

Article

A Comparative Study between Machine Learning Algorithm and Artificial Intelligence Neural Network in Detecting Minor Bearing Fault of Induction Motors

Shrinathan Esakimuthu Pandarakone ^{1,*}, Yukio Mizuno ¹ and Hisahide Nakamura ²

¹ Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Nagoya 466-8555, Japan; mizuno.yukio@nitech.ac.jp

² Research and Development Division, TOENEC Corporation, Nagoya 457-0819, Japan; hisahide-nakamura@toenec.co.jp

* Correspondence: sarushrinathan@gmail.com; Tel.: +81-80-4842-8090

Received: 30 April 2019; Accepted: 30 May 2019; Published: 1 June 2019



Abstract: Most of the mechanical systems in industries are made to run through induction motors (IM). To maintain the performance of the IM, earlier detection of minor fault and continuous monitoring (CM) are required. Among IM faults, bearing faults are considered as indispensable because of its high probability incidence nature. CM mainly depends upon signal processing and fault detection techniques. In recent decades, various methods have been involved in detecting the bearing fault using machine learning (ML) algorithms. Additionally, the role of artificial intelligence (AI), a growing technology, has also been used in fault diagnosis of IM. Taking the necessity of minor fault detection and the detailed study about the role of ML and AI to detect the bearing fault, the present study is performed. A comprehensive study is conducted by considering various diagnosis methods from ML and AI for detecting a minor bearing fault (hole and scratch). This study helps in understanding the difference between the diagnosis approach and their effectiveness in detecting an IM bearing fault. It is accomplished through FFT (fast Fourier transform) analysis of the load current and the extracted features are used to train the algorithm. The application is extended by comparing the result of ML and AI, and then explaining the specific purpose of use.

Keywords: induction motor; minor bearing fault; frequency spectrum analysis; machine learning algorithm; artificial intelligence

1. Introduction

In recent decades, induction motor (IM) applications have been extended to various fields in industry due to their numerous advantages, such as low cost, less maintenance, simple and robust construction, high efficiency with good reliability in operation than any other motors available. If a fault occurs in an IM and is not identified at the earlier stage, it may lead to unplanned downtime and economical loss to the industry, and even sometimes results in catastrophic effects to the industry [1,2]. Thus, some industries have started performing maintenance to safeguard the equipment by detecting the faults at the earliest. On the other hand, maintenance may lead to a production blockage in the industry by consuming more time for diagnosis. Many studies [3,4] keep on suggesting the importance of condition monitoring (CM) and fault diagnosis of IM, which is more recommended to avoid the production blockage for diagnosing the IM. Generally, CM deals with continuous monitoring of the failure progress of respective equipment, and indicates the significant changes observed above the critical level. The CM in IM can be done by monitoring any of its own parameters, like voltage, current, magnetic flux, etc. The online surveillance of IM progressively cuts down the scheduled maintenance cost and automatically increases the production rate by reducing the diagnosis time.

According to IEEE survey [5], bearing failure (44%) is most frequent in IM rather than any other faults, and thus research based on bearing has been taken for the present study. Frequently observed bearing faults are pitting, scratch, flaking of the surface, and abrasion [6]. These faults are mainly due to misalignment, environmental condition, mishandling, voltage fluctuation, contamination, corrosion, insufficient lubrication, overloading, and bearing current [7]. If a defective motor continues to operate, it may exhibit symptoms, like temperature rise, vibration, changes in the harmonic components of current, and variations observed in the electromagnetic field, which gradually reduce its function and result in motor failure [8].

Among the various diagnostic methods, vibration analysis [9,10], motor current signature analysis (MCSA) [11], acoustic emission (AE) [12], and stray flux monitoring [13] are widely practiced. The vibration analysis is a traditional method, but parameters like cost, access, and placement of vibration sensors make it difficult to gain use in practical cases. In the case of MCSA, being a non-invasive technique, it avoids the disadvantage of the vibration method by eliminating direct access and sensor requirements but, again, fails in determining the present state of the bearing condition. Additionally, AE and stray flux methods need to have direct access to the motor. Many studies are attempting to find the ideal and effective technique for diagnosing bearing failure. In addition, non-invasive techniques are preferred because of easy access and low cost. Therefore, it is necessary to develop an effective CM with a non-invasive technique for early identification of the fault.

In recent decades, the application of digital signal processing to detect bearing faults in IMs has been extended [14]. It produces reliable output by interpreting the statistical data. Generally, raw signals collected from the experiment are not enough in identifying the fault [15], therefore, fault features need to be extracted from any of the three methods: time domain [16], frequency domain [17], or time-frequency domain analysis [18]. Many studies have performed bearing failure analysis considering a hole [19] as a faulty factor. On the other hand, scratches in the bearing have a high probability of occurring and should be considered for performing bearing failure analysis. The fault should be detected at the early stage, which is during the minor stage, so that sudden breakdown in industries may be reduced. Thus, studies based on scratch and minor fault detection need greater attention.

Bearing fault diagnosis relies on ML and AI to have a better CM system and, thus, it may increase the reliability of the motor. Generally, ML has been used in telecommunication, medical diagnosis, market analysis, weather prediction, image sensing, etc., where its application gets globalized. In ML, a wide variety of algorithms are available and, based on the application, the algorithms are selected. The different styles of ML algorithms are supervised, unsupervised, and semi-supervised. Supervised learning is about inputting the known data and evaluation is performed based on the probability. In contrast, in unsupervised learning the input data are not known, and the algorithm is designed to detect the structures of the data. In the case of semi-supervised learning, input data are a mixture of labeled and unlabeled values as an input function and the evaluation is performed. The application of ML for the purpose of diagnosis and prediction in other fields are studied [20–22]. Additionally, AI has been a growing technology in all fields which are competitive to ML algorithms. In recent days, the application of AI to fault detection has increased; artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) [23] are mostly investigated. The application of ML and AI is a challenging task, but if the outcome of the approach satisfies the industrial requirement, then data processing through ML and AI will be the best choice in the fault detection of IM.

