Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Severity Classification of Electrical Unbalance in Variable-Speed Induction Motors Using Low-Dimensional MCSA Features: A Data-Driven Framework for Predictive Maintenance in Manufacturing

Abdelfettah BOUSSAID

abdlfettah.BOUSSAID@lec-umc.org

University of Constantine 1: Universite Constantine 1 https://orcid.org/0000-0002-6859-9983

Sabbah ATAYA
Yacine KARMI
Haithem BOUMEDIRI
Omar REFFAS
Abderrahim TOUIL
Rashid KHAN

Research Article

Keywords: MCSA, Induction motor, Voltage unbalance, Variable speed, Fault severity, Machine learning

Posted Date: March 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9140611/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Severity Classification of Electrical Unbalance in Variable-Speed Induction Motors Using Low-Dimensional MCSA Features: A Data-Driven Framework for Predictive Maintenance in Manufacturing

Abdelfettah BOUSSAID[1,2*], Sabbah ATAYA[3], Yacine KARMI[1], Haithem BOUMEDIRI[4], Omar REFFAS[1], Abderrahim TOUIL[2], Rashid KHAN[3]

[1]*Electromechanical Department, Institute of Science and Applied Technics, University of Constantine 1, Constantine, Algeria*

[2]*Laboratoire d'électrotechnique de Constantine, University of Brothers Mentouri-Constantine 1, Constantine, Algeria.*

[3]*Department of Mechanical Engineering, College of Engineering, Imam Mohammad Ibn Saud Islamic, University (IMSIU), Riyadh, 11432, Saudi Arabia.*

[4]*Mechanical Department, Institute of Science and Applied Technics, University of Brothers Mentouri-Constantine 1, Constantine, Algeria.*

Corresponding authors *abdlfettah.boussaid@lec-umc.org

Abstract

Electrical unbalance is a critical supply-related disturbance in three-phase induction motors, as it alters the electromagnetic behavior of the machine, degrades operating stability, and may adversely affect maintenance scheduling and production continuity in industrial environments. In this context, the present study proposes a compact and data-driven framework for the multi-class severity classification of electrical unbalance in variable-speed induction motors using low-dimensional Motor Current Signature Analysis (MCSA) features. A dedicated experimental dataset was established from one healthy operating condition and five unbalance severity levels (U60%, U70%, U80%, U90%, and U110%) acquired at multiple rotational speeds. Three-phase currents and voltages, neutral voltage, and rotational speed were synchronously recorded at 10 kHz over 5 s, and each acquisition was segmented into eight non-overlapping windows to generate supervised learning samples. Six physically interpretable indicators were extracted from each window in order to characterize negative-sequence effects, spectral distortion, and neutral-voltage behavior induced by electrical unbalance. The diagnostic capability of the resulting feature space was then assessed using three supervised classification models, namely a Deep Neural Network (DNN), k-Nearest Neighbors (kNN), and AdaBoost.M2. In addition to the baseline evaluation performed with the complete feature set, a systematic feature-number ablation study was carried out to quantify the effect of input dimensionality on classification performance. The results demonstrate that the proposed descriptors provide strong class separability across variable-speed operating conditions. In particular, kNN achieved perfect discrimination using only two features, whereas the DNN reached perfect performance with four features, thereby highlighting the partial redundancy of the extracted indicators. Overall, the findings confirm that accurate multi-class severity diagnosis of electrical unbalance can be achieved using a compact and physically meaningful MCSA representation, making the proposed approach attractive for real-time condition monitoring and predictive maintenance of induction-motor-driven industrial systems.

Keywords: MCSA, Induction motor, Voltage unbalance, Variable speed, Fault severity, Machine learning.

1. Introduction

Three-phase induction motors are the most widely used electromechanical energy-conversion devices in industrial applications due to their simple construction, mechanical robustness, satisfactory efficiency, and favorable cost–performance ratio [1–3]. They are extensively employed in manufacturing systems, pumping installations, transportation infrastructure, ventilation units, and energy-conversion processes. Despite their inherent reliability, induction motors are continuously subjected to electrical, thermal, and mechanical stresses during operation, which progressively degrade winding insulation, magnetic circuits, and rotating components. Long-term industrial surveys and failure analyses have identified electric motors as a major contributor to unplanned maintenance activities and production interruptions, accounting for approximately 35–45% of total maintenance interventions, with unexpected shutdowns potentially increasing operational expenditures by 15–20% annually [4–6]. These observations highlight the necessity for effective condition-monitoring and fault-diagnosis strategies. Recent large-scale reviews have further confirmed that early-stage electrical fault detection significantly reduces downtime and maintenance costs, particularly when non-invasive current-based techniques are employed [7–10].

Among the different failure modes affecting induction motors, electrical and electromechanical faults represent a significant proportion of reported incidents. Stator winding inter-turn short circuits, broken rotor bars, air-gap eccentricity, insulation degradation, and supply-voltage unbalance are particularly critical because they directly alter the electromagnetic behavior of the machine and may accelerate thermal aging processes [5, 11]. Dash and Subudhi [12] demonstrated that neuro-fuzzy techniques (ANFIS) can enhance detection of stator inter-turn faults. Maraaba et al. [13] confirmed that stator current analysis provides reliable information for diagnosing winding faults under varying operating conditions. Liu et al. [14] experimentally showed that characteristic current components can effectively distinguish inter-turn short circuits under variable load, reinforcing the relevance of current-based diagnostics.

In real industrial environments, induction motors often operate under complex conditions where multiple fault mechanisms may coexist. Electrical faults may appear simultaneously with mechanical anomalies such as rotor asymmetry or eccentricity, leading to overlapping fault signatures and increased diagnostic complexity. To address this challenge, several studies have investigated time–frequency and multi-resolution signal-processing techniques. Menshawy et al. [5] reported that discrete wavelet transforms combined with adaptive neuro-fuzzy inference systems improves detection of broken rotor bar faults under variable load conditions. Cherif et al. [15] demonstrated that wavelet-based feature extraction combined with neural networks enhances stator fault discrimination compared with conventional spectral methods. Allal et al. [16] proposed an advanced residual mapping approach for improved sensitivity and severity estimation of stator faults under real operating conditions.

Stator current monitoring has therefore emerged as one of the most attractive non-invasive approaches for induction motor condition monitoring, as it avoids additional sensors and mechanical modifications. Classical diagnostic techniques based on Motor Current Signature Analysis and Fourier-based spectral analysis have been widely used to detect rotor bar defects and eccentricity-related harmonics [17–19]. Bellini et al. [17] provided a quantitative evaluation of broken rotor bar severity using electrical signature analysis, while Filippetti et al. [18] demonstrated the effectiveness of artificial intelligence techniques in enhancing diagnostic reliability. Nevertheless, conventional spectral methods may exhibit limitations under variable-speed operation and in the presence of non-stationary disturbances.

To overcome these limitations, adaptive signal-processing methods such as Empirical Mode Decomposition (EMD) have been introduced for induction motor fault diagnosis. Niu et al. [20] proposed a framework combining EMD with cyclostationary analysis and sparse representation, achieving improved fault characterization under varying operating conditions. Lin et al. [21] applied EMD and Hilbert modulus spectrum to air-gap eccentricity diagnosis, demonstrating enhanced fault-feature localization. Arifin et al. [22] further confirmed the relevance of EMD-based techniques for extracting characteristics associated with broken rotor bars. Chen et al. [23] showed that combining adaptive decomposition with digital-twin-based residual analysis improves early fault sensitivity under parameter uncertainties.

The integration of Artificial Intelligence techniques has further advanced induction motor fault diagnosis by enabling nonlinear pattern recognition and automated decision-making. Fuzzy logic and neural-network-based approaches have been successfully applied to stator and rotor fault diagnosis using current-based features [24–26]. Cleber et al. [27] demonstrated the applicability of fuzzy logic systems for broken rotor bar detection. More recent studies have explored deep-learning-based solutions: Susanta et al. [28] proposed a wavelet-kernel convolutional neural network for stator inter-turn fault identification, Turker [29] investigated convolutional neural networks for broken rotor bar detection, and Emanuele et al. [30] introduced an unsupervised deep autoencoder framework for electric motor fault detection. Kawady et al. [31] confirmed that ensemble AI models can effectively diagnose stator and rotor faults using stator current signals.

Within this scientific and industrial context, the present study proposes an experimental data-driven framework for the multi-class severity classification of electrical unbalance in variable-speed three-phase induction motors using low-dimensional Motor Current Signature Analysis (MCSA) features. The methodology is developed from a controlled laboratory test bench with synchronized acquisition of three-phase currents, three-phase voltages, neutral voltage, and rotational speed under healthy and multiple unbalance severity conditions. A compact set of six physically interpretable indicators, related to negative-sequence unbalance, spectral distortion, and neutral-voltage behavior, is extracted and used to train and

compare three supervised classifiers, namely Deep Neural Networks (DNN), k-Nearest Neighbors (kNN), and AdaBoost.M2. In addition, a feature-number ablation analysis is conducted to identify the minimum subset of descriptors required to preserve reliable diagnostic performance under variable-speed operation. Unlike broader fault-diagnosis approaches relying on complex multi-stage signal-processing pipelines, the present work intentionally focuses on a low-dimensional and interpretable MCSA representation suitable for accurate and computationally efficient severity diagnosis. From an industrial perspective, the early identification of electrical unbalance is essential not only for fault monitoring, but also for reducing unplanned downtime, improving maintenance scheduling, preserving equipment availability, and supporting production continuity in induction-motor-driven manufacturing systems. In this sense, the proposed framework contributes to predictive maintenance and real-time condition monitoring in line with Industry 4.0 requirements.

2. Materials and Methods

A custom-designed test bench was used to conduct an extensive experimental operation. The setup was carefully designed and calibrated to guarantee the collection of high-fidelity data under tightly regulated operating conditions, enabling a trustworthy assessment of the model's prediction skills.

A 1.1-kW three-phase squirrel cage induction machine made by "FIMET," with 22 rotor bars, two pole pairs, and an accessible neutral point, was the apparatus being tested. A magnetic powder brake was used to apply mechanical loading, allowing for accurate torque and speed control and repeatable experimental conditions in a variety of steady-state and transient regimes.

The integration of a high-performance data gathering and signal processing system intended for real-time operation was a crucial component of the experimental setup, as shown in Figure 1. The dSPACE DS1104 control and acquisition board, which directly connects with voltage and current sensors and a rotary encoder, is the central component of this architecture. Four LV 25-P voltage sensors and three LA 25-NP current sensors, each with a sensitivity of 25 mV/V, an accuracy of 0.9%, and a bandwidth of 100 k Hz, were used to obtain the voltage and current readings.

Experimental tests were performed across various operating speeds. The rotor speed was measured utilizing an OptoPulse EIL580-SC Speed Encoder, capable of delivering up to 5000 pulses per revolution with an accuracy of ±0.01% and a bandwidth of 300 kHz. These specifications guaranteed accurate high-resolution speed tracking and ensured precise synchronization throughout the signal acquisition process.

This setup facilitates the concurrent and synchronized acquisition of electrical parameters, such as three-phase voltages, line currents, and neutral point displacement, alongside mechanical variables like shaft

speed and vibration levels. The system integrates smoothly with a host computer through MATLAB/Simulink and Control-desk platforms, enabling real-time signal observation, dynamic parameter adjustment, and organized data recording throughout the experimental process.

