A Projection-Based Support Vector Machine Algorithm for Induction Motors' Bearing Fault Detection

Narges Khadem
Department of Engineering
University of Cambridge
Cambridge, UK
nk616@cam.ac.uk

Hamid Toshani
Department of Engineering)
University of Cambridge
Cambridge, UK
ht463@cam.ac.uk

Salman Abdi
School of Engineering
University of East Anglia
Norwich, UK
s.abdi-jalebi@uea.ac.uk

Abstract- This paper proposes a binary fault detection algorithm for detecting inner raceway bearing faults in a 4KW induction motor. The algorithm uses Support Vector Machine (SVM) and Projection Recurrent Neural Network (PRNN) techniques and is based on data collected experimentally at different speeds and load conditions. Time and frequency contents of the three-phase stator currents are analysed using Discrete Wavelet Transform (DWT), Power Spectral Density (PSD), and cepstrum analysis. A feature set is obtained using various statistical measures, and feature selection algorithms are used to select the most relevant features. The SVM is then trained using these features, and its optimisation problem is formulated as Constrained Nonlinear Programming (NCP). A PRNN is proposed to solve the NCP and obtain the optimal decision boundary of the SVM. The study demonstrates that the accuracy of the algorithm depends on the type of kernel function and the number of relevant features selected. The results suggest that the proposed algorithm is effective in detecting inner raceway bearing faults in induction motors.

Keywords—Bearing fault, SVM, PRNN, Feature selection.

I. INTRODUCTION

Induction Motors (IMs) are one of the most widely used types of electric motors in the industry due to their reliability, robustness, and relative affordability compared to other motor types. They are commonly used in applications such as pumps, compressors, fans, wind generators, and solar panels [1]. However, over time, IMs may experience wear and tear, overloading, and misalignments, leading to bearing faults. These types of failures are responsible for 51% of the downtime experienced by IMs [2]. Early identification and diagnosis of bearing faults are critical to prevent further damage and avoid costly repairs, downtime, and loss of production. Therefore, it is essential to develop accurate and reliable fault detection and diagnosis techniques for IMs to ensure their efficient and continuous operation.

Mechanical fault detection methods, such as acoustic monitoring, thermal imaging, and vibration analysis, can be utilised to detect potential bearing faults at an early stage before they become severe issues. Acoustic monitoring is a non-invasive method that uses sensors to quantify and analyse the audible emissions from the motor [3]. Thermal imaging can identify various types of bearing faults through thermal image processing, feature extraction, and machine learning algorithms [4]. Vibration analysis is a widely used technique for detecting potential defects or irregularities in bearings. It involves evaluating the vibration characteristics associated with engine bearings for different characteristic fault frequencies using a frequency-based method such as Fast Fourier Transform Spectral Analysis [5]. However, these methods can be affected by external noise and ambient temperature, are subject to sensor failure, and require specialised and expensive equipment. Nevertheless, they remain useful tools for early fault detection and are widely employed in industrial applications.

A commonly used technique for identifying bearing faults is known as Motor Current Signature Analysis (MCSA). The basic idea is to examine the frequency spectrum of both healthy and faulty data to isolate the fault harmonic frequencies. This technique is a non-invasive and low-cost method that can be used to detect outer ring fault frequencies in induction motors using stator current signals and their probability distributions [6]. MCSA is sensitive to speed ripple, voltage variations, and ambient noise [7]. It is unable to detect incipient faults, and it is difficult to discriminate between different types of faults that may have similar signature patterns. These drawbacks can be mitigated by integrating MCSA with techniques such as Discrete Wavelet Transform (DWT) and Fast Fourier Transform (FFT). The first step is to convert the sampled data to the time-frequency domain using wavelet transform and then decompose it using FFT to investigate the frequency components that represent the bearing fault [8]. MCSA can also be combined with noise cancellation and wavelet analysis to estimate fault severity using the fault indexing parameters of the Power Spectral Density (PSD) [9]. These techniques assume that the signal is stationary. They are computationally intensive and are affected by the length of the time window used

Machine learning algorithms, such as SVM, are capable of detecting and classifying various types of defects in IMs. SVM algorithms have been found to achieve high accuracy in fault classification while being robust to overfitting and able to handle noise in the data. Additionally, SVM algorithms require fewer hyperparameters to be tuned than other machine learning algorithms, such as artificial neural networks, making it easier to find the optimal configuration for the classification model. SVM algorithms can handle high-dimensional data and are able to identify the separating hyperplane [10]. In [11], the motor dynamic strain signals are pre-processed using FFT to obtain a dimensionality-reduced dataset by selecting the four highest picks of the PSD and then using Principal Component Analysis. An SVM algorithm is extended to detect the outer raceway bearing fault, with its main contribution being the ability to detect different severity levels of the bearing fault. In [12], an optimised stationary wavelet

packet transform is used to extract the features from the motor current signals. A dataset containing the root mean square of the wavelet coefficients is then used to train and test an SVM-based bearing fault detection algorithm. This method is characterised by optimised wavelet kernel functions. Similarly, in [13], a continuous wavelet transform is used to process the vibration signals and then, after feature extraction, an SVM algorithm based on linear kernel functions is applied to detect the bearing fault. The accuracy of the fault detection is evaluated in terms of the mother wavelet selection. When compared with artificial neural networks, SVM algorithms can achieve more accurate results with a simpler structure.