The present work aims to develop a non-invasive and inexpensive diagnostic tool for detecting a minor bearing fault in an IM at the early stage by using the various ML approaches and AI. Among the ML, support vector machine (SVM), naive Bayes classifier (NBC) algorithm, k-nearest neighbor (k-NN) algorithm, decision tree (DT), and random forest (RF) are selected. In AI, deep learning (DL) with a convolutional neural network (CNN) architecture is selected and discussed. For making the proposed system to be accepted by industry, the rotating speed is not taken into account in the proposed diagnosis method. The scratch and hole are considered as the faulty factors. Initially, an experiment is conducted at different load conditions and the frequency spectrum of the load current is obtained by fast Fourier

transform (FFT) analysis. The features are extracted and used to train the ML and DL algorithms. Finally, the difference between the ML and DL are evaluated and its application towards motor fault detection is discussed. The concept of the method proposed is illustrated in Figure 1.

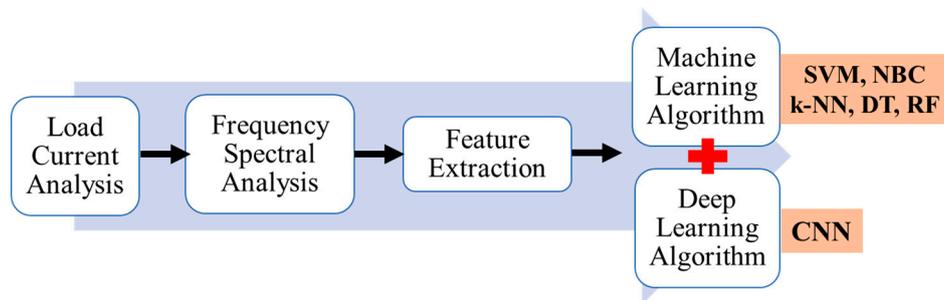


Figure 1. Overview of the proposed method.

2. Experimental Setup and Fault Specification

A specimen of the present study is a three-phase IM (2.2 kW, 200 V, 8.5 A, 1740 min^{-1} , four poles). Figure 2a shows the experimental setup and a powder brake is connected to the IM as a load. The rotating speed of the induction motor can be varied between 1780 min^{-1} and 1765 min^{-1} by changing the load parameters. The current sensors (HIOKI, Type 9696-02), voltage sensors (HIOKI, Type 9666), and a tachometer (ONOSOKKI, Type HT-5500) continuously monitor the load current, line-to-line voltage, and rotating speed, respectively. The developed instrument with seven A/D converters is effectively used to transfer data from these sensors to a desktop computer and shown in Figure 2b. Frequency resolution of the instrument is about 0.76 Hz because the sampling time is designed to be $10 \mu\text{s}$, and the data recording length per channel is 2^{17} . Data transfer is performed in less than 20 s and the data acquisition command for every 30 s. The power supply frequency is 60 Hz.

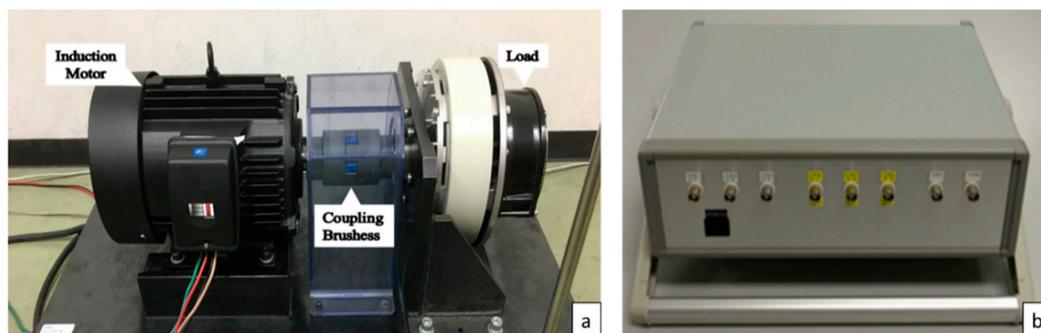


Figure 2. (a) Experimental setup; and (b) measurement equipment.

The collection of bearings with defects from the factories is arduous and consumes time. Alternately, an artificial fault is induced on the outer raceway of the bearing and the diagnosis is performed. Initially, healthy bearing data are taken for reference and the next experiment is carried out after replacing the healthy bearing with a faulty one. The bearing conditions used in the present study are shown in Figure 3. The fault specification details are: the diameter and the depth of a hole are 0.5 mm and 0.5 mm, respectively. A scratch has a size 5 mm in length and 0.5 mm in width and depth. The bearing conditions are indicated as follows: H, healthy; F1, failure with a hole; and F2, failure with a scratch.

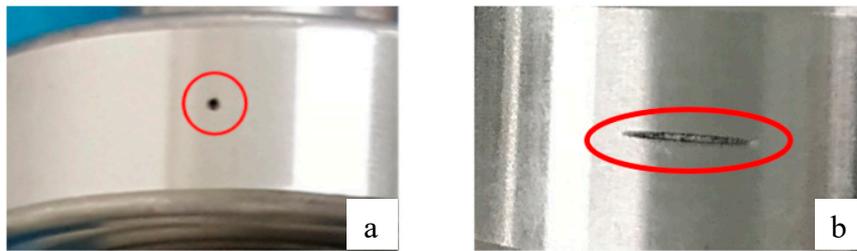


Figure 3. Faulty bearings (a) hole; and (b) scratch.

3. Fast Fourier Transform Analysis

The entire analysis is based on the U-phase load current and the voltages are recorded for the further investigation as a reference.

3.1. Spectral Analysis

For three different bearing conditions, frequency spectra of the load current are obtained by FFT for the rotating speed of 1780, 1775, 1775, and 1765 min^{-1} . For a clear study, frequency spectra in the case of 1765 min^{-1} are plotted for the classes F1F2 and HF1F2 as shown in Figure 4. The normalization of the amplitude is performed, and the maximum amplitude is 0 dB. At frequencies of 30, 90, 120, 150, and 180 Hz clear amplitude differences between classes are recognized. However, the amplitude difference at frequencies of 30 and 90 Hz are only commonly observed in other rotating speeds. Hence, 30 and 90 Hz frequency components are considered as the main features in characterizing the bearing failure study.

Large differences between the features of the healthy and faulty bearings are observed (H and F1, H and F2). These differences show the possibility of identifying the bearing condition and makes the proposed features suitable to determine the present condition of the bearing.

