CIRP Journal of Manufacturing Science and Technology 40 (2023) 53–67

Contents lists available at ScienceDirect

CIRP Journal of Manufacturing Science and Technology

journal homepage: www.elsevier.com/locate/cirpj

Review

Challenges in predictive maintenance – A review

P. Nunes [a,b,*], J. Santos [a,b], E. Rocha [c,d]

[a] Department of Mechanical Engineering, University of Aveiro, Campus de Santiago, Aveiro 3810-193, Portugal
[b] Center for Mechanical Technology and Automation (TEMA), University of Aveiro, Campus de Santiago, Aveiro 3810-193, Portugal
[c] Department of Mathematics, University of Aveiro, Campus de Santiago, Aveiro 3810-193, Portugal
[d] Center of Research and Development in Mathematics and Applications (CIDMA), University of Aveiro, Campus de Santiago, Aveiro 3810-193, Portugal

ARTICLE INFO

Available online 23 November 2022

Keywords:
Predictive maintenance
Prognostics
Predictive models
Review
Industry 4.0

ABSTRACT

Predictive maintenance (PdM) aims the reduction of costs to increase the competitive strength of the enterprises. It uses sensor data together with analytics techniques to optimize the schedule of maintenance interventions. The application of such maintenance strategy requires the cooperation of several agents and involves knowledge and skills in distinct fields, since it encompasses from the averaging of relevant signals in the shop-floor to its processing, transmission, storage, and analysis in order to extract meaningful knowledge. PdM is a broad topic, making it impossible to address all its subtopics in the same paper. Having this into consideration, this paper focuses on the main challenges that hinder the development of a generalized data-driven system for PdM, namely: the existence of noisy or erroneous sensor data in a real industrial environment; the necessity to collect, transmit and process high volumes of data in a timely manner; and the fact that current approaches for PdM are specific for a part or equipment rather than global. This paper connects three different perspectives: anomaly detection, which allows the removal of noisy or erroneous data and the detection of relevant events that can be used to improve the prognostics methods; prognostics methods, which address the models to forecast the condition of industrial equipment; and the architectures, which may allow the deployment of the anomaly detection and prognostics methods in real-time and in different industrial scenarios. Furthermore, the last trends, current challenges and opportunities of each perspective are discussed over the paper.

Contents

* Corresponding author at: Department of Mechanical Engineering, University of Aveiro, Campus de Santiago, Aveiro 3810-193, Portugal.
E-mail address: pnunes@ua.pt (P. Nunes).

https://doi.org/10.1016/j.cirpj.2022.11.004

Introduction

Maintenance plays an important role in industrial sector, since its costs may represent a significant percentage of an enterprise's production costs [1]. Effective maintenance strategies avoid unexpected production stops, reduce the costs, and may even increase the useful lifetime of industrial machines. For these reasons, maintenance approaches have suffered transformations in the result of the concern and efforts of researchers, engineers, technicians and experts. Fig. 1 depicts this evolution over the time. The most primitive strategy is the corrective maintenance (CM), also known as "run-to-failure", which consists in replacing or repairing a part only when it is damaged and the equipment is unable to work without an intervention.

Due to the high costs of having unexpected production stops, CM evolved to proactive approaches. The first one to emerge was the preventive maintenance (PM), which involves periodical inspections made by specialized technicians and the replacement of parts before a critical failure occurs. This replacement is made in equal-spaced periods of time, or after a certain number of working cycles, generally provided by the equipment's manufacturer. This may lead to an earlier or later replacement of components. In the first scenario parts that are in good conditions and could perform a considerably higher number of working cycles are replaced, increasing the costs with maintenance [2], however, the second scenario may have more severe consequences, since a CM action has to be performed [3].

With the technological advances in the field of Industry 4.0 and the development of the Internet of Things (IoT), condition-based maintenance (CBM) approach emerged [4]. The inspections that were made by technicians and experts were automatized by sensors and devices capable of measuring, monitoring, and processing signals that represent physical parameters of industrial equipment, such as acoustic signals, current, voltage, temperatures, forces, vibrations, etc. With this strategy, interventions can be based on sensors' values and actions can be triggered when a value is out of pre-established bounds. More effective and complex approaches arose from the CMB, such as predictive maintenance (PdM), that aligns the paradigms of IoT and Cyber Physical Systems (CPS) with knowledge in the fields of automation, engineering, information technology and data analytics [5] to predict failures and the remaining useful life (RUL) of industrial assets, and schedule maintenance actions accordingly [6]. On the other hand, prescriptive maintenance goes a step further, since it uses the predictions made to make relevant suggestions to address the failure mode and increase the RUL [7,8]. For example, by monitoring the bearing temperature PdM can predict when the equipment is likely to fail, while prescriptive maintenance can suggest a speed reduction to increase its RUL.

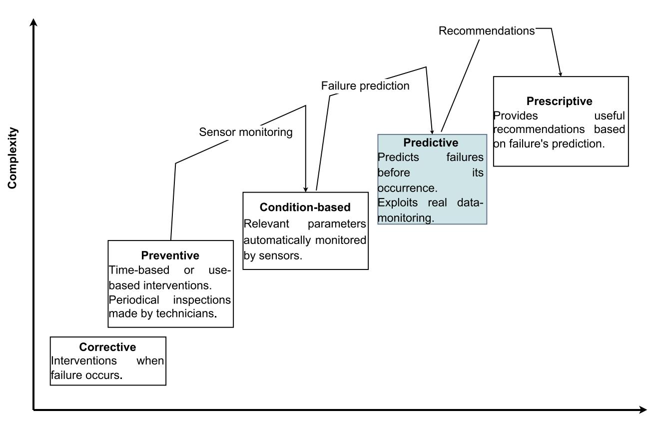

Fig. 1. Evolution of maintenance strategies over time.

Besides the advantages that came with the more recent maintenance approaches, namely the prescriptive maintenance, the implementation of such strategy requires that the CBM and PdM are well consolidated, since it is not possible to implement a prescriptive maintenance approach without a solid PdM, and a PdM cannot be obtained if the necessary apparatus to perform CBM is not well established. For this reason, this paper will focus on PdM, which still has some unsolved issues. One of the challenges is that industrial data may be prone to erroneous measurements, due to harsh environmental conditions or sensor faults. Conversely, the exponential increasing amounts of data that have to be processed, stored and analyzed represent a specific challenge, especially when a real or almost real-time actions are needed [9]. Moreover, the actual industrial environment is composed by very different equipments and flexible management and production systems, which require PdM systems to be effective in very different scenarios. A robust solution should include effective monitoring systems, as well as appropriate analytics techniques to forecast failures and the RUL, and also provide an architecture that can process large amount of data in real-time.

Review approach

PdM requires knowledge from different fields, making it a broad topic. For this reason, it is impossible to address in depth all its subtopics in this paper. In a general way, to develop an effective PdM system, it is necessary to have a careful data treatment to deal with missing data, noise, outliers and other issues that are commonly present in industrial data. Most of the anomaly detection techniques address these issues, while detecting relevant anomalous patterns. On the other hand, the prediction of the RUL of an asset is one of the main features of PdM and is a part of prognostics and health management (PHM) [10]. Often these topics (prognostics and anomaly detection) are addressed individually, however they are both fundamental for PdM, since the anomaly detection outputs can be a valuable input for the predictive models and enhance its performance [11], while contributing to a more generic predictive approach. As depicted in Fig. 2, PdM is often associated with the 5 V's of Big Data; velocity, veracity, value, volume and variety [12]. To accomplish these requirements in an industrial environment the architecture employed plays an important role, especially with respect to volume, velocity and variety, since sensors and devices produce large amounts of data (volume) from different types, such as images, videos, texts, among others (variety), with a high cadence and must be analyzed in real or almost real-time (velocity). The volume plays a key role to develop the predictive models, since typically they require considerable amounts of data to be fitted/trained. The importance of velocity is related to the need to collect data and process data in real-time to forecast problems in a timely manner, so actions can be planned accordingly. Veracity stands for the truthfulness and accuracy of data [13], which is very important in PdM, since sensors and devices work in harsh conditions and are prone to erroneous data. In this scope, anomaly detection can play an important role, by detecting these data and improving the predictive performance of

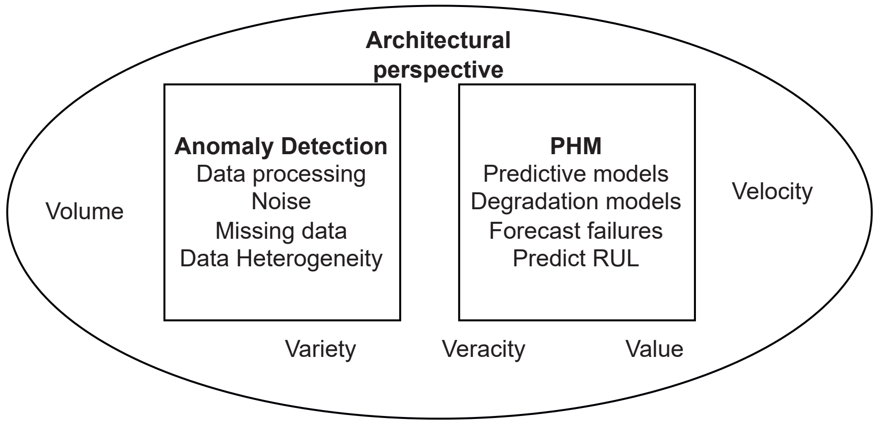

Fig. 2. Schematic of review topics.

prognostics methods. Value represents all the valuable information obtained from data. It strongly depends on the employed methods for the RUL determination and anomaly detection, since they extract knowledge and value from data.

