Computational Model for Electrical Motors Condition Analysis and Monitoring

Álvaro Souza de Melo[1*], Oberdan Pinheiro Rocha[1], Alex Álisson Bandeira Santos[1]
[1]SENAI CIMATEC University Center; Salvador, Bahia Brazil

Predictive maintenance uses several methods to monitor the conditions of electric motors applied in industrial plants. Among these methods, vibration analysis stands out as a widely used method due to the possibility of identifying a wide chance of failures. This work presents the development of a computational model endowed with a classifying algorithm to receive vibration readings from electric motors and determine if the machine has a fault behavior and, if so, which fault. For the algorithm to detect which type of failure, a dataset with readings from several engines in different failure conditions should be developed for training the model.
Keywords: Vibration Analysis. Condition-Based Monitoring. Feature Extraction. Classifier Agent.

Introduction

Maintenance is a critical and constantly evolving point in industries. Even with many technological advances, it is still a challenge to predict failures before happen. Predictive maintenance is essential, especially in the industrial environment, where the failure of some equipment that results in the need for corrective maintenance can cause the partial or total stoppage of the production plant, which causes inconvenience and losses for the company.

Vibration analysis in electric motors mitigates the possibility of failure. This method is an important tool for diagnosing operating conditions. Many engine problems reflect directly on their vibration due to several factors: misalignment, cavitation, and clearance, among others. Each component of a motor has a signature at a different frequency. It is possible to detect the type of fault present in the machine from the analysis of the behavior of the vibration profile of the equipment. Current studies show classification algorithms specialized in a specific failure, which does not allow their implementation in an industrial environment to obtain indications of several failures, since an electric motor can induce defects in different components of this equipment. This research project aims to develop a classifier algorithm that will determine the current condition of an electric motor based on a dataset containing readings from different equipment. To this research is considered feasible, the dataset was obtained in a laboratory with industrial vibration analysis equipment.

Different classification methods were also evaluated to obtain what brings greater precision and accuracy to detect the failure in the shortest execution time. After training the network using the model that best fits the problem solving, the technical feasibility of the proposed model was analyzed based on laboratory tests.

Materal and Methods

Classifier Algorithms

Classifier algorithms aim to predict the class of a new piece of data based on learning about similar data. They are a subcategory of supervised learning algorithms, where the objective is to predict the category of new data based on past observations [1]. We evaluated the performance, accuracy, and precision by three classification algorithms: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Artificial Neural Networks (ANN).

SVM

Support Vector Machine (SVM) is a classification and regression tool that uses machine

Received on 16 January 2021; revised 20 April 2022.
Address for correspondence: Álvaro Souza de Araújo. Av. Orlando Gomes, 1845 - Piatã, Salvador - BA- Brazil. Zipcode: 41650-010. E-mail: alvarosouza00@gmail.com.
DOI 10.34178/jbth.v5i2.206.

J Bioeng. Tech. Health 2022;5(2):100-105.

learning theory to maximize prediction accuracy, while automatically avoiding overfitting the network [2]. This algorithm performs the learning by assigning annotations to the data [3]. Also, according to Noble (2006), SVM can be used in several applications, from fraudulent credit card transactions and detection of handwriting patterns, through images, to applications in biology to recognize anomalies in DNA.

SVM algorithms are well known for their excellent performance in statistical classification. Still, the high computational cost due to the complexity of the cubic runtime is problematic for big data sets: training the SVM classifier requires solving a quadratic optimization problem [4].

Given an N number of training data for, where (x_i, y_i) is the input data and corresponds to its target value, the algorithm tries to find the hyperplane in a manner that the margin (distance perpendicular to the hyperplane) between support vectors is maximized [5]. If the data are linearly separable, there is a hyperplane, which can be expressed in the equation (1), considering the weight parameter.

$$f(x) = w^T x + b = \sum_{i=1}^{N} w_i x_i + b = 0 \qquad (1)$$

The regularization parameter C and kernel parameter y influence the performance of SVM [6]. The C is used to control the trade-off between maximizing the margin and minimizing the training error. For a given problem, if C is too large, SVM may store many support vectors, and it may be overfitting. If C is too small, SVM may again not fit properly, or it is underfitting [6].

RNA

Artificial Neural Networks can be defined as machines designed to model the working principle of a brain to perform a task [7]. The essential elements of this algorithm are the input connections, where each one has a weight to be determined, subsequent layers of neurons, an accumulator element to concentrate the signals, and, finally, an activation function, which can take different formats, to then, present the output value. One of the difficulties in neural networks is choosing the best architecture since this process is experimental and requires a great deal of execution time [8]. Thus, this technique requires extensive tests with different configurations to obtain the model best adapted to the problem.