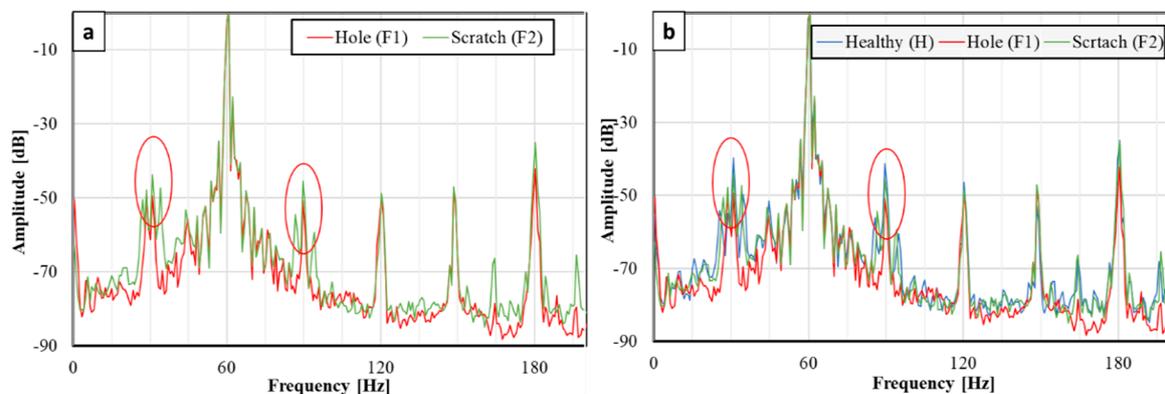


Figure 4. (a) Frequency spectrum of class F1F2 at a rotating speed of 1765 min^{-1} ; and (b) frequency spectrum of class HF1F2 at a rotating speed of 1765 min^{-1} .

3.2. Explanation of Feature Difference at 30 and 90 Hz

The following equation is suggested to explain why the amplitude difference appears at 30 and 90 Hz:

$$F_B = F_L \pm F_R \quad (1)$$

where F_L is the power supply frequency (60 Hz) and F_R is the frequency that depends on the rotating speed. In general, a four-pole IM has a synchronous speed N_s of 1800 min^{-1} . Since the motor revolves 30 times per second, the corresponding frequency becomes 30 Hz. For example, in the case of the rotating speed of 1765 min^{-1} , 30.59 Hz ($60 - 29.41$ Hz), and 89.41 Hz ($60 + 29.41$ Hz) are obtained as F_B

since F_R is 29.41 ($=1765/60$) Hz. Similarly, F_B for other rotating speed is calculated and found to be very close to 30 and 90 Hz, though F_R depends on the rotating speed. The frequency resolution of the measuring equipment is designed as 0.76 Hz regardless of the rotating speed. Thus, it is difficult to distinguish the small difference observed in F_B based on the rotating speed.

4. Characteristics of Extracted Features

To understand the characteristics of the extracted features visually, the features of all bearing conditions are plotted in a two-dimensional plane for each rotating speed and shown in Figure 5. The bearing condition and the rotating speed affect the location of features and shows no overlapping. Thus, detecting the minor fault at the early stage and predicting the present state of the bearing are made possible if the rotating speed is fixed.

However, in the industries, the rotating speed of the motor is not uniform and subjected to change with time and, therefore, the feature distribution is discussed neglecting the rotating speed, and the result is shown in Figure 6. The result of healthy and faulty motors (HF1) are differentiated from each other, but on the other hand, in case of HF2 and F1F2, the overlapping of features is observed which makes the method hard to differentiate the bearing condition. Hence, to match the industrial condition, ML and DL methods are used for diagnosing the bearing failure, which may help to reduce or eliminate the overlapping issues.

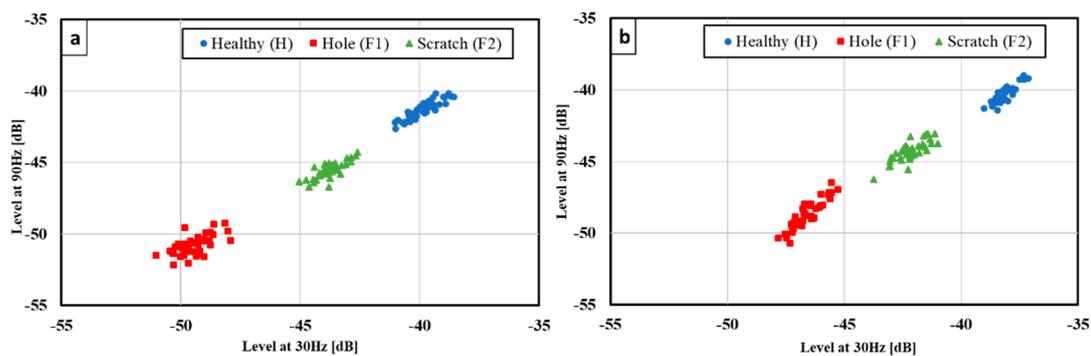


Figure 5. Feature distribution of H-F1-F2 bearing conditions (a) considering rotating speed at 1765 min^{-1} ; (b) considering rotating speed at 1770 min^{-1} .

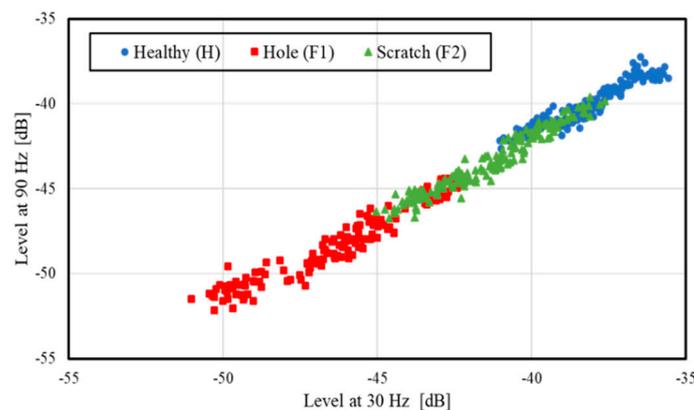


Figure 6. Feature distribution of H-F1-F2 bearing conditions without considering the rotating speeds.

5. Diagnosis Using Machine Learning (ML)

Generally, ML algorithms are classified as supervised and unsupervised algorithms. The supervised algorithms consist of target variables which are to be predicted from a given set of independent variables. These variables are used to generate the function and maps of the input to

obtain the desired output and to achieve the objective. The data are trained to achieve a better accuracy rate. The training process is continued until the model achieves the desired level of accuracy. However, in the case of unsupervised algorithms, there is no target variable and the clustering technique is adopted. The techniques may segment the group and certain levels of diagnosis can be achieved. The role is to identify the current condition of the bearing and from the case study it is known that the supervised machine learning algorithm is suited. The algorithms used in the present study are SVM, NBC, k-NN, DT, and RF, and python is used as a diagnostic tool.

