Production Management and Process Control, Vol. 104, 2023, 29–35
https://doi.org/10.54941/ahfe1003049

Empirical Study of Machine Learning for Intelligent Bearing Fault Diagnosis

Armin Moghadam[1] and Fatemeh Davoudi Kakhki[1,2]

[1]Department of Technology, San Jose State University, San Jose, CA 95192, USA
[2]Machine Learning & Safety Analytics Lab, Department of Technology, San Jose State University, San Jose, CA 95192, USA

ABSTRACT

Bearing failure highly impacts performance and production of manufacturing systems, causes safety incidents, and results in casualties and property loss. According to the current literature, bearing faults cause 30-40% of all failures in induction motors. Therefore, identification of bearing faults, at early stages, is crucial to ensure seamless and reliable operation of induction motors in industrial and manufacturing operations. Faults occur in four components of bearing: inner race, outer race, ball, and cage. Regardless of the component in which fault occurs, it causes changes in vibration signals. Therefore, comparing normal signals with faulty ones is helpful in detecting localized faults in bearings. We use a benchmark publicly-available data set to conduct this analysis. The main challenge in using publicly-available benchmark datasets for fault detection is lack of manual for instruction on analysis experiments on the original data, which leaves researchers with the challenge and opportunity of applying various analytical methods for achieving higher accuracy rates and useful models for fault detection. This study presents a machine learning-based fault detection and classification scheme in induction motors to evaluate the significance and effects of various data preparation and feature extraction methods on accuracy and reliability of fault detection outcomes. The data preparation stage includes discussion of efficient data dimension reduction, and noise eradication, as well as feature extraction methods for induction motor signals. The main methodology is developing a variety of machine learning classifiers for detection and classification of normal bearings versus faulty bearings. Finally, the implications of the methodology and results for early fault diagnosis and enhanced reliability, as well as maintenance planning efforts in manufacturing systems are discussed. This study introduces proper implementation of machine learning models to improve system performance with higher speed and reliability. Furthermore, the methodology and results contribute to planning and undertaking maintenance operation more efficiently. Therefore, the approach, methodology, and results will be beneficial to both researchers and practitioners involved in manufacturing systems reliability analysis and optimization.

Keywords: Machine learning, Monitoring and fault diagnosis, Bearing faults

INTRODUCTION

Condition monitoring of industrial systems is significant in enhancing safety, reliability, performance, and overall quality of industrial processes through providing insights on diagnostics and prognostics for maintenance strategical planning (Hendriks, Dumond and Knox, 2022). Fault detection of rolling bearing in induction motors plays a vital role in efficient condition monitoring (Jallepalli, and Davoudi Kakhki, 2021). Traditionally, model-based approach is used for fault detection in which a complete model of the system is formulated for the purpose of fault diagnosis and prognosis. The main limitations with model-based approach, that make it a challenging and tedious task, include: the variations of the system highly affect its performance, and each component of the system should have a specific formulated model.

A useful alternative for model-based approach for fault diagnosis and prognosis purpose is to use data-driven models (Ziani, Felkaoui and Zegadi, 2017). The data-driven model can be developed using different types of operational data from a system, such as vibration data, thermal image data, etc (Moghadam and Davoudi Kakhki, 2022). A valuable source of data for fault detection studies are publicly available datasets such as the datasets from Case Western Reserve University (CWRU). Due to the different operating conditions under which data was collected, its analysis allows for building various vibration-based analytical models, and therefore, has been widely used in many studies for testing models for efficient fault detection of bearing in induction motors (Smith and Randall, 2015). With the availability of large amounts of operation data, and the flexibility of different modelling algorithms, machine learning (ML) has been used as a popular data-driven approach for various regression and classification problems (Badarinath, Chierichetti and Kakhki, 2021; Choppala *et al.*, 2023). Similarly, ML techniques have been widely used for analysing vibration signals for the purpose of fault detection and classification (Moghadam and Davoudi Kakhki, 2022).

Machine Learning Models for Classification

A brief description of the four Machine Learning (ML) classifiers used in the study follows (Smith and Randall, 2015; Davoudi Kakhki, Freeman and Mosher, 2019; Kakhki, Freeman and Mosher, 2019). ML techniques have been widely used for analysing vibration signals for the purpose of fault detection and classification. The reason for using ML models for developing data-driven models is their ability to generalize on new data and the flexibility in changing the structure of the models and tuning them for obtaining higher accuracy values.

