

Review

# Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0

Zeki Murat Çınar <sup>1</sup>, Abubakar Abdussalam Nuhu <sup>1</sup>, Qasim Zeeshan <sup>1,\*</sup>, Orhan Korhan <sup>2</sup>,  
Mohammed Asmael <sup>1</sup> and Babak Safaei <sup>1,\*</sup>

<sup>1</sup> Department of Mechanical Engineering, Eastern Mediterranean University, Famagusta 99628, North Cyprus via Mersin, Turkey; cinar.zekimurat@gmail.com (Z.M.Ç.); abubakar.nuhu@emu.edu.tr (A.A.N.); mohammed.asmael@emu.edu.tr (M.A.)

<sup>2</sup> Department of Industrial Engineering, Eastern Mediterranean University, Famagusta 99628, North Cyprus via Mersin, Turkey; orhan.korhan@emu.edu.tr

\* Correspondence: qasim.zeeshan@emu.edu.tr (Q.Z.); babak.safaei@emu.edu.tr (B.S.)

Received: 15 August 2020; Accepted: 25 September 2020; Published: 5 October 2020

**Abstract:** Recently, with the emergence of Industry 4.0 (I4.0), smart systems, machine learning (ML) within artificial intelligence (AI), predictive maintenance (PdM) approaches have been extensively applied in industries for handling the health status of industrial equipment. Due to digital transformation towards I4.0, information techniques, computerized control, and communication networks, it is possible to collect massive amounts of operational and processes conditions data generated from several pieces of equipment and harvest data for making an automated fault detection and diagnosis with the aim to minimize downtime and increase utilization rate of the components and increase their remaining useful lives. PdM is inevitable for sustainable smart manufacturing in I4.0. Machine learning (ML) techniques have emerged as a promising tool in PdM applications for smart manufacturing in I4.0, thus it has increased attraction of authors during recent years. This paper aims to provide a comprehensive review of the recent advancements of ML techniques widely applied to PdM for smart manufacturing in I4.0 by classifying the research according to the ML algorithms, ML category, machinery, and equipment used, device used in data acquisition, classification of data, size and type, and highlight the key contributions of the researchers, and thus offers guidelines and foundation for further research.

**Keywords:** predictive maintenance; artificial intelligence; machine learning; industrial maintenance

---

## 1. Introduction

Industries are currently going through “The Fourth Industrial Revolution,” as professionals have called it, a term also known as “Industry 4.0.” (I4.0) Integration amongst physical and digital systems of the production contexts is what mainly concerns Industry 4.0 [1]. With the appearance of I4.0, the concept of prognostics and health management (PHM) has become unavoidable tendency in the framework of industrial big data and smart manufacturing; plus, at the same time, it offers a reliable solution for handling the industrial equipment health status. I4.0 and its key technologies play an essential role to make industrial systems autonomous [2,3] and thus make possible the automatized data collection from industrial machines/components. Based on the collected data type machine learning algorithms can be applied for automated fault detection and diagnosis. However, it is very cruel to select appropriate machine learning (ML) techniques, type of data, data size, and equipment to apply ML in industrial systems. Selection of inappropriate predictive maintenance (PdM) technique, dataset, and data size may cause time loss and infeasible maintenance scheduling. Therefore, this study aims to present a comprehensive literature review to discover existing studies

and ML applications, and thus help researchers and practitioners to select appropriate ML techniques, data size, and data type to obtain a feasible ML application.

The industrial equipment predictive maintenance (PdM) can perceive the degradation performance because it was designed to achieve near-zero; hidden dangers, failures, pollution, and near-zero accidents in the entire environment of manufacturing processes [4].

These huge amounts of data collected for ML contains very useful information and valuable knowledge which can improve the whole productivity of manufacturing processes and system dynamics, and can also be applied into decision support in several areas, mainly in condition-based maintenance and health monitoring [5]. Due to the recent advances in technology, information techniques, computerized control, and communication networks, it is now possible to collect vast volumes of operational and processes conditions data generated from several pieces of equipment in order to be harvested in making an automated Fault Detection and Diagnosis (FDD) [6]. The datasets collected can also be applied to develop more efficient methodologies for the intelligent preventive maintenance activities, similarly known as PdM [7].

ML applications provide some advantages which include maintenance cost reduction, repair stop reduction, machine fault reduction, spare-part life increases and inventory reduction, operator safety enhancement, increased production, repair verification, an increase in overall profit, and many more. These advantages also have a tremendous and strong bond with the procedures of maintenance [1,8–11]. Moreover, fault detection is one of the critical components of predictive maintenance; it is very much needed for industries to detect faults at very early stage [12]. Techniques for maintenance policies can be categorized into the following main classifications [13–17].

1. (R2F) Run 2 Failure: also known as corrective maintenance or unplanned maintenance. It is the simplest amongst maintenance techniques which is performed only when the equipment has failed. It may lead to high equipment downtime and a high risk of secondary faults and thus, create a very large number of defective products in production.
2. Preventive Maintenance (PvM): also known as scheduled maintenance or time-based maintenance (TBM). PvM refers to periodically performed maintenance based on a planned schedule in order to anticipate the failures. It sometimes leads to unnecessary maintenance which increase the operating costs. The main aim here is to improve the efficiency of the equipment by minimizing the failures in production [18].
3. Condition-based Maintenance (CBM): this method of maintenance is based on a constant machine or equipment monitoring or their process health that can be carried out only when they are actually necessary. The maintenance actions can only be carried out when the actions on the process are taken after one or more conditions of degradation of the process. CBM usually cannot be planned in advance.
4. PdM: known as Statistical-based maintenance: maintenance schedules are only taken when needed. It is based on the continuous monitoring of the equipment or the machine, as like CBM. It utilizes prediction tools to measure when such maintenance actions are necessary, hence the maintenance can be scheduled. Furthermore, it allows failure detection at an early stage based on the historical data by utilizing those prediction tools such as machine learning methods, integrity factors (such as visual aspects, coloration different from original, wear), statistical inference approaches, and engineering techniques.

It is required that any maintenance strategy ought to minimize equipment failure rates, must improve equipment condition, should prolong the life of the equipment, and reduce the maintenance costs. An overview for the maintenance classifications is shown in Figure 1. PdM turned out to be one of the most promising strategies amongst other strategies of maintenance that has the ability of achieving those characteristics [19], thus the strategy has been applied recently in many fields of studies. PdM captivates the attention of the industries, hence it has been applied in the era of I4.0 due to its capability of optimizing the use and management of assets [1,20].

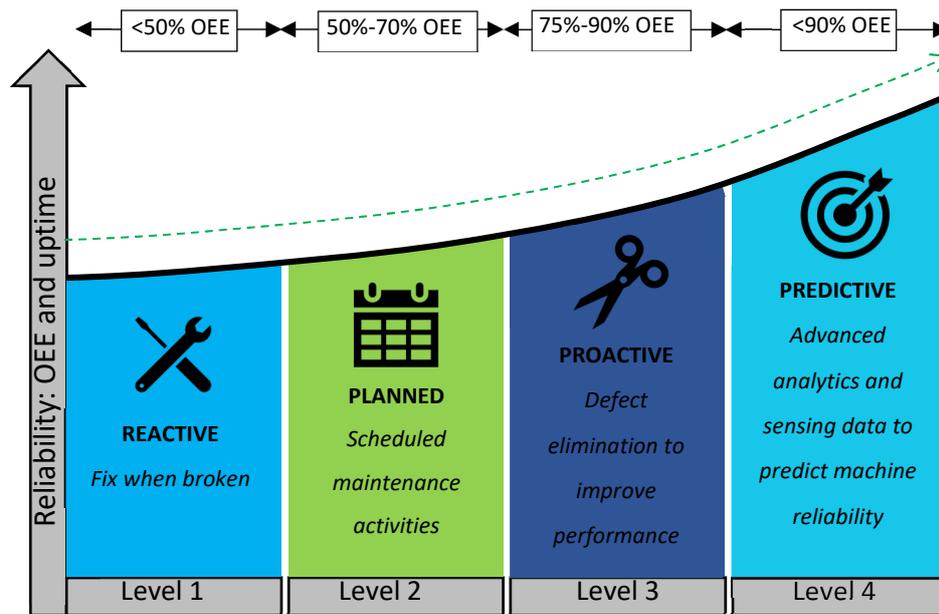


Figure 1. Maintenance types [1].

ML, within the contexts of artificial intelligence (AI) (Figure 1, copyright permission of Figure 1 has taken on 20 September 2020), lately, has appeared to be one of the most powerful tools that can be applied in several applications to develop intelligent predictive algorithms. It has been developed into a wide field of research over the past decades. ML can be defined as a technology by which the outcomes can be forecasted based on a model prepared and trained on past or historical input data and its output behavior [21]. According to Samuel, A.L. [22], ML mainly means that if computers are allowed to solve without specifically being programmed in doing so. ML approaches are known to have tremendous advantages, as they have the ability in handling multivariate, high dimensional data and can extract hidden relationships within data in complex, dynamic, and chaotic environments [1,23,24]. However, depending on the ML approach chosen, the performance and advantages might differ. As of today, ML techniques have been widely applied in several areas of manufacturing (such as maintenance, optimization, troubleshooting, and control) [23]. Consequently, this paper aims to provide the recent advancements of ML techniques applied to PdM from an ample perspective. Predominantly, this ample review uses Scopus database while acquiring and identifying the articles used. From a comprehensive perspective, this paper aims to pinpoint and categorize based on the ML technique considered, ML category, equipment used, device used in data acquiring, applied data description, data size, and data type.

The following describes how the paper is organized: firstly, this section gives a brief introduction on the current field of study. Secondly, Section 2 presents a brief background on PdM and ML techniques. Thirdly, Section 3 explains the methodology employed in the literature while considering and categorizing the papers to review and how they are grouped. Section 4 presents the comprehensive applications of ML techniques applied to PdM. Subsequently, discussions are drawn based on the analysis carried out in the literature of ML algorithms for PdM. Finally, a conclusion and future research guidelines are given.

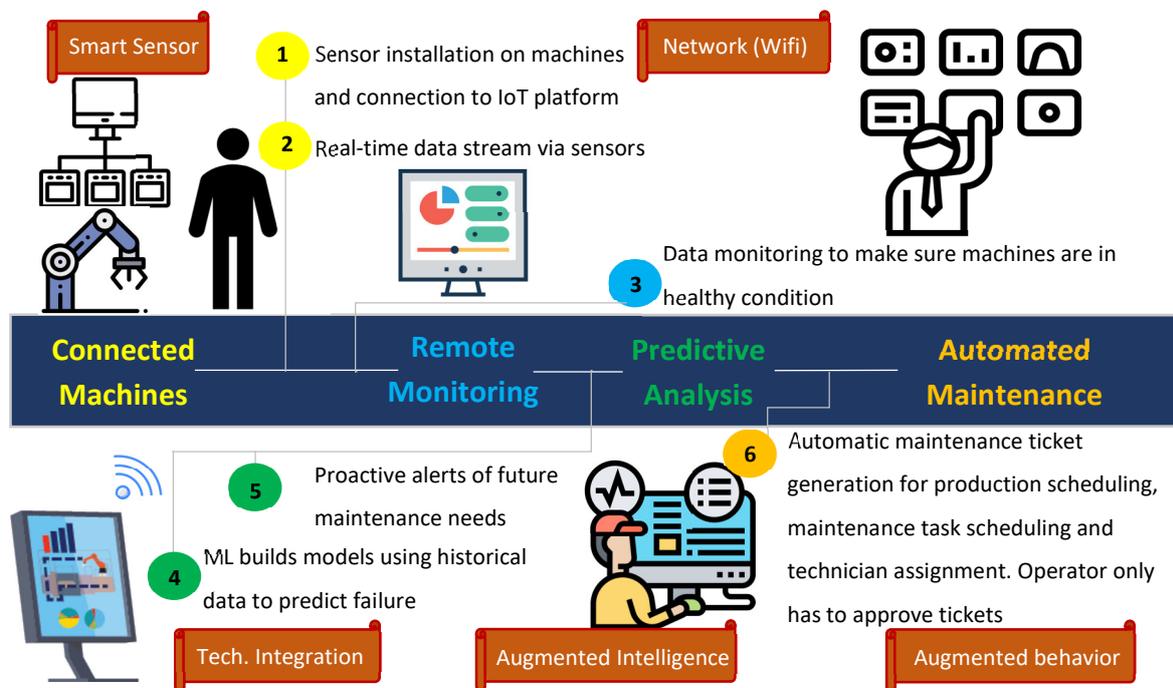
## 2. PdM and ML Techniques

Currently, the PHM system has become a safe-fire method for maintaining the safety status of equipment (e.g., defect detection and Remaining Useful Life (RUL)). It is accomplished by the systematic use of the current testing findings in AI technology and IT technology. [4]. Additionally, PdM cannot only provide reduction in the costs of the maintenance, it can also prolong the RUL [25]. The incipient issues that may lead to disastrous failures can be correctly forecast and appropriate steps can be set in order to avoid these failures on the basis of the prediction outcomes [4].

Nevertheless, at any appropriate time, an industrial equipment can be replaced or repaired before the fault happens, and thus might restore the original condition of the equipment or system after each and every completed maintenance. Moreover, the equipment, component, system, or a machine's health status can be obtained at any instance. Their failures can also be predicted in order to achieve a non-zero downtime performance [26]. PdM mainly focuses on utilizing predictive info in order to accurately schedule the future maintenance operations [27].

The aim of PdM is not only to collect process data and its parameters, but also to collect the physical health aspects of the equipment, machine, or component (such as pressure, vibration, temperature, viscosity, acoustics, viscosity, flow rate data) and many as such. At the same time, this information collected is now widely used for fault identification, early fault detection, equipment health assessment, and predicting the future state of the equipment [12].

According to [12], ML is a subcategory of AI and is defined as any algorithm or a program that has the ability of learning with the smallest or no additional support. ML assists in solving many difficulties such as in vision, big data, robotics, and speech recognition [12]. Moreover, ML techniques are designed to derive knowledge out of existing data [23,28]. PdM process and technologies to drive PdM is shown in Figure 2. Crucial technologies for PdM such as smart sensors, network, artificial intelligence, big data, and cloud systems are also highlighted by [29,30].



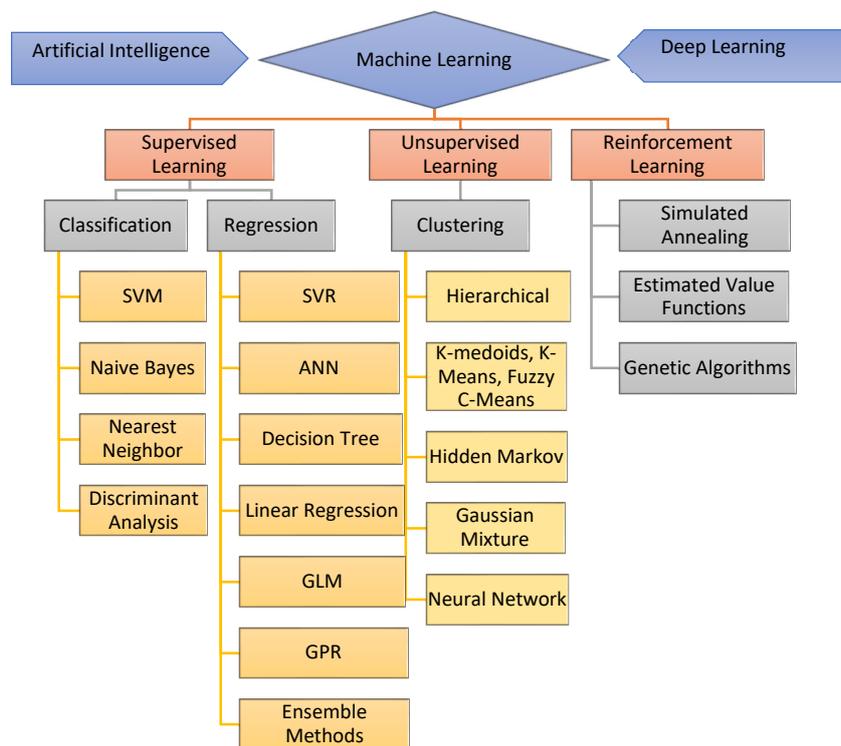
**Figure 2.** PdM process and technologies to drive PdM.

Deloitte [31] classifies the technologies that drive PdM into five different categories. Those are sensors, network, integration, augmented intelligence, and augmented behavior. Smart sensors are used to gather machine information with the use of built-in sensors or environment information with the use of external sensors. The network provides data storage as well as data transfer by using Bluetooth and WiFi [32,33]. Technology integration allows data management and data accumulation via Internet of Things (IoT), augmented intelligence assist data processing, and data analytics [30], whereas augmented behavior allows virtualization, computing and service platform via apps, and tickets to assist the operator [29].

ML algorithms are categorized into three different types; (1) supervised, (2) unsupervised, and (3) reinforcement learning (RL) (see Figure 3) [23,34,35]. The aim is to show how complex the structure can be and the commonly used available learning techniques. Moreover, as stated by [23], different algorithms can be combined together in order to maximize the classification power. To add on, some among the ML algorithms are both applicable to unsupervised and supervised learning.