Neural networks are similar to the human brain in two main aspects: the network acquires knowledge from its environment through the learning process, and connection forces between neurons (synaptic weights) are used to store the acquired knowledge [9]. At the output of each neuron, the generated signal goes through an activation function, which is responsible for weighing the effect of each output on the subsequent layer.

KNN

For pattern recognition, the KNN algorithm is a method to classify objects based on training examples that are closer in space [10]. This model was proposed by Fukunaga and Narendra (1975) and is a simple to implement classifier that can get very accurate results depending on the application [11].

This algorithm sorts data in a dataset based on proximity to already sorted data. Thus, the number of neighbors that must be considered for the classification of later data is determined as a parameter. This rule retains the entire training set during learning and assigns each query a class represented by the label of most of its closest neighbors in the training set [10].

The KNN classification method also has good accuracy and precision compared to the previously mentioned methods. The algorithm consists of estimating the distance between data to determine classification limits. It is important to emphasize that it is possible to graphically represent the data separation limits represented by up to three dimensions. With higher dimensions, the method can obtain the classification, however, the representation of more dimensions does not make physical sense.

To determine the degree of similarity of an entry with its closest neighbors, there are different methods for calculating the distance between the data, including the Euclidean distance, Hamming distance, and Manhattan distance, the first being the most used with input variables of the same type.

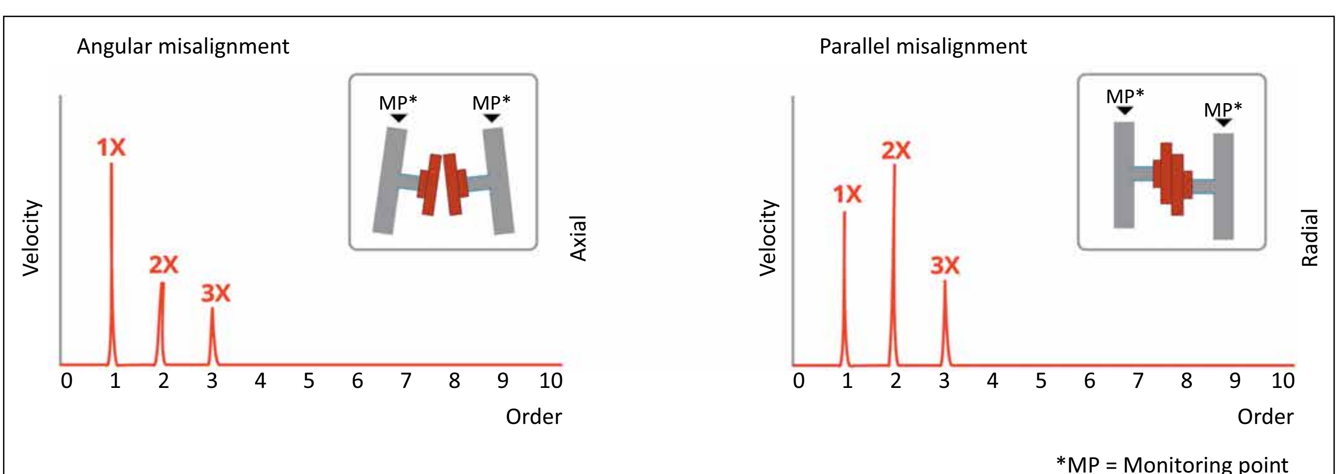
Vibration Analysis

Every machine has noise and vibration due to its operation and external excitations. However, a portion of the vibration is due to minor defects that compromise the performance of the equipment [12]. In this way, each failure behavior has a different frequency spectrum, which allows the distinction of the type of failure and knowing the specifications of the machine operation and the degree of severity. For the detection of failures of mechanical origin, vibration analysis is used, where, using the equipment equipped with an accelerometer sensor, a vibration signal in acceleration, velocity or displacement is measured, to then observe these data in the spectrum of time or frequency. In Figure 1, it is possible to visualize the behavior in the frequency spectrum of two faults: the angular and parallel misalignment, where 1x represents the dominant operating frequency of the motor. If it is powered directly from a 60 Hz electrical grid, which means that this is its dominant frequency.

