# Intelligence Based Condition Monitoring Model

**Kouroush Jenab, Tyler Ward, Cesar Isaza, Jorge Ortega-Moody, and Karina Anaya**

**Abstract** Condition-Based Maintenance (CBM) is a maintenance strategy that reduces equipment downtime, production loss, and maintenance cost based on changes in equipment condition (e.g., changes in vibration, changes in power usage, changes in operating performance, changes in temperatures, changes in noise levels, changes in chemical composition, increase in debris content and changes in the volume of material). In this study, we present the newly developed Condition Monitoring Model (CMM) based on an ensemble machine-learning model that utilizes the random forest, support vector machine, and artificial neural network classifiers, to classify data points from the normal state of a rotating machine. The efficacy of the model in adequately detecting and diagnosing faults in the rotating machine for maintenance planning is discussed in this paper. The developed model can efficiently avoid unnecessary maintenance and make timely actions by analyzing the received vibration signals from the rotating machine. An illustrative example is demonstrated to present the application of the model.

**Keywords** Maintenance management · Condition-based maintenance · Maintenance loss function · Bayesian decision theory · Artificial decision making

K. Jenab (✉) · T. Ward · J. Ortega-Moody
School of Engineering & Computer Science, Morehead State University, Morehead, KY, USA
e-mail: k.jenab@moreheadstate.edu

T. Ward
e-mail: tbward@moreheadstate.edu

J. Ortega-Moody
e-mail: j.ortegamoody@moreheadstate.edu

C. Isaza · K. Anaya
Networking and Telecommunication Academic Program, Polytechnique University of Queretaro, Queretaro, Mexico
e-mail: cesar.isaza@upq.mx

K. Anaya
e-mail: karina.anaya@upq.edu.mx

# 1 Introduction

The efficiency of any industrial setting is dependent on the uninterrupted operation of critical equipment. Different strategies exist for proper machine maintenance scheduling, each with its own benefits and drawbacks. These maintenance strategies can generally be grouped into three main categories: preventive maintenance (PM), corrective maintenance (CM), and CBM [1].

Traditional maintenance strategies largely fall into the corrective or preventative maintenance strategies. CM can be defined as maintenance that occurs after a machine has failed, while PM is conducted on a set schedule to prevent the occurrence of faults in the machinery [1]. While PM is preferable to CM in most cases due to CM generally being the most expensive form of maintenance since it occurs after a machine has already failed [2], PM has its own flaws because as it is conducted on a set schedule, it can result in unnecessary maintenance being conducted on a machine that is not in danger of failing.

While some authors [3] equate CBM and PM, or consider CBM a type of PM, we consider CBM as a distinct maintenance strategy from PM. This is because that while CBM also aims to conduct maintenance on a machine before a failure occurs, it also uses analytical techniques to monitor the actual condition of the machine, and the maintenance is only conducted at the point right before the point of failure. This process reduces the potential for unnecessary maintenance that is often present in PM approaches.

CBM is not without its own limitations, with a high investment cost resulting from the necessity to install adequate monitoring equipment, develop decision-making strategies, and train staff on the new technology and processes [1]. However, with modern advances in machine learning (ML) technology, these drawbacks can be addressed in a way that reduces costs and improves efficiency [4]. ML for maintenance has applications in the fields of Bayesian decision theory and artificial decision-making and can help improve maintenance loss functions.

ML algorithms can be categorized in a variety of ways, but two simple learning categories can describe most ML methods: supervised learning and unsupervised learning. Supervised methods typically involve the process of classification, such as detecting faults in incoming data based on historical examples, while unsupervised methods typically involve clustering, an example being the grouping of data points to attempt to determine when a fault will occur. Ensemble learning is a method that combines both supervised and unsupervised approaches.

This paper presents a Condition-based Maintenance Model (CMM) based on an ensemble model stacking approach to improve the performance of ML for CBM compared to traditional approaches. This model is capable of accurately classifying normal and faulty data from sensors placed on a rotating machine. This capability reduces the possibility of unnecessary maintenance and enables timely automated decision-making.