In unsupervised machine learning, there is no feedback from an external teacher or knowledgeable expert [23]. Based on the existing data, the algorithm identifies the clusters. The main aim here in supervised learning is determining the unknown classes of items by clustering [36], whereas classification is for supervised learning. Unsupervised ML basically defines any ML method that attempts to learn structure in the absence of either an identified output (like supervised ML) or feedback (like Reinforcement learning (RL)) [23].

Clustering, self-organizing maps, and association rules are the basic three main examples of unsupervised learning. In this paper, reviewed articles are categorized into three ML categories, classification, regression, and clustering, as shown in Figure 3. Data utilized in papers are categorized into two data types, real data that are taken from real world machineries, and simulated or synthetic data that are generated to meet specific needs such as model validation in ML. Authors [12,23,32,37,38] has also defined ML categories.



**Figure 3.** Classifications within Machine Learning Techniques.

RL is characterized by the delivery of information on training to the community. Through RL, the learner must discover which actions produce the greatest outcomes (numeric reinforcement signal) by attempting instead of being told. [23]. Nonetheless, some of the researchers considered RL as some sort of special supervised learning, like [34]. Moreover, as stated by RL, problems differ from supervised learning, as they can be described by the absence of labeled examples of “good” and “bad” behavior [23].

There are several available supervised machine learning algorithms, as few can be seen from Figure 3. Each of these algorithms has its own specific advantages as well as limitations regarding the application (either PdM or manufacturing). Selecting the most appropriate and suitable ML algorithm can be a major challenge for the requirements of the PdM problem. It is also important to get good at applied machine learning by practicing on lots of different datasets. Therefore, each problem requires different subtlety, different data preparation, and modelling methods. Datasets are classified into seven categories: multivariate, sequential text, time series, sequential, univariate, text, and domain theory. However, this paper classifies datasets into two categories. One of them is real datasets that are any production data obtained from real production processes and applicable to ML.

Another one is synthetic datasets that are any kind of production data applicable to ML, but they are simulated data rather than direct measurement in the production.

Ethical/legal permission is not required for this study. The study complies with research and publication ethics in obtaining all kinds of data and images.

### 3. Survey Methodology and Analysis

The scholarly or academic databases used for this review include articles from Scopus, ScienceDirect, Institute of Electronic and Electrical Engineers (IEEE), and Google Scholar. The Scopus database was mainly used for this review. In this review, the articles reviewed are categorized into two groups. The first group comprises the articles collected from Scopus and are considered as the main featured articles of the research. The second group comprises the articles that are used as supporting or background work in the contexts of introduction and the study in general. The second group of the articles that are obtained from the four databases stated above helped in building the theoretical foundations to the PdM, ML techniques, and the ML algorithms.

Strategy and keywords used while collecting the article from Scopus are as follows:

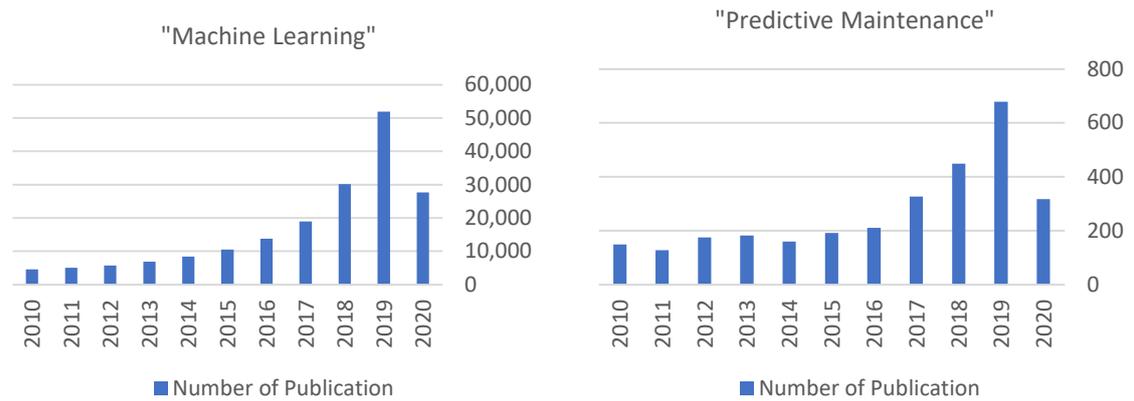
- Firstly, the search was carried out based on “machine learning.”
- Followed by search within search, with “predictive maintenance,” note that the use of quotation here means that the whole phrase was searched entirely, not as separate word by word.

With (TITLE-ABS-KEY (“machine learning”) AND (“predictive maintenance”)), as the search keywords, 788 documents appeared, and the survey was carried out on 30 July 2020.

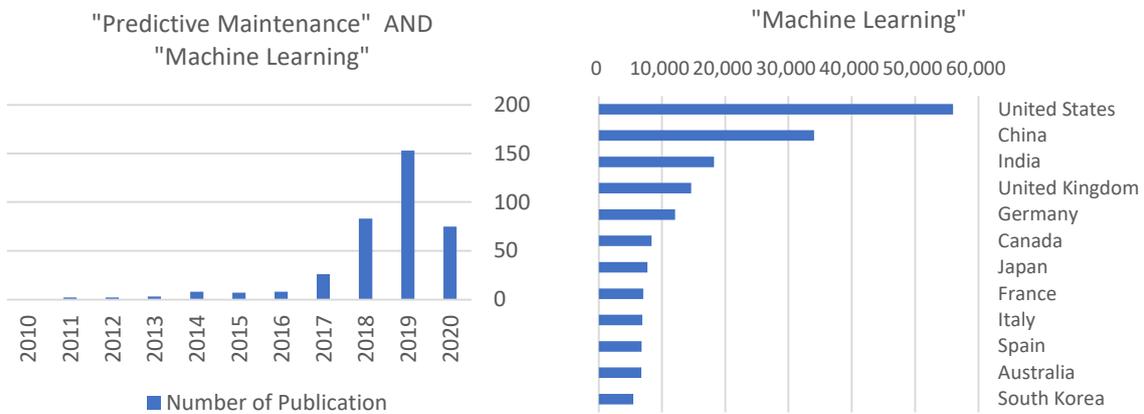
- The documents were then limited to the recent time parameters of 2010 to 2020. All inclusively, the number of the documents reduced to 367.
- Subject area was limited to engineering, energy, and materials science, then the documents reduced to 273.
- Then, from the document type, review and conference review were excluded from the analysis. That left a total of 217 documents. From the language section, the documents were limited to English.

Figure 4 shows the number of documents that are published over the years, between 2010 and 2020. As can be seen from Figure 4, it is confirmed that just recently (i.e., from the last three years) ML techniques in PdM captivates the attention of the researchers. As there are very few papers published in 2010 and 2011 in comparison to how the number of published articles just spiked-up through the year 2017–2018. Thus, it can be concluded that application of ML technique in the field of PdM is a new method with a growing interest in the field of research. This might be due to the increase in the amount of dataset that are generated in industries, by the industrial equipment, system, or components, and at the same time it could be due to the recent advances of ML techniques and their algorithms [1]. Figure 5 shows the information on how the documents are published by type considered in this review, 134 conference papers, 72 articles, and 11 book chapters.

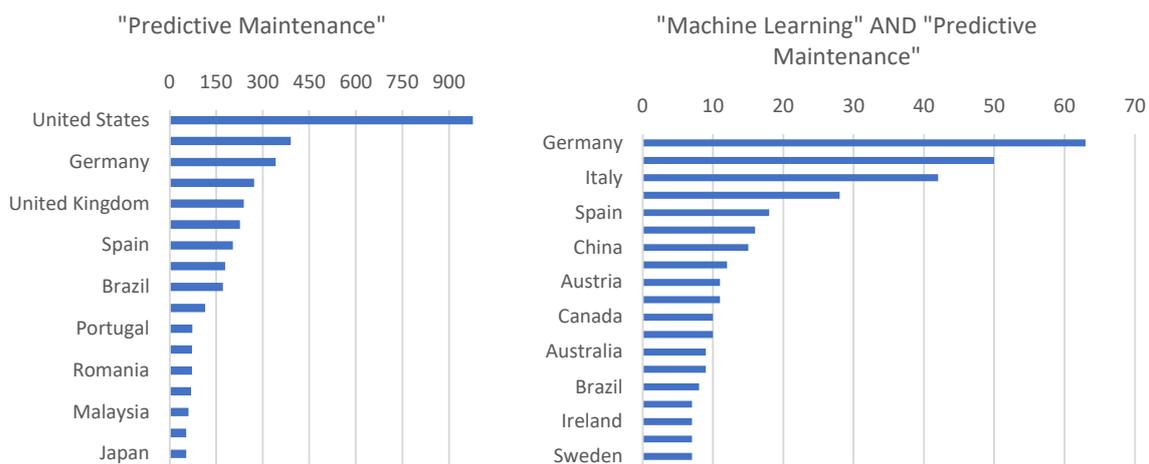
Among these 217 documents, we were able to download 103. These 103 articles and 33 from IEEE, combined together gives a total of 126 articles. Then, they are analyzed and screened for the removal of any duplicates. The number of publications are given in Figure 4. Article selection criteria aforementioned can be summarized and pictured from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart given in Figure 5.



(a) (b)



(c) (d)



(e) (f)

**Figure 4.** Statistics from Scopus database (Date: 30.07.2020), (a) Published documents per year with keywords: (“Machine Learning”), (b) Published documents per year with keywords: (“Predictive Maintenance”), (c) Published documents per year with keywords: (“Machine Learning” AND “Predictive Maintenance”), (d) Published documents by country with keywords: (“Machine Learning”), (e) Published documents by country with keywords: (“Predictive Maintenance”), (f)

Published documents by country with keywords: (“Predictive Maintenance” AND “Machine Learning”).

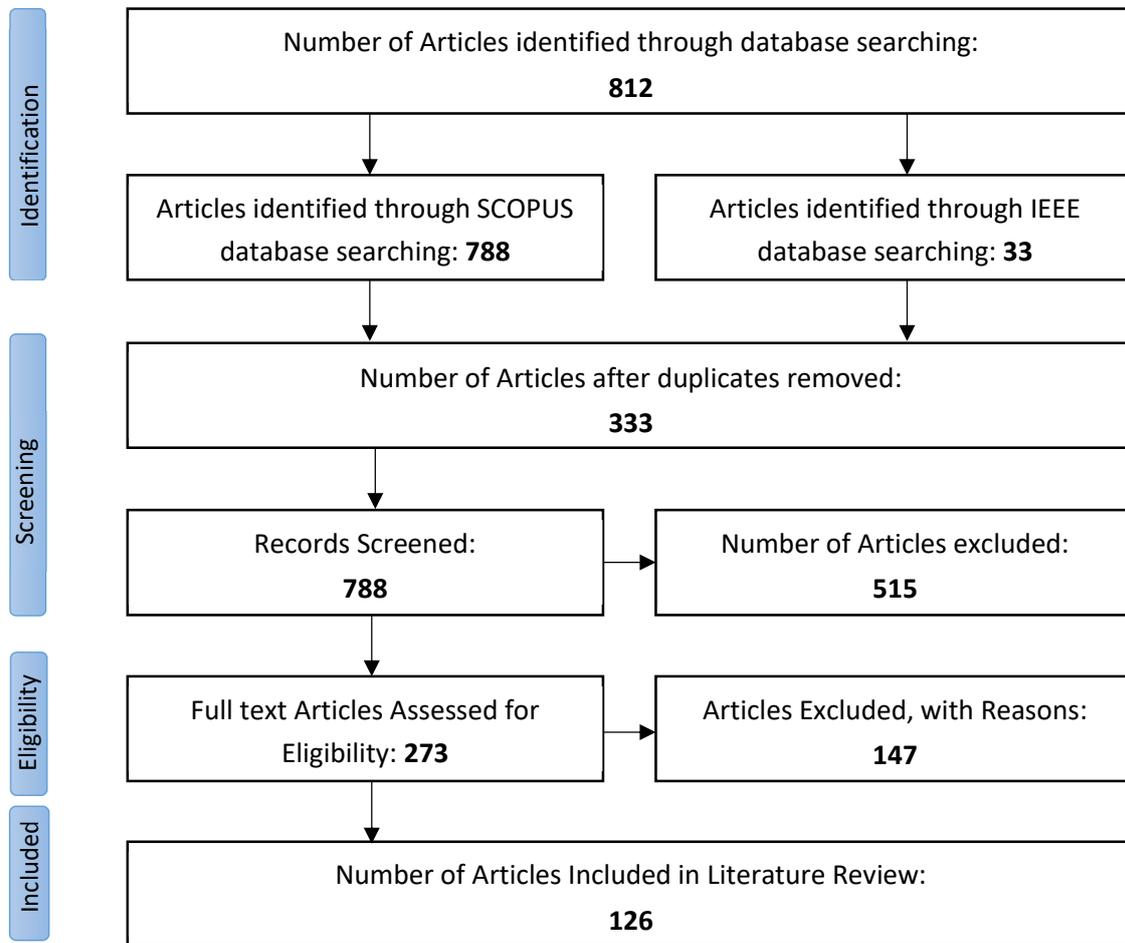


Figure 5. Literature survey flowchart.

#### 4. Applications of ML Algorithms in PdM

ML algorithms can be used to solve several problems with the enormously available data generated from industries, thus, ML has been widely used in computer science and other areas, such as PdM of manufacturing system, tool or machine, and is one of the possible areas of use for data-driven methods (Artificial Neural Network (ANN), Reinforcement Learning (RF), Support Vector Machine (SVM), Logistics Regression (LR), and Decision Tree (DT)) [4,23]. Recently, ML techniques have been widely applied in various fields of study. Selecting the most appropriate, simple, and the most efficient could be of a great concern. ML algorithms usually require collecting huge amounts of data of the failure status scenarios and the health conditions scenarios for model training. These algorithms that mainly require large amounts of data involves Vector Space Model (VSM), LR, DT, and RF. ML algorithm development covers historical data selection, pre-processing data, model selection, model training, model validation, and maintenance. The steps involved in ML algorithm development can be specified as input, feature extraction & selection, features, traditional ML techniques, and output [4]. Similarly, [1] describes the main steps for ML development as historical data, data pre-processing, model selection, training and validation, and model maintenance. Further details on the main steps involved can be found in [1]. PdM has been broadly applied in industries such as manufacturing industries using ML techniques [39] and deep learning [40].

#### 4.1. Artificial Neural Network (ANN)

In fact, ANN is developed from the subject of biology, where the Neural Network (NN) plays a significant role in the human brain. [41]. ANN is an intelligent computational technique that has been inspired by biological neurons [10]. It is a massively parallel computing system consisting of an extremely large number of simple processors with many interconnections. Instead of following the set of laws specified by human experts, ANNs learn the basic laws from the set of given symbolic situations in examples [42]. They are organized in three layers or more, (i.e., input layer, several hidden layers, and an output layer) [43]. Moreover, the analytical activity of these ANNs derives from the relations between the network processing units.

ANN models are broadly applied in many fields of studies due to their capability to learn from examples. To add on, ANNs models in comparison to the other traditional machine learning algorithms have noticeable advantages in addressing random data, fuzzy data, and nonlinear data. ANNs are primarily appropriate for systems with a complex, large scale structure and unclear information [4]. ANNs are widely applied and they are the most common ML algorithms [1], at the same time they have been suggested in several industrial applications involving soft sensing [44], and in predictive control systems [45]. Hesser, D.F. and Markert, B. [46] trained an ANN model to classify tool state of a Computer Numerical Control (CNC) milling machine with acceleration data. The proposed study was based on a retrofitting approach in order to facilitate older machines towards to I4.0. The tool wear was monitored by utilizing a programmable prototyping platform equipped with built-in sensors. The study proves the feasibility of retrofitting older machines. In the study, the performance of the built model was compared and outperformed the performance of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) models.

A methodology was proposed by Sampaio, G.S. et al. [47] to treat and convert the collected data of vibration measurements from a vibration system that simulated a motor and to build a dataset in order to train and test an ANN model capable of predicting the future condition of the equipment, predicting when a failure can happen. The methodology involves the use of frequency and amplitude data by classifying the dataset and defining a way of calculating the vibrating system's failure time. In the study, Multilayer Perceptron (MLP) methodology was used in performing the prediction task, due to its easier implementation with a good generalization index. The ANN model proposed was then compared in terms of its efficiency and based on Root Mean Square Error (RMSE) performance index values with other ML techniques, including Regression Tree (RT), Random Forest (RF), and Support Vector Machine (SVM). Comparative and training results were adequate and showed that ANN was greater than the others. In terms of medium-term and long-term prediction, ANN outperforms the others, whereas generalization in short term predictions between ANN, RF, and RT were equal.

