Int J Syst Assur Eng Manag (April 2024) 15(4):1345–1360
https://doi.org/10.1007/s13198-022-01843-7

REVIEW PAPERS

Condition-based maintenance using machine learning and role of interpretability: a review

Jeetesh Sharma[1] · **Murari Lal Mittal**[1] · **Gunjan Soni**[1]

Received: 30 March 2022 / Revised: 14 October 2022 / Accepted: 13 December 2022 / Published online: 20 December 2022

Abstract This article aims to review the literature on condition-based maintenance (CBM) by analyzing various terms, applications, and challenges. CBM is a maintenance technique that monitors the existing condition of an industrial asset to determine what maintenance needs to be performed. This article enlightens the readers with research in condition-based maintenance using machine learning and artificial intelligence techniques and related literature. A bibliometric analysis is performed on the data collected from the Scopus database. The foundation of a CBM is accurate anomaly detection and diagnosis. Several machine-learning approaches have produced excellent results for anomaly detection and diagnosis. However, due to the black-box nature of the machine learning models, the level of their interpretability is limited. This article provides insight into the existing maintenance strategies, anomaly detection techniques, interpretable models, and model-agnostic methods that are being applied. It has been found that significant work has been done towards condition based-maintenance using machine learning, but explainable artificial intelligence approaches got less attention in CBM. Based on the review, we contend that explainable artificial intelligence can provide unique insights and opportunities for addressing critical difficulties in maintenance leading to more informed decision-making. The analysis put forward encouraging research directions in this area.

Keywords Condition-based maintenance · Machine learning · Anomaly detection · Explainable artificial intelligence · Interpretability · Model-agnostic methods

Abbreviations

AI	Artificial intelligence
ALE	Accumulated local effects
ANFIS	Adaptive neuro-fuzzy inference
ANN	Artificial neural networks
CBM	Condition-based maintenance
CNN	Convolutional neural networks
DT	Decision trees
HBOS	Histogram-based outlier score
ICE	Individual conditional expectation
IF	Isolation Forest
KNN	K-nearest neighbor
LOF	Local outlier factor
LR	Linear regression
ML	Machine Learning
NN	Neural networks
PCA	Principal component analysis
PdM	Predictive maintenance
PDP	Partial dependence plot
PHM	Prognostics and health management
PM	Planned maintenance
RF	Random Forest
RM	Reactive maintenance
RNN	Recurrent neural networks
rPCA	Robust principal component analysis
RUL	Remaining useful life
RVM	Relevance vector machine
SM	Scheduled maintenance

✉ Jeetesh Sharma
2019rme9528@mnit.ac.in

Murari Lal Mittal
mlmittal.mech@mnit.ac.in

Gunjan Soni
gsoni.mech@mnit.ac.in

1 Department of Mechanical Engineering, Malaviya National Institute of Technology, Jaipur 302017, India

SVM Support vector machine
TC Total citations
XAI Explainable artificial intelligence

1 Introduction

Maintenance is crucial in improving asset availability in manufacturing industries. The appropriate choice of maintenance strategy is essential for minimizing costs and downtime. Successive improvement approaches have been tried to overcome the inefficiencies encountered by the previous maintenance approaches (Prajapati et al. 2012). The maintenance strategies have been divided into Reactive maintenance (RM), Scheduled maintenance (SM), Condition-based maintenance (CBM), and predictive maintenance (PdM) (Jardine et al. 2006; Fink 2020).

Reactive maintenance (RM) occurs when a machine component is failed and can no longer operate. The failed machine component needs to be replaced or repaired (Paz & Leigh 1994) for the machine to work again. The main advantage of the RM strategy is that the expenses associated with keeping machines running are low (Swanson 2001; Vanzile & Otis 1992). But, this strategy is risky from the point of view of safety measures and higher costs to restore the catastrophic failures, and a higher amount of time to be repaired. Scheduled maintenance (SM) is a strategy where maintenance is carried out at predecided time intervals. It comprises inspections, adjustments, and planned shutdowns. This strategy depends on the machine components' probability of failing in specified time intervals (Gits 1992). The SM strategy's primary goals are to reduce the cost of reactive maintenance and machine failures. The repair cost, however, is generally less in SM compared to RM. This strategy relies on a predetermined degradation model of a particular machine which sometimes may lead to missed faults caused by various factors. Condition-based maintenance (CBM) and Predictive maintenance (PdM) strategies are different from RM and SM as they are data-driven approaches that help operators in setting times for maintenance activities. Condition-based maintenance involves continuously monitoring an asset and replacing it when it stops functioning normally. On the other hand, Predictive maintenance (PdM) uses proactive, data-driven maintenance techniques to assess equipment conditions and determine when maintenance is necessary. PdM predicts the remaining useful life of a component to designate the point when maintenance has to be performed. PdM results in reduced maintenance costs compared to CBM (Carvalho et al. 2019a, b; Ran et al. 2019).

The CBM and PdM are considered to be an integral part of modern smart manufacturing or Industry 4.0. Industry 4.0 integrates computer science engineering with mechanical and electrical engineering to bring advancement at the technological level (Lasi et al. 2014). It requires manufacturing companies to collect and store operations data to monitor several parameters affecting machinery conditions and detect warning signals in case of breakdown (Bousdekis et al. 2015). Some of the key players in the Industry 4.0 revolution are prognostic and health management (PHM) systems. These systems aim to solve the problems of effectively detecting whether an industrial component has deviated from its normal operating condition or predicting when a fault will occur and execute an intelligent maintenance approach through real-time monitoring and data analysis. Prognostics and health management comprises condition-based and predictive maintenance (Fig. 1). CBM refers to anomaly detection and fault diagnosis. The motive of anomaly detection is to find out whether the health of the system is normal or not. The diagnosis refers to finding out the deviation from the normal behavior of the system and its corresponding degree. Prognosis refers to identifying the trend of deterioration and estimating the remaining useful life (RUL) (Z. Zhao et al. 2021). Effective PHM methods promise to reduce the likelihood of catastrophic failures, thereby enhancing the safety of industrial machines.

As stated previously, as a key player in the fourth industrial revolution, PHM utilizes some of the most recent advancements made in computer science engineering over the past few years. Among them, machine learning is arguably one of the technologies receiving the most remarkable growth. Machine learning is an essential and fundamental technical achievement that enables Industry 4.0 and plays a critical role in industrial innovation by delivering solutions to various difficulties (Ayvaz & Alpay, 2021). Classical machine learning methods such as artificial neural networks (ANN), support vector machine (SVM), K-nearest neighbor (KNN), decision trees (DT), etc., have been favorably implemented in artificial intelligence-enabled maintenance (Wuest et al. 2016; W. Zhang

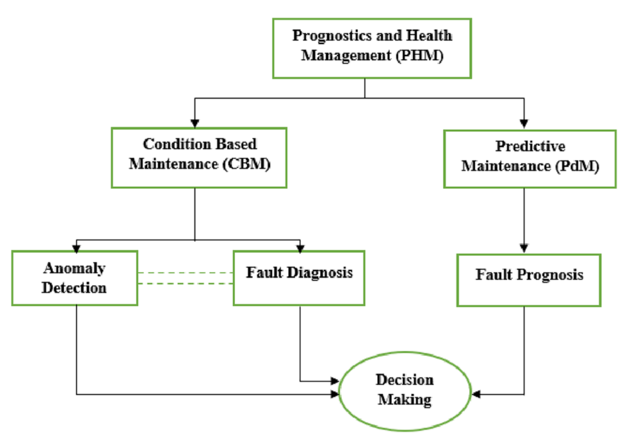


Fig. 1 Classification of prognostics and health management process (Z. Zhao et al. 2021)

et al. 2019a, b). Deep learning (DL) is a subset of machine learning built on artificial neural networks. DL methods such as convolutional neural networks (CNN), recurrent neural networks (RNN), deep reinforcement learning, autoencoders, etc., are emerging as important approaches in the maintenance of various equipment and system (Amihai et al. 2018; Huuhtanen and Jung 2018; Janssens et al. 2017).

Advancements in machine learning techniques have achieved an influential accuracy benchmark in recent years (LeCun et al. 2015). But this has resulted in higher model complexity. Consequently, model interpretability and explainability are reduced. Therefore, it is hard for the users to understand the prediction results. Hence, these models are considered "black-box models" (Rudin C. 2019). Explainable artificial intelligence (XAI) generates more interpretable and user-understandable explanations. The ultimate goal of the XAI approach is to create and modify the existing machine learning techniques with an adequate explanation so as the end-user, who is dependent on the recommendations of the artificially intelligent systems, can understand the overall behavior, weakness, and strength of the system (Gunning 2019). Interpretability is the extent to which an end-user can acknowledge the source of a decision (Kim et al. 2016). It is sufficient to know about prediction accuracy in a low-risk environment. However, there are cases where it is essential to see the model's explanation of how it came to that prediction (Doshi-Velez & Kim 2017). The motivation of interpretability is to create interpretable models, explain the model's complexity, increase the model's fairness, and check the sensitivity of the predictions.

