Article

Mutual Information and Meta-Heuristic Classifiers Applied to Bearing Fault Diagnosis in Three-Phase Induction Motors

Gustavo Henrique Bazan [1,2], **Alessandro Goedtel** [2], **Marcelo Favoretto Castoldi** [2], **Wagner Fontes Godoy** [2], **Oscar Duque-Perez** [3,*] **and Daniel Morinigo-Sotelo** [3]

1   Department of Electromechanics, Federal Institute of Paraná, Av. Doutor Tito, s/n,
     Jacarezinho, PR 86400-000, Brazil; gustavo.bazan@ifpr.edu.br
2   Department of Electrical Engineering, Federal University of Technology-Paraná, Av. Alberto Carazzai, 1640,
     Cornélio Procópio, PR 86300-000, Brazil; agoedtel@utfpr.edu.br (A.G.); marcastoldi@utfpr.edu.br (M.F.C.);
     wagnergodoy@utfpr.edu.br (W.F.G.)
3   ADIRE-ITAP-Department of Electrical Engineering, University of Valladolid, Paseo del Cauce, 59, 47011
     Valladolid, Spain; daniel.morinigo@eii.uva.es
*   Correspondence: oscar.duque@eii.uva.es; Tel.: +34-983-184533

Abstract: Three-phase induction motors are extensively used in industrial processes due to their robustness, adaptability to different operating conditions, and low operation and maintenance costs. Induction motor fault diagnosis has received special attention from industry since it can reduce process losses and ensure the reliable operation of industrial systems. Therefore, this paper presents a study on the use of meta-heuristic tools in the diagnosis of bearing failures in induction motors. The extraction of the fault characteristics is performed based on mutual information measurements between the stator current signals in the time domain. Then, the Artificial Bee Colony algorithm is used to select the relevant mutual information values and optimize the pattern classifier input data. To evaluate the classification accuracy under various levels of failure severity, the performance of two different pattern classifiers was compared: The C4.5 decision tree and the multi-layer artificial perceptron neural networks. The experimental results confirm the effectiveness of the proposed approach.

Keywords: bearing failure diagnosis; mutual information; artificial bee colony; pattern recognition

1. Introduction

Induction machines are widely used in industrial applications due to their simple structure, their robustness in harsh environments, and their adaptability to different types of load, as well as their low acquisition and maintenance cost [1]. However, these machines are subject to various types of electrical and mechanical failures. Recent studies highlight bearing failures, representing between 40% and 52% of all failures in three-phase induction motors (TIMs) [1,2]. These failures can influence the operation of the machines, causing unexpected downtime. Generally, bearing failures are related to contamination, corrosion, inappropriate lubrication, and installation problems [2].

Due to the high possibility of failures, several authors are currently studying the condition monitoring of these machines [1,2]. Methodologies have been proposed to identify incipient defects related to these machines. Proper monitoring offers significant advantages to industrial processes, such as preventing unscheduled downtime, reducing downtime and repair costs, and providing more reliable machines. Thus, it can be seen that the development of a methodology capable of identifying incipient defects arouses great interest in the scientific community [1,2].

Fault diagnosis methodologies for induction motor are classified using the signal processing and pattern classification tools. Different physical measurements are used to analyze the motor, such as: vibration, temperature, acoustic emission, stator current, voltage,

power or magnetic signals, and oil or waste analysis [3]. However, most of these methodologies require invasive sensors that make their application difficult, requiring access to the machines' interior, which is difficult during operation. Thus, several approaches based on the stator current have recently received great attention, as it is a non-invasive method and is considered a reliable technology. The current sensor has advantages such as low-cost, simple access, ease of implementation, and it does not require additional sensors [3].

Signal processing tools have helped monitor induction motors, as they allow the extraction and selection of the most relevant characteristics of the signals. In the literature, conventional methodologies work in the frequency domain, which uses spectral decomposition of signals, such Fourier Transform (FT) [1–6], the Wavelet Transform (WT) [3,7,8], and the Hilbert Transform (HT) [9]. Some researchers are recently employing advanced tools to extract the relevant fault signatures for diagnosing failures in electrical machines.

In Ben Abid et al. [7], a methodology for the diagnosis of bearing defects based on Optimized Standing Wave Packet Transformation (Op-SWPT) and the artificial immune network (aiNet) is presented. The Op-SWPT allows extracting the characteristics of the defects from current signals even with a low acquisition rate and a reduced number of samples. In Saucedo-Dorantes et al. [10], the authors used linear discriminant analysis (LDA) and neural networks (NN) for monitoring multiple failures in TIMs; the main objective was to present a methodology for reducing the size of the data to increase the performance of the diagnostic system. Thus, allied to LDA, the researchers used other optimization tools such as the genetic algorithm, principal component analysis (PCA), and Fisher's scoring technique to select the most significant attributes resulting from the vibration decomposition signals. This set of tools allowed using a simple artificial neural network (ANN) to identify multiple faults.

Recently, the work of Li et al. [11] presents an alternative method based on envelope kurtosis for the selection of intrinsic mode functions (IMFs) generated by Value Mode Decomposition (VMD). The proposed methodology allows selecting the IMFs that contain the most relevant fault information. The results obtained showed that kurtosis outperforms other statistical measurements, such as the cross-correlation coefficient and entropy.

In Lu et al. [12], the researchers applied an adaptive filter based on nonlinear stochastic resonance (SRAF) to diagnose the bearings' defects of direct current motors. This tool filters out acoustic signals from the machine, reducing noise interference and improving the characteristic of defect-induced signatures. The proposed methodology also uses the zero-crossing detection algorithm to directly measure the signal period, identifying the bearing defect type.

In opposition to the methodologies that use the frequency domain, some studies propose alternative techniques, which use low-cost sensors and perform the analysis of the signal in the time domain, as in Nayana and Geethanjali [13], Nayana and Geethanjali [14], that employed time-domain features of the vibration signals, for diagnosing bearing failure in TIMs. The characteristic selection was then carried out using a filter approach, which was implemented using a Laplacian score (LS), and a wrapper approach using a brute-force method. Thus, several pattern classifiers, such as LDA, Naïve Bayes (NB), and Support Vector Machines (SVM), evaluated the selected features.

Information Theory is another methodology that has stood out in fault detection and monitoring area since it can determine the redundancy between the variables involved in the problem [15]. Entropy and Mutual Information (MI) are two of the computer tools used extensively in the diagnosis of failures in electrical machines. The work of Wang et al. [16] presented the Generalized Refined Composite Multiscale Sample Entropy (GRCMSE), which considers a methodology to determine the entropy of machine vibration signals to characterize the complexity of the signals at various scales and then diagnose bearing defects.

In Zheng et al. [17], an alternative non-linear dynamic approach, called Generalized Multi-Scale Composite Permutation Entropy (GCMPE), is presented to extract bearing failure characteristics from the vibration signals of rotating electrical machines. Moreover, Li et al. [18]

used the non-linear entropy characteristics, such as dispersion entropy (DisEn), permutation entropy (PerEn), and sample entropy (SampEn) of IMFs to determine the characteristics of different bearing defect severities.

MI has already been used in several approaches designed to solve pattern recognition problems, such as in the work of Kumar et al. [4] and Romero-Troncoso et al. [15]. Specifically, in the study of Kumar et al. [4], the authors devised the kernel-based mutual information matching (KEMI) function to determine the optimal parameters of Variational Mode Decomposition (VMD) from vibration signals to identify bearing defects in TIMs. The number of modes and the penalty factor greatly influence the decomposition process, so it is essential to adopt the best parameters. The experimental results, which exceeded 91% accuracy, demonstrated the promising capacity of the proposed methodology. MI can identify linear and non-linear dependencies between the time series [19]. Considering that the TIM is a complex non-linear system, in this study, unlike Kumar et al. [4], MI is used as a measurement of the similarity between the currents to extract meaningful information from the signals, which allows an efficient diagnosis of bearing failure in TIMs.

Like signal processing tools, Intelligent Systems (IS) have proven to be important in the diagnosis of failures in electrical machines due to their effectiveness in identifying the fault without the need for complex mathematical models, simplifying the computational implementation and allowing the detection of the origin of the failures, contributing to a more reliable operation of the machines. Among the IS tools, it can be mentioned Decision Trees (DT), Random Forests, ANNs, SVMs, *k*-Nearest Neighbour (*k*-NN), among others [5,7,10,13,16,20]. In the manuscripts of Chen et al. [21] and Chen and Jiang [22], a survey of several methodologies for the fault diagnosis in traction systems of high-speed trains were presented. It is possible to observe the wide applicability of IS tools, as well as signal processing techniques, such as FT, WT, PCA, among others, in the fault detection systems.

