MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 63 – 67
DOI: https://doi.org/10.15282/mekatronika.v2i1.6735

ORIGINAL ARTICLE

The Condition Based Monitoring for Bearing Health

Lim Weng Zhen[1], Anwar P.P Abdul Majeed[1*], Mohd Azraai Mohd Razman[1] and Ahmad Fakhri Ab. Nasir[1]

[1]Faculty of Manufacturing and Machatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.

ABSTRACT – Bearing is a small component that widely uses in industries, either in rotary machines or shafts. Faulty in bearing might cause massive downtime in the industries, which lead to loss of revenue. This paper intends to find the consequential statistical time-domain-based features that can be used in classification from accelerometry signals for the bearing condition. An accelerometer was used as the data logger device to attain the condition signals from the bearing. Machinery Failure Prevention Technology (MFPT) online dataset has three different bearing conditions: baseline condition, inner faulty condition, and outer faulty condition. Extraction of eight statistical time-domain features was done, which is root-mean-square (RMS), minimum (Min), maximum (Max), mean, median, standard deviation, variance, and skewness. The identification of informative attributes was made using a filter-based method, in which the scoring is done by using the Information gain ratio. For the extracted features, the data splitting of training data to testing data was set to the ratio of 70% and 30%, respectively. The selected feature for classification is then fed into various types of classifiers to observe the effect of this feature selection method on the classification performance. From this research, six features were identified as the significant features: variance, standard deviation, Min, Max, mean, and RMS. It is said that the classification accuracy of the training data and the testing data using the filter-based feature selection method is equivalent to the classification accuracy of all the features selected.

ARTICLE HISTORY
Received: 8th May 2020
Revised: 2nd June 2020
Accepted: 10th June 2020

KEYWORDS
Time-Domain Features
Condition Based Monitoring
Feature Selection
Bearing
Classification

INTRODUCTION

Bearing is a mechanical element that restricts relative movement to only required movement and reduces friction between moving parts. It is an important element in a machinery compartment, and without bearing, the machine could not function and rotate in its desired position. There are two main functions in bearing: make the rotation of the compartment smoother and protect the part that supports the rotation by maintaining the correct position for the rotating shaft. Faults in the bearings are the main cause of the breakdown in rotating machinery [1]. The breakdown of the bearing will lead to an increase in production costs for the industries. Worst come to worst, once the end product could not hand out to the customer on time, the industries might need to pay a large amount of compensation. Some research has shown that the breakdown of 40% in the large machinery and 90% in the small machinery systems is caused by bearing defects, making the detection and accurate diagnosis of bearing faults essential [2], [3].

The term Industrial Internet, Smart Manufacturing and Smart Production is widely used during the paradigm of the fourth industrial revolution. All the terms mention just now represent the changes in the industrial model by the irruption of the Internet of Things (IoT) [4]. The Internet of Things or known as IoT, is a system composed of interconnected computing devices, machinery and digital machines, objects, animals, or people. These systems have a unique identifier (UID) and can transmit data through the network without the need for people. Between-human-computer interaction or human-computer interaction. These latest challenges in the industrial sector require technologies previously used to build dynamic, intelligent, flexible and open applications that must work in a real-time environment [5], [6].

RELATED WORK

A summary can be made after reviewing a total of 30 related literature reviews, as shown in Figure 1 with regards to the classifiers used for condition based monitoring. Xu et al. [7] study the earing fault diagnosis by using random forest. Random forest is the popular model among the others models. It is the leading choice used for the condition based monitoring study. Support Vector Machine (SVM) gain the second with 23% followed by Neural Network (NN) with a percentage of 15%. Both NN and Logistic Regression (LR) gain the same percentage of 15%. Other models used by those researchers, the decision forest, decision jungle, boosted becision tree, gradient decision tree, and others. In this study, the *k*-Nearest Neighbor (K-NN), SVM and random forest (RF) is used as the classification models.

