Intelligent predictive maintenance model for rolling components of a machine based on speed and vibration
Ahmad, Baseer; Kumar Mishra, Bhumesh; Ghufran, Muhammad; Pervez, Zeeshan; Ramzan, Naeem

Published in:
The Third International Conference on AI in Information and Communication (ICAIIC 2021)

DOI:
10.1109/ICAIIC51459.2021.9415249

Published: 29/04/2021

Document Version
Peer reviewed version

Link to publication on the UWS Academic Portal

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the UWS Academic Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact pure@uws.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

“© © 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”
Intelligent Predictive Maintenance Model for Rolling Components of a Machine based on Speed and Vibration

Baseer Ahmad  
School of Computing, Engineering and Physical Sciences  
University of West of Scotland  
Paisley, Scotland  
baseer.ahmad@uws.ac.uk

Bhupesh Kumar Mishra  
Faculty of Engineering and Informatics  
University of Bradford  
Bradford, England  
b.mishra@bradford.ac.uk

Muhammad Ghufran  
Department of Electrical Engineering  
Abasyn University  
Peshawar, Pakistan  
ghufran.muhammad@abasyn.edu.pk

Zeeshan Pervez  
School of Computing, Engineering and Physical Sciences  
University of West of Scotland  
Paisley, Scotland  
zeeshan.pervez@uws.ac.uk

Naeem Ramzan  
School of Computing, Engineering and Physical Sciences  
University of West of Scotland  
Paisley, Scotland  
naeem.ramzan@uws.ac.uk

Abstract—Machines have come a long way, from the industrial revolution to a modern-day industry 4.0. In this massive transition, one thing that has never changed within a machine is the moving part. Most industries use rotating machine with different load capacity and speed. These machines run at variable load and variable speed creating vibration bootstrap thus causing machine failure due to an increase in vibrations. Most of the researcher used vibration for fault detection in bearings but sometimes it caused by miss alignment in a shaft due to a fraction of overloading the machine. In this paper, we address it to solve those problems by using two parameters speed and vibration. To verify our approach, we use three different kinds of machine learning algorithms: Support Vector Machine (SVM), Naïve Bays, and Random Forest. By using these machine learning algorithms, we tried to find out the relationship between machine failure due to speed and vibration by predicting good and faulty bearings. After applying these models, we have seen that the SVM has 78% accuracy as compared to Naïve Bays, and Random Forest.

Keywords: Predictive Maintenance, Machine Learning, SVM, Naïve Bays, Random Forests

I. INTRODUCTION

Today's world industry is running on machines and most of them have rotating parts where the main rotating element is the rolling bearing which support the rotating shafts. The health and performance of those machines depend on the defect-free working of rolling bearings. The poor condition of rolling bearing causes several negative impacts such as unexpected bearing failures, reduction in productivity, increase in machine downtime as well as the increase in permanent breakdowns risk [1], [2]. During the past decades, many researchers work to solve these problems by using intelligent system algorithms such as machine learning, AI, and recently deep learning as well as different maintenance strategies such as Run to failure/reactive maintenance [3], Preventive maintenance [4], Predictive maintenance [5] and Condition-based maintenance (CBM) [6] have been used.

Run to failure/reactive maintenance has been used only when a fault occurs in a machine whereas reactive maintenance returns the equipment to its usual working state after the equipment has been broken down. Although there might be a place for reactive maintenance, a well-rounded maintenance plan should not be for all repairs. The drawback of this kind of strategy is increased downtime and adding more cost in terms of stop productivity, and labour cost [4].

Preventive maintenance is another type of maintenance strategy where pre-scheduled maintenance plan for each machine are created based upon the machine usage cycle i.e., how often it is been used and how busy it has been. For example, if a machine is working without any break over a longer period then it may require a regular maintenance check to keep it in a healthy state. This kind of machines more often require preventive scheduling or inspections. Also, it depends on the age of the machine, a new machine requires four to five-time scheduling per year and an old machine may require more inspection in a year [7]. The advantage of this kind of strategy is to minimise permanent breakdown to machines and reduce downtime. However, applying preventative maintenance adds extra labour costs as frequent inspections are required even though the machines are in good running condition [8].

