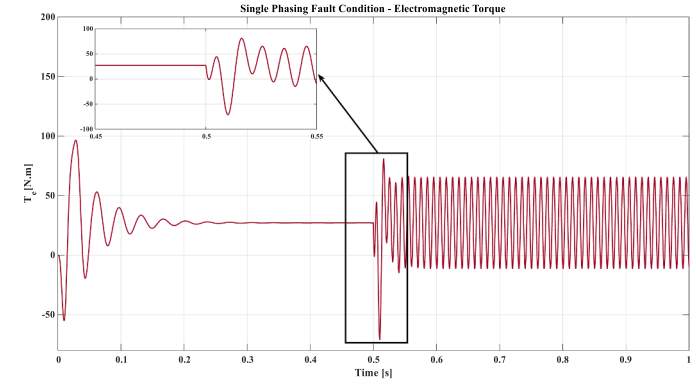
(a)



(b)



(c)



(d)

Fig. 9. Waveforms of IM during SPHF Fault Condition, (a) Stator Current, (b) Rotor Current, (c) Motor Speed, and (d) Torque.

Table 2
Dataset size and fault classification summary.

Fault Class	No-Load (NL)	Half-Load (HL)	Full-Load (FL)	Total Samples
Healthy	150,000	150,000	150,000	450,000
BRB	150,000	150,000	150,000	450,000
ITSC	150,000	150,000	150,000	450,000
OV/UV	150,000	150,000	150,000	450,000
SPH	150,000	150,000	150,000	450,000
Total	**750,000**	**750,000**	**750,000**	**2250,000**

looks stable but attenuated in this condition. Especially at starting or high-demand conditions, this reduction limits motor capacity to manage full loads, resulting in a drop in overall efficiency or a slowed mechanical reaction. Though the torque waveform indicates no notable oscillations, this doesn't necessarily signify that the situation is okay. Reduction in motor speed to 1350 rpm as shown in Fig. 8c and thermal

Table 3
Parameters of NN model.

NN Model Parameter	
Preset	[Medium Neural Network]
Number of fully connected layers	[1]
First Layer Size	[25]
Activation	[ReLU]
Iteration Limit	[1000]
Regularization strength (Lambda)	[0]

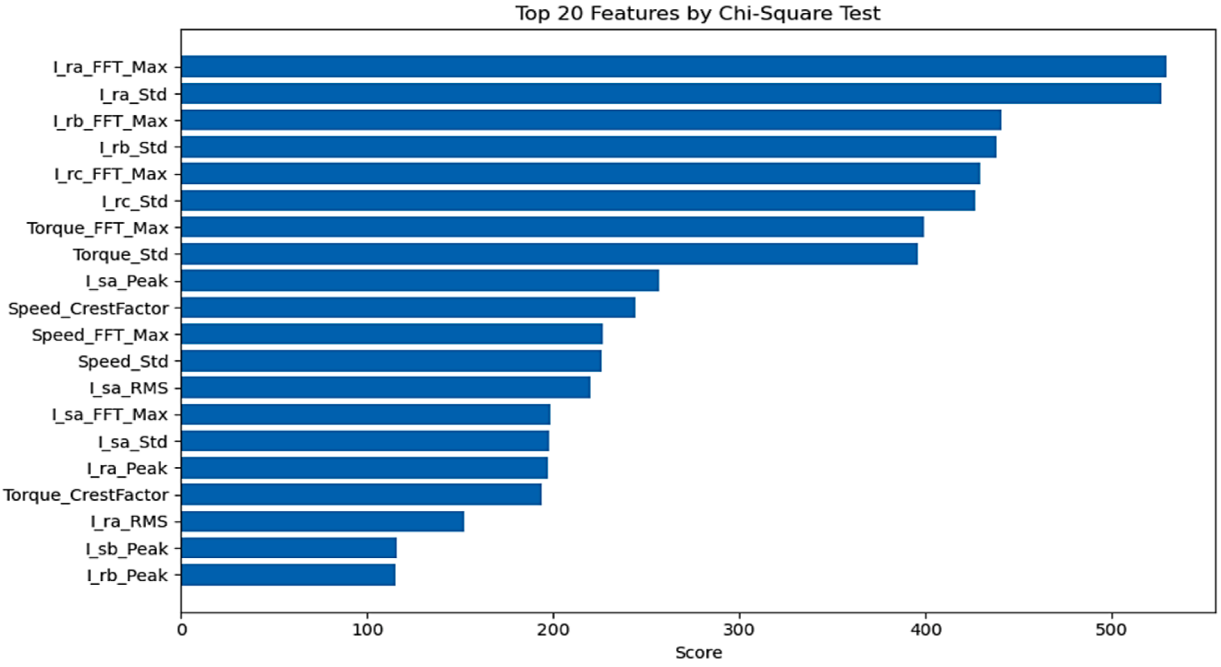


Fig. 10. Feature importance scores.

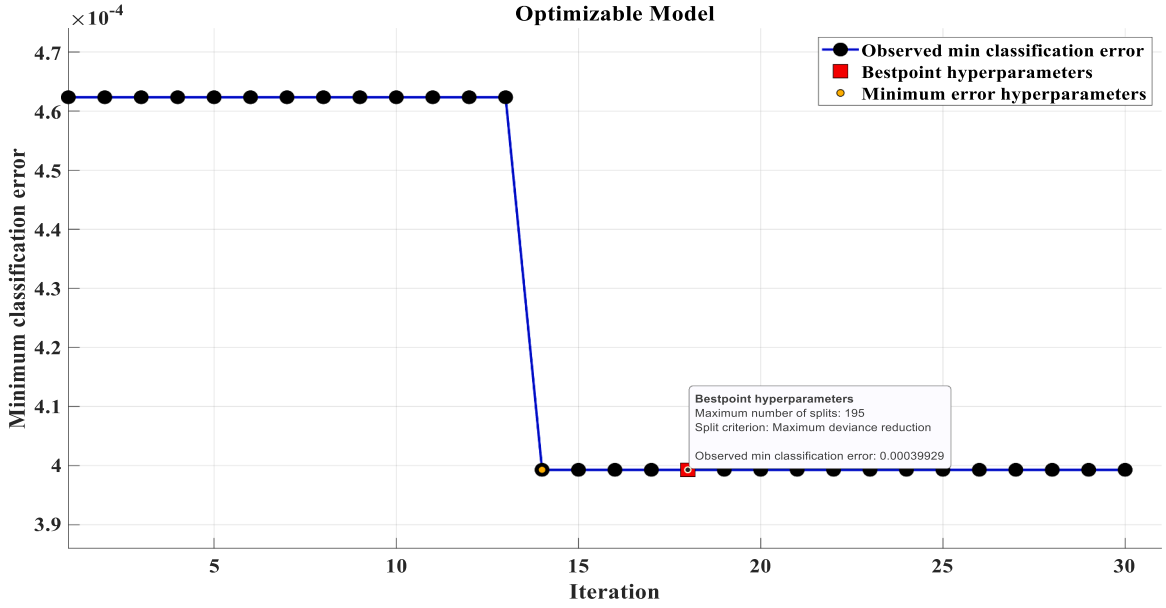
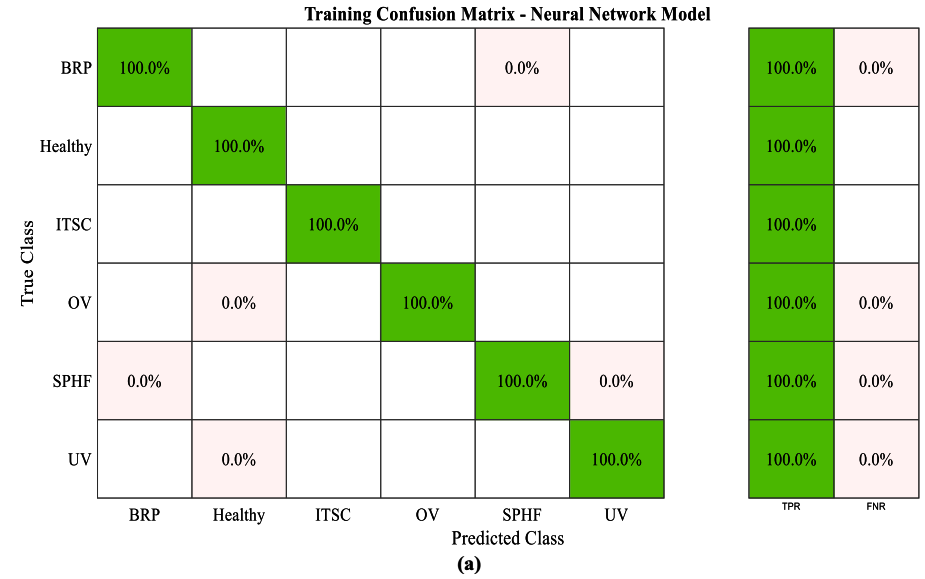
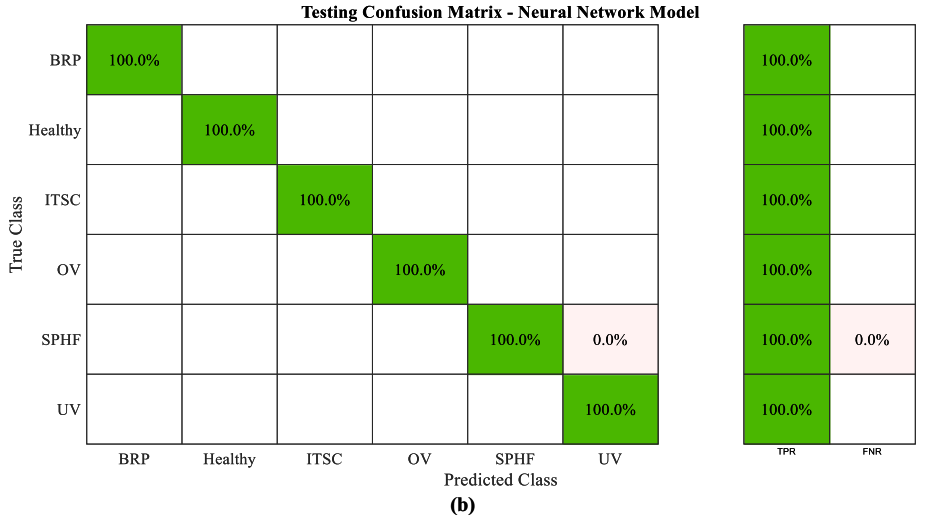