For a large dataset with nonlinear kernel functions, solving the Constrained Optimisation Problem (COM) in SVM can be challenging. Projection recurrent neural networks have the ability to solve a variety of COMs with linear and nonlinear constraints. They are fast, have a simple structure, are easy to implement, and are asymptotically stable [14]. These optimisers are used to solve quadratic programming [15], and Constrained Nonlinear Programming (CNP) [16]. PRNNs can be integrated with the SVM to develop a fast, simple, and optimised classifier for bearing fault detection in the IM.

In this paper, three-phase current signals are collected experimentally at different speeds from no load to full load for both healthy and faulty conditions. A feature set is provided using the time-domain signals, their DWT including detail and approximation coefficients, PSD, and cepstrum. A feature selection algorithm is developed to select the most relevant features for the training and testing datasets. The COM of SVM is formulated as a CNP. A PRNN, including dynamic and algebraic equations, is integrated into the SVM to minimise the CNP and obtain the optimal classification boundary. The rest of this paper is structured as follows: Section II presents the proposed fault detection algorithm. The experiments including the test rig and instruments are described in Section III. The data processing including obtaining DWT, PSD, and cepstrum from the current signals, is presented in Section IV. Sections V and VI present feature extraction and feature selection, respectively. Section VII describes the integration of SVM and PRNN. The classification results and the conclusion are presented in Sections VIII, and IX, respectively.

II. Proposed Bearing Fault Detection Algorithm

As depicted in Figure 1, the bearings are composed of two rings, namely, the outer and inner rings, along with a cage containing evenly spaced balls to prevent any mutual contact [11]. The bearing fault can be characterised as a crack on the inner or outer raceway with early, intermediate, and severe levels. This paper focuses on detecting an intermediate bearing fault associated with the outer raceway. In the proposed method, only the three-phase currents from a 4 KW squirrel cage IM are utilised for detecting the bearing fault. Both healthy and faulty raw data are acquired at various speeds of the IM under different load conditions, ranging from no load to full load. Additionally, the DWT, PSD, and cepstrum of the current signals are computed to investigate the fault characteristics in both time and frequency domains.

Several statistical features are subsequently extracted from the time and frequency signals, which include mean, standard deviation, kurtosis, skewness, crest factor, shape factor, impulse factor, and clearance factor.

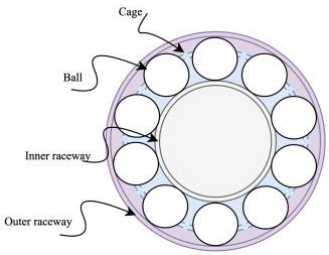


Fig. 1. Main parts of a rolling bearing

A feature selection algorithm is applied to determine the most prominent features. Consequently, a dimensionally reduced feature set is utilised to train the SVM and identify the defect. To simplify the optimisation process and render it suitable for online classification and practical implementation, a PRNN is integrated with the SVM as a numerical optimiser to determine the optimal classification boundary. Figure 2 illustrates the block diagram of the proposed fault detection algorithm, commencing from the data acquisition system and concluding with the detection of the bearing fault.

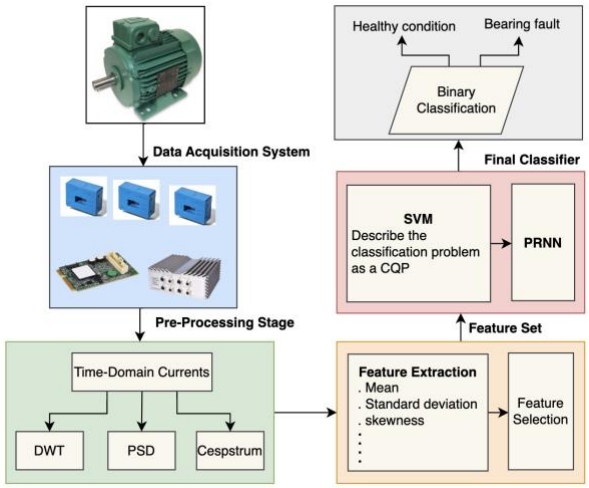


Fig. 2. Block diagram of the fault detection algorithm

III. Experimental Test Rig Set-up

The experiment is carried out using a test rig with a 4KW induction motor coupled to a 4KW load motor as shown in Fig. 3. A torque transducer is mounted between the motors to measure the torque signals. A data acquisition system consisting of three current sensors, three voltage sensors, a Speedgoat and an input/output (I/O) module is set up to measure in particular the instantaneous three-phase currents for both healthy and fault conditions. The Speedgoat is a target machine that has an Intel 2.0 GHz quad-core CPU, Simulink Real-Time operating system, 4GB DDR3 RAM memory and 4 I/O slots. The I/O module is a configurable FPGA board with 13 differential I/O lines.