In this review article, we make the consecutive contributions:

1. We illustrate anomaly detection techniques in the context of CBM and familiarize the readers with it based on the literature using bibliometric analysis.
2. We identify XAI techniques as an essential study field for upkeep in unsolved research problems.
3. To assess the scope of XAI in maintenance, we discuss interpretable models and model-agnostic approaches that are being used and provide their merits and disadvantages in various aspects.

This paper is organized as follows; Sect. 2 discusses the data gathering and research methods and includes a brief discussion of research growth, relevant sources, keyword analysis, etc. Different anomaly detection methods based on machine learning types are described in Sect. 3. The problems with black-box machine learning models are illustrated in Sect. 4, which also briefly introduces interpretable and model-agnostic methods. Section 5 winds up the paper.

2 Data collection and research methodology

We collected the data from the most preferred archive: Scopus, and bibliometric analysis was performed. This analysis reveals the trends in the publication from the perspective of research advancement, which helps in choosing the best journal, crucial keywords, and the most relevant sources.

The keywords used for the research are: "Condition-based maintenance" or "condition-based monitoring" and "learning" performed on 1st October 2021 (Search within article title, abstract, keywords). We considered the data from 2010 to 2021, and the subject area is limited to engineering. Scopus showed 261 documents. Out of the 261 papers, 128 (49.0% of the total documents) were articles. Other categories were conference paper (110), review (10), conference review (7), book chapter (5), and book (1) (Table 1).

2.1 Research growth

CBM, with machine learning, has been gaining expeditious attention since its establishment. Figure 2 shows the total year-wise number of publications in Scopus. As shown in Fig. 2, the publications from 10 in 2015 reached more than 60 in 2020. The search was performed on 1st October 2021, with three months remaining in the year, and covid 19 pandemic could be two of the reasons for the drop in the publications as in the last part of the chart.

2.2 Most relevant sources

In this section, we show (Table 2) the top 10 relevant sources of the publications in the area, sorted by the number of articles. As can be seen, IEEE access is at the top with 13 articles, while Lecture Notes in networks and systems stands at the bottom with four articles. This analysis represents the exploration that CBM has acquired through these publications over the years.

Table 1 List of document types

Document types	Numbers
Article	128
Book	1
Book chapter	5
Conference paper	110
Conference review	7
Review	10
Total documents	261

Fig. 2 The number of documents published by year

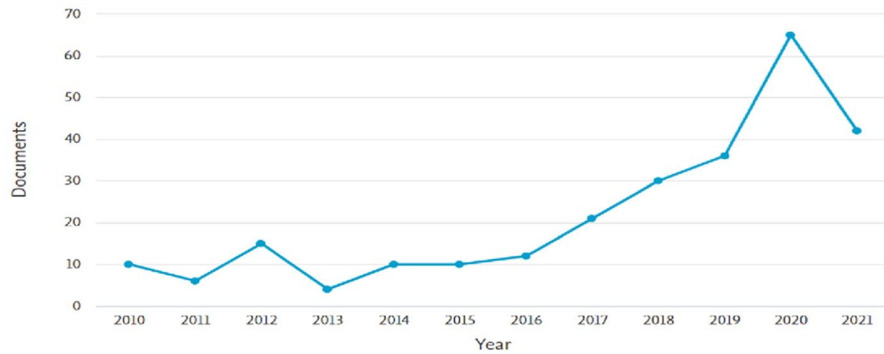


Table 2 List of top 10 most relevant sources

Sources	Articles
IEEE access	13
Mechanical systems and signal processing	12
Ocean engineering	6
Expert systems with applications	5
Lecture notes in mechanical engineering	5
Proceedings of the annual conference of the prognostics and health management society phm	5
Applied sciences (Switzerland)	4
Computers in industry	4
IEEE transactions on reliability	4
Lecture notes in networks and systems	4

Fig. 3 Word Cloud

Table 3 Country scientific production

Region	Freq
USA	93
China	92
Italy	34
France	33
UK	26
Canada	18
Brazil	17
Spain	15
Germany	13
Singapore	12

2.3 Country-specific production

This section lists the top 10 countries according to the total number of publications (Table 3). The USA is at the top of the list with 93 publications, and Singapore is at the bottom with 12 publications.

2.4 Keyword analysis

Figure 3 shows a word cloud which indicates the prominence of the word as per its appearance in the manuscript. It is noted here that "condition-based maintenance" and "learning systems" are two of the most relevant keywords.

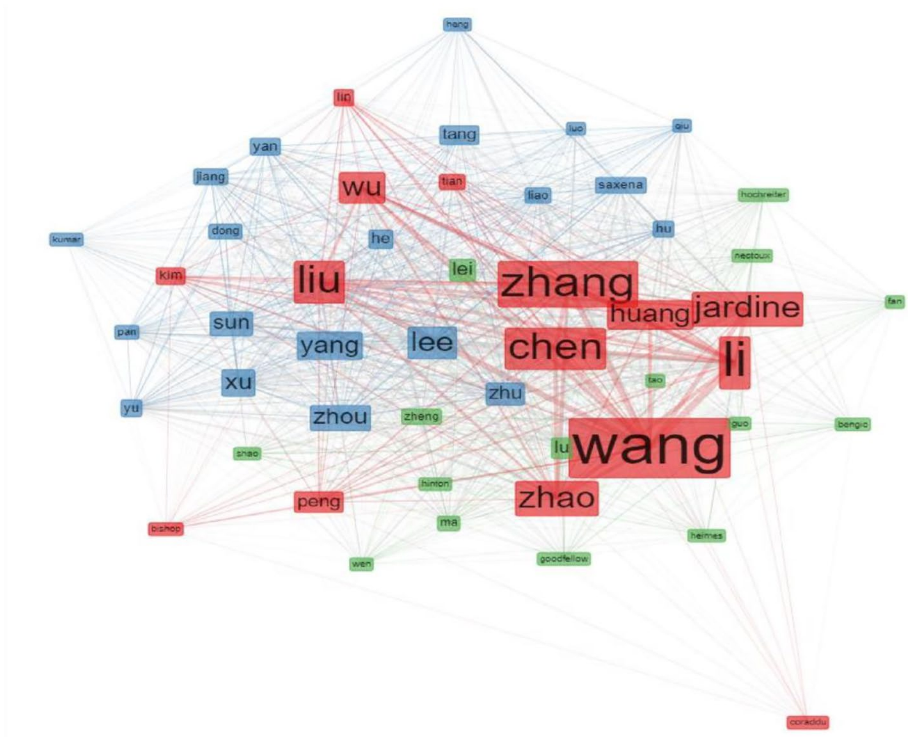
2.5 Co-Citation network

Co-citation is stated as the recurrence with which two articles are cited together by other articles. Figure 4 represents the prominent co-cited authors of CBM research. By co-citation analysis, the predominant knowledge bases of the CBM can be established proficiently and effortlessly from cited references.

3 Condition-based maintenance (CBM)

The maintenance of an industrial asset is an indispensable exercise in the production process. Condition-based maintenance (CBM) is a strategy that observes the health condition

Fig. 4 Co-citation network



of machine elements over time to recognize and prevent possible failures (Vachtsevanos et al. 2006). Planned maintenance (PM) is accomplished through predetermined scheduled intervals, whereas CBM is conducted after a decline in the condition of the equipment. This maintenance strategy effectively reduces the cost of machine failures, minimizes the chances of interference with normal operations, reduces unscheduled downtime due to failures, and enhances machinery reliability (de Jonge et al. 2017).

CBM is a maintenance strategy that upholds maintenance decisions depending on the condition monitoring process (Jardine et al. 2006; Grall et al. 2002). Figure 5 illustrates CBM and PdM's main components, i.e., anomaly detection, fault diagnosis, and prognosis. In CBM, the equipment's operative condition is monitored using sensors to collect various parameters such as vibration, temperature, acoustic emission, lubricating oil, etc. (Campos 2009). CBM follows the data acquisition, extraction, and preprocessing processes to uphold decision-making through information regarding a system's health condition. The information derived from the anomaly detection and fault diagnosis process can be used at the PdM level to describe system health conditions in a better way.

3.1 Anomaly detection

Anomaly detection helps in finding the patterns in the particular data that do not match the expected behavior. These non-compliant arrangements are called anomalies, outliers, exceptions, etc., in different domains (Chandola et al. 2009). An anomaly is detected when the input data show dissimilarities compared to normal machine conditions (Khan and Yairi 2018). Finding an anomaly does not imply erroneous data because it might be due to a new healthy feature that the anomaly detection algorithm has not acclimated to or does not relate to historical data (Yue Zhao et al. 2019). Anomaly detection techniques have been favorably applied for fault revelation in manufacturing (Schneider and Frank 1996), fraud detection in the financial sector (Sadgali et al. 2019), intrusion detection in a computer network (García-Teodoro et al. 2009), finding the potential risk in the health sector (Antonelli et al. 2013), oil and gas industries (Shahir et al. 2015), etc.