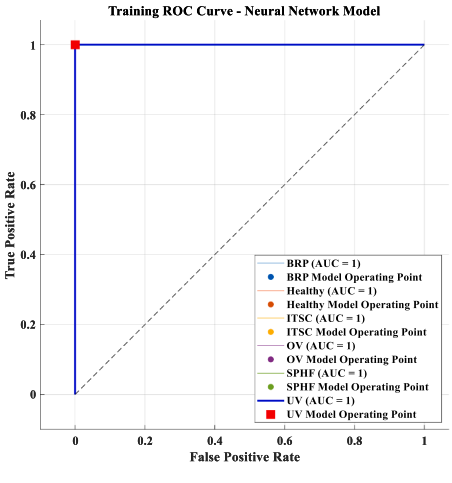
Fig. 11. Sample of hyperparameter optimization.



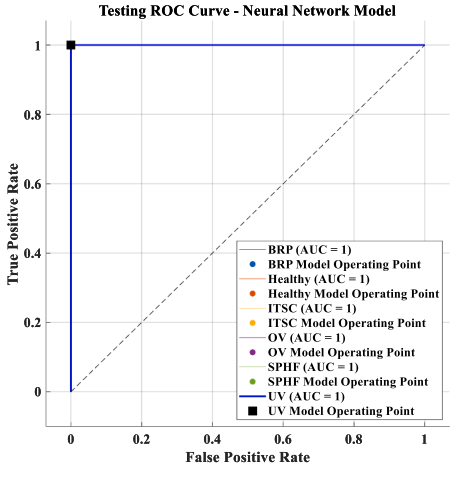
(a)



(b)



(c)



(d)

Fig. 12. Evaluation of NN model based on, (a) Training CM, (b) Training ROC, (c) Testing CM, and (d) Testing ROC.

Table 4
Parameters of KNN model.

KNN Model Parameter	
Preset	[Fine KNN]
Number of neighbours	[1]
Distance Metric	[Euclidean]
Distance Weight	[Equal]

stress brought on by prolonged undervoltage can eventually affect the system's dependability and efficiency.

3.1.5. Single phasing fault condition

The motor suffers from a significant stator and rotor current imbalance in SPH fault condition, as illustrated in Fig. 9a and b, in which one phase is shorted to ground. While the other two phases carry diminished or distorted currents, the faulty phase pulls a dramatic surge in current waveforms, ranging from (−60 to +60 A). This produces a very lopsided waveform with obvious harmonic distortions, particularly on the impacted phase. The loss of the currents sinusoidal shape puts the insulation system at risk and potentially harms delicate electronics. Motor speed instability results from the strong current imbalance, which disturbs the rotating magnetic field. The speed waveform starts to fluctuate ranges from (1200 to 1300 rpm) and falls below the nominal values, as seen in Fig. 9c. The motor may find it difficult to sustain its rotational speed under heavy load conditions, and if the problem continues, it may even stall. A major effect of the SPH defect is torque. The torque waveform as seen in Fig. 9d becomes irregular and unstable, and the motor shows a sudden decline in instantaneous torque ranges from (−5 to +60 N.m). This is brought about by the distorted magnetic field and the unequal distribution of current, which induce mechanical vibrations and significantly reduce drive capacity. If the issue is not fixed immediately, it could result in mechanical failure or complete motor shutdown.

3.2. Dataset preparation and strict data isolation protocol

The comprehensive IM model described in Section 3.1 was created using the MATLAB/Simulink environment to generate high-fidelity time-series data for this study. The model was tested under healthy state as well as other distinct defective states, considering three different load scenarios: FL, HL, and NL as shown in Table 2.

3.2.1. Independent simulation and data generation

To ensure source data independence and avoid any leakage, each operational condition (fault category and load level) was simulated using a completely independent simulation run. Time-series data were gathered for 1 s each scenario to get readings of stator currents (I_{sa}, I_{sb}, I_{sc}), voltages (V_{sa}, V_{sb}, V_{sc}), rotor currents (I_{ra}, I_{rb}, I_{rc}), motor speed (Nm), and electromagnetic torque (T_e). A sampling frequency of 150 kHz was used to precisely capture weak, high-frequency transient components and sideband harmonics associated with the faults. This yielded 150,000 samples per class per second. A thorough dataset of 450,000 samples for each motor scenario was created by combining samples from all load conditions (NL, HL, and FL).

3.2.2. Leakage prevention and statistical validation

To strictly avoid data leaking, a hierarchical data splitting protocol was used. The time-series data were divided into rigorously non-overlapping blocks:

1. **Development Set (80%):** Used solely for training, 5-fold cross-validation, Chi-Square feature ranking, and hyperparameter tuning using Grid Search (Section 4.1 and 4.2).
2. **Final Hold-Out Test Set (20%):** This set was fully separated from the training pipeline and was only used once for the final performance assessment.

The separation ensures that no information from the test distribution influenced the feature selection or the optimization process. While it is acknowledged that the simulation-generated data is inherently noise-free which enhances class separability this strict structural isolation guarantees that the reported near-optimal accuracy is a result of legitimate model generalization rather than artificial data leakage.