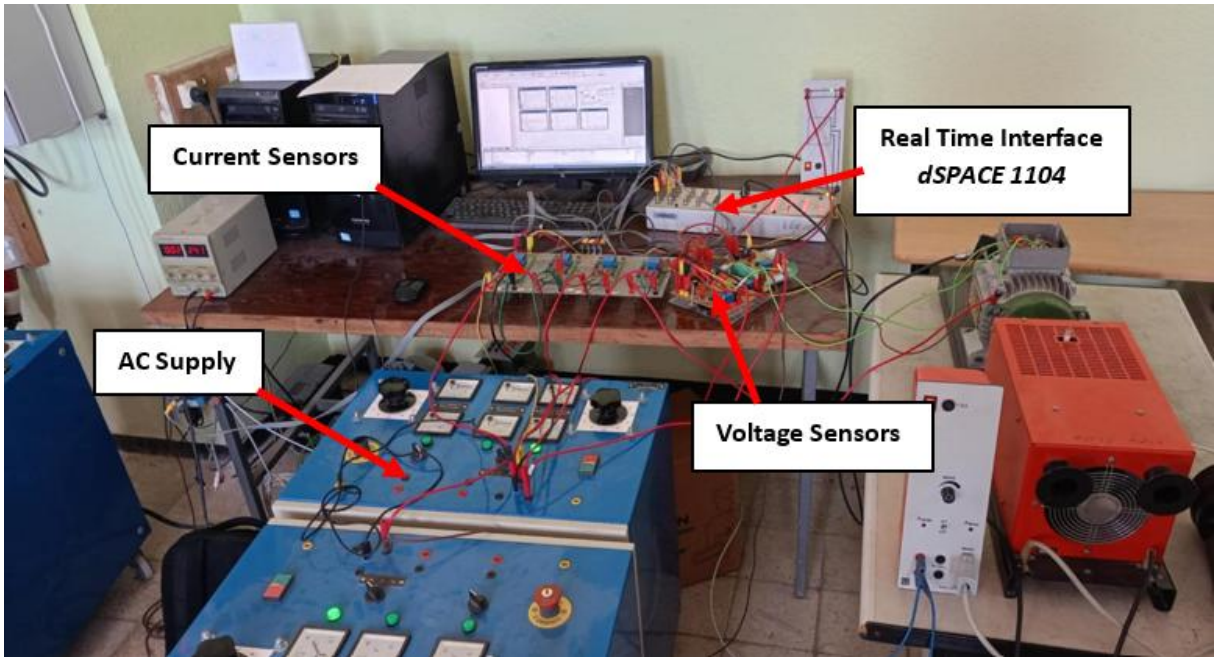


Figure 1. Experimental test bench used for induction motor current acquisition.

2.1. Visualization of unbalance signatures across speed and severity

The influence of rotor unbalance on the electrical response is first illustrated through time-domain and geometric representations, reported in Figs. 2 (a)-(c). These visualizations are provided to (i) verify that the acquired dataset exhibits coherent, fault-dependent signal variations under steady-state conditions, and (ii) support the interpretation of the diagnostic features introduced in the subsequent sections. All plots are generated after removing the initial transient interval $[0, 0.5]$ s, and the displayed time window is restricted to 0–0.25 s to ensure a homogeneous visual comparison across operating points.

Figure 2 reports the phase current overlays $\{i_a(t), i_b(t), i_c(t)\}$for three representative speeds. For each speed, the three subplots correspond to the three phases, while the curves within each subplot are superimposed for the Healthy case and for all unbalance levels. This representation highlights that the current waveforms are not only speed-dependent but also progressively affected by fault severity. In particular, as the unbalance level increases, the waveforms exhibit increasingly visible departures from the nominal balanced pattern, including amplitude redistribution among phases and waveform deformation. Such variations are consistent with the physical expectation that unbalance introduces additional periodic components and asymmetry in the electromechanical system, which are ultimately reflected in the stator

currents. The systematic nature of these changes across the three speeds indicates that the current signals contain reproducible information suitable for subsequent condition discrimination.

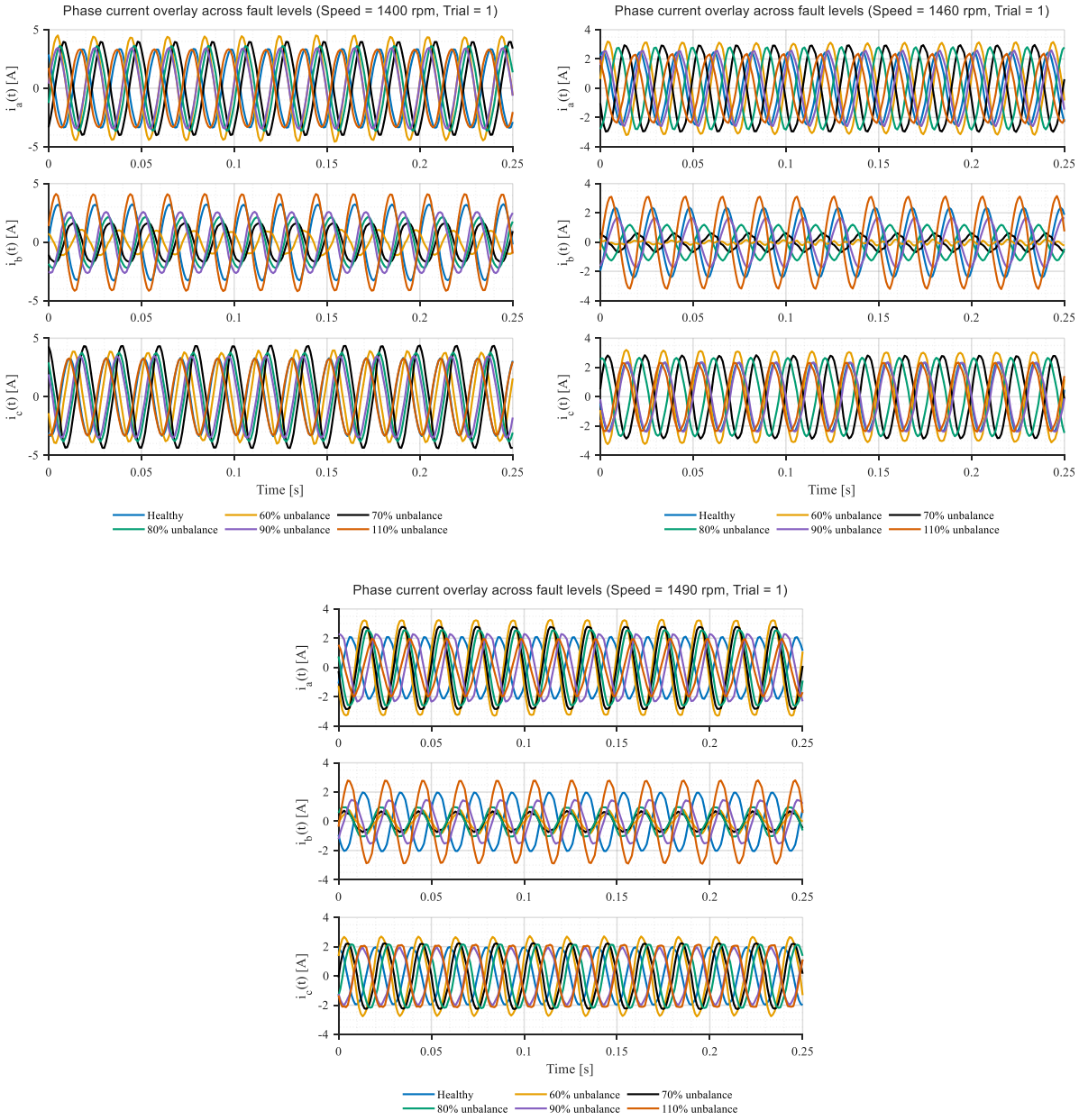


Figure 2. Phase current overlays across unbalance levels at three rotational speeds (Trial 1).

To provide a geometric interpretation of the three-phase behavior, Fig. 3 (a)–(c) depicts the Clarke α–β trajectories for the same three speeds. The transformation is defined as

$$\alpha(t) = \frac{2}{3}\left(i_a(t) - \frac{1}{2}i_b(t) - \frac{1}{2}i_c(t)\right), \beta(t) = \frac{2}{3}\left(\frac{\sqrt{3}}{2}\right)\left(i_b(t) - i_c(t)\right), \qquad (1)$$

which maps the three-phase set into an orthogonal stationary frame. In the Healthy condition, the trajectory forms a comparatively regular closed locus, whereas increasing unbalance levels yield progressive trajectory deformation, anisotropy, and spreading. These geometrical changes are clearly observable across all speeds in Figure 3, indicating that the Clarke plane representation amplifies asymmetry effects that may be less evident in individual phase waveforms. This observation motivates the use of unbalance-sensitive indicators derived from orthogonal components (or from their spectral content) in the proposed diagnostic framework.

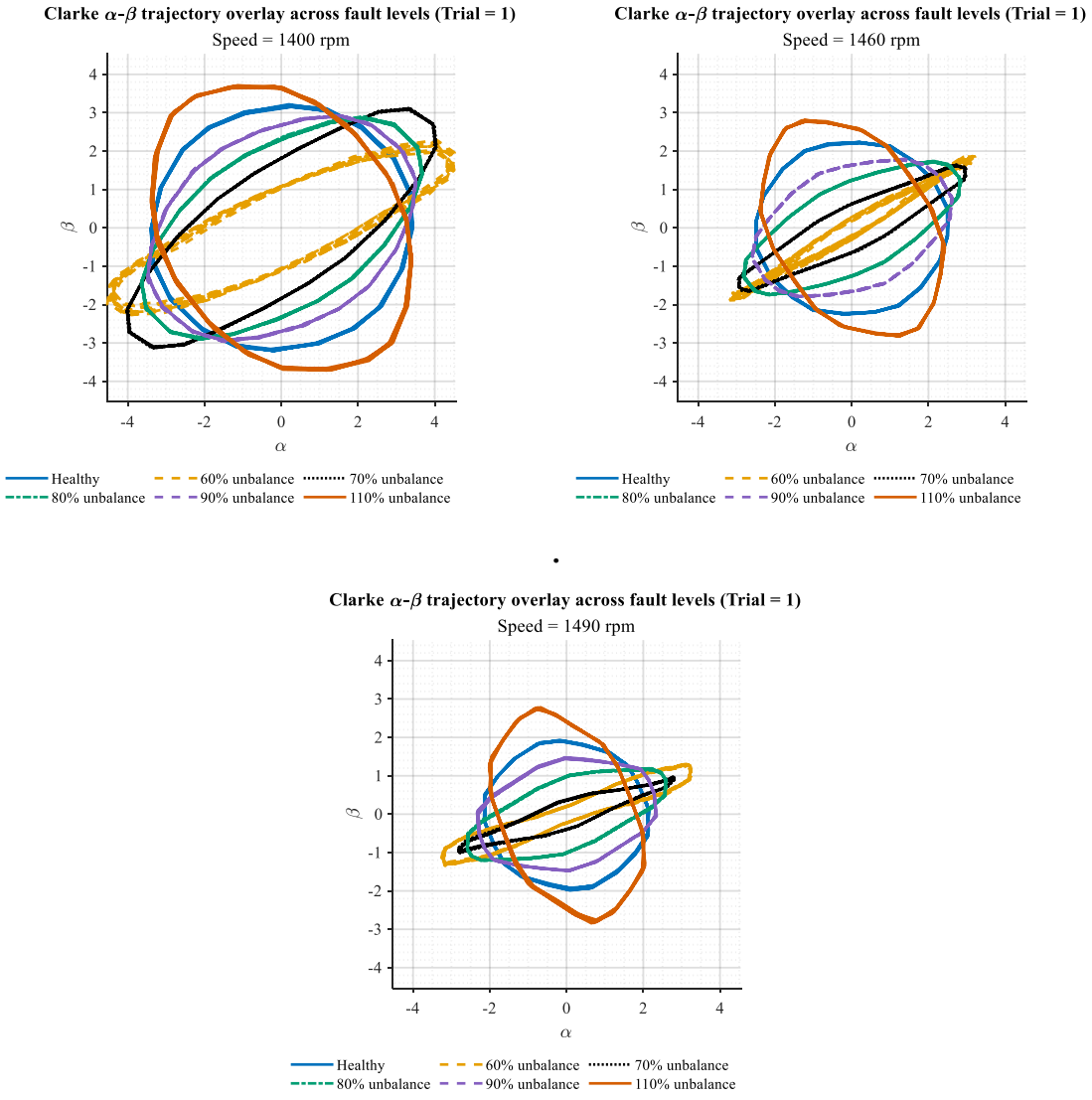


Figure 3. Clarke α–β current trajectories across unbalance levels for three speeds (Trial 1).

