Influence of Dataset Structuring on Condition Monitoring of a Rotating System by Machine Learning

Flavio C. A. Pereira[1], Eduardo S. de Andrade[1], Eduardo M. O. Lopes[1], Eduardo L. O. Batista[2]

[1]*Dept. of Mechanical Engineering, Federal University of Parana*
Rua Cel. Francisco H. dos Santos, 100 – Bloco IV, 81530-900, Curitiba/PR, Brasil
flaviocanfilde@gmail.com, eduardosilvadeandrade97@gmail.com, eduardo_lopes@ufpr.br
[2]*Dept. of Electrical and Electronics Engineering, Federal University of Santa Catarina*
Campus Universitário Trindade, s/n, 88040-900, Florianópolis/SC, Brasil
eduardo.batista@ufsc.br

Abstract. Condition monitoring consists of constant data acquisition from a machine of interest to determine its operational condition, and also to give reasonable predictions regarding its behavior over time. Considering that vibration generated by a machine carries information about internal conditions and is sensitive to structural changes, vibration analysis can be employed to detect faulty components. As some defects have known vibrational responses ("vibrational signatures"), it is possible to infer the type of defect by analyzing the vibration signal characteristics. An algorithm capable of automatically doing this type of analysis could potentially prevent financial or physical harm. In this context, the present study focuses on preprocessing vibrational response data related to induced defects in a rotating system, extracting features of interest, using a machine learning classifier to identify common problems, and segregating troublesome conditions from expected normal operation ones. The processed data was obtained from the Machine Fault Dataset (MaFaulDa/UFRJ). The obtained results show the influence of dataset structuring on the algorithm generalization capability, revealing that bigger datasets do not always lead to superior performance and that an increase in the amount of attributes is not always the most interesting choice.

Keywords: Machine learning, Condition monitoring, Rotating machinery, Vibration analysis.

1 Introduction

Amongst many maintenance strategies in the industry, condition-based maintenance (CBM) has been widely applied due to its economic and life-safety advantages, making it one of the most efficient maintenance strategies [1]. When applied to rotating machines, CBM is referred to as condition monitoring (CM) [2]. This strategy, according to Scheffer and Girdhar [3], consists of constant acquisition of machinery data aiming not only to determine operational conditions, but also to give reasonable predictions of machine behavior over time [1].

CM can be implemented via vibration analysis, since vibration generated by a machine carries information about internal and structural conditions. In other words, the "vibration signature" [2][4] of a machine differs from the standard signature when a defect is present. According to Scheffer and Girdhar [3], CM through vibration analysis is the most effective technique for detecting faults in rotating machines.

In industrial applications, vibration obtained from an operating machine in real time can be compared with historical data from a former regular condition to provide diagnostics concerning the type of fault that is already present or is slowly developing. This is due to the fact that some defects have well known vibrational responses, which makes it possible to categorize the fault only by analyzing the characteristics of the vibration [5].

Machine learning methods have been gaining importance in the context of CM. For instance, Marins et al. [6] introduce a machine learning algorithm capable of classifying failures of rotating machines that could potentially prevent the majority of cases where human error would cause financial or physical harm.

In this study, vibration response data related to purposely induced defects in a rotating system is preprocessed to extract the corresponding features in the frequency domain and use neural networks to build an intelligent algorithm. This algorithm should be capable of identifying problems such as imbalance and misalignment and segregating those conditions from the expected normal operation conditions. The data comes from the Machine Fault Database (MaFaulDa), acquired at the Signal, Multimedia and Telecom Laboratory (UFRJ) [6]. The results show the influence of dataset structuring on the network's generalization capability, resulting in a biased algorithm capable of obtaining good performance when using randomized samples, but performing poorly when using more well distributed samples along the frequency range.

2 Methodology

As previously mentioned, the employed dataset is the MaFaulDa/UFRJ [6]. This database was obtained by introducing different defects on a SpectraQuest's machinery fault simulator, which is capable of reproducing several conditions of rotating machinery dynamics for different rotation speeds [https://spectraquest.com/simulators/details/mfs/]. Each measurement was made in a time interval of 5s. The vibration response was measured by one tachometer, one microphone and two sets of three accelerometers (one set for each bearing), in axial, radial and tangential directions.