Furthermore, the entire training and testing procedure was performed across six independent runs using various random seeds and shuffled partitions to ensure the results' stability and account for statistical changes. The given results (mean accuracy ± standard deviation) in Section 5.1 (Table 14) demonstrate that the high performance is statistically reproducible and not due to a favorable random split.

4. Development of machine learning models

ML models were created using the MATLAB Classification Learner toolbox. The collected features were used to train different algorithms, including NN, DT, KNN, SVM and Ensemble algorithms. The MATLAB Classification Learner application was used to construct ML models. This application offers an easy-to-use interface for training, assessing, and optimizing classification models using the extracted features. The first step was importing the dataset gathered from the motor modelling procedure into MATLAB Simulink. This dataset was recorded in Excel sheets and included values for motor speed, rotor currents, stator currents, stator voltages and electromagnetic torque. The following critical phase was to prepare the data for model training after it had been imported into the Classification Learner application. In accordance with normal ML procedures, the data was divided into 80% training and 20% testing sets. While the testing set was set aside to see how well the trained models performed on unknown data, the training set was used to train the models. This division ensures that the models can generalize effectively to new, unknown data and are not too optimized for the data they were trained on. Cross-validation folds were used to further improve the model generalizability and lower the possibility of overfitting, which occurs when a model performs well on the training set but badly on fresh data. Dividing the training data into multiple smaller groups (or "folds") is known as cross-validation. A subset is used to train the model, which is then tested on the remaining folds several times.

In this study, a comparative performance analysis is conducted using five distinct machine learning architectures: ANN, SVM, KNN, DT, and Ensemble Learning. Rather than combining these models into a hybrid system, each model is trained and evaluated independently. This benchmarking approach is designed to rigorously determine the most effective algorithm for classifying complex electrical and mechanical faults under varying load conditions, ensuring the proposed diagnostic framework is both accurate and computationally efficient. All models are trained and tested using the same dataset, feature set, and validation strategy to ensure a fair and unbiased comparison.

4.1. Feature extraction and selection

To modify and select the most valuable data from unprocessed signals like stator currents, voltages, speed, and torque, the feature extraction method is a vital stage in the ML process for IM fault detection. By removing redundancy and noise, this procedure reduces the dimensionality of the dataset, improving computational efficiency and increasing model accuracy. It indicates that models focus on related patterns, improving differentiation and generalization between motor states that are healthy and those that are faulty. Model accuracy can be affected by underfitting or overfitting caused by improper feature selection. Features were thoroughly assessed across datasets to reduce bias and remedy this [27]. The most important features were found using

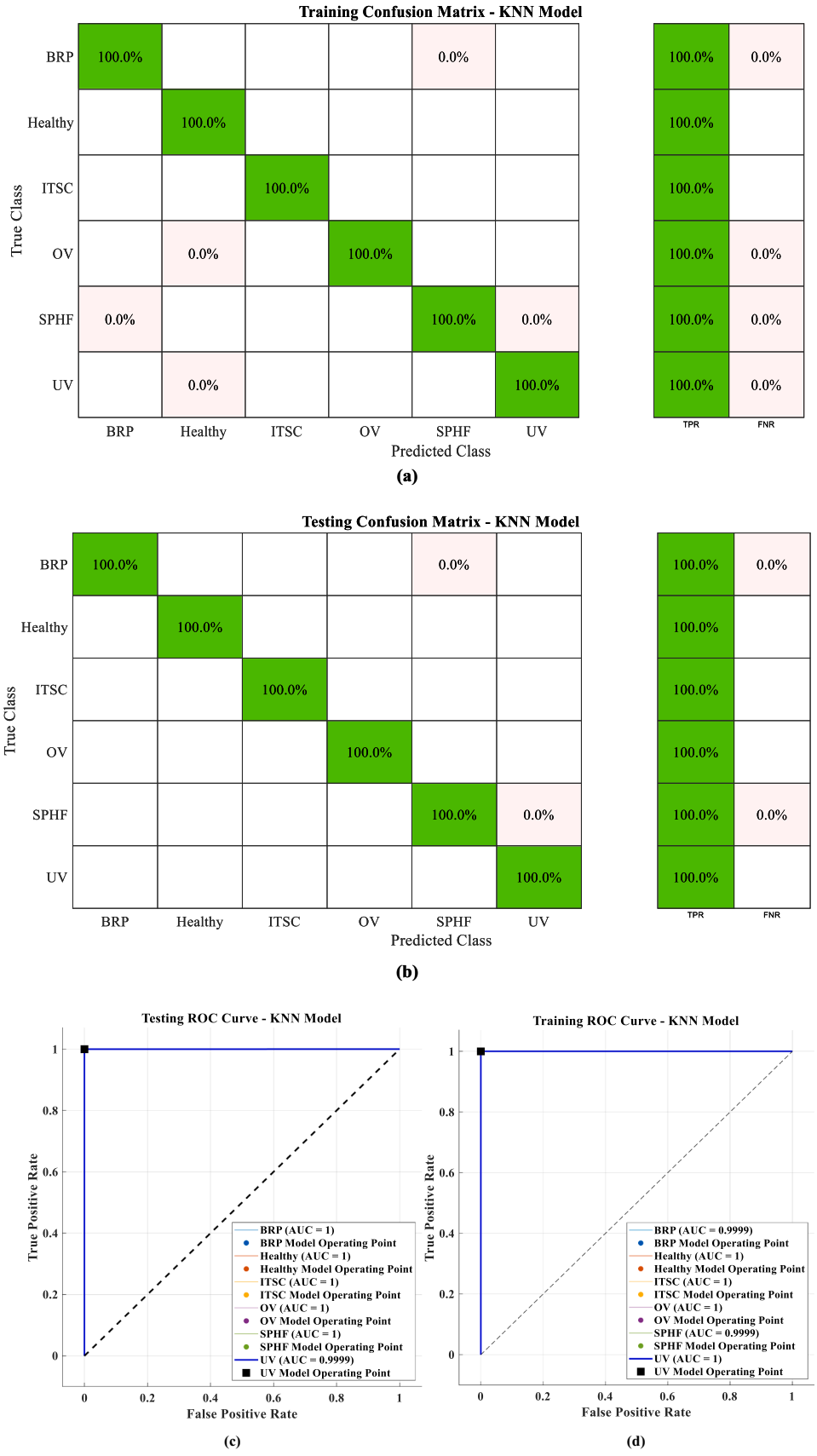


Fig. 13. Evaluation of KNN model based on, (a) Training CM, (b) Training ROC, (c) Testing CM, and (d) Testing ROC.