Recently, deep learning (DL) has attracted researchers' attention in several areas, such as image processing, computer vision, and pattern recognition, due to their better understanding of the intrinsic information of the analyzed data [5]. Besides, DL algorithms are being used in rotating machine failure diagnosis due to their ability to classify data without specific tools for extracting signal characteristics [5,6,23,24].

In Sohaib and Kim [5], Convolutional neural networks (CNN) are used to diagnose bearing defects in TIMs. These networks learn the information in the vibration signals' spectra, thus compacting the data needed for the classification process. On the other hand, Ding et al. [6] presented an alternative model called stacked autoencoders sparse filter rotating component comprehensive diagnosis (SAFC) to identify bearing defects in electric rotating machines. The proposed methodology has the ability to perform clustering and reduce the size of the spectra obtained through machine vibration signals.

In Liu et al. [24], a deep-belief convolutionary network (CBDN) approach to extracting and learning the relevant characteristics of bearing fault signals in electrical machines was presented. The main advantages of CBDN over other DL approaches are that they are quickly processed to perform calculations and extract features. Specifically, in this work, the researchers adopted Adam optimizer to accelerate the training model and the speed of convergence, making it possible to find an optimized model structure. In the experimental results, the classification accuracy values between 96% and 98% demonstrate the ability of this new approach in the diagnosis of bearing defects.

Some works from different areas have recently employed swarm intelligence tools and meta-heuristics techniques to achieve ideal solutions for complex optimization problems. Among them, it can be mentioned particle swarm optimization (PSO), whale optimization (WO), wheel-based differential evolution (WDE), grasshopper optimization (GO), biogeography-based optimization (BBO), beetle antennae search (BAS), and artificial bee colony (ABC) algorithms [8,14,16,17,23,25–28]. In the work of Haidong et al. [23], the researchers used the PSO algorithm to optimize the multiple parameters of the model proposed to diagnose bearing defects in several rotating machines. The PSO algorithm's

use allowed an increase in the efficiency of the bearing failure diagnosis in TIMs subject to several operating conditions, such as variation in the operating speed ranges. It was obtained classification rates higher than 90%, which validates the proposed system.

In Li et al. [25], a new approach called enhanced selective ensemble deep learning method with BAS algorithm was presented to diagnose train locomotive machine bearing faults. In this methodology, the BAS algorithm determines the optimal class-specific thresholds to perform the best classification. The experimental results showed that using this algorithm to the proposed method enabled a more accurate and robust recognition of the bearing fault patterns. The methodology correctly diagnosed more than 97% of all samples. It should be noted that there are no variations in the operating conditions of the machines. Recently, taking advantage of the WO algorithm's good optimization capacity, Miao et al. [26] were able to define the appropriate limits for signal decomposition to extract the bearing failure information needed for proper diagnosis.

In the work of Nayana and Geethanjali [14] previously described, the PSO and WDE algorithms allowed the identification of the most relevant features extracted through the feature selection tools used in the methodology. This approach presented classification rates of approximately 96% even in situations where the TIMs were subject to variations in the load, demonstrating the system's promising diagnostic capacity. Recently in the study of Wang et al. [16], it was employed the GO optimization algorithm to improve the learning and generalization capacity of the SVM classifier. This optimization tool was used to determine two important parameters of the SVM that must be previously defined before its use in pattern classification. The experimental results analyzed that the GO-SVM meta-heuristic classifier is capable of diagnosing bearing failures in rotating electrical machines.

Already in Zhang et al. [8], the outputs of the BBO and DE optimization techniques were used as the initial weights in the training process of the radial basis function (RBF) neural classifier. This DE/BBO algorithm was implemented to the RBF to overcome the drawback of this classifier's slow convergence rate. The experimental results demonstrated the superior performance of the optimized RBF classifier compared to the conventional RBF classifier in the diagnosis of bearing defects in electrical machines.

On the other hand, Huang and Ai [27] have already used the ABC algorithm to identify computer programs failures. The results showed that the proposed approach to fault diagnosis is promising, considering that the ABC methodology can efficiently avoid the optimal location and ensure the validity of fault finding for all software. Moreover, Huang et al. [28] used the ABC algorithm to investigate faults in photovoltaic systems. This tool allowed to reduce the number of samples for efficient diagnosis of defects in these systems.

In the present work, these optimization tools can collaborate to select the most relevant attributes of the failure patterns to reduce the input matrix of the pattern classifiers. Thus, the contribution of this work consists of developing an alternative methodology for the diagnosis of bearing failures in three-phase induction motors based on the optimization of the input matrix of the pattern classifier. First, this approach selects the most relevant attributes of mutual information measurements between the signal processing stage's current signals using the ABC algorithm. Finally, the pattern classifiers, the C4.5 decision tree (C4.5 DT) and the multi-layer artificial neural network (MLP ANN), better identify the bearing fault signatures extracted in the previous stage, are evaluated.

To validate the proposed approach presented in this work, experimental tests were conducted with TIMs running in steady-state, powered by frequency inverters. Several operating conditions found in the industry were emulated, such as the variation of the supply frequency from 20% to 100% concerning the nominal value and the load torque variation ($10\% \leq T_n \leq 120\%$, where T_n is the nominal torque). The analyzed severity levels are based on 15 to 90 min of wear of the bearings of two TIMs of 0.74 and 1.48 kW.

This paper is divided into five sections. Section 2 describes the theoretical background of the proposed methodology, which is fully explained in Section 3. Section 4 shows and discusses the experimental results, and Section 5 brings a conclusion and final remarks.

2. Theoretical Background

This section discusses the tools used in the proposed approach to fault diagnosis in inverter-fed TIMs. The general aspects related to the stages of characteristics extraction, as well as the classification of the failure signatures according to the methodology proposed in this work are presented.

2.1. Mutual Information

MI is defined as the measurement of the stochastic dependence between random variables, which allows the measurement of common information. Still, it reveals how the uncertainty about a random variable decreases when observing the other variable [29,30]. This measurement also describes the relationship between the time series collected simultaneously from the analyzed system [19]. In this context, MI is used in electrical machine signals to quantify the association between variables and monitor motor operating conditions. Equation (1) shows how this measurement can be determined using the probability density functions of the signals in the study.

$$I(X,Y) = \sum_{x \in \chi} \sum_{y \in \gamma} p(x,y) \log_2 \frac{p(x,y)}{p(x) \cdot p(y)} \tag{1}$$

where X and Y are the random variables; $p(x)$ e $p(y)$ are the probability density functions (PDF) marginal of X e Y, respectively; and $p(x,y)$ is the PDF between these two variables X and Y. When the mutual information is calculated with the logarithm in base 2, it is measured in bits. However, when using the Neperian base, its unit is nats.

In the proposed methodology, the signals analyzed are phases A and B of the stator current. Since these phases are out of phase with each other, it is necessary to use the Delayed Mutual Information tool (DMI). DMI is a measurement that has been used in the analysis of dynamic structures of complex systems, as it allows to estimate the similarity between random variables based on a time shift τ [19]. This measurement can be defined using Equation (2) and has been widely used in feature extraction to help solve pattern recognition problems [4,15].

$$I(X,Y^\tau) = \sum_{x_n \in \chi} \sum_{y_{n-\tau} \in \gamma} p(x_n, y_{n-\tau}) \log_2 \frac{p(x_n, y_{n-\tau})}{p(x_n) \cdot p(y_{n-\tau})} \tag{2}$$

Once the feature extraction stage is completed, the next step is to select the relevant mutual information values using the ABC algorithm, whose main characteristics are explained in the following section.

2.2. Artificial Bee Colony Algorithm

The ABC algorithm is based on the swarm meta-heuristic, and was introduced by Karaboga [31] to solve the optimization problems. This tool is based on the intelligent behavior of bees in nature [32]. In this algorithm, the artificial bee colony comprises three groups: worker bees, related to specific food sources; follower bees, which observe the workers inside the hive to choose a food source; and peasant bees, which search for food sources at random. Initially, honey bees discover all positions of food sources. Later, nectar from the food sources is pursued by worker and follower bees, and eventually, this continuous exploration will lead to food depletion. Eventually, the worker bees that chased the exhausted food source turn into the peasant bees to search for food sources again. The ABC algorithm mimics this behavior: The solutions of the problem are represented by the positions of the food sources and the quality of the solutions by the amount of nectar

in that source. Each worker bee is associated with one food source, so there are as many possible solutions as worker bees.