Supervised anomaly detection approaches need a dataset labeled as "faulty" and "normal". Unsupervised anomaly detection approaches discover formerly unseen data points without any prior understanding. A minimal amount of dataset represents an anomaly (usually less than 1%). Hence there are insufficient examples to learn for a supervised learning approach. Therefore, unsupervised learning algorithms are trendy for finding an anomaly. Labels are accessible only for normal classes in semi-supervised learning (Kiran et al. 2018). Here we discuss some widespread industrial anomaly detection approaches. Figure 6 illustrates various anomaly detection approaches based on various machine learning approaches.

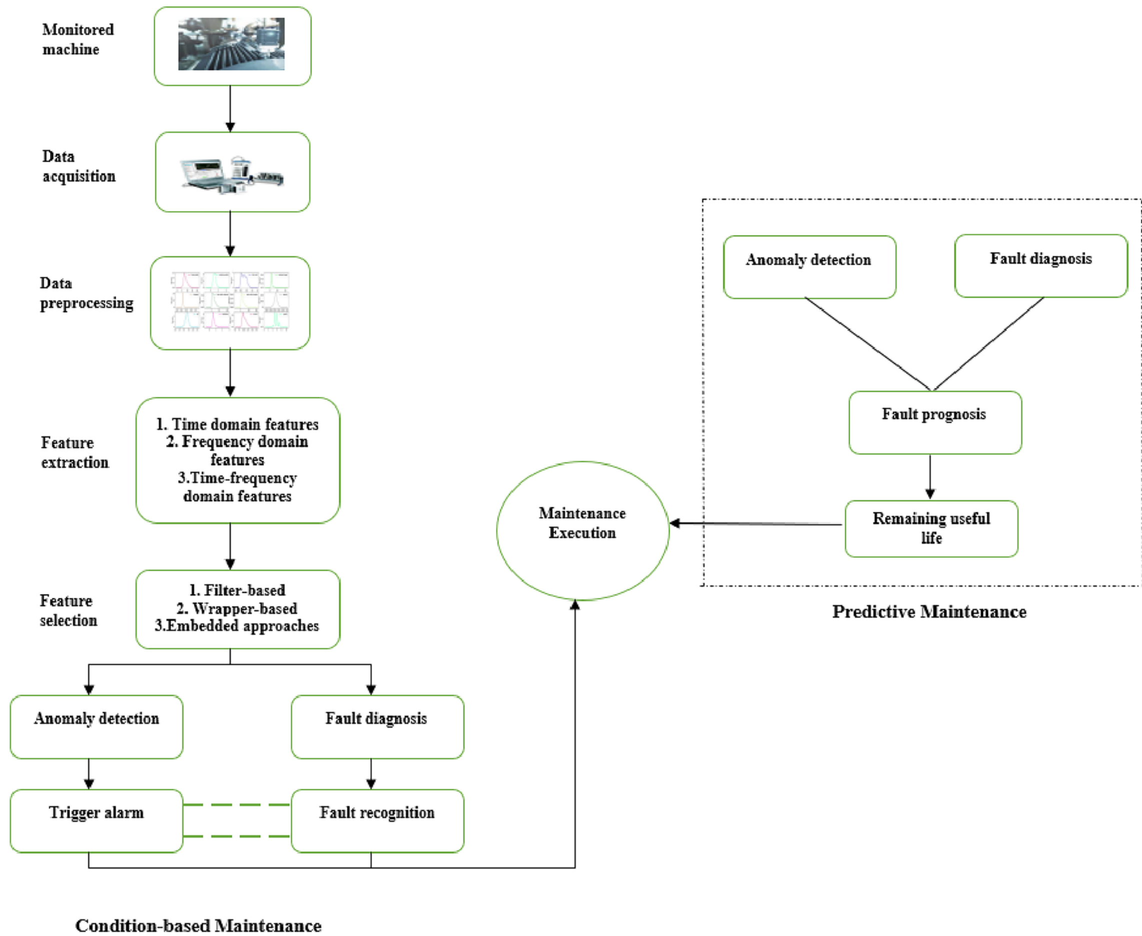
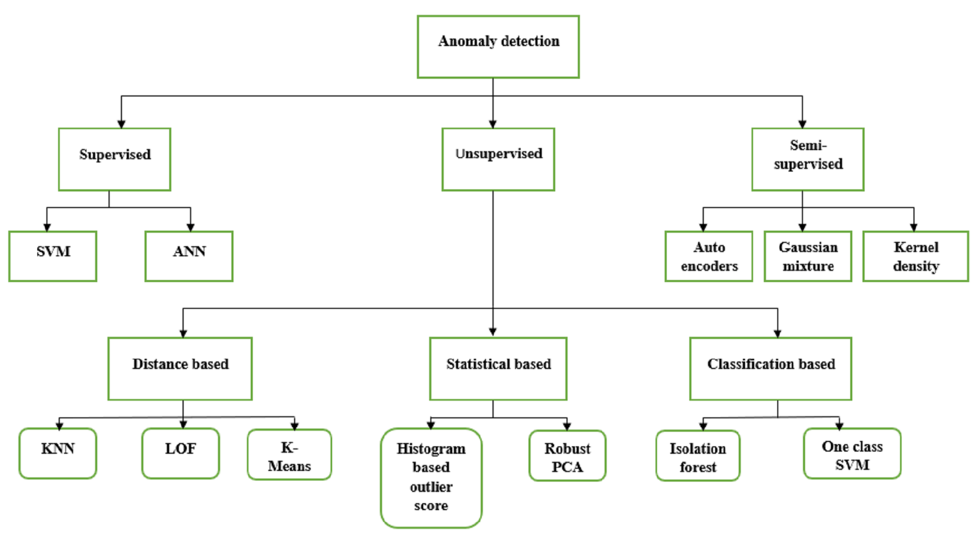


Fig. 5 Flow diagram of condition-based maintenance and predictive maintenance

Fig. 6 Types of anomaly detection techniques



3.1.1 Supervised anomaly detection techniques

Support vector machine (SVM) is the most used supervised learning algorithm in multiclass classification problems. The hyperplane with the highest margin is used in multidimensional space to disparate one class from another (X. Zhang et al. 2006). Classical SVM, and its improved versions, have been successfully implemented for anomaly detection. For example, power signals along with vibrations were used for anomaly detection in wind turbines using SVM (Santos et al. 2015), and a combination of SVM and adaptive neuro-fuzzy inference (ANFIS) was used for fault detection in the case of the steam turbine (Salahshoor et al. 2010), a fusion of relevance vector machine (RVM) and SVM was used for detecting a fault in the case of low-speed bearing (Widodo et al. 2009).

An artificial neural network (ANN) is the constituent of AI that is intended to simulate the human brain. Neurons are connected by the nodes and carry the information. ANN has been implemented in many types of machines for fault detection. A feedforward neural network with a backpropagation algorithm is used to detect the fault in a centrifugal pump (Rajakarunakaran et al. 2008). The ANN approach is developed to identify defects in the cooler water spray system (Subbaraj and Kannapiran 2010).

3.1.2 Unsupervised anomaly detection techniques

3.1.2.1 Distance-based approaches The fundamental assumption of distance-based anomaly detection approaches is that expected data points have a dense region, and outliers are far from their neighbors. K-nearest neighbor is a supervised machine learning approach, but it follows an unsupervised method for anomaly detection because there is no preestablished inlier or outlier; instead, it is based on threshold values. The fundamental of KNN is that similar data points are near each other, and outliers are away from similar data points clusters. The technique has been applied, for instance, in the slot milling cutting tool (Liu et al. 2022), semiconductor manufacturing process (Subbaraj and Kannapiran 2010), motor bearing (Tian et al. 2015), combustion engine (Jafarian et al. 2018), gas sensor arrays (J. Yang et al. 2016), power transformers (Islam et al. 2017), reciprocating compressor (Patil et al. 2022) among others.

K-means clustering is another unsupervised approach in which data points are grouped into distinct clusters. Here K denotes the number of predefined clusters so that all dataset belongs to one group with similar properties. Every cluster is affiliated with a centroid. A threshold value is added to detect outliers. A case is considered an outlier if the interspace between the data point and its nearby centroid exceeds the threshold value. K-means clustering has been extensively applied in fault detection in various machines. It has been used, e.g., for fault detection of commutator motors (Glowacz 2019), rolling element bearing (Yiakopoulos et al. 2011; Yu et al. 2021), wind turbines (H. H. Yang et al. 2015; W. Zhang & Ma 2016), heat pump air-conditioning systems (H. Zhang et al. 2019a, b), cutting tools (Lahrache et al. 2017), stream turbines (Yao et al. 2022), among others.

