Applied Mathematics & Information Sciences
An International Journal

Bayesian Approach to Infer Types of Faults on Electrical Machines from Acoustic Signal

Glaucia Maria Bressan[1,*]*, Beatriz C. Flamia de Azevedo*[2]*, Herman Lucas dos Santos*[1]*, Wagner Endo*[1]*, Cristiano Marcos Agulhari*[1]*, Alessandro Goedtel*[1] *and Paulo Rogerio Scalassara*[1]

[1]Universidade Tecnológica Federal do Paraná (UTFPR), Campus Cornélio Procópio, PR, Brazil
[2]Instituto Politécnico de Bragança, Bragança, Portugal

Abstract: Considering the classification of failures in electrical machines, the present paper aims to use supervised machine learning techniques in order to classify faults in electrical machines, using attributes from audio signals. In order to analyze data and recognize patterns, the considered supervised learning methods are: Bayesian Network, together with the BayesRule algorithm, Support Vector Machine and k-Nearest Neighbor. The performances and the results provided from these algorithms are then compared. The main contributions of this paper are the acquisition process of audio signals and the elaboration of Bayesian networks topologies and classifiers structures using the acquired signals, since the algorithms provide the generalization of the classification model by revealing the network structure. Also, the utilization of audio signals as input attributes to the classifiers is infrequent in the literature. The results show that the Support Vector Machine and k-Nearest Neighbor present a high accuracy. On the other hand, the Bayesian approach is advantageous due to the possibility of showing, through graph representations, the generalized structure to represent the trend of faults in real cases on industry applications.

Keywords: Bayesian Networks Topologies, Supervised Learning Methods, Faults Classification, Audio Signals.

1 Introduction

Electrical machines are frequently used in various productive and industrial sectors, such as production planning, quality control and equipments monitoring. These machines need proper maintenance, since faults can affect production processes and cause losses for the industries where they are used. According to [1], fault diagnosis of electric machines is one of the main tasks that need to be performed to guarantee their right operation in most industrial facilities. Thus, preventing faults can avoid secondary effects, such as overheating, vibration, current and voltage unbalance, reduced efficiency and large financial losses [2,3]. The development of recognition and classification methods to predict motors faults has become an important area of research and interest in the industrial sector, because the occurrence of a fault can lead to severe damage to productive process, as a complete stop of production process and costly machinery repair [4].

Among the techniques available for fault detection, one can highlight the machine learning based methods [5, 6,7,8]. The literature involves recent and important works using machine learning techniques to predict faults in electrical machines and motors. More specifically, intelligent techniques such as Naive Bayes, k-Nearest Neighbor (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN) and Decision Trees algorithms have been reported in many applications. A review of Artificial Intelligence algorithms in rotating machinery fault diagnosis can be found in [9], presenting both the views of theory background and industrial applications.

Considering the recent literature, the author in [10] investigates defect diagnosis of induction motors, based on feature extraction from the envelope of the motor current. Three pattern classifiers – Naive Bayes, KNN, and SVM – are applied for defect classification. In [8], the authors propose an evaluation of pattern classification methods for fault identification in induction motors. The applied methods are: Naive Bayes, KNN, Support Vector

---

* Corresponding author e-mail: glauciabressan@utfpr.edu.br

Machine (Sequential Minimal Optimization), Multilayer Perceptron ANN, Repeated Incremental Pruning to Produce Error Reduction, and C4.5 Decision Tree. The paper [1] provides an experimental comparative evaluation of Naive Bayes, KNN and SVM techniques for rotor fault identifications. The inputs are obtained from current signals of an induction motor with two states of rotor bar degradation under two preestablished load levels.

The KNN method, in particular, has been applied in several types of problems. In [11], KNN is used in the decision phase, associated with an evolution tracking of the system using the trajectories, consequently allowing the diagnosis. A solution to detect broken bars under mechanical load is proposed based on voltage and current signal in the frequency domain. In [12], the KNN is applied to diagnose eccentricity faults in synchronous motors, analyzing the stator current signature in the frequency domain. An early fault diagnosis technique based on acoustic signals is described in [13], in which Nearest Neighbor and Nearest Mean classifiers are used for diagnosis of the single-phase induction motors.

A ANN-based method for the identification of winding failures in induction motors is presented by [5]. The fields with unbalanced currents and short-circuit conditions are identified by studying the harmonic orders of the radiated magnetic fields. Also, using ANN in [6], time and frequency domain parameters are extracted from the vibration and current signals and used to train the ANN and ANFIS models, which are then used to detect and diagnose the severity of the bearing fault. The author in [14] introduces a self-evolving maintenance scheduler framework for maintenance management of wind turbines. The authors proposes an ANN-based condition monitoring approach using data from supervisory control and data acquisition system. An alternative ANN-based method is proposed in [15] for classifying and detecting bearing faults in three-phase induction motors connected directly to the power grid.

In [16], SVM is applied along with continuous wavelet transforms [17] to analyze the frame vibrations during start-up. Motor vibration signals are used as SVM input, which performs the classification conditions. The work [18] employs supervised machine learning techniques (SVM, multilayer ANN and gaussian process regression) to correlate acoustic emission features with corresponding natural wear of slow speed bearings throughout series of laboratory experiments. Analysis of signal intensity estimator and root mean square is undertaken to discriminate individual types of early damage.

Using Bayesian Networks, [19] develops an acoustical damage detection method of the yaw system based on Bayesian network. The sound pressure level features are extracted from the measuring acoustic signal and a three-layer Bayesian Network diagnostic model combined with the structure learning strategy based on Bayesian information criterion which is designed for damage detection. The paper [20] uses Bayesian robust new hidden Markov modeling for bearing fault detection and diagnosis based on its acoustic emission signal. Then, the training set result obtained from is compared to the result from artificial neural network fault detection for same complex system of low speed and varying load conditions which are difficult from a diagnostic perspective.

Considering the important task of classifying failures, it is necessary to employ a signal that represents the motor condition. The present paper aims to use supervised machine learning techniques in order to classify faults in electrical machines, using frequencies and audio signals. The considered techniques in the present paper are: Bayesian Network, together with the BayesRule algorithm [21], SVM and KNN, since these methods consist of supervised learning methods to analyze data and recognize patterns. In addition, the performances and the results provided from these algorithms are compared. In supervised learning, there must be a relationship between the input attributes and the output classes, mapping the system behavior and providing a result that describes the reality [22]. For the motor faults evaluation, attributes can be extracted using microphones or sensors.