ANN and SVM ML algorithms are applied in developing gauge degradation measurements prediction for two types of rail track including straight and curved segments by Falamarzi, A. et al. [48]. Mean squared error and coefficient of determination are used in the performance evaluation of the proposed models, ANN with greater than 0.9 coefficient of determination value. Based on the results obtained from the study, both ANN and SVM models provide satisfactory and slightly similar outcomes, but the performance of ANN models in predicting gauge deviation of straight segments is slightly better than SVM models. Biswal, S. and Sabareesh, G.R. [10] designed and developed a bench top test-rig for investigating the time domain vibration signatures of several critical components in wind turbine by imitating the operating condition of an actual wind turbine and use it for monitoring its condition. In their work, they acquired the healthy and faulty condition vibration signature of the critical components, then developed and applied ANN model to carry out the classification of the faulty and healthy state features. The model developed shows a 92.6% accuracy classification efficiency. Zhang, Y. et al. [49] reported a study on Physics-based Model and Neural Network Model for Monitoring Starter Degradation of Auxiliary Power Unit (APU). In their study, a generic modeling technique is adopted to overcome the limitations of lack of component characteristics. A comparative analysis between back-spreading and forward-feed neural network models has been performed, trained, and tested. Both models are applied under nominal and deteriorated conditions and their

capabilities are validated. Depending on the data collected, their analysis concluded that the physics-based approach produces more consistent outcomes for cases with degraded starters, although the neural network model showed better results with starters in healthy condition.

#### 4.2. Support Vector Machine (SVM)

SVM is a well-known ML technique which is widely used for both classification and regression analysis, due to its high accuracy [1,50,51]. SVM is defined as a statistical learning concept with an adaptive computational learning method. SVM learning algorithm is presented in Figure 6. SVM learning technique employs input vectors to map nonlinearly into a feature space whose dimension is high [52–54].

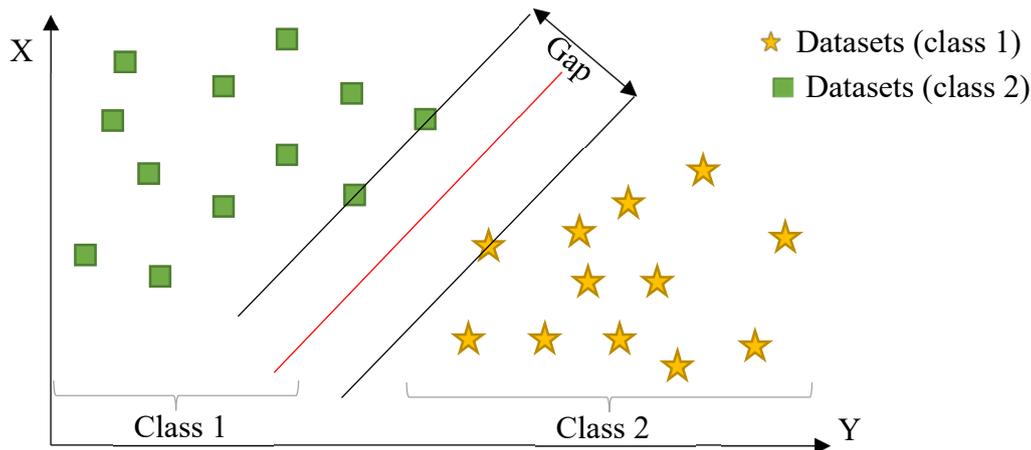


Figure 6. Support vector machine algorithm.

SVM is a supervised ML technique that can perform pattern recognition, classification, and regression analysis. In the PdM of industrial equipment, SVMs have been widely applied for identifying a specific status based on the acquired signal [55]. SVM and ANN ML algorithms are applied in developing gauge degradation measurements prediction for two types of rail track including straight and curved segments by Falamarzi, A. et al. [48], where mean squared error and coefficient of determination are used in the performance evaluation of the proposed models, SVM with greater than 0.75 coefficient of determination value. Based on the results obtained from the study, both ANN and SVM models provide satisfactory and slightly similar outcomes, but, the performance of SVM models in predicting gauge deviation of curved segments is slightly better than ANN models. Moreover, in the study, Melbourne tram network has been used as a case study.

A data driven diagnostics and prognostics framework for machines to increase efficiency and reduce maintenance cost was proposed by Xiang, S. et al. [56]. Moreover, an accurate data labeling methodology is developed for supervised learning via comparing the serial number of target components in the adjacent dates. In the study, vending machine real data was used to validate the proposed framework for three different classifiers including SVM, RF, and Gradient Boosting Machines (GBM). Moreover, two models were developed for PdM, one for diagnostics and the other for two-stage prognostics. Results for the cross-validated simulation obtained shows that the diagnostics model can achieve more than 80% of accuracy, thus the developed model of SVM can be applied for diagnosis and prognostics monitoring of complex vending machines. The prognostics model outperforms one-stage conventional prediction models.

#### 4.3. Decision Tree (DT)

Decision Tree is a network system composed primarily of nodes and branches, and nodes comprising root nodes and intermediate nodes. The intermediate nodes are used to represent a feature, and the leaf nodes are used to represent a class label [52]. DT can be used for feature selection [57]. DT algorithm is presented in Figure 7.

DT classifiers have gained considerable popularity in a number of areas, such as character identification, medical diagnosis, and voice recognition. More notably, the DT model has the potential to decompose a complicated decision-making mechanism into a series of simplified decisions by recursively splitting covariate space into subspaces, thereby offering a solution that is sensitive to interpretation [58,59].

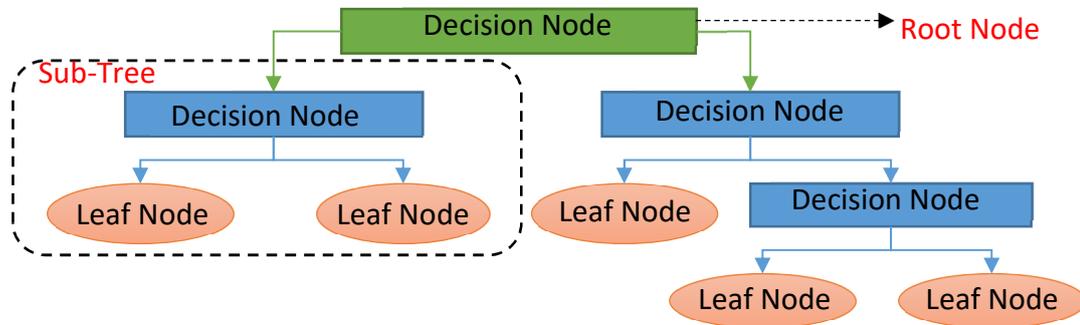


Figure 7. Decision tree algorithm, adapted from.

#### 4.4. Random Forest (RF)

RF was developed by Breiman, L. [60]. This is an ensemble learning algorithm made up of several DT classifiers, and the output category is determined collectively by these individual trees. When the number of trees in the forest increases, the fallacy in generalization error for forests converges. There are also important benefits of the RF. For example, it can manage high-dimensional data without choosing a feature; trees are independent of each other during the training process, and implementation is fairly simple; however, the training speed is generally fast and, at the same time, the generalization functionality is good enough [4]. Random forest algorithm for machine learning has tree predictions, and based on tree predictions, the RF provides random forest predictions [61]. The RF model is visualized in Figure 8.

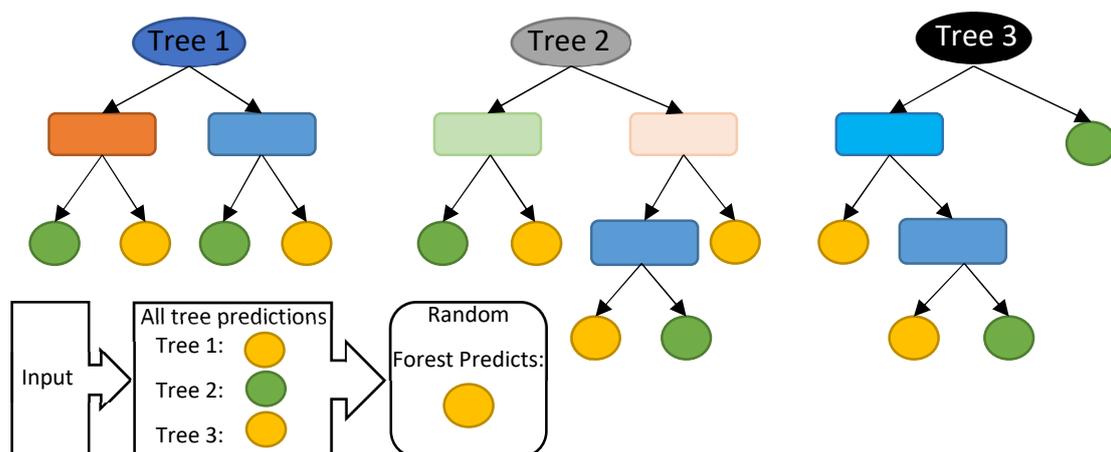


Figure 8. Random forest regression algorithm for machine learning.

A study was reported on forecasting the downtime of a printing machine based on real time predictions of imminent failures [62]. In their study, they used unstructured historical machine data to train the ML classification algorithms including RF, XGBoost, and LR to predict the machine failures. Different metrics were analyzed to determine the fitness of the models. These metrics include empirical cross-entropy, area under the receiver operating characteristic curve (AUC), receiver operating characteristic curve itself (ROC), precision-recall curve (PRC), number of false positives (FP), true positives (TP), false negatives (FN) and true negatives (TN) at various decision thresholds,

and calibration curves of the estimated probabilities. Based on the results obtained, in terms of ROC, all the algorithms performed significantly better and almost similar. But in terms of decision thresholds, RF and XGBoost perform better than LR.

ML algorithms including Linear Regression, RF, and Symbolic Regression (SR) are applied in modeling the condition of a healthy industrial machinery [63]. They proposed a methodology for detecting and predicting drifting behavior (called concept drifts) in continuous data streams. Further, a real-world case study was presented on industrial radial fans. Based on the results obtained using the synthetic data, both the results from concept drift detection and prediction are highly successful. Moreover, based on the conducted real word study, experts on-site at the strained radial fan reported that the principle of drift detection has been successfully deployed. However, due to the lack of continuous deterioration information, the predictability of concept drifts is currently based on assumptions and cannot be measured yet, even though results of the tests are already very promising.

Janssens, O. et al. [64] proposed a multi-sensor device that uses not only infrared thermal imaging data, but also uses vibration measurements for automatic conditioning and fault detection in rotating machines. The feature fusion is utilized where model-driven features are derived from vibration measurements and data-driven features are derived from infrared thermal imaging data. Then, the extracted features are combined together and presented to RF classifiers for actual fault detection. They have demonstrated in the study by mixing these two types of sensor data, a variety of conditions/faults and combinations can be measured more accurately than in the case of individual sensor streams.

Lacaille, J. and Rabenoro, T. [65] developed a learning algorithm that can automatically detect and analyze multidimensional datasets of turbofan engine. The developed model uses a very wide population of pre-treatments and statistic tests on the data and has the ability to select good combinations of tests with higher than 85% pre-identification. Quiroz, J.C. et al. [66] proposed a new approach to diagnose broken rotor bar failure in a Line Start-Permanent Magnet Synchronous Motor (LS-PMSM) using RF. The transient current signal during the motor startup was acquired from a healthy motor and a faulty motor with a broken rotor bar fault. The model was trained using features extracted from thirteen different statistical time-domain features, and these features were used in determining the state of the motor where it is operating under faulty or normal conditions. Feature importance was considered for their feature selection in order to reduce the number of features to very few from the RF. Results have shown that RF categorizes the motor disorder as safe or deficient with an accuracy of 98.8% using all the features and an accuracy of 98.4% using only the mean index and impulsion features. A comparison was carried out between the developed model and other traditional ML algorithms including Decision Tree (DT), Naive Bayes classifier (NBC), LR, linear ridge, and SVM, the RF consistently outperforms these algorithms with having a higher accuracy than the other algorithms. The suggested methodology can be used for electronic tracking and fault detection of LS-PMSM motors in the industry, and the findings can be beneficial for the development of preventive maintenance plans in factories.

Yan, W. and Zhou, J.H. [67] proposed a predictive model using Term Frequency-Inverse Document Frequency (TF-IDF) and RF can forecast faults of high sensitivity in advance by analyzing the historical data of aircraft maintenance systems, and preventive maintenance may be carried out on the basis of the model's prediction performance. TF-IDF has been employed in order to extract the features from raw data in the past consecutive flights. Different priorities were considered in classifying the faults by the proposed RF model. The ROC curve has been adopted as a performance metric as the dataset is highly imbalanced. Compared to the other method, the suggested approach reaches the maximum true positive rating of 100% and the lowest false positive rate of 0.13%. For the testing dataset, the proposed method achieves true positive rate 66.67% and false positive rate 0.13%.

#### 4.5. Logistic Regression (LR)

Binding, A. et al. [62] reported a study on forecasting the downtime of a printing machine based on real time predictions of imminent failures. In their study, they utilized unstructured historical machine data to train the ML classification algorithms including RF, XGBoost, and LR in predicting

the machine failures. Various metrics were analyzed to determine the goodness of fit of the models. These metrics include empirical cross-entropy, area under the receiver operating characteristic curve (AUC), receiver operating characteristic curve itself (ROC), precision-recall curve (PRC), number of false positives (FP), true positives (TP), false negatives (FN), and true negatives (TN) at various decision thresholds, and calibration curves of the estimated probabilities. Based on the results obtained, in terms of ROC, all the algorithms performed significantly better and almost similar. But in terms of decision thresholds, RF and XGBoost perform better than LR. Using a given set of independent variables, linear regression is used to estimate the continuous dependent variations. However, using a given set of independent variables, logistic regression is used to estimate the categorical contingent variations [68]. Graph of the linear regression model and logistics regression model are shown in Figure 9.

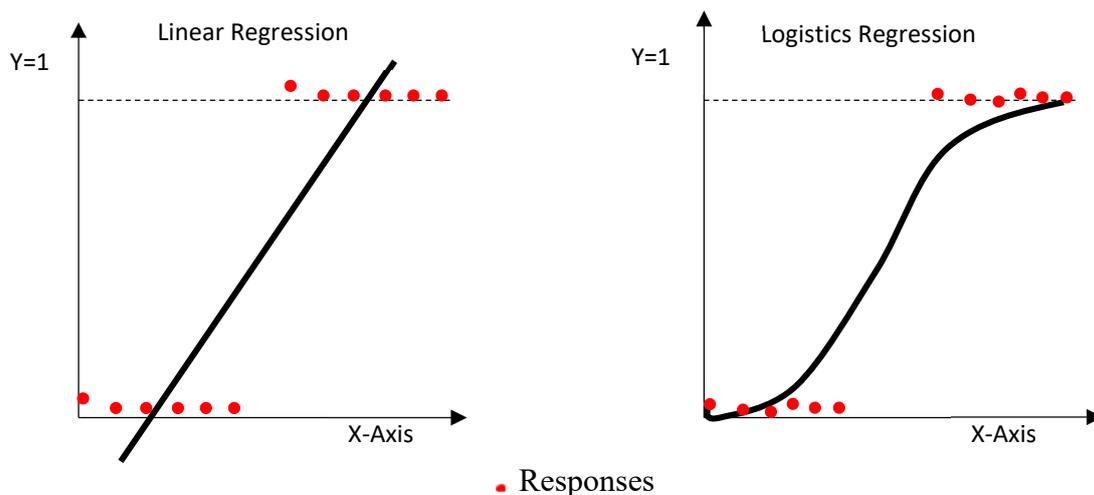
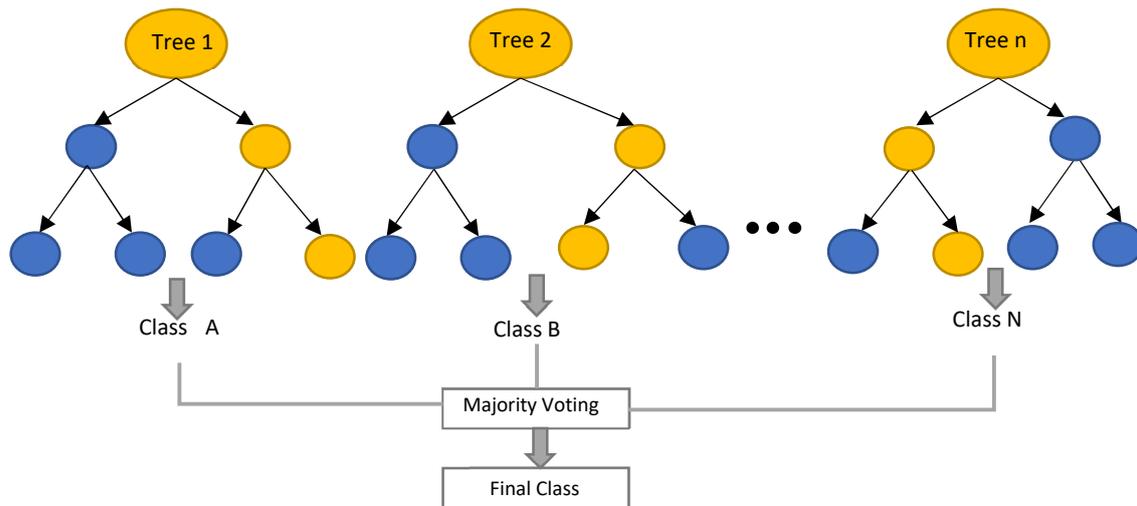


Figure 9. Logistic regression.

