

Explainable machine learning for motor fault diagnosis

Yuming Wang

Department of Electrical and Computer Engineering
University of Kentucky
Lexington, USA
yuming.wang@uky.edu

Peng Wang

Department of Electrical and Computer Engineering &
Department of Mechanical and Aerospace Engineering
University of Kentucky
Lexington, USA
edward.wang@uky.edu

Abstract— Industrial motors have been widely used in various fields such as power generation, mining, and manufacturing. Various motor faults and time-consuming motor maintenance processes will lead to serious economic losses in this context. Different sensing technologies, including acceleration, acoustic, and current sensing can be useful in motor condition monitoring, defect detection, and diagnosis. Regarding sensing data analytics, Machine Learning (ML) and Deep Learning (DL) techniques have been increasingly investigated, because of their promising capabilities in complex data characterization and pattern recognition. However, the explainability of ML and DL models and their decision-making remains a challenge, because of their black-box modeling by nature. Shapley Additive Explanations (SHAP), as a game theoretic approach, provides a way to explain ML and DL modeling results, by allocating credits (known as SHAP values) through local connections to quantify the contributions of input features to model outputs. In this paper, three commonly seen ML techniques, including Support Vector Machine (SVM), Random Forest (RF), and Neural Network (NN) are investigated for vibration-based motor fault diagnosis. Corresponding SHAP explanation methods are applied to the three ML techniques to discover the most important vibration features in detecting motor conditions and differentiating faults. Explanation results from the three ML techniques demonstrate great consensus: average vibration frequency contributes most to motor fault diagnosis. This explanation conclusion matches the physical understanding that fault occurrences would bring in additional frequency components to the spectrum. Improving the physical explainability of ML and DL techniques would significantly improve their credibility and generalizability.

Keywords—*Explainable Machine Learning, Fault Diagnosis, Neural Network, Shapley Additive Explanations*

I. INTRODUCTION

Industrial motors have been widely used in different industrial fields across the globe. An industrial motor mainly consists of several mechanical components, such as a stator, rotor, stator windings, bearing, and terminal box. These mechanical components are prone to failure. The most common motor failures include bowed rotor, broken rotor, unbalanced rotor, and faulted bearing. These failures generally lead to abnormal motor speed, abnormal noise, and vibration, which would hugely influence motor efficiency. That's why motor condition monitoring and fault diagnosis is vital to the whole industry.

Many motor fault diagnosis methods (more related to motor sensing) have been established, such as motor oil analysis, temperature analysis, and vibration analysis. These techniques may or may not be practical. Oil analysis requires sophisticated and expensive instruments. Temperature measurement, because of the low sampling rate of thermometers, has been shown not valuable in capturing motor dynamic behaviors. Vibration analysis has been the

focus of many studies. Vibration sensors are cost-effective, and vibration signals contain lots of crucial information about operating motors. In the last century, some early-stage works discovered the fundamental mechanism of vibration generation (mainly caused by electromagnetic force) in motors and tried to analyze vibration signals for motor fault diagnosis [1-8]. But the accuracy and efficiency of these signal processing approaches were poor. Researchers later investigated more advanced signal processing techniques like Short-Time Fourier Transform (STFT) and Wavelet in vibration data analysis, which increased the performance [9]. However, these techniques are still limited in efficient and automated fault detection and diagnosis.

Machine Learning (ML) techniques have been focused on recently to solve motor fault diagnosis problems because of the explosive growth of computing power. At the beginning of this century, some researchers proposed that some machine learning techniques like Fuzzy Logic (FL), Neural Networks (NN) can be used along with vibration analysis [10]. Saud Altaf et al. discussed using a NN model to diagnose motor Broken Rotor Bars (BRB) faults under two-level load torque [11]. This work was later expanded to a new distributed NN model for detecting BRB fault and air eccentricity [12]. Wavelet Neural Network (WNN) was developed to make classification independent from different load levels, and time-frequency representation (TFR) was used to transform vibration data to low dimension array for performance improvement [13]. Kil Chong et al. used STFT for converting time series quasi-steady vibration signal to continuous spectra and trained a NN model with converted data and Levenberg-Marquardt (LM) algorithm to generate extra fault information [14]. As a new variant of NN, 1-D Convolutional Neural Network (CNN) was designed to have the ability to extract features from vibration or current signal, which avoids manual parameter tuning [15-16]. Recurrent Neural Network (RNN) was developed as an important variant of NN. Bambang et al. classified motor bearing faults using a custom RNN model that is robust to environmental changes because of its recurrent connections [17]. Long Short-Term Memory (LSTM) was a powerful tool over RNN for fault diagnosis because it can remember long-term memory, which made it more adaptive [18]. Support Vector Machine (SVM) was another popular machine-learning technique for motor fault classification. Lane et al. discussed using multiple SVM classifiers to diagnose vibration signals that were processed by Fast Fourier Transform (FFT) [19]. A Random Forest (RF) model was developed with a simpler structure and excellent performance for vibration data analysis [20]. Tapan et al. did a study that compare some unsupervised learning algorithms like k-means clustering, hierarchical clustering, and Expectation-maximization (EM) clustering. These algorithms don't require any labels on data, and EM clustering had the overall best performance among them [21].