Among the different kinds of faults considered in MaFaulDa, the analysis herein is focused on the horizontal misalignment. In order to make the data more treatable for purposes of this study, the data was resampled from its original 50 kHz sampling frequency to 1 kHz. Moreover, microphone data was not taken into consideration. It is important to mention that some small differences were observed between rotation speeds calculated from the tachometer readings and the nominal speeds. In the present work, it was then opted to use the calculated rotation speeds with no loss of generality.

The Fast Fourier Transform (FFT) was applied to each signal aiming at a frequency domain analysis. Regarding the structure of the datasets, the chosen approach was to consider the magnitude of the FFT in order to apply the so-called frequency/spectral analysis [7].

2.1 Frequency Analysis

When analyzing rotating machinery data in the frequency domain, a common practice is to search for peaks of amplitude, usually located at multiples of the fundamental rotation frequency. It is known that the application of FFT to a N point time vector returns a N point frequency vector where the absolute values start repeating after $N/2$ points due to the periodic nature of the FFT. Taking this into account, for each calculated FFT, only the first half of the generated array was included in the dataset.

2.2 Dataset assembly

In the present work, five different classes from the MaFaulDa dataset were retrieved: four classes of horizontal misalignment (0.5mm, 1.0mm, 1.5mm and 2.0mm) and the "normal" class, which represents the vibrational behavior of the machine in good operating conditions. For each of these five classes, there was one 5-second sample for each rotation speed. Then, three different forms of dataset structuring were considered here, resulting in three case studies, which are described in the following.

Case 1: The 5 seconds of data in each sample were taken into consideration without segmentation, resulting in FFTs with N=5000 points, from which the first N/2 FFT magnitude samples were used as attributes in the dataset (samples in the range 0 to 2.5 kHz).

Case 2: Same as Case 1, but using only the first N/5 first magnitude samples of the FFTs (in the range 0 to 500 Hz).

Case 3: A segmentation of each sample was performed, leading to five different 1-second samples for each rotation/class combination, which produced FFTs with N=1000 points. As in Case 1, only the N/2 first FFT magnitude samples were used as attributes (range 0 to 500 Hz).

For each of these 3 cases, 6 different classification tasks were performed, each considering data from a single accelerometer isolated. A seventh classification task was also performed for each case considering the data from all accelerometers together.

2.3 Machine Learning Algorithm Implementation

A machine learning (ML) approach based on a multilayer neural network (multilayer Perceptron - MLP) [8] was applied to all cases considered. Such an approach was implemented using the Scikit-Learn Python module [9][10] and trained using the Adam method [11].

When structuring the model, the default Scikit-Learn's MLP hyperparameters were used, except for the ones specified in the following. For the classification tasks set up with data from a single accelerometer, the model consisted of 1000, 700 and 500 neurons for each the three hidden layers, respectively; for the tasks involving features from all the 6 accelerometers together, the model consisted of 2000, 1000 and 500 neurons for each hidden layer, respectively. Moreover, the maximum number of training epochs was set to 500, the L2 regularization penalty was set to 0.01 and the maximum number of epochs after meeting the default tolerance was equal to 30. These values of hyperparameters were obtained from preliminary experiments.

3 Results and Discussion

The results obtained for Case 1 are presented in Table 1. These results show that, although the model generally has a very good performance (88.29%) when considering all accelerometer attributes together as features, it performs even better when taking only the external radial accelerometer (AccRad2) into account (93.69%). With respect to Case 2, from the results shown in Table 2, one can notice that better accuracy is obtained when the accelerometers are taken together (77.03%). For this case, the highest accuracy reached using a single accelerometer was 71.17% (AccRad2).

Table 1. Results for Case 1 (5 seconds window, 0 to 2500 Hz)

Accelerometer	Average of the best 3 train scores	Average of the best 3 test scores	Test variance for all tests
Axial 1	100%	59.01%	0.37%
Radial 1	100%	74.32%	0.42%
Tangent 1	100%	49.10%	0.27%
Axial 2	100%	55.41%	0.23%
Radial 2	100%	93.69%	0.33%
Tangent 2	100%	47.30%	0.22%
All	100%	88.29%	0.53%

Table 2. Results for Case 2 (5 seconds window, 0 to 500Hz)