The next two sections are dedicated to explaining the classifiers used. The C4.5 Decision Tree and the Artificial Neural Network MLP are chosen since they have been widely applied in recent work on induction motor diagnosis [5,10,20] and have shown that they work very well for this purpose.

2.3. C4.5 Decision Tree

According to Fürnkranz et al. [33], the decision tree is a classification model whose structure consists of a certain number of nodes and arcs, representing the rules that will help in the classification of unknown samples. The classification of these samples is given by following the decision tree from top to bottom. One of the topologies is the C4.5 DT, which was idealized in Quinlan [34]. This algorithm works by recursively dividing the training data set into subsets using statistical data as the selection criterion. Thus, Entropy, Information Gain, and Information Gain Ratio are calculated for each sample's attributes present in the training subset. Subsequently, the algorithm tries to minimize the amount of information needed for the classification of a sample and ensure that a simplified tree is obtained.

2.4. MLP Artificial Neural Network

In this pattern classifier, the network structure's information flow follows a strict path, starting at the input layer and then passing through at least one hidden layer and ending at the output layer [35,36]. In the MLP ANN training process, the backpropagation algorithm is used with its forward and backward stages. In the first stage, the input vectors of the training set are propagated from the input to the output layer, taking into account the neurons' synaptic weights and thresholds. The obtained outputs are compared with the desired values, so this first stage is a supervised learning process. In the second stage, the synaptic weights and the neuron thresholds' values are adjusted according to the results of the previous stage. These two stages are performed recursively, leading to a progressive decrease of the sum of the differences between the network outputs and the desired values [35,36].

3. Proposed Approach for the Bearing Failure Diagnosis

In the industrial environment, TIMs are subject to different operational situations, such as load torque variations, unbalanced supply voltage, installation in a harsh environment, and electrical and/or mechanical defects, among others [1,3]. Proper monitoring to identify incipient failures related to these machines provides significant benefits to industrial processes, such as reduction of unscheduled downtime, reduced maintenance costs and more reliable machine operation.

Two TIMs were tested to simulate harsh operating conditions (Motor 1 of 0.74 kW and Motor 2 of 1.48 kW). The motors operated in steady-state, with and without bearing failures. These failures were created to emulate the process of bearing degradation through overuse, under lubrication and overload. First, the bearing was cleaned, and then an abrasive slurry was used in place of the lubrication to wear the parts. Then, the motor was run for the previously set times of 15, 30, 60, and 90 min to produce the bearing's degradation. This way, four classes of faults were obtained, one for each time of operation. Once the degradation of the bearings was produced, they were cleaned and lubricated to acquire the data properly.

The two TIMS were tested fed by two different inverters Siemens Sinamics G110 (G110), and Siemens MicroMaster 440 (MM440). To emulate situations common in an industrial environment, the coupled mechanical load and the supply frequency were varied. For the tests carried out with Motor 1, 808 samples were acquired, varying the load torque at 0.5 N.m intervals, from 0.5 N.m to 5.0 N.m. For Motor 2, 771 samples were collected with the load varying from 1.0 N.m to 9.0 N.m, in 1.0 N.m intervals. For both

TIMS, the supply frequency varied from 20% to 100% of the nominal value at 10% intervals. These conditions were considered for each level class of bearing failure.

Figure 1 shows the experimental workbench used to perform the experimental data. This workbench was designed to monitor electrical and mechanical measurements, such as voltage, current, vibration, torque and speed. In addition, it is possible to emulate many of the operating conditions mentioned above.
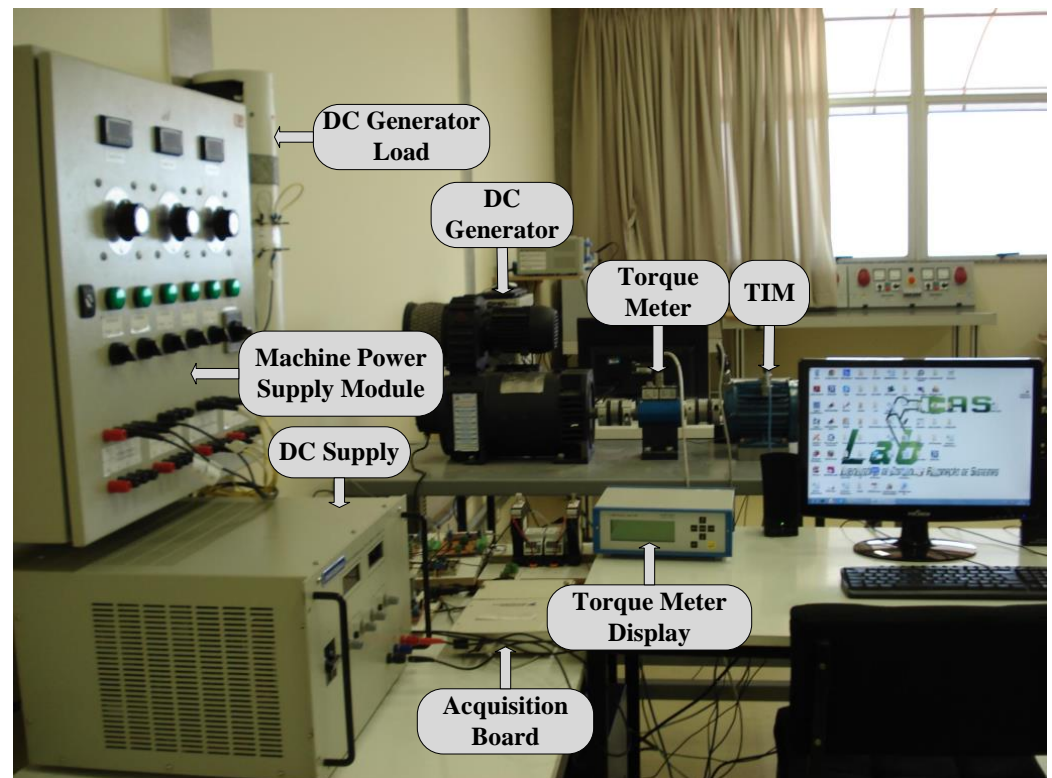


Figure 1. Workbench used in the experimental tests.

The workbench allows the variation of the supply frequency and a mechanical torque imposition on the shaft because the TIM is connected to a direct current generator. Moreover, there is a torque meter with an integrated speed sensor that can read torque and speed signals. Hall sensors are responsible for collecting and conditioning the stator current signals transferred to the data acquisition board's analog inputs connected to a microcomputer.

An algorithm developed in MATLAB software using the concepts described in Section 2.1 extracts relevant information from the collected data and reduces the number of inputs to the pattern classifiers. Figure 2 shows the procedures of the stages of acquisition, processing, and classification of signals used in this work.