It is important to highlight that paper [23] presents supervised learning approaches to classify three-phase induction motors faults, applying Decision Trees and Random Forest algorithms. This is another approach to this problem. One of the main contributions of the present paper is the elaboration of Bayesian networks topologies and classifiers structures using audio signals. Analyzing the literature, the method most often used, based on the Bayes' theorem, is the Naive Bayes. The present paper proposes the application of Bayesian Network structures, considering the relationship between the input attributes. The BayesRule algorithm [21] is then applied in each network in order to extract linguistic rules, which provide a more comprehensible classification method. In this proposed methodology, the algorithms provide the generalization of the classification model for the decision-making process by revealing the network structure, allowing the inspection of the now available inner components. Another contribution is the acquisition process of audio signals, since the use of audio signals from motor noises as input attributes to the classifiers is infrequent in the literature. Usually, the works in the literature use current, voltage and vibration signals for fault classification [24]. Finally, another contribution is the performance evaluation of classifiers based on data mining techniques in order to predict faults in electrical machines.

This paper is organized as follows: Section 2 presents the Machine Learning techniques for electrical machines faults classification. The database composed by audio signals is presented in Section 3. The results are presented in Section 4. The analysis and the conclusion are introduced in Section 5.

2 Machine Learning Techniques for Classification Task

The supervised machine learning techniques considered in this paper are presented in this section. The considered techniques, applied to classify faults in electrical machines, are the Bayesian Networks with the BayesRule algorithm, SVM and KNN. These methods are characterized by their use of training data and validation to create a classification model. After that, the model can be used to classify a new instance of related data. A brief description of the algorithms is presented in the following subsections.

2.1 Bayesian Networks

The Bayesian inference model consists in analyzing conditioned variables in a probabilistic way, by using *Bayes' theorem*, as described in [25]: Suppose that $y' = (y_1,...y_n)$ is a vector of n observations in which the probability distribution $p(y \mid \theta)$ depends on the k parameters values $\theta' = (\theta_1,...\theta_k)$. Consider that θ has probability distribution $p(\theta)$. Then, given an observation y, the conditional probability of θ is represented by equation (1), known as *Bayes' theorem*.

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)}. \qquad (1)$$

In (1), $p(\theta)$, so called *a priori* distribution, provides the knowledge about θ without the knowledge about data and $p(\theta \mid y)$, called *a posteriori* distribution of θ given y, provides the knowledge about θ given the knowledge about data. Considering the information about y is known, $p(y \mid \theta)$ can be considered as a function of θ, known as the *likelihood* of θ given y. The Bayes theorem then states that the *a posteriori* distribution of θ given y is proportional to the product of *a priori* distribution of θ given the knowledge of the data by the likelihood of θ given y. Therefore, *posteriori* $\propto$ *likelihood* $\times$ *priori*.

In most real systems, the variable of interest is dependent or is subject to several others. Although it can be extended to more than one variable, the Bayes' theorem is not recommended in cases involving multiple variables due to the complexity of the calculations [26]. Faced with this difficulty, Pearl developed in [27] networks that evaluate the interconnections of variables through their structures, known as *Bayesian Networks*.

Bayesian networks can be defined as a probabilistic graphical model used to represent knowledge about the data domain. These networks learn cause and consequence relationships and can combine *a priori* knowledge with patterns learned from the data. In addition, the user can interfere in the nodes of the network and insert a knowledge that propagates in the other nodes. The networks are then composed of a structure consisting

of a directed acyclic graph and a set of probability tables. The nodes of the network structure represent the variables and the arcs between nodes represent dependency relations between the corresponding variables. An arc starting at a A node and ending at a B node establishes A as the parent of B and B as the child of A. A Bayesian network can be used as a classifier by calculating the conditional probability of a node, called a class node, given the values of the probabilities of the other nodes.

In [28], a Bayesian Network is represented by $BN = \langle N,A,\Theta \rangle$, in which $\langle N,A \rangle$ is a directed acyclic graph where each node $x_i \in N$ represents a variable and each arc $a \in A$ between nodes represents a probabilistic dependency between the associated nodes. A conditional probability distribution θ_i is associated with each node $x_i \in N, i = 1,...,n$, collectively represented by $\Theta = \theta_i$, which quantifies how a node depends on its parents [28]. As described in [27], the conditional independence, which is the Markov condition, allows the calculation of the joint distribution of all variables, given by

$$P(x_1,x_2,\ldots,x_n \mid BK) = \prod_{i=1}^{n} P(x_i \mid \pi_{xi}, BK), \qquad (2)$$

where BK represents the antecedent knowledge, x_i is the ith variable or node, and π_{xi} is the set of x_i parents. Therefore, a Bayesian network can be used as a representation of the knowledge that allows inferences. The nodes that are part of the Markov Blanket of the M node are those that influence the conditional probability distribution of M, which are the parent nodes of M, the child nodes of M and the parents of the children of M [27].

Bayesian networks consist of an efficient way of working with complex information; however, the knowledge represented is not as comprehensible as other forms of representation that express knowledge similar to human reasoning as rules of classification. Hence, the method called *BayesRule* was developed by Hruschka Jr *et al.* in [21]. It uses the concept of maximum *a posteriori* probability to extract a set of "if-then" probabilistic rules, which describe the classification.

The *BayesRule* method uses the algorithm K2, which employs a heuristic search to induce a Bayesian network using as input a database and an ordered list of variables. The algorithm uses a greedy method to search for the best network structure that represents the data [21]. In addition to using maximum *a posteriori* probability, the *BayesRule* method employs the Markov Blanket concept to reduce the number and complexity of classification rules generated by the extraction process. This allows the reduction in the time required to construct models when more than one variable is designated as a class. In a Bayesian classifier, the propagation algorithm must be used to propagate the values of the variables and infer the value of the class variable. Rule confidence can be defined using inferential results. Thus, the probability given to the inferred class can be used as a confidence

value and is embedded in the inference algorithm. The *BayesRule* algorithm can be seen in [21].

2.2 Support Vector Machine (SVM)

SVMs can be used for numeric prediction as well as classification. The SVM algorithm uses a nonlinear mapping to project the original training data into a higher dimension. Within this new dimension, a linear optimal hyperplane capable of separating the data into two classes is computed. According to [29], it is always possible to obtain such hyperplane with an appropriate nonlinear mapping to a sufficiently high dimension, and the hyperplane is obtained using the so-called support vector and margins (defined by the support vectors).