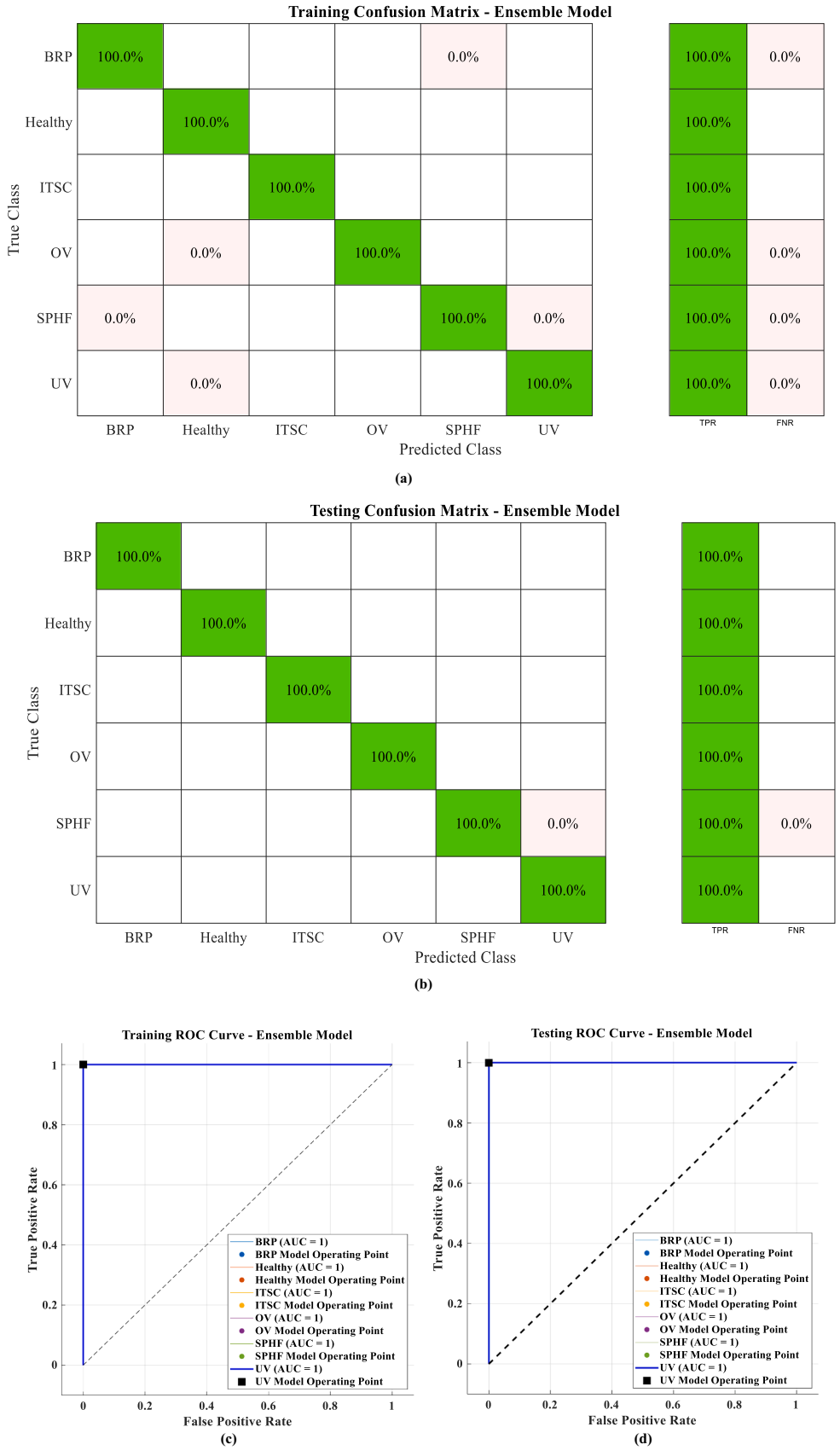


Fig. 14. Evaluation of Ensemble Model based on, (a) Training CM, (b) Training ROC, (c) Testing CM, and (d) Testing ROC.

Table 5
Parameters of ensemble model.

Ensemble Model Parameter	
Preset	[Subspace KNN]
Ensemble Method	[Subspace]
Learner Type	[Nearest neighbours]
Number of Learners	[30]
Subspace dimension	[4]

Table 6
Parameters of SVM model.

SVM Model Parameter	
Preset	[Cubic SVM]
Kernel Function	[Cubic]
Kernel Scale	[Automatic]
Box constraint level	[1]
Multi-class coding	[One-vs-One]

MATLAB Chi-Square (Chi2) approach, as depicted in Fig. 10, which improved the accuracy and resilience of ML models under various operating conditions.

Fig. 10 illustrates the results of the Chi-Square feature importance analysis for the top 20 extracted features. The analysis reveals a strategic diversity in the feature set, spanning both time and frequency domains. Notably, the maximum frequency components (FFT_Max) and statistical dispersion (Std) of the rotor currents (Ira, Irb, Irc) and electromagnetic torque (Te) emerged as the most significant predictors. Other vital features such as the Stator Current Peak (Isa_Peak) and Speed Crest Factor (Speed Crest Factor) also contribute to the multidimensional decision boundaries of the ML classifiers. This rigorous selection process ensures that only the most relevant physical signatures are utilized, thereby enhancing model generalization and reducing computational redundancy.

4.2. Hyperparameter optimization

Model parameters were fine-tuned using optimization approaches to ensure optimal accuracy and low error rates. This step included Grid search, Bayesian optimization, and Cross-validation to prevent overfitting.

There are many parameters to modify in a ML model, which can increase the model's performance. Hyperparameter tuning is the most effective way for evaluating a classifier's effectiveness by executing a variety of parameter combinations. Overfitting, a fundamental ML problem, occurs when a classifier is evaluated using training data. Overfitting occurs when a model performs poorly on test data but exceptionally well on rainy data. To optimize hyperparameters, cross-validation is employed in conjunction with the grid search approach. The grid search method is a technique for determining the optimal parameters of a classifier such that a model may accurately predict some unlabeled data. The Grid Search method is used to tune hyperparameters that cannot be learned directly through the training process. The classification model has numerous hyperparameters, and determining the appropriate combination of these parameters is a difficult task. Grid Search is one of the most effective strategies for this purpose [28]. This optimization technique deeply increased model accuracy and reliability across many operating conditions, allowing for precise and effective fault detection, as illustrated in Fig. 11.

4.3. Model evaluation

Receiver Operating Characteristic (ROC) curve and Confusion Matrix (CM), precision, recall, and F1-score were among the key indicators used to evaluate the effectiveness of ML models in classification and detection of faults. To find specific misclassifications, the CM indicates true and false predictions. ROC curve establishes the model capability to classify between healthy and fault types across different thresholds, where the area under the curve (AUC) represents overall performance of ML model. When mixed, CM and ROC provide a strong evaluation mechanism that ensures the ML-based fault detection system's high performance, accuracy, reliability, and effectiveness. A more detailed view of a predictive model efficiency can be obtained with precision, recall, and F1-score indicators in classification scenarios, where models predict categorical outputs. Precision model is determined by the percentage of true positive predictions among all positive predictions. It is calculated as the ratio of true positives to the total of false positives and true positives. The percentage of accurate positive predictions in datasets that are true positive predictions is known as Recall. Recall shows how many of the real positive samples it can accurately detect. A high recall means that the model is detecting most positive outcomes. F1-score is the harmonic means of precision and recall. It balances between precision and recall by combining the two into a single statistic. The F1-score is a useful indicator for unbalanced datasets, where precision and recall may be skewed, because it takes into consideration both false positives and false negatives [29].

The 5-fold cross-validation technique was employed exclusively during the hyperparameter optimization phase (Grid Search) on the Training Set. Crucially, cross-validation was not used for the final assessment. The definitive performance metrics and generalization capabilities reported in Section 5 were derived from the model's assessment on the 20% independent Hold-Out Set, which remained untouched throughout the training and tuning procedures.

5. Results and discussion

In this study, an EV with the specifications listed in [30], was considered. The chosen scenario is a 2022 Star EV Sirius 2 + 2 golf cart driven by a 5.4 HP AC IM with thermal protection. The investigation focused on motor operation, with the system being investigated under constant speed driving conditions. A thorough simulation model of the motor was created to depict its dynamic activity under both healthy and defective conditions. Using the acquired data, various ML techniques were deployed and assessed for fault detection. The results from these models are provided and discussed in the sections below.

The current framework utilizes a 5.4 HP motor model, which served as a robust, well-defined platform for establishing the performance ceiling of our proposed diagnostic methodology. The rationale for this selection was rooted in modeling fidelity and the need to confirm the efficacy of our feature set against fundamental fault signatures. While this motor falls within the light EV category (e.g., golf carts, utility vehicles), the principles of fault signature detection remain physically consistent across varying motor sizes. We assert that the optimized ML framework developed here is fundamentally scalable. The primary challenge in applying this framework to high-power EV traction motors lies in managing increased thermal and mechanical dynamics. Future research will therefore focus on validating this optimized model on experimental data generated by higher-power IMs to fully assess the model's robustness and practical generalizability across the full spectrum of Electric Vehicle applications.