Since the signals of the analyzed currents are out of phase, the delayed version of the mutual information described in the Section 2.1 is used to verify the similarity measurement. This measurement is calculated using the motor phase stator current signals, IA and IB, and the displacement value τ. The phase B stator current must be shifted sample by sample until the displacement value τ is reached. To calculate the DMI measurement, for each iteration of the algorithm, the marginal *pdfs* of the stator currents, and the joint *pdf* between these currents must be calculated. This procedure is performed recursively until the displacement value τ is reached. In this study, a displacement value of τ of 150 samples because using this value, the currents are in phase for at least one cycle. Figure 3 shows phases A and B of the current for Motor 1 without bearing failures, coupled load of 5 N.m, and supply frequency of 60 Hz; Figure 4 shows the calculated MI values.
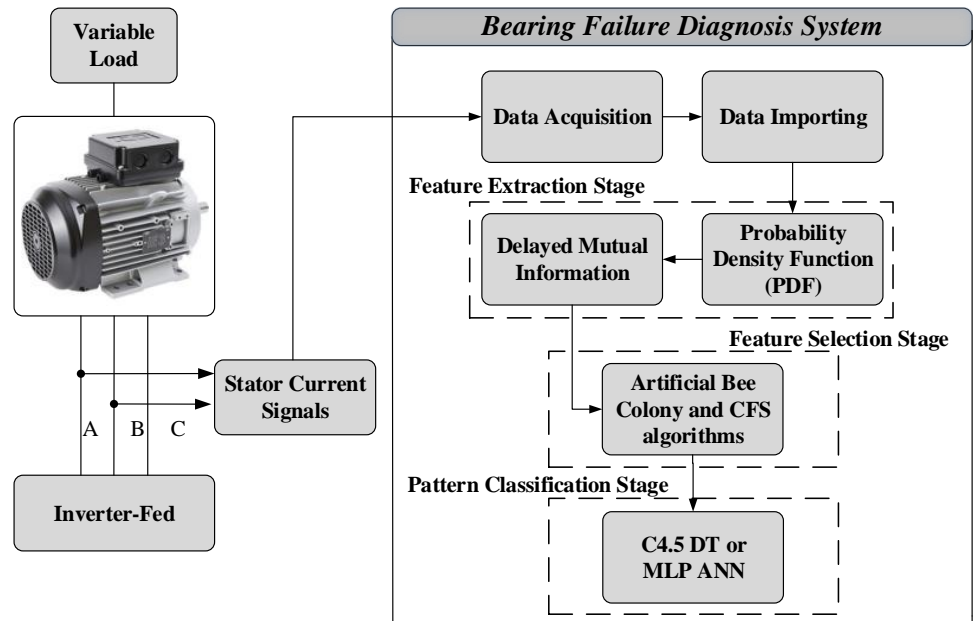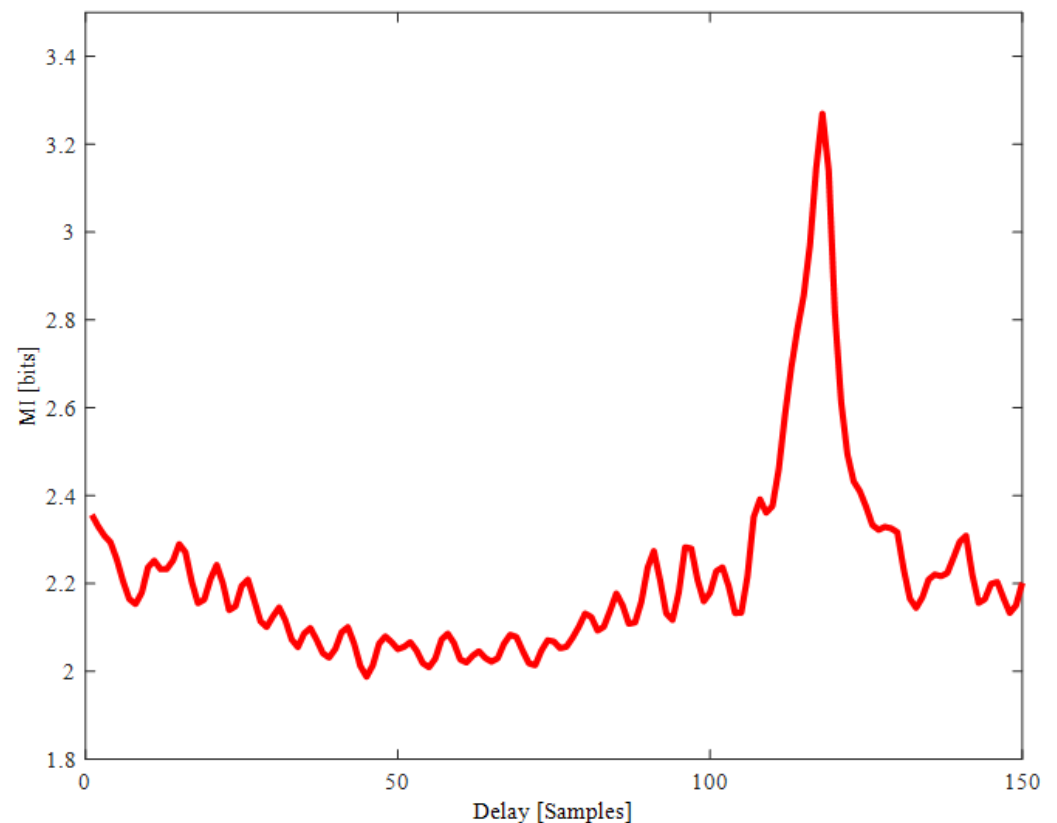
Figure 2. Block diagram of signal acquisition, data processing, and classification of the three-phase induction motor (TIM) bearing diagnosis system.
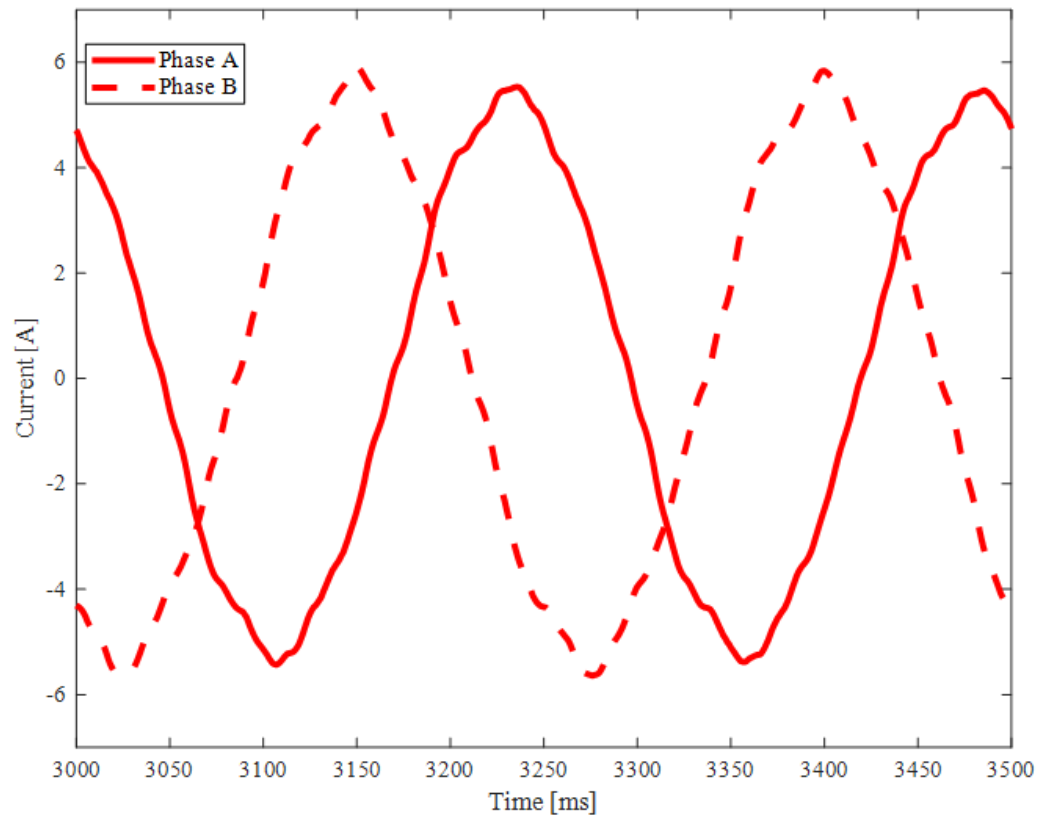
Figure 3. Stator current signals of Motor 1, without bearing faults, coupled load of 5 N.m. and supply frequency of 60 Hz.
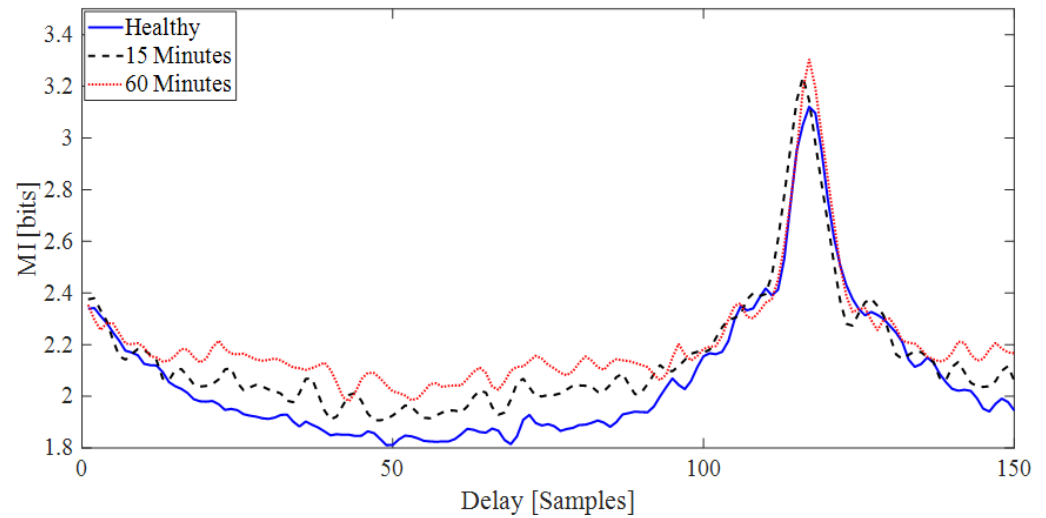
Figure 4. Delayed mutual information signature of Motor 1, without bearing faults, coupled load of 5 N.m. and supply frequency of 60 Hz.