*CORRESPONDING AUTHOR |Anwar P.P Abdul Majeed | amajeed@ump.edu.my

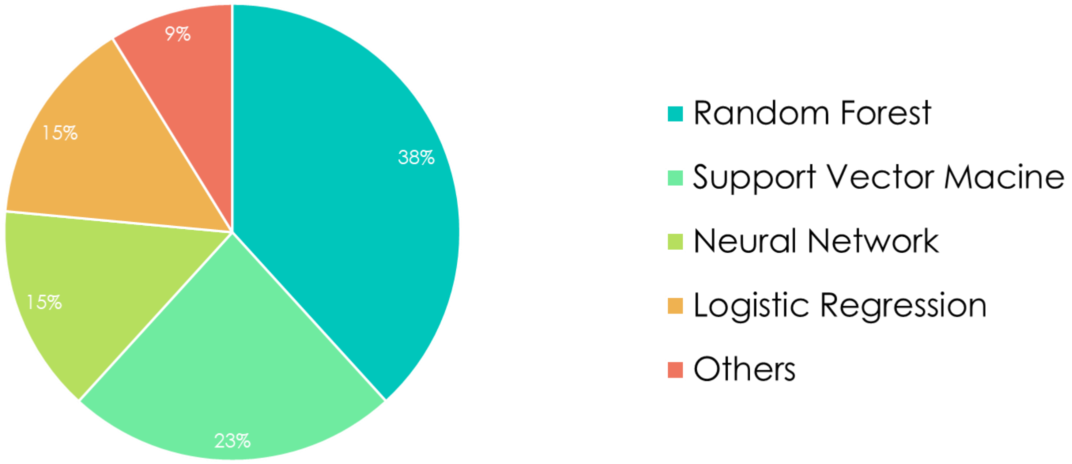

Figure 1. Classifiers Used for Condition Based Monitoring (2016 – 2020)

METHODOLOGY

Data Acquisition and Data Collecting Device (Accelerometer)

For the bearing signal recording, an accelerometer was used. It is used to record or measure the acceleration forces acting on the bearing. The data collected is then classified into three different classes labelled as the baseline, outer race fault and inner race fault, respectively. The sampling rate for the Machinery Failure Prevention Technology (MFPT) online dataset is 16,276 samples per second (sps), which will be used in this research. A total of 2070 data is extracted from the raw data and used in this study. The specimen data is collected from bearing of a machinery compartment. The reading is collected from operating machinery by using an accelerometer. The data set is collected under the same input shaft rate which are 25 Hz. There are two different sample rates in this data set which are 97,656 sps for baseline with a duration of 6 seconds and 48,828 sps for both outer race fault and inner race fault for approximately 3 seconds. The bearing parameters for bearing data set collection are 31.62 mm pitch diameter, 5.97 mm ball diameter with 8 number of elements.

The raw bearing signal is then gone through the preprocessing phase by using MATLAB. Figure 2 to Figure 4 illustrates each of the samples of the preprocessed bearing signal for baseline, outer race faculty and inner race faulty, respectively. The preprocessing phase is then split according to the signal window. The data is then split into different categories according to their classes. For the class baseline, a total of 950 windows can be spilt from the dataset while inner race faulty and outer race faulty can only manage to get 690 windows, respectively. To reduce the chances of bias between these signals, the windows' size was standardized to 690 windows. From the filtered samples, features are extracted after the preprocessing phase to acquire eight statistical time-domain [8], [9]. The extracted features are mean, median, standard deviation, variance, root-mean-square (RMS), minimum (Min), maximum (Max) and skewness for all three classes.

In this research, a filter-based method was implemented in the feature selection phase. Rather than depends on the machine learning algorithms, the filter-based method scored the features through some statistical tests to identify the correlation between the input feature and the output variable and the information gain ratio is chosen as the scoring method.