The effect of fault severity is further emphasized by the neutral voltage overlays shown in Fig. 4 (a)–(c) for 1400, 1460, and 1490 rpm, respectively. Compared with phase currents, the neutral voltage $v_n(t)$is particularly sensitive to deviations from three-phase symmetry; consequently, it provides a compact indicator of imbalance-induced asymmetry. As illustrated in Figure 4, the increase in unbalance severity is associated with a clear growth of waveform distortion and amplitude excursions, while the overall waveform morphology remains speed-dependent. Importantly, the adoption of identical time limits and consistent scaling across subpanels allows a direct visual appraisal of how the same fault level manifests differently as speed changes, supporting the need for speed-aware evaluation.

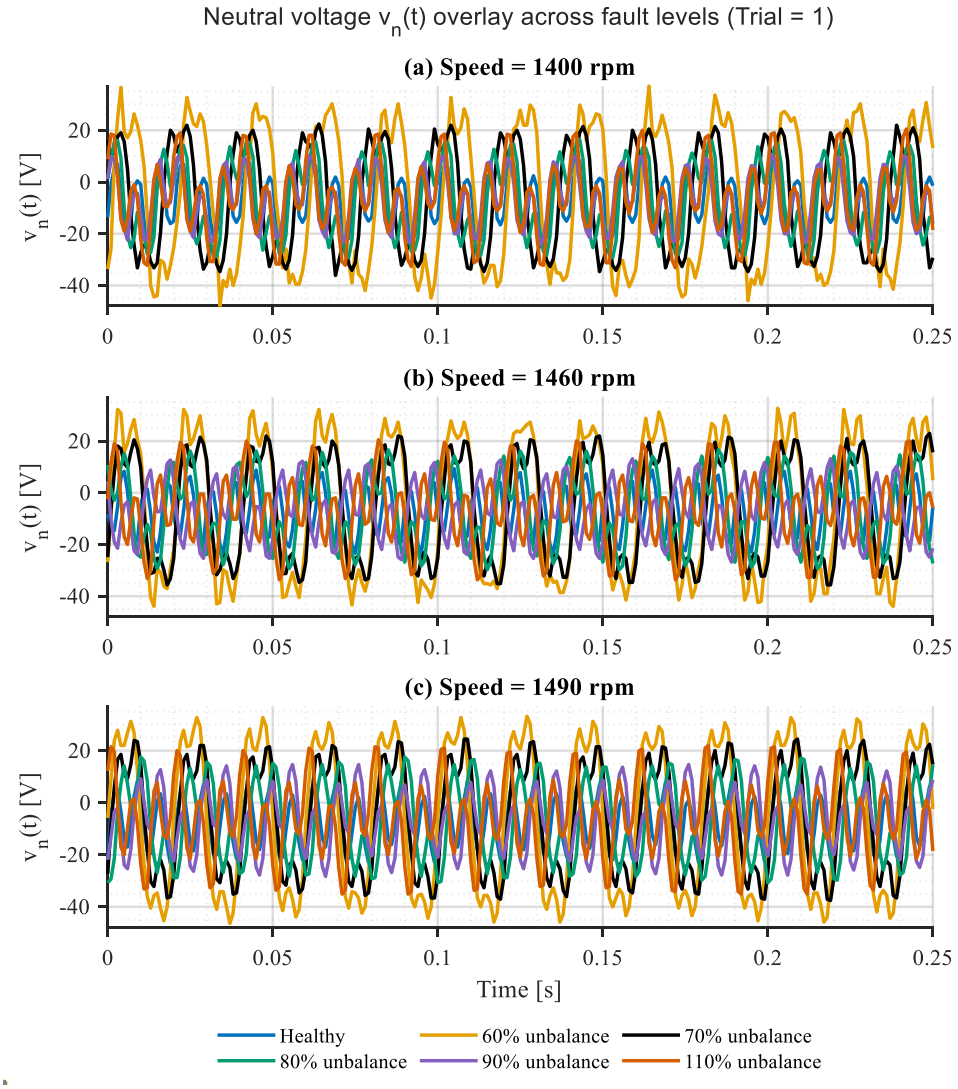


Figure 4. Neutral voltage $v_n(t)$overlays across unbalance levels at three rotational speeds (Trial 1).

2. MCSA-based feature extraction

To quantify the signatures induced by voltage/current unbalance, we adopted Motor Current Signature Analysis (MCSA). Feature extraction is performed from the stator currents (and optionally voltage) using frequency-domain and sequence-component representations commonly employed in recent fault diagnosis literature.

3.1. Clarke/Park transformation and sequence-related indicators

First, three-phase currents are transformed into stationary orthogonal components via Clarke transform:

$$\begin{aligned} i_\alpha(t) &= \frac{2}{3}\left(i_a(t) - \frac{1}{2}i_b(t) - \frac{1}{2}i_c(t)\right), \\ i_\beta(t) &= \frac{2}{3}\left(\frac{\sqrt{3}}{2}i_b(t) - \frac{\sqrt{3}}{2}i_c(t)\right). \end{aligned} \tag{2}$$

A complex current vector can be defined as

$$\underline{i}(t) = i_\alpha(t) + j\, i_\beta(t) \tag{3}$$

In addition, voltage unbalance severity can be described using symmetrical components. Let $\underline{V}_1$ and $\underline{V}_2$ be the positive- and negative-sequence voltage phasors, then the Voltage Unbalance Factor (VUF) is:

$$\mathrm{VUF}(\%) = 100 \times \frac{|\,\underline{V}_2\,|}{|\,\underline{V}_1\,|} \tag{4}$$

Similarly, the Current Unbalance Factor (IUF) is:

$$\mathrm{IUF}(\%) = 100 \times \frac{|\,\underline{I}_2\,|}{|\,\underline{I}_1\,|} \tag{5}$$

These indicators are widely used to characterize unbalance conditions.

3.2. Harmonic distortion feature

To capture spectral distortion, we compute the Total Harmonic Distortion of a representative current magnitude (or of the $\alpha\beta$ current vector). A generic THD definition is:

$$\mathrm{THD}_I(\%) = 100 \times \frac{\sqrt{\sum_{h=2}^{H} I_h^2}}{I_1} \qquad (6)$$

where I_1 is the fundamental RMS component and I_h are harmonic RMS components up to order H.

3.3. Time-domain RMS indicator

Since neutral voltage is sensitive to asymmetry and unbalance, we also compute a neutral-voltage energy measure such as:

$$\mathrm{RMS}_{V_n} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} v_n\,[n]^2} \qquad (7)$$

3.4 Park-vector / spectral ratio indicator

A compact unbalance descriptor can be formed by comparing energy around frequency components associated with negative sequence or characteristic sidebands. In this work, we used a ratio-type feature such as:

$$\mathrm{Park2f}_{\mathrm{ratio}} = \frac{E_{2f}}{E_f} \qquad (8)$$

where E_f and E_{2f} denote spectral energy in narrow bands around f and $2f$ extracted from the Park/Clarke-domain current representation (band integration in the power spectrum).

3. Dataset organization and sample definition

The experimental dataset was acquired under one healthy reference state and multiple electrical unbalance severity levels. Each experimental trial corresponds to a single operating point defined by (severity level × speed setpoint × repetition). Trials include synchronized measurements of three-phase currents $i_a(t), i_b(t), i_c(t)$, three-phase voltages $v_a(t), v_b(t), v_c(t)$, neutral voltage $v_n(t)$, and motor speed $\omega_m(t)$ (and vibration, when available). Signals were recorded for $T = 5$s with sampling frequency $f_s = 10$kHz, yielding $N = Tf_s + 1$ samples per channel.

To increase the number of learning samples and promote local quasi-stationarity of spectral descriptors, each trial is segmented into $W = 8$ non-overlapping windows. The window length is therefore:

$$T_w = \frac{T}{W} = 0.625 \text{ s}, N_w = T_w f_s \approx 6250 \text{ samples} \quad (9)$$

A single window constitutes one observation for ML:

$$(\mathbf{x}_{i,w}, y_i), w \in \{1, \dots, W\} \quad (10)$$

where $\mathbf{x}_{i,w}$ is the feature vector extracted from window $w$of trial i, and the label y_i is inherited by all windows from that trial. The final multi-class label set is:

$$\mathcal{C} = \{\text{Healthy,U60\%,U70\%,U80\%,U90\%,U110\%}\} \quad (11)$$

4.1. **Feature extraction**

Each window is mapped to a compact vector of MCSA-related features commonly used for unbalance detection. The complete feature set used in this study contains up to six variables (the paper can report the final selection used in experiments):

$$\boldsymbol{x} = \left[RMS_{V_n}, CUF_{RMS}, \quad THD_I, \quad Park2f_{ratio}, \quad IUF, \quad VUF\right] \quad (12)$$

All features are computed per window, producing a supervised dataset matrix $\mathbf{X} \in \mathbb{R}^{M \times d}$, with M windows and d the number of selected features.

4.2. Preprocessing and standardization

Because the extracted features have different physical units and magnitudes, a standardization step is applied prior to training. Let $x_{i,k}$ be the k-th feature of sample i. Standardization is performed using the mean and standard deviation computed only on the training set:

$$\tilde{x}_{i,k} = \frac{x_{i,k} - \mu_k}{\sigma_k} \quad (13)$$

The same μ_k and σ_k are then applied to validation and test sets. This is critical for (i) distance-based models such as kNN, and (ii) gradient-based optimization in neural networks. Constant features (zero variance) are handled by enforcing $\sigma_k = 1$ to avoid numerical instability.

4.3. Data splitting strategy

The dataset is partitioned into training, validation, and test subsets using a fixed random seed for reproducibility. A typical split used in this work is 70% training / 10% validation / 20% test (or 70/20/10 depending on the experiment) [32, 33]. Importantly, splitting is performed at the trial level so that each trial and its windows belong exclusively to one subset.

The validation set is used for hyperparameter selection and early stopping, while the test set is reserved exclusively for the final reporting of performance metrics.

4. Classification models

5.1. Deep Neural Network (DNN)

A deep feedforward neural network (multilayer perceptron) is used to learn a nonlinear mapping from the standardized feature vector $\tilde{\mathbf{x}}$ to class posterior probabilities [34, 35]. For L hidden layers, the network computes:

$$\mathbf{h}^{(1)} = \phi\left(\mathbf{W}^{(1)}\tilde{\mathbf{x}} + \mathbf{b}^{(1)}\right), \mathbf{h}^{(l)} = \phi\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \quad (14)$$

followed by an output layer:

$$\mathbf{z} = \mathbf{W}^{(L+1)}\mathbf{h}^{(L)} + \mathbf{b}^{(L+1)} \quad (15)$$

Class probabilities are obtained via softmax:

$$p(y = c \mid \tilde{\mathbf{x}}) = \frac{\exp(z_c)}{\sum_{k \in \mathcal{C}} \exp(z_k)} \quad (16)$$

Prediction is:

$$\hat{y} = \arg\max_{c \in \mathcal{C}} p(y = c \mid \tilde{\mathbf{x}}) \quad (17)$$

Training objective and optimization

The DNN is trained by minimizing multi-class cross-entropy:

$$\mathcal{L}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c\in\mathcal{C}} \mathbb{1}(y_i = c)\log\, p(y = c \mid \tilde{\mathbf{x}}_i) \qquad (18)$$

using backpropagation and gradient-based optimization. A validation set is used to monitor generalization and avoid overfitting via early stopping. In practice, the architecture (number of layers and units) is selected to balance expressiveness and stability given the dataset size.

5.2. AdaBoost.M2 (Multi-class Boosting Ensemble)

AdaBoost.M2 is a stage-wise ensemble learning method designed for multi-class classification, where a strong classifier is built by combining a sequence of weak learners [36]. Unlike bagging-based ensembles, boosting is sequential: each new learner is trained to correct the mistakes of the previously trained learners by increasing the importance (weights) of difficult samples. This mechanism is particularly well suited for fault severity diagnosis, where misclassifications often occur between adjacent unbalance levels (e.g., U70% vs. U80%) and where speed variations can create overlapping feature distributions.