Accelerometer	Average of the best 3 train scores	Average of the best 3 test scores	Test variance for all tests
Axial 1	100%	52.25%	0.32%
Radial 1	100%	50.00%	0.15%
Tangent 1	100%	41.44%	0.22%
Axial 2	100%	47.30%	0.21%
Radial 2	100%	71.17%	0.17%
Tangent 2	100%	42.79%	0.28%
All	100%	77.03%	0.43%

The results for Case 3 (with segmented signals) are shown in Table 3. From this table, one can notice that, when applying the MLP classifier to each single accelerometer, the best performance is obtained using AccRad2 (71.17%). However, the performance is even better when considering all the accelerometers together (81.98%), even exceeding the result obtained for Case 2 (without segmentation and frequency range of 0 to 500 Hz). The results of Tables 1-3 are related below to those of Figure 1, for classification tasks.

Table 3. Results for Case 3 (1 second window, 0 to 500 Hz)

Accelerometer	Average of the best 3 train scores	Average of the best 3 test scores	Test variance for all tests
Axial 1	100%	51.35%	0.31%
Radial 1	100%	49.55%	0.14%
Tangent 1	100%	43.69%	0.28%
Axial 2	100%	50.00%	0.26%
Radial 2	100%	71.17%	0.19%
Tangent 2	100%	39.19%	0.20%
All	100%	81.98%	0.28%

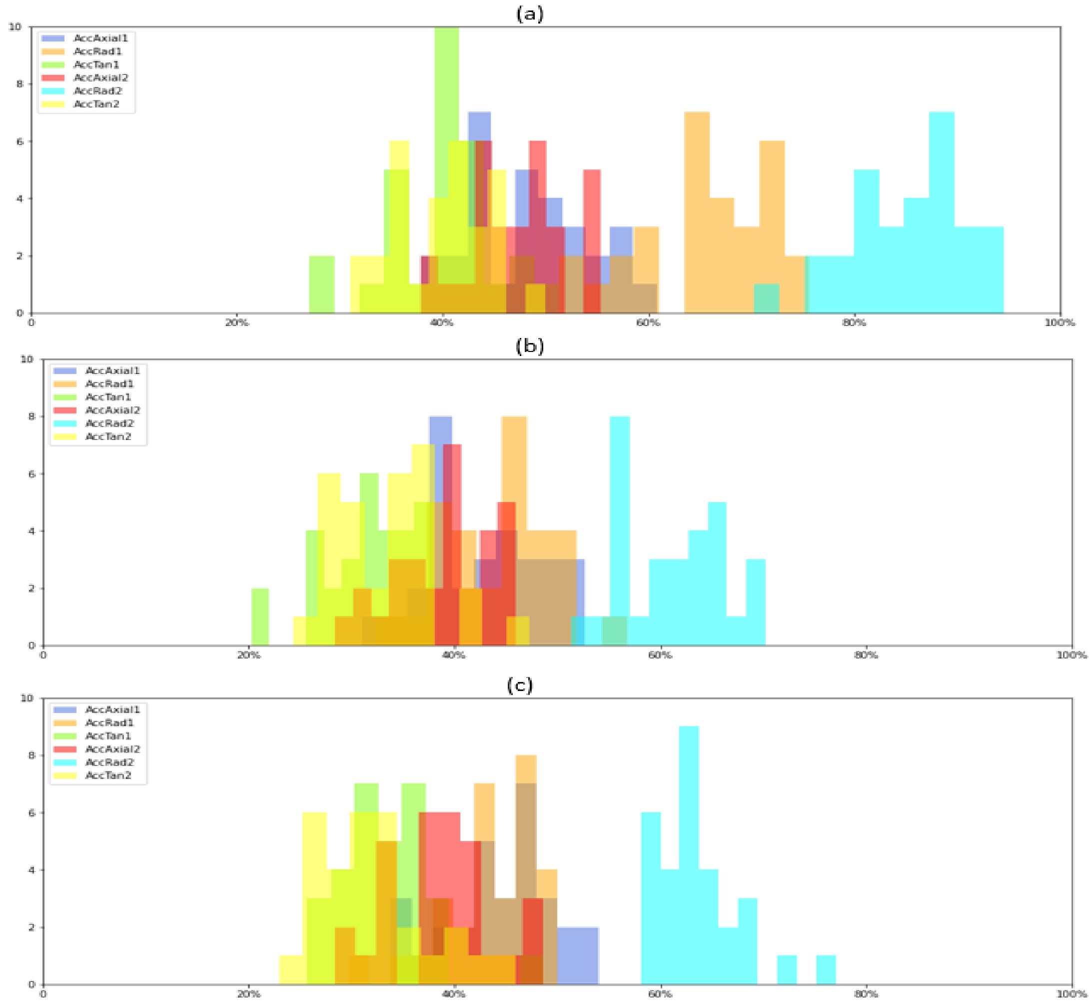

Figure 1. Score test distribution of all 30 executions for: a) Case 1 (5 seconds window, 0 to 2500 Hz); b) Case 2 (5 seconds window, 0 to 500 Hz); c) Case 3 (1 second window, 0 to 500 Hz)