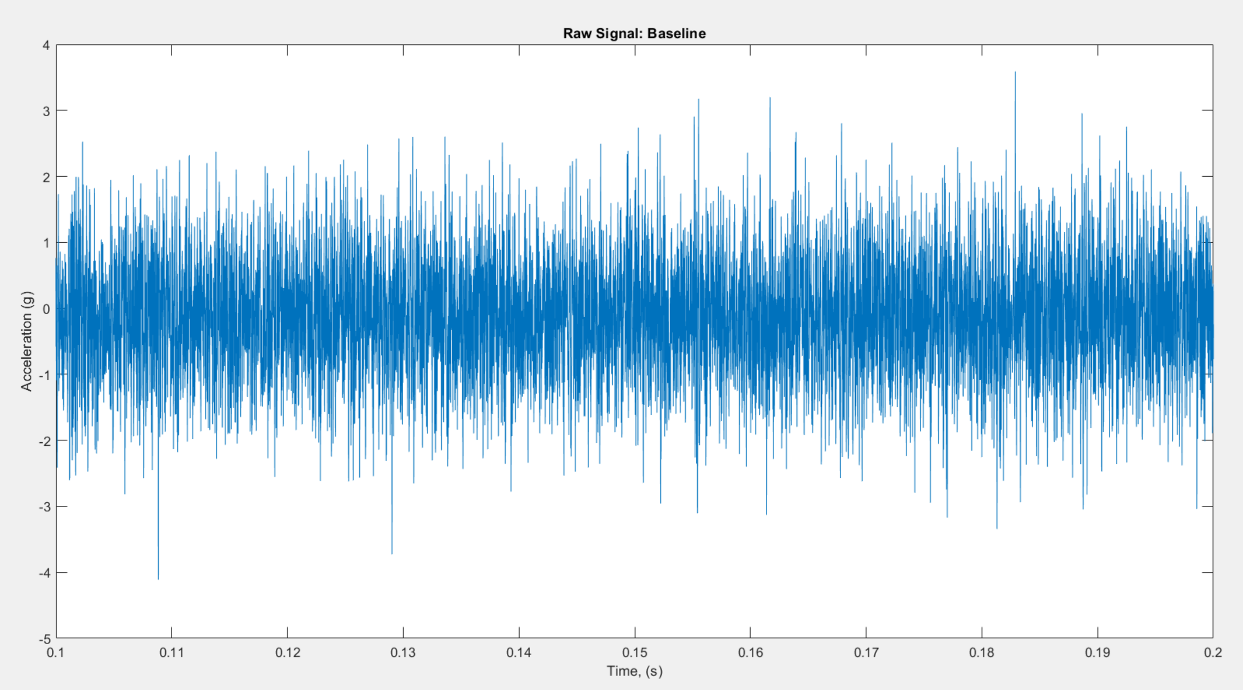

Figure 2. Preprocessed baseline signal.