Most ML models have been called “black box” models because the causal reasoning underlying the model and modeling results cannot be physically explained. ML models make decisions without any interpretation, which reduces their credibility, applicability, and generalizability. Surrogate methods can explain a machine learning model by constructing new models. Scott et al. created an outstanding surrogate method called Shapley Additive Explanations (SHAP) to explain machine learning models by calculating the SHAP value for each feature. SHAP value shows how features contribute to the model results [22].

Motivated by studies on the recent SHAP model explanation method, this paper tries to apply SHAP explanations to three commonly applied ML techniques, including SVM, RF, and NN in the context of motor fault diagnosis, and compare the explanation results of these three techniques to identify the most valuable features of vibration data in detecting the motor fault and differentiating fault types.

II. EXPLAINABLE MACHINE LEARNING

Three different ML techniques are reviewed and compared in II. A. Corresponding SHAP explanation methods are introduced and explained in II. B.

A. Machine Learning Techniques

In general, ML techniques process input data X and output predictions $\hat{y}$ upon different input-output functions.

1) Neural Network (NN)

The NN model was inspired by the human brain, and it usually belongs to the supervised learning category, which means it needs labeled data for training. A general NN consists of three parts, an input layer, one or multiple hidden layers, and an output layer. The input layer is where data is fed in, and the output layer is where models return the results. A simple three-layer NN can be expressed as

$$\hat{y} = f\left(\sum_i w_i * g\left(\sum_j x_j * w_j + b_1\right) + b_2\right) \quad (1)$$

where $\hat{y}$ is the output, x_j is the input, w_i, w_j are weights and b_1, b_2 biases. f and g are the activation functions which is used to add nonlinearity to the network. The most common activation function is ReLU function, which is expressed as $f(x) = \max(0, x)$.

2) Support Vector Machine (SVM)

The essence of SVM is more like a road with a middle line that separates cars on the left side and right side. SVMs use hyperplanes to separate different classes instead. The distance between classes is called the margin of SVM. Thus, the optimization process of SVMs is to find the best position of each hyperplane and maximize the margin. Lagrange multipliers are used to solve the optimization process, and the Kernel trick is used to map the input to a higher dimension for non-linear classification.

3) Random Forest

RF is a useful machine learning tool based on the decision tree classifier, and it can address the overfitting issue that the decision tree has. Bootstrapping is used for making subsets of data and multiple trees are raised based on each subset. Each decision trees use the traditional Gini index to get its own results. The ultimate result of RF will be decided by using majority voting.

B. SHAP-based ML Model Explanation

SHAP explainers are built after ML models. It can explain an ML model by generating SHAP values to quantify the contributions of individual input features to the model outputs. Then, the model explanation results can be cross-checked with physical/empirical domain knowledge to make ML decision-making physically explainable. SHAP explanation of a general ML model can be expressed as [23]:

$$y = f(x_1, x_2, \dots) = \phi_0 + \sum_{i=1}^M \phi_i \quad (2)$$

where M is the number of input features in feature set $\{x_i, i = 1:M\}$ and y is the output. ϕ_0 is the average prediction and ϕ_i is the SHAP value of the i^{th} feature.

SHAP explanations stem from game theory. SHAP values are generated considering the interaction between features in affecting the modeling results. In other words, subsets of features are exhausted, and the marginal contribution of a specific feature is obtained from averaging the contributions of subset features before and after removing that specific feature. The marginal contribution of a feature is also called a SHAP value, which is generally expressed as

$$\phi_i(f, x) = \sum_{z' \subseteq x'} \frac{|z'|! (M - |z'| - 1)!}{M!} [f_x(z') - f_x(z' \setminus i)] \quad (3)$$

where i is the serial number of features and f is the model that needs to be explained. z' are the subsets that are recurrently calculated and x' is simplified data input. M is the total number of features.