The current sensors referred to in this study are single-channel Hall effect sensors with a closed-loop bi-directional module specifically designed to measure electric currents of up to 100A AC/DC. The measurement data obtained from these sensors is collected using the Simulink Real-Time toolbox of MATLAB, with a sampling time of one millisecond, and stored on the computer.

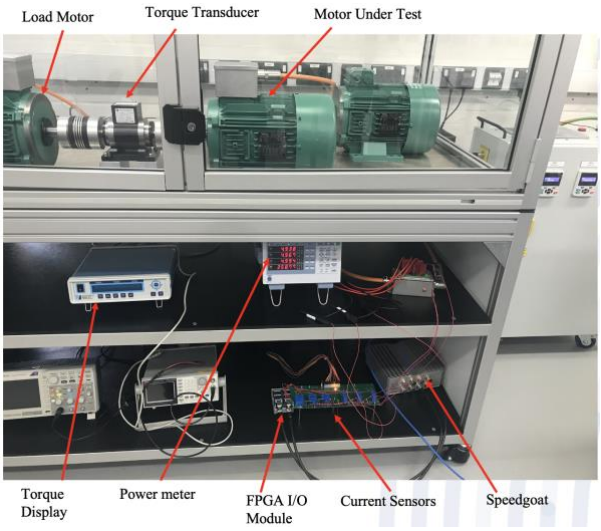


Fig. 3. A view of the experimental rig and instrumentation system

IV. Monitoring Signal Processing Method

Different motor faults may exhibit distinct characteristics in either the time or spectral domain of the monitoring signal. For instance, a sudden change in signal amplitude can be easily recognised as a sharp spike or step change when observed in the time domain. Conversely, irregularities that cause alterations in the frequency content of signals are more readily identified by examining their spectra. By simultaneously scrutinising both types of time and frequency analysis, it is possible to identify and diagnose a broader range of potential faults within the given system. In this paper, a high-performance data acquisition system was used to initially acquire three-phase current signals from a 4KW IM. This initial dataset was obtained at various speeds and load conditions with a sampling time of 1 millisecond. A time window of 20 seconds was considered for each operating mode.

For a real-time function $x(t)$, DWT is given by [8]

$$DWT(m,n)=\frac{1}{\sqrt[m]{a_0}}\sum_k x(k)\psi\left(\frac{k-nb_0a_0^m}{a_0^m}\right) \tag{1}$$

where a_0 and b_0 are the scale and translation factors, respectively, m is the decomposition level, n is the translation step, ψ is the wavelet function and $t=kT$, (k=0,1,…) where T is the sampling time. Low-pass and high-pass wavelet decomposition filters based on the biorthogonal spline wavelet are used to obtain the approximation and detail coefficients. Welch's spectrum estimation is employed to calculate the PSD, where the current signals are initially divided into overlapping segments. The periodogram for each segment is subsequently computed, and the outcomes are averaged to provide an overall estimation of the PSD as follows [17]:

$$PSD(w)=\frac{1}{n_s}\sum_{j=1}^{n_s}\left(\frac{1}{MP}\left|\sum_{k=1}^{M}v(k)x_j(k)e^{-iwk}\right|^2\right) \tag{2}$$

where n_s is the number of samples, M is the number of segments, $v(k)$ is the Hamming window as

$$v(k)=0.54-0.46cos\left(2\pi\frac{k}{N}\right),0\le k\le N \tag{3}$$

where N+1 is the window length, and P denotes the power of Hamming window as

$$P=\frac{1}{M}\sum_{k=1}^{M}\left|v(k)\right|^2 \tag{4}$$

To calculate the cepstrum of a signal x(t), the initial step involves taking the Fourier transform of x(t). Subsequently, the magnitude of the Fourier transform is calculated, followed by taking the logarithm of that magnitude. Finally, the logarithm is subjected to the inverse Fourier transform [18]:

$$C_x(n)=\frac{1}{N}\sum_{k=0}^{N-1}log\left|\sum_{j=0}^{N-1}x(n)e^{\frac{-i2\pi nj}{N}}\right|e^{\frac{i2\pi nk}{N}} \tag{5}$$

Figure 4 demonstrates an instance of the three-phase currents and their wavelet detail coefficients, PSD, and cepstrum at one of the operational modes of the defective motor.