The DMI signatures for all samples collected in the experimental tests were obtained using this feature extraction tool. These DMI curves are similar regardless of the machine's operating conditions. Figures 5 and 6 illustrate the behavior of the DMI values with the evolution of bearing defects. It is analyzed that the characteristic patterns of the DMI are similar regardless of the load in the axis of the machines.

Firstly, Figure 5 was obtained for signals for Motor 1 with a supply frequency of 60 Hz and with a coupled load torque of 0.5 N.m. In this figure, it is possible to notice that the evolution of wear on the machine's bearings causes an increase in mutual information's maximum value. In conclusion, as the fault progresses, the peak of the DMI increases. These characteristic patterns are also similar and can be analyzed as the machine's load torque changes. Figure 6 shows the DMI curves for Motor 1 with a supply frequency of 60 Hz and a load torque of 5.0 N.m. Again, there is a similar behavior observed in Figure 5. With the evolution of the defect, there is an increase in the maximum values of DMI signatures.

All the previous analyses were also performed for Motor 2, and the same behavior presented in Figures 5 and 6 was observed regardless of the machine's supply frequency. Once calculated, the DMI values were normalized using the maximum of the mutual information (MI_{max}) found in all TIM operating situations. Then, the most relevant attributes of mutual information were selected using the ABC algorithm to reduce the input matrix of the C4.5 and MLP ANN classifiers. The ABC algorithm was selected because it is widely employed in pattern classification problems, as Huang and Ai [27] and Huang et al. [28]. It was also observed that the ABC algorithm could optimize the data matrix more efficiently, considering that this matrix has been reduced with a smaller amount of data.

Figure 5. Delayed mutual information signatures change with the evolution of the bearing faults-Motor 1 with a supply frequency of 60 Hz and coupled load of 0.5 N.m.
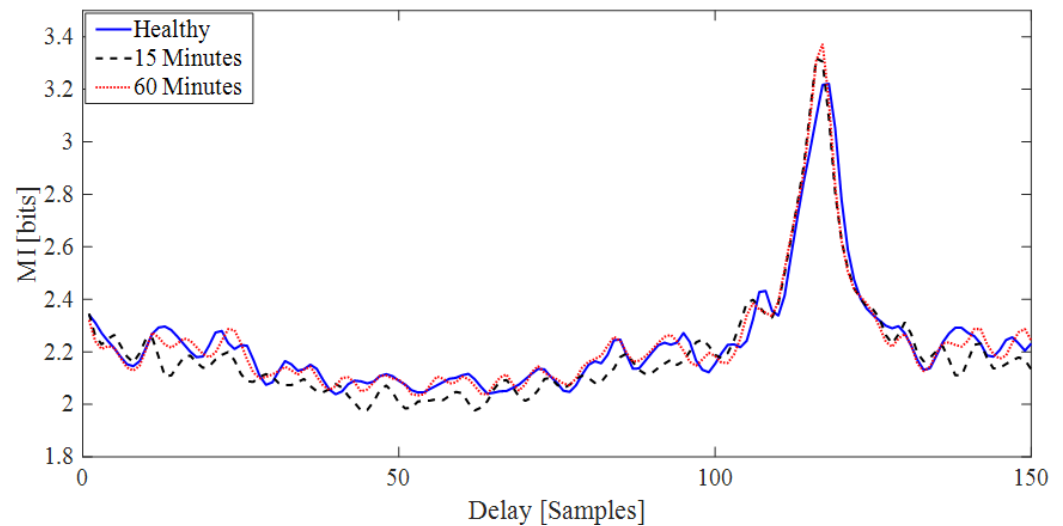
Figure 6. Delayed mutual information signatures change with the evolution of the bearing faults-Motor 1 with a supply frequency of 60 Hz and coupled load of 5 N.m.

In the step of attribute selection using the ABC algorithm, this tool employs the Correlation Feature Selector (CFS) algorithm, a technique that correlates the data set attributes based on a heuristic function. This algorithm also considers each attribute's individual predictive ability and the correlation degree between those attributes, including the class [37].

This methodology evaluates a subset of attributes, using the concept of relevance and redundancy. The relevance of an attribute is measured by the correlation of that attribute with the class, and the redundancy is measured by the correlation of that attribute with the other attributes of the subset. These concepts are intuitive, and they use the two main objectives of selecting attributes: choosing attributes with greater predictive power and decreasing the dimensionality by removing attributes with redundant information. Thus, the CFS method seeks to choose attributes with greater relevance and less redundancy.

The implementation is carried out using a local search strategy named forward, which only adds an attribute A to a subset S if there is no other attribute belonging to this set that has a greater correlation with class C than the attribute A. Therefore, the search for subsets of attributes is completed when the addition of attributes does not improve the

assessment value than the current best subset. Thus, both insignificant and redundant attributes should be discarded because they have a low correlation with the class studied and are probably highly related to one or more other attributes. Equation (3) defines the calculation for evaluating subsets using the CFS algorithm.

$$M_s = \frac{k\overline{r_{c_f}}}{\sqrt{k + k(k-1)\overline{r_{f_f}}}} \tag{3}$$

where M_s is the "merit" of a subset S of k features, r_{c_f} is the main correlation among class features, while r_{f_f} is the average among the all feature correlations.

Finally, by maximizing the function described in Equation (3), the input data from the TIM's bearing failure diagnosis system was obtained. These processed data have labeled the healthy motor signals received with the desired 0 output tag and the bearing failure signals with the desired 1 output tag according to their classes.

4. Experimental Results

A k-fold cross-validation technique was employed to obtain the experimental results to allow the training data set to be sufficiently representative. This technique works by randomly dividing the dataset into k subsets of the same size, and in each iteration, $k - 1$ subsets are used for training the classifiers, and the remaining subset is used for validation. For each test, the incorrect classification index of the validation data set is calculated. When all the k tests are completed, the classifier's performance is estimated according to the mean error. The following results were obtained with $k = 10$, according to studies of Kohavi [38] and Witten et al. [39].

The experimental tests were performed on the Waikato Environment for Knowledge Analysis (WEKA) software, which provides several intelligent algorithms, allowing the resolution of pattern classification problems [40]. The works of Kankar et al. [41], Pandya et al. [42], Konar and Chattopadhyay [43] and Konar et al. [44] employed this tool for pattern classification, clustering, and feature selection in problems related to motor failure diagnosis, which demonstrates the acceptance of this tool in the scientific community. In the selection of the most relevant attributes, it was used the ABC algorithm aforementioned. The following parameters were defined: an initial population of 30 bees, a maximum number of 20 iterations, bit-flip mutation type, and mutation probability of 0.01, which are default values for the computational tool WEKA.

The MLP ANN and the C4.5 DT described in Section 2 were used in the pattern classification stage. The MLP ANN was configured with a training learning rate of 0.3, a momentum of 0.2, and a maximum number of epochs for the training of 500, which is the standard configuration of the MLP ANN available in the computational tool WEKA. For the hidden and output layers, the logistics and linear activation function were used, respectively. The number of neurons was also defined as 10 neurons for cases where all 150 DMI attributes are used as the input matrix. On the other hand, an MLP ANN with 5 neurons was used when the ABC algorithm was applied to select the most relevant characteristics. In the construction of the C4.5 DT model, binary divisions in the nominal attributes were not considered. A seed was defined to randomize the samples, and a minimum number of attributes per leaf of 2 was considered. However, to build a more simplified and generalist decision tree, the pruning method was used, with a pruning confidence factor of 0.25 and an amount of data used to reduce pruning errors of 3. Moreover, it is worth mentioning that the results presented in this section were separated by frequency inverters used as supply source of the motors since each inverter has a specific switching characteristic, according to Martin-Diaz et al. [45].