## 2  Literature Review

CBM is a maintenance strategy that focuses on real-time condition monitoring of industrial machines, only recommending action be taken if machine failure is imminent, to eliminate unnecessary maintenance and costs. CBM would not be possible if not for the variety of sensors and other indicators that continuously send signals for analysis. There are two main factors in the analysis of these signals: diagnosis and prognosis [5].

Fault diagnosis can be defined as determining the type of fault that has occurred based on factors such as abnormal sounds or vibrations [6], while fault prognosis is defined as the estimation of the operating time before a failure occurs [7]. Traditionally, these tasks have been performed by large teams of engineers, which can cause low efficiency and suspect accuracy. These factors make fault diagnosis and prognosis ripe for ML support [6].

A survey of ML algorithms in Industry 4.0 literature [4] found that artificial neural networks (ANN), random forest (RF), and support vector machines (SVM) were the most used, with RF being the most prevalent. Decision tree (DT), Gradient boosting machine (GBM), and eXtreme gradient boosting (XGBoost) algorithms were also discovered to be common in the reviewed literature. Other algorithms discussed include logistic regression (LR), linear regression, convolutional neural network (CNN), and deep neural network (DNN).

RF's prevalence comes from its improved performance when compared to contemporary algorithms. A CBM approach based on a comparison of ANN, DT, and RF's performance on the prediction of faults in a high-speed packing machine found that RF achieved the highest accuracy rating, with the caveat that the false positive rate is higher with RF than the other approaches [8]. A possible solution to this issue is to create a hybrid model that combines RF with other ML algorithms to strengthen the performance of each model and reduce any flaws that are present [4, 8].

Model stacking is an approach that aims to strengthen the performance of ML models by using previous predictions of individual models to form a hybrid model based on the best performing classifiers of each model [9]. Model stacking has been shown to achieve better classification accuracy for maintenance related tasks when compared to individual models by themselves [10]. LR is frequently used as the meta-model to combine the multiple ML models [9].

The improved accuracy of model stacking techniques is demonstrated in [11], where predictions from SVM, k-Nearest Neighbors (KNN), Bayes' classifier (NB), and LR were stacked using RF as the meta-model. This study showed that the stacked model achieved better accuracy on the task of fault diagnosis on roller bearings than the individual models themselves. The stacking method used in this study is a heterogeneous approach, meaning that the algorithms stacked were different [12]. In contrast, homogeneous methods stack the same algorithm multiple times [12], such as in [13].

The research discussed in this paper focuses on an intelligence-based CMM for rotating machines. This is because the use of rotating machines since the Industrial

Revolution has increased drastically [14], and failures of these machines can lead to severe economic or social impacts [15]. Because of the potential for severe negative impacts, early detection of faults is crucial, meaning that a CBM approach is preferable to different maintenance strategies [1].

## 3  Methodology

This research project proposes an intelligence-based CMM for rotating machines. The methodology to create this CMM consists of four main steps: downsampling collected data, training individual ML models, stacking those models, and using the resultant ensemble model to diagnose machine faults. A block diagram of the system is shown in Fig. 1. The steps are described in the following sections.

### 3.1  Machinery Fault Database

A public database for machine faults was used for this research [16]. This database consists of data from ten simulated machine states: normal operation, imbalance fault, horizontal and vertical misalignment faults, and outer race, rolling element, and inner race faults for underhang and overhang bearings. The machine used to collect the data is shown in Fig. 2. The data was collected from three accelerometers, one triaxial accelerometer, one analog tachometer, one microphone, and two four-channel analog acquisition modules.