In the multi-class setting, AdaBoost.M2 maintains an adaptive distribution that penalizes samples that are repeatedly misclassified and, more generally, samples whose correct class is not sufficiently separated from competing classes. Consequently, subsequent learners focus on the most ambiguous observations, leading to improved class discrimination and often stronger macro-averaged performance.

Weak learners and model capacity control. In this study, each weak learner $\mathrm{h_t}$ is a shallow decision tree trained under the current sample weights. Trees are used because they naturally model threshold-based relations in diagnostic features (e.g., unbalance factors, THD, spectral ratios). Importantly, weak learners are constrained in complexity (e.g., limited number of splits and minimum leaf size) to preserve the weak-learning regime and reduce overfitting.

After T boosting rounds, the final ensemble decision is obtained via a weighted vote:

$$F(\tilde{x}) = \arg\max_{c\in C} \sum_{t=1}^{T} \alpha_t \ \mathbf{1}\left(h_t(\tilde{x}) = c\right) \qquad (19)$$

where T is the number of boosting iterations and $\alpha_t > 0$ is the contribution weight of learner t, typically increasing when the weighted classification quality of h_t is higher. The weights α_t are derived from a multi-class boosting loss (pseudo-loss), ensuring that learners performing better than random guessing receive more influence in the final classifier.

5.3. k-Nearest Neighbors (kNN)

kNN is a non-parametric classifier that performs classification based on local neighborhood structure in the feature space. It is particularly useful as a baseline for engineered diagnostic features, because strong performance indicates that the extracted features yield naturally separable clusters between operating conditions (healthy vs. unbalance severities)[37].

Given a standardized test vector $\tilde{x}$, kNN computes the distance to all training samples. Using Euclidean distance:

$$d(\tilde{x}, \tilde{x}_i) = \| \tilde{x} - \tilde{x}_i \|_2 \tag{20}$$

the k closest training samples define the neighborhood $N_k(\tilde{x})$.

The predicted class is obtained by majority voting:

$$\hat{y} = \arg \max_{c \in C} \sum_{i \in N_k(\tilde{x})} \mathbf{1}(y_i = c) \tag{21}$$

Optionally, to increase robustness to borderline samples, a **distance-weighted voting** scheme can be used:

$$\hat{y} = \arg \max_{c \in C} \sum_{i \in N_k(\tilde{x})} w_i \ \mathbf{1}(y_i = c), w_i = \frac{1}{d(\tilde{x}, \tilde{x}_i) + \epsilon} \tag{22}$$

where ϵ is a small constant to avoid division by zero.

5. Model evaluation and statistical reporting

Confusion matrix analysis

A confusion matrix $\mathbf{M} \in \mathbb{N}^{K \times K}$ is computed for each classifier. The element M_{ij} counts samples from true class i predicted as class j. This analysis reveals where confusion occurs (commonly between neighboring severity levels).

Accuracy

$$\mathrm{Acc} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(\hat{y}_i = y_i) \tag{23}$$

Precision, recall and Macro-F1

For each class c:

$$\mathrm{Prec}_c = \frac{TP_c}{TP_c + FP_c}, \mathrm{Rec}_c = \frac{TP_c}{TP_c + FN_c} \tag{24}$$

The class-wise F1 is:

$$\mathrm{F1}_c = \frac{2\,\mathrm{Prec}_c\,\mathrm{Rec}_c}{\mathrm{Prec}_c + \mathrm{Rec}_c} \tag{25}$$

and the macro-average is:

$$\mathrm{F1}_{macro} = \frac{1}{K}\sum_{c=1}^{K} \mathrm{F1}_c \tag{26}$$

Macro-F1 is reported because it treats all severity levels equally and is robust to class imbalance.

One-vs-rest ROC and Macro-AUC

For each class c, a one-vs-rest ROC curve is derived from model scores $s_c(\tilde{\mathbf{x}})$. The AUC is computed per class and averaged:

$$\mathrm{AUC}_{macro} = \frac{1}{K}\sum_{c=1}^{K} \mathrm{AUC}_c \tag{27}$$

This provides a threshold-independent measure of separability.

6. Results and discussion

7.1.Statistical Characterization of the Extracted MCSA Features

Table 1 reports the mean ± standard deviation of the extracted MCSA-related features, computed at the window level and then aggregated over all speeds and repetitions. The Healthy condition exhibits consistently low values for all unbalance-sensitive indicators, whereas faulty conditions show clear increases, confirming the relevance of the proposed feature set for severity discrimination.

Figure 5 illustrates the class-wise distributions of the six extracted MCSA-related features using boxplots. Each box represents the interquartile range (IQR), the central line indicates the median, whiskers extend to 1.5×IQR, and isolated points correspond to outliers. The distributions are computed from all windowed samples and aggregated over the different operating speeds and repetitions.

Table 1. Statistical Analysis of Low-Dimensional MCSA Features Across Variable Rotational Speeds and Unbalance Severities.

Severity Level	VUF (%)	IUF (%)	$Park_{2f}$ Ratio	THD_{Ivec} (%)	CUF_{RMS}	RMS_{Vn} (V)
Rotational Speed: n=1400 rpm						
Healthy	0.0088±0.0002	0.0296±0.0017	0.0390±0.0020	11.74±0.63	1.0234±0.0030	11.96±0.01
U60%	0.1618±0.0001	1.5297±0.0138	0.5973±0.0036	320.56±40.19	4.0245±0.0742	27.47±0.03
U70%	0.0996±0.0002	0.4877±0.0031	0.4856±0.0027	155.68±3.69	2.4808±0.0192	21.74±0.06
U80%	0.0545±0.0002	0.3052±0.0010	0.3202±0.0015	75.70±6.59	1.6871±0.0027	15.66±0.28
U90%	0.0251±0.0003	0.1735±0.0017	0.1890±0.0012	38.73±3.15	1.3075±0.0054	12.28±0.02
U110%	0.0443±0.0006	0.1467±0.0008	0.1371±0.0035	38.17±3.81	1.2770±0.0024	15.15±0.04
Rotational Speed: n=1420 rpm						
Healthy	0.0092±0.0001	0.0299±0.0004	0.0405±0.0006	11.50±0.69	1.0236±0.0007	11.85±0.01
U60%	0.1604±0.0002	1.3624±0.0076	0.6320±0.0023	393.87±52.47	5.3685±0.1066	27.39±0.06
U70%	0.0987±0.0004	0.5608±0.0007	0.5281±0.0050	165.44±21.83	2.9785±0.0056	21.59±0.04
U80%	0.0538±0.0003	0.3375±0.0009	0.3445±0.0021	82.95±16.98	1.7971±0.0034	15.22±0.02
U90%	0.0251±0.0002	0.1916±0.0004	0.2042±0.0031	40.11±1.62	1.3512±0.0012	12.14±0.08
U110%	0.0438±0.0004	0.1619±0.0021	0.1537±0.0027	37.64±1.80	1.3016±0.0047	15.10±0.03
Rotational Speed: n=1440 rpm						
Healthy	0.0088±0.0001	0.0355±0.0007	0.0491±0.0006	11.65±0.41	1.0328±0.0010	11.79±0.03
U60%	0.1601±0.0003	1.2003±0.0014	0.6510±0.0080	321.43±130.19	9.1240±0.0259	27.35±0.03
U70%	0.0977±0.0002	0.6288±0.0009	0.5703±0.0017	145.49±5.50	3.6118±0.0099	21.71±0.03
U80%	0.0538±0.0003	0.3797±0.0007	0.3761±0.0036	76.72±6.94	1.9476±0.0022	14.92±0.02
U90%	0.0252±0.0001	0.2174±0.0006	0.2290±0.0007	38.57±0.97	1.4006±0.0018	12.49±0.01
U110%	0.0439±0.0005	0.1911±0.0011	0.1728±0.0030	37.66±2.07	1.3331±0.0024	15.25±0.02
Rotational Speed: n=1460 rpm						
Healthy	0.0095±0.0002	0.0332±0.0019	0.0481±0.0022	9.93±0.84	1.0309±0.0026	11.86±0.02
U60%	0.1586±0.0003	1.0787±0.0014	0.6692±0.0038	343.30±96.00	19.63±0.24	27.37±0.05
U70%	0.0973±0.0004	0.7275±0.0007	0.5974±0.0092	163.41±32.69	4.9556±0.0119	21.78±0.06
U80%	0.0554±0.0017	0.4464±0.0153	0.4172±0.0124	80.47±16.19	2.2442±0.0800	15.66±0.36
U90%	0.0257±0.0003	0.2458±0.0016	0.2511±0.0004	35.10±2.01	1.4958±0.0050	12.72±0.05
U110%	0.0441±0.0004	0.2102±0.0012	0.1943±0.0013	40.42±3.06	1.3578±0.0027	15.44±0.03
Rotational Speed: n=1480 rpm						
Healthy	0.0087±0.0001	0.0370±0.0009	0.0544±0.0011	7.63±0.50	1.0397±0.0014	12.20±0.01
U60%	0.1427±0.0005	1.1633±0.0449	0.6480±0.0123	173.12±27.46	6.0928±2.0543	29.77±0.03
U70%	0.0971±0.0001	0.7694±0.0004	0.6223±0.0010	121.87±12.00	5.2893±0.0094	21.96±0.03
U80%	0.0560±0.0004	0.4997±0.0012	0.4513±0.0054	75.20±8.28	2.5205±0.0124	16.30±0.34
U90%	0.0249±0.0002	0.2630±0.0004	0.2622±0.0004	33.05±1.52	1.5471±0.0011	13.00±0.03
U110%	0.0433±0.0001	0.2271±0.0011	0.2177±0.0012	35.48±1.97	1.4407±0.0014	15.91±0.01
Rotational Speed: n=1490 rpm						
Healthy	0.0083±0.0002	0.0380±0.0011	0.0535±0.0011	5.19±0.60	1.0413±0.0010	12.33±0.01
U60%	0.1426±0.0002	1.2007±0.0024	0.6450±0.0011	186.61±21.61	4.7099±0.0651	29.62±0.36
U70%	0.0956±0.0003	0.8367±0.0012	0.6179±0.0014	98.54±16.19	4.1133±0.0055	22.07±0.01
U80%	0.0556±0.0004	0.5245±0.0014	0.4557±0.0050	89.88±24.77	2.5477±0.0093	16.27±0.03
U90%	0.0242±0.0002	0.2710±0.0012	0.2590±0.0013	30.49±3.23	1.5500±0.0044	13.23±0.04
U110%	0.0438±0.0003	0.2379±0.0011	0.2272±0.0029	42.42±8.96	1.5108±0.0025	16.29±0.02

The boxplots confirm a clear statistical separation between the *Healthy* condition and the unbalance severity levels. In particular, VUF and IUF show strong discrimination power: both

indicators remain close to zero under healthy operation and increase significantly under unbalance, reflecting the emergence of negative-sequence components. Similarly, THD_{Ivec} exhibits a pronounced rise for faulty cases especially for U60%, indicating substantial spectral distortion caused by supply asymmetry. The $Park2f_{ratio}$ also provides consistent separation across severity levels, suggesting that the second-harmonic-related energy content in the transformed current signature is highly sensitive to unbalance.