SVM relies on the backpropagation training algorithm that uses optimization of a convex quadratic function for maximum practical performance. Using the input values, the algorithm then elaborates a nonlinear boundary hyperplane that divides the feature space in two distinct regions. As an optimization problem, the SVM equation can be expressed as in (3), which is a Lagrange dual optimization problem.

$$\max W(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \lambda_i \lambda_j (x_i x_j) \qquad (3)$$

$$\text{subject to: } 0 \leq \lambda_i \leq C$$

$$\sum_{i=1}^{N} \lambda_i y_i = 0; \quad i = 1, 2, \ldots, N$$

The variable C in (3) is the threshold of error, N is the number o samples, λ_i is Lagrange multipliers which must follow the direction of $\partial W / \partial \lambda$, y is the desired output and x are the input samples. Maximizing W provides the separation border between the classes. According to [29], although the training time of even the fastest SVMs can be extremely slow, they are highly accurate due to their ability to model complex nonlinear decision boundaries.

2.3 K-Nearest Neighbors (KNN)

K-Nearest Neighbors is a supervised machine learning algorithm that can be used to solve both classification and regression problems. The KNN classification approach consists in fixing the number k of samples and let the width change, such that each region contains exactly k samples [22]. The KNN process starts at the test point and expands a region until it encloses k training samples, labeling the test point **x** by a majority vote of these samples. If the majority of samples closest to the unknown sample are from a specified class, the sample will be assigned to that class.

For two classes, the value of k should be odd to avoid a tie, and larger values are more likely to resolve ties. In fact, the larger the value of k, the smoother will be the classification boundary, and smaller values for k results on a more convoluted boundary [22]. It is important to highlight that the region will be circular if the data are normalized.

The similarity is defined by the Euclidean distance between data, given by

$$d(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \qquad (4)$$

where X and Y are the set of n samples.

There is essentially no training involved in the KNN method, so it is considered a lazy learning algorithm. In general the KNN defers data processing until it receives a request to classify an unlabeled (test) example, and then discards any intermediate results. The main KNN advantages are that it is intuitive, analytically tractable and simple to be implemented [22].

3 Experimental Details

In order to classify electrical motor faults, it is necessary to extract attributes that represent acoustic emission signals and frequencies. Therefore, time domain and frequencies characteristics, that are very reliable for fault detection, are measured.

The data set used in this paper is obtained from a three-phase induction motor assembled at the Intelligent Systems Laboratory (LSI), together with Signal Processing and Applications Laboratory (LPSA), located in the Federal University of Technology of Paraná, in Cornélio Procópio city, Paraná, Brazil. The data are composed by acoustic emissions recorded using two Behringer ECM8000 condenser type omni-directional microphones, being the acoustic emission sensors running to a *Focusrite Scarlett 2i2* audio interface, which is responsible for the data acquisition. The microphones were positioned close to the electric motor, with 9.53 cm between them, as shown in Figure 1, in order to avoid any spatial ambiguity. The best distance was obtained experimentally.

The input attributes for the supervised machine learning techniques consist of the peak magnitudes of the signal frequency spectrum at predetermined frequencies and the total signal power. The *frequency peaks* were extracted using the *Fast Fourier Transformer* algorithm and *Hanning* windowing, considering the frequencies of 30, 60, 120 and 2500 Hz.

Electrical unbalance generates vibration in the power supply frequency first harmonic [30]. Broken rotor bars fault causes vibration around the harmonics of the rotor speed related to supply frequency [31]. Bearing faults are commonly detected by modulations in high frequency

Fig. 1: Equipment used for obtaining the data set.
Font: authors

Table 1: Output data set.

Mechanical Conditions	Electrical Conditions	Outputs
Healthy	balanced	output 1
	unbalanced	output 2
Bearing outer race fault	balanced	output 3
	unbalanced	output 4
Broken rotor bar	balanced	output 5
	unbalanced	output 6

cleaned, properly lubricated with grease (in order to simulate normal operation) and inserted into the motor. Finally, the last class consists of signal samples provided by motor with two broken rotor bars. The bars were damaged using a drilling machine in two adjacent bars to emulate the fault.

4 Results and Discussion

In this section, we present the results from the topologies of Bayesian networks (together with the BayesRule algorithm) and from the structures of classifiers proposed in this work, based on SVM and KNN. The hardware used in computational experiments is composed by an Intel Core I7 processor, RAM memory 16GB executing operational system Microsoft Windows 10.

The classifiers consist of supervised machine learning techniques to classify faults in electrical machines using frequency components from acquired audio signals. In order to build the structures of the proposed classifiers, tests were performed considering the relationship of the 5 input attributes: freq30, freq60, freq120, freq2500 (the peak magnitudes of the signal frequency spectrum at the predetermined frequencies of 30, 60, 120 and 2500 Hz) and the total signal power.

Data set is composed of 570 instances and 6 columns, of which the first 5 columns are the input attributes and the last one is the output of the classification models, since the proposed classifiers are supervised, as described in Table 1. Data set is randomly divided in training set, composed of 80% of the data, and the test set, composed of the remaining 20%. Then, the results provided from the test set are analyzed. If the model correctly classified the sample of the test set (comparing it with the known output), then one success is counted; otherwise, one error is considered. This method is known in the literature as *holdout* [34]. With this procedure, the performance of classifiers are evaluated based on data mining techniques in order to predict faults in electrical machines. After the tests performed, we propose 6 topologies of Bayesian Networks, 6 structures of SVM and 6 structures of KNN based classifiers, described in Table 2, which also presents the best relationship of the input attributes.

[32]. Finally, winding faults vibrations are related to the power supply frequency harmonics and slip [33]. Therefore, five numerical input attributes are considered: peak values at 30 Hz, 60 Hz, 120 Hz, 2500 Hz and the *signal power*. The signal power is estimated using the autocorrelation of the signals in point 0, obtained using the cross-correlation function of the signal with itself, given by

$$\hat{R}_{x_1 x_2}(m) = \sum_{n=0}^{N-m-1} x_1(n+m)x_2(n), \quad m = 0, 1, \dots, N-1,$$
(5)

being $x_1(n)$ and $x_2(n)$ signals with length N. The signal power features are calculated as $\hat{R}_{x_1 x_1}(m)$ and $\hat{R}_{x_2 x_2}(m)$.