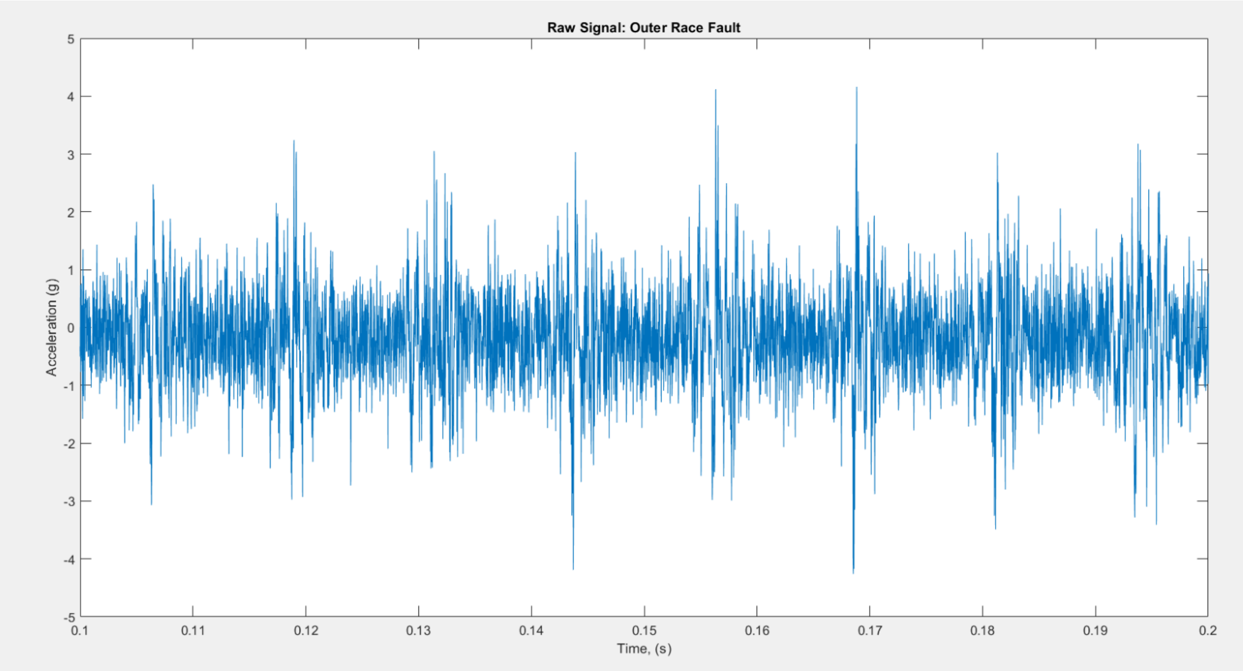

Figure 3. Preprocessed outer faulty signal.

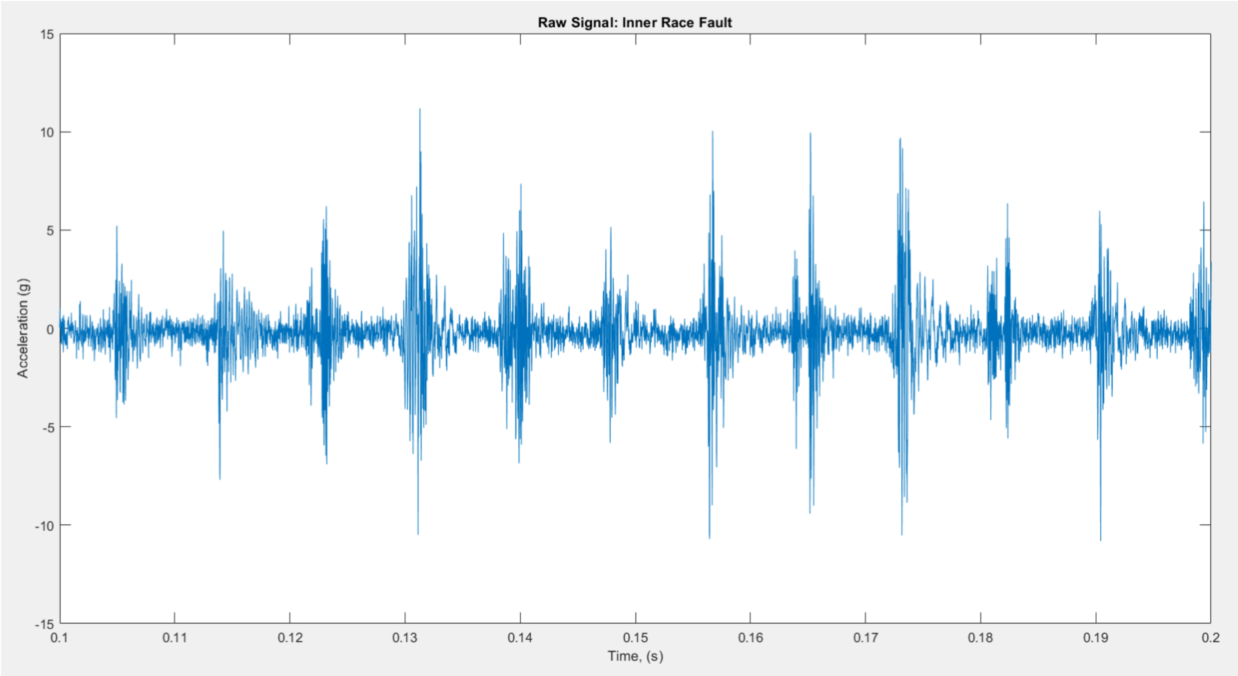

Figure 4. Preprocessed inner faulty signal.