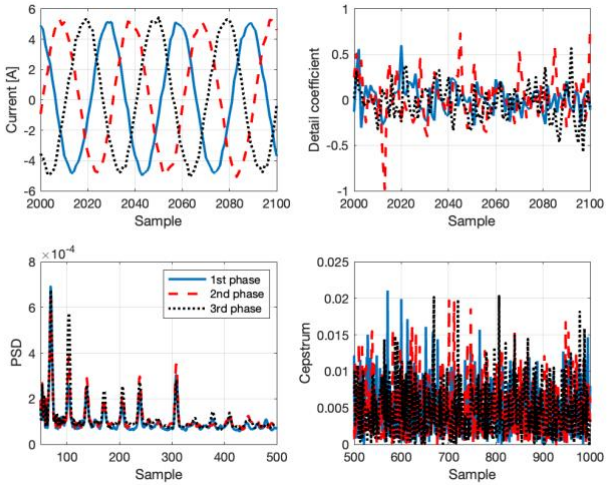


Fig. 4. Three-phase stator currents of the defective motor, their DWT, PSD and cepstrum

V. Feature Extraction for Bearing Fault Detection

Feature extraction is a critical step in fault detection because it transforms raw data into a set of relevant features that can be used to train machine learning models. It can reduce the dimensionality of the data, improve the signal-to-noise ratio, and identify fault signatures. In this paper, several statistical features are applied to the processed data which are listed in Table I. The feature set is provided by applying these features to three-phase currents, wavelet approximation and detail coefficients, PSD and cepstrum signals. This means that there are 40 features corresponding to each operating mode for each phase of either a healthy or faulty motor. An example of the features obtained for the first phase of the faulty motor is shown in Figure 5.

VI. Feature Selection Process

The objective of feature selection is to identify the most pertinent and informative features in a given dataset for the development of defect detection machine learning models. This process can enhance model effectiveness whilst reducing complexity during training and testing. In this study, an optimal set of features that are mutually and maximally dissimilar and can effectively represent the response variable is determined using the Minimum Redundancy Maximum

Relevance (MRMR) algorithm [19]. The MRMR algorithm evaluates the relevance and redundancy of each feature to the target variable before selecting a subset of information that maximises relevance whilst minimising redundancy.

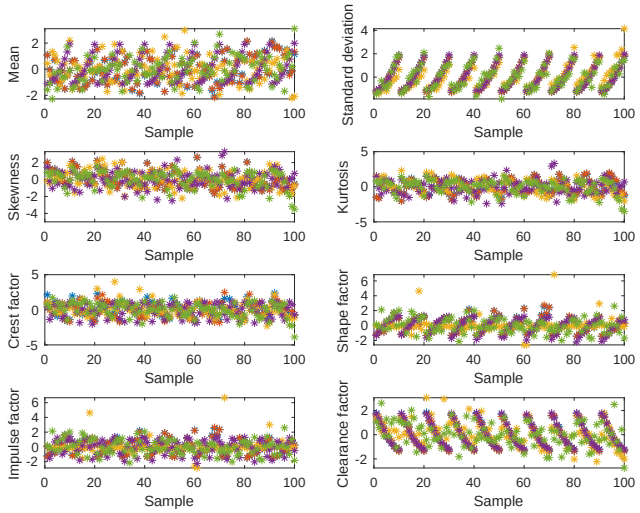


Fig. 5. Features related to the first phase of the faulty motor

TABLE I. STATISTICAL FEATURES WITH THEIR DESCRIPTION AND DEFINITION [3]

Feature	Description	Definition
Mean	The average value of a signal over time.	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
Standard deviation	A signal's dispersion in terms of variance around the mean.	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_{i-} - \mu)^2}{N}}$
Kurtosis	A measure of distribution peakedness.	$k_x = \frac{\sum_{i=1}^{N}(x_{i-} - \mu)^4}{\sigma^4}$
Skewness	A measure of distribution asymmetry.	$s_x = \frac{\sum_{i=1}^{N}(x_{i-} - \mu)^3}{\sigma^3}$
Crest factor	The peak-to-average ratio of a signal.	$crf = \frac{\max_i x_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert^2}}$
Shape factor	A measure of the shape of a signal (RMS/mean).	$sf = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert^2}}{\mu}$
Impulse factor	A measure of the sharpness of a signal.	$If = \frac{\max_i x_i}{\mu}$
Clearance factor	A measure of the alignment of motor components.	$cf = \frac{\max_i x_i}{\left[\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert\right]^2}$

In general, the mutual information between two variables is a measure of the extent to which uncertainty in one variable can be reduced by knowing the other variable. For two random variables x and y, it is defined as follows [19]:

$$I(x,y) = \sum_{i,j} P(x = x_i, y = y_i) log \frac{P(x = x_i, y = y_i)}{P(x = x_i,) P(, y = y_i)} \quad (6)$$