The local outlier factor (LOF) is a density-based unsupervised anomaly detection approach for finding local anomalies. It calculates the local density variation of a data point corresponding to its neighbors. A sample with a lower density than its neighbors is considered an outlier. LOF has been extensively implemented for anomaly detection. For example, it has been applied to fault detection of batteries for electric vehicles (Yang Zhao et al. 2017), detection of abnormal rail wear (Famurewa et al. 2017), detection of abnormal behavior in lithium-ion batteries (Diao et al. 2020; Fan et al. 2022), anomaly detection in the diffusion process of semiconductor manufacturing (Chang et al. 2021), etc.

3.1.2.2 Statistically based approaches Statistics-based anomaly detection approaches fit a statistical model for the expected behavior of given data points which can be used to determine whether the unseen data point belongs to this model or not. Histogram-based outlier score (HBOS) is an unsupervised learning approach. HBOS determines the degree of outliers by constructing histograms for every feature, and then histogram density is measured for every feature (Chang et al. 2021). This has been applied to detect anomalies in multiphase flow meters used in the oil and gas industries (Barbariol et al. 2019).

Principal component analysis (PCA) is a dimensionality reduction approach. It converts the multidimensional data into a lower dimension for more straightforward analysis (Sapra 2010). Traditional PCA is sensitive to anomalies and can mislead in results in the existence of anomalies. To reduce the sensitivity, a covariance matrix is replaced by its robust variants in robust principal component analysis (rPCA) (Tharrault et al. 2008). The rPCA model-based approach is introduced to detect faults for helical coil steam generator systems (K. Zhao & Upadhyaya 2006) and fault detection in weld inspection (Cassels et al. 2019).

3.1.2.3 Classification based approaches The training stage learns a classifier using the available training data in the classification-based anomaly detection approach. The testing stage classifies the unseen data point as normal or abnormal. Classification-based anomaly detection approaches are divided into one-class and multiclass anomaly detection approaches. In one-class anomaly detection techniques, it is presumed that entire training data points have only one class label, and a borderline around the normal data points is formed. If any instance falls outside the borderline, it is considered an anomaly (Schölkopf et al. 2001). The multiclass anomaly detection technique presumes that training data points belong to different normal classes, and classifiers learn to differentiate between all normal classes. If any instance is not considered normal by any classifier, it is considered an anomaly (Barbará et al. 2001).

One-class support vector machine (One class SVM) is an unsupervised anomaly detection approach that learns a decision function for anomaly detection. A segmentation algorithm with one-class SVM is introduced to detect anomalies in petroleum industry turbomachines (Martí et al. 2015). A fault detection approach is introduced using one-class SVM in electro-mechanical machines from vibration quantification (Shin et al. 2005).

Isolation Forest (IF) is an unsupervised learning approach based on decision trees. Subsampled data is prepared in a tree form based on the selected features. The samples finish up in shorter branches and designate anomalies. Thus, it is easy to separate anomalies from the tree. The samples that go deeper into the tree are less likely to be anomalies. Isolation forest has been successfully applied in various areas. For instance, an integrated hybrid manifold learning and IF technique is developed for fault monitoring in marine diesel engines (R. Wang et al. 2021), anomaly detection technique using IF in the diffusion process of semiconductor manufacturing is introduced (Chang et al. 2021), IF based fault detection approach is developed for hydroelectric generators (Hara et al. 2020), IF based fault detection technique is developed for fault detection in heavy haul railway operations (Oliveira et al. 2019), etc.

3.1.3 Semi-supervised anomaly detection techniques

Semi-supervised techniques combine supervised and unsupervised learning processes where unlabelled data is used for training a model. Autoencoders are a specific type of neural network where the output is the same as the input. An autoencoder comprises two main components: an encoder that plans the input into the code and a decoder that plans the code to reform the input. The key idea behind autoencoders is to determine a low-level representation of the input data. Autoencoders have been extensively applied for fault detection in various areas. For example, it has been used for fault detection of bearings (Sun et al. 2017; H. Liu et al. 2018a, b; Meng et al. 2018; C. Li et al. 2017; X. Li et al. 2020; Sohaib and Kim 2018), gearboxes (Jiang et al. 2017; G. Liu et al. 2018a, b; Yu 2019), electric motors (Principi et al. 2019), gas turbine (Luo and Zhong 2017), transformers (Ou et al. 2019), induction motors (J. Wang et al. 2017), building automation systems (Choi and Yoon 2021), etc.

The gaussian model-based anomaly detection approach assumes that the data arise from a Gaussian distribution. By using maximum likelihood estimates, a gaussian can be fit. The distance calculated from the mean in standard deviation is an anomaly score for a data point. The gaussian mixture model is used for fault detection in industrial gas turbines (Y. Zhang et al. 2017). A vital advantage of the gaussian mixture models is their applicability for anomaly detection when there is insufficient foreknowledge of fault patterns.

4 Open challenges

4.1 Interpretability

One of the strong condemnations of most machine learning techniques is their black-box nature. An absolute mathematical description of most machine learning and deep learning approaches is tough to acquire. This pessimistic property of the machine learning approaches represents a notable limitation in maintenance. The problem is that a single metric, such as classification accuracy, is an incomplete description of most real-world tasks (Ribeiro et al. 2016a).

Figure 7 shows the trade-off between the model's interpretability and predictive accuracy (Morocho-Cayamcela et al. 2019). A simple Linear Regression (LR) model has the highest level of interpretability, but the predictive accuracy level is generally low. That is why simple LR models are straightforward to explain. In contrast, Neural Networks (NN) are very effective in predictive accuracy, but they are considered black-box models, so they can not be interpreted.

In predictive modeling, it is not enough to know what is predicted. Instead, we want to see why this particular prediction was made. In some cases, it is adequate to see the prediction results only, not the explainability of the results, because of the low-risk environment. Nevertheless, in other cases, the explainability of models helps to understand the data and the problem more precisely. Whenever there is an inadequacy of problem clarification, the need for explainability arises because the prediction results only moderately

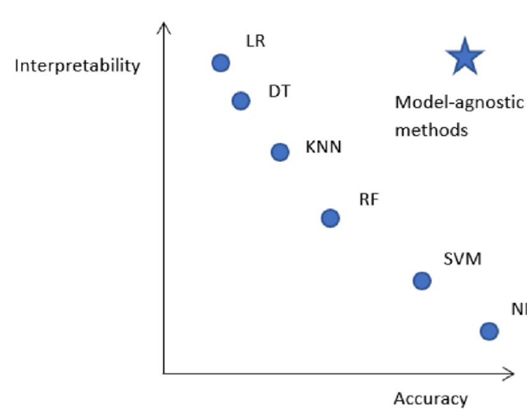


Fig. 7 Interpretability v/s Accuracy of different machine learning algorithms

solve the issues. The following are the needs for interpretability and explainability:

Human understanding and learning: The fundamental goal of humans is to find out the meaning and gain knowledge. Applicants may be unsatisfactory or objectionable when a particular machine learning model rejects something or anticipates low prediction accuracy.

Scientific understanding and learning: In today's environment, most problems have an extensive dataset and are solved with black-box machine learning models. Interpretability and explainability extract further knowledge acquired by the model.

Safety measures: It is tough to create whole scenarios where the structure may fail in complex tasks. Listing all inputs and outputs are analytically infeasible, and we can not indicate all unenviable outputs.

If a machine learning model has explainability, then the model contains fairness because of unbiased predictive results, privacy due to data protection, and reliability because if we make minor changes in inputs, it does not lead to significant changes in outputs and trust because a human can trust explainable model compared to a black-box model. Figure 8 represents the difference between a standard machine learning model and an interpretable machine learning model. Because of the interpretable results, human feedback improves the data and model, which gives more substantial predictions.

In machine learning and AI, interpretability and explainability are generally used interchangeably in machine learning systems. Explainability is the extent to which humans can understand the internal mechanism of a complex machine learning system. In comparison, interpretability is the extent to which we can predict what will happen if the model's input parameters have been changed.

4.2 Interpretable models

The simplest way to accomplish interpretability is to use algorithms that generate interpretable models. Logistic regression, Linear regression, Generalized Linear Models, Generalized Additive Models (Hastie 2017), Decision trees (Kingsford and Salzberg 2008), and Decision rules (Apté and Weiss 1997) are some of the most widely used interpretable models. Table 4 shows the expected advantages and disadvantages of interpretable models.