The reasons mentioned motivated the writing of this paper, which explores three different subjects in PdM, namely, anomaly detection, prognostics, and architectural perspectives. Since we could not address all the topics of PdM in this paper, only the most fundamental topics to deploy a generalized and flexible data-driven PdM system are addressed. Prognostics and the RUL forecasting are a central topic in PdM, and anomaly detection is connected to it, since it can improve the prognostics model by detecting and removing noisy or erroneous data, and by providing new inputs in the form of events automatically retrieved from data. The architectural perspectives are connected to these subjects, since not only the models are required to be generalized, i.e., the efficient deployment of these models in different industrial scenarios strongly depends on the architectural aspects. The main objective of this work is to connect the three topics, in order to discuss approaches and methods to develop a generic and efficient PdM solution.

We searched in two scientific databases, Scopus and Web of Science. The reviewed papers were published in journals or conferences in the last six years. Concerning the anomaly detection topic, papers that did not consider sensor data or industrial use cases were excluded. Moreover, the chosen papers were divided based on the employed architecture, which is an important feature when the integration with prognostics models is intended. The prognostics' topic focus on data-driven approaches, however physics-based and knowledge-based approaches are defined, and the most relevant papers are cited even if their publication date is older than six years, because it is pertinent for a better understanding of multi-model approaches that exploit data-driven approaches together with physics or knowledge-based models. The papers reviewed from the architectural perspective topic were chosen based on their novelty, namely which concerns the interaction between cloud, fog and edge layers and their potential to address the 5 V's requirements in the industrial sector. The remaining of this document is organized as follows. Section 3 addresses anomaly detection in sensor data and presents the most recent approaches in this area. Section 4 presents models employed to determine the RUL and/or the degradation of industrial assets, while, Section 5 explores different architectures employed in PdM systems. Finally, Section 6, summarizes the contributions obtained from this paper.

Anomaly detection

Anomaly detection is concerned with identifying data values that are considerably deviated from a typical behavior. The anomaly may be caused by several factors; some of them are related to errors in the acquisition system, such as sensor malfunction, low battery, errors during data transmission, while other anomalies may be caused due to an industrial equipment malfunction, or event, such as changes in the production line or a curative stop [14]. While anomalies caused by machinery's events have relevant information for the analyzer, anomalies caused by sensor errors do not provide any relevant information, and could lead to misinterpretation of data. These anomalies may be designated as noise, however, as discussed by [15], the characterization of anomalies and noise is different for different types of data.

As depicted by Fig. 3, anomalies may be classified in: point anomalies, when one data point is considerably different from its neighbors; behavioral or collective anomalies, when a data pattern is different from an expected behavior; and contextual anomalies, when a data pattern may be expected but in a different context [14,16]. As shown in Fig. 3, there are two types of approaches to handle data anomalies regarding the architecture of the solution; centralized approaches, where the computing process is full-carried in the same equipment (e.g. remote server); and distributed solutions, where the several steps of the computing process are carried in different components, (e.g. edge devices, cloud, etc.). Regarding the methodologies for anomaly detection, they may be divided in two big groups: statistical and machine learning (ML) approaches. Statistical approaches aim to determine anomalies based on the distribution of variables during the working process, while ML is within the field of artificial intelligence (AI), and provides methodologies to handle high dimensional data and extract hidden relationships between data in non-linear and complex environments [17].

Since there are different types of anomalies, and they may be triggered by several factors, most of recent researches employ data from several sensors and exploit the correlation between them [16,18–21–24], as can be noted in Table 1. The exploited correlations may be temporal [20,21], spatial [16], or multi-variate [18,19]. In most of the researches, correlation values are employed together with other techniques, such as the exponential moving average (EMA) [25], artificial neural networks (ANN) [26,27], or fuzzy techniques [28,29]. Besides the exploitation of the correlated sensors, cluster techniques [30], such as the fuzzy clustering [31], the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [32], the Principal Component Analysis (PCA) [33], or the Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) algorithm [34] have been widely applied in recent researches in order to detect anomalous data. The further subsections detail the most recent methods and techniques to detect anomalies that are summarized in Table 1. Note that this paper presents anomaly detection in the specific context of PdM and sensor data. For this reason, the

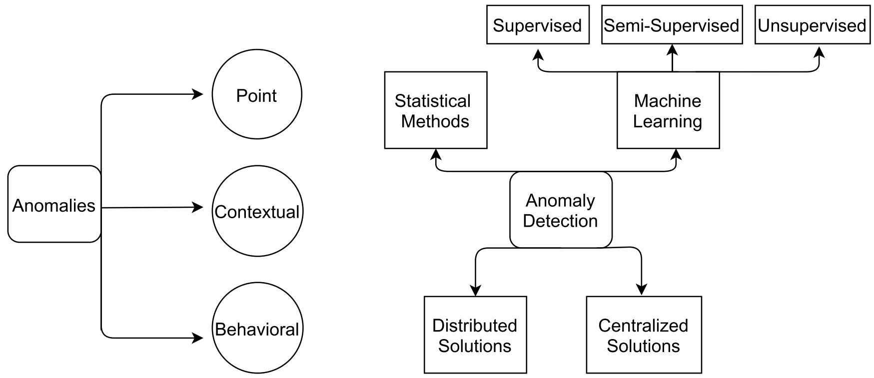

Fig. 3. Overview of anomaly types according to [14,16], and detection techniques and architectures.

methodologies presented are focused on data-driven approaches, which are categorized according to their architecture. To have a more detailed view of broader definitions and concepts, we point the reader to the review made by [14].

Centralized approaches

This subsection presents recent centralized solutions that have been developed to detect anomalies. Note that the researches that do not explicitly propose an architecture to implement the proposed approach in a real environment are encompassed in this subsection.

The authors in [16] introduced the context of shared area, in which sensors are divided by type and area. The continuous data are grouped in blocks and a sliding window is used during a training phase to construct a probabilistic detector, based on the number of times that a sequence of a pair of blocks occurs. The results demonstrated that this approach has better accuracy detection compared to other statistical approaches, such as the K-Nearest-Neighbor probabilistic [39] and the K-means-Gaussian [40].

In [18] the authors employed the Piecewise Aggregate Approximation [41] to aggregate the data in fixed intervals of time. The correlation matrix is calculated for a training set, in order to detect outliers when the correlation of two highly correlated sensors decreases more than a defined threshold. When a sequence has more than a defined limit of outliers, it is considered abnormal, and a minimal number of outliers is removed to respect the speed constraint used by the authors (speed of increasing/decreasing of a sensor's value). The experimental results are compared with two other approaches to detect anomalies, namely Window-median (WM) predict model [42] and the autoregressive (AR) predict model [43]. The proposed approach has a slightly higher time consumption compared to the WM and a considerable higher time consumption compared to the AR, however, globally it performs better than the other approaches, considering metrics as precision and recall in a real dataset.

The authors in [19] proposed a majority voting system. First, three existing algorithms that use multi variable correlation are applied to detect outliers, namely the Elliptic Envelope [44], the Isolation Forest [45] and the Local Outlier Factor [46]. Then, points that are classified as outliers by two or more methods are classified as outliers by the final algorithm. The results show that the approach has very stable results, however, one of the approaches employed in the majority voting system has better results and less computational overheads compared with the voting system that needs the execution of the three techniques.

A supervised method to detect faults was proposed by [20]. The approach assumes that the correlation between the variables of the system changes due to equipment's failure. The method consists in exploiting a large dataset containing normal behavior data, to compute the correlation matrix [47]. The variables whose correlation coefficient is between 0.5 and 0.95 are grouped in 2-D clusters and their data are selected to perform the anomaly detection, by calculating the Mahalanobis distance between the incoming data and the clusters built in the training phase, i.e., an anomaly is detected when the data point is outside the elliptic cluster.

The authors in [35] used a polynomial regression model provided by the manufacturer of wind generators to separate valid sensor data from invalid. A fuzzy clustering technique is applied to the invalid data and a regression model is deployed for each statistically relevant cluster. The final output separates the derating curves and the normal operation curve from incoherent data.