RESULT AND DISCUSSION

In Figure 5, the score of information gain ratio was computed and the features of variance, standard deviation, Min, Max, mean and RMS are most significant with the score of 0.462, 0.393, 0.339, 0.272 and 0.262, respectively. The classification accuracy of both training and testing for different types of classification models are illustrated in Figure 6. It can be observed that with selected features or all features, the test Classification Accuracy (CA) and train CA for RF differ from other classifiers in which it has the high train CA but low in test CA. For selected features, k-NN has the highest test CA among all classifiers, 90.7%, followed by RF with the test CA 88.6%. SVM achieved the lowest test CA of 83.1%.

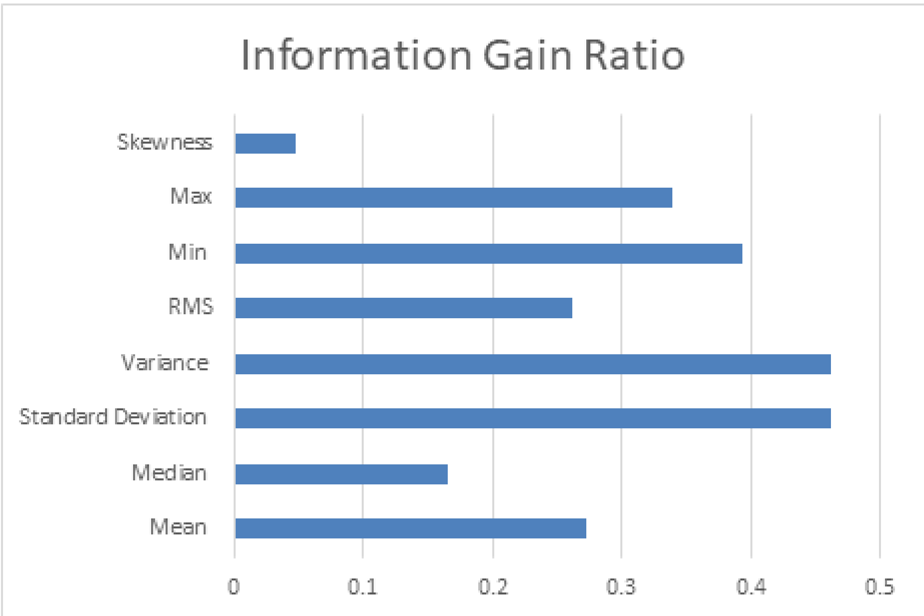

Figure 5. Features Extraction for Information gain ratio via filter-based method

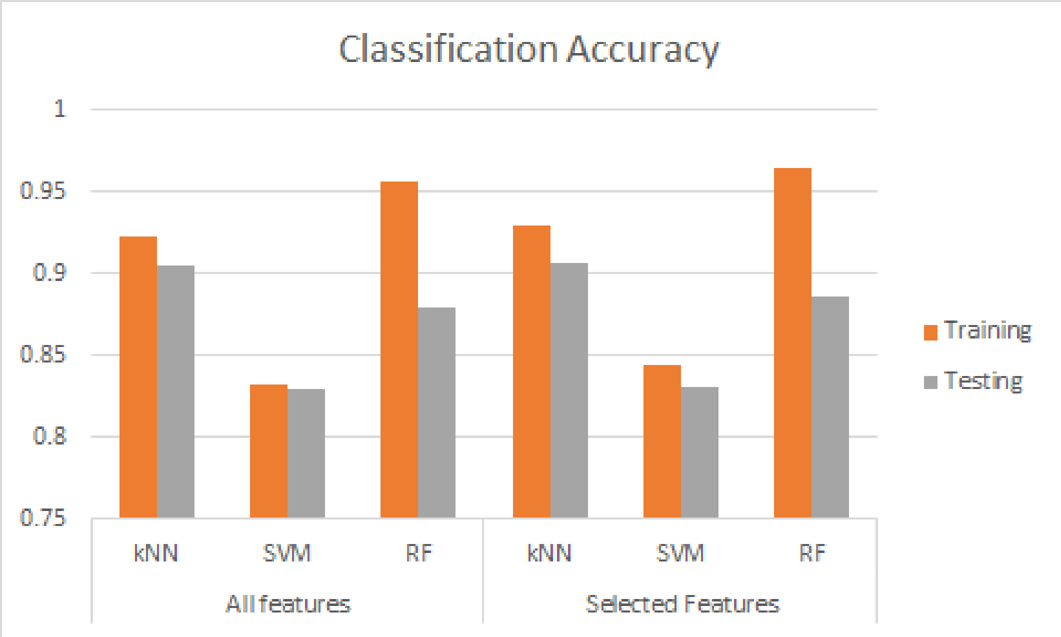

Figure 6. CA based on the different feature set

CONCLUSION

For this research, bearing signals were used as the raw signal, and these signals underwent preprocessing stage to improve the purity of the signal. Therefore, a total of 8 time-based statistical features are extracted from the preprocessed position signal. Then, the feature selection method based on the filter is used to find the information-rich features. Since these three features scored the highest, only the six features were selected from the information gain ratio. As shown in the results, it is proved that the bearing signal can be classified only by selecting essential features. Compared with the classification using all the features, comparable results can be obtained. Reducing functions by filter-based methods can reduce calculation costs as this is very important in real life for reducing machine downtime.

REFERENCES