5.1. Machine learning results

5.1.1. Neural network model

The NN network model (parameters specified in Table 3) capability to identify and categorize faults was evaluated using essential performance metrics obtained from the training and testing outcomes. The results of the training and testing stages are illustrated, as well as the model evaluation utilizing CM, ROC curve, F1 score, precision, and recall for every operating state, as indicated in Fig. 12 and Table 8. The results show that the model effectively identified and classified every

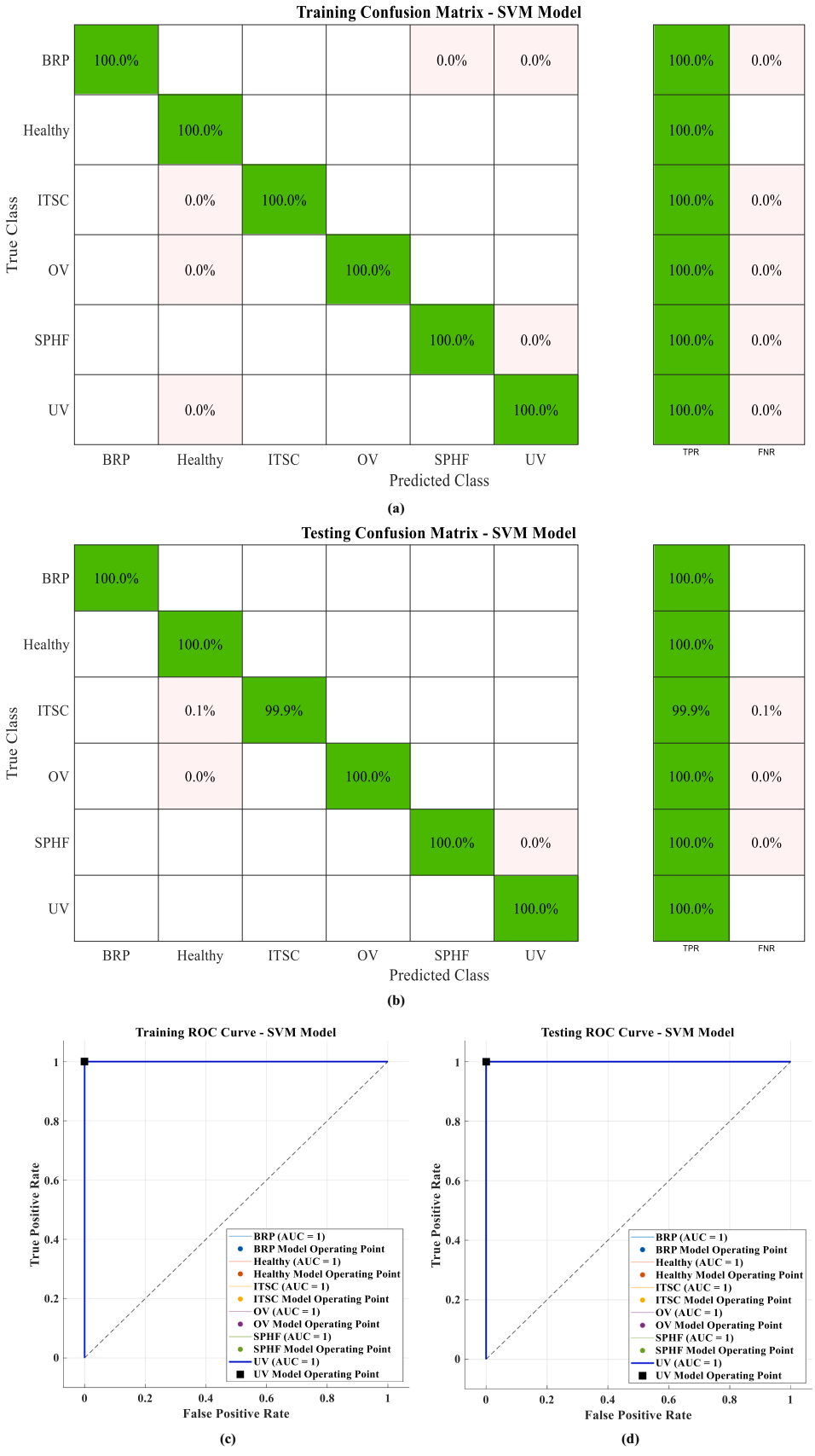


Fig. 15. Evaluation of SVM model based on, (a) Training CM, (b) Training ROC, (c) Testing CM, and (d) Testing ROC.

Table 7
Parameters of DT model.

DT Model Parameter	
Maximum number of splits	[178]
Split criterion	Maximum deviance reduction

type of fault with an accuracy close to 100% (99.99%) and an error of 1.156×10^{-5} as mentioned in Table 13, even with data that had not been seen before, confirming its generalization capacity and reliability.

5.1.2. K-Nearest neighbours model

The efficacy of the KNN model (parameters in Table 4) for fault detection and classification by the same tools will be assessed. The model underwent initial training and testing under diverse loading and operational situations, as depicted in Fig. 13 and Table 9. The outcomes show that the model successfully detects and classifies each type of fault with an accuracy of 99.996% and an exceedingly low error of 3.4687×10^{-5} as depicted in Table 13.

5.1.3. Ensemble model

To determine the Ensemble Model ability for fault detection and classification, it was trained, tested, and then evaluated using the previously described performance methods. As can be seen from Fig. 14 and Table 10, the model with parameters seen in Table 5 performed almost flawlessly in recognizing and categorizing every fault type, producing extremely precise results under various operating scenarios with accuracy (99.99%) and error (4.625×10^{-5}) as mentioned in the comparison between ML models listed in Table 13.

5.1.4. Support vector machine model

The same metrics as previously stated were used to train, test, and assess the SVM model with parameters depicted in Table 6. The model showed exceptional accuracy in detecting and classifying the faults, as evidenced by the results achieved under diverse operating conditions. The SVM model achieves accuracy of 99.9% and an error of 9.481×10^{-4} as depicted in Fig. 15, Tables 11, and Table 13.

5.1.5. Decision tree model

The DT model for fault detection and classification accuracy was validated through the recorded data under different operating situations. The DT with characteristics outlined in Table 7 exhibited efficient processing and interpretable decision paths, although its structure is simpler than that of other models. This model achieves an accuracy (99.97%) with a very small percentage of error (3.122×10^{-4}) as illustrated in Fig. 16,Tables 12, and Table 13.

A comparative analysis of the five assessed ML models for fault detection and classification is provided in Table 13. Based on accuracy, computational efficiency, and error rates, the evaluation positions the NN model in the first rank owing to its exceptional accuracy and minimal error rate. The subsequent models, in descending sequence, are KNN, Ensemble, DT, and SVM.

Strong stability and little sensitivity to data reshuffling or initialization are illustrated by the findings shown in Table 14, which reveal remarkably low standard deviation values for all models. Additionally, the very strict 95% confidence intervals indicate that the model performance is constant over multiple executions.

The proposed machine learning framework demonstrates very high classification performance across all tested models, with Artificial Neural Networks (ANN) achieving the highest accuracy. This strong performance can be attributed to several key factors. First, the use of a multi-source feature set, including stator and rotor currents, speed, and torque, provides rich and complementary information about the motor operating condition. This improves class separability and enhances the learning capability of all models. Second, the application of feature selection using the Chi-square method reduces redundancy and improves the quality of the input data, leading to more stable and accurate classification. Third, hyperparameter optimization using grid search ensures that each model operates under near-optimal configurations, maximizing performance and generalization capability. Finally, the use of a large-scale dataset with multiple fault types and varying load conditions significantly improves the robustness of the trained models.