#### 4.6. Extreme Gradient Boosted Trees (XGBoost)

XGBoost was developed by Chen, T. & Guestrin, C. [69], a scalable tree boosting system that is widely used by data scientists and provides state-of-the-art results on many problems. Open source C++ was utilized in the implementation of XGBoost algorithm on forecasting the downtime of a printing machine based on real time predictions of imminent failures [62] and used unstructured historical machine data to train the ML classification algorithms including RF, XGBoost, and LR in predicting the machine failures. Various metrics were analyzed to determine the goodness of fit of the models. These metrics include; empirical cross-entropy, area under the receiver operating characteristic curve (AUC), receiver operating characteristic curve itself (ROC), precision-recall curve (PRC), number of false positives (FP), true positives (TP), false negatives (FN) and true negatives (TN) at various decision thresholds, and calibration curves of the estimated probabilities. Based on the results obtained, in terms of ROC all the algorithms performed significantly better and almost similar. But in terms of decision thresholds, XGBoost and RF perform better than LR. XGBoost algorithm tree uses majority voting technique to define final class [70]. XGBoost algorithm tree is presented in Figure 10.



**Figure 10.** XGBoost algorithm tree.

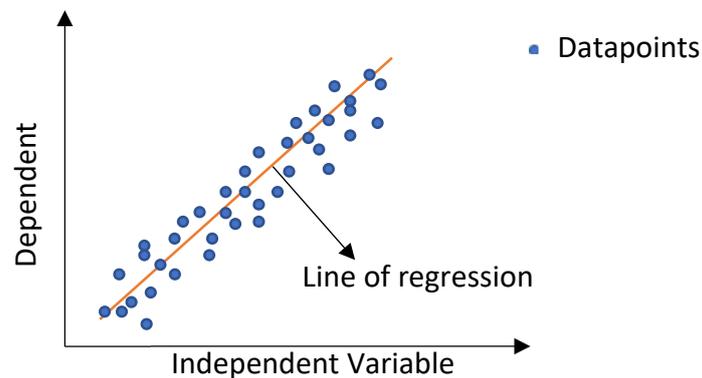
#### 4.7. Gradient Boosting Machines (GBM)

GBM is a family of powerful machine-learning techniques that have shown considerable success in a wide range of practical applications. It is also an assembly-based model that learns to update prediction results on new models consecutively [71,72].

A data driven diagnostics and prognostics framework for machines to increase efficiency and reduce maintenance cost was proposed by Xiang, S. et al. [56]. Moreover, an accurate data labeling methodology is developed for supervised learning via comparing the serial number of target components in the adjacent dates. In the study, vending machine real data was used to validate the proposed framework for three different classifiers including SVM, RF, and Gradient Boosting Machines (GBM). Moreover, two models were developed for PdM, one for diagnostics and the other for two-stage prognostics. Results for the cross-validated simulation obtained shows that the diagnostics model can achieve more than 80% of accuracy, thus the developed model of GBM can be applied for diagnosis and prognostics monitoring of complex vending machines. The prognostics model outperforms one-stage conventional prediction models.

#### 4.8. Linear Regression

Linear regression refers to a multivariate linear combination of regression coefficients [63]. The coefficients are calculated by the generalized least square technique. Linear regression is deterministic and the parameter is less, there is no need to adjust something other than the data break for model training and testing. Linear regression and Random Forest Regression are very common regression algorithms in general and time series regression algorithms that have already been used in the field of predictive maintenance [73]. Linear regression model in machine learning is presented in Figure 11.



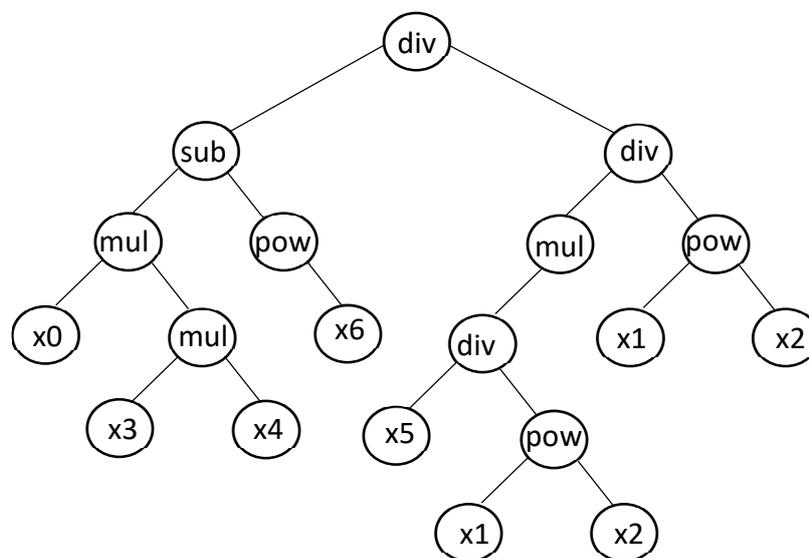
**Figure 11.** Linear regression in machine learning.

ML algorithms including linear regression, RF, and Symbolic Regression (SR) are applied in modeling the condition of a healthy industrial machinery [63], where they proposed a methodology for detecting and predicting drifting behavior (called concept drifts) in continuous data streams. Further, a real-world case study was presented on industrial radial fans. Based on the results obtained using the synthetic data, both the results from concept drift detection and prediction are highly successful.

#### 4.9. Symbolic Regression (SR)

SR refers to models in the form of a syntax tree composed of arbitrary mathematical symbols (terminals: constants and variables, non-terminals: mathematical functions) that can be easily converted into simple mathematical functions. Top-down syntax trees are reviewed for target estimation [63]. Syntax trees are developed using the stochastic genetic programming technique from the field of evolutionary algorithms [74].

SR has been applied in modeling the condition of a healthy industrial machinery [63]. The study proposed concept drifts methodology in continuous data streams. Further, a real-world case study was presented on industrial radial fans. Based on the results obtained using the synthetic data, both the results from concept drift detection and prediction were highly successful. Sample of SR algorithm is shown in Figure 12.



**Figure 12.** Example of SR algorithm.

#### 4.10. Other Machine Learning Algorithms

Janssens, O. et al. [75] reported a study on DL for Infrared Thermal Image Based Machine Health Monitoring, where in the study they considered Convolutional Neural Network (CNN), a Feature Learning (FL) tool, in detecting the various conditions of the machine. Moreover, FL was considered because it requires no feature extraction nor expert knowledge. Transfer Learning (TL) is a means to reuse layers of a pretrained Deep Neural Network (DNN). This was mitigated in their study. Case studies have been carried out on machine-fault detection and the oil-level prediction; in both cases, results have shown that CNN outperforms the classical FE methods. They added that the proposed method has the potential to improve online CM, like offshore wind turbines. Another potential application is the monitoring of bearings in manufacturing lines. Using thermal imaging together to the trained CNN allows identifying the location of the faults in the manufacturing lines.

Huhtanen, T. and Jung, A. [76] proposed a study on DL for predictive maintenance of photovoltaic panels. CNN was applied for monitoring the operation of photovoltaic panels. In fact, they estimate the regular electrical power curve of the photovoltaic panel depending on the power curves of the neighboring panels. An unusually broad difference between the predicted and the actual (observed) power curve can be used to suggest a malfunctioning panel. By the means of numerical experiments, they are able to demonstrate that the proposed method is able to predict accurately the power curve of a functioning panel and the method out-performs the existing methods that are based on simple interpolation filters.

Pan, Z. et al. [77] proposed a modular cognitive acoustics analytics service for IoT that provides customers with an incremental learning approach to improve their analytical capabilities for non-intuitive and unstructured acoustic data through a combination of acoustic signal processing. They pointed out that different types of data formats created from complicated acoustic environments can go through pre-processing and noise reduction stages and then feed into higher-level analytics platforms. The model allows for acoustic signal-based anomaly detection, acoustic grouping, acoustic signal processing, acoustic array processing, and other features. In classification, the model uses a baseline algorithm when a small amount of data is used, while when a huge amount of data is used, this model utilizes a technique based on the DNN to perform a more accurate classification. This service will include signal processing data, such as sound intensity, spectral centroid, frequency, etc., and can support numerous applications. Eventually, the service can also detect several sources of sound that allows detection and enhancement of the acoustic source. Experimental findings show that this service achieve excellent performance. The application case for the diagnosis of a washing machine is defined.

Jimenez-Cortad, A. et al. [78] carried out a case study based on the application of predictive maintenance to a real machining process. The aim of their study is to increase tool life of the machine by application of ML methods for RUL prediction. Real-time data obtained from the computer and then approximation of the data performed in the analysis. Their study conducted linear and quadratic regression models to perform the design application for RUL estimation. Finally, accurate results were found in their study to predict RUL for comparison. Luo, W. et al. [79] employed predictive maintenance approach for machine tool driven by digital twin to avoid faults and causalities. In their study, a hybrid approach was utilized to calculate RUL results that show the prediction error ratio (between actual value and predicted value).

In this section, a summary on the applications of ML algorithms in PdM will be given. Table 1 summarizes the analysis and comparison among these algorithms that are mainly applied in the field of PdM according to ML techniques, ML categories, equipment systems, and type of data.

**Table 1.** Recent Applications of machine learning (ML) in Predictive Maintenance.

ML Techniques	ML Cat.	Equipment/System	Device Used for Data Acquisition	PdM Data Description	Data Size	Data Type	Key-Findings	Ref
ANN	C	Tool wear for CNC-MM, Deckel Maho DMU 35M	Bosch XDK sensor	Acceleration data	3-dimensional input vector	Real data	<ul style="list-style-type: none"> <li>• Tool wear monitoring of a CNC-MM with equipped built-in sensors.</li> <li>• Applicable to older machines that can be utilized in I4.0.</li> <li>• Explore and enable a rapid adaptation to new environmental conditions.</li> <li>• It can be used to predict RUL of the tool.</li> </ul>	[46]
ANN	C	AK-FN059 with 12cm cooling fan	MMA8452Q-Accelerometer	Motor vibration measurements	9180 observations with 4 attributes	Synthetic data	<ul style="list-style-type: none"> <li>• Generates a training dataset based on vibration measurements.</li> <li>• ANN trained to predict equipment failure time.</li> <li>• k-fold cross-validation and model generalization performed.</li> <li>• Compared with other ML techniques. Resulted that ANN performs better.</li> <li>• In comparison to RT, RF, and SVM, ANN shows better results.</li> </ul>	[47]
LR XGBoost RF	C	Printing machine	-	Machine's operational status data	100 operational variables/minute	Real data	<ul style="list-style-type: none"> <li>• The fit of the models was determined by various metrics.</li> <li>• Based on decision thresholds, RF and XGBoost perform better than LR.</li> <li>• All the algorithms performed similarly better in terms of ROC.</li> </ul>	[62]
ANN SVM	C, R	Rail-Tram track, 250 km of double tracks and 25 routes	Non-contact optical laser	Track geometry data-gauge measurements data	-	Real data	<ul style="list-style-type: none"> <li>• Used for tracking gauge deviation and measurements prediction.</li> <li>• Slightly ANN models perform better in predicting the gauge deviation of straight segments.</li> <li>• SVM models are better in predicting gauge deviation of curved segments.</li> </ul>	[48]

ANN	C	Wind turbine at 1200rpm	Accelerometer sensor	Vibration signals data	243 datasets with 10,000 sample length	Real data	<ul style="list-style-type: none"> <li>Investigated the time domain vibration signatures for critical components.</li> <li>Acquired the healthy and faulty condition vibration signature.</li> <li>Model classifies faulty and healthy state features with 92.6% classification efficiency.</li> </ul>	[10]
DNN-CNN, DL	C	Rotating machinery	IRT image, temperature sensors	Accelerometer, thermocouple, and thermal camera measurements	5 REBs × 8 conditions for IRT and 5 REBs × 12 temperature measurements	Real data	<ul style="list-style-type: none"> <li>CNN algorithm applied to detect several conditions of rotary machinery.</li> <li>The technique is able to improve online CM in offshore wind turbines.</li> <li>Can be applied in manufacturing lines to monitor bearings.</li> <li>Oil-level prediction and machine-fault detection CNN outperforms.</li> <li>FL technique provides 6.67% better compared to the FE technique.</li> </ul>	[75]
CNN	C	Photovoltaic panels	Wireless sensor	Electrical power signals	400 samples at 1-min sample interval for 1,000 days	Synthetic data and real data	<ul style="list-style-type: none"> <li>CNN techniques can be applied in PdM of PV panels.</li> <li>Validated with the means of numerical experiments, that can accurately predict the power-curve of functioning panel.</li> </ul>	[76]
CNN	C	IoT- IBM system x3650 M4 machine that has 16 hardware threads with 192GB memory	Acoustic sensor	Acoustic sensor data	Dataset of >20hrs-audio within 7-sounds of classes	Real data	<ul style="list-style-type: none"> <li>Model allows acoustic classification, signal based anomaly detection, signal processing, sensor array processing.</li> <li>Uses deep-NN to provide more precise classification when large amount of data is used.</li> <li>Presented excellent performance and it can provide signal analysis results.</li> </ul>	[77]
NN	-	APU-gas turbine, GTCP 331- 430kW	Fleet of Airbus	Gas path measurements	310 routine records, with 14- starter nominal and 12- starter degraded conditions	Real data	<ul style="list-style-type: none"> <li>Developed an NN and physics-based models for starter degradation monitoring.</li> </ul>	[49]

							<ul style="list-style-type: none"> <li>Both models are able to monitor the health of the starter and can illustrate symptoms of degradation.</li> <li>NN provides more accurate results with the starters at healthy condition.</li> <li>Physics-based provides better results on the starters with growing degradations.</li> </ul>	
BN	C	TASMI-01	-	Semiconductor manufacturer	1890-time intervals with 24 failure occurrences	Real data	<ul style="list-style-type: none"> <li>Failure predictions performed by BN for event-driven maintenance data.</li> <li>The promising results obtained from offline prediction of the case study reveals the significance to outspread it for real time predictions.</li> <li>The BN model can be used for fault diagnosis not only for failure inference.</li> </ul>	[80]
RF	C	Rotating machinery with SEW Euro drive 1.1kW motor	IRT imaging and Vibration-based sensors	Thermal imaging and vibration data	5 bearings × 8 conditions for IRT and 5 bearings × 12 measurements	Real data	<ul style="list-style-type: none"> <li>A multi-sensor system for rotating machinery using a feature fusion method.</li> <li>Multi-sensor technique can compensate the shortcomings of the heat or vibrations.</li> <li>Offers a significant increase in fault detection performance.</li> </ul>	[64]
RF, DT, NB, BC	C	Turbofan-engine	Phonic wheel-pressure, fuel metering valve, temperature sensors	Turbofan-engine multidimensional data	1616 observations of abnormal and healthy engines	Real data	<ul style="list-style-type: none"> <li>Developed a learning algorithm that automatically detects and analyzes multidimensional dataset of turbofan-engine.</li> <li>The model has the ability to select good combinations of tests with higher than 85% pre-identification.</li> </ul>	[65]
RF, -DT, NBC, SVM, LR, linear ridge	C	Rotor bar- LS-PMSM	Torque and speed sensors	Transient current signal data	320 tests for healthy and faulty motor	Real data	<ul style="list-style-type: none"> <li>Uses RF to detect a broken rotor bar failure in LS-PMSM.</li> <li>RF classifiers results outperformed compared to other techniques result.</li> <li>Validity and reliability of the RF technique for fault detection performed.</li> </ul>	[66]

							<ul style="list-style-type: none"> <li>The technique attained 98.8% correct rate of diagnosis.</li> <li>Impulsion and mean-index are identified.</li> </ul>	
RF, TF-IDF - SVM, LR	C	Aircrafts	-	Historical data of aircraft maintenance systems	1542 events with 131 fault types for 2yrs	Real data	<ul style="list-style-type: none"> <li>ML model for fault prediction using TF-IDF and RF of aircraft.</li> <li>Used TF-IDF to extract the features of the raw-dataset from past flights.</li> <li>RF achieves 66.67% true positive rate with a low of 0.13%.</li> </ul>	[67]
RF, and XGBoost	C	Production lines and semiconductor	-	Bosch and SECOM manufacturing datasets	Bosch: 1.2mill observations and over 4,000 features.	Real data	<ul style="list-style-type: none"> <li>Analytics framework applied in fault detection to Bosch and SECOM datasets.</li> </ul>	[81]
GLM, RF, GBM, DL	-	Semiconductor	Tool sensors, process recipe and wafer count.	Past maintenance, tool sensors, and process datasets	21 tool-sensors, 3 process recipe and wafer count	Real data	<ul style="list-style-type: none"> <li>Ensemble learning selected as best fit technique for PdM for semiconductors.</li> <li>PdM strategies on ML models and equipment data.</li> <li>Emphasized the need to adopt cross validation on ensemble PdM based models.</li> </ul>	[82]
SAFE	R	Semiconductor Ion Implantation process	31 sensors, - ion-implantation tool	Maintenance cycles dataset	33 R2F maintenance cycles with 31 variables	Real data	<ul style="list-style-type: none"> <li>An approach that dealt with time-series data extraction for PdM.</li> <li>Applied SAFE methodology in time-series PdM.</li> <li>The technique outperforms the classical feature extraction methods.</li> </ul>	[83]
RF	C	Industrial pumps	Accelerometer and temperature sensors - ABB WiMon 100	Industrial pumps vibration data	30 chemical plant industrial pumps for 2.5yrs with 1066 features.	Real data	<ul style="list-style-type: none"> <li>A case study based on real data of industrial plant.</li> <li>RF algorithm that can detect faults in vibration data for up to 7 days.</li> <li>The model was validated using data collected for over 2.5 years.</li> </ul>	[84]