Apart from “Run-to-failure reactive maintenance” and “Preventative Maintenance”, another strategy is Predictive maintenance. This type of maintenance is done by using predictive algorithms such as SVM (Support Vector Machine), Naïve Bayes etc. The basic operation of this method is to collect data from various sensors, which is then used for training to build a machine learning model [5]. Data with healthy and unhealthy trends and patterns are used to train a model. The trained model is used to predict current health conditions based on machine current state data. Whenever, the model detects any pattern which indicates the machine is going towards an unhealthy state, the model creates an alert before its breakdown [9].
Condition-based maintenance (CBM), modification of predictive maintenance, is another strategy to collect real-time data information from intelligent sensors to monitor the machine’s health conditions and make a decision based on collected information [10]. The CBM aims to make useful life prediction of a machine based on degradation in trends along with historical data [11]. This kind of maintenance technique helps to plan maintenance activities effectively and efficiently according to the data received from the real-time monitoring system. It decreases downtime and increases the life of the machine. All these maintenance strategies’ objective is to increase the productivity and decrease downtime of machines and decreases the periodic maintenance cost [12].

In all of the above-mentioned strategies predictive and Condition-based maintenance are good from preventive and run to failure. Because of using historical data for predicting faults by comparing different patterns from realtime sensors. While Run to failure only apply when faults occur which leads to an increase in machine downtime and at another side the preventive maintenance increase the labour cost by servicing machines multiple times a year depends on the type of machine without even breakdown.

Predictive maintenance models mostly use vibration parameters to predict a fault in a mechanical rotating part. However, speed is also an important parameter that can indicate a fault in a machine. According to the research, there is a direct relationship between speed and vibration [13]. A model which is trained with only vibration parameter can detect a faulty bearing, nevertheless, it will still give a wrong prediction after replacing bearing due to an unbalanced shaft which is connected and supported by that bearing. A predictive model trained with speed and vibration will be more effective than a model trained only based on vibrations. In this paper, we applied a comparative study on three predictive models which are trained on speed and vibration to successfully detect a faulty and healthy bearing.

To achieve the goal, we used three machine-learning algorithms SVM, Naïve Bayes and Random Forest. These machine learning models are widely used for predictive maintenance [14]. We get higher precision and fewer errors in the SVM. We discussed all three models in detail in section IV of this paper.

The rest of the paper is organised as follows. In Section II, the literature on related works is presented. In Section III, our approach of fault detection using a different machine learning approach is presented. In Section IV, computational experiment and result analysis are presented and in Section V, the conclusion is future work is presented

II. LITERATURE REVIEW

Over the years, several prediction techniques have been deployed for early failure alerts generating for machines. Vibration-based maintenance technique commonly used for prediction of the faulty bearing. Where the remaining useful life is mainly used for health monitoring and predictive maintenance with artificial intelligence, machine learning and deep learning techniques [15].

Literature shows multiple faults can occur in a bearing with different frequencies as shown in equation (1-4) which represents the faults that normally occurs in the bearings [16]. Figure 1 shows the bearing mechanism of a machine [17].

Normally bearing has 4 elements which are rolling Ball, inner ring, cage (which hold the bearing inside the inner and outer ring) and outer ring.

![Figure 1 Bearing structure](image)

- $D_b =$ Ball diameter
- $D_c =$ Pitch diameter
- $N_b =$ Number of rolling elements
- $\beta =$ Ball contact angle
- $F_R =$ Rotating shaft frequency

A. Equations

1. Cage fault frequency: $F_c = \frac{1}{2} \left(1 - \frac{D_b}{D_c} \cos \beta \right) F_R$
2. Inner race fault frequency: $F_i = \frac{N_b}{2} \left(1 + \frac{D_b}{D_c} \cos \beta \right) F_R$
3. Outer race fault frequency: $F_o = \frac{N_b}{2} \left(1 - \frac{D_b}{D_c} \cos \beta \right) F_R$
4. Ball fault frequency: $F_b = \frac{N_b}{2} \left(1 - \left(\frac{D_b}{D_c} \cos \beta \right)^2 \right) F_R$

In the above equation, the $F_c$ represents the cage fault which is not often encountered and occurred when a defect of speed variation founds [18]. $F_i$ represents any scratches or degradation that occurred to the inner ring from inside which create a fraction with balls where $N_b$ is the number of balls used inside the bearing [19], this fault is commonly found in machines because of a direct link with the shafts FR. Fo is the same as $F_i$ however it’s not encountered often and create less vibration. $F_o$ is also an uncommon fault and the ratio of this fault is less than other faults.