Since supervised methods are considered, the output data set is divided into three classes, called Mechanical Conditions: healthy motor, motor with bearing outer race fault and motor with two broken rotor bars. Each one of these three Mechanical Conditions are subdivided into two other subclasses, called Electrical Conditions, which correspond to the balancing or unbalancing of the supply voltage phases, as shown in Table 1. Each one of the six rows of Table 1 corresponds to an output of the classification system, as shown in the last column of the table. The unbalance generates an excessive current flow in one or more phases, so the motor has the voltage unbalance, causing abnormal behavior in the rotor and irregular vibration.

The healthy motor data were obtained from a three-phase, 1HP and 4 pole electric motor with 220/380 V – 60 Hz power supply, considering nominal speed of approximately 1700 RPM.

In order to obtain the bearing outer race faults data, a corrosive slurry was placed in the outer race of the bearing and, after the action of the pulp, the bearing was

Table 2: Proposed structures for the classifiers.

Proposed Structure	Input 1	Input 2	Input 3
Structure 1	freq30	freq2500	signal power
Structure 2	freq30	freq120	signal power
Structure 3	freq30	freq60	signal power
Structure 4	freq30	freq60	freq2500
Structure 5	freq30	freq60	freq120
Structure 6	freq30	freq120	freq2500

Table 3: Accuracies of the SVMs.

Classifier	Input 1	Input 2	Input 3	Accuracy
SVM 1	freq30	freq2500	signal power	100%
SVM 2	freq30	freq120	signal power	100%
SVM 3	freq30	freq60	signal power	100%
SVM 4	freq30	freq60	freq2500	85.83%
SVM 5	freq30	freq60	freq120	95%
SVM 6	freq30	freq120	freq2500	95.33%

4.1 Inferring a priori knowledge

To validate the faults in electrical machines, we applied support vector machine and k-Nearest Neighbor. Both of techniques have a high accuracy to compare the results. In advantage, Bayesian inference shows through graph representation the generalized structure to represent the trend of faults in real cases on industry applications.

4.1.1 Support Vector Machine

The data classification was also implemented in SVM classifier. In order to compare the performances of techniques, the same training and test sets were employed. The numerical values of the signal features were normalized in their own domain and employed as inputs of the classifier. Since the features in the dataset were chosen in each motor condition classification, as described in Table 1, specialist SVMs were trained for each one. The training and operation phase were implemented in MATLAB.

Six specialist SVMs were implemented, with outputs described, as follows: SVM 1 classifies if the signal represents a healthy motor with balanced power supply. SVM 2 corresponds to the classification for healthy motors with unbalanced power supply. SVM 3 is dedicated to bearing fault and balanced power supply. SVM 4 presents the classification for bearing fault motor and unbalanced power supply. The last two structures are correspondent to motors containing broken rotor bars with normal power supply, SVM 5, and unbalanced power supply, SVM 6.

The SVMs were trained using radial basis as kernel function. Also, the hyperparameters kernel scale and box constraint were optimized to obtain the higher accuracy. The results are shown in Table 3.

4.1.2 k-Nearest Neighbor

The KNN classifier was implemented following the same rules as SVM. The structure was applied by designing a specialist classifier for each motor condition. Altogether, six KNN were employed in the classification. The number of neighbors were tested from one to twenty and the best results are shown. The best result was obtained using 7

neighbors, which provide the average accuracy of 91.94% using the classifiers KNN1, ..., KNN6, as discriminated in Table 4.

Table 4: Accuracies of the kNN

Classifier	Input 1	Input 2	Input 3	Accuracy
kNN 1	freq30	freq2500	signal power	90%
kNN 2	freq30	freq120	signal power	99.17%
kNN 3	freq30	freq60	signal power	88.33%
kNN 4	freq30	freq60	freq2500	93.33%
kNN 5	freq30	freq60	freq120	93.33%
kNN 6	freq30	freq120	freq2500	87.5%

Although KNN and SVM both yield better overall accuracies, there is no concrete information about the knowledge of which features contribute the most to the pattern recognition of the faults. These information have a major contribution in early fault detection, which is scope for future works regarding new strategies for efficient diagnosis systems.

4.2 Bayesian Networks infers the generalized structure

Since the classification of electrical machine faults involves different input attributes, we propose a probabilistic structure approach, based on Bayesian Networks, considering the relationship between input attributes. If new information is known, it can easily be incorporated to the model, contributing to improve the reliability of the results.

Due to the performed tests considering the combination of the input attributes, the obtained networks present significant relationships, express the best results on the parameter of inference and allow the use of *BayesRule* algorithm to extract linguistic rules.

Bayesian Networks and the *BayesRule* algorithm use linguistic input attributes to extract linguistic rules. For this reason, the numerical input attributes must be subject to a process of discretization. In other words, they must be transformed into linguistic variables, partitioned into three classes defined in this work as: low, medium and high, as

shown in Table 5. The ranges for each class are obtained using the software GENIE[1].

Also, using GENIE and the algorithms *K2* and *Greedy Thick Thinning* [35], the 6 structures of the Bayesian Networks were generated, according to Figure 2, where the *k*-th Network corresponds to the *k*-th Structure of Table 2. After that, the *BayesRule* algorithm is executed in order to extract the linguistic "if-then" rules. From the combination of the input attributes and their classes, 27 linguistic rules are generated from each of the 6 Bayesian Network structures. Since Bayesian Networks present a probabilistic structure, all the generated linguistic rule presents a probability of occurrence. Following, one rule of each network is given as example:

Network 1: *IF (freq30 is high) AND (freq2500 is low) AND (signal power is medium) THEN output 2 (Healthy and unbalanced)* with 75.862% of probability.

Network 2: *IF (freq30 is medium) AND (freq120 is high) AND (signal power is medium) THEN output 2 (Healthy and unbalanced)* with 91.53% of probability.

Network 3: *IF (freq30 is low) AND (freq60 is low) AND (signal power is high) THEN output 4 (Bearing and unbalanced)* with 83.333% of probability.

Network 4: *IF (freq30 is high) AND (freq60 is high) AND (freq2500 is low) THEN output 2 (Healthy and unbalanced)* with 75% of probability.

Network 5: *IF (freq30 is medium) AND (freq60 is low) AND (freq120 is low) THEN output 1 (Healthy and balanced)* with 74% of probability.