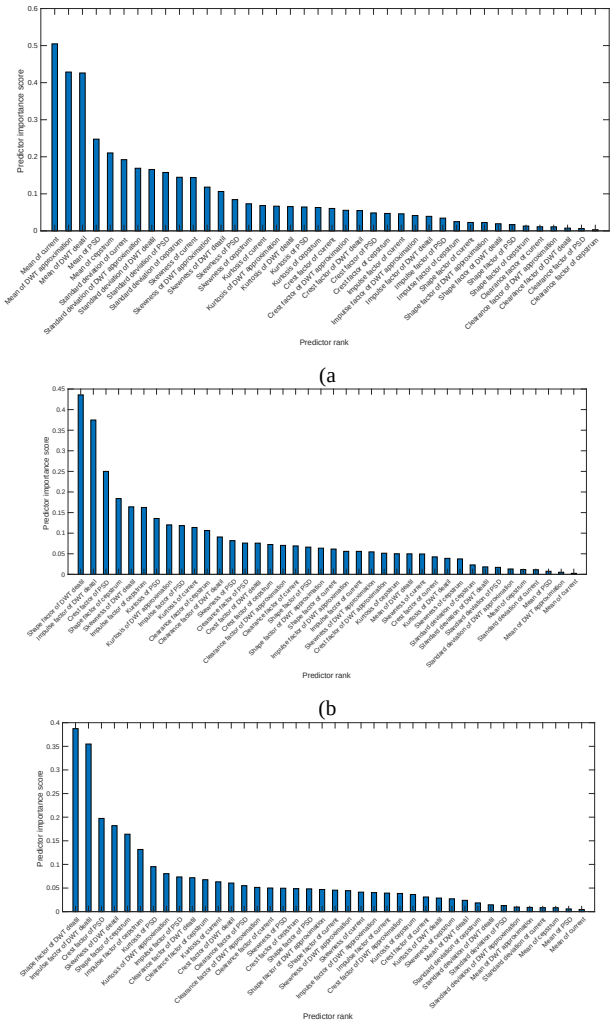
where $P(.)$ is the probability function. In feature selection, the mutual information of a feature with the target variable determines its relevance, but the mutual information between the features determines its redundancy. The algorithm first determines the relevance V_S and redundancy W_S of each feature in a set S with respect to the target variable, y, as

$$V_S = \frac{1}{|S|}\sum_{x \in S} I(x,y) \quad (7)$$

$$W_S = \frac{1}{|S|^2}\sum_{x,z \in S} I(x,z) \quad (8)$$

where $|S|$ is the number of features in S. The MRMR aims to find the most optimal set of features (S) that maximises V_S, and minimises W_S. After calculating these values, the algorithm selects the features that have the largest Mutual Information Quotient (MIQ) value ($\max\left\{\frac{V_{x \in S}}{W_{x \in S}}\right\}$) with nonzero relevance and redundancy in S.

To rank the features based on the MIQ scoring function, a matrix of 40 features is initially provided, with each feature having 200 values composed of both healthy and faulty data. The first eight features relate to time domain signals, while the remaining four groups are associated with the DWT approximation coefficients, DWT detail coefficients, PSD, and cepstrum respectively. Figure 5 displays the feature ranks for the different phases of the IM. Based on these results, the most relevant features can be selected for training and testing the error detection algorithm. For instance, the impulse factor of the PSD, the shape factor of the cepstrum, and the shape factor of the DWT detail coefficients are the three most prominent features of the first phase of the IM.



(a)

(b)

(c)

Fig. 6. Ranking of features for each phase of the IM, a) 1st phase, 2) 2nd phase and 3) 3rd phase

VII. PRNN-SVM-BASED FAULT DETECTION METHOD

SVM is a technique for supervised machine learning which is utilised for regression analysis and classification. The underlying principle behind SVM involves the determination of an appropriate decision boundary or hyperplane for data point classification. The SVM algorithm achieves this objective by maximising the margin or distance between the hyperplane and the nearest points from each class, which are commonly referred to as support vectors [20].

In the case of a set of training data $\{(x_1, z_1), \ldots, (x_n, z_n)\}$ where x_i represents the ith input vector and y_i denotes its corresponding class label, the objective of SVM is to identify a hyperplane $(\mathbf{w}, h)$ that satisfies the following optimisation problem:

$$\begin{aligned} &\min_{\mathbf{w}^i} \mathbf{w}^T(\mathbf{w}) \\ &s.t. \quad z_i\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\mathbf{w}+h\right) \geq 1,\ i=1,2,...,n \end{aligned} \tag{9}$$

where $\mathbf{w}$ is the weight vector perpendicular to the hyperplane, h is the bias term and $\boldsymbol{\psi}(\boldsymbol{x_i})$ is the kernel function. Consider the dual of (9) using the Lagrange function as