There were 1,951 total sequences in the database with 250,000 samples each, meaning that the total number of samples for the dataset is 487,750,000. Different parameters were used for each of the machine state simulations: 49 normal sequences were simulated with a fixed rotation speed between 737 to 3686 rpm, 333 imbalanced sequences were simulated with a load range from 6 to 35 g, 197 sequences had a horizontal misalignment with motor shaft shifts from 0.5 to 2.0 mm, 301 sequences had a vertical misalignment with motor shaft shifts from 0.51 to 1.90 mm, and 1,071
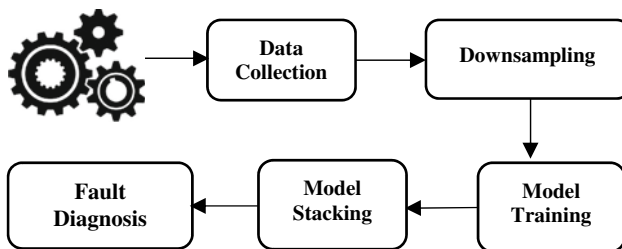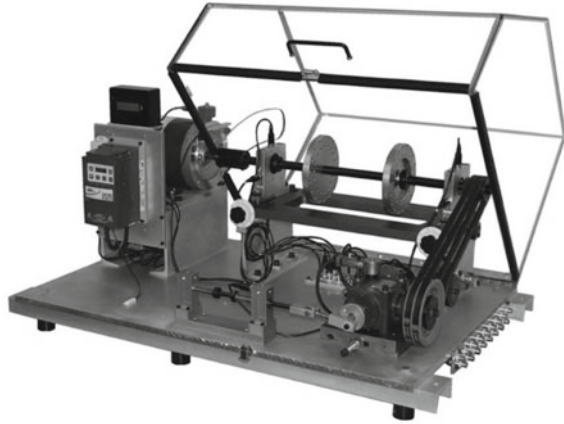


**Fig. 1** Proposed methodology for intelligence-based CMM

**Fig. 2** The fault simulator
used to generate the data



sequences were simulating either underhang or overhang bearings with roller element
or inner and outer race faults with a mass between 0 to 35 g added [16].

## 3.2 Data Pre-processing

Due to computational limits, the workable size of the database was downsampled to
80,000 samples, with 8,000 samples per machine state. Data for the imbalance state
was taken from the sequence with a 6 g load, data for the horizontal and vertical
misalignment states were taken from sequences with 2.0- and 1.90-mm motor shaft
shifts, respectively, and data from the underhang and overhang bearings were taken
from bearings with 6 g masses applied.

## 3.3 Model Training

So far, research into intelligence-based CMMs has focused on the application of
one ML model to a problem. We propose a model stacking approach to increase the
accuracy of fault diagnosis in a rotating machine. The ensemble model generated
from this approach is based on three of the most popular, accurate, and efficient ML
algorithms in use.

The system starts with the input data from the machine being split into training
and testing sets so the ML models can be trained. The train and test split utilized
by this research is 80% of the data in the training data, and 20% of the data in the
testing set. Each model is trained individually and then combined by the metamodel
to form the ensemble model that shapes the CMM.

### 3.3.1 Random Forest

RF is an ensemble method based on the combination of results of multiple DTs. Traditional DT methods have long been employed in both traditional and intelligence-based maintenance strategies, due to their ability to aid data analysis and numbers-based decision-making. The general process for using DTs in maintenance consists of the following steps: visually represent complex decisions and their potential outcomes, calculate a "cost" value for each decision and their consequences, and then compare the final values to determine the best course of action. Done by hand, this can be a time-consuming process, which is why the most common form of DT analysis in Industry 4.0 is ML-based DT.

There exist several types of ML DT algorithms: classification trees, regression trees, and classification and regression trees. Each type of tree is named after the task it enables. Although DT is a very simple and powerful approach to implement, the chance of overfitting the data is high, especially in cases where a data set is very large. The RF approach is a solution to this issue.

As a classification algorithm, the RF approach assembles a forest of classification trees to enable its predictions. One method to reduce data overfitting as well as data variance for DTs is called bootstrap aggregation (bagging). This approach repeatedly samples the training set with replacement and trains individual models on each of the samples, before averaging each prediction to generate the final predictions of the model. While generally effective in accomplishing its goal and generating accurate predictions, bagging can result in trees becoming correlated in the event that a strong estimator is found. RF extends the bagging approach by training the classification trees on a random sample of features from the training set, a process called the random subspace method. A DT taken from a subset of data in the RF of the CMM is shown in Fig. 3.