The above SHAP value generation method applies only to linear models. For nonlinear modeling techniques, more advanced SHAP methods are needed, and different ML techniques require different types of SHAP explainers. For NN models, the SHAP values need to be calculated through the DeepSHAP explainer which is based on the DeepLIFT approximation method [24]. The average prediction ϕ_0 in DeepLIFT is the reference point $y^0 = f(x_0^1, x_0^2, \dots)$. ϕ_i is represented by $m\Delta x_i \Delta y_i$. It means that DeepLIFT explains the difference in output from the reference point in the field of the difference in input from the reference point. For a simple NN with one hidden layer, the SHAP value for the i^{th} feature can be written as:

$$m\Delta x_i \Delta y = \sum_j m\Delta x_i \Delta h_j m\Delta h_j \Delta y \quad (4)$$

where h represents the outputs of hidden neurons. Eq. (4) is based on the chain rule for multipliers. To differentiate positive and negative contributions of individual input features SHAP RevealCancel rule [26] can be applied. If the output of a hidden neuron is $h = f(x)$, the SHAP values can be evaluated in terms of positive and negative output differences Δh^+ and Δh^- through the RevealCancel rule:

$$\begin{aligned} \Delta h^+ &= \frac{1}{2} \left(f(x_i^0 + \Delta x_i^+) - f(x_i^0) \right) \\ &+ \frac{1}{2} \left(f(x_i^0 + \Delta x_i^+ + \Delta x_i^-) - f(x_i^0 + \Delta x_i^-) \right) \end{aligned} \quad (5)$$

$$\begin{aligned} \Delta h^- &= \frac{1}{2} \left(f(x_i^0 + \Delta x_i^-) - f(x_i^0) \right) \\ &+ \frac{1}{2} \left(f(x_i^0 + \Delta x_i^- + \Delta x_i^+) - f(x_i^0 + \Delta x_i^+) \right) \end{aligned} \quad (6)$$

The SHAP values then can be approximated by calculating the positive and negative contribution of input features to the positive and negative differences in the output:

$$m\Delta x_i^+ \Delta h^+ = \frac{\Delta h^+}{\Delta x_i^+}; \quad m\Delta x_i^- \Delta h^- = \frac{\Delta h^-}{\Delta x_i^-} \quad (7)$$

KernelSHAP explainer [25] and TreeSHAP explainer [26] are used for SVM models and random forest models respectively. They use different methods to explain models but the basic idea behind them is similar.

III. EXPERIMENTAL EVALUATION

The experimental study is carried out to evaluate the proposed SHAP explanation methods for the three ML models in the context of motor fault diagnosis. This section introduces details related to the experimental setup, vibration signal collection, ML model construction, and training.

A. Experimental Setup

This study focuses on diagnosing four types of common motor faults, which are bowed rotor, broken rotor, unbalanced rotor, and faulted bearing. Thus, datasets are obtained from five different motors (including a normal motor). The experimental setup is shown in Fig. 1. These motors were mounted on the experiment platform one by one to obtain vibration data under motor operation. Two vibrations sensors (Horizontal and vertical, normal to each other) are applied to capture electromagnetic forces in different directions. 7 different rotating speeds from 1200rpm to 3000rpm have been tested, generating 7 sets of vibration data from each motor. A consistent load was set with all motor speeds. The load is very light compared to the weight and power of motors so the influence of load on motor torque can be ignored.

B. Data sampling and Feature extraction

Statistics features are extracted from both the time domain and frequency domain of motor vibration signals. Feature extraction helps reduce the dimensionality of data processing, and these features have been demonstrated effective in characterizing motor operating conditions [27]. In this paper, 16 features are selected and extracted from one channel of the vibration data, including 9 time-domain features and 7 frequency-domain features, as listed in Table I.

The sampling rate of the vibrations sensor is 10kHz. From individual motors running under a specific speed, 60 seconds of data (i.e., 600,000 data points from each vibration channel) have been collected. Sliding windows are applied to pre-