Despite these advantages, several limitations should be acknowledged. Although the simulation-based dataset provides controlled and balanced data, it may not fully capture all complexities of real industrial environments such as sensor noise, parameter uncertainty, and unmodeled dynamics. In addition, while multiple machine learning models were evaluated, their performance may vary when deployed in real-time systems with limited computational resources, especially in embedded or edge applications. Furthermore, the high classification accuracy achieved in this study is based on offline training and testing, which may not directly reflect performance under continuous online monitoring conditions.

In practical applications, additional challenges may arise, including sensor degradation, communication delays, and data drift over time. These factors can affect model reliability and require periodic retraining or adaptive learning strategies. Therefore, future work should focus on real-time implementation, hardware validation, and transfer learning techniques to further enhance industrial applicability.

While the results obtained in this study demonstrate the exceptional diagnostic potential of the proposed ML framework, it is important to acknowledge that the dataset was generated within a high-fidelity simulation environment. Simulation-based data, while mathematically rigorous and effective for capturing fault signatures, lacks the stochastic noise and measurement uncertainties inherent in physical industrial settings. Therefore, this study is positioned as a comprehensive proof-of-concept and a theoretical benchmark. Future research will focus on validating these models using experimental data from a physical induction motor test-bed to ensure the framework's robustness against real-world environmental disturbances and sensor inaccuracies.

The high diagnostic accuracy achieved across all ML models (approaching 100% in some cases) warrants a careful discussion regarding potential overfitting. In this study, overfitting was mitigated through a multi-stage validation protocol, including 5-fold cross-validation and the use of an independent hold-out test set. The results remained consistent across both sets, indicating robust generalization. The primary reason for such high performance lies in the feature engineering stage; the Chi-Square analysis confirmed that frequency-domain features of the rotor current and torque provide highly discriminative signatures that clearly separate fault categories. However, it is important to note that these results represent a theoretical benchmark within a controlled, high-fidelity simulation environment. In real-world applications, factors such as measurement noise, sensor degradation, and stochastic load fluctuations might reduce these margins, which will be the focus of our future experimental validation.

Despite the robust statistical validation, it is crucial to acknowledge the inherent limitations of this study. The current framework relies exclusively on simulation-only data generated under steady-state, constant speed conditions. Consequently, it does not fully replicate the stochastic complexities of real-world EV powertrains, particularly the presence of high-frequency inverter switching harmonics, sensor measurement noise, and dynamic speed variations.

6. Conclusion and future work

This study established a comprehensive, comparative machine learning framework for the accurate diagnosis of Induction Motor (IM) faults in EV platforms. Utilizing a high-fidelity MATLAB/Simulink environment, a high-resolution, multi-fault dataset was generated at a 150 kHz sampling frequency. This dataset specifically incorporated critical fault conditions namely BRB, ITSC, OV, UV, and SPH across full, half, and no-load scenarios.

Fig. 16. Evaluation of DT model based on, (a) Training CM, (b) Training ROC, (c) Testing CM, and (d) Testing ROC.

Table 8
Evaluation of NN Model during different load conditions.

Evaluation of NN Model															
Fault Type	Full Load Condition					Half Load Condition					No Load Condition				
	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.
BRP	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
Healthy	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
ITSC	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
OV	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
SPHF	1	0.999833333	0.99991666	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
UV	0.999833361	1	0.999916674	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%

Table 9
Evaluation of KNN Model during different load conditions.

Evaluation of KNN Model															
Fault Type	Full Load Condition					Half Load Condition					No Load Condition				
	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.
BRP	1	0.999833333	0.99991666	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
Healthy	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
ITSC	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
OV	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
SPHF	0.999833333	0.999833333	0.999833333	100%	100%	0.99980008	1	0.99990003	100%	100%	1	1	1	100%	100%
UV	0.999833361	1	0.999916674	100%	100%	1	0.9998	0.99989999	100%	100%	1	1	1	100%	100%

Table 10
Evaluation of Ensemble Model during different load conditions.

Evaluation of Ensemble Model															
Fault Type	Full Load Condition					Half Load Condition					No Load Condition				
	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.
BRP	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
Healthy	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
ITSC	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
OV	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
SPHF	1	0.999833333	0.99991666	100%	100%	0.9998	0.99960008	0.99970003	100%	100%	1	1	1	100%	100%
UV	0.999833361	1	0.999916674	100%	100%	0.99960008	0.9998	0.99970003	100%	100%	1	1	1	100%	100%

Table 11
Evaluation of SVM Model during different load conditions.

Evaluation of SVM Model															
Fault Type	Full Load Condition					Half Load Condition					No Load Condition				
	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.
BRP	1	1	1	100%	100%	1	1	1	100%	100%	1	1	1	100%	100%
Healthy	0.998253057	1	0.999125765	100%	100%	0.999001248	1	0.999500375	100%	100%	0.999666889	1	0.999833417	100%	100%
ITSC	1	1	1	100%	100%	1	0.999403579	0.9997017	100%	100%	0.993932999	0.999469496	0.996693559	100%	99%
OV	1	1	1	100%	100%	1	1	1	100%	100%	1	0.999733333	0.999866649	100%	100%
SPHF	1	0.999833333	0.99991666	100%	100%	1	0.9990002	0.99949985	100%	100%	0.990224571	0.999466667	0.994824154	99%	99%
UV	0.999833361	1	0.999916674	100%	100%	0.999000799	0.9998	0.99940024	100%	100%	0.99891687	0.983995732	0.991400161	99%	100%

Table 12
Evaluation of DT Model during different load conditions.

Evaluation of Decision Tree Model													
Fault Type	Full Load Condition					Half Load Condition				No Load Condition			
	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.	Precision	Recall	F1-score	Training Acc.	Testing Acc.
BRP	0.998336106	1	0.999167361	100%	100%	1	1	1	100%	100%	0.999733191	0.999733191	0.999733191	100%	100%
Healthy	1	1	1	100%	100%	1	1	1	100%	100%	0.999333999	1	0.99966889	100%	100%
ITSC	0.999869944	1	0.999934968	100%	100%	1	0.999602386	0.999801153	100%	100%	1	0.999734748	0.999867356	100%	100%
OV	0.999833361	1	0.999916674	100%	100%	1	0.9996	0.99979996	100%	100%	0.999466809	0.999733333	0.999600053	100%	100%
SPHF	0.999833333	0.999833333	0.999833333	100%	100%	0.999200639	0.99980004	0.999500025	100%	100%	0.999466667	0.999466667	0.999466667	100%	100%
UV	0.999833361	1	0.999916674	100%	100%	0.99980004	1	0.99990001	100%	100%	0.99973312	0.999199787	0.999466382	100%	100%

Table 13
Comparison between ML models based on percentage of accuracy and error.

ML Model	Accuracy %	Percentage of Error %
NN Model	99.998844%	0.000011562431%
KNN Model	99.996531%	0.000034687294%
Ensemble Model	99.995375%	0.000046249725%
SVM Model	99.905188%	0.000948119371%
DT Model	99.968781%	0.000312185646%

Table 14
Statistical evaluation of the ML Models over six independent runs using new unseen data.

ML Model	Mean Accuracy (%)	Std (±)	95% Confidence Interval	Min Accuracy (%)	Max Accuracy (%)
NN Model	99.99812	± 0.00042	[99.99721 – 99.99889]	99.9975	99.9988
KNN Model	99.99584	± 0.00063	[99.99456 – 99.99679]	99.9948	99.9967
Ensemble Model	99.99497	± 0.00071	[99.99346 – 99.99601]	99.9940	99.9961
SVM Model	99.90344	± 0.00112	[99.90122 – 99.90589]	99.9015	99.9057
DT Model	99.96712	± 0.00054	[99.96614 – 99.96831]]	99.9663	99.9680

Five distinct ML algorithms (ANN, KNN, SVM, DT, and Ensemble) were independently evaluated and optimized using Grid Search. A rigorous hierarchical data isolation protocol was enforced, successfully eliminating bias and preventing information leakage during the multi-domain feature selection and validation phases. Driven by the high discriminative power of the extracted time-frequency features (particularly from rotor currents and electromagnetic torque), the models achieved near 100% classification accuracy.