The authors in [21] presented a method that uses sliding windows to segment time series data, and the Mahalanobis distance between sliding windows to detect spike-shaped noisy data. To detect temporary noisy data, a piece of normal behavior data is selected and divided in time windows to build a background model. If the Mahalanobis distance between a time window and the background model is higher than a threshold, it is considered noise. The approach was tested in real industrial IoT (IIoT) data with real noise and anomalies, and it achieved better results in detecting noise compared with outlier detection methods.

Another approach, that intends to differentiate noise or data anomaly caused by errors from anomalous data from events, was proposed by [22]. The authors normalize the data and apply the EMA to smooth the time series. An anomaly is detected when the value of a variable is higher than a threshold. When an anomaly is detected, it is considered an event if the previous and next values are detected as anomaly, otherwise, the anomaly detection is performed for correlated sensors and the classification depends on the results obtained for these sensors. The experimental results on a real traffic dataset reveal good prediction accuracy in detecting noise (96%), however, they only achieve 70% in event detection.

Distributed approaches

In most industrial scenarios, it is of utmost importance to detect anomalies in real-time, or almost real-time. Due to the harsh environmental conditions and communication latency, in some cases, the detection solution has to be distributed between sensor nodes or data acquisition systems, intermediate gateways or base stations and cloud, in order to achieve higher accuracy in a timely manner.

A two-stage anomaly detection was proposed by [23]. The first detection stage is performed in the sensor node, using fuzzy theory. It takes two arguments: the degree of abnormality of a sensor value, and the past values averaged for that sensor. The second stage is performed in the base station, and it considers the data from sensors of the same type in different locations and their correlation coefficient. The goal of the multi-stage detection is to detect local

Table 1
Summary of contributions in anomaly detection.

	Ref.	Techniques	Experimental datasets	Results
Centralized Approaches	[16]	Spatial correlation Statistical methods	Spatial correlation Statistical methods	Better detection metrics compared to other statistical methods
	[18]	Time-series analysis Correlation between sensors	Real wind turbine dataset, errors randomly introduced	Better detection metrics compared to other approaches, slightly higher time cost
	[19]	Majority voting system Correlation between sensors	Sensor data from Intel Berkeley database	More stable performance than using other techniques alone
	[20]	Cluster of high correlated sensors	Experimental data from hydraulic test equipment	Detect faults in several equipments, enable the analysis of physical meaning of anomalies
	[35]	Fuzzy clustering technique Polynomial regression	Real data from wind generator	Remove incoherent data from wind generator dataset
	[21]	Mahalanobis distance Correlation between sensors	IIoT data with real noise and anomalies	Data reconstruction with better results after the cleansing process
Distributed Solutions	[22]	EMA Correlation between sensors	Real traffic dataset	High accuracy in detecting noise Low accuracy detecting events
	[23]	Fuzzy theory Correlation between sensors	Temperature, wind speed, gas parameters Abnormal data generated by manual setting	Better detection metrics compared to other approaches, slightly higher time cost
	[36]	Unsupervised clustering, based on DBSCAN	UCI-ML repository	Better detection metrics compared to other DBSCAN approaches
	[37]	Micro clustering Agglomerative clustering	Mixing various rare events, from multiple sources, with different backgrounds	Good detection metrics No comparative results
	[24]	Unsupervised neural network Correlation between sensors	Sensor network Disturbed manually	More consistent results
	[38]	Variant of PCA	Dataset from IBRK Synthetic anomalies generated	Better detection metrics compared with local detection

abnormal points in each sensor, and compare the result with correlated sensors (in the same base station) to determine if the abnormal detection is induced by external conditions, or due to a sensor malfunction. The results show that the anomaly detection method for multi source data achieves higher accuracy compared to traditional methods using single source data. However, it has higher computational costs, which according to the authors is worthwhile compared to the improvement in detection. In comparison with the DBSCAN technique, this approach achieved better performance regarding detection accuracy and computation time for large amounts of data.

A multi-stage anomaly detection was also proposed by [36]. In the first stage, the data are processed using the Boruta algorithm [48], which extracts the most relevant features from raw data. In the second stage, the data are divided into clusters using the extended k-medoid partitioning algorithm [49] together with the firefly inspired partitioning [50], which is employed to determine the number of clusters. In the third stage, an algorithm based on the DBSCAN is employed to detect anomalies. This algorithm does not force all the data to be located in one cluster, thus sparse points which are not in any cluster are considered anomalous data. The approach was tested on 6 different datasets from online available repositories and the results show better performance, regarding detection metrics such as F1-score, accuracy, false positive rate, and detection rate. Furthermore, the necessity to properly set some initialization parameters in the DBSCAN is overcome, because it is employed a methodology to automatically compute these parameters.

A two-stage unsupervised method to detect abnormal data from acoustic sensors was addressed by [37]. First the authors divide the acoustic data in time windows, and then employ three different techniques to extract features from raw data, namely linear predictive coding (LPC) [51], Mel-frequency cepstral coefficients (MFCCs) [52], and Gammatone frequency cepstral coefficients (GMCC) [53]. In the first stage, the BIRCH algorithm is employed to aggregate the data in several micro clusters according to the computed features. In the second stage the clusters are merged using the distance between clusters' centroids, i.e., if the distance between two centroids is lower than a threshold, they are merged, otherwise they are not. The final output consists of two clusters: a dense one containing normal behavior points, and the other containing rare events. The experiments were performed on background noise records, where the sounds of a gun shot, a glass break, sirens and screams were added as anomalous data. The approach shows good results, namely, the metrics precision, recall and F1-score reached values above 90%. The F1-score and precision metrics outperformed a single stage approach, where only a macro-clustering technique is employed.

The authors in [24] proposed a decentralized architecture in which short-term anomaly detection is performed in the sensor nodes, using an unsupervised ANN algorithm. In this stage, all the data from a sensor node are handled together, in order to ensure data fusion. The second detection stage is triggered by the short-term detection, and is performed in the cloud. It analyzes the correlation changes between high correlated sensors. Thus, the short-term detection identifies potential anomalies and temporal windows that may contain relevant information, while cloud computation enables more complex methods and uses historical data to compute correlation changes between sensors and identify anomalies. Note that due to the limited computational capabilities of the sensor node, the authors make use of the concept of generative replay [54], more specifically, the Restricted Boltzmann Machine (RBM) [55] is employed to allow the ANN to be trained in the cloud and avoid recording data on the edge device, thus, only the parameters of the algorithm are stored in the sensor nodes. A supervised and distributed method was proposed by [38]. The architecture of the proposed solution encompasses clusters of sensor nodes with low computational capacities, and a cluster head, which is a device with more computational capabilities in the neighborhood of the sensor nodes. Initially, the first detection model is developed during the offline training phase, by applying a variant of PCA [56,57] to compute the eigenvector matrix and eigenvalues for the training dataset. A dissimilarity measure for each point (based on the eigenvector matrix and eigenvalues) is calculated in this phase. The anomaly detection is performed by detecting dissimilarity values that are not between the limits computed during the training phase. The sensor nodes average data and constantly build new models with the upcoming data that is sent to the cluster head, which uses the information from the several sensor nodes to periodically build a new global model and send it back to the nodes. The approach was tested using an Intel Berkeley Research Lab (IBRL) sensor network, and synthetic anomalies were generated and introduced in the dataset. The results show that the combination of local and global models to detect anomalies achieves better results compared to local models. The approach was also compared with other spatial and temporal correlation based distributed outlier detection methods, showing that only one of the techniques achieved better results in both, detection rate and false positive rate metrics, however the authors demonstrated that their approach has less computational overheads.

Current challenges, trends and future directions

One of the main challenges in anomaly detection is related to the fact that the abnormal behavior of data can have several causes. While an anomaly caused by a machine malfunction or degradation contains valuable information, anomalies caused by a sensor's malfunction, low battery or other external disturbance, are considered noise and may cause miss-interpretations [58]. For this reason the majority of recent proposed techniques [16,18–24] uses correlation between different sensors, which minimizes erroneous data due to sensor problems, however, only the approaches proposed by [21] and [22] focus explicitly on distinguish noise from abnormal data that describes relevant events. However, these techniques often require the fine-tuning of several parameters, like thresholds and bounds. In addition, these models have to be updated, due to the degradation of equipment's and new external conditions, which requires specialized engineering knowledge, as stated by [21]. Moreover, when using the variation of correlation between sensors, it is assumed that an anomaly causes a considerable difference in the correlation values, however, it may change due to the machine working state or other non-abnormal situation. On the other hand, supervised techniques, as the ones proposed by [20] and [38] need large amounts of labeled data to build the models.

One of the challenges to develop and test anomaly detection techniques is the lack of labeled data. For this reason, some authors employ unsupervised techniques [24,36,37]. Besides the results presented for these techniques being good, the datasets used to obtain the metric performances are built by adding explicit perturbations. For example, in [37], the authors introduced sounds, such as gun shot, glass break, screams and sirens no normal background audio, while in [24], sensors were disturbed with lighted bulbs, and a silicon bag. In a real industrial environment, one may have perturbations that are not so explicit. A real labeled dataset with noise, events and machines' degradation need a lot of specialized engineering work. For this reason, such a rich dataset is still a lack in the field of anomaly detection. Thus, the majority of datasets used in the researches addressed in this subsection introduce synthetic errors to the normal behavior dataset, which may not represent accurately the real behavior of anomalous data.