Network 6: *IF (freq30 is medium) AND (freq2500 is low) AND (freq120 is high) THEN output 2 (Healthy and unbalanced)* with 91.53% of probability.

In order to validate the classifier, the instances of the test set were discretized using the same ranges of training set, shown in Table 5. Each test instance is applied to the corresponding network and the instances correctly classified, according to the "if-then" rules generated by *BayesRule* algorithm, are counted up. Table 6 presents the accuracy of the classification process considering the application of the test set to each network structure.

It is important to highlight that although the accuracies presented in Table 6 are lower when compared to the results obtained by the application of SVM and KNN, the Bayesian Networks show the interconnection of the attributes and their importance for each classification. These interconnections, which are hidden when considering the SVM and KNN methods, may bring interesting qualitative analysis to the fault classification problem. For instance, by analyzing the presented linguistic rules, it can be seen that the freq30 component

by itself is insignificant for classification, but if it has a medium value and the component freq120 is high, there is a great probability that the machine is healthy and unbalanced, according to Networks 2 and 6. Thus, it is important not only to inspect the inferences between the attributes, but also to combine the obtained networks into a single generalized structure.

Such a general structure is presented in Figure 3. The main contribution from this model is to determine the structured learning process using all graphs obtained from the Bayesian Network approach. Since the general graph is oriented, not only the relations between parameters, but also the statistical dependencies are shown. In Figure 3, the gray edges labeled as *a*, *b* and *c* show the inner dependencies and relations, while the blue, green and red edges represent how the related variables generate the classification output with, respectively, low, medium and high probability. The full comprehension if this information is useful on real industry applications, since it offers a simple and qualitative way, in terms of numerical effort, to evaluate the occurrence of faults.

With the Bayesian Networks, it is also possible to measure a trend for the beginning of the fault occurrence. Figure 4 presents the potential of the Bayesian Networks to measure the evolution to a fault. To generate the curves, each sample test was applied to the networks in Figure 2 and, for each resultant output, the probability of the sample being correctly classified was stored. Such probabilities are the values displayed in Figure 4. One can see that the probability can gradually change with the values of the attributes, showing the potential of the technique to track subtle changes from the machine that can lead to a fault. This feature is of utmost importance in industrial settings, since predictive maintenance actions can significantly lower the costs and improve the production process.

Another possible improvement for industry is to recommend, for initially healthy machines, to test the linguistic rules for each machine. Such procedure allows the storage of the typical audio signals obtained by each specific machine, improving the accuracy of classification patterns since the healthy pattern for each machine is already known, eliminating inaccurate classifications due to the presence of background sound.

5 Conclusion

The application of Bayesian Network structures for fault classification of electrical machines using audio signals, and a comparison with other machine learning techniques, was proposed in this paper. The proposed method consists of combining the input attributes to generate a set of networks, and then applying the BayesRule algorithm to extract linguistic rules which, through a set of "if-else" conditions, perform the desired fault classification.

---

[1] Developed by Pittsburgh University (https://dslpitt.org/genie/)

Table 5: Discretization of input attributes

Input Attributes	Ranges		
	Low	Medium	High
Freq 30	0 ; 0.196501	0.196501 ; 0.33865	0.33865 ; greater numerical value of data
Freq 60	0 ; 0.153198	0.153198 ; 0.25833	0.25833 ; greater numerical value of data
Freq 120	0 ; 0.073761	0.073761 ; 0.128693	0.128693 ; greater numerical value of data
Freq 2500	0 ; 0.0300057	0.0300057 ; 0.052686	0.052686 ; greater numerical value of data
Signal power	0 ; 17077.1	17077.1 ; 56241.8	56241.8 ; greater numerical value of data
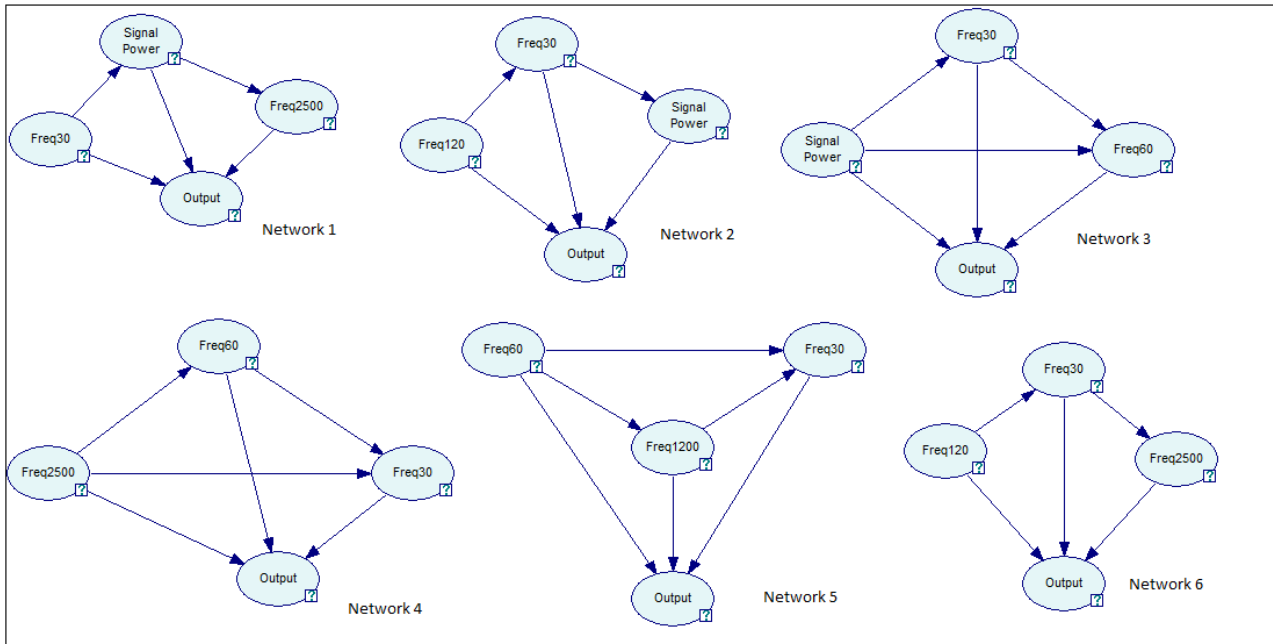
**Fig. 2:** The six Bayesian Network structures generated.