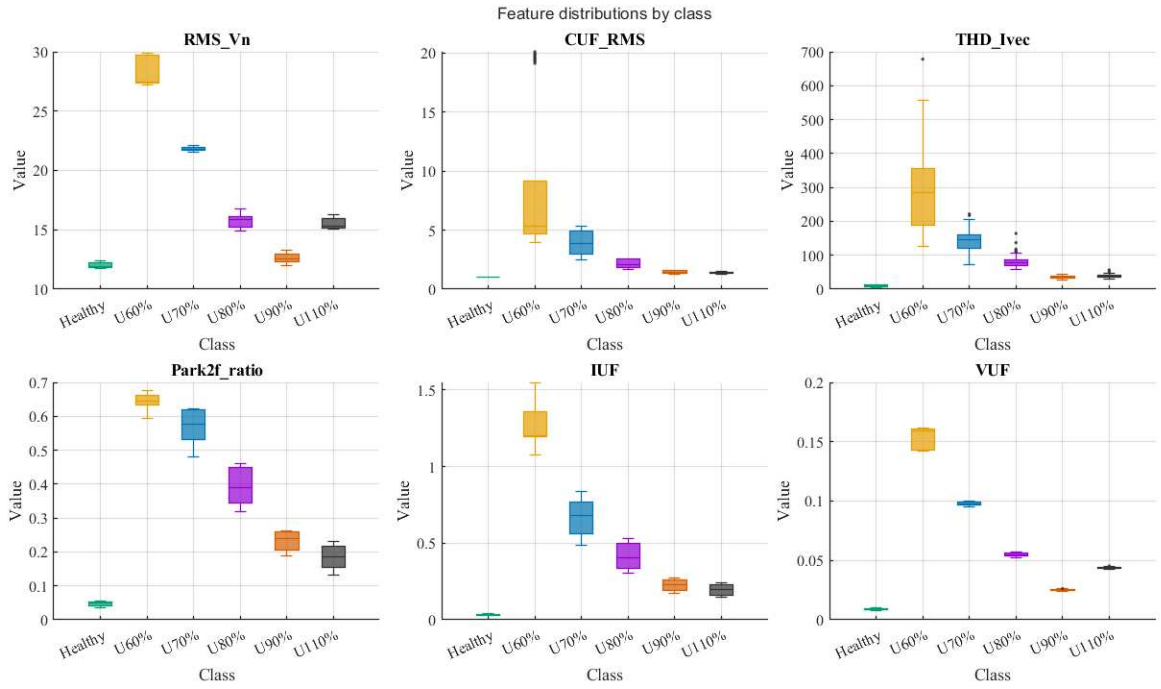


Figure 5. Class-wise boxplots of the extracted MCSA-related features aggregated across operating speeds and repetitions. Boxes indicate the interquartile range (IQR), the central line is the median, whiskers extend to 1.5×IQR, and markers denote outliers.

In contrast, CUF_{RMS} presents larger dispersion and occasional outliers in the most severe condition(s), highlighting the influence of operating variability (speed-dependent effects and experimental fluctuations). Nevertheless, its median values still follow the expected trend, supporting its role as a complementary descriptor rather than a standalone indicator. Finally, RMS_{Vn} exhibits an evident gap between *Healthy* and faulty conditions, confirming the sensitivity of neutral voltage to electrical asymmetry.

Figure 6 presents the Pearson correlation heatmap of the extracted MCSA-based features $\{\mathrm{RMS}_{V_n}, \mathrm{CUF}_{RMS}, \mathrm{THD}_{Ivec}, \mathrm{Park2f_ratio}, \mathrm{IUF}, \mathrm{VUF}\}$. This analysis aims to quantify the level of redundancy between features and to support the interpretation of the feature-number study.

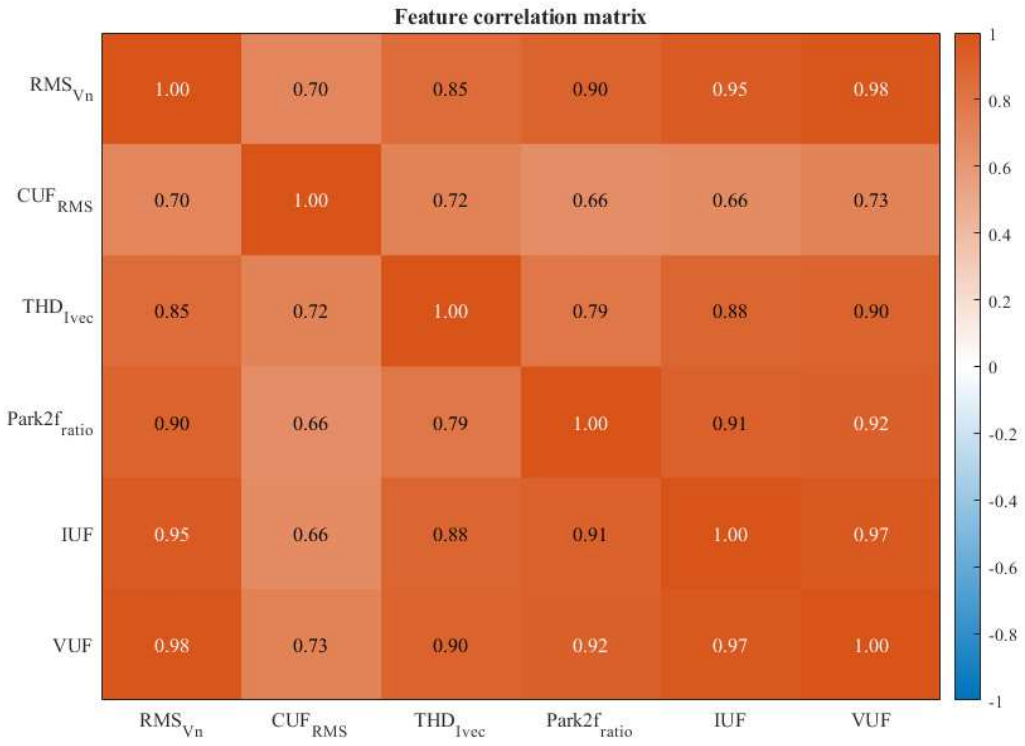


Figure 6. presents the Pearson correlation heatmap of the extracted MCSA-based features $\{\mathrm{RMS}_{V_n}, \mathrm{CUF}_{RMS}, \mathrm{THD}_{Ivec}, \mathrm{Park2f_ratio}, \mathrm{IUF}, \mathrm{VUF}\}$.

the heatmap reveals that most variables are strongly positively correlated, indicating that they capture closely related physical effects induced by electrical unbalance. In particular, VUF and IUF exhibit a very high correlation ($\rho \approx 0.97$), which is expected because both metrics are derived from symmetrical components and reflect the magnitude of negative-sequence content. Similarly, RMS_{V_n} is highly correlated with VUF($\rho \approx 0.98$) and IUF($\rho \approx 0.95$), confirming that the neutral voltage RMS is a reliable proxy for supply asymmetry.

Spectral descriptors also show strong associations with sequence-related indicators: THD_{Ivec} correlates with VUF($\rho \approx 0.90$) and with IUF($\rho \approx 0.88$), suggesting that increasing unbalance not only increases negative-sequence components but also enhances harmonic distortion. The

Park2f_ratiois also strongly correlated with VUF and IUF($\rho \approx 0.92$ and $\rho \approx 0.91$, respectively), supporting the interpretation that modulation effects captured in the $\alpha\beta$ representation intensify with unbalance severity.

Among all features, CUF_{RMS} shows comparatively lower correlations (approximately 0.66–0.73) and thus provides more complementary information, potentially improving robustness when used jointly with the other descriptors.

Implication for modeling. The strong correlations indicate that the six-feature set contains redundant information, which explains why high classification performance may be achieved using only a subset of features. This observation motivates the incremental feature-number analysis (ablation study) and supports the use of models capable of handling correlated predictors (e.g., boosted ensembles and neural networks).

Figure 7 shows the two-dimensional PCA projection of the extracted MCSA feature vectors. The first principal component (PC1) explains **86.7%** of the total variance, while PC2 explains **7.4%**, indicating that most of the information is captured along PC1. The projection reveals well-structured class clusters: The **Healthy** samples form a compact group clearly separated from unbalance conditions, confirming that the selected features strongly discriminate normal operation from faulty states.

Moreover, the unbalance severity levels exhibit an ordered separation mainly along PC1, suggesting that PC1 is closely associated with the magnitude of supply asymmetry and negative-sequence effects. In particular, the **U60%** class appears far from the other clusters and shows the largest spread, which is consistent with the highest disturbance level and with stronger sensitivity to operating variations (e.g., speed-dependent effects and window-to-window variability).

The intermediate classes (U70%–U110%) remain distinguishable and form compact clusters, supporting the suitability of the feature set for multi-class severity classification. Overall, this PCA visualization provides qualitative evidence that the proposed low-dimensional MCSA descriptors yield a separable feature space, which explains the near-perfect performance achieved by DNN and kNN classifiers.

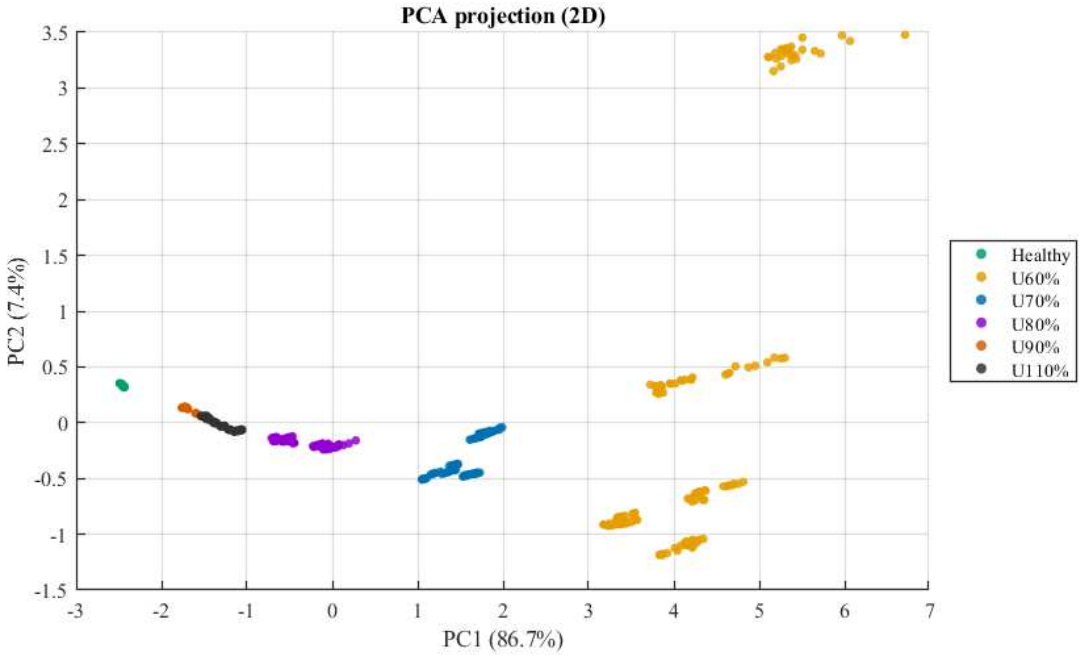


Figure 7. 2D PCA projection of the extracted MCSA feature vectors for the six classes.

7.2.Baseline Model Comparison Using the Full Feature Set ($k = 6$)

This subsection analyzes the baseline classification performance obtained when using the complete MCSA feature vector ($k = 6$), by comparing three supervised models (DNN, AdaBoost.M2, and kNN) through (i) global metrics, (ii) confusion matrices, and (iii) macro-averaged ROC curves.

The bar chart in Figure 8 summarizes the overall performance of each classifier. Two models DNN and kNN achieve near-perfect results, with Accuracy and Macro-F1 close to 1, and Macro-AUC equal to 1. This indicates that the six selected features provide a highly discriminative representation where class boundaries are either strongly non-linear but learnable (DNN) or already well separated in feature space (kNN).

In contrast, AdaBoost.M2 exhibits substantially lower performance (very low Accuracy and Macro-F1 and Macro-AUC close to 0.5). In practice, an AUC close to 0.5 indicates behavior close to random ranking. This gap suggests that the boosting configuration (weak learner complexity, number of cycles, learning rate, or class-score calibration) did not converge to a meaningful separation of the six classes under the current settings.