There is a growing concern with industrial Big Data, and the requirements of modern industry compel systems to be efficient and give timely responses. For these reasons, multi-stage and decentralized systems have gained popularity, since they distribute the processing load by several agents. Furthermore, if incoherent or irrelevant data are detected in the sensor network, it does not need to be stored or transmitted, optimizing the resources' exploitation. It must be pointed that anomaly detection plays an important role, since the results of these techniques may be explored two-fold in further steps of PdM systems: the noisy data with no value can be removed; and the relevant anomalies, such as changes in production, curative stops, oil refills, etc., can be detected automatically from sensor data and employed as additional features for the models that forecast the RUL, which can potentially improve the accuracy of predictive models.

Prognostics methods

In the context of manufacturing, prognosis refers to the forecasting of the expected state of degradation of a machine or its components to estimate its RUL [59]. In the modern industry, prognostics is seen as a service that has a key role in the maintenance field [60], since it allows the schedule of long and short time actions, according to the predictions made by the models to determine the RUL. In this scope, literature refers to different approaches, namely knowledge-based, physics-based and data-driven models [60,61], as depicted in Fig. 4. Data-driven models may exploit statistical or ML approaches and are often combined with knowledge-based and physics-based models, in order to address more complex problems [62].

Anomaly detection was addressed in the previous section of this review, since it can be exploited to enrich the predictive models. For that purpose, anomaly detection has to be performed earlier, and for this reason, the presented approaches were categorized according to the employed architecture, since it is relevant for their integration with more complex prognostics models. Prognostics is in general a more complex subject compared with anomaly detection, and for reason, this section categorizes prognostics techniques according to the type of approach used, while architectural perspectives are discussed with more detail in section 5. Since one of the main goals of this paper is to connect anomaly detection with prognostics and architectural perspectives in the context of PdM, data-driven approaches will be addressed more in depth, compared with other approaches.

Knowledge-based approaches

Knowledge-based approaches are settled on expert knowledge that is usually gathered during years of experience working with a specific equipment or part. This knowledge can be employed to detect faults, determine the state of degradation of equipment or determine the root cause of faults (diagnostics). These approaches are applied in industry since the beginning of 1990s [63–65],

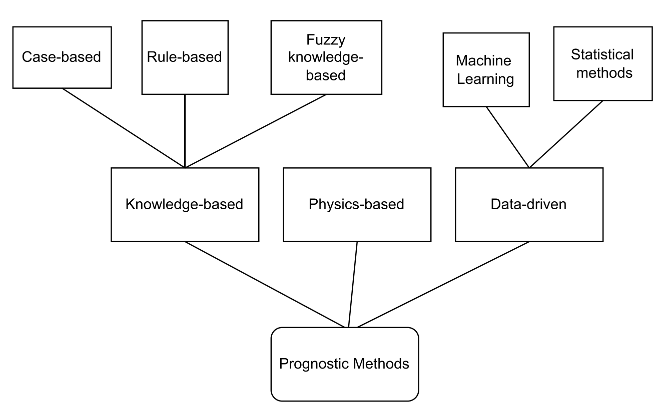

Fig. 4. Overview of different prognostic methods, according to [60].

because they rely mainly on the performance of specialized engineers and technicians, with no need to acquire specialized equipment for monitoring industrial assets. The exploitation of experts inputs evolved with the arising of computational models and techniques to enable the automation of diagnostics and fault detection. In a general way, knowledge-based models may be rule-based, case-based or fuzzy knowledge-based [62]. The most recent research on these topics, concerning maintenance use cases, is summarized in Table 2.

Rule-based models intend to code human reasoning in the form of rules *"IF-THEN"*. These rules are gathered by experts and stored in knowledge bases, that can be exploited when necessary to infer the causes of a fault or stop. Similar to the rule-based models, the fuzzy knowledge-based models also employ rules in the form of *"IF-THEN"*, but differently from the previous approaches which use the boolean logic, i.e, a rule statement may assume only two values, True or False, these rules follow the fuzzy logic [28,29], meaning that the statements may assume continuous values according to its degree of truthfulness. This approach intends to express human perception. For example, the vibration level of an asset can be low/high and its degradation can be low/high according to the different perceptions of each expert. Several researches have exploited these techniques, especially regarding the diagnostics in different industrial use cases [66–69]. A similar approach, which uses a Bayesian approach to estimate the expert's beliefs, was developed by [70]. It is important to note that these models may be used together with data-driven approaches, which will be discussed further in this document.

The Case-based knowledge based uses experiences to address current situations. For example, in the research developed by [71] authors use a case-base reasoning [72], where several cases are stored in a knowledge base. These include the feature's values and the actions taken to overcome that situation (usually a failure). Each time a similar situation occurs, the system suggests the solution adopted to previous situations.

In general knowledge-based methods are interesting for diagnostics, however they are not widely employed to predict a machine's failure or its RUL, without being fused with data-driven methods. Expert knowledge is expensive and may not be available for all components or machines in a factory, but when it exists, these models should exploit and automate expert decision and provide more insights to other approaches such as the data-drive ones.

Physics-based models

Physics-based approaches exploit mathematical models to describe physical processes, which have direct or indirect impact on the health condition of equipment. These models are often applied to mechanical and structural components, since studies concerning physical models to describe fatigue and crack propagation are widely explored in the literature [73–75]. These approaches, like the knowledge-based models are domain specific and require deep knowledge on mathematics together with expertize on the physical behavior of machinery's elements, which is expensive, time-consuming, and the necessary knowledge is often scarce for the majority of the components.

Table 2
Summary of knowledge-based approaches.

Ref.	Technique (s)	Prognostics	Diagnostics	Use case
[66]	Fuzzy logic	no	yes	Car production line
[67]	Fuzzy logic	no	yes	Semi-conductor manufacturing process
[68]	Fuzzy logic	no	yes	Grinding wheels
[69]	Fuzzy logic	no	yes	Bearings
[70]	Bayesian Networks	no	yes	Pipeline leak detection
[71]	Case-base reasoning	no	yes	Rolling mill geraboxes

Data-driven approaches

Data-driven approaches exploit information collected by sensors and actuators in a factory, in order to extract meaningful knowledge from it. The proliferation of technologies, such as IIoT and CPS [76] contributed to enhance the importance of these approaches. The collected data may be used to study the degradation of components, or to create behavioral models from data and estimate its RUL. In the era of Big Data, the capabilities of data-driven approaches have been a topic of exploration in the literature in recent years, mostly due to their capability to be applied together with other approaches in a multi-model approach. For example, collected data may be used with statistical or ML techniques to build a behavioral model of the system, or to deduce the physical model of a particular component. Multi-model approaches allow addressing the complexities of modern industry and the exploitation of existing knowledge about the system, since expert knowledge may be combined with data-driven technique to enhance the approaches' effectiveness.

Data-driven approaches, as the anomaly detection approaches, may be classified in two different types, statistical and ML. In the literature there are several statistical models and techniques applied to PdM, namely hidden Markov models (HMM) [77], Wiener process model (WPM) [78], gamma process model [79], proportional hazards model [80], autoregressive-moving-average (ARMA) models [81], among others. ML techniques comprise ANN [82,83] and its variations, support vector machines (SVM) [84], random forests (RF) [85], xGBoost [86], self organized map (SOM) [87], among others. The remaining of this section addresses data-drive approaches in the prognostics field, as well as the main contributions of each one, which are summarized in Table 3.

Statistical methods

The authors in [88] and [89] employed sensor-based degradation models to predict RUL. This approach is mentioned as sensory-updated degradation-based model (SUDM). First, the degradation is described as a stochastic process by degradation features extracted from sensor data, and then, a vector of stochastic parameters is used to measure the unit-to-unit variability (UtUV), which is a parameter that describes the differences of the machine's health state due to external conditions. In the research developed by [88], the RUL is the necessary time for the degradation to reach a threshold value, and it is updated according to a Bayesian approach [100]. On the other hand, authors in [89] aim to take advantage of the statistical lifetime distribution (SLD), which is the methodology typically employed in preventive maintenance to estimate the time between equally spaced maintenance interventions, and SUDM. The SUDM is applied to determine the RUL and SLD is used to determine when the parameters have to be updated. The results presented demonstrate that by combining SUDM and SLD approaches, fewer parameters have to be defined in comparison to the research developed by [88], because the time increments to update SUDM model are obtained with the inclusion of SLD model. However, the approach proposed by [88], demonstrated that if the most suitable time increments are chosen, the system's availability is higher.