$$\mathrm{L}\left(\mathbf{w}, h, \lambda^i\right) = \frac{1}{2}(\mathbf{w})^T(\mathbf{w}) - \sum_{i=1}^{n}\left(\lambda^i\left(z_i\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\mathbf{w}+h\right)-1\right)\right) \tag{10}$$

where $\lambda^i \geq 0$ is the Lagrange multiplier. Derivative of (10) to $\mathbf{w}$ and h returns:

$$\begin{cases} \dfrac{\partial \mathrm{L}\left(\mathbf{w}, h, \lambda^i\right)}{\partial \mathbf{w}} = w - \sum_{i=1}^{n}\lambda^i z_i\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)^T = 0 \\ \dfrac{\partial \boldsymbol{L}\left(\mathbf{w},\ \boldsymbol{h}, \lambda^i\right)}{\partial h^i} = \sum_{i=1}^{n}\lambda^i z_i = 0, i=1,2,...,n \end{cases} \tag{11}$$

Substituting (11) to (10) yields:

$$\mathrm{L}\left(\mathbf{w}, h, \lambda^i\right) = \sum_{i=1}^{n}\lambda^i - \frac{1}{2}\sum_{i,j=1}^{n} z_i z_j \lambda^i \lambda^j \left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)^T \tag{12}$$

Assuming λ^{i*} is the optimal solution of (11), the CNP of SVM can be expressed as follows [20]:

$$\begin{aligned} &\max_{\lambda^i} \sum_{i=1}^{n}\lambda^i - \frac{1}{2}\sum_{i,j=1}^{n} z_i z_j \lambda^i \lambda^j \left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)^T \\ &s.t. \sum_{i=1}^{n} z_i \lambda^i = 0,\ \ \lambda^i \geq 0, i, 1, 2, ..., n \end{aligned} \tag{13}$$

The optimal weight vector and bias term are as follows:

$$\begin{aligned} \mathbf{w}^* &= \sum_{i=1}^{n} z_i \lambda^{i*}\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)^T \\ h^* &= -\frac{\max_{z_i=1}\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\mathbf{w}^*\right) + \min_{z_i=-1}\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\mathbf{w}^*\right)}{2} \end{aligned} \tag{14}$$

To solve (13), consider the following general CNP:

$$\begin{aligned} &\boldsymbol{min} \quad \boldsymbol{f}(\mathbf{x}) \\ &\text{s.t. } \mathbf{g}(\mathbf{x}) \leq 0,\ \mathbf{h}(\mathbf{x}) = 0 \end{aligned} \tag{16}$$

where $\mathbf{x} \in R^n$, $f: R^n \to R$, $\mathbf{g}: R^n \to R^m$, and $\mathbf{h}: R^n \to R^r$. The Karush-Kuhn-Tucker (KKT) optimal conditions for optimal point $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is given by [16]

$$\begin{cases} \mathbf{y} \geq 0, \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{h}(\mathbf{x}) = 0 \\ \nabla f(\mathbf{x}) + \nabla \mathbf{g}^T(\mathbf{x})\mathbf{y} + \nabla \mathbf{h}^T(\mathbf{x})\mathbf{z} = 0, \mathbf{y}^T\mathbf{g}(\mathbf{x}) = 0 \end{cases} \tag{17}$$

where $\mathbf{y}$ and $\mathbf{z}$ are the Lagrange multipliers. For a closed set $\Omega = \{\mathbf{x} \in R^n | \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{h}(\mathbf{x}) = 0\}$, $\mathbf{x}^*$ is the optimal solution of (14) if and only if the following equality holds:

$$Pr_{\Omega}\left(\mathbf{x} - \alpha \mathbf{F}(\mathbf{x})\right) = \mathbf{x} \tag{18}$$

where $\alpha > 0$ is a positive scalar and $Pr_{\Omega}(.)$ is the projection operator, which is given $Pr_{\Omega}(x) = arg\min_{v \in \Omega} x - v$.

By corresponding (13) and (16), it can be written:

$$\begin{aligned} f(\boldsymbol{\lambda}) &= \frac{1}{2}\sum_{i=1}^{n} z_i z_j \lambda^i \lambda^j \left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)\left(\boldsymbol{\psi}(\boldsymbol{x}_i)\right)^T - \sum_{i=1}^{n}\lambda^i \\ h(\boldsymbol{\lambda}) &= \sum_{i=1}^{n} z_i \lambda^i = 0,\ \ g(\boldsymbol{\lambda}) = -\lambda^i \leq 0, i, 1, 2, ..., n \end{aligned} \tag{19}$$