**Fig. 3** A sample decision tree based on a subset of the data
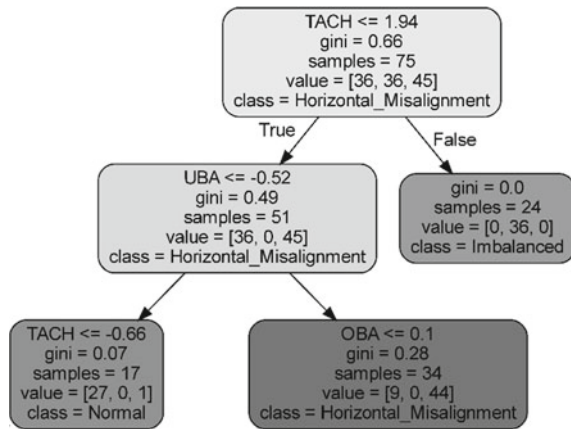
### 3.3.2 Support Vector Machine

SVM is one of the most popular classification algorithms, and offers robust prediction capabilities. Given a high-dimensional set of data, the SVM approach operates by finding a hyperplane, or decision boundary, that best divides the data points into distinct classes. The hyperplane is chosen based on the maximum distance between the closest data points, also called support vectors, from each class. Initially, this approach was only effective on a linearly separable data set, but advancements have been made to enable the classification of non-linear data via a process called the kernel trick.

The kernel trick transforms non-linear data into a higher dimensional space where it can become linearly separable. This process works by computing the dot products between data points in the new higher dimensional space without explicitly computing the coordinates of the data. There exist a variety of kernel functions, each with their own benefits and drawbacks. This research utilizes a kernel function called radial basis function (RBF), which operates by measuring the similarity between two points in the higher dimensional space based on their radial distance from the origin [17].

As with kernels, there are also different implementations of the SVM algorithm itself. This research utilizes C-Support Vector Classification (C-SVC), where C is a parameter that determines the cost of misclassifications. By default, C-SVC only solves binary classification problems, so a one-versus-one approach is used to enable multi-class classification. The one-versus-one system considers a series of binary classifications, before selecting the classifier that gave the highest prediction confidence score when making its final predictions.

SVM is one of the most common ML algorithms used in maintenance. One of the major contributing factors to this is that SVM is based on much of the same statistical learning theory as many of the traditional methods employed in quality control. It is a data-driven algorithm, and has been found to be more generalizable than techniques such as ANN. It is because of this generalizability that SVM finds many uses in maintenance and quality control, because it can be applied to many different situations without much editing of the process [18]. A visual representation of the decision boundaries for the SVM approach in the CMM on a subset of data is shown in Fig. 4.

### 3.3.3 Artificial Neural Network

ANNs are unique compared to RF and SVM in that they are based on biological processes, specifically animal brains. ANNs generally consists of three layers: input, hidden, and output. The input layer is the first layer and is responsible for receiving information signals. These signals are passed to the hidden layers where patterns are extracted and analyzed. Finally, in the output layer, neurons process information from the previous layers to produce and present the final outputs of the ANN [19]. A feedforward ANN is an ANN where the connection between the nodes do not form
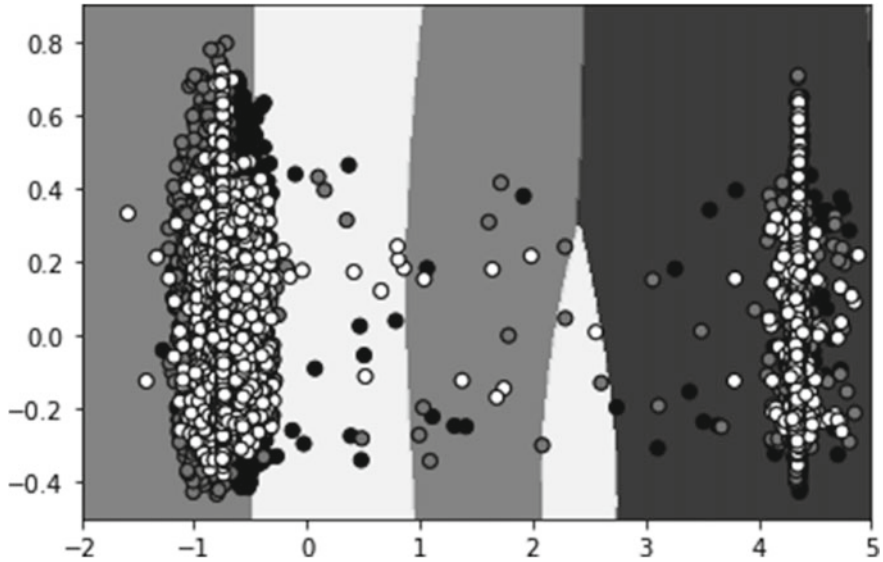
**Fig. 4** Decision boundaries of SVM on a subset of the data

a cycle. This research utilizes a feedforward ANN called the multi-layer perceptron (MLP).