4.1. Experimental Results without ABC Algorithm

This section presents the experimental results of the cases in which the ABC algorithm was not used to select the most relevant attributes of the DMI signatures. Therefore, all DMI

values (150 attributes) are used in the feature extraction stage as the C4.5 DT and MLP ANN classifiers' input data. As described above, several operating conditions found in industrial environments are emulated, such as variations of frequency, load and bearing failure levels. Table 1 shows the classification results for this first case performed on Motor 1.

Table 1. Experimental results using all the Delayed Mutual Information (DMI) values—0.74 kW motor.

Data Set	MM440 Inverter		G110 Inverter	
Classifier	C4.5	MLP	C4.5	MLP
Global Accuracy (%)	94.3	99.8	90.5	98.1
Healthy Accuracy (%)	85.1	100	64.4	93.2
Faulty Accuracy (%)	96.6	99.7	97.0	99.3
Kappa Index	0.82	0.99	0.67	0.94
Time to build the model (s)	0.19	10.7	0.08	9.56

Table 1 shows that the best results are obtained using the MLP classifier independently of the supply motor's frequency inverter. Particularly for the MM440 inverter, the classification rate is 99.8% of the overall data set, 100% of the healthy samples, and 99.7% of the faulty ones. The Kappa index of 0.99 supports the results described above. This shows that the use of mutual information combined with the MLP classifier can effectively separate classes. Similarly, it is possible to observe in Table 1 that the C4.5 classifier achieves results below the MLP classifier. However, 94.3% of all samples are correctly identified,which represents a satisfactory result. However, it is noteworthy that the C4.5 classifier takes 0.19 s to build the model, which is much shorter than the MLP network, which takes 10.7 s. Therefore, the C4.5 decision tree is also presented as a promising proposal for the diagnosis of bearing failures in the TIM fed by the MM440 inverter.

Analyzing the results presented for Motor 1 fed by G110 inverter, it can be seen that over 90.5% of the overall samples are properly identified. The RNA MLP to diagnose bearing failures achieves a classification accuracy of 98.1%, 93.2%, and 99.3% of the overall, healthy, and faulty samples, respectively. The Kappa index of 0.94 confirms the results obtained. As for the C4.5 classifier, it is possible to verify inferior results than the MLP ANN, where 90.5% of the samples are correctly classified. The system has greater difficulty in diagnosing healthy samples according to Table 1, only 64.4% of the samples are diagnosed. The Kappa index of 0.67 represents a significant agreement with the results obtained with this classifier.

After identifying the bearing failures in Motor 1, further experimental tests are performed with Motor 2. Table 2 shows the classification results considering the four classes of bearing failure and using the two frequency inverters.

Table 2. Experimental results using all the DMI values—1.48 kW motor.

Data Set	MM440 Inverter		G110 Inverter	
Classifier	C4.5	MLP	C4.5	MLP
Global Accuracy (%)	94.9	99.2	92.6	99.7
Healthy Accuracy (%)	86.1	100	73.7	100
Faulty Accuracy (%)	97.1	99.0	97.4	99.7
Kappa Index	0.84	0.98	0.76	0.99
Time to build the model (s)	0.06	9.75	0.06	9.33

According to Table 2, when the MM440 frequency inverter feeds the motor, the best separation between healthy and faulty classes is obtained using the MLP network. In this case, a classification accuracy of 99.2% of the overall samples is achieved, in addition to 100% and 99.0% of the healthy and faulty samples, respectively. When using the C4.5 classifier, lower classification rates are achieved, 94.9% of the overall samples. For this

situation, the Kappa coefficient is 0.84, which confirms the results obtained. It is noteworthy that the C4.5 DT presents promising results since the classifier model's construction time is lower than for the MLP network.

When the G110 inverter feeds Motor 2, the MLP classifier has the best classification accuracy. 99.7% of the global samples are properly diagnosed. Healthy and faulty classification rates above 99.7% are achieved, and the Kappa index is greater than 0.99, confirming the results. The C4.5 classifier diagnoses adequately 92.6% of the total samples, as shown in Table 2. Therefore, the MLP is the most suitable classifier.

In this section, the experimental results have shown the proposed methodology's capability based on information measurements and intelligent classifiers in diagnosing bearing faults in TIMs fed by frequency inverters and subject to various operating conditions. The following section evaluates the system's performance when the ABC algorithm is used to select the relevant DMI attributes and reduce the input data matrix of the pattern classifiers, decreasing the calculation time to build the methodology model.

4.2. Experimental Results Using the ABC Algorithm

This section presents the experimental results when meta-heuristic classifiers are used to identify failures. Similar to the tests made in Section 4.1, all data sets include variations in the power supply frequency from 20% to 100% of the nominal frequency and variations in the mechanical load torque of 10% to 120% of its nominal value.

As described in Section 3, the 150 attributes of the DMI signature are evaluated individually using the CFS algorithm present in ABC, which uses meta-heuristic merit to predict classes. The attributes with the greatest merit are identified as the most relevant for fault diagnosis. Therefore, they are used as input data for the construction of the classification model. For the Motor 1 data set fed by the MM440 frequency inverter, the ABC meta-heuristic selected 23 attributes: 29, 51, 55, 60, 66, 68, 82, 84, 86, 91, 92, 96, 106, 108, 110, 113, 114, 116, 119, 120, 121, 142 and 144, which constitute the matrix of 150 attributes of the DMI signature. In the case in which the G110 inverter was used to feed Motor 1, 17 attributes were selected using the ABC algorithm, among which are 49, 52, 54, 62, 74, 75, 89, 100, 109, 110, 114, 121, 127, 133, 138, 145, 147 of the 150 DMI attributes. Table 3 presents the experimental results for the new tests of Motor 1.

Table 3. Experimental results using the artificial bee colony (ABC) algorithm—0.74 kW motor.

Data Set	MM440 Inverter		G110 Inverter	
Classifier	C4.5	MLP	C4.5	MLP
Global Accuracy (%)	93.6	98.4	90.5	95.7
Healthy Accuracy (%)	81.6	95.4	56.2	89.0
Faulty Accuracy (%)	96.6	99.1	99.0	97.3
Kappa Index	0.80	0.95	0.65	0.86
Time to build the model (s)	0.13	1.44	0.06	1.31

By operating the ABC algorithm on the attribute selection for the Motor1 data set fed by the MM440 inverter, it is observed that there was a decrease of 0.70% and 1.40% in the accuracy for the C4.5 and MLP classifiers, respectively, compared to the tests presented in Section 4.1. However, the accuracy of the classification is still satisfactory, as can be seen in Table 3. Using the MLP pattern classifier, 98.4%, 95.4%, and 99.1% of the overall, healthy, and faulty samples were correctly identified, respectively. In this case, the Kappa index of 0.95 confirms the results obtained. For the C4.5 meta-heuristic classifier, 93.6% of all samples, 81.6% of the healthy samples, and 96.6% of the faulty ones were correctly detected. The Kappa coefficient of 0.80 confirms the agreement with the results, since values equal to or higher than 0.80 indicate this relation.

It is also possible to note in Table 3 that using the G110 inverter for feeding Motor 1. The accuracy rates are slightly lower compared to the tests using the 150 DMI attributes.

The proposed methodology correctly identifies 95.7% of the total samples using the MLP meta-heuristic classifier. Even so, 89.0% and 97.3% of the healthy and faulty samples are properly diagnosed. The Kappa coefficient of 0.86 confirms these promising results. Moreover, when the C4.5 meta-heuristic classifier was used, 90.5% of the global samples are correctly classified, according to Table 3. It is possible to note that this classifier had greater difficulty in identifying healthy samples. The Kappa index represents a substantial agreement with the obtained results.