RF	C	Industry 4.0 - cutting machine	Machine PLCs and communication protocols sensors	PLCs and communication protocols sensors data	530,731 data readings for 15 different machine features	Real data	<ul style="list-style-type: none"> <li>Permits dynamical decision rules for PdM.</li> <li>Shows a suitable performance in predicting the machine stages with good precision.</li> <li>Predicts different stages of machine with 95% of high accuracy.</li> </ul>	[85]
RF	C	Refrigeration systems of supermarket	Temperature sensors	Temperature, work-order data and defrost state	2,265 refrigeration cases across 17 stores for 2 months	Real data	<ul style="list-style-type: none"> <li>ML based model for early detection in refrigeration system.</li> <li>Validated with a real data for 2265 different refrigerators.</li> <li>Achieved 89% in precision.</li> <li>A lead time of 7 days.</li> <li>A recall of 46% when evaluated on unseen cases.</li> </ul>	[86]
RF	C	HDD	-	Historical data	Dataset of 16,862 failures and 47,677 non-failures for 2yrs	Real data	<ul style="list-style-type: none"> <li>PdM of HDD failure detection based on Apache Spark.</li> <li>The model can assist IT staff by making them more proactive and productive by identifying imminent disk failure quickly.</li> </ul>	[87]
RF	C	Induction motor, 2.2 kW	Datalogger with WIFI communication capability	Voltage and Current waveforms data	3-phases dataset of 1,159-(357-healthy and 802-faulty)	Real data	<ul style="list-style-type: none"> <li>Analyzed single and double classifier approach.</li> <li>Methods could effectively be used in detecting the inter-turn short-circuits using few numbers of data points.</li> <li>Double classifier approach produces a better result than single.</li> </ul>	[88]
RF	C	Wind turbine	Alarms activations and deactivations	Alarm types and operational dataset for 17-turbines	448-alarm types and 104-operational data (2yrs-1,787,040 dataset)	Real data	<ul style="list-style-type: none"> <li>Approach to generate a predictive data driven models based upon historical dataset.</li> <li>A front-end where the status of the turbine can be visualized in real time.</li> <li>Experiments proved that the implementation has gained an optimal overall success.</li> </ul>	[89]
RF	C	Trucks and buses air compressors	Logged on-board	LVD and VSR dataset	1,250 unique features	Real data	<ul style="list-style-type: none"> <li>Commercial vehicles air compressor to detect forthcoming faults.</li> </ul>	[90]

				with 65,000 European Volvo trucks			<ul style="list-style-type: none"> <li>• Generalizes to repair several components of a vehicle.</li> <li>• Used RF and two techniques for feature selection.</li> <li>• The ML feature-based model outperform compared to the human ones.</li> <li>• Usage sets and Beam search 1 were found to be the best features.</li> <li>• A positive profit was shown by all the features at final evaluation.</li> </ul>	
LSTM	Classifier	Turbofan engine	C-MAPSS tool – multiple sensors	NASA Ames Prognostics Data Repository	A dataset of 4 subsets with 708 trajectories by 21 columns for 21 sensors	Synthetic data	<ul style="list-style-type: none"> <li>• A DPM framework can be used for prognostics decisions and prognostics estimations.</li> <li>• No specific degradation model or a specific RUL function.</li> <li>• Does not predict the RUL.</li> <li>• It provides the probabilities of when the system will fail into several time intervals.</li> <li>• Model that permits maintenance costs and inventory evaluation.</li> </ul>	[91]
LSTM	Clustering	Engine	Operational and Sensors	C-MAPSS NASA simulation dataset of engine degradation	14-inputs and 4-outputs of (21 sensor measurements and 3-operational settings)	Synthetic data	<ul style="list-style-type: none"> <li>• Implemented an LSTM model to Apache Spark on a large-scale dataset for current life condition predictions of an engine.</li> <li>• Deals with sequential data.</li> <li>• LSTM network output is to decide the current state of a component before their life ends.</li> <li>• Trained and tested on an engine degradation open source data.</li> <li>• It can be used in industries to detect break downs before they occur.</li> </ul>	[92]
1NNC, LS-SVM-linear kernel	C	Rolling bearing	TFI recognition	Vibration signals data	80 samples for a test with 1024 data points.	Real data	<ul style="list-style-type: none"> <li>• Developed an approach for fault monitoring in bearings based on sparse TFIs recognition.</li> </ul>	[93]

							<ul style="list-style-type: none"> <li>• Traditional time–frequency drawbacks can be overcome by applying STFA-PD.</li> <li>• STFA-PD method shows a promising diagnosing performance.</li> <li>• Used to obtain high-quality TFIs and also used in analyzing non-stationary signals.</li> </ul>	
SVM, - PF, GAPF	R	Aircraft	LVDT and RVDT	Actuator internal oil leakage fault data	2,400- training samples and 480- testing samples datasets	Real data	<ul style="list-style-type: none"> <li>• Studied a novel oil leakage fault prognostics method for actuators.</li> <li>• Presented a hybrid SVM-PF framework for actuator fault prognostics.</li> <li>• Tested and the results obtained as a satisfactory performance.</li> <li>• SVM-PF hybrid framework has better prognostics accuracy.</li> <li>• Higher fault resolution was observed compared to traditional ML.</li> </ul>	[94]
SVM, RF, GBM	R	Vending machines	-	System logs, operation logs, context, process and performance logs data.	3,450 malfunction machines data	Real data	<ul style="list-style-type: none"> <li>• Framework for machines to increase efficiency and reduce maintenance cost.</li> <li>• Used vending machines real data to validate the proposed framework.</li> <li>• Two models developed for PdM, one for diagnostics and the other for two-stage prognostics.</li> <li>• The diagnostics model can achieve more than 80% of accuracy.</li> <li>• The prognostics model outperforms one-stage conventional prediction models.</li> </ul>	[56]
SVM, Binary LR, Gamma process	C, R, D	Track	-	Track geometry data of railroads	6,500 red and 17,500 yellow tag defects records of dataset	Real data	<ul style="list-style-type: none"> <li>• Suggested that ensemble methodology outperform the other techniques.</li> <li>• A technique for forecasting track geometry defects by developing an ensemble classifier.</li> <li>• In all cases, ensemble classifier proved to outperform the individual algorithms.</li> </ul>	[95]

							<ul style="list-style-type: none"> <li>The technique can be applied to any type of tracks and other type of defects.</li> </ul>	
SVM, RF, LDA	C	Railways - Mile track	-	Track geometry data	1 mile of track with 28 inspection dates and 31 features by 31 parameters	Real data	<ul style="list-style-type: none"> <li>Investigated the possibility of minimizing multivariate track geometry indices into a low-dimensional form.</li> <li>SVM was found to be the most effective technique and predicts better track defects.</li> <li>Measured the performance of the model using TPR and FPR.</li> </ul>	[96]
SVM-Regression kernel	R	Gas turbine engine of an aircraft	Time series sensor	Sensor measurements data of the time series from CMAPSS dataset of aircrafts	CMAPSS dataset-14 inputs with 21 outputs	Synthetic data	<ul style="list-style-type: none"> <li>Model for SVM in prognostics prediction of multiple time series tasks.</li> <li>Testing was carried out with simplified time-series simulated data.</li> <li>Improvements were shown from the results obtained over the conventional SVM results.</li> </ul>	[97]
1st and 2nd class SVM, DT, RF, LSTM-NN	-	Power Transformer	-	Insulating oil tests dataset: Mechanical and chemical measurements	15,031 tests with 30 variables	Real data	<ul style="list-style-type: none"> <li>Used insulating oil tests dataset for failure prediction in power transformers using several ML techniques.</li> <li>SVM produces best performance among the other tested techniques.</li> <li>SVM achieved 77% recall performance with 35% FPR.</li> <li>LSTM did not give better performance because the data was collected at low frequency.</li> <li>Collecting the data at an adequate frequency is better than large number of observations.</li> </ul>	[98]
SVM, K-means, K-NN, Euclidean Distance, and CRA	C	Wind turbine bearings	Accelerometer, displacement, velocity, and torque sensors	Vibration signals data	-	Real data	<ul style="list-style-type: none"> <li>Applied several ML methods for bearing fault detection.</li> <li>Investigated the similarity of the ML models results.</li> </ul>	[99]

							<ul style="list-style-type: none"> <li>• Vibration signal analysis was performed while studying and extracting the pattern-behavior of bearings.</li> <li>• CRA can recommend a fault with 93% accuracy.</li> </ul>	
SVM, MLP-ANN	C	Electrical power systems	-	Partial Discharge samples and noise samples dataset	100 thousand datasets	Synthetic data & Real data	<ul style="list-style-type: none"> <li>• Parallelism and pipelining methods can give a significant maximization of sample rates.</li> <li>• The method provides better performance compared to other classifiers.</li> <li>• Use low number of self-configuration capabilities and configuration parameters.</li> <li>• The technique can deal with uneven dimensional input spaces.</li> <li>• The model can process with simple evaluation function.</li> </ul>	[100]
SVM, k-NN, MC	C	Tungsten filament - Ion implantation	Ion-implantation tool	Maintenance cycles dataset	N = 33 R2F maintenance cycles with 31 variables and 3671 batches	Real data	<ul style="list-style-type: none"> <li>• Approach for integral faults type.</li> <li>• Can be used in high-dimensional and censored data problems.</li> <li>• Can be applied to any maintenance problems as R2F data is available.</li> <li>• The model can be applied in health factor indicator.</li> <li>• Shows better performance than other PdM classical methods and SVM.</li> <li>• MC-PdM - k-nn outperforms PvM approaches.</li> <li>• SVM offers better performance than k-nn.</li> </ul>	[14]
Dynamic regression	R	Bearing	PRONOSTIA platform	Vibration measurements data	2,560 samples amplitude of the vibration signals	Real data	<ul style="list-style-type: none"> <li>• Method for rolling element bearing.</li> <li>• Predicts the bearing health and its URL using regression models.</li> <li>• Used dimensionless quantity as bearing health indicator.</li> <li>• ABT was used to determine TSP.</li> <li>• Validated and tested on PRONOSTIA dataset.</li> </ul>	[101]

							<ul style="list-style-type: none"> <li>Achieved a reasonable result compared to the existing techniques.</li> </ul>	
GA-ANN, SVM	C	Rotating machine-Gear box	Acoustic emission and vibration sensors	Acoustic emission and vibration signals	(16 time-domain and 6 frequency-domain	Real data	<ul style="list-style-type: none"> <li>Investigated an early detection of the potential failures of rotating machine.</li> <li>Early misalignment detection was very hard using frequency analysis technique.</li> <li>The feature analysis method can detect a growth in fault.</li> </ul>	[102]
SMDP	-	Ram feed-Boring machine	-	Feed straightness, positioning and hydrostatic pressure units	18 degradation histories and 16 failure histories	Real data	<ul style="list-style-type: none"> <li>Developed a unit-level and system-level maintenance of a boring machine.</li> <li>Achieved low maintenance cost in comparison to other techniques.</li> <li>The model can be generalized to a wide variety of systems with a particular failure mode.</li> </ul>	[103]
HC, k-means, PCA, model-based and Fuzzy C-Means clustering	Clustering	Exhaust fan	Vibration monitor sensor	Vibration data	The vibration was collected at every 240 min for 12days with 41 observations	Real data	<ul style="list-style-type: none"> <li>Early fault detection of exhaust fan using several ML algorithms.</li> <li>T<sup>2</sup> statistic method is the best for faults detection only.</li> <li>Clustering algorithms is the best for fault detection under different levels.</li> <li>PAC produced better results compared to model-based techniques.</li> </ul>	[12]
k-means	Clustering	Laser melting	Temperature & pressure sensors	Laser melting machine sensor data	206 manufacturing processes data with 3D- matrix (7 × 3 × 206)	Real data	<ul style="list-style-type: none"> <li>Analyzing and visualizing offline-data from several sources.</li> <li>The clusters were identified by utilizing three sensors.</li> <li>Three faulty and normal operation stages were identified by the sensors.</li> <li>Implemented a CMS that permits machine tools for PdM solutions.</li> </ul>	[104]
k-means, -DL, RNN,	C	-	Vibration sensors	Vibration data and	51 vibration sensors for 2.5yrs	Real data	<ul style="list-style-type: none"> <li>Sequential data with categorical inputs were used in machine health predictions.</li> </ul>	[105]

				categorical metadata.				<ul style="list-style-type: none"> <li>Validated using an ablation study.</li> <li>Can be used without making any assumptions to the data on any dataset.</li> </ul>	
PCA, k-means	C, Clustering	Oil-immersed power transformer	-	Dissolved gas concentrations data	46-observations with 6-variables	Real data		<ul style="list-style-type: none"> <li>A study for automatically extracting classes in dissolved gases.</li> <li>Class interpretation was according to the information obtained for each gas and TDCG.</li> <li>Permits to identify the major working periods of the power transformer.</li> </ul>	[106]
FURIA, -ANN, SVM, RF, BayesNet, LogitBoost	C	Gas turbine	-	Big dataset of gas turbine	11,934 instances with 16 independent variables for 49 measurements	Synthetic data		<ul style="list-style-type: none"> <li>Big data analytics maintenance optimizing through CBM.</li> <li>The FURIA model produces better decision boundaries.</li> <li>Simulation dataset of a sophisticated gas turbine propulsion plant is used.</li> </ul>	[107]
DL	C	CNC machine	Accelerometer sensor	Vibrational data	10,000 samples with 1,024 data points for 288 days	Real data		<ul style="list-style-type: none"> <li>An early fault detection DL model under time varying conditions for CNC machine.</li> <li>Developed from vibration signals to select impulse responses.</li> <li>Experiment show that the proposed technique is not affected by time-varying conditions.</li> <li>The model has the potential to detect an early fault in manufacturing.</li> <li>The method can significantly detect the machine tool health status.</li> </ul>	[108]
DT, ANN, RF, GNB, BNB	C	Anode production of industrial equipment	Process sensor	Process sensor data within the period of operation	29 features and 96 faults with 4,852,153 data points	Real data		<ul style="list-style-type: none"> <li>PdM technique for fault forecast in real time before occurrence of an industrial equipment.</li> <li>Predicts faults in industrial equipment 5–10 min before it occurs.</li> </ul>	[109]
LR, DT, SVM, RF, kNN, K-	-	Turbofan engine	Sensor run-to-failure measurements	Repository dataset of NASA for	Dataset collected from 100 to 250 engines, each engine	Real data		<ul style="list-style-type: none"> <li>A comparative study for ML techniques in predicting RUL of aircraft-turbofan engine.</li> <li>The models can be applied in predicting faults before their occurrence.</li> </ul>	[21]

Means, GBM, AdaBoost				turbofan-engine	with 21 sensor values.		<ul style="list-style-type: none"> <li>Developed the ML models using NASA prognostics data of turbofan-engine.</li> <li>Tested and validated, results obtained have shown a promising result.</li> </ul>	
MGGP	C	Metal lathe machine	Accelerometer and noise level meter	Vibration and acoustic signals data	42 data samples	Real data	<ul style="list-style-type: none"> <li>Developed for probable faults detection of metal lathe machine using vibro-acoustic condition monitoring.</li> <li>Proposed MGGP framework based on it is new measure of complexity.</li> <li>The model outperforms the classical MGGP models.</li> </ul>	[110]
RLS, SVM	R	LM 2,500 - Gas turbine	Sophisticated simulator of a gas turbine	CODLAG propulsion plant data	9 × 51 experiments	Synthetic data	<ul style="list-style-type: none"> <li>Techniques for the degradation forecasting of propulsion plant.</li> <li>Tested using realistic and sophisticated simulator of a gas-turbine.</li> <li>Effective solution in real application of maritime for CBM purposes.</li> </ul>	[111]
HC, K-medoids, K-means	Clustering	Semiconductor manufacturing - CVD process	CVD process	Sensors dataset	80-sensors for 6-months: 6,912,000 data points	Real data	<ul style="list-style-type: none"> <li>Validated the model using CVD process sensor data.</li> <li>Outperforms the model with full sensors.</li> <li>OPTICS was found to be the best of all the 5 algorithms.</li> <li>The model can be applied in engineers making decision of CBM.</li> </ul>	[13]
ARIMA, GB, RF and RNN	R, Clustering	Machine tool of CNC Machine	Machine tool signal recorder	Spindle load, piece number, and tool position	2 years data. 1,100 series and 550,000 pieces were stored.	Real data	<ul style="list-style-type: none"> <li>Developed for simulation of RUL in a machining phase based on a linear regression model.</li> <li>Obtained more precise findings to predict the RUL for comparison.</li> <li>Can be applicable to other processes in production.</li> </ul>	[78]
GBM, K-nearest Neighbors,	C	Engine equipped with	MEMS triaxial accelerometer	Vibration, noise, pressure,	4,000 acceleration values	Real data	<ul style="list-style-type: none"> <li>Maximum peak in the Naïve Bayes algorithm with a precision loss of 5%.</li> </ul>	[112]