Each neuron in the MLP performs computations based on the weighted sum of its inputs and a non-linear activation function. This research uses the rectified linear unit (ReLU) as the activation function. ReLU will output the input, $i$, if it is positive, and 0 otherwise. Research has shown that ANN is one of the most applied algorithms for maintenance [20]. A visual representation of the decision boundaries for the ANN approach in the CMM on a subset of data is shown in Fig. 5.

### 3.4 Model Stacking

Once RF, SVM, and ANN have been trained on the training set of data, we use a model stacking approach to combine them into an ensemble model for improved results. LR is used as the meta-model to facilitate this hybridization. As a statistical model designed to analyze independent variables that lead to an outcome, LR is an ideal choice for a meta-model, as it can learn the non-linear relationships between the predictions of each model being stacked and the actual value.

The LR model processes the three sets of predictions generated by RF, SVM, and ANN as input features, and the actual values as output features. The most accurate predictions from the three ML models are identified, and combined by the LR meta-model. The goal of this approach is to improve on the overall performance metrics for each of the three ML models (Fig. 6).
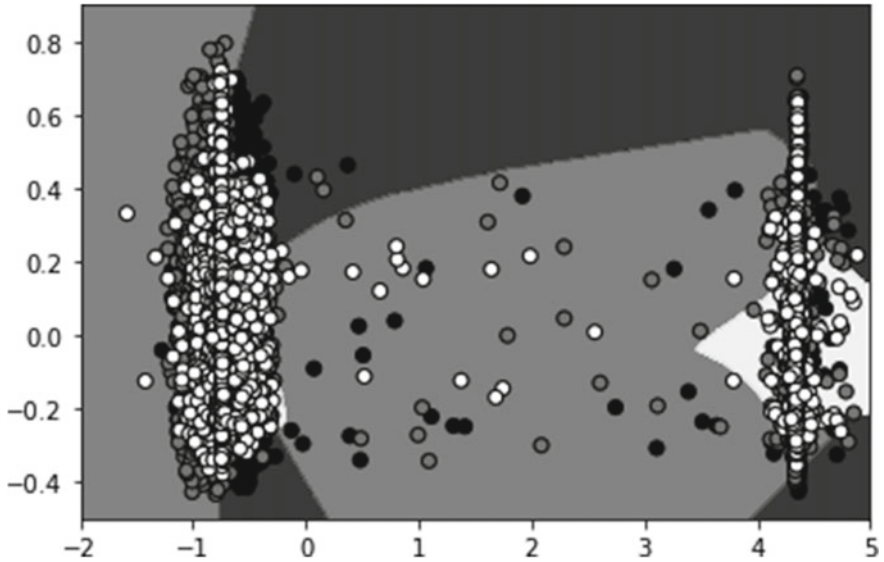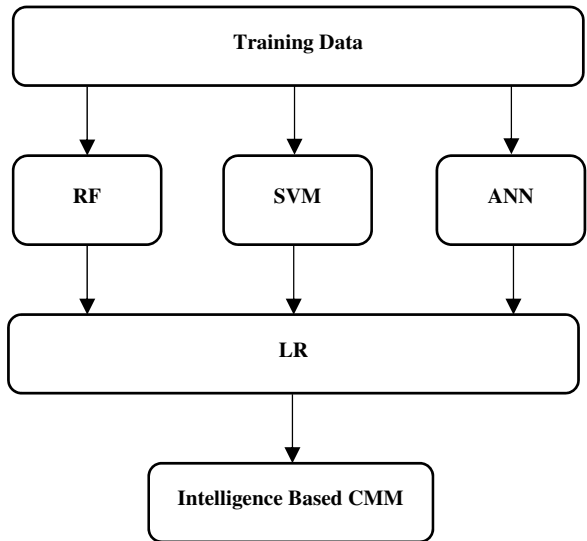
**Fig. 5** Decision boundaries of ANN on a subset of the data

**Fig. 6** Block diagram of the model stacking approach



## 4 Results

Table 1 shows the performance parameters of the ensemble model compared to traditional machine learning methods. The performance metrics used are given by the following formulas.