Table 6: Accuracies of the proposed Bayesian Networks.

Proposed Structure	Input 1	Input 2	Input 3	Accuracy
Network 1	freq30	freq2500	signal power	65%
Network 2	freq30	freq120	signal power	58%
Network 3	freq30	freq60	signal power	71%
Network 4	freq30	freq60	freq2500	65%
Network 5	freq30	freq60	freq120	63%
Network 6	freq30	freq120	freq2500	67%

The input attributes stem from audio signals acquired from microphones positioned around the electrical motor to be analyzed, which is less invasive than other kinds of sensors usually considered for this task. The Bayesian Network structures were obtained from a training procedure, and the obtained results are then compared with SVM and KNN techniques.

Analyzing the classification accuracy from each method, the accuracies obtained in this paper are in accordance with the performances presented in the literature for similar problems, as can be seen in the review [9]. In addition, classification methods presented in the literature use only the Naive-Bayes structure [8,10] and do not explore Bayesian networks topologies. The algorithms proposed in this current paper provide the generalization of the classification model for the decision-making process, by revealing the network structures. Therefore, the main advantage of these Bayesian Network structures, which is one important contribution of this paper, is to explicit the relationships and dependencies between the input attributes, allowing a qualitative analysis of each fault and how the attributes can be interconnected to perform the classification. As a result, the obtained structures can be used to perform a numerically simple classification procedure, or be applied along with other technique to improve the results.
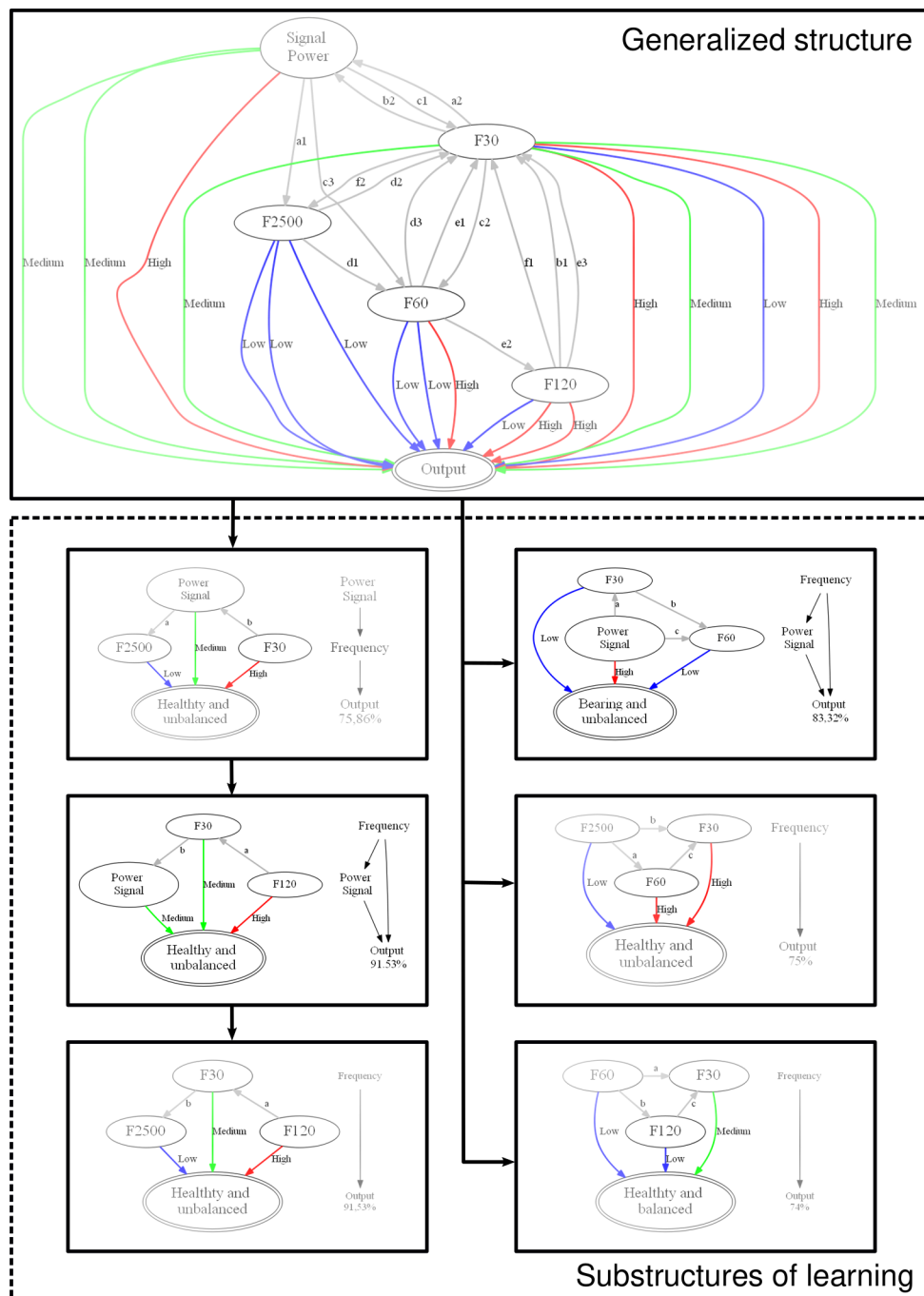
Fig. 3: Generalized structure to infer the relations of features from audio frequencies and signal power. In the generalized structure and in the substructures of learning, the gray edges show the inner dependencies and relations, while the blue, green and red edges represent how the related variables generate the classification output with, respectively, low, medium and high probability.
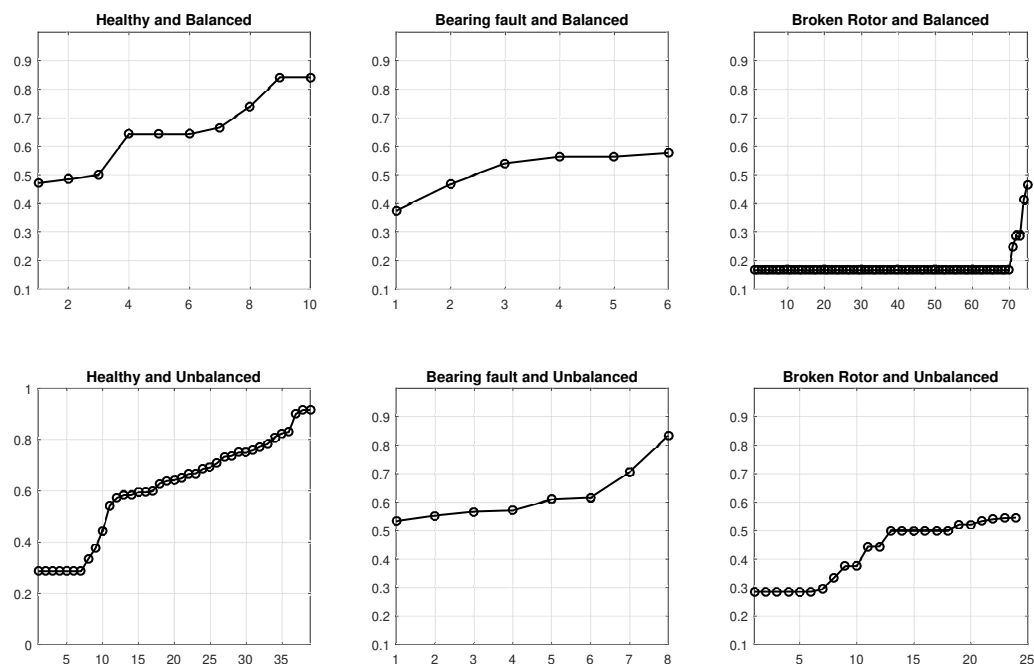
Fig. 4: Trend graphs to describe the inference about accuracy of the substructures from each fault network generated.