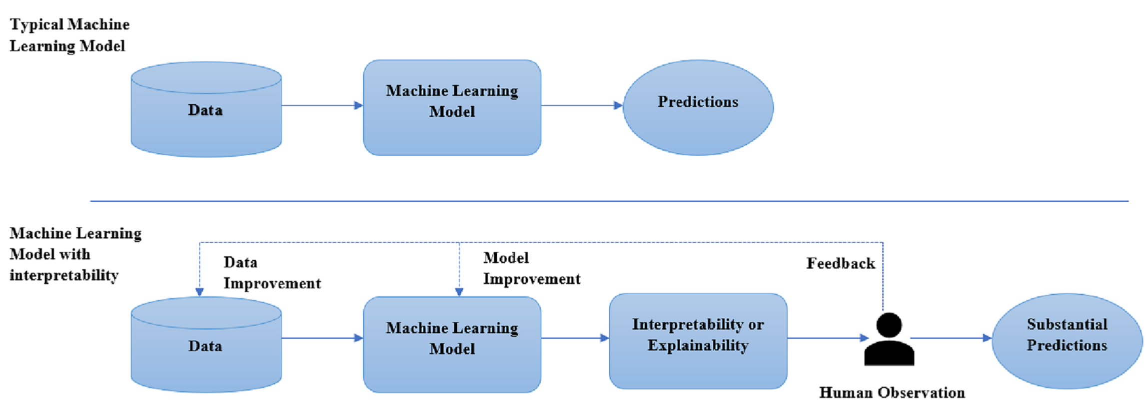


Fig. 8 Typical machine learning model and an interpretable machine learning model

Table 4 Advantages and disadvantages of interpretable models

Model	Advantages	Disadvantages
Linear regression	Transparent about the prediction, Assurance to detect optimal weights	Only perform linear relationships, Not up to the mark concerning predictive performance
Logistic regression	Not just a classification model provides probabilities; also, it can be used for multiclass classification	Restrictive expressiveness, Explainability is difficult due to multiplicative weights
Generalized linear models and generalized additive models	Problems faced in linear models can be fixed in these extensions, Models can be used for conclusion in addition to prediction	Less interpretable, Depending upon the assumptions of data generating procedure
Decision trees	Instinctive visualization with nodes, Generates better explanations, Easy to acknowledge	Fails to handle the linear relationship, Insufficiency of smoothness, Unstable because the entire tree structure changes with different feature
Decision rules	Easy to elucidate, Fast prediction, More compact than decision trees, Only pertinent features are selected	Used for classification, not for regression, Generally used for categorical features, Fails to handle the linear relationship

4.3 Model-agnostic methods

The main idea of the model-agnostic methods is to split up the elucidations from the machine learning models (Ribeiro et al. 2016a). One critical recognition of the model-agnostic methods over model-specific interpretable methods is the applicability to any model and its flexibility. Model-agnostic methods can be used for any kind of model. The main disadvantage of using model-specific interpretable methods is the low predictive accomplishment compared to other machine learning models and the limitation of using a specific model. Model-agnostic methods are further divided into global model-agnostic methods and Local- Model agnostic methods (Fig. 9).

4.4 Global model-agnostic methods

Global model-agnostic methods represent the overall average behavior of the model. These methods identify the patterns in general and characterize the effect of input features on prediction (Doshi-Velez and Kim 2017). Global interpretability is challenging to implement in highly complex machine learning models. Partial dependence plot (PDP) (Friedman 2001), Accumulated local effects (ALE) (Apley and Zhu 2020), Feature interaction (H-statistic), Functional decomposition, Permutation feature importance (Breiman 2001), and Global surrogate are some global interpretation methods. As it is not feasible to explain each method in detail, we list the advantages and disadvantages of each method (Table 5).

4.5 Local model-agnostic methods

Local interpretation techniques explain individual predictions. These methods estimate the model's behavior in a small region and assume that machine learning prediction for the neighbor instance can be proximate by an interpretable model. Individual conditional expectation curves (Goldstein et al. 2015), Local surrogate models (Ribeiro et al. 2016b), Counterfactual explanation (Wachter et al. 2017), Scoped rules (Ribeiro et al. 2018), Shapely values, and Shapely additive explanations (Lundberg 2017) are some local interpretation methods. The advantages and disadvantages of Local Model-agnostic methods are presented in Table 6.

5 Summary and future work

This paper presents an overview of different anomaly detection techniques in the context of CBM, emphasizing where these techniques have been applied in decision-making. A bibliometric research analysis on CBM is performed with information associated with the most productive authors, countries, and relevant keywords. Scopus database was used for the data collection in this analysis. Prognostics and health management (PHM) is a crucial indicator of industrial automation in industry 4.0, which comprises CBM and PdM. This review discusses anomaly detection techniques and their applications based on supervised, unsupervised (distance-based, statistically-based, classification-based), and semi-supervised machine learning approaches. AI, Machine learning, and data science are the dominant and fundamental tools that allow industrial innovation and technological advancement in maintenance.

Despite the significant success of machine learning methods in industrial maintenance, the black-box nature (low interpretability) and generalization insufficiency are substantial shortcomings of the machine learning methods. To address the issue of not explaining the decisions to a system, XAI has become a research field focusing on machine learning interpretability, providing a more transpicuous AI by maintaining the level of predictive performance. Production delays have high costs because of mechanical problems. XAI

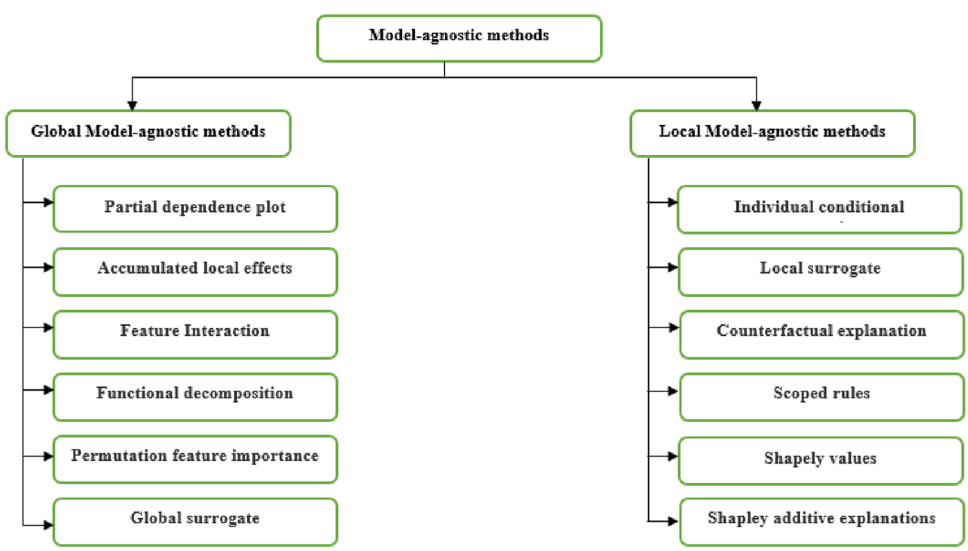


Fig. 9 Types of Model-agnostic methods

Table 5 Advantages and disadvantages of Global Model-agnostic methods

Global Model-agnostic methods	Advantages	Disadvantages
Partial dependence plot (PDP)	Computation is instinctual, and interpretation is understandable if features are not correlated, Uncomplicated to implement	The assumption that features are not correlated, Diversified effects of the feature values may be concealed, Two-dimensional representation of the features
Accumulated local effects (ALE)	Works fine if features are correlated, Rapid computation than PDP, Interpretation is more understandable than PDP	Implementation is more strenuous than PDP plots, Interpretation becomes challenging in the case of strongly correlated features, Less instinctual than PDP
Feature interaction	Discover multitudinous interactions, Can analyze interactions robustness between three or more features	Lengthy computation, Unsteady results, It does not show the appearance of the interactions
Functional decomposition	It gives theoretical reasons for individual feature effects, Provides better apprehension of other methods also	Computation of all feature interactions is highly time-consuming, The drawback of being a manual procedure
Permutation feature importance	Automatically consider all feature interactions, Provides better discernment regarding the model's behavior, No need to retrain the model	Variation in results on permutation repetition, Can show biased results if features are correlated
Global surrogate	Easy implementation, More flexible, Easily understandable	Whatever interpretable model, we choose all the drawbacks to come with it

Table 6 Advantages and disadvantages of Local Model-agnostic methods

Local Model-agnostic methods	Advantages	Disadvantages
Individual conditional expectation (ICE)	Easy to understand than PDP, Can reveal diversified relationships	Unveil one feature only, It may not work correctly if features are correlated Many ICE curves together become congested
Local surrogate	Works also for text and images, Easy to use and explain	Explanations can be manipulated, Uncertainty of explanations, Ignorance of feature correlation
Counterfactual explanation	Clear interpretation, Can work without the use of machine learning, Easy implementation	Various counterfactual explanations for an instance
Scoped rules	Easy to interpret, Highly efficient because of batch sampling, Works efficiently when predictions are nonlinear	Thoroughly variable because its runtime depends on the model's performance, Data discretization needs to be used carefully not to acquire poor results
Shapely values	The essential method to convey a full explanation, Can compare the predictions to the average forecast of a subset	High computation time, Requirement of the data to compute the shapely value for a new data instance
Shapley additive explanations	Impartially distribution of prediction amidst the feature values, Quick implementation and computation	Avoid feature dependence, Requirement of the data to compute the shapely value for a new data instance

in maintenance avoids the issues before they arise, thereby diminishing the impacts of the downtime with a more definite and translucent system to record its performance. XAI is the way to earn customers' assurance and trust. XAI makes the system more interpretable and constructive by tracking its performance, integrity, and inaccuracy.