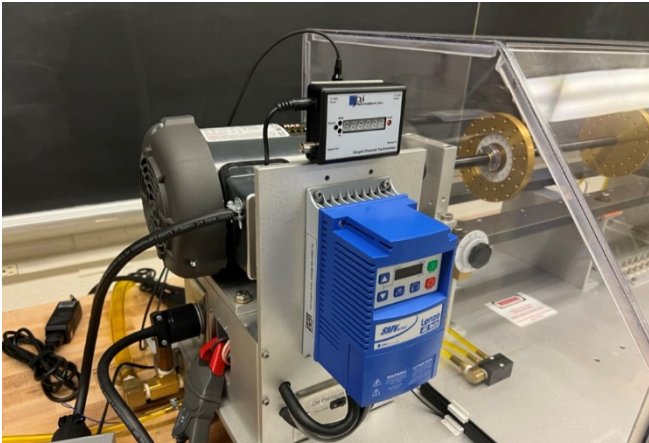


Fig. 1. Experiment gearbox setup

process the raw vibration data and generate features. In this case, a window is set across 1000 data points, and there is no overlapping between subsequent windows. In other words, a sample would be created to contain 16 features extracted from 1000 horizontal vibration data points and 16 features from vertical vibration data. In total, Thus 21,540 samples are generated for the 5 motor fault types and 7 rotating speeds, and each sample contains 32 features.

C. Model Construction and Training

The NN model used in this study has 5 layers. The input layer has 32 neurons corresponding to 32 features, and three hidden layers have 128, 64, and 32 neurons respectively. The output layer has 5 layers corresponding to 5 types of motor faults. The learning rate is set to 0.01, and the Adam algorithm is used for the optimizer. The construction of SVM and RF models is similar. To improve the training robustness, 5-fold cross-validation is applied during the training process, which split the dataset into 5 pieces and uses each piece as a test set and the rest as a training set so that the model will be trained 5 times. Each time will return a result and the final result will be the mean value of all results. Three SHAP explainers are applied to interpret the classification results of the three ML models, once they are done with training.

Training of the three ML models is evaluated in 5 different scenarios, corresponding to 5 sets of training data. In the first round of experiment, training data proportionally come from all rotating speeds. In the second round of the experiment, only slow motor speed data from 1200rpm to 2100rpm are used for training, which would tell whether the motor speed influences the performance of models. In the third round of the experiment, only selected interval motor speed data, which are 1200rpm, 1800rpm, 2400rpm, and 3000rpm are used for training, which would also further observe the influence of motor speed on the performance of models. In the fourth round of the experiment, only the 10 most important features selected from the SHAP explanation results of the first round of the experiment are used for model training. In the last round of experiments, only 16 features extracted from the horizontal vibration signal are used for model training.

IV. RESULTS AND DISCUSSION

The model accuracy comparison for five rounds of experiments is shown in Fig. 2. The results indicate that the first standard experiment (i.e., training data come from all tested rotating speeds) produced the best results, with all three different models achieving 100% accuracy. The accuracy of the three models is very low in the second round of experiments that only uses data from slow motor speeds as