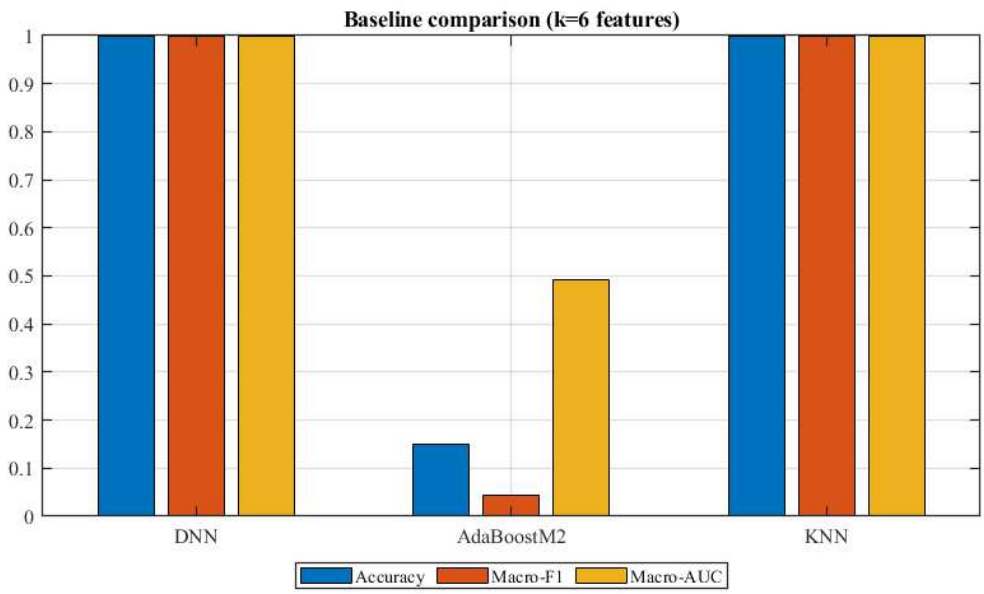


Figure 8. Bar-chart comparison of the three classifiers (DNN, AdaBoost.M2, and kNN) in terms of Accuracy, Macro-F1, and Macro-AUC on the test set when using the complete MCSA feature vector (k = 6).

Figure 9 provides the confusion matrices, which allow a fine-grained interpretation of the misclassifications. DNN (left matrix), the matrix is almost perfectly diagonal, meaning each class is predicted correctly. The absence of off-diagonal counts confirms that the learned decision boundaries generalize well across the test samples and do not confuse neighboring severity levels (e.g., U70% vs. U80%). For kNN (right matrix), similarly, the matrix is perfectly (or near-perfectly) diagonal. This is a strong indication that the classes form compact and well-separated clusters in the standardized feature space, so that a local neighborhood rule is sufficient to classify all samples. AdaBoost.M2 (middle matrix), the matrix shows a highly degenerate behavior, with most samples being assigned to a single class (dominant column/row effect). This pattern typically occurs when the ensemble collapses to predicting one label due to poor weak-learner discrimination, an unsuitable number of boosting rounds, excessive regularization, or issues in score-to-class conversion. From a diagnostic perspective, this means AdaBoost.M2 (in its current configuration) fails to capture the class structure present in the feature space.

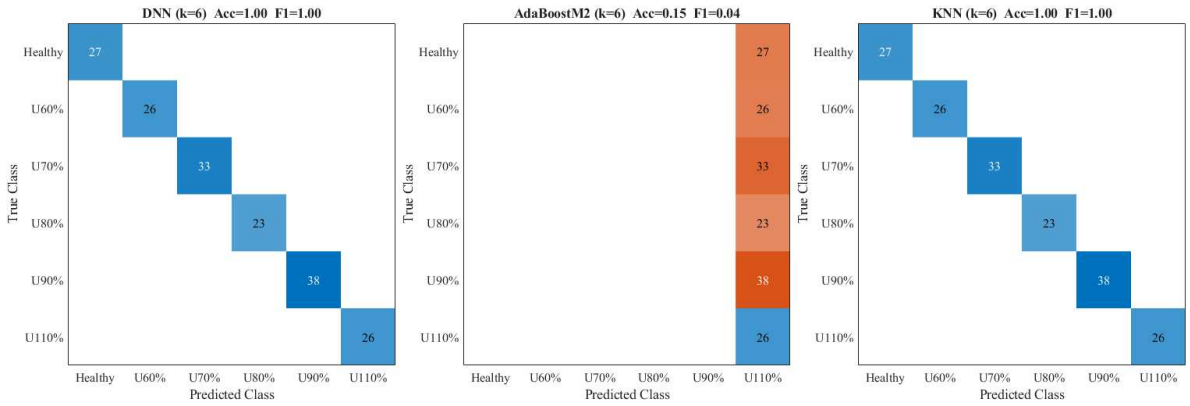


Figure 9. Test-set confusion matrices obtained with DNN, AdaBoost.M2, and kNN using the full feature set (k = 6).

The macro-averaged ROC curves in Figure 10 confirms the previous observations. Both DNN and kNN produce ROC curves close to the top-left corner, with Macro-AUC = 1.0, indicating perfect separability and ranking of the correct class against the rest (one-vs-rest, then macro-averaged). This supports the conclusion that the selected MCSA indicators are extremely informative for the considered unbalance severity classes. Conversely, the AdaBoost.M2 ROC curve follows the diagonal line, with Macro-AUC around 0.49, consistent with near-random discrimination. This further indicates that the boosting model, as currently tuned, did not learn a useful mapping from features to class labels.

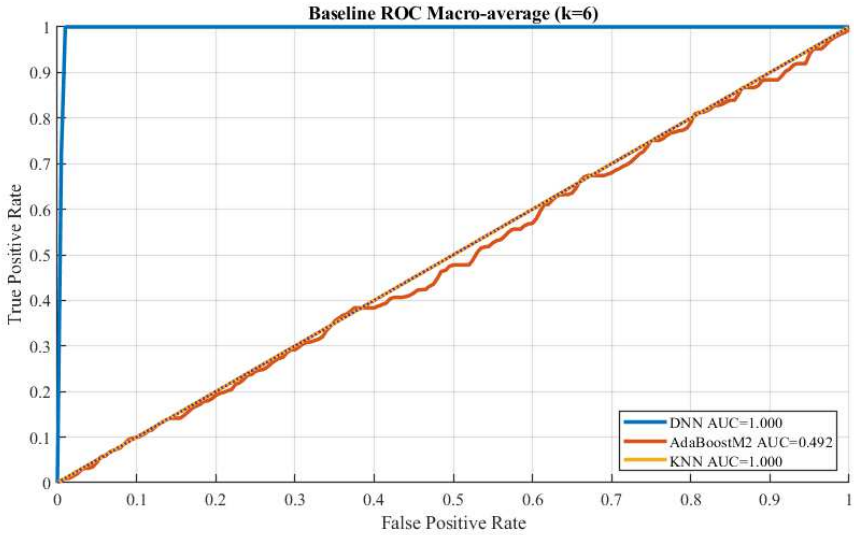


Figure 10. Macro-averaged one-vs-rest ROC curves on the test set for DNN, AdaBoost.M2, and kNN using the complete feature set (k = 6).

7.3.Effect of the Number of Input Features on Classification Performance

This subsection evaluates how the number of input features k affects the classification performance of DNN, AdaBoost.M2, and kNN on the test set. Starting from a single feature and progressively adding features up to the full set ($k = 6$), the objective is to identify the minimum feature subset that preserves high accuracy and to assess the robustness of each model when the input information is reduced.

The confusion matrices (Figure 11) obtained using the best-performing k for each classifier indicate that both DNN and kNN can reach perfect discrimination with a reduced feature set. The DNN achieves a strictly diagonal confusion matrix with $k = 4$, meaning that all severity levels are correctly classified with no confusion between neighboring classes. kNN reaches the same behavior already with $k = 2$, suggesting that the feature space becomes linearly separable at the local scale with only two carefully chosen indicators. In contrast, AdaBoost.M2 attains its best performance at $k = 1$, but still presents off-diagonal errors, highlighting weaker robustness to multi-class separation under the current ensemble configuration.

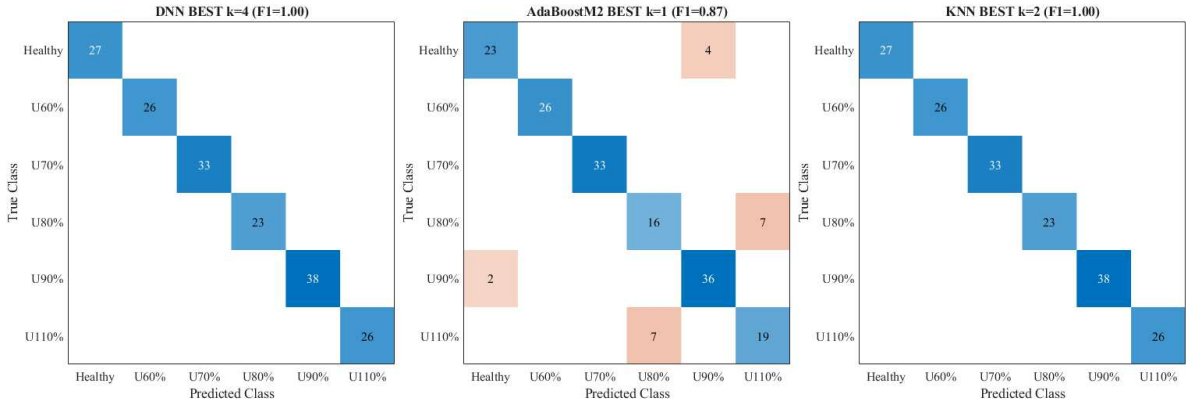


Figure 11. Best-performing confusion matrices using the optimal number of features for each model.

The evolution of Macro-F1 as a function of k presented in Figure 12 confirms these observations. kNN rapidly saturates and reaches Macro-F1 close to 1.00 from $k = 2$ onward, indicating that a small subset of MCSA features is sufficient to produce compact class clusters after standardization. The DNN improves progressively with increasing k, reaching near-perfect Macro-F1 around $k = 4$, which suggests that the neural network benefits from combining complementary descriptors to

refine decision boundaries between close severity levels. Conversely, AdaBoost.M2 shows a sharp degradation when additional features are added: after $k \geq 2$, Macro-F1 collapses to very low values, indicating that the current boosting settings do not exploit multi-dimensional correlated inputs effectively.

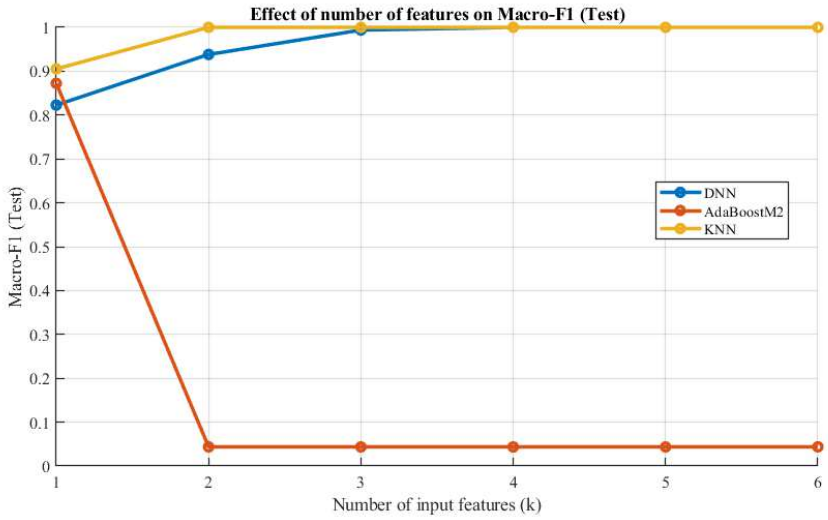


Figure 12. Effect of the number of input features on Macro-F1 (test set).

The accuracy presented in Figure 13 trends follow the same pattern as Macro-F1. kNN attains near-perfect accuracy with only two features and remains stable thereafter, while DNN reaches perfect accuracy once k approaches four. The agreement between Accuracy and Macro-F1 for these two models suggests that performance is not dominated by a single class but is consistent across all severity levels. The heatmap in Figure 14 summarizing Macro-F1 across models and feature numbers further highlights early saturation for kNN, gradual improvement for DNN, and the instability of AdaBoost.M2 once the feature dimension increases.