In [91], the authors developed a methodology based in the WPM to estimate the RUL of turbofan engines. They also employ the UtUV parameter to describe the differences of machine's health state due to the difference of health and operational conditions. This parameter is updated during the process via particle filtering (PF) [101] and fuzzy resampling algorithms. The results showed that the RUL converges to the ground truth when both algorithms are used to update the UtUV during the working cycles. One of the drawbacks of this approach is the fact that it is more effective when the UtUV

Table 3
Summary of data-driven approaches in prognostics.

	Ref.	Techniques	Case study	Main contributions
Statistical methods	[88]	SUDM	Simulation model	Updates degradation models with sensor data
	[89]	SUDM	Simulation model	Advantages of SLD-based models and SUDM
	[90]	PF; WMQE	Bearing tests	New health indicator
	[91]	WPM; PF	Turbo fan engines	RUL based on age and state
	[92]	KS; Chi2	Insulation of electrical machines	Automatic data labeling
	[93]	WPM	Lead-acid batteries	Considers hard failures
Machine learning methods	[94]	ANN	Wind turbine gearbox	Normal behavior model created by ANN
	[95]	RBM; SOM	Bearing data	RBM extracts features that improve RUL prediction
	[96]	PF; RF	Aircraft bearings; aircrafts engines	Ensemble method
	[97]	SVM; ANN	MEP components	Extended life of MEP components
	[98]	DBF	Hot rolling machine	Fusion of expert knowledge, real-time information and configuration parameters
	[99]	XGBoost; RF	Consumer goods factory	Applied in real environment

plays a more important role than other resources, such as age and state-dependent WPMs. On the other hand, theoretically, it allows the methodology to be employed in different contexts by changing the degradation model, which brings the opportunity to use physics-based models when they are available.

There are two different types of failure with respect to degradation signals. One is the soft failure, in which an asset condition is given by a degradation signal and its failure occurs when a predefined degradation threshold is reached [102]. The other type is the hard failure, which occur instantaneously instead of having an established degradation threshold [93,102]. The authors in [93] use the WPM to model the degradation of automotive components (lead-acid batteries) subjected to hard failures, which increase the complexity of the problem due to its unpredictability nature. These failures are addressed by assuming that the hazard rate [103] follows a Weibull distribution, whose parameters are determined from historical data. The experimental results show that the RUL predictions improve when the batteries are near to failure, which may not be enough for a proper maintenance, due to the short period of time that the decision maker has available. Moreover, the approach only considers hard failures, when real systems often can suffer from both: degradation induced soft failures, and hard failures. Another drawback of the model is the fact that authors fit the hazard rate to a Weibull distribution in order to comply with statistical theoretical models, instead of using its real distribution.

The authors in [90] developed a health indicator based on the vibration's data collected from bearing tests. The weighted minimum quantization error (WMQE) was developed from [104] and used as a health indicator, while a PF algorithm is used to predict the RUL. The results of this research showed that the developed indicator improves the prediction of the RUL. However, the degradation process was described using a variant of the Paris-Erdogan model [105], which describes the propagation of micro fatigue cracks in mechanical components, making the approach specific to some mechanical components. This research shows how physics-based models may be employed together with data-driven approaches, but unfortunately, this kind of approach can not to be applied in a wide range of contexts, since the failure model is often unknown.

A statistical method was employed by [92] to assess the degradation phenomena in the insulation of electrical machines. First, the authors describe an empirical model of the normal behavior of components [106], and then, Chi-Square (Chi2) and Kolmogorov-Smirnov (KS) tests are employed to determine if the considered data corresponds to a normal behavior. Besides the strategy for the PdM, the authors propose an IoT-based system architecture and algorithms to embed the strategy in low-cost devices, such as Raspberry Pi. The fact that the approach is computationally light and unsupervised leverages a wide range of capabilities, for example, labels extracted from each apparatus may be concatenated in a central unit, thus more knowledge may be extracted from these data. However, in the present research, the exploitation of such capabilities was not explored.

Machine Learning methods

ML approaches were applied by [94], who employed an ANN algorithm to build a normal behavior model of a wind turbine gearbox, which was used to assess the health state of the system. The sensor values predicted by the normal model were compared to the actual values, in order to determine maintenance actions. ML techniques, such as ANN are able to model very complex systems, however, in this case study, the considered variables are well-known and in a short number. In more complex systems with dozens of sensors, other techniques, such as data reduction, feature extraction, and other data treatments may have to be performed in order to extract meaningful knowledge. All these tasks may cause a heavy burden in a centralized cloud system, as the one proposed by the authors, especially when this strategy is applied to several equipments.

In [95], authors applied an RBM to extract relevant features from data of operating bearings, to train a SOM. The online process consists in comparing a vector of testing data with the weight vectors of all the units in the baseline map, using the minimum quantization error. The results show that the features extracted by RBM improve the RUL prediction. The feature extraction and the RUL prediction are addressed as two separated tasks, which allows the application of other prediction techniques, but, on the other hand, parameter tuning and model selection, have to be performed for both tasks, increasing the computational effort.

On the other hand, [96] proposed an ensemble method that combines an ML methodology, the similarity-based interpolation (SBI) [107] with a statistical technique, the PF. Both algorithms are trained offline and their outputs are weighted according to the predicted degradation stage. The output of this linear combination is considered the final prediction for RUL. The results of the tests made on aircraft bearings and engines showed better predictive performance compared to the use each one of the methods alone. However, the application of two different methods involves more computational effort and complexity, especially, when the weights for each method are dynamic, i.e., each degradation level has different weights.

The authors in [97] exploited the building information modeling (BIM) together with an IoT system that collects relevant data, concerning electrical and plumbing (MEP) components. The condition forecasting is performed by using two different algorithms, whose choice strongly depends on the developer's experience, according to the author. Furthermore, the predicted deterioration curves of MEP components depend on a lot of parameters, which forces the model to be trained for each MEP component, meaning that the models are specific for each MEP component (even if the components are similar).

On the other hand, [98] developed a model in which expert knowledge is exploited together with a discrete Bayesian filter (DBF) [108] to predict the degradation of a hot rolling machine. The degradation is discretized, and the model is trained with expert's beliefs [109], concerning the degradation stage at each moment. According to the results, this approach has higher performance compared to the preventive maintenance schema and also performs better than the traditional state-of-the-art ML algorithms. Besides exploiting expert's knowledge is a good practice that contributes to the effectiveness of predictive models, this kind of approach can not be applied to all industrial equipments, because this knowledge is not always available, especially in complex equipment with many parts subjected to different kinds of failure.

Finally, [99] tested several ML algorithms to predict the RUL of components of a production factory. The results showed that XGBoost and RF were the most effective techniques. Besides the results from the models, the authors proposed and deployed an architecture in a real industrial environment. This fact highlights the necessity to distribute tasks, such as data collection, data processing and training models, by several layers to achieve the computational performance required by modern industry. In terms of the employed models, the methodology was applied at one factory, thus, it can not be generalized, as pointed out by the authors.

Current challenges and future directions

As can be concluded, the most recent state of the art on prognostics in PdM combines data-driven approaches with models or expert knowledge to achieve better predictive results. Although the research developed in this area, there are some challenges unsolved. The presented methods are applied to a specific part or equipment, making the solutions specific rather than global. There are papers that present good results in different use cases, for example [110] presented accuracies higher than 90% assessing the condition of cutting tools and rolling bearings, however, it was achieved with different ML models, namely SVM and a type of ANN. Furthermore, the influence of each part's degradation on the others is not accounted in the solutions. Furthermore, risk metrics or models such as Dynamic Fault Trees (DFT) [111] are not applied to model the failure's cause and the interaction between the parts.

One of the components widely studied are the rolling bearings [112,91,90,95], since there are plenty of tests to model the failure of these components, thus enough expert knowledge have been gathered. Moreover, to develop models from data is easier for simple components, compared to an entire complex industrial machine, such as presses, or molding machines. Also, the datasets with synthetic data typically do not represent the heterogeneity present in real data [88,89], thus the validation of these models is not as strong as the ones tested in a real industrial environment.

With respect to statistical methods, they have a very strong theoretical basis, however, they assume that parameters follow known distributions, such as exponential, Weibull, or others, instead of using the real distributions, which may be a rough approximation of the real behavior. Besides that, sometimes statistical approaches are based on fatigue or crack failure models that are known for some components, but the methodology can not be replicated on other industrial assets. On the other hand, ML methods have the capability to represent highly heterogeneous and non-linear models, but they need large amounts of data to train the models, and require high computational capabilities.