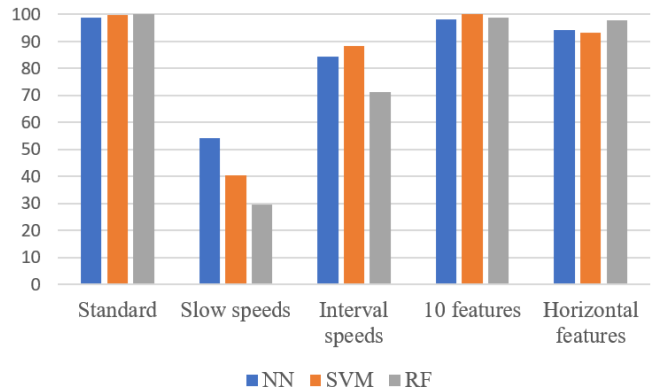


Fig. 2. Model accuracy comparison

TABLE I. FEATURE EXTRACTION EQUATIONS

Time-domain features		Frequency-domain features	
Feature	Equation	Feature	Equation
Mean (M)	$c_1 = \frac{1}{n} \sum_{i=1}^n x_i$	Mean frequency (MF)	$p_1 = \sqrt{\frac{\sum_{i=1}^K f_i^2 \cdot S(f_i)}{\sum_{i=1}^K S(f_i)}}$
Root mean square (RMS)	$c_2 = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$	Average frequency that is wave shape of signal crosses the mean of time-domain signal (AF)	$p_2 = \sqrt{\frac{\sum_{i=1}^K f_i^4 \cdot S(f_i)}{\sum_{i=1}^K f_i^2 \cdot S(f_i)}}$
Standard deviation (SD)	$c_3 = \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}}$	Stabilization factor of wave shape (SFW)	$p_3 = \frac{\sum_{i=1}^K f_i^2 \cdot S(f_i)}{\sqrt{\sum_{i=1}^K S(f_i) \sum_{i=1}^K f_i^4 \cdot S(f_i)}}$
Skewness (SK)	$c_4 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - c_1}{c_2} \right)^3$	Coefficient of variability (CV)	$p_4 = \frac{\sigma}{\bar{f}}$
Kurtosis (K)	$c_5 = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - c_1}{c_2} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)} \right\}$	Frequency-domain skewness (FSK)	$p_5 = \frac{\sum_{i=1}^K (f_i - \bar{f})^3 \cdot S(f_i)}{\sigma^3 K}$
Crest factor (CF)	$c_6 = \frac{\max(x)}{c_2}$	Frequency-domain kurtosis (FK)	$p_6 = \frac{\sum_{i=1}^K (f_i - \bar{f})^4 \cdot S(f_i)}{\sigma^4 K}$
Latitude factor (LF)	$c_7 = \frac{\max(x)}{((1/n) \sum_{i=1}^n x ^2)^{\frac{1}{2}}}$	Root-mean-square ratio (RMSR)	$p_7 = \frac{\sum_{i=1}^K \sqrt{(f_i - \bar{f})} \cdot S(f_i)}{\sqrt{\sigma} K}$
Shape factor (SF)	$c_8 = \frac{c_2}{((1/n) \sum_{i=1}^n x_i)}$	where $S(f_i)$ is power spectrum density obtained by the Welch method.	
Impulse factor (IF)	$c_9 = \frac{\max(x)}{((1/n) \sum_{i=1}^n x_i)}$	$\bar{f} = \frac{\sum_{i=1}^K f_i \cdot S(f_i)}{\sum_{i=1}^K S(f_i)}$ $\sigma = \sqrt{\frac{\sum_{i=1}^K (f_i - \bar{f})^2 \cdot S(f_i)}{K}}$	

where x_i is sampled vibration signal for $i = 1, 2, \dots, n$.

training data, which preliminarily indicates that the training data coverage is very important for model performance. The main reason for such poor results is that motor operating dynamics under low and high rotating speeds are different, and fault-related patterns learned by the three ML models from low-speed data cannot be well generalized to high-speed data. Training data coverage is also a major concern in affecting the generalizability of ML models. The accuracy of the three models in the third round of experiments that uses interval motor speeds as training data, although not high, is still better than that of the previous round of experiments. By seeing the data from the lowest and highest speed limits, the identified fault-related patterns by the ML models are more generalizable to data from interval speeds.

The fourth round of experiments is based on the 10 most important features that are obtained from the SHAP explanation results of the first round of experiments. This round of results is very close to the first round of experiments. This indicates that the 10 most important features identified

by SHAP explanation contain almost all necessary information out of 32 features in detecting and differentiating motor faults. This also demonstrates the effectiveness of SHAP explanation in identifying critical data features. The fifth round of experiments uses only 16 features extracted from the horizontal signal. The accuracy of the three models is good but lower than that of models trained with the 10 most important features, indicating that the quality of features is more important than the quantity. However, this experiment still shows that the model can be trained well using only the data obtained from one sensor.

The SHAP feature importance plot for three machine learning models in the standard experiment is presented in Fig. 3. (a) ~ (c). The vertical axis of plots shows the names of features, and the suffix (v or h) of each name indicates whether the feature is extracted from the horizontal or vertical sensor data. The horizontal axis represents the corresponding mean SHAP value. The higher the SHAP value, the greater the contribution of this feature to model classification results. Five

different colors represent the contribution of a particular feature to a particular motor fault class during model training. The SHAP value of AF(h) (Average frequency extracted from the horizontal signal) in the three models is dominant, indicating that AF(h) contributes the most among all the features, as it carries information on how signal components vary with motor faulty conditions. Another finding is that the SHAP explanation results of three different ML models have

great consensus: the first two most critical features from the three explanations are the same. This proves the generalizability of SHAP explanations in relative to specific ML models. Fig. 4 are SHAP summary plots that show the decomposition of feature contributions for a specific fault type. The horizontal axis of each plot is the SHAP value rather than the mean SHAP value. Unlike previous plots, these scatter plots consisting of many points, and each point