5.1. Support Vector Machine (SVM)

SVM [24] is a well-known pattern recognition algorithm mainly used for classification and regression problems. The basic parameters of SVM are the hyperplane and margin. The hyperplane separates the datasets and performs the classification task and the margin identifies the support vectors of the datasets. SVM performs classification by finding the optimum hyperplane that maximizes the margin width between the classes. The high margin width avoids the overlapping issues between the classes. Generally, the margin is classified into two types: soft and hard margins. Since the present diagnosis deals with the non-linear classification problem, a soft margin is used.

The accuracy of the SVM will be affected by three factors: the threshold function, the cost parameter (C), and the kernel function. Recognition ability can be improved by the threshold function. The cost parameter adjusts the tradeoff between the smooth decision of the boundary condition and classifying the training points. Low bias and high variance can be achieved when a large value is used as the cost parameter. The main role of kernel functions is to map the input at high dimensional features so that non-linear classification is available. Radial basis function (RBF) has been selected in the present study. In RBF kernel, data classification is affected by gamma parameter. The gamma parameter sets the pattern contrast to the cost parameter. The values of the cost and gamma parameters should not be very high because of overfitting, and they should not be very small because of underfitting problems. These parameters can be tuned through programming by selecting the optimum ranges.

The selection of optimal hyperplane is mainly based on the training features. Techniques, such as cross-validation, re-sampling, and grid search, help in selecting the values of the cost and gamma parameters automatically during the diagnosis. Cross-validation in ML helps to train the model using optimal hyper-parameters. Re-sampling is a series of methods used to reconstruct the sample data sets, including training and testing datasets. A grid search is an iterative method of choosing the best parameter value for an ML model. The optimized values of C and γ with the highest accuracy rate are used in the present study and the corresponding hyperplane is constructed. Specifications of the SVM used in the present study are described in [25] and explained in Table 1.

Table 1. SVM specification.

SVM Type	Soft Margin
Kernel	Radial basis kernel function
Gamma parameter	2^{-4}
Cost parameter	2^{-1}
Number of support vectors	8
Number of classes	3
Threshold in training	0
Threshold in prediction	0.5

5.2. Naive Bayes Classifier (NBC) Algorithm

The probability of occurrence of an event can be analyzed by the NBC theorem [26] using evidence or data. The conditional probability and the assumption of attributes are independent of each other and should be considered in the NBC theorem. In spite of its simplicity, NBC enables quick and effective analysis by using high dimensional datasets. NBC is suitable to the case where large supervised or

unsupervised data are processed and the classification is performed in a limited time. In the present study, a naive Gaussian is selected because it matches our prescribed condition.

In the present study, a NBC algorithm is created considering the basis of conditional independence $P(X|Y)$, where $X = (X_1, X_2, \dots, X_n)$ which is used to specify the number of parameters in training and Y is used to emphasize the selected condition of X . The value of each subset $P(X|Y)$ is calculated according to the input X and the problem of estimating the training data is neglected.

In the current study, X is 3 (the number of the bearing condition) and Y is 2 (amplitude of 30 and 90 Hz components). To map the input data in a three-dimensional network, a kernel function is used. Training of the algorithm is performed to obtain a high accuracy rate by importing the data selected randomly.

5.3. K-Nearest Neighbor Algorithm (k-NN)

k-NN [27] is non-parametric and versatile learning algorithm used for both classification and regression problems. Generally, this algorithm memorizes the training datasets instead of learning the discriminative function. The instance-based learning helps in avoiding errors by memorizing the training sets. In the model, the non-parametric is not fixed in advance and it varies based on the data size. The disadvantages of k-NN is its large memory storage, long prediction time, and unnecessary sensitivity to irrelevant features.

k-NN performs classification on testing data based on the k-nearest training samples around the test data. Mainly k-NN depends on two things: (1) a distance metric which is used to compute the distance between two points; and (2) the value of "k" which is used to define the number of neighbors. The value of k decides the shape of the decision boundary. The boundary becomes smoother if the neighbor selection increases the value of k. Technically, small k values result in a hard boundary condition, but it still gives a more flexible fit with low bias and high variance.

In the present study, the value of k is selected in the range of 1–50. With this condition, the boundary becomes smoother and the boundary condition becomes adoptable to any classes of condition. The main objective to get the maximum accuracy rate with high variance condition. Thus, the low bias condition is made possible, making the method suitable to identify the present condition of the bearings.

5.4. Decision Tree Algorithm (DT)

A DT [28] is a dendritic classification model used for both classification and regression problems. The classification is performed by the breakdown of data into smaller subsets and mainly based on the feature selection. The final structure is like a tree with branches and leaf nodes. Each node represents a feature (attribute), each branch represents a decision (rule), and each leaf represents an outcome (categorical or continuous value). A given set of data is divided into several sets, stepwise. The models are recursively constructed from the root node and every possible outcome of the condition is analyzed. If a condition passes through at one root junction, the decision is made at that moment and the output is displayed. If the condition becomes false, it moves to the next stated condition and the process gets repeated until the output is decided.

The constructed DT reaches its optional depth depending on parameter n. Additionally, it is important how many times these branches containing the parameter-n is used as root junction to the analysis. If n is too small, the accuracy may be low. However, when the optimization is performed for larger values of n, the count of missing data is increased. Thus, considering the condition, the value of n is selected as 5 in the present study.

5.5. Random Forest (RF)

The RT [29] is a classification method where the output is decided by the majority rule using the results of plural DTs. First, a DT is constructed by random data extracted from the training data. Then the technique named tree bagging is performed n number of times. Each DT is branched by an explanatory selected variable which has the highest purity. More DTs are constructed, and the

class function is established. The output is concerned with majority of the voting and the final class is declared. Thus, the RF has a better output when compared to the DT algorithm but takes more time for commutation. The class division and the selecting/deciding the result of the majority voting is performed by the bagging function, which is nothing but the tree bagging. For the point of reliability, the value of n is chosen in such a way that an acceptable accuracy rate should be obtained, and the count of missing data should be less.

In the present study, three classes of DTs are selected and framed to form the RF. The DT takes the n value 4 and the bagging technique is performed. Between the three classes of DTs, selection of the variable is done randomly, and the major voting is performed using the tree bagging function to carry out the diagnosis.