BNB, NN, DT		a rotating shaft		temperature, humidity				<ul style="list-style-type: none"> <li>• Only NN algorithm behaves differently compared to other algorithms.</li> <li>• Best results obtained with isolation forest algorithm by increasing the training time with 20%.</li> </ul>	
LR, DTR, RFR, SVR	R	CNC machine tool	Accelerometer, Dynamometer, AE sensor	Vibration, force, temperature, federate, cutting depth	3 sensors utilized for data collection	Real data		<ul style="list-style-type: none"> <li>• Hybrid method predicts value close to actual value with small error.</li> <li>• Data driven method provides less accuracy.</li> <li>• Hybrid approach driven by Digital Twin provides more precise predicted values, 6.27% at the end stage.</li> <li>• Approach verified by a case study of the CNCMT tool life prediction.</li> </ul>	[79]
K-means, PCA	Clustering	Machine motor	Sensor, PLC, Production monitoring system	Power, torque, vibration, temperature	13 datasets for machine1, 9 datasets for machine 2	Real data		<ul style="list-style-type: none"> <li>• There are inliers within the dataset which are seen in low frequency.</li> <li>• Observed that torque measurement deviates from normal range and achieves the peak value. So that, the spindle of Machine1 may not work within the normal range.</li> <li>• There are many inliers found and they can give diagnostic information.</li> </ul>	[113]
ANN, SVM	C, R	Building facilities	IoT sensor network	Temperature, pressure, flow rate	300 datasets for condition prediction.	Real data		<ul style="list-style-type: none"> <li>• BIM and IoT facilitated the implementation of predictive maintenance to improve the feasibility of FMM process.</li> <li>• Consists of two layers: information layer and application layer.</li> <li>• Four modules are applied for maintenance: 1) condition prediction, 2) maintenance planning, 3) condition monitoring, 4) condition assessment.</li> </ul>	[30]
ANN	C	Packaging robot	Sensor	Vibration, temperature, humidity	157 failure for 2 years. 8 input, 20 hidden and 4 output nodes	Real data		<ul style="list-style-type: none"> <li>• MLP structure can cope with unplanned downtime occurrence.</li> </ul>	[114]

							<ul style="list-style-type: none"> <li>• Can reduce the unplanned production downtime costs immensely.</li> <li>• Theoretical and practical comparison of failures provided.</li> </ul>
SVM, LR	C, R	Nuclear power plant	Temperature sensor, pressure sensor	Temperature, power, current, speed	-	Synthetic data	<ul style="list-style-type: none"> <li>• ML algorithm to predict maintenance of nuclear infrastructure.</li> <li>• Power consumption is monitored.</li> <li>• Temperatures within electrical panels are captured.</li> </ul>
FFNN, RF, LR	C, R	Oil analysis of gearbox	Data obtained from oil analysis company	Oil condition (loss of additives or contamination detection)	26 features with 887,255 samples collected from 126,644 gearboxes	Synthetic data	<ul style="list-style-type: none"> <li>• Demonstrated the potential use of RF as a diagnostic tool in PdM.</li> <li>• The RF models managed to achieve high recall, but the precision was low.</li> <li>• RF achieved for LEAK a mean accuracy of 0.9837.</li> <li>• For OVHT, POIL, and OTHR, none of the classification models showed significant classification ability.</li> </ul>
DT, RF, LR	C, R	Backblaze data center	SMART sensor	vibration	232,662 recordings in 2018.	Real data	<ul style="list-style-type: none"> <li>• Simulated the operation of several predictive maintenance systems.</li> <li>• Method used for bearing vibration failure data.</li> <li>• RF technique performs the best in terms of predictive accuracy.</li> </ul>
DNN, KNN	C, R	Turbofan engine	Multi-sensor	temperature, rotation speed, pressure	21 featured data sets.	Real data	<ul style="list-style-type: none"> <li>• C-RE has proven to be a robust feature selection method.</li> <li>• Identifies faults or occurrence of faults during the asset's life cycle.</li> </ul>
NN, weighted NN, DT	C	Boiler and heat pump of (HVAC) system	Sensor in each boiler and IoT platform	Temperature, number of boilers, heating requests, and cycle	The data was collected for 16 months, 1,000 appliances of ten different models.	Real data	<ul style="list-style-type: none"> <li>• Model exhibited best performance when the LSTM2 with 3 hidden layers of 50 neurons and the LSTM3 with 1 hidden layer of 25 neurons.</li> <li>• Models exhibited worse performance than other techniques.</li> </ul>

							<ul style="list-style-type: none"> <li>Weighted NNs has poor precision compared with the NN models.</li> </ul>	
GBM, RF, XGBoost, NN classifier	C	Woodworking industrial machines	Big Data provided by machine tool data log systems	Vibration, current, temperature	Dataset divided into two groups: a training set (70%) and testing set (30%)	Real data	<ul style="list-style-type: none"> <li>up to 98.9%, 99.6%, and 99.1% accuracy, recall, and precision respectively.</li> <li>PdM implemented to Big Data stream processing system.</li> <li>Screens log files and forecasts the computer's status every 24 h.</li> </ul>	[120]
RNN, SVM, k-nearest neighbor	C	Switchgear,	Infrared Sensor, thermopile array sensor	Temperature, voltage, current	-	Real data	<ul style="list-style-type: none"> <li>The combination of novel sensor technology and ML methods.</li> <li>Provides predictive maintenance solutions for medium voltage switchgear.</li> </ul>	[121]
RF	C	Aircraft	Aircraft advanced sensors	Pre-defined parameters, failure messages	More than 7 years' worth of data collected.	Real data	<ul style="list-style-type: none"> <li>Model has a precision of more than 70%.</li> <li>Model predicts aircraft failure that leads to component replacement.</li> <li>Predicts more than 50% of aircraft component replacement.</li> </ul>	[122]
LR	R	Jet engine blades	Flight sensor	Temperature, stress, strain	-	Real data	<ul style="list-style-type: none"> <li>87% or higher in life consumption prediction for 19 out of 21 blades that had failed at visual inspection.</li> </ul>	[123]
MLP, LR, GBT, SVM	R, C, Clustering	Wind turbine	Public dataset library, SCADA	Rotation, temperature, pitch angle, wind speed		Real data	<ul style="list-style-type: none"> <li>MLP produces the most promising model for predicting failures on the given dataset.</li> <li>MLP models is a good way of producing a model with a lower variance than the individual base models.</li> </ul>	[124]

CNC-MM: CNC milling machine; PF: Particle filter; LVDT: Linear variable differential transformer; RVDT: Rotary variable differential transformer; GAPF: Genetic algorithm particle filter; DBN: Dynamic Bayesian Network; FDD: Fault detection and diagnosis; BC: Bayesian classification; DT: Decision tree; GMB: Gradient boosting machines; DNN: Deep Neural Network; GA: Genetic Algorithms; LS-PMSM: Line start-permanent magnet synchronous motor; DT: Decision tree; NBC: Naïve Bayes classifier; LR: logistic regression; 1NNC: 1-nearest neighbor classifier; TFI: Time–frequency image; IRT: Infrared thermal; STFA-PD: Sparse time–frequency analysis method based on the first-order primal-dual algorithm; SMDP: Semi-Markov decision process; APU: Auxiliary power unit; TF-IDF: Term Frequency-Inverse Document Frequency; HC: Hierarchical clustering; PCA: Principle component analysis; GLM: Generalized linear model; GBM: Gradient boosting machine; FURIA: Fuzzy unordered rule induction algorithm; D: Deterioration; SLM: Selective laser melting; GNB: Gaussian Naïve Bayes; BNB: Bernoulli Naïve Bayes; RNN: Recurrent neural network; RUL: Remaining useful life; LDA: Linear discriminant analysis; HDD: Hard Disk Drive; LSTM: Long Short Term Memory; SAFE: Supervised Aggregative Feature Extraction;

CRA: Collaborative Recommendation Approach; MLP: Multilayer Perceptron; LVD: Logged Vehicle Data; VSR: Volvo Service Records; MC: Multiple classifier; MGGP: multi-gene genetic programming; BN: Bayesian Network; TASMI: Thermal Treatment equipment; R2F: Run to failure; RLS: Regularized least square; REB: Rolling element bearings; MOM: Manufacturing operation measurement; CVD: Chemical vapor deposition; DBSCAN: Density based spatial clustering of applications with noise; OPTICS: Ordering points to identify the clustering structure; FE: Finite element; FL: Feature learning; TPR: True Positive Rate; FPR: False Positive Rate; ABT: Alarm bound technique; TSP: Time to start prediction; CMS: Condition Monitoring System; CODLAG: COmbined Diesel eLectric And Gas propulsion plant; DPM: Dynamic Predictive Maintenance.

#### 4.11. Commercial Platforms available for Machine Learning in Smart Manufacturing Industry 4.0

Data Science and Machine Learning Platforms offer platforms for the development, implementation, and analysis of machine learning algorithms. Such systems integrate intelligent algorithms for decision taking with data, thereby enabling developers to build a business solution. Several platforms provide pre-constructed algorithms and simple workflows with functionality such as drag and drop modeling and visual interfaces that quickly link required data to the end solution, whereas others need further programming and coding skills. In addition to other machine learning applications, these algorithms have functionalities for image recognition, natural language processing, speech recognition, and recommendation systems [125]. The most common platforms are mentioned in Table 2.

**Table 2.** Most common commercial platforms for ML.

Software	Remarks
TensorFlow	<ul style="list-style-type: none"> <li>Open source software library for numerical computation</li> </ul>
IBM Watson Studio	<ul style="list-style-type: none"> <li>ML framework designed for an AI-powered company.</li> </ul>
RapidMiner	<ul style="list-style-type: none"> <li>Combines the whole lifecycle of data science from data planning to machine learning to predictive algorithm implementation.</li> </ul>
Google Cloud AI Platform	<ul style="list-style-type: none"> <li>Machine learning on any data, of any size.</li> </ul>
Box Skills	<ul style="list-style-type: none"> <li>Create structure and extract insights from your data at scale.</li> </ul>
Google Cloud AutoML	<ul style="list-style-type: none"> <li>Train high-quality models specific to their enterprise needs.</li> <li>Google's transfer learning and Neural Architecture Search technology.</li> </ul>
SAS Enterprise Miner	<ul style="list-style-type: none"> <li>Streamlines the data mining process to develop models quickly.</li> <li>Find the patterns that matter most for processes.</li> </ul>
MATLAB	<ul style="list-style-type: none"> <li>Programming, modeling, and simulation tool developed by MathWorks.</li> </ul>
IBM Watson Machine Learning	<ul style="list-style-type: none"> <li>Use of existing data to create, train, and deploy machine learning and deep learning models.</li> <li>An automated, collaborative workflow to grow intelligent business.</li> </ul>
Anaconda Enterprise	<ul style="list-style-type: none"> <li>Harness data science, machine learning, and AI.</li> </ul>
Amazon SageMaker	<ul style="list-style-type: none"> <li>Quickly build, train, and deploy machine learning models at any scale.</li> </ul>
IBM Decision Optimization	<ul style="list-style-type: none"> <li>Combines mathematical and AI techniques to help decision-making such as operational, tactical/strategic planning, and scheduling processes.</li> </ul>
IBM Cloud Pak for Data	<ul style="list-style-type: none"> <li>Modernizes data collection and data analytics to infuse AI across their organization.</li> </ul>
BigML	<ul style="list-style-type: none"> <li>Programmatic machine learning.</li> </ul>
H2O	<ul style="list-style-type: none"> <li>Decrease fraud and money laundering risks.</li> <li>Improve product design, marketing and business innovation.</li> <li>Early disease detection, drug discovery, personalized medicine.</li> </ul>
Oracle Data Science Cloud Service	<ul style="list-style-type: none"> <li>Train, deploy, and manage models on the Oracle Cloud.</li> </ul>
Domino	<ul style="list-style-type: none"> <li>Develop and deploy predictive models more rapidly.</li> </ul>
Deep Cognition	<ul style="list-style-type: none"> <li>Design, train, and deploy ML models without coding.</li> </ul>
KNIME Analytics	<ul style="list-style-type: none"> <li>Open source data analytics, reporting and integration platform.</li> </ul>
Qubole	<ul style="list-style-type: none"> <li>Provides a simple and secure data lake platform for ML, streaming, and ad-hoc analytics.</li> </ul>

## 5. Discussion and Conclusions

The literature review is categorized based on ML techniques, ML categories, equipment used, device used in data acquiring, description of the applied data, data size, and data type. Based on performed comprehensive literature review, predictive maintenance continues to be an important method for improving efficiency in all kinds of environments where machines that wear down over time are involved. The possibilities of manufacturing and placing cheap, connected sensors will continue to increase with the rise of IoT. As the amount of data increases with the number of sensors, so will the possibilities of applying machine learning algorithms to perform predictive maintenance.

This paper presents a comprehensive review of ML techniques applied in PdM of industrial components. Recent applications within the timeframe of ten years (i.e., 2010–2020) for several ML algorithms were reviewed and presented. Finally, some discussions have been drawn based on the literature review performed.

It is observed that predictive maintenance has enormous market opportunities, and that machine learning is an innovative solution to predictive maintenance implementation. Yet, according to a PwC survey, only 11% of the companies surveyed have “realized” predictive maintenance based on ML [126]. There are some challenges implementing ML algorithms for PdM in I4.0 and those are identified in Table 3.

**Table 3.** Challenges in implementing ML for Industry 4.0 (I4.0).

Challenges	Remarks	References
Identification of required data to collect	<ul style="list-style-type: none"> <li>• Launch of connected machines.</li> <li>• Unclear evidence of data that provide value.</li> <li>• Unclear business goal and planning.</li> </ul>	[33]
Getting required dataset	<ul style="list-style-type: none"> <li>• Without input data, it is not possible to run ML algorithm.</li> <li>• Much time and resources to establish ML solutions.</li> <li>• Choosing wrong ML algorithm causes loss time and loss in cost.</li> </ul>	[29]
Enhanced data science	<ul style="list-style-type: none"> <li>• Determine an appropriate method of analyzing the data.</li> <li>• Choosing a correct method of presenting the data-driven insights.</li> </ul>	[126]
Security	<ul style="list-style-type: none"> <li>• Safeguarding admission to critical equipment.</li> <li>• Proactive approach to cybersecurity whilst protecting connected assets</li> </ul>	[29,33,126]

SVM, RF, and ANN, are the most used ML algorithms in reviewed literature. They have been successfully applied in several field of PdM applications. Some authors [10,30,46–48,75–77,107,114,118–121] focused on ANN ML algorithms. Some other authors [64–67,81,82,84–88,90,116,117] studied RF technique. In the last years, use of SVM technique has received attention from authors [14,21,48,56,66,79,93–100,102,107,111,115,121,124]. Some authors [21,56,82,112,120,124] have given attention to GBM technique. Some authors [65,66,79,109,112,117] carried out studies by use of DT technique. However, it is observed that there is less consideration given to XGBoost technique and there are only a few studies found in the literature [62,81,120].

Based on the literature, it is observed that RF algorithm is the most extensively applied ML technique for PdM, as it has been applied on various industrial equipment, components, or systems, including rotating machineries, turbofan engines, rotor bar-LS-PMSM, aircrafts, production lines, semiconductors, industrial pumps, cutting machines, refrigeration systems of supermarket, hard drive disk (HDD), wind turbine, vending machines, and many others. However, authors usually

focused on CNC Machines [46,78,79], wind turbine [10,89,99,124], aircraft [67,94,97,122,123], and semiconductor [13,81–83].

Most of the data type used is real data; very few studies applied simulated or synthetic data while developing the ML algorithms. Public data applied in developing ML algorithm PdM models include Bosch data, SECOM data, Repository dataset of NASA for turbofan-engine, CMAPSS dataset of aircrafts, CMAPSS NASA simulation dataset of engine degradation.

- Vibration signals acquired using accelerometer are the most used data.
- The most applied ML category is classification.

According to existing research, a couple of limitations are highlighted. The limitations are the following: (1) Although the classifiers have presented excellent accuracy in the distinction between states, they are required to be trained with a complete dataset of all the faults. (2) Algorithms are selected base on developer's experience and this situation can have influence on variable of prediction results. (3) Carrying out a study with a single prediction method may not present excellent results. Therefore, application of other methods to provide comprehensive results between methods can give better understanding about the study. (4) Cross validation for models can be unsuccessful due to lack of RAM memory.