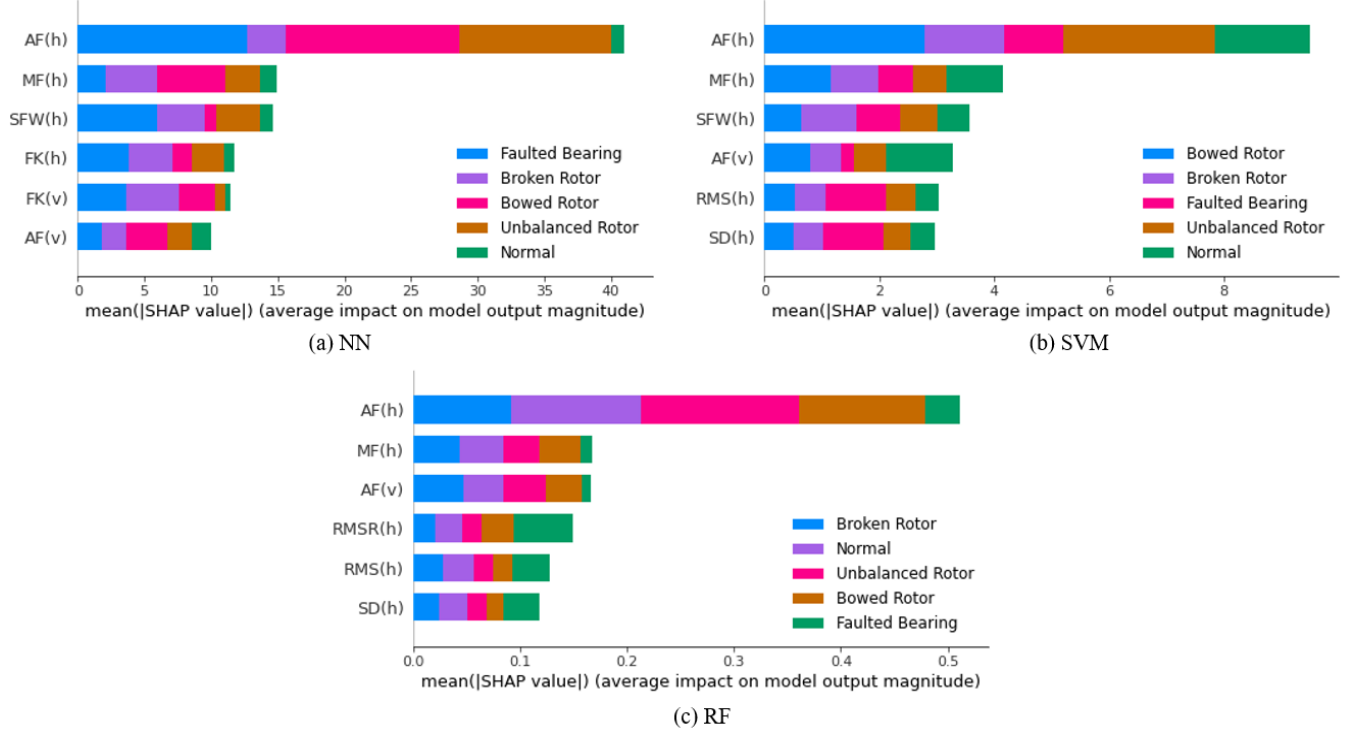


Fig. 3. SHAP model importance plot

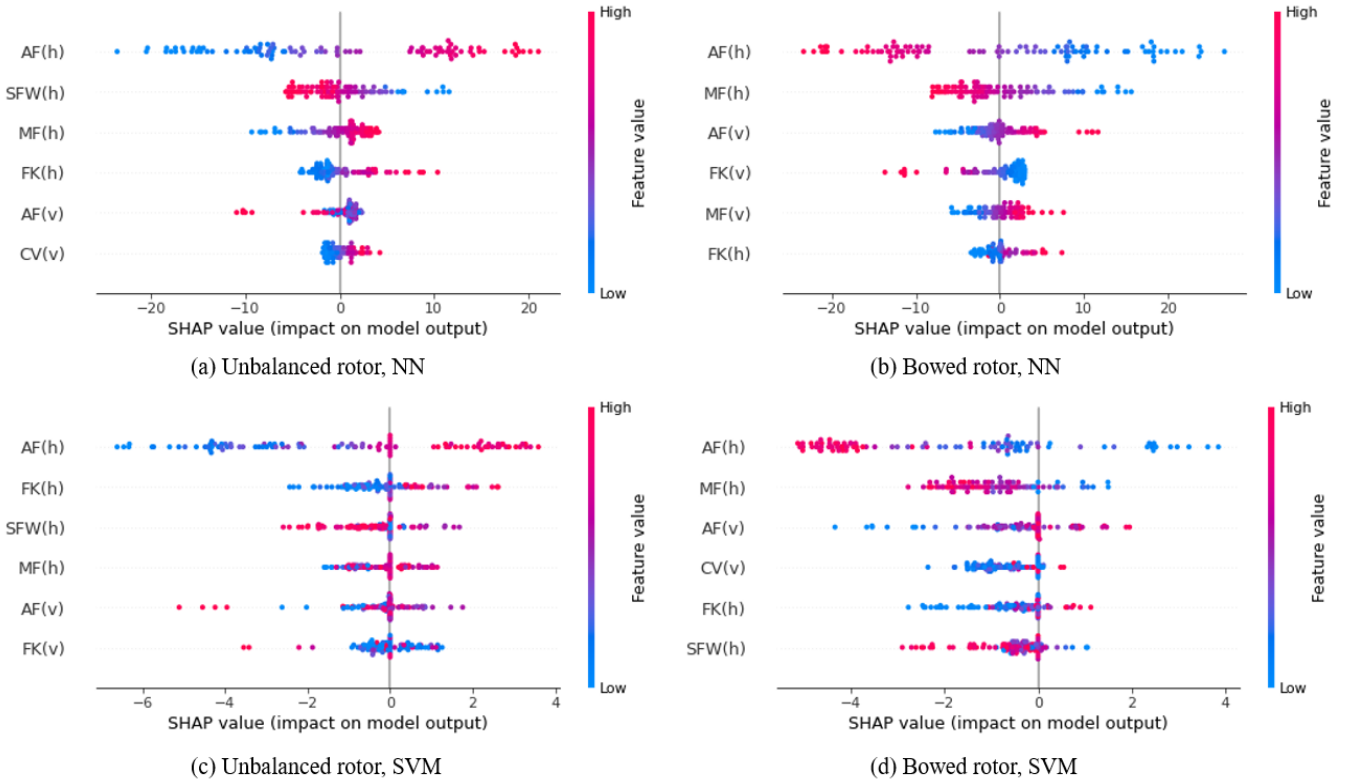


Fig. 4. SHAP summary plot for two types of faults

represents a sample. The color of points represents feature value. Fig. 4. (a) ~ (b) is the summary plots of NN for two different fault types. As mentioned earlier, the feature in the first row is AF because it contributes the most to the model. A high value of AF(h) would make the NN model to be more likely to classify that sample as an unbalanced rotor fault and against other fault types. So that in Fig. 4. (a) high values of AF(h) are explained as positive contributions to unbalanced rotor fault. Similarly, as shown in Fig. 4. (b), a low value of AF(h) would make the model classify the sample as bowed rotor fault; hence, low AF(h) values have a positive contribution to the identification of this specific motor fault type. Fig. 4. (c) ~ Fig. 4. (d) are summary plots of SVM for two different fault types. Similar explanation results in Fig. 4 again prove the generalizability of SHAP explanation of different ML models. SHAP explanation would advance our physical understanding of the motor operation and fault occurrence, e.g., what type of fault leads to what changes in electromechanical force in motor operating and vibration measurement.

V. CONCLUSION

This paper presents a SHAP-based explanation of three commonly used ML techniques (SVM, RF, and NN) in detecting and classifying motor faults. The explanation helps identify the critical vibration sensing features in differentiating fault types, not only reducing the dimension of data analysis but also advancing the physical understanding of motor fault occurrences. The ML models are evaluated in different training scenarios and explained by different SHAP explainers. Some observations are drawn from the analysis:

- All three ML techniques could achieve great fault diagnosis if training data has full coverage.
- Uncomprehensive training data coverage greatly reduces model performance; including data from upper and lower operating limits is necessary.
- Explanation results of the three ML models, although performed by different SHAP explainers, show a great consensus that average vibration frequency is the most critical feature in motor fault diagnosis.
- Including features with the highest contributions achieve comparable performance as including all features.

Future work will explain more types of ML and DL models like CNN and RNN.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation under Grant No. 2015889.