The occurrence of these faults depends on the variation of load and speed [20]. The research also shows that vibration increases when it is passing through a faulty region due to an unbalanced shaft. Each machine has its environmental factors and conditions, which predominantly define the work-able life of the rolling bearing. On the bases of those conditions which is mentioned in equation (1-4), companies need to be careful about the selection of maintenance policies as discussed in the introduction section which suits the machine behaviour for its better performance and its life span.
Recurrent Neural Network Health Indicator (RNN – HI) technique was proposed to achieve the goal of early prediction of bearing failures [21]. The construction of their model was based on three main stages: sensitive features, related similarity or feature extraction, and RNN-HI construction. They used two accelerometers on the machine in a horizontal and vertical orientation for detecting vibration on the experimental platform with different speeds [22]. During the experiment, the vibration was periodically increased with respect to time. For training, the model took the initial data of a new bearing and then the final data was taken once the bearing showed the signs of vibration due to wear and tear. In their test results, there was a linear relationship between the real-time vibration data from sensors and actual physical damage which verified their data-driven and physics-based method.

In [23] a dynamic predictive maintenance framework (DPFM) has been applied where Long Short-Term Memory (LSTM) network was used based on RNN architecture. Memory retaining capacity and learning over long time sequences was the main advantage of this architecture. The author used NASA’s turbofan engine degradation simulation dataset [24] for verification of their model. For performance evaluation, they compare the cost rate of the two policies “Predictive maintenance policy” and “Ideal predictive maintenance policy (IPM)”. In recent prognostics health studies, Predictive maintenance and Ideal predictive maintenance policy gain more attractions [25].

In the following, two policies are used by Nguyen et al., [23] to make a new predictive maintenance framework.

Periodic maintenance policy: This policy is based on the historical data of equipment for prediction of early fault detection which decreases downtime and also decreases cost with the below information

$$T_F = \left[ \frac{r_F}{\Delta T} \right]^* \Delta T$$

where $T_F$ is time to failure obtained from historical data and preventive maintenance with the cost of $C_p$ (preventive maintenance cost) performed at $TR$ (equipment is available for maintenance).

Ideal predictive maintenance policy (IPM): This policy is based on the hypothesis of accurate failure time prediction by assuming the correct lifetime during the inspection. The cost rate value can be minimising based on this perfect prediction information given below:

$$CR_{IPM} = \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{C_p}{T_R - 1/\Delta T} \right] \Delta T$$

Where parts are available for preventive maintenance $[T_R - 1/\Delta T]$ after inspection. So, the cost is minimised.

Researchers in [26] used the same NASA dataset [24] with a different approach for prediction and health prognostic by vanilla LSTM method.

Figure 2 represents the vanilla LSTM method used in [26] - where $(Xt - 1)$ is a combination of multiple inputs and $(Ht - 1)$ outputs of the previous time, $Ct$ is kind of classic RNN cell state where $(Ct-1)$ is an old state of a previous time and $(Ct+1)$ is a new state of the current time. This approach showed better accuracy for the prediction of multiple faults in a complex machine system for high precision RUL (Remaining Useful Life) estimation [27].

III. FAULT DETECTION USING SPEED AND VIBRATION: OUR APPROACH

In a machine, there is a direct link between speed and vibration, which can be used to detect faulty behaviour of a moving part i.e., bearing [13]. For example, a faulty bearing vibration and speed are directly proportional, whereas a normal bearing will not show any signs of vibration regardless of its speed. The literature has also illustrated that the same model or the same device from the same manufacturers has different vibration patterns [13]. Therefore, there is a need for the development of multiple or dedicated prediction models for each device. In this work, we present a comparative study between three approaches to detect a fault in a bearing.