While these exceptional results serve as a theoretical upper-bound benchmark demonstrating the algorithms' maximum potential on noise-free simulated data, they provide a vital proof-of-concept for early fault detection in EV powertrains. Future research will focus on overcoming real-world implementation challenges such as industrial sensor noise and coupled faults by validating this framework using experimental Hardware-in-the-Loop (HIL) testbeds, alongside exploring deep learning architectures to build a fully robust, physical diagnostic system.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

CRediT authorship contribution statement

Mohamed Sharawy: Writing – review & editing, Supervision, Methodology, Conceptualization. **Adel El-Nahas:** Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation. **M.A. Alahmar:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Fahd A. Banakhr:** Writing – review & editing, Funding acquisition. **Mohamed I. Mosaad:** Writing – review & editing, Visualization, Funding acquisition. **Mohamed Selmy:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

The authors declare that the paper is original and not under review in

any other journal.

Appendix

Parameters of IM

Parameter	Value
Rated Power	(5.4 HP)
Rated Voltage	(400 v) (Line Voltage)
Frequency	(50 Hz)
Number of Poles	(4 Poles)
Rotor Type	(Squirrel Cage)
Stator Resistance (R_s)	(1.405 Ω)
Rotor Resistance (R_r)	(1.395 Ω)
Inertia Constant (J)	(0.0131 kg.m^2)
Rated Speed (N)	(1430 rpm)
Stator Inductance (L_s)	(0.005839 H)
Rotor Inductance (L_r)	(0.005839 H)
Mutual Inductance (L_m)	(0.1722 H)
Total number of rotor bars (N_b)	(32 bars)
Number of broken rotor bars (n_{bb})	(5 bars)
Ratio of short-circuited turns (k_{sf})	(5%)
Short-circuited fault resistance (R_f)	(0.1 Ω)

Definition of parameters used in IM model

θ = Electrical rotor angle.	Ψ_{ds} = Stator flux linkage in d − axis
Ψ_{qs} = Stator flux linkage in q − axis	Ψ_{dr} = Rotor flux linkage in d − axis
Ψ_{qr} = Rotor flux linkage in q − axis	V_{ds} = Stator voltage in d − axis
V_{qs} = Stator voltage in q − axis	V_{dr} = Rotor voltage in d − axis
V_{qr} = Rotor voltage in q − axis	R_s = Stator resistance
R_r = Rotor resistance	X_{ls} = Stator leakage reactance
X_{lr} = Rotor leakage reactance	X_{ml} = Magnetizing reactance
ω_e = Electrical angular speed of supply	ω_r = Electrical angular speed of rotor
ω_b = Base angular speed	Ψ_{mq} = Mutual flux linkage in q − axis
Ψ_{md} = Mutual flux linkage in d − axis	I_{qs} = Stator current component in q − axis
I_{ds} = Stator current component in d − axis	I_{qr} = Rotor current component in q − axis
I_{dr} = Rotor current component in d − axis	T_e = Electromagnetic torque developed (N.m)
P = Number of poles of the machine	J = Moment of inertia of the rotor

Definition of parameters used in ITSC model

I_{sa} = Stator phase A current	I_f = Fault current
i_r = Rotor current	V_{sa} = Stator phase A voltage
R_s = Stator resistance	R_f = Fault winding resistance
R_r = Rotor resistance	λ_{sa} = Stator flux linkage of phase A
λ_f = Fault flux linkage	L_{sabc} = Stator inductance matrix
L_{ff} = Fault inductance	L_r = Rotor inductance
L_{sf}, L_{fs} = Mutual inductance between fault and stator	L_{rf}, L_{fr} = Mutual inductance between fault and rotor
L_{rs}, L_{sr} = Mutual inductance between stator and rotor	k_{sf} = Fault turns ratio coefficient
n_{sf} = Number of shorted turns	R_{sf} = Stator to fault resistance vector
R_{ff} = Total fault resistance	R_{fs} = Fault to stator resistance vector
L_{sar} = Stator to rotor mutual inductance vector	L_{ms} = Stator mutual leakage inductance

References

[1] U. Fesli, M.B. Ozdemir, Electric vehicles: a comprehensive review of technologies, integration, adoption, and optimization, IEEE Access 12 (2024) 140908–140931.

[2] Z. Wang, T.W. Ching, S. Huang, H. Wang, T. Xu, Challenges faced by electric vehicle motors and their solutions, IEEE Access 9 (2021) 5228–5249.

[3] P. Bhatt, H. Mehar, M. Sahajwani, Electrical motors for electric vehicle – A comparative study, SSRN Electron. J. (2019).

[4] K. Yatsugi, S.E. Pandarakone, Y. Mizuno, H. Nakamura, Common diagnosis approach to three-class induction motor faults using stator current feature and support vector machine, IEEE Access 11 (2023) 24945–24952.

[5] S. Karmakar, S. Chattopadhyay, M. Mitra, S. Sengupta, Induction motor and faults. Induction Motor Fault Diagnosis, Power Systems, Springer, Singapore, 2016.

[6] T. Garcia-Calva, D. Morinigo-Sotelo, V. Fernandez-Cavero, R. Romero-Troncoso, Early detection of faults in induction motors—a review, Energies 15 (21) (2022) 7855.

[7] C. Terron-Santiago, J. Martinez-Roman, R. Puche-Panadero, A. Sapena-Bano, A review of techniques used for induction machine fault modelling, Sensors 21 (14) (2021) 4855.

[8] P. Kumar, A.S. Hati, Review on machine learning algorithm based fault detection in induction motors, Arch. Comput. Methods Eng. (2020).

[9] A.F.A.H. Alnuaimi, T.H.K. Albaldawi, An overview of machine learning classification techniques, Bio Web Conf./BIO Web Conf. 97 (4) (2024) 00133.

[10] Y.O. Lee, J. Jo, J. Hwang, Application of deep neural network and generative adversarial network to industrial maintenance: a case study of induction motor fault detection, in: *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA, USA, 2017, pp. 3248–3253.

[11] H. Idris, G. Loganathan, Induction motor fault monitoring and fault classification using deep learning probablistic neural network, Solid State Technol. 63 (2020) 2196–2213.

[12] S. Esakimuthu Pandarakone, Y. Mizuno, H. Nakamura, A comparative study between machine learning algorithm and artificial intelligence neural network in detecting Minor bearing fault of induction motors, Energies 12 (11) (2019) 2105.

[13] M. Mohamed, E. Mohamed, A.-A. Mohamed, M. Abdel-Nasser, M.M. Hassan, Detection of inter turn short circuit faults in induction motor using artificial neural network, in: *2020 26th Conference of Open Innovations Association (FRUCT)*, Yaroslavl, Russia, 2020, pp. 297–304.

[14] H.D.L. Rações, F.J.T.E. Ferreira, J.M. Pires, C.V. Damásio, Application of different machine learning strategies for current- and vibration-based motor bearing fault detection in induction motors, in: *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, Lisbon, Portugal, 2019, pp. 68–73.

[15] I. Ullah, N. Khan, S.A. Memon, W.-G. Kim, J. Saleem, S. Manzoor, Vibration-based anomaly detection for induction motors using machine learning, Sensors 25 (3) (2025) 773.