REFERENCES

- [1] J. P. Den Hartog, "Vibration of frames of electrical machines," *Trans. Am. Soc. Mech. Engrs.* 50, pp. 1-6 and 9-11. 1928.
- [2] M. Krongl, *Noise of Electrical Machinery*, CIGRE, Paris, France (in French), 1933.
- [3] L. E. Hildebrand, "Quiet induction motors," *AIEE Transactions*, Vol. 49, pp. 848-856, July 1930.
- [4] H. Jordan, *The Low Noise Electric Motors*, Verlag W. Girardet, Essen, Germany (in German), 1950.
- [5] E. Erdelyi, *Predetermination of the Sound Pressure Levels of Magnetic Noise in Medium Induction Motors*, University of Michigan, Engineering College, Industry Program, 1955.
- [6] P. L. Alger, and E. Erdelyi, "Calculation of the magnetic noise of polyphase induction motors," *The Journal of the Acoustical Society of America*, vol. 28, issue 6, pp.1063-1067, Nov 1956.
- [7] B. Heller, V. Hamata, *Harmonic Field Effects in Induction Machines*. Academia, Prague, 1977. (Elsevier Scientific Publishing Company).
- [8] S. J. Yang, *Low Noise Electrical Motors*, Oxford, Clarendon Press, 1981.
- [9] M. R. W. Group, "Report of large motor reliability survey of industrial and commercial installations, part II," *IEEE Transactions on Industry Applications*, no. 21, pp. 865-872, 1985.
- [10] R. P. Panadero, J. P. Llinares, V. C. Alarcon, and M. P. Sanchez, "Review diagnosis methods of induction electrical machines based on steady state current" in *Proceedings of the 11th Spanish Portuguese Conference on Electrical Engineering*, 2009.
- [11] S. Altaf, S. W. Soomro, and M. S. Mehmood, "Fault Diagnosis and Detection in Industrial Motor Network Environment Using Knowledge-Level Modelling Technique," *Modelling and Simulation in Engineering*, vol. 2017, pp. 1-10, 2017.
- [12] S. Altaf, A. Al-Anbuky, and H. G. Hosseini, "Fault signal propagation in a network of distributed motors," *IEEE 8th International Power Engineering and Optimization Conference*, pp. 59-63, March 2014.
- [13] T. Boukra, A. Lebaroud and G. Clerc, "Neural-Network approaches for classification of induction machine faults using optimal time-frequency representations," *2011 10th International Conference on Environment and Electrical Engineering*, pp. 1-4, 2011.
- [14] Su, Hua, Kil To Chong, and R. Ravi Kumar, "Vibration Signal Analysis for Electrical Fault Detection of Induction Machine Using Neural Networks," *Neural Computing and Applications*, vol. 20, no. 2, pp. 183-194, 2011.
- [15] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-Time Motor Fault Detection by 1-d Convolutional Neural Networks," *IEEE Trans. on Industrial Electronics*, vol. 63, no. 11, pp. 7067-7075, 2016.
- [16] R. Mukhopadhyay, P. S. Panigrahy, G. Misra, and P. Chattopadhyay, "Quasi 1d CNN-Based Fault Diagnosis of Induction Motor Drives," *2018 5th International Conference on Electric Power and Energy Conversion Systems (EPECS)*, 2018.
- [17] B. Purwahyudi, "Recurrent neural network (RNN) based bearing fault classification of induction motor employed in home water pump system," *Journal of electrical engineering and computer sciences*, vol.3(1), June 2018.
- [18] D. Y. Xiao, Y. X. Huang, X. D. Zhang, H. T. Shi, C. L. Liu, and Y. M. Li, "Fault diagnosis of asynchronous motors based on LSTM neural network," *Prognostics and system health management conference. IEEE*, 2018.
- [19] L. Baccarini, V. Silva, B. Menezes, and W. Caminhas, "SVM practical industrial application for mechanical faults diagnostic," *Expert Systems with Applications* 38.6, pp. 6980-6984, 2011.
- [20] R. Patel and V. K. Giri, "Feature selection and classification of mechanical fault of an induction motor using random forest classifier," *Perspectives in Science*, pp. 334-337, 2016.
- [21] T. Mekaroonkamon and S. Wongsu, "A comparative investigation of the robustness of unsupervised clustering techniques for rotating machine fault diagnosis with poorly-separated data," *Eighth International Conference on Advanced Computational Intelligence. IEEE*, 2016.
- [22] Lundberg, M. Scott, and S. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems* vol. 30, 2017.
- [23] E. Štrumbelj and I. Kononenko, "Explaining Prediction Models and Individual Predictions with Feature Contributions," *Knowledge and information systems*, vol. 41(3), pp. 647-665. 2014.
- [24] A. Shrikumar, P. Greenside, and A. Kundaje, "Learning Important Features through Propagating Activation Differences," *International Conference on Machine Learning, ICML*, pp. 4844-4866, 2017.
- [25] K. Aas, M. Jullum, and A. Løland, "Explaining individual predictions when features are dependent: More accurate approximations to Shapley values," *Artificial Intelligence*, vol. 298, 2021.
- [26] M. Scott, "Explainable AI for trees: From local explanations to global understanding," *arXiv preprint arXiv*, 2019.
- [27] Z. Xu, Y. Li, Z. Wang, and J. Xuan, "A novel clustering method combining ART with Yu's norm for fault diagnosis of bearings," *Shock and Vibration*, 2016.