A. Support Vector Machine-based model

SVM is a commonly used supervised machine learning model. It is used for classification and regression. Its implementation is unique as compared to other machine learning algorithms. It can handle categorical and multiple continues variables. It can represent different classes in multidimensional space on a hyperplane. For minimizing error, it iteratively generates a hyperplane. To find maximum marginal hyperplane (MMH) SVM divide the dataset into different classes as shown in figure3 [28].

![Figure 3 concepts in SVM](image-url)
B. Naïve Bayes-based model

The Naïve Bayes classifier is a Bayes theorem-based algorithm, it belongs to the statistical family of the classifier and considered to be a supervised machine learning algorithm. It has a strong independence assumption between the features. This classifier is highly scalable, it only requires a few linear parameters to predict features in machine learning problems. Literature shows Naïve Bayes has a variety of models like Independence Bayes and simple Bayes [29].

\[
P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}
\]

The Bayes rules work from P(Y|X), which is known from training datasets to find P(Y|X). by replacing A and B with X and Y in Bayes rule, with the feature of x and return value of y. For testing, the observation data x is known while y is unknown. To compute the probability of y given the x has already happened in each row of the dataset. In another case, if y has more than two types than we compute the probability of each y class and let the peak win.

C. Random Forest-based model

Random forest is used for regression as well as classification. It is also a supervised machine learning algorithm like SVM and other classifiers. It is widely used for the classification of machine learning problems. In real-world forest has many trees which make it more robust. The same concept is applied in a random forest algorithm, which creates decision trees based on samples of data and gets prediction results from each of them utilizing voting to ensure the best possible solution. By taking average result it reduces over fittings, this method makes it better from a single decision tree. Figure 4 shows a working model of a Random Forest [30].

![Figure 4 Structure of the random forest classifier](image)

IV. COMPUTATIONAL EXPERIMENT AND RESULT ANALYSIS

In this section, the computational experiment is presented to demonstrate the efficiency of SVM, NB, and RF in detecting faculty bearing using speed and vibration. The experiment is run on a machine with 8 core i9 processors along with an RTX2080 Graphics card and 32GB of RAM. All these models are coded in R language.

A. Dataset Pre-processing and Data Labelling

The dataset used in this paper is taken from [31] - on time-varying analysis on vibration data. The dataset consists of 36 different speed varying datasets of good and faulty bearings. Two datasets in which speed within the range of 14Hz to 29Hz are selected for our model training. We selected this dataset for verification of our approach to demonstrate the relationship between speed and vibration, whilst ensuring a high detection rate of faculty behaviour. Because the test platform linearly increased the speed of the motor from 14Hz to 29Hz without any jerk effect, we called it soft start. This prevented creating a high vibration peak at the starting of the machine. The chosen dataset consists of 2 million observations. One dataset consists of good bearing vibration and the other consists of faulty bearing vibration. The dataset is combined for training machine learning models. By using the CARET package in R, the dataset is split into 80% training data and 20% test data.

B. Relationship between Vibration and Speed

At first, the relationship between vibration and speed has been plotted as shown in figure 5. The vibration pattern can be observed for good and faulty bearings against the speed. The blue lines representing the faulty vibration and the green representing the good bearing vibration. The speed range is denoted from 0 to 4 representing the increase in the speed level from 14HZ to 29Hz.

![Figure 5 Speed and vibration](image)

C. Prediction Models

After verification of the relationship between speed and vibration, three different models (1) support vector machine SVM, (2) Navies Bayes, And (3) decision tree are trained and tested.

1) Support Vector Machine (SVM) –based prediction

For the best-optimised SVM model, we used the linear kernel which has most commonly used for prediction when we have a large dataset with a number of features. It is faster than other kernels like Polynomial, Gaussian (RBF) and Sigmoid. It only required one tuning parameter which is C Regularisation that controls the trade-off between accomplishing a low training error and a low testing error that is the capability to simplify your classifier to hidden information. On the other hand, when selecting other kernels for prediction they also required the Y parameter along with the C parameter which consumes more time. Because they are using the grid search.

By achieving the best result, we change the (C) cost value from 5 to 10 for smooth decision surface low cost is
used and for classifying more points correctly high cost is widely used. These optimised values achieved an accuracy of 78% on detecting faulty and good bearings.