According to [16], the optimal $(\boldsymbol{\lambda}^*, \mathbf{y}^*, \mathbf{z}^*)$ is the solution of the following conditions:

$$\begin{cases} \mathbf{y} = Pr\left(\mathbf{y} + \mathbf{g}(\boldsymbol{\lambda})\right), z^T\boldsymbol{\lambda} = 0 \\ \nabla f(\boldsymbol{\lambda}) + \nabla \mathbf{g}^T(\boldsymbol{\lambda})\mathbf{y} - \mathbf{y}^T\mathbf{g}(\boldsymbol{\lambda}) = 0 \end{cases} \tag{20}$$

where $\mathrm{Pr}(\delta) = \max\{0, \delta\}$. The dynamic model of PRNN to meet conditions in (20) is given below:

$$\frac{d}{dt}\begin{pmatrix} \boldsymbol{\lambda} \\ \mathbf{y} \\ \boldsymbol{\lambda} \end{pmatrix} = \varepsilon \begin{pmatrix} -\nabla f(\boldsymbol{\lambda}) - \nabla \boldsymbol{g}^T(\boldsymbol{\lambda})\mathbf{y} + \mathbf{z}^T\boldsymbol{\lambda} \\ -\mathbf{y} + \mathrm{Pr}\left(\mathbf{y} + \mathbf{g}(\boldsymbol{\lambda})\right) \\ \mathbf{z}^T\boldsymbol{\lambda} \end{pmatrix} \tag{21}$$

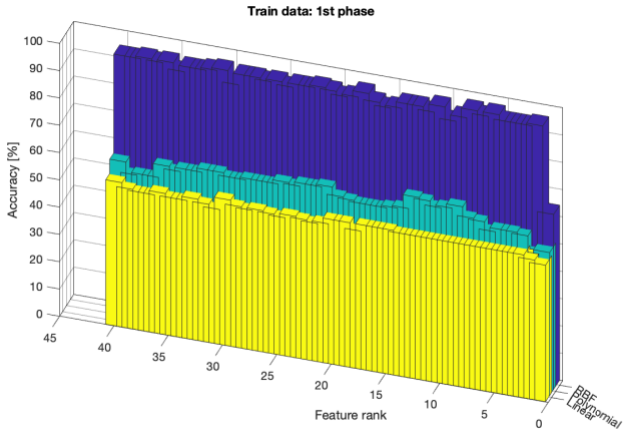
where $\varepsilon > 0$ is the convergence rate of PRNN. According to (20), the structure of the optimiser is simple, it is asymptotically stable and easy to implement. Using (21), the optimal $\boldsymbol{\lambda}^*$ is first obtained. The optimal weight vector $\mathbf{w}^*$ and bias h^* are then calculated using (13). Finally, the decision boundary is obtained as $f(\mathbf{x}) = \boldsymbol{\psi}(\mathbf{x})\mathbf{w}^* + h^*$.

VIII. FAULT DETECTION RESULTS

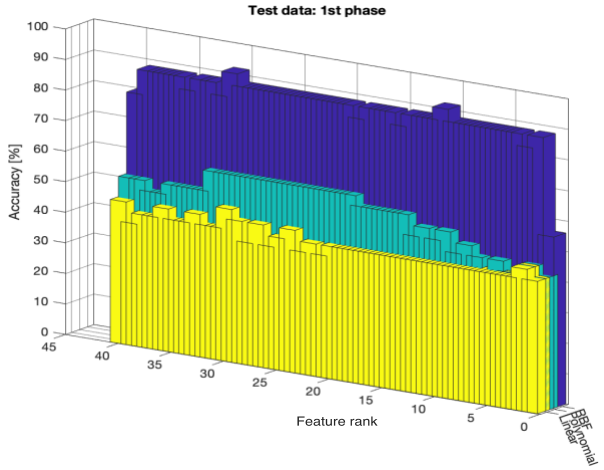
To evaluate the proposed fault detection algorithm, 70% of the dataset, comprising randomly mixed healthy and faulty features, is used for training the PRNN-based SVM. The remaining 30% of the data is reserved for testing the algorithm's accuracy in detecting bearing faults by comparing predicted and true labels. The accuracy of the algorithm is assessed using various feature ranks based on Section V and for three different kernels: linear, polynomial and Radial Basis Function (RBF) functions. Figure 7 illustrates the results of the first phase of the IM for both the training and test phases. It shows that the RBF kernel function achieves the highest accuracy. Table II indicates that the highest average accuracy in both training and testing is obtained using the RBF kernel function. This is mainly due to the high non-linearity of the feature space, where a non-linear kernel function is better suited to discriminate between healthy and faulty features. The algorithm achieves the best error detection accuracy in the training phase (approximately 97%) using the RBF kernel function with only the six most relevant features. For the testing phase, the algorithm with the RBF kernel function and 13 prominent features achieves the highest accuracy of 92%.