We have covered the benefits and significance of interpretability in this review. We have highlighted the benefits and drawbacks of interpretable machine learning models. The advantages of model-agnostic procedures over model-specific interpretable methods have been discussed. We also talked about the advantages and disadvantages of local and global model agnostic approaches. Table 7 includes a list of the top 10 highly referenced articles on XAI and interpretable machine learning models, as well as the top 20 highly cited articles on CBM (Table 8). This investigation, like most studies, offers great insights but has some pretentious limits as well. Only the Scopus database was taken into account when gathering the data, but there are other well-known databases as well; hence, an additional study using other databases could be done in the future. We have covered CBM only. In subsequent papers, we intend to cover interpretability in the PdM and various RUL approaches. The discernment obtained from this research analysis has implications for academic scholars.

Funding No funding was received for conducting this study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Human and Animal rights This article does not contain any study involving human and animals and performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix A

See Table 7.

Table 7 Top 20 papers of condition-based maintenance in Scopus

S.No	Authors and year	Title	Source	TC
1	Tran et al. (2012)	Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine	Mechanical Systems and Signal Processing	165
2	Li et al. (2014)	Improving rail network velocity: A machine learning approach to predictive maintenance	Transportation Research Part C: Emerging Technologies	93
3	Corradu et al. (2016)	Machine learning approaches for improving condition-based maintenance of naval propulsion plants	Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment	66
4	Zhang et al. (2018)	Transfer learning with deep recurrent neural networks for remaining useful life estimation	Applied Sciences (Switzerland)	58
5	Kumar et al. (2018)	A big data driven sustainable manufacturing framework for condition-based maintenance prediction	Journal of Computational Science	56
6	Yang et al. (2011)	A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis	Expert Systems with Applications	52
7	Zhao et al. (2017)	Research advances in fault diagnosis and prognostic based on deep learning	Proceedings of 2016 Prognostics and System Health Management Conference, PHM-Chengdu 2016	49
8	Mortada et al. (2014)	Fault diagnosis in power transformers using multi-class logical analysis of data	Journal of Intelligent Manufacturing	43
9	Liu et al. (2019)	Review on Applications of Artificial Intelligence Driven Data Analysis Technology in Condition Based Maintenance of Power Transformers	Gaodianya Jishu/High Voltage Engineering	42
10	Bousdekis et al. (2018)	Review, analysis and synthesis of prognostic-based decision support methods for condition-based maintenance	Journal of Intelligent Manufacturing	39
11	Khoa et al. (2014)	Robust dimensionality reduction and damage detection approaches in structural health monitoring	Structural Health Monitoring	39
12	Robles et al. (2016)	Multiple partial discharge source discrimination with multiclass support vector machines	Expert Systems with Applications	37
13	Sadoughi et al. (2019)	Physics-Based Convolutional Neural Network for Fault Diagnosis of Rolling Element Bearings	IEEE Sensors Journal	32
14	Accorsi et al. (2017)	Data Mining and Machine Learning for Condition-based Maintenance	Procedia Manufacturing	28
15	Mathur et al. (2001)	Reasoning and modeling systems in diagnosis and prognosis	Proceedings of SPIE—The International Society for Optical Engineering	28
16	Sezer et al. (2018)	An Industry 4.0-Enabled Low-Cost Predictive Maintenance Approach for SMEs	2018 IEEE International Conference on Engineering, Technology and Innovation, ICE/ITMC 2018—Proceedings	24
17	Wang et al. (2019)	Degradation evaluation of slewing bearing using HMM and improved GRU	Measurement: Journal of the International Measurement Confederation	21
18	Xu et al. (2015)	Big Data Analytics Framework for System Health Monitoring	Proceedings—2015 IEEE International Congress on Big Data, BigData Congress 2015	20
19	Martin-del-Campo and Sandin (2017)	Online feature learning for condition monitoring of rotating machinery	Engineering Applications of Artificial Intelligence	19
20	Hong and Zhou (2012)	Remaining useful life prognosis of bearing based on Gauss process regression	2012 5th International Conference on Biomedical Engineering and Informatics, BMEI 2012	19

Appendix B

See Table 8.

Table 8 Top 10 papers of XAI and interpretable machine learning

S.No	Authors and year	Title	Source	TC
1	Carvalho et al. (2019a, b)	Machine learning interpretability: A survey on methods and metrics	Electronics (Switzerland)	150
2	Hohman et al. (2019)	Gamut: A design probe to understand how data scientists understand machine learning models	Conference on Human Factors in Computing Systems—Proceedings	53
3	Rosenfeld and Richardson (2019)	Explainability in human-agent systems	Autonomous Agents and Multi-Agent Systems	43
4	Linardatos et al. (2021)	Explainable ai: A review of machine learning interpretability methods	Entropy	20
5	Ji et al. (2019)	Survey on Techniques, Applications, and Security of Machine Learning Interpretability I	Jisuanji Yanjiu yu Fazhan/Computer Research and Development	10
6	Ponce et al. (2017)	Interpretability of artificial hydrocarbon networks for breast cancer classification	Proceedings of the International Joint Conference on Neural Networks	10
7	Kim et al. (2020)	Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information	Decision Support Systems	9
8	Elshawi et al. (2020)	Interpretability in healthcare a comparative study of local machine learning interpretability techniques	Proceedings—IEEE Symposium on Computer-Based Medical Systems	9
9	Blanco-Justicia et al. (2020)	Machine learning explainability via microaggregation and shallow decision trees	Knowledge-Based Systems	8
10	Blanco-Justicia and Domingo-Ferrer (2019)	Machine Learning Explainability Through Comprehensible Decision Trees	Lecture Notes in Computer Science	7

References

Accorsi R, Manzini R, Pascarella P, Patella M, Sassi S (2017) Data mining and machine learning for condition-based maintenance. Procedia Manuf 11:1153–1161

Amihai I, Chioua M, Gitzel R, Kotriwala AM, Pareschi D, Sosale G, Subbiah S (2018) Modeling machine health using gated recurrent units with entity embeddings and k-means clustering. In: 2018 IEEE 16th international conference on industrial informatics (INDIN), IEEE. pp 212–217

Antonelli D, Bruno G, Chiusano S (2013) Anomaly detection in medical treatment to discover unusual patient management. IIE Trans Healthc Syst Eng 3(2):69–77. https://doi.org/10.1080/19488300.2013.787564

Apley DW, Zhu J (2020) Visualizing the effects of predictor variables in black box supervised learning models. J R Stat Soc Series B (Stat Methodol) 82(4):1059–1086

Apté C, Weiss S (1997) Data mining with decision trees and decision rules. Futur Gener Comput Syst 13(2–3):197–210

Ayvaz S, Alpay K (2021) Predictive maintenance system for production lines in manufacturing: a machine learning approach using IoT data in real-time. Expert Syst Appl 173:114598

Barbará D, Couto J, Jajodia S, Wu N (2001) ADAM: a testbed for exploring the use of data mining in intrusion detection. ACM SIGMOD Rec 30(4):15–24

Barbariol T, Feltresi E, Susto GA (2019) Machine learning approaches for anomaly detection in multiphase flow meters. IFAC-PapersOnLine 52(11):212–217

Blanco-Justicia A, Domingo-Ferrer J (2019) Machine learning explainability through comprehensible decision trees. In: international cross-domain conference for machine learning and knowledge extraction. Springer, Cham. pp 15–26

Blanco-Justicia A, Domingo-Ferrer J, Martinez S, Sanchez D (2020) Machine learning explainability via microaggregation and shallow decision trees. Knowl-Based Syst 194:105532

Bousdekis A, Magoutas B, Apostolou D, Mentzas G (2015) A proactive decision making framework for condition-based maintenance. Ind Manag Data Syst 115:1225–1250

Bousdekis A, Magoutas B, Apostolou D, Mentzas G (2018) Review, analysis and synthesis of prognostic-based decision support methods for condition based maintenance. J Intell Manuf 29(6):1303–1316

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Campos J (2009) Development in the application of ICT in condition monitoring and maintenance. Comput Ind 60(1):1–20

Carvalho DV, Pereira EM, Cardoso JS (2019a) Machine learning interpretability: a survey on methods and metrics. Electronics 8(8):832

Carvalho TP, Soares FA, Vita R, Francisco RDP, Basto JP, Alcalá SG (2019b) A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 137:106024

Cassels B, Shark LK, Mein SJ, Nixon A, Barber T, Turner R (2019) Robust principal component analysis of ultrasonic sectorial scans for defect detection in weld inspection. In: Multimodal Sensing: Technologies and Applications. International Society for Optics and Photonics. Vol. 11059, p 110590E

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv (CSUR) 41(3):1–58

Chang K, Yoo Y, Baek JG (2021) Anomaly detection using signal segmentation and one-class classification in diffusion process of semiconductor manufacturing. Sensors 21(11):3880