Moreover, some of the works conducted by this research employ regular machine learning methods without parameter tuning. Perhaps this is due to the fact that PdM is a new subject for industry experts and is beginning to be explored. It is also important to point out that it is appropriate to have the R2F and PvM strategies already applied in its process to collect data for PdM modeling in order to obtain good results of a PdM strategy in a plant. Based on that data, designing and validating a PdM strategy becomes feasible. During this study, it was noted that there is incremental application of machine learning techniques to develop PdM applications. Integrating PdM and machine learning in some applications provides cost reduction. However, the incorporation of PdM techniques with the new sensor technologies can be seen as avoiding unnecessary replacement of equipment, saving costs and improving process safety, availability, and performance.

This paper presents a comprehensive review of ML techniques in PdM by identifying the most used ML techniques, industrial areas where ML is applied, and utilized data type for ML applications, and proposes a way forward, and provides a foundation for further research. Below, remarks for further researches are made to motivate authors and practitioners.

- Extraction of real time data using intelligent data acquisition system can help to automate predictive maintenance.
- Combination of more than one ML models can provide better prediction compared to use of individual model.
- ML model implementations based on cloud can be further studied.
- Classification and Anomaly Detection algorithms can be combined to maintain precision of classification models without losing Anomaly detection advantages. By this way, PdM can be applied to equipment or system which does not have large dataset.

**Author Contributions:** Z.M.Ç. was responsible for data curation, resources, visualization. B.S. was responsible for investigation. Q.Z. and O.K. were responsible for supervision, methodology. B.S., Q.Z., O.K., M.A., A.A.N., and Z.M.Ç. were responsible for writing—original draft and writing—review and editing. B.S., Q.Z., and O.K. were responsible for project administration. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Carvalho, T.P.; Soares, F.A.; Vita, R.; Francisco, R.D.; Basto, J.P.; Alcalá, S.G. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024, doi:10.1016/j.cie.2019.106024.
- Cinar, Z.M.; Nuhu, A.A.; Zeeshan, Q.; Korhan, O. Digital Twins for Industry 4.0: A Review. In *Industrial Engineering in the Digital Disruption Era. GJCIE 2019. Lecture Notes in Management and Industrial Engineering*; Calisir, F., Korhan, O., Eds.; Springer: Cham, Switzerland, 2020; doi:10.1007/978-3-030-42416-9\_18.
- Cinar, Z.M.; Zeeshan, Q.; Solyali, D.; Korhan, O. Simulation of Factory 4.0: A Review. In *Industrial Engineering in the Digital Disruption Era. GJCIE 2019. Lecture Notes in Management and Industrial Engineering*; Calisir, F., Korhan, O., Eds.; Springer: Cham, Switzerland, 2020; [https://doi.org/10.1007/978-3-030-42416-9\\_19](https://doi.org/10.1007/978-3-030-42416-9_19).
- Zhang, W.; Yang, D.; Wang, H. Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Syst. J.* **2019**, *13*, 2213–2227, doi:10.1109/jsyst.2019.2905565.
- Borgi, T.; Hidri, A.; Neef, B.; Naceur, M.S. Data analytics for predictive maintenance of industrial robots. In Proceedings of the 2017 International Conference on Advanced Systems and Electric Technologies (IC\_ASET), Hammamet, Tunisia, 14–17 January 2017; pp. 412–417, doi:10.1109/ASET.2017.7983729.
- Dai, X.; Gao, Z. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2226–2238, doi:10.1109/TII.2013.2243743.
- Lee, J.; Lapira, E.; Bagheri, B.; Kao, H.A. Recent advances and trends in predictive manufacturing systems in big data environment. *Manuf. Lett.* **2013**, *1*, 38–41, doi:10.1016/j.mfglet.2013.09.005.
- Peres, R.S.; Rocha, A.D.; Leitao, P.; Barata, J. IDARTS—Towards intelligent data analysis and real-time supervision for industry 4.0. *Comput. Ind.* **2018**, *101*, 138–146, doi:10.1016/j.compind.2018.07.004.
- Sezer, E.; Romero, D.; Guedea, F.; MacChi, M.; Emmanouilidis, C. An Industry 4.0-Enabled Low Cost Predictive Maintenance Approach for SMEs. In Proceedings of the 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Stuttgart, Germany, 17–20 June 2018; pp. 1–8, doi:10.1109/ICE.2018.8436307.
- Biswal, S.; Sabareesh, G.R. Design and development of a wind turbine test rig for condition monitoring studies. In Proceedings of the 2015 International Conference on Industrial Instrumentation and Control (ICIC), Pune, India, 28–30 May 2015; pp. 891–896, doi:10.1109/IIC.2015.7150869.
- Wan, J.; Tang, S.; Li, D.; Wang, S.; Liu, C.; Abbas, H.; Vasilakos, A.V. A Manufacturing Big Data Solution for Active Preventive Maintenance. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2039–2047, doi:10.1109/TII.2017.2670505.
- Amruthnath, N.; Gupta, T. A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In Proceedings of the 2018 5th International Conference on Industrial Engineering and Applications, ICIEA, Singapore, 26–28 April 2018; pp. 355–361.
- Yoo, Y.; Park, S.H.; Baek, J.G. A Clustering-Based Equipment Condition Model of Chemical Vapor Deposition Process. *Int. J. Precis. Eng. Manuf.* **2019**, *20*, 1677–1689, doi:10.1007/s12541-019-00177-y.
- Susto, G.A.; Schirru, A.; Pampuri, S.; McLoone, S.; Beghi, A. Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Trans. Ind. Inform.* **2015**, *11*, 812–820, doi:10.1109/TII.2014.2349359.
- Susto, G.A.; Beghi, A.; De Luca, C. A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. *IEEE Trans. Semicond. Manuf.* **2012**, *25*, 638–649, doi:10.1109/TSM.2012.2209131.
- Ahmad, R.; Kamaruddin, S. An overview of time-based and condition-based maintenance in industrial application. *Comput. Ind. Eng.* **2012**, *63*, 135–149, doi:10.1016/j.cie.2012.02.002.
- Munirathinam, S.; Ramadoss, B. Big data predictive analytics for proactive semiconductor equipment maintenance. In Proceedings of the 2014 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 27–30 October 2014; pp. 893–902, doi:10.1109/BigData.2014.7004320.
- Rivera Torres, P.J.; Serrano Mercado, E.I.; Llanes Santiago, O.; Anido Rifón, L. Modeling preventive maintenance of manufacturing processes with probabilistic Boolean networks with interventions. *J. Intell. Manuf.* **2018**, *29*, 1941–1952, doi:10.1007/s10845-016-1226-x.
- Jezzini, A.; Ayache, M.; Elkhansa, L.; Makki, B.; Zein, M. Effects of predictive maintenance (PdM), Proactive maintenance (PoM) & Preventive maintenance (PM) on minimizing the faults in medical instruments. In Proceedings of the 2013 2nd International Conference on Advances in Biomedical Engineering, Tripoli, Lebanon, 11–13 September 2013; pp. 53–56, doi:10.1109/ICABME.2013.6648845.

20. Kumar, A.; Chinnam, R.B.; Tseng, F. An HMM and polynomial regression based approach for remaining useful life and health state estimation of cutting tools. *Comput. Ind. Eng.* **2019**, *128*, 1008–1014, doi:10.1016/j.cie.2018.05.017.
21. Mathew, V.; Toby, T.; Singh, V.; Rao, B.M.; Kumar, M.G. Prediction of Remaining Useful Lifetime (RUL) of turbofan engine using machine learning. In Proceedings of the 2017 IEEE International Conference on Circuits and Systems (ICCS), Thiruvananthapuram, India, 20–21 December 2017; pp. 306–311, doi:10.1109/ICCS1.2017.8326010.
22. Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **2000**, *44*, 206–226.
23. Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.D. Machine learning in manufacturing: Advantages, challenges, and applications. *Prod. Manuf. Res.* **2016**, *4*, 23–45, doi:10.1080/21693277.2016.1192517.
24. Testik, M.C. Expert Systems with Applications A review of data mining applications for quality improvement in manufacturing industry. *Expert Syst. Appl.* **2011**, *38*, 13448–13467, doi:10.1016/j.eswa.2011.04.063.
25. Qiao, W.; Member, S.; Lu, D.; Member, S. A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis—Part I: Components and Subsystems. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6536–6545, doi:10.1109/TIE.2015.2422112.
26. Lee, J.; Wu, F.; Zhao, W.; Ghaffari, M.; Liao, L.; Siegel, D. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mech. Syst. Signal Process.* **2014**, *42*, 314–334, doi:10.1016/j.ymsp.2013.06.004.
27. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510.
28. Kwak, D.S.; Kim, K.J. A data mining approach considering missing values for the optimization of semiconductor-manufacturing processes. *Expert Syst. Appl.* **2012**, *39*, 2590–2596, doi:10.1016/j.eswa.2011.08.114.
29. Deloitte. *Making Maintenance Smarter*; Deloitte University Press, New York, USA, 2017.
30. Cheng, J.C.P.; Chen, W.; Chen, K.; Wang, Q. Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Autom. Constr.* **2020**, *112*, 103087, doi:10.1016/j.autcon.2020.103087.
31. Deloitte Touche Tohmatsu Limited (DTTL). *Industry 4.0: Challenges and Solutions for the Digital Transformation and Use of Exponential Technologies*; Finance, Audit Tax Consulting Corporate: Zurich, Swiss, 2015.
32. Nasir, T.; Asmael, M.; Zeeshan, Q.; Solyali, D. Applications of Machine Learning to Friction Stir Welding Process Optimization Applications of Machine Learning to Friction Stir Welding Process Optimization. *J. Kejuruter.* **2020**, *32*, 171–186, doi:10.17576/jkukm-2020-32(2)-01.
33. Hakeem, A.A.A.; Solyali, D.; Asmael, M.; Zeeshan, Q. Smart Manufacturing for Industry 4.0 using Radio Frequency Identification (RFID) Technology. *J. Kejuruter.* **2020**, *32*, 31–38.
34. Centre, M.E. Machine-learning techniques and their applications in manufacturing. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2005**, *219*, 395–412, doi:10.1243/095440505X32274.
35. Monostori, L. AI and machine learning techniques for managing complexity, changes and uncertainties in manufacturing. *Eng. Appl. Artif. Intel.* **2003**, *16*, 277–291.
36. Flynn, P.J. Data Clustering: A Review. *ACM Comput. Surv.* **2000**, *31*, 264–323.
37. Merizalde, Y.; Hernández-Callejo, L.; Duque-Perez, O.; Alonso-Gómez, V. Maintenance models applied to wind turbines. A comprehensive overview. *Energies* **2019**, *12*, 1–41, doi:10.3390/en12020225.
38. MatlabWorks What Is Machine Learning? Available online: <https://www.mathworks.com/discovery/machine-learning.html> (accessed on 2 August 2020).
39. Jiang, Y.; Yin, S. Recursive Total Principle Component Regression Based Fault Detection and Its Application to Vehicular Cyber-Physical Systems. *IEEE Trans. Ind. Inform.* **2018**, *14*, 1415–1423, doi:10.1109/TII.2017.2752709.
40. You, D.; Gao, X.; Katayama, S. WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM. *IEEE Trans. Ind. Electron.* **2015**, *62*, 628–636, doi:10.1109/TIE.2014.2319216.
41. Orille-fernández, Á.L.; Khalil, N.; Rodríguez, S.B. Failure Risk Prediction Using Artificial Neural Networks for Lightning Surge Protection of Underground MV Cables. *IEEE Trans. Power Deliv.* **2006**, *21*, 1278–1282.
42. Mao, J. Why Artificial Neural Networks? *Computer.* 1996, *29*, 31–44, doi:10.1109/2.485891

43. Saritas, M.M.; Yasar, A. Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification. *Int. J. Intell. Syst. Appl. Eng.* **2019**, *7*, 88–91.
44. Gomes Soares, S.; Araújo, R. An on-line weighted ensemble of regressor models to handle concept drifts. *Eng. Appl. Artif. Intell.* **2015**, *37*, 392–406, doi:10.1016/j.engappai.2014.10.003.
45. Shin, J.H.; Jun, H.B.; Kim, J.G. Dynamic control of intelligent parking guidance using neural network predictive control. *Comput. Ind. Eng.* **2018**, *120*, 15–30, doi:10.1016/j.cie.2018.04.023.
46. Hesser, D.F.; Markert, B. Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. *Manuf. Lett.* **2019**, *19*, 1–4, doi:10.1016/j.mfglet.2018.11.001.
47. Scalabrini Sampaio, G.; de Vallim Filho, A.R.A.; da Santos Silva, L.; da Augusto Silva, L. Prediction of Motor Failure Time Using An Artificial Neural Network. *Sensors* **2019**, *19*, 4342, doi:10.3390/s19194342.
48. Falamarzi, A.; Moridpour, S.; Nazem, M.; Cheraghi, S. Prediction of tram track gauge deviation using artificial neural network and support vector regression. *Aust. J. Civ. Eng.* **2019**, *17*, 63–71, doi:10.1080/14488353.2019.1616357.
49. Zhang, Y.; Liu, J.; Hanachi, H.; Yu, X.; Yang, Y. Physics-based Model and Neural Network Model for Monitoring Starter Degradation of APU. In Proceedings of the 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), Seattle, WA, USA, 11–13 June 2018; pp. 1–7, doi:10.1109/ICPHM.2018.8448521.
50. Sexton, T.; Brundage, M.P.; Hoffman, M.; Morris, K.C. Hybrid datafication of maintenance logs from AI-assisted human tags. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1769–1777, doi:10.1109/BigData.2017.8258120.
51. Chang, C.C.; Lin, C.J. LIBSVM: A Library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 604–611, doi:10.1145/1961189.1961199.
52. Ahmad, W.M.T.W.; Ghani, N.L.A.; Drus, S.M. Data mining techniques for disease risk prediction model: A systematic literature review. *Adv. Intell. Syst. Comput.* **2019**, *843*, 40–46, doi:10.1007/978-3-319-99007-1\_4.
53. Parikh, U.B.; Das, B.; Maheshwari, R. Fault classification technique for series compensated transmission line using support vector machine. *Int. J. Electr. Power Energy Syst.* **2010**, *32*, 629–636, doi:10.1016/j.ijepes.2009.11.020.
54. DataFlair Team Support Vector Machines Tutorial—Learn to implement SVM in Python. *Data Flair* **2019**. Available online: <https://data-flair.training/blogs/svm-support-vector-machine-tutorial/> (accessed on 19 September 2020)
55. Widodo, A.; Yang, B.S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2560–2574.
56. Xiang, S.; Huang, D.; Li, X. A Generalized Predictive Framework for Data Driven Prognostics and Diagnostics using Machine Logs. In Proceedings of the TENCON 2018—2018 IEEE Region 10 Conference, Jeju, Korea, 28–31 October 2018; pp. 695–700, doi:10.1109/TENCON.2018.8650152.
57. Sugumaran, V.; Muralidharan, V.; Ramachandran, K.I. Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. *Mech. Syst. Signal Process.* **2007**, *21*, 930–942, doi:10.1016/j.ymsp.2006.05.004.
58. Rasoul, S.; David, L. A Survey of Decision Tree Classifier Methodology. *IEEE Trans. Syst. Man. Cybern.* **1991**, *21*, 660–674.
59. Silipo, R. From a Single Decision Tree to a Random Forest. *Dataiversity* **2019**. Available online: <https://www.dataiversity.net/from-a-single-decision-tree-to-a-random-forest/> (accessed on 19 September 2020)
60. Breiman, L. *Random Forests*; Working Paper CISL: Cambridge, UK, 2001; Volume 15, pp. 1–19; ISBN 9783110941975. doi: 10.1007/s12103-013-9223-5
61. Polamuri, S. How Does the Random Forest Algorithm Work in Machine Learning. Available online: <https://opendatascience.com/how-does-the-random-forest-algorithm-work-in-machine-learning/> (accessed on 30 July 2020).
62. Binding, A.; Dykeman, N.; Pang, S. Machine Learning Predictive Maintenance on Data in the Wild. In Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 15–18 April 2019; pp. 507–512, doi:10.1109/wf-iot.2019.8767312.
63. Zenisek, J.; Holzinger, F.; Affenzeller, M. Machine learning based concept drift detection for predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106031, doi:10.1016/j.cie.2019.106031.