5.6. Diagnosis Procedure

Using features, i.e., amplitudes of 30 and 90 Hz components, shown in Figure 6, diagnosis is performed. Features are divided into two categories randomly in the ratio of 70:30. Seventy percent of features are used for training the algorithm and the remaining 30% for evaluation. The same procedure is repeated for all the algorithms described above. The entire diagnosis is carried out in python for all ML algorithms. Based on the evaluation, the accuracy rate of each algorithm is obtained using the formula defined in the programming in python, i.e., the accuracy rate in the present paper is given by:

$$\text{Accuracy rate (\%)} = \frac{\text{Number of data diagnosed properly}}{\text{Total number of data used in diagnosis}} \times 100 \quad (2)$$

5.7. Diagnosis Result

To distinguish the faulty bearing from the healthy one, diagnosis for the bearing condition of healthy (H) and hole (F1), and healthy (H) and scratch (F2) is performed. Diagnosis of conditions F1 and F2 is also carried out to discuss the possibility of fault identification. The rotating speed is not considered for any diagnosis. In case of HF1 combination, a total of 320 data points are selected, where 240 data points are used for training and 80 data points are used for evaluation. In the same manner, data selection is done for other bearing combinations (HF2 and F1F2). Table 2 shows the accuracy rate of the diagnosis using various ML algorithms.

Table 2. Diagnosis result of ML algorithms.

Bearing Conditions	Accuracy Rate (%)				
	SVM	NBC	k-NN	DT	RF
HF1 (Healthy and Hole)	100	100	100	100	100
HF2 (Healthy and Scratch)	81.96	84.31	86.25	78.75	80.75
F1F2 (Hole and Scratch)	91.25	90.88	93.71	90.31	92.88

Each ML algorithm stands unique in its characteristics and were found to be more effective in detecting the bearing faults accurately. A 100% accuracy rate is obtained between HF1 bearing conditions in all the algorithms. In the case of HF2 bearing conditions, the k-NN algorithm acquires a higher diagnosis rate than any other ML algorithms. The DT algorithm produces a low accuracy rate of 78.75%. SVM, being a powerful algorithm, produces an accuracy rate of 81.96%. The other algorithms, k-NN, NBC, and RF, produce diagnosis rates of more than 80%. The average diagnosis rate of each algorithm is high in the case of diagnosing minor faults in bearings and to be accepted practically. In the case of diagnosing F1F2, all the algorithms produced high diagnosis rates of more than 90%. Thus, no further study is required in the case of F1F2 bearing conditions.

The diagnosis accuracy results in the case of F1F2 for SVM and k-NN are shown in Figures 7 and 8, respectively. The result of SVM (Figure 7) shows the variation of accuracy rate with respect to the cost and gamma parameter. The optimization of parameters is performed, and the highest accuracy

rate is obtained. Figure 8 represents the k-NN diagnosis result. The value of k selected between 1 and 50 is indicated and the maximum accuracy rate is achieved with a k value 20. Both the SVM and k-NN show acceptable diagnosis rates.

While using ML, not only should the accuracy rate, but also the missing data calculation, be considered. NBC, DT, and RF are probability-based algorithms. In that case, the amount of missing data is said to be high and the reliability of the acquired result shall be low. On the other hand, both the SVM and k-NN are independent of the probability and the algorithm truly depends on the tuning parameter like the cost and gamma parameters, the threshold function, and k value. The accuracy rate depends on how the tuning is performed and how the hyperplane is constructed. The amount of missing data are low and the reliability and versatility of the algorithm is said to be high. Thus, among the various ML algorithms, k-NN and SVM take priority and can be employed in diagnosing the motor failure in the industrial environment.

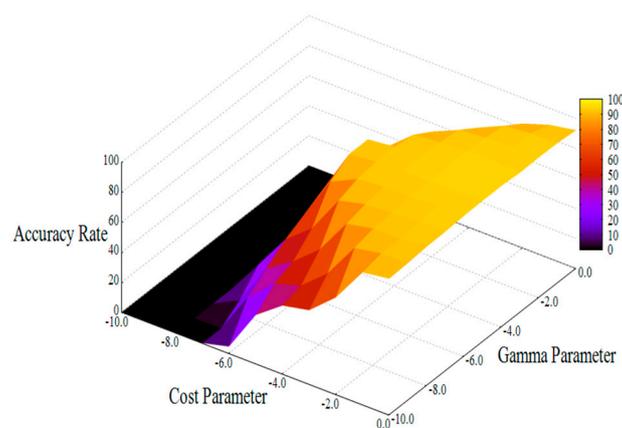


Figure 7. Diagnosis accuracy result of F1F2 in the case of SVM.

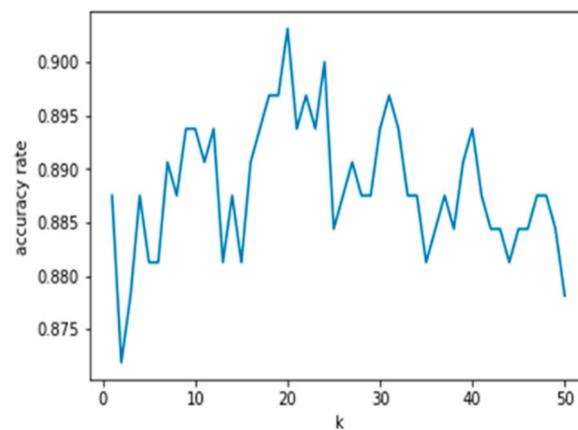


Figure 8. Diagnosis accuracy result of F1F2 in the case of k-NN.

6. Diagnosis Using Deep Learning (DL)

DL [30–32] has wide application in medical fields and is rarely used in the motor field. DL is also said to a form of ML but, in detail, it belongs to the category of AI. The network can be used for both supervised and unsupervised applications. In simple words, the network is formed with multiple layers that contain hidden layers which are used to train the input. The denoising auto-encoders, deep belief network, and convolutional neural network (CNN) are the standard models of DL. Due to advantages like shift-variance, weight sharing, high accuracy rate, and data encoding, among the various architectures, a CNN is selected in the present study.