Finally, the macro-averaged ROC curves (Figure computed using the best k per model show that DNN ($k = 4$) and kNN ($k = 2$) achieve Macro-AUC equal to 1.00, confirming perfect class separability in a one-vs-rest setting. AdaBoost.M2 improves in its best-k scenario compared to the full-feature baseline, but remains below DNN and kNN, indicating incomplete discrimination capability.

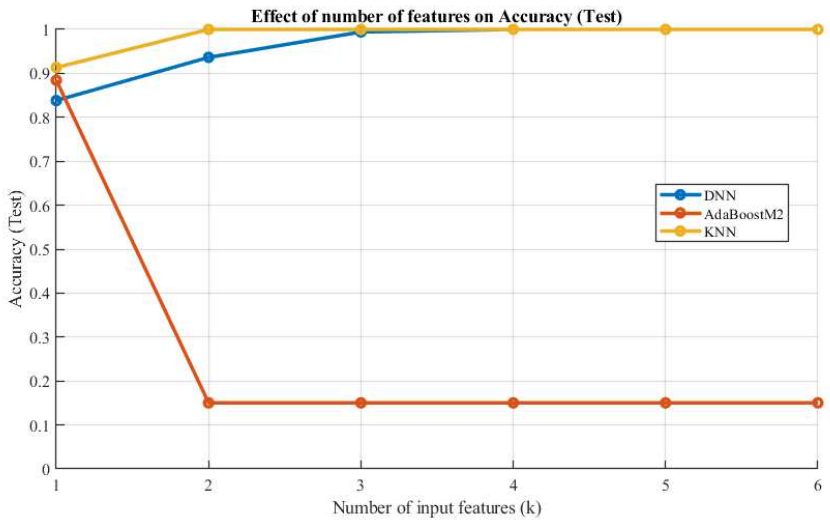


Figure 13. Effect of the number of input features on Accuracy (test set).

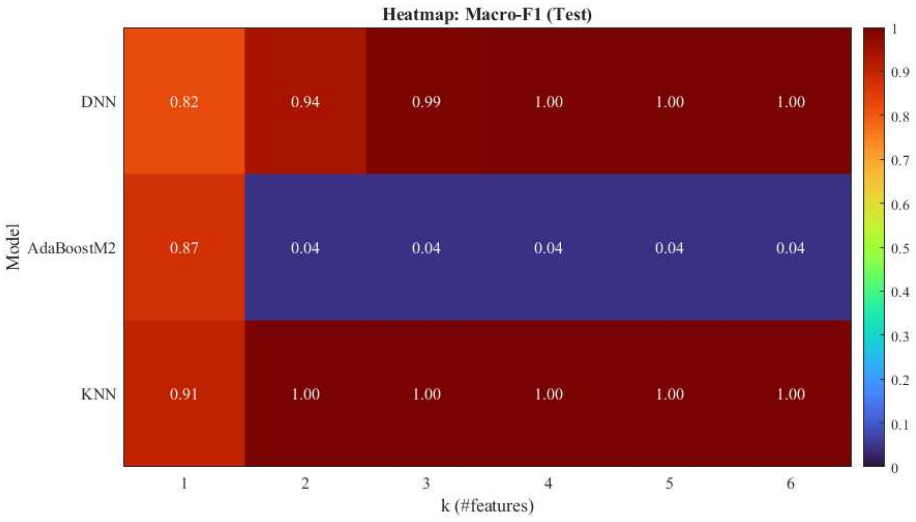


Figure 14. Heatmap summary of Macro-F1 as a function of model type and feature number (test set).

these results in Figure 15, demonstrate that the proposed MCSA feature set is highly informative but partially redundant: near-optimal diagnosis can be achieved with a reduced number of features, which is beneficial for real-time deployment. In this dataset, kNN is the most efficient model in terms of feature economy (perfect results with $k = 2$), while DNN requires slightly more information ($k \approx 4$) to reach the same ceiling. The observed sensitivity of AdaBoost.M2 suggests

that boosting requires careful hyperparameter tuning (learning rate, number of cycles, and weak learner complexity) to reliably benefit from multi-feature inputs.

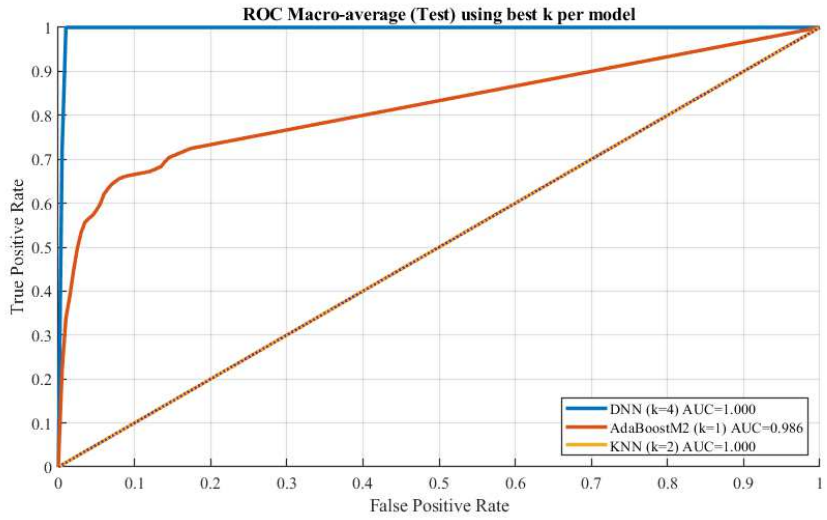


Figure 15. Macro-averaged ROC curves on the test set using the best feature subset per model.

The PCA decision-map in Figure 16 provides an intuitive view of how well the selected best-k = 4 MCSA features separate the six operating conditions. Each point represents one windowed sample, projected from the original 4-D feature space onto the first two principal components. The horizontal axis (PC1) explains 82.8% of the total variance and therefore captures the dominant structure of the dataset, while PC2 explains an additional 9.9%. Together, PC1 and PC2 retain roughly 92.7% of the information contained in the selected features, meaning this 2-D projection is a reliable summary of the original feature space.

A clear class organization is observed mainly along PC1, indicating that the primary discrimination between conditions is driven by one dominant combination of the extracted MCSA descriptors. In particular, the U60% samples appear far to the right (large positive PC1) and form well-separated subclusters, suggesting a strong and consistent signature of this unbalance level that may also vary across operating regimes (e.g., speed). The U70% class is clustered around intermediate positive PC1 values and remains relatively compact, which reflects stable feature behavior under this severity level. The U80% samples concentrate near PC1 close to zero, indicating a moderate

deviation from the healthy reference but still separable from both U70% and the more severe conditions.

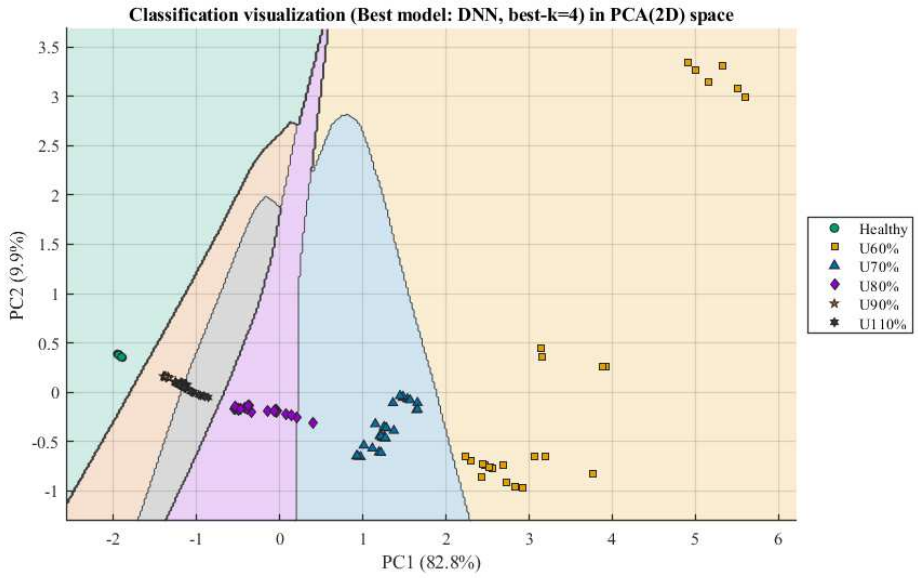


Figure 16. PCA projection (PC1–PC2) of the test samples using the best-k selected MCSA features (k=4), together with the DNN decision regions.

On the left-hand side of the map (negative PC1), the Healthy cluster is compact and clearly isolated from the majority of faulty classes, confirming that the extracted MCSA features are sensitive to the presence of electrical unbalance. The U90% and U110% samples occupy a nearby region around negative PC1 and exhibit partial proximity in this 2-D representation, which suggests that these two high-severity conditions may share similar spectral characteristics and therefore constitute one of the more challenging separations. The spread of these classes in PC2 further indicates that secondary variability—captured by PC2—contributes to refining boundaries between the most similar severities.

The colored background corresponds to the DNN decision regions learned in the same PCA plane and complements the scatter plot by showing the model's global partitioning of the feature space. The fact that most samples fall inside their corresponding decision zones confirms that the selected low-dimensional feature set is highly discriminative and that the DNN operates with large margins for most classes. The decision boundaries become tighter mainly in areas where classes are closest

(typically between adjacent severities), which is consistent with the physical expectation that electrical unbalance levels evolve gradually and may produce partially overlapping signatures.

8. Conclusion

This study proposed a compact and data-driven framework for the multi-class severity classification of electrical unbalance in variable-speed induction motors using low-dimensional MCSA features. The methodology was validated on an experimental dataset built from one healthy operating condition and five unbalance severity levels, using synchronized measurements of electrical variables and rotational speed. By combining physically interpretable indicators with supervised learning, the proposed framework was designed to assess whether a reduced and meaningful feature space can preserve reliable diagnostic performance under variable-speed conditions.

- The statistical analysis of the extracted features demonstrated a clear separation between the healthy condition and the faulty classes, confirming that the selected low-dimensional descriptors effectively capture the progressive impact of electrical unbalance on motor behavior. In particular, the most unbalance-sensitive indicators exhibited consistent and structured variations across the considered severity levels.
- The correlation analysis showed that several descriptors are strongly related, indicating that the six-feature set contains partially redundant information. This result is important because it justifies the feature-number study and supports the possibility of reducing the dimensionality of the diagnostic representation without significantly degrading classification performance.
- The PCA-based feature-space visualization confirmed that the proposed descriptors generate a well-structured and separable representation of the six operating conditions. The healthy state formed a compact and clearly isolated cluster, whereas the faulty classes remained ordered mainly along the first principal component, which reflects the progression of unbalance severity.
- In the baseline comparison performed with the complete feature set, both DNN and kNN achieved near-perfect classification results, with Accuracy and Macro-F1 close to unity and Macro-AUC equal to 1.0. The associated confusion matrices were almost perfectly diagonal,

showing that the proposed features preserve sufficient information to discriminate not only healthy and faulty conditions, but also neighboring unbalance severity levels under variable-speed operation.

- By contrast, AdaBoost.M2 showed substantially lower performance in its current configuration. Its confusion matrix revealed a degenerate prediction behavior and the corresponding ROC performance remained close to random discrimination, indicating that this model did not adequately exploit the structure of the extracted feature space under the present tuning settings.
- The feature-number analysis provided one of the most significant outcomes of the study. It demonstrated that perfect or near-perfect diagnosis can be maintained with a reduced subset of descriptors: kNN reached perfect classification with only two features, while the DNN achieved the same performance with four features. This confirms that the proposed low-dimensional MCSA representation is not only discriminative, but also compact and suitable for computationally efficient implementation.
- The evolution of Macro-F1, Accuracy, and Macro-AUC with the number of input features further confirmed that kNN rapidly saturates with a very small subset of indicators, whereas DNN benefits from progressively combining complementary features. These results highlight the robustness of the proposed framework and show that accurate severity diagnosis can be achieved without relying on high-dimensional feature engineering.