[16] P. Pohakar, R. Gandhi, S. Hans, G. Sharma, P.N. Bokoro, Analysis of multiple faults in induction motor using machine learning techniques, e-Prime - Adv. Electr. Eng. Electron. Energy 12 (2025) 101007.

[17] D. Gonzalez-Jimenez, J. del-Olmo, J. Poza, F. Garramiola, I. Sarasola, Machine learning-based fault detection and diagnosis of faulty power connections of induction machines, Energies 14 (16) (2021) 4886.

[18] S. Sobhi, M. Reshadi, N. Zarft, A. Terheide, S. Dick, Condition monitoring and fault detection in small induction motors using machine learning algorithms, Information 14 (6) (2023) 329.

[19] M. Samiullah, H. Ali, S. Zahoor, and A. Ali, "Fault diagnosis on induction motor using machine learning and signal processing," *arXiv.org*, 2024.

[20] M.-C. Kim, J. Lee, D.-H. Wang, I.-S. Lee, Induction motor fault diagnosis using support vector machine, neural networks, and boosting methods, Sensors 23 (5) (2023) 2585.

[21] I. Ullah, N. Khan, S.A. Memon, W.-G. Kim, J. Saleem, S. Manzoor, Vibration-based anomaly detection for induction motors using machine learning, Sensors 25 (3) (2025) 773.

[22] M. Aishwarya, R.M. Brisilla, Design and fault diagnosis of induction motor using ML-based algorithms for EV application, IEEE Access 11 (2023) 34186–34197.

[23] A. Abdulkareem, T. Anyim, O. Popoola, J. Abubakar, A. Ayoade, Prediction of induction motor faults using machine learning, Heliyon 11 (1) (2025) e41493.

[24] P.L. Ratnani, A.G. Thosar, Mathematical modelling of a 3 phase induction motor using MATLAB/Simulink, Int. J. Mod. Eng. Res. (IJMER) 4 (6) (2014) 62–67.

[25] C. Kozkurt, M. Akar, Simulation of broken rotor bar fault in induction motors, in: 5th International Conference on Advanced Technology & Sciences (ICAT'17), 2017.

[26] J. Tang, J. Chen, K. Dong, Y. Yang, H. Lv, Z. Liu, Modeling and evaluation of stator and rotor faults for induction motors, Energies 13 (1) (2019) 133.

[27] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1157–1182.

[28] S. George, B. Sumathi, Grid search tuning of hyperparameters in random forest classifier for customer feedback sentiment prediction, IJACSA) Int. J. Adv. Comput. Sci. Appl. 11 (9) (2020).

[29] M. Muntean and Florin-Daniel Militaru, "Metrics for evaluating classification algorithms," pp. 307–317, 2023, doi: https://doi.org/10.1007/978-981-19-6755-9_24.

[30] D. Spike, 2022 Star EV Sirius 2+2 | Golf Cars of Houston superstore, *Golfcarsofhouston.com*, https://www.golfcarsofhouston.com/New-Inventory-2022-Star-EV-Golf-Cart-Sirius-2-2-South-Houston-Golf-Cars-16444859, 2022. accessed Aug. 27, 2025.

[31] R.A. Patel, B.R. Bhalja, Condition monitoring and fault diagnosis of induction motor using support vector machine, Electr. Power Compon. Syst. 44 (6) (2016) 683–692.

[32] S. Shao, R. Yan, Y. Lu, P. Wang, R.X. Gao, DCNN-based multi-signal induction motor fault diagnosis, IEEE Trans. Instrum. Meas. 69 (6) (2020) 2658–2669.

[33] P. Chattopadhyay, N. Saha, C. Delpha, J. Sil, Deep learning in fault diagnosis of induction motor drives, in: *2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, Chongqing, China, 2018, pp. 1068–1073.

[34] V. Kavana, M. Neethi, Fault analysis and predictive maintenance of induction motor using machine learning, in: *2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, Msyuru, India, 2018, pp. 963–966.

[35] M. Benninger, M. Liebschner, C. Kreischer, Fault detection of induction motors with combined modeling- and machine-learning-based framework, Energies 16 (8) (2023) 3429.

[36] J.-H. Lee, J.-H. Pack, I.-S. Lee, Fault diagnosis of induction motor using convolutional neural network, Appl. Sci. 9 (15) (2019) 2950.

[37] Q. He, D.-M. Du, Fault diagnosis of induction motor using neural networks, in: *2007 International Conference on Machine Learning and Cybernetics*, Hong Kong, China, 2007, pp. 1090–1095.

[38] V.G. de Araújo, et al., Monitoring and diagnosing faults in induction motors' Three-phase systems using NARX neural network, Energies 17 (18) (2024) 4609.

[39] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs), Appl. Soft. Comput. 11 (6) (2011) 4203–4211.

[40] M.O. Okelola, O.E. Olabode, Detection of voltage unbalance on three phase induction motor using artificial neural network, Int. J. Emerg. Trends Eng. Dev. 4 (8) (2018).

[41] P. Taweewat, W. Suwan-ngam, K. Songsuwankit, P. Konghuayrob, A small deep learning model for fault detection of a broken rotor bar of an induction motor, Sens. Mater. 36 (4) (2024) 1419.

[42] I. Martin-Diaz, D. Morinigo-Sotelo, O. Duque-Perez, R.J. Romero-Troncoso, An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions, IEEE Trans. Ind. Appl. 54 (3) (2018) 2215–2224.

[43] H.D.L. Rações, F.J.T.E. Ferreira, J.M. Pires, C.V. Damásio, Application of different machine learning strategies for current- and vibration-based motor bearing fault detection in induction motors, in: *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, Lisbon, Portugal, 2019, pp. 68–73.

[44] N. Dutta, P. Kaliannan, P. Shanmugam, Application of machine learning for inter turn fault detection in pumping system, Sci. Rep. 12 (1) (2022).

[45] R.N. Toma, A.E. Prosvirin, J.-M. Kim, Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers, Sensors 20 (7) (2020) 1884.

[46] S.E. Pandarakone, M. Masuko, Y. Mizuno, H. Nakamura, Deep neural network based bearing fault diagnosis of induction motor using fast fourier transform analysis, in: 2018 IEEE Energy Conversion Congress and Exposition (ECCE), 2018.

[47] L.P. Chisedzi, M. Muteba, Detection of broken rotor bars in cage induction motors using machine learning methods, Sensors 23 (22) (2023) 9079.

[48] S. Quabeck, W. Shangguan, D. Scharfenstein, W. De, Detection of broken rotor bars in induction machines using machine learning methods, IEEJ J. Ind. Appl. 10 (6) (2021) 688–693.

[49] M. Skowron, T. Orlowska-Kowalska, M. Wolkiewicz, C.T. Kowalski, Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor, Energies 13 (6) (2020) 1475.

[50] S. Shao, W. Sun, P. Wang, R.X. Gao, R. Yan, Learning features from vibration signals for induction motor fault diagnosis, in: *2016 International Symposium on Flexible Automation (ISFA)*, Cleveland, OH, USA, 2016, pp. 71–76.

[51] V. Kavana, M. Neethi, Fault analysis and predictive maintenance of induction motor using machine learning, in: International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2018.

[52] M. Hussain, T. Din Memon, I. Hussain, Z. Ahmed Memon, D. Kumar, Fault detection and identification using deep learning algorithms in induction motors, Comput. Model. Eng. Sci. 133 (2) (2022) 435–470.

[53] K. Laadjal, R. Amaral, M. Sahraoui, A.J. Marques, Machine learning based method for impedance estimation and unbalance supply voltage detection in induction motors, Sensors 23 (18) (2023) 7989.

[54] A.A. Silva, A.M. Bazzi, S. Gupta, Fault diagnosis in electric drives using machine learning approaches, in: 2013 International Electric Machines & Drives Conference, Chicago, IL, USA, 2013, pp. 722–726.