6.1. CNN Construction

The general structure of a CNN consists of three layers, input, hidden, and output layers, as shown in Figure 9. Since the present study deals with unsupervised learning, the construction of the CNN is based on an auto-encoder system and comprised of a feed-forward neural network. The layers are interconnected to each other and since the overlapping of features is observed between the motor bearing conditions, the pooling layer is inserted between the hidden layers to reduce the cause due to overlapping. The pooling layers carry out the data preprocessing function and, further, to increase the accuracy rate, the auto-encoder system is added to the pooling layer. Figure 10 shows the auto-encoder system which consists of an encoder and decoder; the encoder for extracting the feature from the input data and the decoder for passing the sampling data from the hidden layer. The x , h , y , and W in Figure 10 represent the input layer, hidden layer, output layer, and weight matrix, respectively. Considering the present bearing failure conditions and to obtain a high accuracy rate even with the presence of overlapping, two hidden layers and two pooling layers are constructed in the CNN architecture and, further, the auto-encoder system is constituted to the CNN structure, as shown in Figure 11.

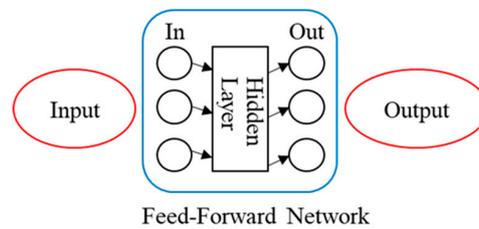


Figure 9. Basic structure of the CNN model.

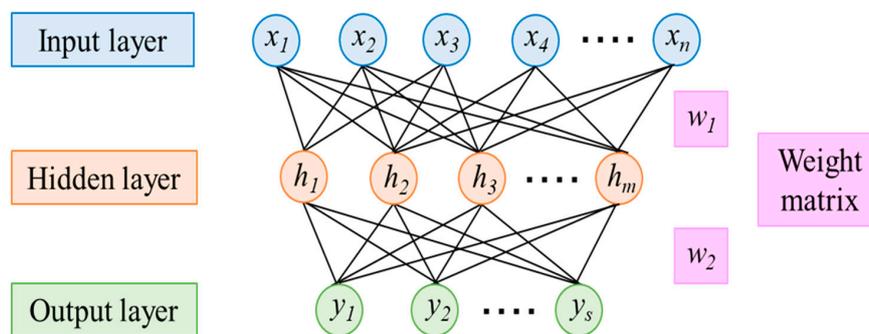


Figure 10. Simple auto encoding system.

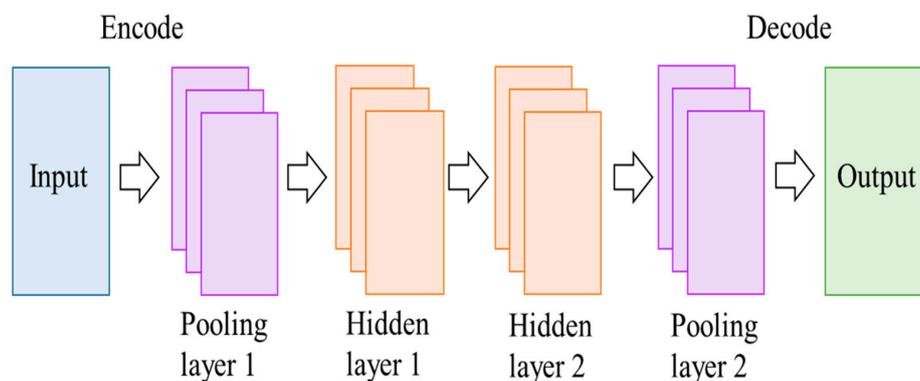


Figure 11. Architecture of the proposed CNN model using an auto encoding system.

6.2. Diagnosis Procedure of CNN

The application of the DL with a CNN architecture to motor bearing failure diagnosis is attempted in this study. Diagnosis is carried out to distinguish the faulty bearing from a healthy one (HF1 and HF2) and to specify the fault condition (F1F2). The same procedure of data selection is adopted, which is described in Section 5.6. To the CNN structure, a stack flow input method is employed and stack values of 0, 1, and 2 are allocated to bearing conditions of healthy (H), hole (F1), and scratch (F2), respectively. The diagnosis is performed using python as a programming tool.

6.3. Discussion

Practically acceptable diagnosis results are obtained and shown in Table 3. The faulty bearing is almost distinguished from the healthy one regardless of the fault configuration; the accuracy rate is 100% for the hole and 82.70% for the scratch. Differentiation of fault type, hole or scratch, is also accomplished with an accuracy rate of 87.29%. Figure 12 shows the change of accuracy and model loss during diagnosis of the bearing condition F1F2. Epoch means the number of iterations during the time of training the neural network, which contains one forward and one backward pass for all of the training samples. As the number of epochs increases, the loss or the error decreases and the accuracy increases. However, at the same time, an increase in epoch may also result in the overfitting problem which makes the diagnosis result worse and it is necessary to determine the number of epochs carefully based on the batch size, the number of training samples in one forward/backward pass.

Table 3. CNN diagnosis result.

Bearing Conditions	Accuracy Rate in CNN (%)
HF1 (Healthy and Hole)	100
HF2 (Healthy and Scratch)	82.70
F1F2 (Hole and Scratch)	87.29

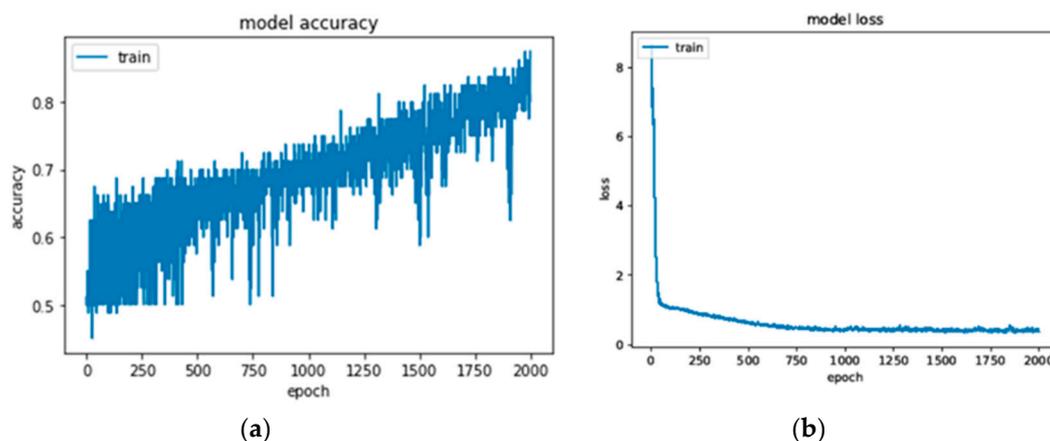


Figure 12. Diagnosis (a) accuracy result and (b) loss of the proposed system.