Similar to the tests performed on Motor 1, the ABC meta-heuristic algorithm is used to select the relevant DMI attributes from the Motor 2 data set. These selected attributes are used as inputs to the MLP network and C4.5 classifiers. Using this technique, 20 DMI attributes were selected for the MM440 inverter data set: 1, 30, 41, 42, 56, 67, 70, 103, 109, 114, 115, 116, 117, 119, 121, 122, 125, 127, 148, 149, which constitute the matrix of 150 attributes of the DMI curve. For the G110 inverter data set, 13 attributes of the DMI were selected (46, 62, 74, 88, 89, 109, 110, 111, 117, 118, 121, 125, 145). Table 4 illustrates the classification results for the Motor 2 tests.

Table 4. Experimental results using the ABC algorithm—1.48 kW motor.

| Data Set	MM440 Inverter		G110 Inverter	
Classifier	C4.5	MLP	C4.5	MLP
Global Accuracy (%)	93.9	98.5	94.7	98.7
Healthy Accuracy (%)	82.3	96.2	82.9	94.7
Faulty Accuracy (%)	96.8	99.0	97.7	99.7
Kappa Index	0.81	0.95	0.83	0.96
Time to build the model (s)	0.04	1.22	0.05	0.92

Table 4 shows that the best results are obtained with the MLP meta-heuristic classifier regardless of the inverter. When the MM440 inverter supplied Motor 2, 98.5% of the samples were correctly identified. Using the G110 as a supply source of Motor 2, the proposed method achieved a classification accuracy of 98.7% of the global samples. Kappa indexes over 0.95 confirm the full results in both situations. Moreover, when the C4.5 meta-heuristic classifiers were used in the MM440 and G110 data set, slightly lower classification rates were obtained. Kappa coefficients over 0.81 confirm the satisfactory performance of the proposed system using the C4.5 meta-heuristic classifier.

From Tables 1–4, it was observed that the use of the ABC algorithm resulted in a reduction of a maximum of 2% in the classification rates of the global samples. Specifically using the MLP meta-heuristic classifier, it was achieved a global accuracy of over 95.7%.

Analyzing the healthy samples, it is possible to verify that there was greater difficulty identifying these samples, mainly employing the C4.5 meta-heuristic classifier. In some cases, there is a reduction of approximately 8% in the correct classification rates. Again, the MLP meta-heuristic classifier shows superior results.

Moreover, there is a small difference in the classification accuracy of the samples with bearing faults. Especially using the MLP meta-heuristic classifier, more than 97.3% of these samples were correctly classified. In general, it is noteworthy that using the ABC algorithm in selecting attributes, a similar accuracy was obtained with less computational time to build the model compared to previous tests. Thus, considering the computational scope, these tools are more suitable for the bearing failure diagnosis in TIMs, since satisfactory results are reached with a lower computational cost.

Figure 7 illustrates the performance comparison of the proposed method using the meta-heuristic classifiers for bearing fault diagnosis in TIMs fed by frequency inverters.
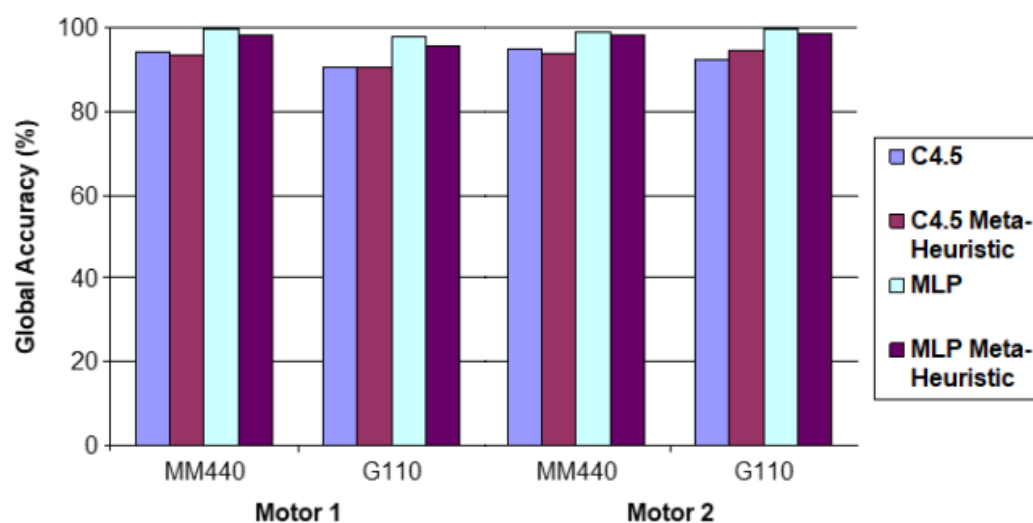
Figure 7. Performance comparison of the proposed method using meta-heuristic classifiers—Motors 1 and 2.

From the results presented in Sections 4.1 and 4.2, it was demonstrated that the proposed methodology presents a promising performance in the diagnosis of bearing failures in inverter-fed TIMs, regardless of the operating frequency and load.

5. Conclusions

This study has proposed a methodology for the diagnosis of bearing failures in TIMs fed by frequency inverters. The proposed approach is based on information measurements, artificial bee colony algorithm, and intelligent pattern classifiers. The experimental tests considered different operating conditions that the TIMs might face in industrial environments, such as variations in the load, supply frequency and four bearing failure configurations.

In the feature extraction stage, the Delayed Mutual Information was used to estimate the similarity measurements between two phases of the stator current to extract the bearing failure signatures. Subsequently, these signatures were evaluated by the pattern classifiers, C4.5 DT, and MLP ANN. The experimental results confirmed that the MLP ANN is a suitable classifier for the feature extraction approach, as it provides promising results for the sample classification, regardless of the operating conditions. However, it is important to highlight that using the C4.5 decision tree, the methodology also presents satisfactory results, achieving a classification rate of 94.3% of all data sets.

When the ABC algorithm was used to select the most relevant attributes of the DMI signatures, it was verified that the C4.5 and MLP meta-heuristic classifiers present promising results, higher than 93.6%. Using this ABC algorithm, the 23 most relevant attributes of the DMI signatures were selected, achieving results close to the tests without the ABC algorithm. This shows that, using the ABC algorithm for selecting the most significant attributes, satisfactory results can be achieved with a lower computational cost since less time is spent to build the classifier model.