- Support Vector Machines (SVM): SVM models are used for classification problems for outputs that are dichotomous or multi-class. SVM algorithm allow for working with non-linearly separable data by creating a linear separation through converting data from an input space to a feature space, using a variety of kernel function.
- Logistic Regression (LR): LR is a powerful statistical method for analysing multi-class output variables, based on a logit model. LR is able of calculating the probability of the occurrence of an outcome of interest based on a linear combination of the input variables.

- Naïve Bayes (NB): NB algorithm is part of the Bayesian classifiers that implements conditional probabilities for classifying an output variable based on independent input variables.
- Adaptive Boosting Decision Trees (AdaBoost): This decision tree model is an iterative algorithm which is constructed based on a linear combination of input variables, in form of trees, for classification of a target categorical variable.

Objective of the Study

The purpose of this study is to evaluate the performance of different ML classification models for fault detection of rolling bearing in induction motors. We use four different machine learning models for this purpose including: support vector machines, logistic regression, Naïve Bayes, and adaptive boosting decision trees. Furthermore, we evaluate the roll of data preparation through dimension reduction in improving the performance of the machine learning classifiers in distinguishing multi-levels of faults in rolling bearing element of the induction motor based on a benchmark dataset.

DATA SUMMARY AND TRANSFORMATION

We used a benchmark dataset for this research, which is publicly available through Case Western Reserve University (CWRU). The vibration data was generated and collected through two accelerometers. We used a subset of the data, generated and collected for a motor load of 1 horsepower and motor speed of 1772 rpm for sampling frequencies of 48kHz. The data was collected on drive end of bearing with single point fault in bearing inner race, bearing outer race, and bearing ball with three different sizes (0.007, 0.014, 0.021-inch diameter).

The type of input variables and data preparation method has a big impact on the performance of ML models. In the current work, we used a sampling gap of 2048 readings with an overlap of 230 points to segment the dataset. This results in 230 sample points per evolution for each bearing fault type with a length of 2048. The faults are label by the size and location. We have three locations and three sizes for the faults plus the normal status data. This was used to create a target variable that has 10 levels; 9 indicating localized faults with different sizes and one normal label. This segmented dataset is used as the first dataset (raw data) for developing data-driven models later.

In the next step, to reduce the dimension of the data, the already segmented data was transformed based on nine statistical time domain features for each row. In other words, for each row of the segmented data, we derived nine various statistics including: *minimum*, *average*, *standard deviation*, *root mean square*, *skewness*, *kurtosis*, *crest factor*, *form factor*, and *maximum* values. This additional step significantly reduces the dimension of the data to only nine input for classification of the target variable. This reduced-dimension data is used as the second dataset (time domain features) for developing ML models later.

ML MODEL DEVELOPMENT

Upon preparation of both datasets, one segmented data, and one transformed data, we split the whole data based on a 70–30 split ratio for training and testing. Two groups of ML models were developed on the training data: group 1 models were developed using the raw segmented data; and group 2 models were developed using the extracted statistical values as inputs. The output for both groups is the 10 classes of localized faults with different sizes and the normal faults. The 30% test data was then used to evaluate the performance of already constructed ML classifiers for detecting and labelling a raw of data as either faulty or normal, using the previously mentioned fault labels. The overall research methodology used in this study is shown in Figure 1.

In order to compare and assess the performance of ML classifiers, and to evaluate the role of data preparation and dimension reduction on those, we used two performance measurement metrics: overall accuracy and F-score. Both are calculated based on the multi-level confusion matrix that is generated as the results of the classification for both correct and incorrect instances. For example, if a fault in bearing inner race with size 0.007 is detected and classifies as exactly the same, it is labelled as a correct instance, or else it is considered incorrect. The number of accurate instances using ML predictive models compared to the observations in original data, is used for developing a confusion matrix and calculating overall model accuracy and the F-score metrics. Overall accuracy values are between 0-100% and F-score measure between 0-1, the closer values to 1, the stronger is the predictive model in disfiguring among various levels of the target variable.

Figure 1: Research methodology used in this study.