In diagnosis of HF2 and F1F2 bearing conditions, the CNN produces lesser accuracy than the ML. This is due to a smaller number of data points used for training the CNN. Additionally, as more data are given as inputs to the training phase, then the CNN will produce a high accuracy rate. For variation, an experiment is made using a motor with the same rating and a scratch in the bearing. Once again diagnosis using the CNN is performed. For the case of HF2 combination, a total of 640 data points is selected, where 480 data points are used for training and 160 data points are used for evaluation. The results of diagnosis are shown in Table 4 and are compared with the results obtained using SVM and k-NN of ML. Figure 13a,b shows the change of accuracy with the number of epochs in the CNN

and the k value dependence of accuracy in k-NN, respectively. The CNN gives a higher accuracy rate compared with ML algorithms. Application of AI to fault diagnosis seem more realistic.

Table 4. Diagnosis result comparison.

Bearing Conditions HF2 (Healthy and Scratch)	Accuracy Rate (%)		
	SVM	k-NN	CNN
Normal Data Count (320)	81.96	86.25	82.70
Increase Data Count (640)	83.04	87.85	89.26

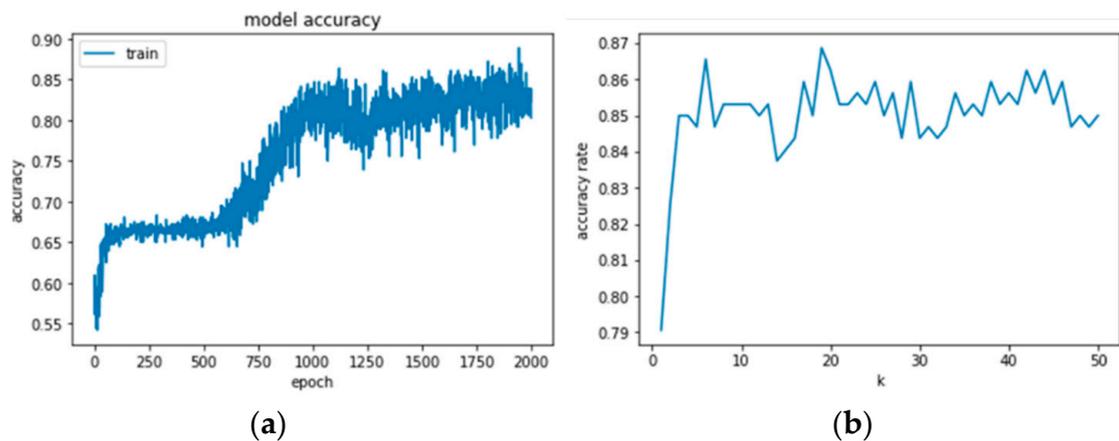


Figure 13. Diagnosis accuracy result of HF2; CNN (a) and k-NN (b).

7. Conclusions

In this paper, a detailed study about the application of ML algorithms and the DL to machine failure diagnosis is carried out. The hole or scratch is selected as a fault for the analysis and the diagnosis is performed without taking the rotating speed in to account.

The diagnosis results of all ML algorithms belong to the same range, yet SVM and k-NN have higher accuracy rates. From the study, the diagnosis method can be selected according to the application. For example, if a large number of bearing conditions is discussed, k-NN and SVM are the preferred algorithm. If the number of bearing conditions is less, for example two or three classes, it is better to go with RF, DT, or NBC. These RF, DT, or NBC algorithms become complex in the case of multiple classes; this is because of the probability-based procedure, as it takes more time for commutation and produces a lower accuracy rate and low reliability.

Finally, a trail to bring the DL algorithm to motor fault diagnosis is achieved and the results are promising and acceptable. It has many advantages over ML algorithms, as it can be trained for any kind of application. The time consumption is less, and programming skill is not required to tune the parameters. The only disadvantage of the CNN is the requirement of the large data count to achieve diagnosis with high accuracy.

Author Contributions: Conceptualization, S.E.P., Y.M. and H.N.; data curation, S.E.P.; formal analysis, S.E.P.; funding acquisition, Y.M. and H.N.; investigation, S.E.P.; methodology, S.E.P.; project administration, S.E.P., Y.M. and H.N.; software, S.E.P.; supervision, Y.M. and H.K.; validation, S.E.P.; visualization, S.E.P.; writing—original draft, S.E.P.; writing—review & editing, Y.M. and H.N.

Funding: This research received no external funding.

Acknowledgments: The authors would gratefully acknowledge the contributions of Santhosh Gunasekaran and Keisuke Asano during the experimental work and data analysis.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

IM	Induction motor
CM	Condition monitoring
FFT	Fast Fourier transform
ML	Machine learning
AI	Artificial intelligence
SVM	Support vector machine
NBC	Naive Bayes classifier
Kk-NN	K-nearest neighbor
DT	Decision tree
RF	Random forest
DL	Deep learning
CNN	Convolutional neural network