From an applied and industrial perspective, these findings confirm that low-dimensional MCSA-based diagnosis constitutes a promising solution for predictive maintenance of variable-speed induction motor systems. Early identification of electrical unbalance severity can contribute to better maintenance planning, improved machine availability, and reduced unexpected stoppages in production environments. Future work should therefore focus on extending the framework to wider load conditions, combined fault scenarios, and additional machines, while also integrating real industrial datasets and online deployment constraints. A further relevant perspective would be to explicitly combine the present current-based descriptors with vibration-related indicators in order to strengthen the relationship between electrical fault severity, electromechanical disturbances, and production-oriented condition monitoring.

Author Contributions Conceptualization: Yacine KARMI, AbdelFettah BOUSSAID, Omar REFFAS, and Abderrahim TOUIL. Methodology: Yacine KARMI and Haithem BOUMEDIRI. Software: Yacine KARMI and Haithem BOUMEDIRI. Validation: AbdelFettah BOUSSAID, Haithem BOUMEDIRI, Sabbah ATAYA, and Rashid KHAN. Formal analysis: AbdelFettah BOUSSAID, Yacine KARMI, and Omar REFFAS. Investigation: AbdelFettah BOUSSAID, Yacine KARMI, and Abderrahim TOUIL. Resources: AbdelFettah BOUSSAID, Sabbah ATAYA, and Rashid KHAN. Data curation: Yacine KARMI, Haithem BOUMEDIRI, and Omar REFFAS. Writing – original draft preparation: AbdelFettah BOUSSAID, Yacine KARMI, Haithem BOUMEDIRI, and Omar REFFAS. Writing – review and editing: Haithem BOUMEDIRI, Sabbah ATAYA, and Rashid KHAN. Visualization: Yacine KARMI and Omar REFFAS. Supervision: AbdelFettah BOUSSAID, Yacine KARMI, and Haithem BOUMEDIRI. Project administration: Haithem BOUMEDIRI, Sabbah ATAYA, and Rashid KHAN. All authors have read and agreed to the published version of the manuscript.

Funding This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2603).

Availability of data and materials the data are not publicly available due to the Laboratory restrictions.

Declarations

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

References

1. Talhaoui H, Menacer A, Kessal A, Tarek A (2017) Experimental diagnosis of broken rotor bars fault in induction machine based on Hilbert and discrete wavelet transforms. Int J Adv Manuf Technol 2017 951 95:1399–1408. https://doi.org/10.1007/s00170-017-1309-7
2. Attoui I, Oudjani B, Boutasseta N, et al (2020) Novel predictive features using a wrapper model for rolling bearing fault diagnosis based on vibration signal analysis. Int J Adv Manuf Technol 2020 1067 106:3409–3435. https://doi.org/10.1007/s00170-019-04729-4
3. Gangsar P, Bajpei AR, Porwal R (2022) A review on deep learning based condition monitoring and fault diagnosis of rotating machinery. Noise Vib Worldw 53:550–578. https://doi.org/10.1177/09574565221139638
4. Bahgat BH, Elhay EA, Sutikno T, Elkholy MM (2024) Revolutionizing motor maintenance: a comprehensive survey of state-of-the-art fault detection in three-phase induction motors. Int J Power Electron Drive Syst 15:1968–1989. https://doi.org/10.11591/ijpeds.v15.i3.pp1968-1989
5. Mohamed MA, Mohamed AAA, Abdel-Nasser M, et al (2021) Induction motor broken rotor bar faults diagnosis using ANFIS-based DWT. Int J Model Simul 41:220–233. https://doi.org/10.1080/02286203.2019.1708173
6. Benbouzid MEH (2000) A review of induction motors signature analysis as a medium for faults detection. IEEE Trans Ind Electron 47:984–993. https://doi.org/10.1109/41.873206

7. Gherghina IS, Bizon N, Iana GV, Vasilică BV (2025) Recent Advances in Fault Detection and Analysis of Synchronous Motors: A Review. Mach 2025, Vol 13, Page 815 13:815. https://doi.org/10.3390/machines13090815
8. Xu X, Qiao X, Zhang N, et al (2020) Review of intelligent fault diagnosis for permanent magnet synchronous motors in electric vehicles. Adv Mech Eng 12:. https://doi.org/10.1177/1687814020944323
9. Xie P, Zhang Z, Liao H, et al (2025) Industrial motorized spindle systems: challenges and trends in bearing, thermal, and fault diagnosis. Int J Adv Manuf Technol 2025 1419 141:4701–4746. https://doi.org/10.1007/s00170-025-16905-w
10. Li H, Zhu ZQ, Azar Z, et al (2025) Fault Detection of Permanent Magnet Synchronous Machines: An Overview. Energies 2025, Vol 18, Page 534 18:534. https://doi.org/10.3390/en18030534
11. Thomson WT, Fenger M (2001) Current signature analysis to detect induction motor faults. IEEE Ind Appl Mag 7:26–34. https://doi.org/10.1109/2943.930988
12. Dash RN, Subudhi B (2010) Stator inter-turn fault detection of an induction motor using neuro-fuzzy techniques. Arch Control Sci 20:363–376. https://doi.org/10.2478/v10170-010-0022-7
13. Maraaba L, Al-Hamouz Z, Abido M (2018) An Efficient Stator Inter-Turn Fault Diagnosis Tool for Induction Motors. Energies 2018, Vol 11, Page 653 11:653. https://doi.org/10.3390/en11030653
14. Wang X, Zhou J, Qin J, et al (2025) Stator inter-turn shortcircuit fault diagnosis of induction motor based on characteristic current. COMPEL - Int J Comput Math Electr Electron Eng 44:19–33. https://doi.org/10.1108/COMPEL-07-2024-0280
15. Cherif H, Benakcha A, Laib I, et al (2020) Early detection and localization of stator inter-turn faults based on discrete wavelet energy ratio and neural networks in induction motor. Energy 212:118684. https://doi.org/10.1016/j.energy.2020.118684
16. Allal A, Lammouchi Z, Khechekhouche A, et al (2024) Advanced Residual Optimal Mapping Approach for Precise Detection of Stator Faults in Induction Motors. IEEE Access 12:143515–143530. https://doi.org/10.1109/ACCESS.2024.3442671
17. Bellini A, Filippetti F, Franceschini G, et al (2001) Quantitative evaluation of induction motor broken bars by means of electrical signature analysis. IEEE Trans Ind Appl 37:1248–1255. https://doi.org/10.1109/28.952499
18. Filippetti F, Franceschini G, Tassoni C, Vas P (1998) AI techniques in induction machines diagnosis including the speed ripple effect. IEEE Trans Ind Appl 34:98–108. https://doi.org/10.1109/28.658729
19. Alimardani R, Rahideh A, Hedayati Kia S (2025) Mixed Eccentricity Fault Detection of Induction Motors Based on Variational Mode Decomposition of Current Signal. Mach 2025, Vol 13, Page 968 13:968. https://doi.org/10.3390/machines13100968
20. Niu Y, Fei J, Li Y, Wu D (2020) A novel fault diagnosis method based on EMD, cyclostationary, SK and TPTSR. J Mech Sci Technol 2020 345 34:1925–1935. https://doi.org/10.1007/s12206-020-0414-y
21. Lin Z, An-quan S, Xin-yu Y, et al (2013) Air-gap Eccentricity Fault Diagnosis of Induction Motors Using EMD. Indones J Electr Eng Comput Sci 11:4168–4174. https://doi.org/10.11591/telkomnika.v11i7.2936
22. Arifin MS, Wang W, Uddin MN (2024) A Modified EMD Technique for Broken Rotor Bar Fault Detection in Induction Machines. Sensors 2024, Vol 24, Page 5186 24:5186. https://doi.org/10.3390/s24165186
23. Chen Y, Zhao L, Li L, et al (2025) Digital Twin-Based Online Diagnosis Method for Inter-Turn Short Circuit Fault in Stator Windings of Induction Motors. Energies 2025, Vol 18, Page 3063 18:3063. https://doi.org/10.3390/en18123063

24. Zhu H, Jia Z, Song X, Sun W (2022) An approach to detect broken rotor bars based on instantaneous frequency of the fault characteristic harmonic during the start-up transient. Int J Adv Manuf Technol 2022 12411 124:4107–4119. https://doi.org/10.1007/s00170-022-08829-6
25. Liu C, Cichon A, Królczyk G, Li Z (2021) Technology development and commercial applications of industrial fault diagnosis system: a review. Int J Adv Manuf Technol 2021 11811 118:3497–3529. https://doi.org/10.1007/s00170-021-08047-6
26. Chen Y, Liang S, Li W, et al (2019) Faults and Diagnosis Methods of Permanent Magnet Synchronous Motors: A Review. Appl Sci 2019, Vol 9, Page 2116 9:2116. https://doi.org/10.3390/app9102116
27. Dias CG, da Silva LC, Chabu IE (2019) Fuzzy-Based Statistical Feature Extraction for Detecting Broken Rotor Bars in Line-Fed and Inverter-Fed Induction Motors. Energies 2019, Vol 12, Page 2381 12:2381. https://doi.org/10.3390/en12122381
28. Ray S, Ganguly B, Dey D (2020) Identification and Classification of Stator Inter-Turn Faults in Induction Motor Using Wavelet Kernel Based Convolutional Neural Network. Electr Power Components Syst 48:1421–1432. https://doi.org/10.1080/15325008.2020.1854384
29. Ince T (2019) Real-time broken rotor bar fault detection and classification by shallow 1D convolutional neural networks. Electr Eng 2019 1012 101:599–608. https://doi.org/10.1007/s00202-019-00808-7
30. Principi E, Rossetti D, Squartini S, Piazza F (2019) Unsupervised electric motor fault detection by using deep autoencoders. IEEE/CAA J Autom Sin 6:441–451. https://doi.org/10.1109/JAS.2019.1911393
31. Kawady TA, Khater WM, Abu-Faty HG, et al (2025) AI driven fault diagnosis approach for stator turn to turn faults in induction motors. Sci Reports 2025 151 15:20125-. https://doi.org/10.1038/s41598-025-04462-x
32. Reffas O, Boumediri H, Karmi Y, et al (2025) Statistical analysis and predictive modeling of cutting parameters in EN-GJL-250 cast iron turning: application of machine learning and MOALO optimization. Int J Adv Manuf Technol 2025 1373 137:1991–2009. https://doi.org/10.1007/s00170-025-15098-6
33. Touati S, Boumediri H, Karmi Y, et al (2025) Performance analysis of steel W18CR4V grinding using RSM, DNN-GA, KNN, LM, DT, SVM models, and optimization via desirability function and MOGWO. Heliyon 11:e42640. https://doi.org/10.1016/j.heliyon.2025.e42640
34. Karmi Y, Boumediri H, Reffas O, et al (2025) Integration of Hybrid Machine Learning and Multi-Objective Optimization for Enhanced Turning Parameters of EN-GJL-250 Cast Iron. Cryst 2025, Vol 15, Page 264 15:264. https://doi.org/10.3390/cryst15030264
35. Selami S, Ataya S, Karmi Y, et al (2026) Hybrid IGWO-DNN approach for bearing fault diagnosis using time domain features from empirical mode decomposition. Int J Adv Manuf Technol 2026 1–24. https://doi.org/10.1007/s00170-026-17817-z
36. Nair N, Johnston H, Du D (2022) A Machine Learning Model for Post-Transplant Lymphoproliferative Disorder Using the UNOS Database. J Hear Lung Transplant 41:S321–S322. https://doi.org/10.1016/j.healun.2022.01.799
37. Siam AI, Zayed AA, Nacpil EJC, Jeon I (2025) A more computationally efficient bearing fault diagnosis system using deep learning classifiers for rotating machinery health monitoring. Int J Adv Manuf Technol 2025 13911 139:5727–5746. https://doi.org/10.1007/s00170-025-16170-x