Although the advantages of data-driven methodologies, they require large amounts of data, and often, the less frequent modes of failure are not considered, because there is not enough data to predict them. Moreover, the researches do not use anomaly or event detection together with degradation models. As discussed in section 3, anomaly detection is employed to detect failures or events that may cause a prompt alert, however, anomaly detection could be applied to detect events such as the filling of oil level, changing the machine's parameters, as well as the changing of work conditions. Moreover, anomaly detection can be employed to detect events that may be generalized and used as features for predictive models, enabling a general method for a PdM system. Finally, anomaly detection can be employed to remove noisy data that deteriorates the accuracy of predictive models. In fact, the potentiality of fusing anomaly detection and classification methods was demonstrated in the research presented by [11], where the authors exploited a combination of time segmentation and anomaly detection techniques together with ML classification algorithms to predict mechanical failures on stamping presses. The simultaneous application of anomaly detection, segmentation and ML allowed an improvement rate up to 22.971% of the metric F1-score, which brings exciting prospects concerning the application of similar methodologies to forecast the RUL of other industrial assets.

Another issue observed in the presented researches, is that, the degradation models tend to differ from the ground truth RUL when the failure is far from happen, converging to the ground truth RUL when the degradation state is near the failure, which may not give enough time to the stakeholders to plan maintenance activities in a timely manner.

Architectural perspective

The increasing complexity of industrial processes, and the necessity to deal with high volumes of heterogeneous data, motivated the necessity to develop technologies and architectures to tackle these challenges in the context of PdM. The developed approaches have suffered several modifications and evolutions over the years. The first architectures proposed were centralized cloud-based approaches [113]. These, typically explore the computational capabilities of a centralized cloud server to handle the most complex tasks, such as data processing, training predictive, or detect anomalies [114–116], while edge devices are responsible only by data collection (sensor networks) and data transmission to the cloud. Besides being a simple and effective approach in some scenarios, with the increasing number of sensor nodes on the shopfloor, the amount of data to process increased in a large scale, and handling all these tasks in the cloud may not be feasible, due to the required network low latency, fast response of the system, and due to the cost of constantly transmitting large amounts of data.

For the mentioned reasons, decentralized architectures exploiting the computational capabilities of edge devices, fog nodes, or the capabilities of several cloud servers, are being the focus of recent researches in the field of industrial maintenance. Fig. 5 (a) depicts the different types of architectures and respective layers, while Fig. 5 (b) gives an overview of the computing layers that may be exploited. Edge devices have computational capabilities, allowing them to preprocess complex data, creating and selecting relevant data features, or even running algorithms [117]. In this scope, the Bosch white paper [118] highlights the advantages, and presents features to consider when building edge computing solutions, namely the capability to run analytics and ML, availability and reliability, local persistent storage, remote management and update, among others.

The fog computing offers some cloud services near to the place where data are generated. It encompasses storage and networking capabilities, which reduces latency and alleviates the burden on the network [119]. The next subsection highlights some recent approaches, focusing on architectural aspects of PdM. These may exploit one or more layers represented in Fig. 5, as can be seen in Table 4, which summarizes the contributions in this field.

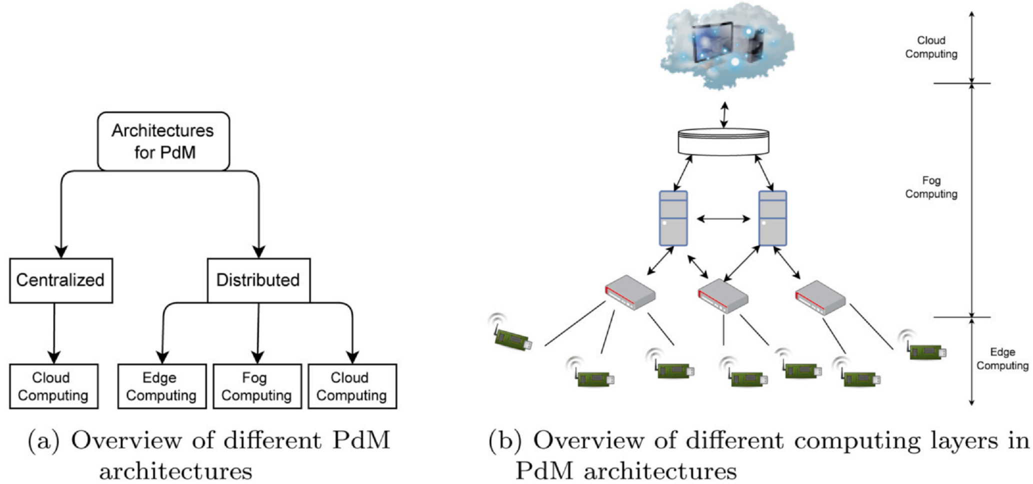

Fig. 5. Architectures for PdM. (a) Overview of different PdM architectures [119]; (b) Overview of different computing layers in PdM architectures. (adapted from [119]).

Main recent contributions

A modular framework designated IDARTS was developed by [120] to support the pluggability of new agents and reduce the network latency. It is based on the monitoring approach developed by [129], and is composed by three main components, the cyber-physical production system, where data are collected, processed and evaluated. The real-time data analysis component, focused on early fault detection of components, and the knowledge management component, which is responsible for updating models with historical and incoming data. The platform was tested using a cluster with four nodes containing three Java Agent Development Framework (JADE) agents [130] and a message queue based on Apache Kafka [131]. The main features achieved by this research are: the support for pluggability, as well the scalability and flexibility, since virtual resources can be added or removed; different configurations were successfully tested; and different models were trained. Besides the mentioned achievements, the architecture requires four different machines with cloud computing capabilities, instead of exploiting the capabilities of edge devices and handling locally with data.

In the research developed by [121], a MapReduce-based Random Matrix Mode algorithm [132] was employed at edge devices for anomaly detection in a decentralized architecture with three layers. The bottom layer comprises the environment where the big data are generated, above are the IoT networks, composed by several edge devices that collect data from the environment, acting as simple network management protocol (SNMP) agents, i.e., they act as SNMP clients and enable the communication between the central network management station and the shop-floor. The fact that the detection technique may run on edge devices is very interesting, however, the presented research is in a very embryonic stage, instead of employing a real case study with real-time data collection, a simulated virtual sensor was employed to validate the approach.

Guided by the features highlighted in the Bosch white paper [118] the authors in [122] developed a modular platform focused on assuring the features mentioned in the white paper. The modularity of the proposed architecture allows the use of cloud computing or edge computing. The edge computing capabilities are improved with an AI accelerator hardware, which allows running complex deep learning modules, and the use of standard libraries such as TensorFlow. The research is very interesting, in the way it achieves almost all features mentioned in [118], especially the capability to run ML models. However, the data treatment and the fine-tuning of parameters such as the window size are not automatic. Furthermore, it was used an existing dataset for validation of the model, thus, no data collection from the edge devices was performed.

With focus on edge computing, the authors in [123] developed a compressed recurrent neural network (RNN) [82], by employing techniques such as quantization and pruning [133]. The main focus of this research is the ability to run ML algorithms on low power edge devices, in order to predict bearing faults of induction motors. Also concerned with edge devices computational performance, in [124], the authors explored diverse sensor's configurations and tested different fault detection algorithms running on different edge devices, such as Raspberry Pi 3, NVIDIA Jetson TX2 and NVIDEA Xavier boards, being the last board the one who achieved better performance concerning the computational time. These researches highlight the fact that even when the power consumption or computational performance are critical factors, it is possible to process data, run ML models, while collecting and transmitting data, using edge devices. However, the capabilities of an architecture for PdM may be extended with the exploitation of both, cloud server and

Table 4
Summary of architectures for PdM.

Ref.	Cloud layer	Fog layer	Edge layer	Main contributions
[120]	Yes (4 nodes)	No	No	Support for pluggability, scalability and flexibility
[121]	Yes	No	Yes	Anomaly detection runs directly on edge devices
[122]	Yes	No	Yes	Flexibility to use edge or cloud to run several ML algorithms
[123]	Yes	No	Yes	Compressed a RNN to run on low power edge devices
[124]	No	No	Yes	Tested different algorithms on different low power edge devices
[125]	Yes	Yes	Yes	High flexibility and scalability, allowed by docker micro-services
[126]	Yes	Yes	Yes	Optimized computing resources and brings analytics closer to the shop-floor
[37]	Yes	Yes	Yes	Tasks divided by the different layers allow more complex event detection methods
[127]	Yes	No	Yes	Plug-and-play gateway to extract and structure field-level data in different scenarios
[128]	Yes	Yes	Yes	Classifiers are trained near the shop-floor

edge devices. For example, several data frameworks and data visualization tools, as well as more complex data analysis may be enabled by cloud computing.

The authors in [125] proposed an architecture where the cloud capabilities are explored to train complex predictive models, which are pushed to the edge devices through a fog network. The developed framework encompasses several microservices, such as predictive analytics, data visualization and maintenance schedule. These microservices are built within docker containers, allowing the scalability of all the resources. A similar architecture based on docker containers was employed by [134] in robotics industry. This kind of service-based architecture, theoretically allows the implementation of all the functionalities mentioned in [118], however, there are important features that were not reported by these researches.