2) Naïve Bayes–based prediction

The same percentage of data is used for the training and testing of this model. By splitting the test and training data we store the data in variable x and classes data in y according to the Bayes algorithm as discussed earlier in Section III. In the background of our vibration data, we are seeking the probability of bearing belonging to a faulty and good class Ck (where Cyes=good and Cno= faulty) given that its predictor values are x1,x2,…,xp. This can be written as P(Ck|x1,…,xp). By apply 10-fold cross-validation tuning parameters. The model has an accuracy of 82% on detecting faulty and good bearings. By changing the tuning parameters results do not change and remain the same.

3) Random Forest-based prediction

The same x and y variable are used in this model which we create for Naïve Bays. Random Forest algorithm has two main tuning parameters to achieve higher accuracy. “MTRY” is the number of randomly sampled variables at each split, and “NTREE” is the number of trees to produce, by default it is 500. Due to my pc memory limitation... we change the “mtry” parameter to achieve the highest accuracy of 78.2% on detecting faulty and good bearings.

D. Comparative Analysis

Comparative study of these three models, based on SVM, NB, and RF have been analysed. For the comparative analysis, a confusion matrix for all the models have been generated and different performance measures are compared for each model. Below are the tables for the confusion matrix of SVM, Naïve, Random forest-based models respectively.

Table 1 SVM Prediction Model Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Faulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>71.63%</td>
<td>28.36%</td>
</tr>
<tr>
<td>Faulty</td>
<td>9.83%</td>
<td>90.16%</td>
</tr>
</tbody>
</table>

Table 2 Naïve Bays Prediction Model Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Faulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>72.56%</td>
<td>27.43%</td>
</tr>
<tr>
<td>Faulty</td>
<td>12.23%</td>
<td>87.76%</td>
</tr>
</tbody>
</table>

Table 3 Random Forest Prediction Model Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Faulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>78.66%</td>
<td>21.33%</td>
</tr>
<tr>
<td>Faulty</td>
<td>13.35%</td>
<td>86.64%</td>
</tr>
</tbody>
</table>

As it can be seen that in the above confusion matrix diagram the best accuracy for true positive is achieved in the Random Forest as compared to the Naïve Bays and SVM but on another side, we have a high accuracy of true negative in SVM as compare to the Random Forest and Naïve bays. Also, SVM has less error chance to predict good bearing as faulty if we see the ratio is 9.83 percent. While Naïve Bays and Random Forest has a 12.2 and 13.3 percent chance of predicting a good bearing as a faulty bearing.

![Figure 6 Comparative plot of three approaches performance measures](image)

If we take another look at figure 6 where precision and recall for Good and Faulty bearing classes, SVM has less precision to avoid predicting of Good class but has high Recall to avoid errors in predicting Faulty bearing class. From Figure 6, Random forest has high precision but lower recall than SVM. Also, Naïve Bays has an average of both precision and recall as compared to the other models.

If we compare the time consumed for each model training the highest number is 15 hours for SVM because of the selection of higher-cost value to classify higher accurate points as we discussed in Section IV, that's why we have a good accuracy percentage of true negative. While Naïve Bays and Random Forest take 6 and 2 hours respectively. Because of less access to the tuning parameters and limitation of dataset observations like in Random Forest we were limited to 500 trees due to resource limitations.

V. CONCLUSION AND FUTURE WORKS

Initially when any machine starts the first element which has a direct link with friction is bearing which causing heat factors and then leads to vibration. So, in that case, the motor is spinning at a higher speed the more fraction it faces and it creates more vibration which finally leads to breakdown. As we discussed in section IV the speed is slightly increasing and there is much difference in the graph between the good bearing and faulty bearing vibration. In this paper, we introduced the vibration and speed relationship to differentiate between a good and a faulty bearing. By testing widely used models of SVM, Naïve Bays, and Random Forest verified our approach towards speed and vibration relationship. However, it further needs more polishing to achieve the best results, as we mentioned in each algorithm that they have different kernels with different tuning parameter, for this in future we focus on more model testing with different kernels and tuning parameters as well as finding some other features along with speed and vibration. Also, we will use deep learning and artificial neural network algorithm as we see in literature like LSTM, and others because of supporting millions of observations for training models.
VI. REFERENCES


[31] H. Huang and N. Baddour, “Bearing Vibrati-