When the classification tasks with a single accelerometer are compared, some interesting patterns emerge. Figure 1 shows histograms of accuracy results obtained running the training of the classifier thirty times for each single-accelerometer classification task. From this figure, one can notice that, for all cases, best performance was obtained using either radial accelerometer 1 (AccRad1) or 2 (AccRad2), which confirms the results from Table 1, Table 2 and Table 3. It is important to mention that this is the behavior reported as expected [12].

Another important result worth highlighting is that larger numbers of features do not necessarily result in superior performance. This is evident from the results obtained for Case 1 (Table 1), where the best results were obtained using a single accelerometer instead of all accelerometers together. The opposite was observed for Cases 2 and 3, where the best results were obtained when all accelerometers were accounted together for the classification task. It is also worth noting that the use of signal segmentation to enlarge the dataset did not deliver better results, contradicting the authors' expectations. That can be observed by comparing the results obtained for Case 3 with those from Case 1.

Thus, these results show that, for the cases considered, dataset size was not a determinant factor to ML algorithm performance and neither the number of features considered. Therefore, further investigation is required to find the most relevant attributes in a way to improve the classifier reliability and efficiency.

4 Conclusions

In this study, different dataset structures were submitted to a MLP neural network machine learning algorithm, revealing a clear influence of the dataset structure on the ML algorithm performance. Also, neither dataset size nor quantity of attributes were noted to be determinant on the algorithm's reliability. Further studies should be carried out to find the most relevant attributes concerning the vibrational response of rotating machines aiming to feed the algorithm with relevant features as well as to dismiss the irrelevant ones.

Acknowledgements. The authors would like to thank Agência Nacional do Petróleo (ANP), Financiadora de Estudos e Projetos (FINEP) and Programa de Recursos Humanos 12.1, as well as Universidade Federal do Paraná (UFPR) and Universidade Federal de Santa Catarina (UFSC). Eduardo M. O. Lopes would like to thank the National Council for Scientific and Technological Development – CNPq.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] R. B. Randall. Vibration-based Condition Monitoring: industrial, aerospace and automotive applications. West Sussex: John Wiley & Sons, Ltd, 2011.
[2] C. R. Farrar and K. Worden. Structural Health Monitoring, 1 ed. United Kingdom: John Wiley & Sons, Ltd, 2013.
[3] C. Scheffer and P. Girdhar. Practical Machinery Vibration Analysis and Predictive Maintenance. Oxford, United Kingdom: Newnes, 2004.
[4] S. Braun. Mechanical Signature Analysis: theory and applications. London: Academic Press, 1986.
[5] A. W. Lee. Vibration Problems in Machines: Diagnosis and Resolution. Florida: Taylor & Francis Group, 2016.
[6] M. A. Marins et al., "Improved similarity-based modeling for the classification of rotating-machine failures". *Journal of the Franklin Institute*, vol. 335, n. 4, pp. 1913-1930, 2018.
[7] A. Brandt. Noise and Vibration Analysis: Signal analysis and experimental procedures. West Sussex: John Wiley & Sons, Ltd, 2011.
[8] R. O. Duda et al. Pattern Classification, 2 ed. US: Wiley-Interscience, 2009.
[9] F. Pedregosa, "Scikit Learn: Machine Learning in Python". *Journal of Machine Learning Research*, vol. 12, n. 85, pp. 2285-22830, 2011.
[10] https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html
[11] D. P. Kingma and J. L. Ba, "ADAM: A Method For Stochastic Optimization". *3rd International Conference on Learning Representations (ICLR)*, 2015.
[12] J. C. A. Correa and A. A. L. Guzman. Mechanical Vibrations and Condition Monitoring. Academic Press, 2020.