**Table 1** Accuracy of ensemble model versus traditional methods

Fault detection and diagnosis

| Classifier | Accuracy (%) | Recall (%) | Precision (%) | F1 (%) | Jaccard (%) |
|---|---|---|---|---|---|
| LR | 28.11 | 28.40 | 28.23 | 24.75 | 15.14 |
| GNB | 74.30 | 74.20 | 75.30 | 73.69 | 60.76 |
| GBM | 76.54 | 76.44 | 76.95 | 76.42 | 63.94 |
| SVM | 78.33 | 78.21 | 79.32 | 78.51 | 67.37 |
| KNN | 82.77 | 82.67 | 83.21 | 82.82 | 72.48 |
| DT | 88.13 | 88.03 | 88.01 | 88.02 | 79.74 |
| ANN | 89.99 | 89.90 | 89.91 | 89.82 | 82.62 |
| RF | 91.85 | 91.78 | 91.80 | 91.76 | 85.61 |
| Ensemble | 92.38 | 92.31 | 92.27 | 92.27 | 86.40 |

(1) Accuracy: the percentage of true positive (TP) and true negative (TN) predictions, with false positive (FP) and false negatives (FN) included:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

(2) Precision: the percentage of correctly classified positive predictions relative to predictions classified as positive:

$$P = \frac{TP}{TP + FP} \tag{2}$$

(3) Recall: the percentage of true positive predictions that were correctly classified:

$$R = \frac{TP}{TP + FN} \tag{3}$$

(4) f1-Score: the harmonic mean of the precision and recall:

$$f1 = 2 \times \frac{P \times R}{P + R} \tag{4}$$

(5) Jaccard Score: the similarity of predicted labels (PL) to true labels (TL):

$$J(PL, TL) = \frac{PL \cap TL}{PL \cup TL} \tag{5}$$

Specifically, the ensemble model is compared to traditional implementations of LR, GNB, GBM, SVM, KNN, DT, ANN, and RF.

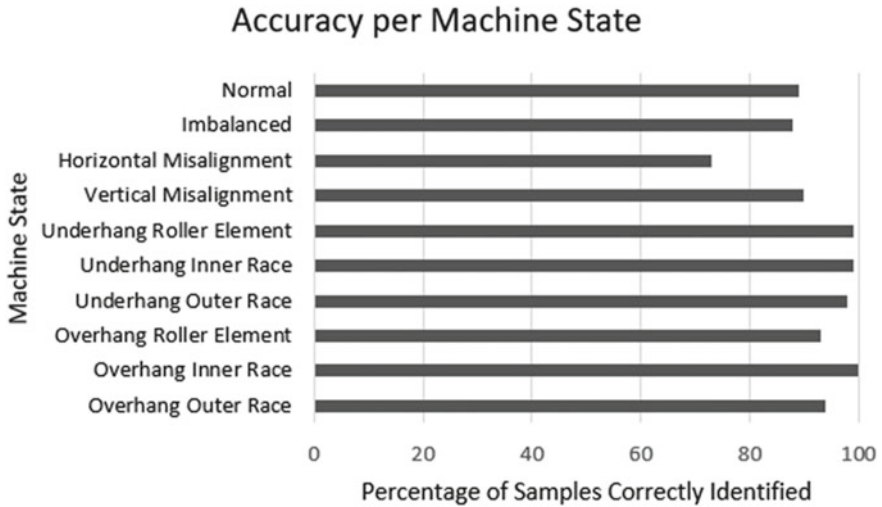A bar chart showing the accuracy breakdown per machine state is shown in Fig. 7.