Since numerical experiments were performed, it is important to mention the runtime of the algorithms process on solving the proposed problems. In order to run each one of the 6 proposed structures (according to Table 2), the average runtime of the algorithms is: KNN - 3min, Bayesian networks - 6min, SVM - 7min.

Finally, for future works, one can improve the obtained results to develop a system to track the evolution of a fault within a given machine, since this paper shows the potential of the proposed technique for this task.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] I. Martin-Diaz, D. Morinigo-Sotelo, O. Duque-Perez, and R. J. Romero-Troncoso, "An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions," *IEEE Transactions on Industry Applications*, vol. 54, no. 3, pp. 2215–2224, 2018.

[2] A. J. Bazurto, E. C. Quispe, and R. C. Mendoza, "Causes and failures classification of industrial electric motor," *IEEE ANDESCON*, 2016.

[3] A. Bellini, F. Filippetti, C. Tassoni, and G. Capolino, "Advances in diagnostic techniques for induction machines," *IEEE Transactions on Industrial Eletronics*, vol. 55, no. 10, pp. 4109–2008, 2008.

[4] I. Aydin, M. Karakose, and E. Akin, "An approach for automated fault diagnosis based on a fuzzy decision tree and boundary analysis of a reconstructed phase space," *ISA Transactions*, vol. 53, pp. 220–229, 2014.

[5] M. Barzegaran, A. Mazloomzadeh, and O. Mohammed, "Fault diagnosis of the asyn-chronous machines through magnetic signature analysis using finite-elementmethod and neural networks," *IEEE Transactions on Energy Conversion*, vol. 28, no. 4, p. 1064–1071, 2013.

[6] H. Ertunc, H. Ocak, and C. Aliustaoglu, "ANN- and ANFIS-based multi-staged decisionalgorithm for the detection and diagnosis of bearing faults," *Neural Computing and Applications*, vol. 22, no. 1, p. 435–446, 2013.

[7] M. Hajiaghajani, H. Toliyat, and I. Panahi, "Advanced fault diagnosis of a DCmotor," *IEEE Transactions on energy conversion*, vol. 19, no. 1, p. 60–65, 2004.

[8] R. H. C. Palácios, I. N. Silva, A. Goedtel, and W. F. Godoy, "A comprehensive evaluation of intelligent classifiers for faultidentification in three-phase induction motors," *Electric Power Systems Research*, vol. 127, no. 1, p. 249–258, 2015.

[9] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[10] J. Wang, S. Liu, R. Gao, and R. Yan, "Current envelope analysis for defect iden-tification and diagnosis in induction

motors," *Journal of Manufacturing Systems*, vol. 31, no. 4, p. 380–387, 2012.

[11] O. Ondel, E. Boutleux, and G. Clerc, "A method to detect broken bars in inductionmachine using pattern recognition techniques," *IEEE Transactions on Industry Applications*, vol. 42, no. 4, p. 916–923, 2006.

[12] B. Ebrahimi, M. J. Roshtkhari, J. Faiz, and S. Khatami, "Advanced eccentricity fault recognition in permanent magnet synchronous motors usingstator current signature analysis," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 4, p. 2041–2052, 2014.

[13] A. Glowacz, W. Glowacz, Z. Glowacz, and J. Kozik, "Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals," *Measurement*, vol. 113, pp. 1–9, 2018.

[14] P. Bangalore and L. B. Tjernberg, "An artificial neural network approach for early fault detection of gearbox bearings," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 980–987, 2015.

[15] W. S. Gongora, A. Goedtel, A. A. O. da Silva, and C. L. Graciola, "Neural approach to fault detection in three-phase induction motors," *IEEE Latin America Transactions*, vol. 14, no. 3, pp. 1279–1288, 2016.

[16] P. Konar and P. Chattopadhyay, "Bearing fault detection of induction motor usingwavelet and support vector machines (SVMs)," *Applied Soft Computing*, vol. 11, no. 6, p. 4203–4211, 2011.

[17] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley: Wellesley-Cambridge Press, revised ed., 1997.

[18] M. Elforjani and S. Shanbr, "Prognosis of bearing acoustic emission signals using supervised machine learning," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5864–5871, 2018.

[19] B. Chen, L. Xie, Y. Li, and B. Gao, "Acoustical damage detection of wind turbine yaw system using bayesian network," *Renewable Energy*, vol. 160, pp. 1364–1372, 2020.

[20] H. O. Omoregbee and P. S. Heyns, "Fault detection in roller bearing operating at low speed and varying loads using bayesian robust new hidden markov model," *Journal of Mechanical Science and Technology*, vol. 32, no. 9, pp. 4025–4036, 2018.

[21] E. R. Hruschka, M. C. Nicoletti, V. A. Oliveira, and G. M. Bressan, "Markov-blanket based strategy for translating a bayesian classifier into a reduced set of classification rules," *7th International Conference on Hybrid Intelligent Systems*, pp. 192–197, 2007.