RESULTS

The results of the analysis are presented in Table 1. When developing ML models on raw segmented data, the overall accuracy is variating between 48% to 54%, with LR and AdaBoost models having the lowest and highest accuracy, respectively. Looking at the F-score, except for the AdaBoost classifier with F-score as low as 0.513, the other three ML models have a higher F-score between 0.585 and 0.610. When developing the ML models on the transformed data, where time domain features were used as input, the accuracy increased to values between 91% to 96.4% across all models. The same increase was observed in the F-score measure, with all ML models having values over 0.90, and a maximum value of 0.957 for the AdaBoost model.

The results presented in Table 1 show that the data transformation method that was used for reducing the dimension of the training dataset is highly

Table 1. ML model performance on raw data and transformed data.

Model	Accuracy-Raw Data	Accuracy-Time Domain Features	F-Score-Raw Data	F-Score-Time Domain Features
SVM	0.553	0.964	0.601	0.936
LR	0.484	0.948	0.585	0.906
NB	0.531	0.914	0.61	0.952
AdaBoost	0.542	0.976	0.513	0.957

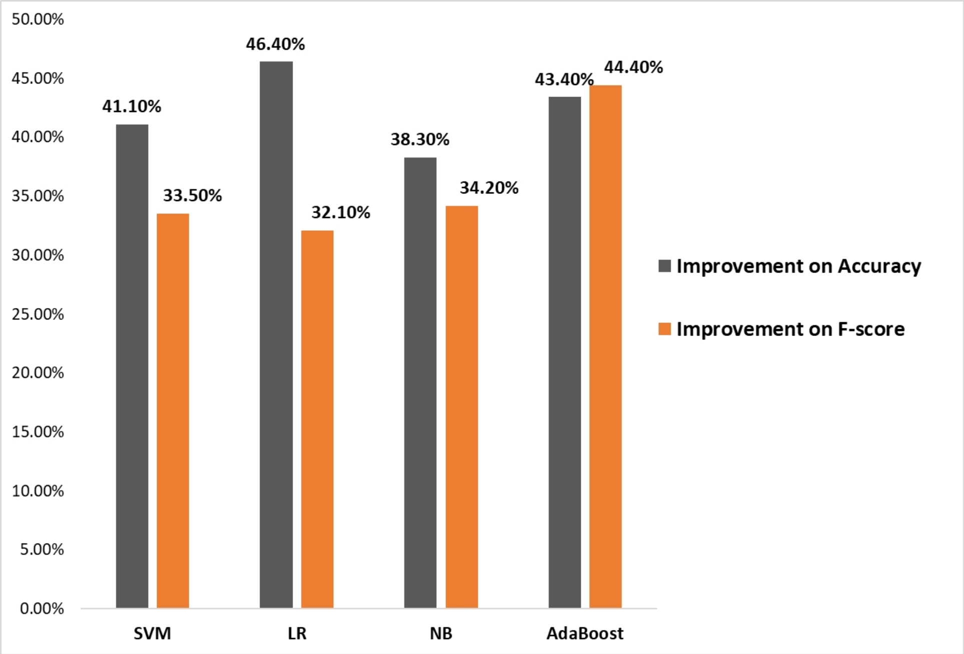

Figure 2: Improvement in ML model performance when using transformed data.

effective in improving the model predictive power and outcomes. This emphasizes the role of data preparation approaches in enhancing the performance of data-driven models for creating more useful and reliable results. As presented in Figure 2, the data transformation approach, using time domain features as input, has increased the overall accuracy and F-score for an average of 42% and 0.36, respectively.

APPLICATION IN IMPROVING MANUFACTURING PERFORMANCE

The availability of big data in manufacturing sector plus the advancement for storing, processing and analysing data provides valuable opportunities for manufacturing intelligence and improving smart manufacturing (Tao *et al.*, 2018). Developing efficient and reliable data-driven models that are capable of identifying and classifying faults in a manufacturing system allows for real-time analysis and reliable prediction of future downtimes in the overall system. Therefore, maintenance planning strategies can be developed based on the data-driven module incorporated in the analytics of a smart manufacturing framework.

Furthermore, the use of data-driven models for fault detection eradicates the need for crating model-based analytical approach for a system which is limited to the system, not generalizable to other systems, and needs to be redone for various components of the system. Data-driven models, such as machine learning, are generalizable and can be tested on new data from the same system, regardless of the system model formula. The application of data-driven models in creating smart ground for manufacturing allows for enhancement in productivity and efficiency of manufacturing sector performance. Using real-time data analytics for fault and anomaly detection is beneficial to the overall performance of the system by creating opportunity for enhanced safety, efficiency, real-time control, remote diagnosis and prognosis, and maintenance (Bauernhansl *et al.*, 2016).