64. Janssens, O.; Loccufer, M.; Van Hoecke, S. Thermal Imaging and Vibration-Based Multisensor Fault Detection for Rotating Machinery. *IEEE Trans. Ind. Inform.* **2019**, *15*, 434–444, doi:10.1109/TII.2018.2873175.
65. Lacaille, J.; Rabenoro, T. A Trend Monitoring Diagnostic Algorithm for Automatic Pre-identification of Turbofan Engines Anomaly. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing), Chongqing, China, 26–28 October 2018; pp. 819–823, doi:10.1109/PHM-Chongqing.2018.00146.
66. Quiroz, J.C.; Mariun, N.; Mehrjou, M.R.; Izadi, M.; Mison, N.; Mohd Radzi, M.A. Fault detection of broken rotor bar in LS-PMSM using random forests. *Meas. J. Int. Meas. Confed.* **2018**, *116*, 273–280, doi:10.1016/j.measurement.2017.11.004.
67. Yan, W.; Zhou, J.H. Predictive modeling of aircraft systems failure using term frequency-inverse document frequency and random forest. In Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, 10–13 December 2017; pp. 828–831, doi:10.1109/IEEM.2017.8290007.
68. ODSC Logistic Regression with Python. Available online: <https://medium.com/@ODSC/logistic-regression-with-python-ed39f8573c7> (accessed on 12 August 2020).
69. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 13–17 August 2016; pp. 785–794, doi:10.1145/2939672.2939785.
70. TIBCO Community Template for Using XGBoost in TIBCO Spotfire. Available online: [https://www.google.com/imgres?imgurl=https%3A%2F%2Feng.uber.com%2Fwp-content%2Fuploads%2F2019%2F12%2FTwitter.png&imgrefurl=https%3A%2F%2Feng.uber.com%2Fproductionizing-distributed-xgboost%2F&tbid=JkiedoDnC48yVM&vet=12ahUKEwiVjOOv\\_XqAhWWwoUKHedmD6IQMygLe](https://www.google.com/imgres?imgurl=https%3A%2F%2Feng.uber.com%2Fwp-content%2Fuploads%2F2019%2F12%2FTwitter.png&imgrefurl=https%3A%2F%2Feng.uber.com%2Fproductionizing-distributed-xgboost%2F&tbid=JkiedoDnC48yVM&vet=12ahUKEwiVjOOv_XqAhWWwoUKHedmD6IQMygLe) (accessed on 19 September 2020).
71. Cao, H.; Nguyen, M.N.; Phua, C.; Krishnaswamy, S.; Li, X.L. An integrated framework for human activity recognition. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, 5–8 September 2012, Pittsburgh, PA, USA, 2012; pp. 621–622, doi:10.1145/2370216.2370334.
72. Akira AI Gradient Boosting. Available online: <https://www.akira.ai/glossary/gradient-boosting/> (accessed on 30 July 2020).
73. Mattes, A.; Schopka, U.; Schellenberger, M.; Scheibelhofer, P.; Leditzky, G. Virtual Equipment for benchmarking Predictive Maintenance algorithms. In Proceedings of the 2012 Winter Simulation Conference (WSC), Berlin, Germany, 9–12 December 2012; doi:10.1109/WSC.2012.6465084.
74. Koza, J.R. Genetic programming as a means for programming computers by natural selection. *Stat. Comput.* **1994**, *4*, 87–112, doi:10.1007/BF00175355.
75. Janssens, O.; Van De Walle, R.; Loccufer, M.; Van Hoecke, S. Deep Learning for Infrared Thermal Image Based Machine Health Monitoring. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 151–159, doi:10.1109/TMECH.2017.2722479.
76. Huuhtanen, T.; Jung, A. Predictive Maintenance of Photovoltaic Panels via Deep Learning. In Proceedings of the 2018 IEEE Data Science Workshop (DSW), Lausanne, Switzerland, 4–6 June 2018; pp. 66–70, doi:10.1109/DSW.2018.8439898.
77. Pan, Z.; Ge, Y.; Zhou, Y.C.; Huang, J.C.; Zheng, Y.L.; Zhang, N.; Liang, X.X.; Gao, P.; Zhang, G.Q.; Wang, Q.; et al. Cognitive Acoustic Analytics Service for Internet of Things. In Proceedings of the 2017 IEEE International Conference on Cognitive Computing (ICCC), Honolulu, HI, USA, 25–30 June 2017; pp. 96–103, doi:10.1109/IEEE.ICCC.2017.20.
78. Jimenez-Cortadi, A.; Irigoien, I.; Boto, F.; Sierra, B.; Rodriguez, G. Predictive maintenance on the machining process and machine tool. *Appl. Sci.* **2020**, *10*, 224, doi:10.3390/app10010224.
79. Luo, W.; Hu, T.; Ye, Y.; Zhang, C.; Wei, Y. A hybrid predictive maintenance approach for CNC machine tool driven by Digital Twin. *Robot. Comput. Integr. Manuf.* **2020**, *65*, 101974, doi:10.1016/j.rcim.2020.101974.
80. Abu-Samah, A.; Shahzad, M.K.; Zamai, E.; Ben Said, A. Failure prediction methodology for improved proactive maintenance using Bayesian approach. *IFAC-PapersOnLine* **2015**, *28*, 844–851, doi:10.1016/j.ifacol.2015.09.632.
81. Carbery, C.M.; Woods, R.; Marshall, A.H. A new data analytics framework emphasising preprocessing of data to generate insights into complex manufacturing systems. *Proc. Inst. Mech. Eng. Part. C J. Mech. Eng. Sci.* **2019**, *233*, 6713–6726, doi:10.1177/0954406219866867.

82. Butte, S.; Prashanth, A.R.; Patil, S. Machine Learning Based Predictive Maintenance Strategy: A Super Learning Approach with Deep Neural Networks. In Proceedings of the 2018 IEEE Workshop on Microelectronics and Electron Devices (WMED), Boise, ID, USA, 20 April 2018; pp. 1–5, doi:10.1109/WMED.2018.8360836.
83. Susto, G.A.; Beghi, A. Dealing with time-series data in Predictive Maintenance problems. In Proceedings of the 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), Berlin, Germany, 6–9 September 2016; pp. 1–4, doi:10.1109/ETFA.2016.7733659.
84. Amihai, I.; Gitzel, R.; Kotriwala, A.M.; Pareschi, D.; Subbiah, S.; Sosale, G. An industrial case study using vibration data and machine learning to predict asset health. In Proceedings of the 2018 IEEE 20th Conference on Business Informatics (CBI), Vienna, Austria, 11–14 July 2018; Volume 1, pp. 178–185, doi:10.1109/CBI.2018.00028.
85. Paolanti, M.; Romeo, L.; Felicetti, A.; Mancini, A.; Frontoni, E.; Loncarski, J. Machine Learning approach for Predictive Maintenance in Industry 4.0. In Proceedings of the 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), Oulu, Finland, 2–4 July 2018; pp. 1–6, doi:10.1109/MESA.2018.8449150.
86. Kulkarni, K.; Devi, U.; Sirighee, A.; Hazra, J.; Rao, P. Predictive Maintenance for Supermarket Refrigeration Systems Using only Case Temperature Data. In Proceedings of the 2018 Annual American Control Conference (ACC), Milwaukee, WI, USA, 27–29 June 2018; pp. 4640–4645, doi:10.23919/ACC.2018.8431901.
87. Su, C.J.; Huang, S.F. Real-time big data analytics for hard disk drive predictive maintenance. *Comput. Electr. Eng.* **2018**, *71*, 93–101, doi:10.1016/j.compeleceng.2018.07.025.
88. Dos Santos, T.; Ferreira, F.J.T.E.; Pires, J.M.; Damasio, C. Stator winding short-circuit fault diagnosis in induction motors using random forest. In Proceedings of the 2017 IEEE International Electric Machines and Drives Conference (IEMDC), Miami, FL, USA, 21–24 May 2017; doi:10.1109/IEMDC.2017.8002350.
89. Canizo, M.; Onieva, E.; Conde, A.; Charramendieta, S.; Trujillo, S. Real-time predictive maintenance for wind turbines using Big Data frameworks. In Proceedings of the 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), Dallas, TX, USA, 19–21 June 2017; pp. 70–77, doi:10.1109/ICPHM.2017.7998308.
90. Prytz, R.; Nowaczyk, S.; Rögnvaldsson, T.; Byttner, S. Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. *Eng. Appl. Artif. Intell.* **2015**, *41*, 139–150, doi:10.1016/j.engappai.2015.02.009.
91. Nguyen, K.T.P.; Medjaher, K. A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliab. Eng. Syst. Saf.* **2019**, *188*, 251–262, doi:10.1016/j.res.2019.03.018.
92. Aydin, O.; Guldamlasioglu, S. Using LSTM networks to predict engine condition on large scale data processing framework. In Proceedings of the 2017 4th International Conference on Electrical and Electronic Engineering (ICEEE), Ankara, Turkey, 8–10 April 2017; 281–285, doi:10.1109/ICEEE2.2017.7935834.
93. Du, Y.; Chen, Y.; Meng, G.; Ding, J.; Xiao, Y. Fault severity monitoring of rolling bearings based on texture feature extraction of sparse time-frequency images. *Appl. Sci.* **2018**, *8*, 1–23, doi:10.3390/app8091538.
94. Guo, R.; Sui, J. Prognostics for an actuator based on an ensemble of support vector regression and particle filter. *Proc. Inst. Mech. Eng. Part. I J. Syst. Control. Eng.* **2019**, *233*, 642–655, doi:10.1177/0959651818806419.
95. Cárdenas-Gallo, I.; Sarmiento, C.A.; Morales, G.A.; Bolivar, M.A.; Akhavan-Tabatabaei, R. An ensemble classifier to predict track geometry degradation. *Reliab. Eng. Syst. Saf.* **2017**, *161*, 53–60, doi:10.1016/j.res.2016.12.012.
96. Lasisi, A.; Attoh-Okine, N. Principal components analysis and track quality index: A machine learning approach. *Transp. Res. Part. C Emerg. Technol.* **2018**, *91*, 230–248, doi:10.1016/j.trc.2018.04.001.
97. Mathew, J.; Luo, M.; Pang, C.K. Regression kernel for prognostics with support vector machines. In Proceedings of the 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 12–15 September 2017; pp. 1–5, doi:10.1109/ETFA.2017.8247740.
98. Arias, P.A. Planning Models for Distribution Grid: A Brief Review Piedy. *U. Prto J. Eng.* **2017**, *4*, 42–55, doi:10.24840/2183-6493.
99. Durbhaka, G.K.; Selvaraj, B. Predictive maintenance for wind turbine diagnostics using vibration signal analysis based on collaborative recommendation approach. In Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 21–24 September 2016; pp. 1839–1842, doi:10.1109/ICACCI.2016.7732316.

100. Machado, R.G.V.; De Oliveira Mota, H. Simple self-scalable grid classifier for signal denoising in digital processing systems. In Proceedings of the 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), Boston, MA, USA, 17–20 September 2015; pp. 1057–1061, doi:10.1109/GlobalSIP.2015.7418359.
101. Ahmad, W.; Khan, S.A.; Islam, M.M.M.; Kim, J.M. A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models. *Reliab. Eng. Syst. Saf.* **2019**, *184*, 67–76, doi:10.1016/j.ress.2018.02.003.
102. Ha, J.M.; Kim, H.J.; Shin, Y.S.; Choi, B.K. Degradation Trend Estimation and Prognostics for Low Speed Gear Lifetime. *Int. J. Precis. Eng. Manuf.* **2018**, *19*, 1099–1105, doi:10.1007/s12541-018-0130-z.
103. Duan, C.; Deng, C.; Wang, B. Optimal maintenance policy incorporating system level and unit level for mechanical systems. *Int. J. Syst. Sci.* **2018**, *49*, 1074–1087, doi:10.1080/00207721.2018.1432782.
104. Uhlmann, E.; Pontes, R.P.; Geisert, C.; Hohwieler, E. Cluster identification of sensor data for predictive maintenance in a Selective Laser Melting machine tool. *Procedia Manuf.* **2018**, *24*, 60–65, doi:10.1016/j.promfg.2018.06.009.
105. Amihai, I.; Chioua, M.; Gitzel, R.; Kotriwala, A.M.; Pareschi, D.; Sosale, G.; Subbiah, S. Modeling Machine Health Using Gated Recurrent Units with Entity Embeddings and K-Means Clustering. In Proceedings of the 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, 18–20 July 2018; pp. 212–217, doi:10.1109/INDIN.2018.8472065.
106. Eke, S.; Aka-Ngnui, T.; Clerc, G.; Fofana, I. Characterization of the operating periods of a power transformer by clustering the dissolved gas data. In Proceedings of the 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Tinos, Greece, 29 August–1 September 2017; pp. 298–303, doi:10.1109/DEMPED.2017.8062371.
107. Kumar, A.; Shankar, R.; Thakur, L.S. A big data driven sustainable manufacturing framework for condition-based maintenance prediction. *J. Comput. Sci.* **2018**, *27*, 428–439, doi:10.1016/j.jocs.2017.06.006.
108. Luo, B.; Wang, H.; Liu, H.; Li, B.; Peng, F. Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification. *IEEE Trans. Ind. Electron.* **2018**, *66*, 509–518, doi:10.1109/TIE.2018.2807414.
109. Kolokas, N.; Vafeiadis, T.; Ioannidis, D.; Tzovaras, D. Forecasting faults of industrial equipment using machine learning classifiers. In Proceedings of the 2018 Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki, Greece, 3–5 July 2018; pp. 1–6, doi:10.1109/INISTA.2018.8466309.
110. Garg, A.; Vijayaraghavan, V.; Tai, K.; Singru, P.M.; Jain, V.; Krishnakumar, N. Model development based on evolutionary framework for condition monitoring of a lathe machine. *Meas. J. Int. Meas. Confed.* **2015**, *73*, 95–110, doi:10.1016/j.measurement.2015.04.025.
111. Coraddu, A.; Oneto, L.; Ghio, A.; Savio, S.; Anguita, D.; Figari, M. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proc. Inst. Mech. Eng. Part. M J. Eng. Marit. Environ.* **2016**, *230*, 136–153, doi:10.1177/1475090214540874.
112. Manfre, M. Creation of a Machine Learning Model for the Predictive Maintenance of an Engine Equipped with a Rotating Shaft. Master's Thesis, Politecnico Di Torino, Torino, Italy, 2020.
113. Bekar, E.T.; Nyqvist, P.; Skoogh, A. An intelligent approach for data pre-processing and analysis in predictive maintenance with an industrial case study. *Adv. Mech. Eng.* **2020**, *12*, 1–14, doi:10.1177/1687814020919207.
114. Koca, O.; Kaymakci, O.T.; Mercimek, M. Advanced Predictive Maintenance with Machine Learning Failure Estimation in Industrial Packaging Robots. In Proceedings of the 2020 International Conference on Development and Application Systems (DAS), Suceava, Romania, 21–23 May 2020; pp. 1–6, doi:10.1109/DAS49615.2020.9108913.
115. Gohel, H.A.; Upadhyay, H.; Lagos, L.; Cooper, K.; Sanzetenea, A. Predictive Maintenance Architecture Development for Nuclear Infrastructure using Machine Learning. *Nucl. Eng. Technol.* **2020**, *52*, 1436–1442, doi:10.1016/j.net.2019.12.029.
116. Keartland, S.; Van Zyl, T.L. Automating predictive maintenance using oil analysis and machine learning. In Proceedings of the 2020 International SAUPEC/RobMech/PRASA Conference, Cape Town, South Africa, 29–31 January 2020; doi:10.1109/SAUPEC/RobMech/PRASA48453.2020.9041003.
117. Kaparathi, S.; Bumblauskas, D. Designing predictive maintenance systems using decision tree-based machine learning techniques. *Int. J. Qual. Reliab. Manag.* **2020**, *37*, 659–686, doi:10.1108/IJQRM-04-2019-0131.
118. Aremu, O.O. Achieving a Representation of Asset Data Conducive to Machine Learning driven Predictive Maintenance (PhD Thesis). The University of Queensland, Brisbane, Australia, 2019.

119. Fernandes, S.; Antunes, M.; Santiago, A.R.; Barraca, J.P.; Gomes, D.; Aguiar, R.L. Forecasting appliances failures: A machine-learning approach to predictive maintenance. *Information* **2020**, *11*, 1–13, doi:10.3390/INFO11040208.
120. Calabrese, M.; Cimmino, M.; Fiume, F.; Manfrin, M.; Romeo, L.; Ceccacci, S.; Paolanti, M.; Toscano, G.; Ciandrini, G.; Carrotta, A.; et al. SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0. *Information* **2020**, *11*, 1–17, doi:10.3390/INFO11040202.
121. Hoffmann, M.W.; Wildermuth, S.; Gitzel, R.; Boyaci, A.; Gebhardt, J.; Kaul, H.; Amihai, I.; Forg, B.; Suriyah, M.; Leibfried, T.; et al. Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. *Sensors (Switzerland)* **2020**, *20*, 1–24, doi:10.3390/s20072099.
122. Dangu, M.D.; Skaf, Z.; Jennions, I.K. An integrated machine learning model for aircraft components rare failure prognostics with log-based dataset. *ISA Trans.* **2020**, doi:10.1016/j.isatra.2020.05.001.
123. Karlsson, L. Predictive Maintenance for RM12 with Machine Learning. Master's Thesis, Halmstad University, Halmstad, Sweden, 2020.
124. Eriksson, J. Machine Learning for Predictive Maintenance on Wind Turbines. Master's Thesis, Linköping University, Linköping, Sweden, 2020.
125. G2 Best Data Science and Machine Learning Platforms. Available online: <https://www.g2.com/categories/data-science-and-machine-learning-platforms> (accessed on 13 August 2020).
126. Seebo, Why Predictive Maintenance is Driving Industry 4.0: The Definitive Guide; 2019, pp. 1–13. Available online: <https://files.solidworks.com/partners/pdfs/why-predictive-maintenance-is-driving-industry-4.0405.pdf>, 15.08.2020.



© 2020 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/>).