Author Contributions: All of the authors equally contributed to the conception of the idea, the design of experiments, the analysis and interpretation of results, and the manuscript's writing. Writing—original draft preparation, G.H.B., A.G. and O.D.-P.; writing—review and editing, M.F.C., W.F.G. and D.M.-S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Frosini, L.; Harlisca, C.; Szabo, L. Induction Machine Bearing Fault Detection by Means of Statistical Processing of the Stray Flux Measurement. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1846–1854. [CrossRef]
2. Leite, V.C.M.N.; da Silva, J.G.B.; Veloso, G.F.C.; da Silva, L.E.B.; Lambert-Torres, G.; Borges da Silva, E.L.; de Oliveira, L.E. Detection of Localized Bearing Faults in Induction Machines by Spectral Kurtosis and Envelope Analysis of Stator Current. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1855–1865. [CrossRef]
3. Trachi, Y.; Elbouchikhi, E.; Choqueuse, V.; Benbouzid, M.E.H. Induction Machines Fault Detection Based on Subspace Spectral Estimation. *IEEE Trans. Ind. Electron.* **2016**, *63*, 5641–5651. [CrossRef]
4. Kumar, A.; Zhou, Y.; Xiang, J. Optimization of VMD using kernel-based mutual information for the extraction of weak features to detect bearing defects. *Measurement* **2021**, *168*, 108402. [CrossRef]
5. Sohaib, M.; Kim, J. Fault Diagnosis of Rotary Machine Bearings Under Inconsistent Working Conditions. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 3334–3347. [CrossRef]
6. Ding, R.; Li, S.; Lu, J.; Xu, K.; Wang, J. A Novel Stacked Auto Encoders Sparse Filter Rotating Component Comprehensive Diagnosis Network for Extracting Domain Invariant Features. *Appl. Sci.* **2020**, *10*, 6084. [CrossRef]
7. Ben Abid, F.; Zgarni, S.; Braham, A. Distinct Bearing Faults Detection in Induction Motor by a Hybrid Optimized SWPT and aiNet-DAG SVM. *IEEE Trans. Energy Convers.* **2018**, *33*, 1692–1699. [CrossRef]
8. Zhang, Q.; Gao, J.; Dong, H.; Mao, Y. WPD and DE/BBO-RBFNN for solution of rolling bearing fault diagnosis. *Neurocomputing* **2018**, *312*, 27–33. [CrossRef]
9. Batista, F.B.; Monteiro Lamim Filho, P.C.; Pederiva, R.; Augusto Diniz Silva, V. An Empirical Demodulation for Electrical Fault Detection in Induction Motors. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 559–569. [CrossRef]
10. Saucedo-Dorantes, J.J.; Delgado-Prieto, M.; Osornio-Rios, R.A.; de Jesus Romero-Troncoso, R. Multifault Diagnosis Method Applied to an Electric Machine Based on High-Dimensional Feature Reduction. *IEEE Trans. Ind. Appl.* **2017**, *53*, 3086–3097. [CrossRef]
11. Li, H.; Liu, T.; Wu, X.; Chen, Q. An optimized VMD method and its applications in bearing fault diagnosis. *Measurement* **2020**, *166*, 108185. [CrossRef]
12. Lu, S.; He, Q.; Yuan, T.; Kong, F. Online Fault Diagnosis of Motor Bearing via Stochastic-Resonance-Based Adaptive Filter in an Embedded System. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 1111–1122. [CrossRef]
13. Nayana, B.R.; Geethanjali, P. Analysis of Statistical Time-Domain Features Effectiveness in Identification of Bearing Faults From Vibration Signal. *IEEE Sens. J.* **2017**, *17*, 5618–5625. [CrossRef]
14. Nayana, B.R.; Geethanjali, P. Improved Identification of Various Conditions of Induction Motor Bearing Faults. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 1908–1919. [CrossRef]
15. Romero-Troncoso, R.J.; Saucedo-Gallaga, R.; Cabal-Yepez, E.; Garcia-Perez, A.; Osornio-Rios, R.A.; Alvarez-Salas, R.; Miranda-Vidales, H.; Huber, N. FPGA-Based Online Detection of Multiple Combined Faults in Induction Motors Through Information Entropy and Fuzzy Inference. *IEEE Trans. Ind. Electron.* **2011**, *58*, 5263–5270. [CrossRef]
16. Wang, Z.; Yao, L.; Cai, Y. Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine. *Measurement* **2020**, *156*, 107574. [CrossRef]
17. Zheng, J.; Pan, H.; Yang, S.; Cheng, J. Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis. *Mech. Syst. Signal Process.* **2018**, *99*, 229–243. [CrossRef]
18. Li, R.; Ran, C.; Zhang, B.; Han, L.; Feng, S. Rolling Bearings Fault Diagnosis Based on Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Nonlinear Entropy, and Ensemble SVM. *Appl. Sci.* **2020**, *10*, 5542. [CrossRef]
19. Melia, U.; Guaita, M.; Vallverdú, M.; Embid, C.; Vilaseca, I.; Salamero, M.; Santamaria, J. Mutual information measures applied to EEG signals for sleepiness characterization. *Med. Eng. Phys.* **2015**, *37*, 297–308. [CrossRef]
20. Palácios, R.H.C.; da Silva, I.N.; Goedtel, A.; Godoy, W.F. A novel multi-agent approach to identify faults in line connected three-phase induction motors. *Appl. Soft Comput.* **2016**, *45*, 1–10. [CrossRef]
21. Chen, H.; Jiang, B.; Ding, S.X.; Huang, B. Data-Driven Fault Diagnosis for Traction Systems in High-Speed Trains: A Survey, Challenges, and Perspectives. *IEEE Trans. Intell. Transp. Syst.* **2020**. [CrossRef]
22. Chen, H.; Jiang, B. A Review of Fault Detection and Diagnosis for the Traction System in High-Speed Trains. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 450–465. [CrossRef]
23. Haidong, S.; Ziyang, D.; Junsheng, C.; Hongkai, J. Intelligent fault diagnosis among different rotating machines using novel stacked transfer auto-encoder optimized by PSO. *ISA Trans.* **2020**, *105*, 308–319. [CrossRef] [PubMed]
24. Liu, S.; Xie, J.; Shen, C.; Shang, X.; Wang, D.; Zhu, Z. Bearing Fault Diagnosis Based on Improved Convolutional Deep Belief Network. *Appl. Sci.* **2020**, *10*, 6359. [CrossRef]
25. Li, X.; Jiang, H.; Niu, M.; Wang, R. An enhanced selective ensemble deep learning method for rolling bearing fault diagnosis with beetle antennae search algorithm. *Mech. Syst. Signal Process.* **2020**, *142*, 106752. [CrossRef]

26. Miao, Y.; Zhao, M.; Makis, V.; Lin, J. Optimal swarm decomposition with whale optimization algorithm for weak feature extraction from multicomponent modulation signal. *Mech. Syst. Signal Process.* **2019**, *122*, 673–691. [CrossRef]

27. Huang, L.; Ai, J. Automatic software fault localization based on artificial bee colony. *J. Syst. Eng. Electron.* **2015**, *26*, 1325–1332. [CrossRef]

28. Huang, J.; Wai, R.; Yang, G. Design of Hybrid Artificial Bee Colony Algorithm and Semi-Supervised Extreme Learning Machine for PV Fault Diagnoses by Considering Dust Impact. *IEEE Trans. Power Electron.* **2020**, *35*, 7086–7099. [CrossRef]

29. Cover, T.; Thomas, J. *Elements of Information Theory*, 2nd ed.; Wiley-Interscience: New York, NY, USA, 2006.

30. Principe, J.C. *Information Theoretic Learning-Renyi's Entropy and Kernel Perspectives*; Springer: New York, NY, USA, 2010.

31. Karaboga, D. *An Idea Based on Honey Bee Swarm for Numerical Optimization, Technical Report-TR06*; Technical Report; Erciyes University: Kayseri, Turkey, 2005.

32. Basturk, B.; Karaboga, D. An Artificial Bee Colony (ABC) Algorithm for Numeric function Optimization. In Proceedings of the IEEE Swarm Intelligence Symposium, Indianapolis, IN, USA, 12–14 May 2006; IEEE: Indianapolis, IN, USA, 2006.

33. Fürnkranz, J.; Gamberger, D.; Lavrac, N. *Foundations of Rule Learning*; Springer: Berlin/Heidelberg, Germany, 2012.

34. Quinlan, J. *C4.5: Programs for Machine Learning*; Morgan Kaufmann: San Francisco, CA, USA, 1993.

35. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson: Upper Saddle River, NJ, USA, 2008.

36. Silva, I.N.; Spatti, D.H.; Flauzino, R.A.; Liboni, L.H.B.; Alves, S.F.R. *Artificial Neural Networks: A Practical Course*, 2nd ed.; Springer: New York, NY, USA, 2017.

37. Hall, M.A. Correlation-based feature selection for discrete and numeric class machine learning. In Proceedings of the Seventeenth International Conference on Machine Learning (ICML'00), Stanford, CA, USA, 29 June–2 July 2000; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2000; pp. 359–366.

38. Kohavi, R.A. Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Montreal, QC, Canada, 20–25 August 1995.

39. Witten, I.H.; Frank, E.; Hal, M.A.; Pal, C.J. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: San Francisco, CA, USA, 2016.

40. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *ACM SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [CrossRef]

41. Kankar, P.; Sharma, S.C.; Harsha, S. Fault diagnosis of ball bearings using machine learning methods. *Expert Syst. Appl.* **2011**, *38*, 1876–1886. [CrossRef]

42. Pandya, D.; Upadhyay, S.; Harsha, S. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. *Expert Syst. Appl.* **2013**, *40*, 4137–4145. [CrossRef]

43. Konar, P.; Chattopadhyay, P. Multi-class fault diagnosis of induction motor using Hilbert and Wavelet Transform. *Appl. Soft Comput.* **2015**, *30*, 341–352. [CrossRef]

44. Konar, P.; Sil, J.; Chattopadhyay, P. Knowledge extraction using data mining for multi-class fault diagnosis of induction motor. *Neurocomputing* **2015**, *166*, 14–25. [CrossRef]

45. Martin-Diaz, I.; Morinigo-Sotelo, D.; Duque-Perez, O.; Arredondo-Delgado, P.; Camarena-Martinez, D.; Romero-Troncoso, R. Analysis of various inverters feeding induction motors with incipient rotor fault using high-resolution spectral analysis. *Electr. Power Syst. Res.* **2017**, *152*, 18–26. [CrossRef]