CONCLUSION

Main challenge for modelling CWRU data and consequently various results is that original data do not have a manual for instruction on classification experiments leaves researchers with the challenge of selection of feature extraction methods for achieving higher accuracy rates and useful models. In this study, we analysed the effects of data transformation and dimension reduction on improving the performance and reliability of machine learning models for classification of localized faults with various sizes in rolling bearing element of induction motor based on a benchmark dataset. When comparing the performance of the same machine learning models on segmented raw data with models developed on time domain feature-extracted data, the latter showed considerably higher performance with all models having an accuracy of over 91% and a F-score of over 0.90. The results gained in this study emphasize the significance of data preparation methods in overall performance of data-driven models. The accuracy and F-score values are compatible with or better than similar studies. The further direction of this research is applying other methods for data dimension reduction and feature extraction methods to develop machine learning and deep learning models for intelligent fault detection and classification based on vibration data.

ACKNOWLEDGMENT

This publication was supported by the Professional Development Grant for research dissemination from the College of Engineering at San Jose State University.

REFERENCES

Badarinath, P. V., Chierichetti, M. and Kakhki, F. D. (202a1) 'A machine learning approach as a surrogate for a finite element analysis: Status of research and application to one dimensional systems', *Sensors*, 21(5), pp. 1–18. doi: 10.3390/s21051654.

Bauernhansl, T. *et al.* (2016) 'CIRP Annals - Manufacturing Technology Cyber-physical systems in manufacturing', 65, pp. 621–641. doi: 10.1016/j.cirp.2016.06.005.

Choppala, S. *et al.* (2023) 'Optimal Sensor Location and Stress Prediction on a Plate using Machine Learning', in *AIAA SCITECH 2023 Forum*. the American Institute of Aeronautics and Astronautics, Inc., pp. 1–12. doi: https://doi.org/10.2514/6.2023-0370.

Davoudi Kakhki, F., Freeman, S. A. and Mosher, G. A. (2019) 'Evaluating machine learning performance in predicting injury severity in agribusiness industries', *Safety Science*, 117(July 2018), pp. 257–262. doi: 10.1016/j.ssci.2019.04.026.

Hendriks, J., Dumond, P. and Knox, D. A. (2022) 'Towards better benchmarking using the CWRU bearing fault dataset', *Mechanical Systems and Signal Processing*, 169(February 2021), p. 108732. doi: 10.1016/j.ymssp.2021.108732.

Jallepalli, D., Davoudi Kakhki, F. (2021). Data-Driven Fault Classification Using Support Vector Machines. In: Russo, D., Ahram, T., Karwowski, W., Di Bucchianico, G., Taiar, R. (eds) Intelligent Human Systems Integration 2021. IHSI 2021. Advances in Intelligent Systems and Computing, vol. 1322. Springer, Cham. https://doi.org/10.1007/978-3-030-68017-6_47

Kakhki, F. D., Freeman, S. A. and Mosher, G. A. (2019) 'Use of logistic regression to identify factors influencing the post-incident state of occupational injuries in agribusiness operations', *Applied Sciences (Switzerland)*, 9(17), p. 3449. doi: 10.3390/app9173449.

Moghadam, A. and Davoudi Kakhki, F. (2022) 'Comparative Study of Decision Tree Models for Bearing Fault Detection and Classification', in *Intelligent Human Systems Integration (IHSI 2022) Integrating People and Intelligent Systems*. doi: 10.54941/ahfe100968.

Smith, W. A. and Randall, R. B. (2015) 'Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study', *Mechanical Systems and Signal Processing*, 64–65, pp. 100–131. doi: 10.1016/j.ymssp.2015.04.021.

Tao, F. *et al.* (2018) 'Data-driven smart manufacturing', *Journal of Manufacturing Systems*, 48, pp. 157–169. doi: 10.1016/j.jmsy.2018.01.006.

Ziani, R., Felkaoui, A., Zegadi, R.: Bearing fault diagnosis using multiclass support vector machines with binary particle swarm optimization and regularized fisher's criterion. J. Intell. Manuf. (2017). https://doi.org/10.1007/s10845-014-0987-3