Cloud, fog and edge layers were considered in the architecture proposed by [126]. The edge layer is mainly responsible for sensing and transmitting the data and run low complexity analytics, while fog layer employs technologies such as Apache NiFi and is responsible for gathering the data sent by the edge layer and route it to the cloud. In addition, it can aggregate data and run more complex analytics than the edge layer. The cloud layer enables the storage of all the data and further data analysis. Besides being an interesting architecture, it is in a very embryonic stage, since it lacks of evaluation tests.

Another approach that explores edge and fog capacities was developed by [37], who proposed a framework to detect anomalies in audio stream data, where data is collected by edge devices, feature extraction is performed in the fog layer and an ML method is computed in the cloud. However, the edge layer is under exploited, when one consider all the features highlighted by [118], since it only collects data and sends it to the fog layer.

In some cases, industrial equipment has already built-in field-level sensors, with interfaces that allow users to gather the data produced by machines. However, it requires knowledge about distinct industrial communication interfaces, since manufacturers provide different protocols for their machines. The authors in [127] developed a plug-and-play gateway to extract field-level data [135], and transform it according to a data schema, to store it in databases on the cloud. The gateway was tested in two different scenarios with different communication interfaces. Gathering the data in a well-structured schema is very important for PdM, however, the gateway is under-exploited, since other tasks such as data processing, and event persistent storage could be performed, instead of handling all these tasks on the cloud.

A modular approach for PdM was proposed by [128]. The research encompasses a sensing module to collect raw data, a feature module to extract and select relevant features, a learning module to train classifiers, and a monitoring module. The proposed framework explores the advantages of fog computing paradigm to have these services running on edge devices, near the end users. However, the computational performance and latency were not tested in this research. The fog computing is also exploited by [136], who proposed a decentralized architecture and a genetic algorithm to schedule maintenance interventions [137]. However, the edge devices only collect data, while the fog nodes send it to the cloud, where all the data processing and the development of models are handled.

Challenges and future directions

On top of the different architectures available in the literature and presented above, there is a trend to exploit decentralized architectures in order to distribute the workload by several components. While some authors, such as [120] employ a network of server machines to optimize the data flow, others use a fog layer [37,126,136,128], in order to distribute the computational burden and bring more computational capabilities near to the shop floor. Moreover, the authors in [127] address another issue with this type of architecture, namely gathering data from industrial built-in field-level sensors.

Edge devices play an important role in several researches, [121–125,134], because they allow the execution of several tasks, from data collection and processing, to data analysis using complex ML models. In some mentioned researches, these resources are unexploited, because they are merely used to collect data and send it to storage units, when for example, a data analysis could be performed in order to store only relevant data. Besides some promising researches have been presented, none of them accomplished all the features highlighted by the Bosch white paper [118], which may be a good guide for the development of futures PdM architectures based on edge devices. Moreover, the interaction between edge devices and cloud servers is not fully exploited. While edge devices can have built-in models and analyze the data from a machine, the meaningful data could be gathered in the cloud, which, on the other hand, can extract more information by analyzing together relevant data from several industrial assets. Another important aspect that should be considered in future researches, is the connection between enterprise information systems, such as manufacturing execution system (MES) and IIoT devices, as stated in [138].

The need for an easy and practical distribution of relevant services leads to the adoption of technologies, such as Docker [125,134] and devices with more computational power [124], that allow a more flexible distribution of features as services. A growing importance feature is related to the capability to run models and algorithms as near as possible from the end users. For this reason, not only the edge device's capabilities have been explored, but efforts in developing lighter algorithms have been made [122,123].

Conclusion

This review presents a state of the art on topics that present challenges to the implementation of a generalized data-driven system in the PdM context. These challenges encompass the existence of noisy or erroneous data from harsh industrial scenarios, the lack of generalization of prognostic models, and the necessity to collect and process data in a timely and effective way in very distinct industrial scenarios. Three main perspectives are presented: anomaly detection; prognostics; and architectural perspectives. Anomaly detection techniques have the potential to improve prognostics models two-fold: by removing noisy or erroneous data; and by detecting relevant events that can be used as new input features for the prognostics models, making it possible to have more generalized models. Architectural perspectives are connected to these techniques, since they require an efficient and flexible architecture to allow their deployment in distinct industrial scenarios. The three mentioned perspectives offer the basis for developing and deploying a generalized and effective PdM system. For this reason, they were chosen over other PdM topics to be the focus of this paper.

Along this document, the importance of anomaly detection is described, and the approaches applied in the last years are reviewed from a critical point of view and categorized according to their architecture. Then, a pivotal subject is addressed, prognostics and RUL forecasting, where the different approaches are categorized according to the type of model employed. Then, the open issues and the connection with the previous subject (anomaly detection) are presented. Finally, the most recent advances concerning architectures are reviewed, connected to the previous subjects, and the current challenges are presented from a critical analysis of the presented researches.

In the field of anomaly detection in industrial scenarios, one of the major challenges is the fact that the anomalies may be caused by several events, such as a machine malfunction, curative stops, changes in work conditions, sensors' malfunctions induced by harsh environmental conditions, transmission errors, among others. In this context, the approaches proposed by [21] and [22] are interesting, since they focus explicitly on this challenge by using correlation between sensors. Nevertheless, the outputs of these approaches are not exploited in the prognostic approaches, namely in the models for RUL forecasting. The anomaly detection together with the prognostic techniques has the potential to improve the degradation models, since real data may be less prone to errors when using events or anomalies as inputs instead of raw data. Moreover, erroneous data can be handled to not influence the final output of the predictive models.

The majority of approaches for prognostics are specific for a part or equipment, rather than solutions validated in different industrial scenarios. For example, the researches developed by [112,90,91,95] consider only the failure of rolling bears. Furthermore, the influence of each part's degradation on the others is not accounted, which jointly with the employment of synthetic datasets [88,89] contributes to the loss of accuracy when trying to represent the real behavior of industrial machinery. The inclusion of these interactions in the global predictive model would improve the accuracy of the predictive models, since the proposed models to determine the RUL in the literature are effective only when the degradation stage is near to a failure point.

The growing production of data in industrial environment requires the exploitation and integration of edge and cloud capabilities, as well as enterprises' information systems, such as MES. Moreover, the architectures employed for PdM should be flexible and allow the integration of more components and the prompt modification of the existing ones. In this field, the researches proposed by [125,134] are interesting, since they propose a scalable service distribution based on docker microservices. Also, the exploitation of edge devices in decentralized and distributed architectures is very interesting, since it allows a more efficient treatment of data near the shop-floor, while alleviating the burden in the cloud. The Bosch white paper [118] provides valuable insights, concerning the requirements of an edge device, and the research developed by [122] showed interesting results, since the authors achieved the majority of the features mentioned in the white paper.

Future research should address the current challenges in PdM, namely integrate anomaly detection in the developed prognostics models and account with the interaction of the several components, in order to achieve more generalized models to be applied in complex equipments, such as injection molding machines, presses, among others. Furthermore, future research should exploit computational resources of cloud and edge devices and their interaction in order to achieve the best computational and predictive performance of models, to address real-time and flexibility required by modern industries in the context need of Industry 4.0.