[22] G. Dougherty, *Pattern Recognition and Classification, An Introduction*. California - USA: Springer, 2013.

[23] B. C. F. de Azevedo, G. M. Bressan, C. M. Agulhari, H. L. dos Santos, and W. Endo, "Three-phase induction motors faults classification using audio signals and decision trees," *Applied Mathematics and Information Sciences*, vol. 13, no. 5, pp. 847–858, 2019.

[24] P. S. Bhowmik, S. Pradhan, and M. Prakash, "Fault diagnostic and monitoring methods of induction motor : A review," *Int. J. Appl. Control. Electr. Electron. Eng.*, vol. 1, no. 1, pp. 1–18, 2013.

[25] G. E. P. Box and G. C. Tiao, *Bayesian inference in statistical analysis*. Canadá: John Wiley and Sons, 1992.

[26] R. E. Neapolitan, *Learning bayesian networks*. Prentice Hall Series in Artificial Intelligence, USA: Pearson Prentice Hall, 2003.

[27] J. Pearl, *Probabilistic reasoning in intelligent Systems: networks of plausible inference*. USA: Morgan Kaufmann, 1988.

[28] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, "Learning bayesian networks from data: An information-theory based approach," *Artificial Intelligence*, vol. 137, pp. 43–90, 2002.

[29] J. Han, M. Kamber, and J. Pei, *Data mining: Concepts and Techniques*. USA: Morgan Kaufmann, 2012.

[30] M. Campbell and G. Arce, "Effect of motor voltage unbalance on motor vibration: Test and evaluation," in *2016 Petroleum and Chemical Industry Technical Conference (PCIC)*, (Philadelphia, Pennsylvania, USA), pp. 1–7, Institute of Electrical and Electronics Engineers (IEEE), Sept 2016.

[31] P. A. Delgado-Arredondo *et al.*, "Methodology for fault detection in induction motors via sound and vibration signals," *Mechanical Systems and Signal Processing*, vol. 83, no. Supplement C, pp. 568 – 589, 2017.

[32] V. Rai and A. Mohanty, "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2607 – 2615, 2007.

[33] S. Djurović, D. S. Vilchis-Rodriguez, and A. C. Smith, "Investigation of wound rotor induction machine vibration signal under stator electrical fault conditions," *The Journal of Engineering*, vol. 2014, no. 5, pp. 248–258, 2014.

[34] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Appers in the International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1137–1143, 1995.

[35] B. Lerner and R. Malka, "Investigation of the k2 algorithm in learning bayesian network classifiers," *Applied Artificial Intelligence*, vol. 25, pp. 74–96, 2011.

Glaucia Maria Bressan received a teaching degree in Mathematics with emphasis in computation from Federal University of Sao Carlos (UFSCar), Brazil, in 2000, a Master's degree in Computational and Applied Mathematics from the University of Sao Paulo (USP) in 2003 and her PhD degree in Electrical Engineering at University of Sao Paulo (USP) in 2007. She did a Postdoctoral research in Electrical Engineering at University of Sao Paulo (USP) in 2008. She works as a professor and researcher at the Federal University of Technology of Parana (UTFPR), Brazil, in the Department of Mathematics, where she has been a professor since 2012. research interests are in the areas of Applied Mathematics, as Operational Research, Machine Learning, Bayesian Networks and Fuzzy Logic.
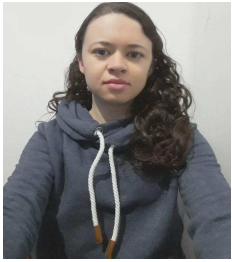
Beatriz Flamia Azevedo obtained her bachelor degree in Control and Automation Engineering at Federal University of Technology - Parana (UTFPR-Brazil). She is master in Industrial Engineering at Polytechnic Institute of Braganca (IPB-Portugal). Currently, she is working as a researcher at IPB and she is PhD student at Universty of Minho. Her interests are relate to optimization, mathematics modeling and machine learning.

Herman L. dos Santos received the B.S. in Automation and Control Engineering and M.Sc. in Electrical Engineering degrees from Federal University of Technology of Parana, Parana, Brazil, in 2016 and 2019, respectively. Since 2020, he has been with Londrina State University (UEL), Parana, Brazil in Electrical Engineering PhD program.

Wagner Endo received the B.S. degree in Electrical Engineering in 2004, the M.Sc. degree in Electrical Engineering in 2006 both from the State University of Londrina, Londrina, Brazil and the Ph.D. degree in electrical engineering from the University of Sao Paulo (USP), Sao Carlos, Brazil, in 2014. Currently, he is an Associate Professor with the Federal Technological University of Parana, Cornelio Procopio, Brazil. His research interests are within the fields of signal processing and automation of systems.

Cristiano Marcos Agulhari received the B.S. degree in Computer Engineering from Universidade Estadual de Campinas (2006), master's at Electric Engineering (2009) and doctorate at Electric Engineering (2013), both from Universidade Estadual de Campinas. Doctorate was done in cotutelle with the Institut National des Sciences Appliquees (INSA), Toulouse. Currently teaching at Universidade Tecnologica Federal do Parana, acting on the following subjects: Signal processing, wavelets, data compression, biomedical signal compression, control systems, robust control, control of LTV systems, linear matrix inequalities (LMIs).

Alessandro Goedtel received the B.S. degree in electrical engineering from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 1997, the M.Sc. degree in industrial engineering from Sao Paulo State University, Sao Paulo, Brazil, in 2003, and the Ph.D. degree in electrical engineering from the University of Sao Paulo, Sao Paulo, in 2007. He is currently a Full Professor with the Department of Electrical Engineering, Federal University of Technology?Parana, Cornelio Procopio, Brazil. His research interests include electrical machinery, intelligent systems, faults identification, and power electronics.

Paulo Rogério Scalassara received his B.S. and M.S. degrees in Electrical Engineering from the State University of Londrina, Brazil, in 2003 and 2005, respectively. He obtained his Ph.D. from the University of São Paulo in 2009. Since 2010, he has been with the Department of Electrical Engineering of the Federal University of Technology ? Paraná, in Cornélio Procópio, as an Associate Professor. In 2016, he was a Postdoctoral Researcher at the Computational NeuroEngineering Laboratory of the University of Florida, Gainesville, USA. His interests are in information theoretical learning, signal processing, machine learning, and embedded systems.