Choi Y, Yoon S (2021) Autoencoder-driven fault detection and diagnosis in building automation systems: Residual-based and latent space-based approaches. Build Environ 203:108066

Coraddu A, Oneto L, Ghio A, Savio S, Anguita D, Figari M (2016) Machine learning approaches for improving condition-based maintenance of naval propulsion plants. Proc Inst Mech Eng Part M J Eng Marit Environ 230(1):136–153

de Jonge B, Teunter R, Tinga T (2017) The influence of practical factors on the benefits of conditionbased maintenance over time-based maintenance. Reliab Eng Syst Saf 158:21–30

Diao W, Naqvi IH, Pecht M (2020) Early detection of anomalous degradation behavior in lithiumion batteries. J Energy Storage 32:101710

Doshi-Velez F, Kim B (2017) Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608

ElShawi R, Sherif Y, Al_Mallah M, Sakr S (2020) Interpretability in healthcare: a comparative study of local machine learning interpretability techniques. Computational Intelligence

Famurewa SM, Zhang L, Asplund M (2017) Maintenance analytics for railway infrastructure decision support. J Qual Maint Eng 23:310–325

Fan Z, Zi-xuan X, Ming-hu W (2022) Fault diagnosis method for lithium-ion batteries in electric vehicles using generalized dimensionless indicator and local outlier factor. J Energy Storage 52:104963

Fink O (2020) Data-driven intelligent predictive maintenance of industrial assets. Women in industrial and systems engineering. Springer, Cham, pp 589–605

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232

Garcia-Teodoro P, Diaz-Verdejo J, Maciá-Fernández G, Vázquez E (2009) Anomaly-based network intrusion detection: techniques, systems and challenges. Comput Secur 28(1–2):18–28

Gits CW (1992) Design of maintenance concepts. Int J Prod Econ 24(3):217–226

Glowacz A (2019) Acoustic fault analysis of three commutator motors. Mech Syst Signal Process 133:106226

Goldstein A, Kapelner A, Bleich J, Pitkin E (2015) Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. J Comput Graph Stat 24(1):44–65

Grall A, Dieulle L, Bérenguer C, Roussignol M (2002) Continuous-time predictive-maintenance scheduling for a deteriorating system. IEEE Trans Reliab 51(2):141–150

Gunning D, Aha D (2019) DARPA's explainable artificial intelligence (XAI) program. AI Mag 40(2):44–58

Hara Y, Fukuyama Y, Murakami K, Iizaka T, Matsui T (2020) Fault Detection of Hydroelectric Generators using Isolation Forest. In: 2020 59th annual conference of the society of instrument and control engineers of Japan (SICE). IEEE. pp 864–869

Hastie TJ (2017) Generalized additive models. In: Statistical models in S. Routledge. pp 249–307

Hohman F, Head A, Caruana R, DeLine R, Drucker SM (2019) Gamut: a design probe to understand how data scientists understand machine learning models. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–13

Hong S, Zhou Z (2012) Remaining useful life prognosis of bearing based on Gauss process regression. In: 2012 5th international conference on biomedical engineering and informatics. IEEE. pp 1575–1579

Huuhtanen T, Jung A (2018) Predictive maintenance of photovoltaic panels via deep learning. In: 2018 IEEE Data Science Workshop (DSW). IEEE. pp 66–70

Islam MM, Lee G, Hettiwatte SN (2017) A nearest neighbour clustering approach for incipient fault diagnosis of power transformers. Electr Eng 99(3):1109–1119

Jafarian K, Mobin M, Jafari-Marandi R, Rabiei E (2018) Misfire and valve clearance faults detection in the combustion engines based on a multi-sensor vibration signal monitoring. Measurement 128:527–536

Janssens O, Van de Walle R, Loccufier M, Van Hoecke S (2017) Deep learning for infrared thermal image based machine health monitoring. IEEE/ASME Trans Mechatron 23(1):151–159

Jardine AK, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process 20(7):1483–1510

Ji S, Li J, Du Tianyu LB (2019) Survey on techniques, applications and security of machine learning interpretability. J Comput Res Dev 56(10):2071

Jiang G, He H, Xie P, Tang Y (2017) Stacked multilevel-denoising autoencoders: a new representation learning approach for wind turbine gearbox fault diagnosis. IEEE Trans Instrum Meas 66(9):2391–2402

Khan S, Yairi T (2018) A review on the application of deep learning in system health management. Mech Syst Signal Process 107:241–265

Khoa NL, Zhang B, Wang Y, Chen F, Mustapha S (2014) Robust dimensionality reduction and damage detection approaches in structural health monitoring. Struct Health Monit 13(4):406–417

Kim B, Khanna R, Koyejo OO (2016) Examples are not enough, learn to criticize! criticism for interpretability. Advances in neural information processing systems, 29

Kim B, Park J, Suh J (2020) Transparency and accountability in AI decision support: explaining and visualizing convolutional neural networks for text information. Decis Support Syst 134:113302

Kingsford C, Salzberg SL (2008) What are decision trees? Nat Biotechnol 26(9):1011–1013

Kiran BR, Thomas DM, Parakkal R (2018) An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. J Imaging 4(2):36

Kumar A, Shankar R, Thakur LS (2018) A big data driven sustainable manufacturing framework for condition-based maintenance prediction. J Comput Sci 27:428–439

Lahrache A, Cocconcelli M, Rubini R (2017) Anomaly detection in a cutting tool by K-Means clustering and support vector machines. Diagnostyka 18(3):21–29

Lasi H, Fettke P, Kemper HG, Feld T, Hoffmann M (2014) Industry 4.0. Bus Inf Syst Eng 6(4):239–242

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Li C, Zhang WEI, Peng G, Liu S (2017) Bearing fault diagnosis using fully-connected winnertake-all autoencoder. IEEE Access 6:6103–6115

Li H, Parikh D, He Q, Qian B, Li Z, Fang D, Hampapur A (2014) Improving rail network velocity: a machine learning approach to predictive maintenance. Transp Res Part C Emerg Technol 45:17–26

Li X, Jiang H, Niu M, Wang R (2020) An enhanced selective ensemble deep learning method for rolling bearing fault diagnosis with