**Fig. 7** Bar graph showing the accuracy per machine state

## 5    Conclusion

In this research, we proposed an ensemble machine-learning model to classify data points from sensors on a rotating machine. With this ensemble model, we enhance maintenance planning by adequately detecting and diagnosing faults in the rotating machine. When tested against traditional machine learning methods of LR, ANN, GBM, SVM, DT, and RF, the ensemble method proposed in this research demonstrated improved accuracy, recall, precision, F1-score, and Jaccard score on every traditional method. As a result, the ensemble model can reduce unnecessary maintenance and enhance automated decision-making. For future work in this area, the model stacking technique described in this paper can be applied to different use cases, and combinations of different machine learning algorithms can be tested to attempt to improve the accuracy of the ensemble model further.

## References

1. Shin J-H, Jun H-B (2015) On condition based maintenance policy. J Comput Des Eng 2(2):119–127
2. Gackowiec P (2019) General overview of maintenance strategies–concepts and approaches. Multidiscip Asp Prod Eng 2(1):126–139

3. Okoh C, Roy R, Mehnen J (2017) Predictive maintenance modelling for through-life engineering services. Procedia CIRP 59:192–201
4. Çınar ZM, Nuhu AA, Zeeshan Q, Korhan O, Asmael M, Safaei B (2020) Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. Sustainability 12(19):8211
5. Goyal D, Pabla BS (2015) Condition based maintenance of machine tools–a review. CIRP J Manuf Sci Technol 10:24–35
6. Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK (2020) Applications of machine learning to machine fault diagnosis: a review and roadmap. Mech Syst Signal Process 138:106587
7. Zhong K, Han M, Han B (2020) Data-driven based fault prognosis for industrial systems: a concise overview. IEEE/CAA J Automatica Sinica 7(2):330–345
8. Accorsi R, Manzini R, Pascarella P, Patella M, Sassi S (2017) Data mining and machine learning for condition-based maintenance. Proced Manuf 11:1153–1161
9. Pavlyshenko B (2018) Using stacking approaches for machine learning models. In: 2018 IEEE second international conference on data stream mining & processing (DSMP)
10. Kanaway A, Sane A (2017) Machine learning for predictive maintenance of industrial machines using IoT sensor data. In: 2017 8th IEEE international conference on software engineering and service science (ICSESS)
11. Peng L, Xu W, Jianglin X (2021) Research on fault diagnosis method of roller bearing based on model stacking. In: 2021 IEEE vehicle power and propulsion conference (VPPC)
12. Rajadurai H, Devi Ghandi U (2020) A stacked ensemble learning model for intrusion detection in wireless network. Neural Comput Appl 34:15387–15395
13. Wang X, Han T (2020) Transformer fault diagnosis based on stacking ensemble learning. IEEJ Trans Electr Electron Eng 17(2):400–417
14. Khadersab A, Shivakumar S (2018) Vibration analysis techniques for rotating machinery and its effect on bearing faults. Proced Manuf 20:247–252
15. Luo C, Mo Z, Wang J, Jiang J, Dai W, Miao Q (2020) Multiple discolored cyclic harmonic ratio diagram based on meyer wavelet filters for rotating machine fault diagnosis. IEEE Sens J 20(6):3132–3141
16. Ribeiro FML (2016) MAFAULDA:: Machinery fault database. https://www02.smt.ufrj.br/~off shore/mfs/page_01.htmlSEC2. Accessed 25 Jan 2023
17. Kowalczyk A (2017) "Kernels". In: Support vector machines succinctly. Syncfusion, pp 72–81
18. Huang H-Z, Wang H-K, Li Y-F, Zhang L, Liu Z (2015) Support vector machine-based estimation of remaining useful life: current research status and future trends. J Mech Sci Technol 29(1):151–163
19. da Silva IN, Hernane Spatti D, Andrade Flauzino R, Liboni LHB, dos Reis Alves SF (2017) Artificial neural network architectures and training processes. In: Artificial neural networks. Springer, pp 21–28
20. Zonta T, André da Costa C, da Rosa Righi R, José de Lima M, Silveira da Trindade E, Pyng Li G (2020) Predictive maintenance in the Industry 4.0: a systematic literature review. Comput Indust Eng 150:106889