Dataset

A dataset MaFaulDa [13] was selected to implement and train this model, which contains 1,951 multivariable series containing vibration sensor readings. This dataset covers the following equipment conditions: normal operation, unbalance, axial and radial misalignment, and internal and external bearing failures. Each reading is composed of three single-axis accelerometers, model IMI 601A01, in this dataset, each positioned on the radial, axial, and tangential axes, a triaxial accelerometer, model IMI 604B31, an analog tachometer, model MT-190, and a microphone Shure SM81 model. The dataset contains data referring to multiple rotation frequencies, ranging from 737 to 3,686 rpm. The first sensor, the uniaxial accelerometer, has a sensitivity of 100 mV/g, which can measure accelerations between -50 g and +50 g, and with a frequency spectrum from 0.27 Hz to 10,000 Hz. The operating principle of this equipment is a piezoelectric sensor, where a mass is in direct contact with the piezoelectric element and, when subjected to an acceleration, this mass exerts a mechanical force on the load that can be converted

Figure 1. Angular and parallel misalignment and the frequency behaviors.



into an electrical signal, thus being readable through a converter analog to digitally connected to a controller. This type of sensor is suitable for industrial applications that require high-frequency responses while maintaining stable responses in varying temperature environments. The triaxial accelerometer is the second sensor, which also has the same sensitivity as the previous one (100 mV/g) and the same measurement range (-50 g to +50g). However, its frequency response varies between 0.5 and 5,000 Hz. It has a ceramic sensor element, which results in low interference from external noise, suffering minimal reading variations when subjected to different temperatures.

Proposed Model

The first step to the entire dataset is to perform the Fast Fourier Transform in all signals to extract the frequency spectrum. Then, the preprocessing algorithm will retrieve statistic features, such as standard deviation, kurtosis, mean, skewness, and variance. ARFF file stores all this information that can train the classifiers. With Weka machine learning software [14], the next step performs an analysis to determine which features are relevant to the models. Therefore, another ARFF file is generated, including the most relevant features. Then, the filtered dataset is used to train classifiers using KNN, SVM, and RNA algorithms. The RNA algorithm was configured with a 0.3 learning rate, using one hidden layer containing 23 neurons, as described in equation (2). Furthermore, the classifier was limited to 500 epochs as stopping criteria.

$$n_{neurons} = \frac{(classes + attributes)}{2} \qquad (2)$$

Since the KNN classifier is a simpler type of classifier, it has fewer parameters to adjust. The model used 5 nearest neighbors, with the Euclidian distance. At last, the SVM model used the polynomial kernel, with the C parameter set to 0.1.

Results and Discussion

After evaluating the accuracy of all three models, the SVM presented the best accuracy for the dataset, as per Table 1. It is important to note that the time required to train the dataset using this classifier is comparable to the RNA, but considerably larger than KNN, due to the nature of the algorithm, which demands higher computational power as the number of neurons and layers increases.

Table 1. Accuracy of the classifier algorithms.

Algorithm	Accuracy
SVM	95.6433%
RNA	94.0031%
KNN	91.3378%

Comparing the results, it is possible to note that the SVM algorithm performed the best, considering the accuracy. Furthermore, in the confusion matrix, the entries, in general, were classified with minimal error, with the class IMBALANCE reporting the lowest score of 93.09% of correct classifications. Other classes such as UNDERHANG_BEARING_CAGE_FAULT and NORMAL contained, respectively, 93.61% and 93.09% of accuracy, which slightly decreased the overall accuracy of the model, to 95.64%. However, the other classes were correctly classified with over 96% of certainty (Figure 2).

The RNA classifier provided the second most accurate result. In its confusion matrix, it showed that the algorithm obtained 55.11% of correct classifications for the class NORMAL, and it negatively impacted the lower global accuracy of the model. However, the other classes presented over 91% accuracy, as shown on the confusion matrix for this algorithm (Figure 3).

At last, the KNN classifier presented the lowest accuracy of the three but was close to the previous RNA values. Differently from the prior results, this model could classify 42.89% of the entries from the NORMAL class correctly. However, this fact still indicates that this model encounters difficulty distinguishing NORMAL and other categories,

Figure 2. Confusion matrix from the results of the SVM classifier.

a	b	c	d	e	f	g	h	i	j	<-- classified as
46	1	1	0	1	0	0	0	0	0	a = NORMAL
5	187	1	0	2	0	0	2	0	0	b = HORIZONTAL_MISALIGNMENT
0	10	285	0	6	0	0	0	0	0	c = VERTICAL_MISALIGNMENT
1	6	7	310	8	0	0	1	0	0	d = IMBALANCE
4	4	0	1	176	1	0	1	1	0	e = UNDERHANG_BEARING_CAGE_FAULT
2	0	0	0	0	182	0	0	0	0	f = UNDERHANG_BEARING_OUTER_RACE
1	0	0	0	0	0	185	0	0	0	g = UNDERHANG_BEARING_BALL_FAULT
4	1	0	0	0	0	0	178	5	0	h = OVERHANG_BEARING_CAGE_FAULT
1	1	0	0	2	0	0	4	180	0	i = OVERHANG_BEARING_OUTER_RACE
0	0	0	0	0	0	0	0	0	137	j = OVERHANG_BEARING_BALL_FAULT

Figure 3. Confusion matrix from the results of the RNA classifier.