beetle antennae search algorithm. Mech Syst Signal Process 142:106752
Linardatos P, Papastefanopoulos V, Kotsiantis S (2021) Explainable ai: a review of machine learning interpretability methods. Entropy 23(1):18
Liu G, Bao H, Han B (2018) A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis. Math Problems Eng 2018:1–10
Liu H, Zhou J, Zheng Y, Jiang W, Zhang Y (2018b) Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders. ISA Trans 77:167–178
Liu Y, Xu Z, Li G, Xia Y, Gao S (2019) Review on applications of artificial intelligence driven data analysis technology in condition based maintenance of power transformers. High Volt Eng 45(2):337–348
Liu Y, Zhang J, Hu X, Sun S (2022) Sensor data anomaly detection and correction for improving the life prediction of cutting tools in the slot milling process. Int J Adv Manuf Technol 119(1):463–475
Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Advances in neural information processing systems, 30
Luo H, Zhong S (2017) Gas turbine engine gas path anomaly detection using deep learning with Gaussian distribution. In: 2017 prognostics and system health management conference (PHM-Harbin). IEEE. p 16
Martí L, Sanchez-Pi N, Molina JM, Garcia ACB (2015) Anomaly detection based on sensor data in petroleum industry applications. Sensors 15(2):2774–2797
Martin-del-Campo S, Sandin F (2017) Online feature learning for condition monitoring of rotating machinery. Eng Appl Artif Intell 64:187–196
Mathur A, Cavanaugh KF, Pattipati KR, Willett PK, Galie TR (2001) Reasoning and modeling systems in diagnosis and prognosis. In: Component and Systems Diagnostics, Prognosis, and Health Management. International Society for Optics and Photonics. Vol. 4389, pp 194–203
Meng Z, Zhan X, Li J, Pan Z (2018) An enhancement denoising autoencoder for rolling bearing fault diagnosis. Measurement 130:448–454
Morocho-Cayamcela ME, Lee H, Lim W (2019) Machine learning for 5G/B5G mobile and wireless communications: potential, limitations, and future directions. IEEE Access 7:137184–137206
Mortada MA, Yacout S, Lakis A (2014) Fault diagnosis in power transformers using multiclass logical analysis of data. J Intell Manuf 25(6):1429–1439
Oliveira DF, Vismari LF, de Almeida JR, Cugnasca PS, Camargo JB, Marreto E, Doimo DR, de Almeida LP, Gripp R, Neves MM (2019) Evaluating unsupervised anomaly detection models to detect faults in heavy haul railway operations. In: 2019 18th IEEE international conference on machine learning and applications (ICMLA). IEEE. pp 1016–1022
Ou M, Wei H, Zhang Y, Tan J (2019) A dynamic adam based deep neural network for fault diagnosis of oil-immersed power transformers. Energies 12(6):995
Patil A, Soni G, Prakash A (2022) A BMFO-KNN based intelligent fault detection approach for reciprocating compressor. Int J Syst Assur Eng Manag 13(2):797–809
Paz NM, Leigh W (1994) Maintenance scheduling: issues, results and research needs. Int J Operations Prod Manag 14:47–69
Ponce H, de Lourdes Martinez-Villaseñor M (2017) Interpretability of artificial hydrocarbon networks for breast cancer classification. In: 2017 international joint conference on neural networks (IJCNN). IEEE. pp 3535–3542
Prajapati A, Bechtel J, Ganesan S (2012) Condition based maintenance: a survey. J Qual Maint Eng 18:384–400
Principi E, Rossetti D, Squartini S, Piazza F (2019) Unsupervised electric motor fault detection by using deep autoencoders. IEEE/CAA J Autom Sin 6(2):441–451
Rajakarunakaran S, Venkumar P, Devaraj D, Rao KSP (2008) Artificial neural network approach for fault detection in rotary system. Appl Soft Comput 8(1):740–748
Ran Y, Zhou X, Lin P, Wen Y, Deng R (2019) A survey of predictive maintenance: systems, purposes and approaches. arXiv preprint arXiv:1912.07383
Ribeiro MT, Singh S, Guestrin C (2016a) Model-agnostic interpretability of machine learning. arXiv preprint arXiv:1606.05386
Ribeiro MT, Singh S, Guestrin C (2016b) "Why should i trust you?" Explaining the predictions of any classifier. In": Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. pp 1135–1144
Ribeiro MT, Singh S, Guestrin C (2018) Anchors: high-precision model-agnostic explanations. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 32, No. 1
Robles G, Parrado-Hernández E, Ardila-Rey J, Martínez-Tarifa JM (2016) Multiple partial discharge source discrimination with multiclass support vector machines. Expert Syst Appl 55:417–428
Rosenfeld A, Richardson A (2019) Explainability in human–agent systems. Auton Agent Multi-Agent Syst 33(6):673–705
Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1(5):206–215
Sadgali I, Sael N, Benabbou F (2019) Performance of machine learning techniques in the detection of financial frauds. Procedia Comput Sci 148:45–54
Sadoughi M, Hu C (2019) Physics-based convolutional neural network for fault diagnosis of rolling element bearings. IEEE Sens J 19(11):4181–4192
Salahshoor K, Kordestani M, Khoshro MS (2010) Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. Energy 35(12):5472–5482
Santos P, Villa LF, Reñones A, Bustillo A, Maudes J (2015) An SVM-based solution for fault detection in wind turbines. Sensors 15(3):5627–5648
Sapra SK (2010) Robust vs. classical principalcomponent analysis in the presence of outliers. Appl Econ Lett 17(6):519–523
Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC (2001) Estimating the support of a high-dimensional distribution. Neural Comput 13(7):1443–1471
Sezer E, Romero D, Guedea F, Macchi M, Emmanouilidis C (2018) An industry 4.0enabled low cost predictive maintenance approach for smes. In: 2018 IEEE international conference on engineering, technology and innovation (ICE/ITMC). IEEE. pp 1–8
Shahir HY, Glasser U, Shahir AY, Wehn H (2015) Maritime situation analysis framework: vessel interaction classification and anomaly detection. In: 2015 IEEE international conference on big data (Big Data). IEEE. pp 1279–1289
Shin HJ, Eom DH, Kim SS (2005) One-class support vector machines—an application in machine fault detection and classification. Comput Ind Eng 48(2):395–408
Sneider H, Frank PM (1996) Observer-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation. IEEE Trans Control Syst Technol 4(3):274–282
Sohaib M, Kim JM (2018) Reliable fault diagnosis of rotary machine bearings using a stacked sparse autoencoder-based deep neural network. Shock Vib 2018:1–11
Subbaraj P, Kannapiran B (2010) Artificial neural network approach for fault detection in pneumatic valve in cooler water spray system. Int J Comput App 9(7):43–52

Sun J, Yan C, Wen J (2017) Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning. IEEE Trans Instrum Meas 67(1):185–195

Swanson L (2001) Linking maintenance strategies to performance. Int J Prod Econ 70(3):237–244

Tharrault Y, Mourot G, Ragot J (2008) Fault detection and isolation with robust principal component analysis. In: 2008 16th mediterranean conference on control and automation. IEEE. pp 59–64

Tian J, Morillo C, Azarian MH, Pecht M (2015) Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. IEEE Trans Industr Electron 63(3):1793–1803

Tran VT, Thom Pham H, Yang BS, Tien Nguyen T (2012) Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine. Mech Syst Sig Proc 32:320–330. https://doi.org/10.1016/j.ymssp.2012.02.015

Vachtsevanos GJ, Vachtsevanos GJ (2006) Intelligent fault diagnosis and prognosis for engineering systems, vol 456. Wiley, Hoboken

Vanzile D, Otis I (1992) Measuring and controlling machine performance. Handbook of Industrial Engineering, John Wiley, New York, NY

Wachter S, Mittelstadt B, Russell C (2017) Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harv JL Tech 31:841

Wang J, Sun C, Zhao Z, Chen X (2017) Feature ensemble learning using stacked denoising autoencoders for induction motor fault diagnosis. In: 2017 prognostics and system health management conference (PHM-Harbin). IEEE. pp 1–6

Wang R, Chen H, Guan C, Gong W, Zhang Z (2021) Research on the fault monitoring method of marine diesel engines based on the manifold learning and isolation forest. Appl Ocean Res 112:102681

Wang S, Chen J, Wang H, Zhang D (2019) Degradation evaluation of slewing bearing using HMM and improved GRU. Measurement 146:385–395

Widodo A, Kim EY, Son JD, Yang BS, Tan AC, Gu DS, Choi BK, Mathew J (2009) Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. Expert Syst Appl 36(3):7252–7261

Wuest T, Weimer D, Irgens C, Thoben KD (2016) Machine learning in manufacturing: advantages, challenges, and applications. Prod Manuf Res 4(1):23–45

Xu B, Kumar SA (2015) Big data analytics framework for system health monitoring. In: 2015 IEEE international congress on big data. IEEE. pp 401–408

Yang HH, Huang ML, Yang SW (2015) Integrating auto-associative neural networks with hotelling T2 control charts for wind turbine fault detection. Energies 8(10):12100–12115

Yang J, Sun Z, Chen Y (2016) Fault detection using the clustering-kNN rule for gas sensor arrays. Sensors 16(12):2069

Yang Y, Liao Y, Meng G, Lee J (2011) A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis. Expert Syst Appl 38(9):11311–11320

Yao K, Fan S, Wang Y, Wan J, Yang D, Cao Y (2022) Anomaly detection of steam turbine with hierarchical pre-warning strategy. IET Generation, Transmission & Distribution

Yiakopoulos CT, Gryllias KC, Antoniadis IA (2011) Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. Expert Syst Appl 38(3):2888–2911

Yu J (2019) A selective deep stacked denoising autoencoders ensemble with negative correlation learning for gearbox fault diagnosis. Comput Ind 108:62–72

Yu K, Lin TR, Ma H, Li X, Li X (2021) A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning. Mech Syst Signal Process 146:107043

Zhang A, Wang H, Li S, Cui Y, Liu Z, Yang G, Hu J (2018) Transfer learning with deep recurrent neural networks for remaining useful life estimation. Appl Sci 8(12):2416

Zhang H, Chen H, Guo Y, Wang J, Li G, Shen L (2019a) Sensor fault detection and diagnosis for a water source heat pump air-conditioning system based on PCA and preprocessed by combined clustering. Appl Therm Eng 160:114098

Zhang W, Ma X (2016) Simultaneous fault detection and sensor selection for condition monitoring of wind turbines. Energies 9(4):280

Zhang W, Yang D, Wang H (2019b) Data-driven methods for predictive maintenance of industrial equipment: a survey. IEEE Syst J 13(3):2213–2227

Zhang X, Gu C, Lin J (2006) Support vector machines for anomaly detection. In: 2006 6th world congress on intelligent control and automation. IEEE. Vol. 1, pp 2594–2598

Zhang Y, Bingham C, Martínez-García M, Cox D (2017) Detection of emerging faults on industrial gas turbines using extended Gaussian mixture models. Int J Rotating Mach 2017:1–9

Zhao K, Upadhyaya BR (2006) Model based approach for fault detection and isolation of helical coil steam generator systems using principal component analysis. IEEE Trans Nucl Sci 53(4):2343–2352

Zhao Y, Liu P, Wang Z, Zhang L, Hong J (2017) Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods. Appl Energy 207:354–362

Zhao Y, Nasrullah Z, Li Z (2019) Pyod: a python toolbox for scalable outlier detection. arXiv preprint arXiv:1901.01588

Zhao Z, Wu J, Li T, Sun C, Yan R, Chen X (2021) Challenges and opportunities of AI-enabled monitoring, diagnosis & prognosis: a review. Chin J Mech Eng 34(1):1–29