[1] V. Vakharia, V. K. Gupta, and P. K. Kankar, "Ball Bearing Fault Diagnosis using Supervised and Unsupervised Machine Learning Methods," *Int. J. Acoust. Vib.*, vol. 20, no. 4, 2015, doi: 10.20855/ijav.2015.20.4387.

[2] M. R. Shahriar, P. Borghesani, and A. C. C. Tan, "Electrical Signature Analysis-Based Detection of External Bearing Faults in Electromechanical Drivetrains," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5941–5950, 2018, doi: 10.1109/TIE.2017.2782240.

[3] D. T. Hoang and H. J. Kang, "A Motor Current Signal-Based Bearing Fault Diagnosis Using Deep Learning and Information Fusion," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3325–3333, 2020, doi: 10.1109/TIM.2019.2933119.

[4] E. Hern, S. Rodr, T. S. Mart, and A. Gonz, "Machine Learning Predictive Model," *Springer*, vol. 1, pp. 501–510, doi: 10.1007/978-3-319-95204-8.

[5] M. Glez-Bedia, J. M. Corchado, E. S. Corchado, and C. Fyfe, "Analytical model for constructing deliberative agents," *Int. J. Eng. Intell. Syst. Electr. Eng. Commun.*, vol. 10, no. 3, pp. 173–185, 2002.

[6] J. M. Corchado, E. S. Corchado, J. Aiken, C. Fyfe, F. Fernandez, and M. Gonzalez, "Maximum likelihood hebbian learning based retrieval method for CBR systems," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2689, pp. 107–121, 2003, doi: 10.1007/3-540-45006-8_11.

[7] G. Xu, M. Liu, Z. Jiang, D. Söffker, and W. Shen, "Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning," *Sensors (Switzerland)*, vol. 19, no. 5, 2019, doi: 10.3390/s19051088.

[8] R. N. Toma, A. E. Prosvirin, and J. M. Kim, "Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers," *Sensors (Switzerland)*, vol. 20, no. 7, 2020, doi: 10.3390/s20071884.

[9] C. Sobie, C. Freitas, and M. Nicolai, "Simulation-driven machine learning: Bearing fault classification," *Mech. Syst. Signal Process.*, vol. 99, pp. 403–419, 2018, doi: 10.1016/j.ymssp.2017.06.025.