a	b	c	d	e	f	g	h	i	j	<-- classified as
27	7	3	6	5	0	0	0	1	0	a = NORMAL
4	175	11	0	2	1	0	2	1	1	b = HORIZONTAL_MISALIGNMENT
0	4	286	1	4	0	3	2	0	1	c = VERTICAL_MISALIGNMENT
0	3	1	323	1	0	2	1	1	1	d = IMBALANCE
1	3	5	6	167	1	0	4	1	0	e = UNDERHANG_BEARING_CAGE_FAULT
0	0	0	0	0	184	0	0	0	0	f = UNDERHANG_BEARING_OUTER_RACE
0	0	0	0	0	0	186	0	0	0	g = UNDERHANG_BEARING_BALL_FAULT
0	1	7	2	3	2	0	172	1	0	h = OVERHANG_BEARING_CAGE_FAULT
1	1	1	0	2	0	0	6	177	0	i = OVERHANG_BEARING_OUTER_RACE
0	0	0	0	0	0	0	0	0	137	j = OVERHANG_BEARING_BALL_FAULT

such as HORIZONTAL_MISALIGNMENT, with 78.71% accuracy. The other classes presented over 89% of accuracy, as shown in Figure 4.

Conclusion

The proposed model could correctly identify different mechanical faults on electrical motors using the vibration analysis readings. To obtain the result, an initial processing step is required to transform the signals to the frequency spectrum, extract relevant statistical features, and filter the amount of data provided to train the classifiers.

We trained the model to detect faults in multiple rotational speeds on the motor to provide better accuracy in a real-world

Figure 4. Confusion matrix from the results of the KNN classifier.

a	b	c	d	e	f	g	h	i	j	<-- classified as
21	11	6	0	11	0	0	0	0	0	a = NORMAL
9	155	20	2	5	0	0	4	2	0	b = HORIZONTAL_MISALIGNMENT
0	8	288	3	0	0	1	1	0	0	c = VERTICAL_MISALIGNMENT
0	2	14	313	2	0	0	1	1	0	d = IMBALANCE
3	10	1	7	160	1	0	3	3	0	e = UNDERHANG_BEARING_CAGE_FAULT
0	0	0	0	0	183	0	1	0	0	f = UNDERHANG_BEARING_OUTER_RACE
0	0	0	0	0	0	186	0	0	0	g = UNDERHANG_BEARING_BALL_FAULT
0	7	0	0	2	1	0	169	9	0	h = OVERHANG_BEARING_CAGE_FAULT
1	2	0	1	1	0	0	13	170	0	i = OVERHANG_BEARING_OUTER_RACE
0	0	0	0	0	0	0	0	0	137	j = OVERHANG_BEARING_BALL_FAULT

scenario, since, in many applications, the frequency inverters control equipment's speed. Furthermore, the classifier was tested under three different algorithms to evaluate which one provides a better result: SVM, KNN, and RNA. Tests were done to determine the best parameters for all three classifiers. From the results, it is possible to note that the SVM algorithm is better suited for classifying this dataset and detecting with better accuracy which failure a vibration signal contains, despite the higher computational requirements, especially since the other algorithms such as KNN and RNA presented a lower accuracy with some classes that are relevant to the failure detection.

References

1. Raschka S. Python machine learning. Packt publishing ltd, 2015.
2. Jakkula V. Tutorial on support vector machine (svm). School of EECS, Washington State University 2006;37.
3. Noble WS. What is a support vector machine? Nature Biotechnology 2006;24(12):1565-1567.
4. Chapelle O et al. Choosing multiple parameters for support vector machines. Machine Learning 2002;46(1):131-159.
5. Kumar A, Kumar R. Time-frequency analysis and support vector machine in automatic detection of defect from vibration signal of centrifugal pump. Measurement 2017;108:119-133.
6. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST) 2011;2(3):1-27.
7. Jr, Ribeiro RF, De Almeida FA, Gomes GF. Fault classification in three-phase motors based on vibration signal analysis and artificial neural networks. Neural Computing and Applications 2020;32(18):15171-15189.
8. Haykin SS, Van Veen B. Sinais e sistemas. Bookman, 2001.
9. Fleck L et al. Redes neurais artificiais: Princípios básicos. Revista Eletrônica Científica Inovação e Tecnologia 2016;1(13):47-57.
10. Imandoust SB, Bolandraftar M. Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. International Journal of Engineering Research and Applications 2013;3(5):605-610.
11. Fukunaga K, Narendra PM. A branch and bound algorithm for computing k-nearest neighbors. IEEE Transactions on Computers 1975;100(7):750-753.
12. Da Silva LEB et al. Desenvolvimento de um Sistema de Detecção de Defeitos Precoces em Motores de Indução.
13. Ribeiro FML et al. Rotating machinery fault diagnosis using similarity-based models, 2017.
14. Frank E et al. The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition, 2016.