References

- Choi, S.; Pazouki, E.; Baek, J.; Bahrami, H.R. Iterative Condition Monitoring and Fault Diagnosis Scheme of Electric Motor for Harsh Industrial Application. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1760–1769. [[CrossRef](#)]
- Riera-Guasp, M.; Antonino-Daviu, J.A.; Capolino, G.A. Advances in Electrical Machine, Power Electronic, and Drive Condition Monitoring and Fault Detection: State of the art. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1746–1759. [[CrossRef](#)]
- Zhang, P.; Du, Y.; Habetler, T.G.; Lu, B. A Survey of Condition Monitoring and Protection Methods for Medium-Voltage Induction Motors. *IEEE Trans. Ind. Appl.* **2011**, *47*, 34–46. [[CrossRef](#)]
- Dalvand, F.; Kalantar, A.; Safizadeh, M.S. A Novel Bearing Condition Monitoring Method in Induction Motors Based on Instantaneous Frequency of Motor Voltage. *IEEE Trans. Ind. Electron.* **2016**, *63*, 364–376. [[CrossRef](#)]
- Motor Reliability Working Group. Report of large motor reliability survey of industrial and commercial installations: Part 3. *IEEE Trans. Ind. Appl.* **1987**, *IA-23*, 153–158. [[CrossRef](#)]
- Gunasekaran, S.; Esakimuthu Pandarakone, S.; Asano, K.; Mizuno, Y.; Nakamura, H. Condition Monitoring and Diagnosis of Outer Raceway Bearing Fault using Support Vector Machine. In Proceedings of the International Conference on Condition Monitoring and Diagnosis (CMD 2018), Perth, Australia, 23–26 September 2018; pp. 1–6.
- Bellini, A.; Filippetti, F.; Tassoni, C.; Capolino, G. Advances in Diagnostic Techniques for Induction Machines. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4109–4126. [[CrossRef](#)]
- Thorsen, O.V.; Dalva, M. A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries. *IEEE Trans. Ind. Appl.* **1995**, *31*, 1186–1196. [[CrossRef](#)]
- Thorsen, O.V.; Dalva, M. Failure Identification and Analysis for High-Voltage Induction Motors in the petrochemical Industry. *IEEE Trans. Ind. Appl.* **1999**, *35*, 810–818. [[CrossRef](#)]
- Javed, K.; Gouriveau, R.; Zerhouni, N.; Nectoux, P. Enabling health monitoring approach based on vibration data for accurate prognostics. *IEEE Trans. Ind. Electron.* **2015**, *62*, 647–656. [[CrossRef](#)]
- Singh, S.; Kumar, N. Detection of Bearing Faults in Mechanical system using Stator Current Monitoring. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1341–1349. [[CrossRef](#)]
- Elforjani, M.; Shanbr, S. Prognosis of Bearing Acoustic Emission Signals Using Supervised Machine Learning. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5864–5871. [[CrossRef](#)]
- Frosini, L.; Harlisca, C.; Szabo, L. Induction Machine Bearing Fault Detection by Means of Statistical Processing of the Stray Flux Measurement. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1846–1854. [[CrossRef](#)]
- Soualhi, A.; Razik, H.; Clerc, G.; Doan, D.D. Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System. *IEEE Trans. Ind. Electron.* **2014**, *61*, 2864–2874. [[CrossRef](#)]
- Gao, Z.; Cecati, C.; Ding, S.X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques-Part I: Fault Diagnosis with Model-Based and Signal Based Approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [[CrossRef](#)]
- Zhou, W.; Habetler, T.G.; Harley, R.G. Bearing Fault Detection Via Stator Current Noise Cancellation and Statistical Control. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4260–4269. [[CrossRef](#)]

17. Ibrahim, A.; El Badaoui, M.; Guillet, F.; Bonnardot, F. A New Bearing Fault Detection Method in Induction Machines Based on Instantaneous Power Factor. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4252–4259. [[CrossRef](#)]
18. Rahman, M.M.; Uddin, M.N. Online Unbalanced Rotor Fault Detection of an IM Drive Based on Both Time and Frequency Domani Analysis. *IEEE Trans. Ind. Appl.* **2017**, *53*, 4087–4096. [[CrossRef](#)]
19. Leite, V.C.M.N.; da Silva, J.G.B.; Veloso, G.F.C.; da Silva, L.E.B.; da Silva, L.E.B.; Lambert-Torres, G.; Bonaldi, E.L.; de Oliveria, L.E.D.L. Detection of Localized Bearing Faults in Induction Machines by Spectral Kurtosis and Envelope Analysis of Stator Current. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1855–1865. [[CrossRef](#)]
20. Raúl, P.; Jordi, F.; Jordi, C.R. Predicting Energy Generation Using Forecasting Techniques in Catalan Reservoirs. *Energies* **2019**, *12*, 1832. [[CrossRef](#)]
21. Jiaying, D.; Wenhai, Z.; Xiaomei, Y. Recognition and Classification of Incipient Cable Failures Based on Variational Mode Decomposition and a Convolutional Neural Network. *Energies* **2019**, *12*, 2005. [[CrossRef](#)]
22. Lijun, Z.; Kai, L.; Yufeng, W.; Zachary, B.O. Ice Detection Model of Wind Turbine Blades Based on Random Forest Classifier. *Energies* **2018**, *11*, 2548. [[CrossRef](#)]
23. Piantsof Mbo'o, C.; Hameyer, K. Fault Diagnosis of Bearing Damage by Means of the Linear Discriminant Analysis of Stator Current Features from the Frequency Selection. *IEEE Trans. Ind. Appl.* **2016**, *52*, 3861–3868. [[CrossRef](#)]
24. Liplowitz, K.B.; Cundari, T.R. Applications of Support Vector Machines in Chemistry. In *Reviews in Computational Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; Chapter 6; Volume 23.
25. Esakimuthu Pandarakone, S.; Mizuno, Y.; Nakamura, H. Frequency spectrum Investigation and Analytical Diagnosis Method for Turn-to-Turn Short-circuit Insulation Failure in Stator Winding of Low Voltage Induction Motor. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 3249–3255. [[CrossRef](#)]
26. You, W.; Qian, K.; Lo, D.; Bhattacharya, P.; Guo, M.; Qian, Y. Web Service-enabled Spam Filtering with Naïve Bayes Classification. In Proceedings of the 2015 IEEE First International Conference on Big Data Computing Service and Applications (CSDL), Redwood, CA, USA, 30 March–2 April 2015; pp. 99–104.
27. Machine Learning, Theory of k-nearest Neighbor Method. Available online: https://dev.classmethod.jp/machine-learning/2017ad_20171218_knn (accessed on 25 May 2019).
28. Machine Learning Algorithm of Decision Tree. Available online: <http://darden.hatenablog.com/entry/2016/12/09/221630> (accessed on 25 May 2019).
29. Importance of Feature Selection and Theory of Random Forest. Available online: <http://drilldripper.hatenablog.com/entry/2016/10/04/211245> (accessed on 25 May 2019).
30. Shao, H.D.; Jiang, H.K.; Wang, F.A.; Zhao, H.W. An enhancement deep feature fusion method for rotating machinery fault diagnosis. *Knowl. Based Syst.* **2017**, *119*, 200–220. [[CrossRef](#)]
31. Tran, V.T.; AlThobiani, F.; Ball, A. An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. *Expert Syst. Appl.* **2014**, *41*, 4113–4122. [[CrossRef](#)]
32. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccufier, M.; Verstockt, S. Convolutional neural network-based fault detection for rotating machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).