TABLE II. AVERAGE AND BEST ACCURACY OF THE ALGORITHM

Kernel function	Average accuracy (%) in the train phase	Best accuracy (%) in the test phase	Average accuracy (%) in the train phase	Best accuracy (%) in the test phase
RBF/	93.9368	97.0588	86.7000	92
Feature rank	40	6	40	11
Polynomial/	59.1029	64.1176	53.9167	60.23
Feature rank	40	13	40	19
Linear/	51.7353	53.5294	43.6667	50.12
Feature rank	40	30	40	30



a)



b)

Fig. 7. Accuracy of the proposed algorithm with different feature ranks and kernel functions during training and test phases. Charts from front to behind correspond to the linear, polynomial, and RBF kernel functions, respectively.

IX. CONCLUSION

In this paper, a machine learning algorithm consisting of SVM and a numerical optimizer called PRNN was developed to detect inner ring faults in the bearing of an induction motor. The initial time-frequency data set was obtained using three-phase currents and their DWT, PSD, and cepstrum, from which several statistical features were computed to form a feature set. A feature selection algorithm was used to identify the most relevant features, and the SVM optimization problem was formulated as a nonlinear programming problem and solved using PRNN. The results demonstrate that the best error detection accuracy can be achieved using an RBF kernel function and only six and thirteen prominent features for the training and testing phases, respectively.

REFERENCES

[1] P. Kazmierkowski, M. (2004). Electric motor drives: Modeling, analysis and control, R. Krishan, Prentice ‐ Hall, Upper Saddle River, NJ, 2001, xxviii+ 626 pp. ISBN 0 ‐ 13 ‐ 0910147.

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

[3] Rajapaksha, N., Jayasinghe, S., Enshaei, H., & Jayarathne, N. (2021, December). Acoustic analysis based condition monitoring of induction motors: A review. In 2021 IEEE Southern Power Electronics Conference (SPEC) (pp. 1-10). IEEE.

[4] Choudhary, A., Goyal, D., & Letha, S. S. (2020). Infrared thermography-based fault diagnosis of induction motor bearings using machine learning. IEEE Sensors Journal, 21(2), 1727-1734.

[5] Zarei, J., Tajeddini, M. A., & Karimi, H. R. (2014). Vibration analysis for bearing fault detection and classification using an intelligent filter. Mechatronics, 24(2), 151-157.

[6] Aviña-Corral, V., Rangel-Magdaleno, J., Morales-Perez, C., & Hernandez, J. (2021). Bearing fault detection in adjustable speed drive-powered induction machine by using motor current signature analysis and goodness-of-fit tests. *IEEE Transactions on Industrial Informatics*, *17*(12), 8265-8274.

[7] Yang, M., Chai, N., Liu, Z., Ren, B., & Xu, D. (2019). Motor speed signature analysis for local bearing fault detection with noise cancellation based on improved drive algorithm. IEEE Transactions on Industrial Electronics, 67(5), 4172-4182.

[8] Lau, E. C., & Ngan, H. W. (2010). Detection of motor bearing outer raceway defect by wavelet packet transformed motor current signature analysis. *IEEE Transactions on Instrumentation and measurement*, *59*(10), 2683-2690.

[9] Kompella, K. D., Mannam, V. G. R., & Rayapudi, S. R. (2016). DWT based bearing fault detection in induction motor using noise cancellation. *Journal of Electrical Systems and Information Technology*, *3*(3), 411-427.

[10] Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing*, *21*(6), 2560-2574.

[11] Brusamarello, B., da Silva, J. C. C., de Morais Sousa, K., & Guarneri, G. A. (2022). Bearing fault detection in three-phase induction motors using support vector machine and fiber Bragg grating. *IEEE Sensors Journal*, *23*(5), 4413-4421.

[12] Abid, F. B., Zgarni, S., & Braham, A. (2018). Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and aiNet-DAG SVM. *IEEE Transactions on Energy Conversion*, *33*(4), 1692-1699.

[13] Konar, P., & Chattopadhyay, P. (2011). Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Applied Soft Computing*, *11*(6), 4203-4211.

[14] Xia, Y., Leung, H., & Wang, J. (2002). A projection neural network and its application to constrained optimization problems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, *49*(4), 447-458.

[15] Xia, Y., & Wang, J. (2015). A bi-projection neural network for solving constrained quadratic optimization problems. *IEEE transactions on neural networks and learning systems*, *27*(2), 214-224.

[16] Xia, Y., Wang, J., & Guo, W. (2019). Two projection neural networks with reduced model complexity for nonlinear programming. *IEEE transactions on neural networks and learning systems*, *31*(6), 2020-2029.

[17] Stoica, P., & Moses, R. L. (2005). *Spectral analysis of signals* (Vol. 452, pp. 25-26). Upper Saddle River, NJ: Pearson Prentice Hall.

[18] Hwang, Y. R., Jen, K. K., & Shen, Y. T. (2009). Application of cepstrum and neural network to bearing fault detection. *Journal of mechanical science and technology*, *23*, 2730-2737.

[19] Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, *3*(02), 185-205.

[20] Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge university press.