Acknowledgments

Thanks are due to the University of Aveiro, FCT/MCTES for the financial support of TEMA research unit (FCT Ref. UIDB/00481/2020 & UIDP/00481/2020) and CENTRO01-0145-FEDER-022083 - Regional Operational Program of the Center (Centro2020), within the scope of the Portugal 2020 Partnership Agreement, through the European Regional Development Fund. The authors also acknowledge FCT - Fundação para a Ciência e a Tecnologia, I.P. for the PhD grants ref. 2020.06926. BD. The third author was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA), through the Portuguese Foundation for Science and Technology, reference UIDB/04106/2021.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Bevilacqua, M., Braglia, M., 2000, Analytic Hierarchy Process Applied to Maintenance Strategy Selection. Reliability Engineering and System Safety, 70/1: 71–83. https://doi.org/10.1016/S0951-8320(00)00047-8.
[2] Mobley, R.K., 2002, An Introduction to Predictive Maintenance. Elsevier. https://doi.org/10.1016/B978-0-7506-7531-4.X5000-3.
[3] Hao, Q., Xue, Y., Shen, W., Jones, B., Zhu, J., 2010, A Decision Support System for Integrating Corrective Maintenance, Preventive Maintenance, and Condition-Based Maintenance. Construction Research Congress 2010 American Society of Civil Engineers, Reston, VA: 470–479. https://doi.org/10.1061/41109(373)47. (pp.).
[4] Lee, J., Kao, H.A., Yang, S., 2014, Service innovation and smart analytics for Industry 4.0 and big data environment. Vol. 16 Procedia CIRP Elsevier: 3–8. https://doi.org/10.1016/j.procir.2014.02.001. Vol. 16.
[5] Bagheri, B., Yang, S., Kao, H.A., Lee, J., 2015, Cyber-physical systems architecture for self-aware machines in industry 4.0 environment. Vol. 28 IFAC-PapersOnLine: 1622–1627. https://doi.org/10.1016/j.ifacol.2015.06.318. Vol. 28.
[6] Sreedharan, R., Unnikrishnan, V.A., 2017, Moving Towards Industry 4.0: A Systematic Review. International Journal of Pure and Applied Mathematics, 117/20: 929–936.
[7] Fox, H., Pillai, A.C., Friedrich, D., Collu, M., Dawood, T., Johanning, L., 2022, A Review of Predictive and Prescriptive Offshore Wind Farm Operation and Maintenance. Energies, 15/2: 504. https://doi.org/10.3390/en15020504.
[8] Lepenioti, K., Bousdekis, A., Apostolou, D., Mentzas, G., 2020, Prescriptive Analytics: Literature Review and Research Challenges. International Journal of Information Management, 50:57–70. https://doi.org/10.1016/J.IJINFOMGT.2019.04.003.
[9] O'Donovan, P., Leahy, K., Bruton, K., O'Sullivan, D.T., 2015, Big Data in Manufacturing: A Systematic Mapping Study. Journal of Big Data, 2/1: 1–22. https://doi.org/10.1186/s40537-015-0028-x.
[10] Zonta, T., daCosta, C.A., daRosaRighi, R., de Lima, M.J., daTrindade, E.S., Li, G.P., 2020, Predictive Maintenance in the Industry 4.0: A Systematic Literature Review. Computers & Industrial Engineering, 150/August:106889 https://doi.org/10.1016/j.cie.2020.106889.
[11] Coelho, D., Costa, D., Rocha, E.M., Almeida, D., Santos, J.P., 2022, Predictive Maintenance on Sensorized Stamping Presses by Time Series Segmentation, Anomaly Detection, and Classification Algorithms. Procedia Computer Science, 200:1184–1193. https://doi.org/10.1016/j.procs.2022.01.318.
[12] BazzazAbkenar, S., HaghiKashani, M., Mahdipour, E., Jameii, S.M., 2021, Big Data Analytics Meets Social Media: A Systematic Review of Techniques, Open Issues, and Future Directions. Telematics and Informatics, 57/September 2020:101517 https://doi.org/10.1016/j.tele.2020.101517.
[13] Bello-Orgaz, G., Jung, J.J., Camacho, D., 2016, Social Big Data: Recent Achievements and New Challenges. Information Fusion, 28:45–59. https://doi.org/10.1016/j.inffus.2015.08.005.
[14] Erhan, L., Ndubuaku, M., DiMauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., Liotta, A., 2021, Smart Anomaly Detection in Sensor Systems: A Multi-perspective Review. arXiv:2010.14946 Information Fusion, 67/September 2020: 64–79. https://doi.org/10.1016/j.inffus.2020.10.001.
[15] E. Keogh, J. Lin, A. Fu, HOT SAX: Efficiently finding the most unusual time series subsequence, in: Proceedings - IEEE International Conference on Data Mining, ICDM, IEEE, 2005, pp. 8–15.10.1109/ICDM.2005.79.
[16] Zhao, B., Li, X., Li, J., Zou, J., Liu, Y., 2020, An Area-Context-Based Credibility Detection for Big Data in IoT. Mobile Information Systems, 2020:1–12. https://doi.org/10.1155/2020/5068731.
[17] Carvalho, T.P., Soares, F.A.A.M.N., Vita, R., Francisco, R.D.P., Basto, J.P., Alcalá, S.G.S., 2019, A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance. Computers & Industrial Engineering, 137:106024 https://doi.org/10.1016/J.CIE.2019.106024.
[18] Z. Li, X. Ding, H. Wang, An Effective Constraint-Based Anomaly Detection Approach on Multivariate Time Series, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12318 LNCS (2020)61–69.10.1007/978–3-030–60290-1_5.
[19] Roig, M., Catalan, M., Gastón, B., 2019, Ensembled outlier detection using multi-variable correlation in WSN through unsupervised learning techniques. Ramachandran WGMVMCV, Walters R. M, (Eds.) IoTBDS 2019 - Proceedings of the 4th International Conference on Internet of Things, Big Data and Security SciTePress: 38–48. https://doi.org/10.5220/0007657400380048. (pp.).
[20] Yoo, Y.J., 2020, Data-driven Fault Detection Process Using Correlation Based Clustering. Computers in Industry, 122:103279 https://doi.org/10.1016/j.compind.2020.103279.
[21] Liu, Y., Dillon, T., Yu, W., Rahayu, W., Mostafa, F., 2020, Noise Removal in the Presence of Significant Anomalies for Industrial IoT Sensor Data in Manufacturing. IEEE Internet of Things Journal, 7/8: 7084–7096. https://doi.org/10.1109/JIOT.2020.2981476.
[22] ElMenshawy, D., Helmy, W., 2018, A Correlation based Approach to Differentiate between an Event and Noise in Internet of Things. International Journal of Advanced Computer Science and Applications, 9/12: 79–83. https://doi.org/10.14569/IJACSA.2018.091212.
[23] Peng, Y., Tan, A., Wu, J., Bi, Y., 2019, Hierarchical Edge Computing: A Novel Multi-Source Multi-Dimensional Data Anomaly Detection Scheme for Industrial Internet of Things. IEEE Access, 7:111257–111270. https://doi.org/10.1109/ACCESS.2019.2930627.
[24] Cauteruccio, F., Fortino, G., Guerrieri, A., Liotta, A., Mocanu, D.C., Perra, C., Terracina, G., TorresVega, M., 2019, Short-long Term Anomaly Detection in

Wireless Sensor Networks Based on Machine Learning and Multi-parameterized Edit Distance. Information Fusion, 52:13–30. https://doi.org/10.1016/j.inffus.2018.11.010.
[25] Anderson, D.R., Sweeney, D.J., Williams, T.A., 2010, Descriptive Statistics: Numerical Measures. Statistics for Business and Economics. 11th edition Cengage South-Western: 85–146. Ch. 3, pp..
[26] Ripley, B.D., 1996, Pattern Recognition and Neural Networks. Cambridge University Press. https://doi.org/10.1017/CBO9780511812651.
[27] Jain, A., Mao, Jianchang, Mohiuddin, K., 1996, Artificial Neural Networks: A Tutorial. Computer, 29/3: 31–44. https://doi.org/10.1109/2.485891.
[28] Rao, S.S., 2009, Optimization of Fuzzy Systems. Engineering Optimization: Theory and Practice Fourth edition John Wiley & Sons, Inc., Hoboken, NJ, USA: 722–727. https://doi.org/10.1002/9780470549124. Ch. 13.6, pp..
[29] Hanss, M., 2005, Applied Fuzzy Arithmetic. Springer-Verlag, Berlin/Heidelberg. https://doi.org/10.1007/b138914.
[30] Xu, R., WunschII, D., 2005, Survey of Clustering Algorithms. IEEE Transactions on Neural Networks, 16/3: 645–678. https://doi.org/10.1109/TNN.2005.845141.
[31] Bezdek, J.C., Ehrlich, R., Full, W., 1984, FCM: The Fuzzy C-means Clustering Algorithm. Computers & Geosciences, 10/2–3: 191–203. https://doi.org/10.1016/0098-3004(84)90020-7.
[32] Schubert, E., Sander, J., Ester, M., Kriegel, H.P., Xu, X., 2017, DBSCAN Revisited, Revisited. ACM Transactions on Database Systems, 42/3: 1–21. https://doi.org/10.1145/3068335.
[33] Abdi, H., Williams, L.J., 2010, Principal Component Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2/4: 433–459. https://doi.org/10.1002/wics.101.
[34] Zhang, T., Ramakrishnan, R., Livny, M., 1996, BIRCH: An Efficient Data Clustering Method for Very Large Databases. SIGMOD Record (ACM Special Interest Group on Management of Data), 25/2: 103–114. https://doi.org/10.1145/235968.233324.
[35] De Caro, F., Vaccaro, A., Villacci, D., 2018, Adaptive Wind Generation Modeling by Fuzzy Clustering of Experimental Data. Electronics, 7/4: 47. https://doi.org/10.3390/electronics7040047.
[36] Garg, S., Kaur, K., Batra, S., Kaddoum, G., Kumar, N., Boukerche, A., 2020, A Multi-stage Anomaly Detection Scheme for Augmenting the Security in IoT-enabled Applications. Future Generation Computer Systems, 104:105–118. https://doi.org/10.1016/j.future.2019.09.038.
[37] Janjua, Z.H., Vecchio, M., Antonini, M., Antonelli, F., 2019, IRESE: An Intelligent Rare-event Detection System Using Unsupervised Learning on the IoT Edge. Engineering Applications of Artificial Intelligence, 84/May: 41–50. https://doi.org/10.1016/j.engappai.2019.05.011.
[38] Rassam, M.A., Maarof, M.A., Zainal, A., 2018, A Distributed Anomaly Detection Model for Wireless Sensor Networks Based on the One-class Principal Component Classifier. International Journal of Sensor Networks, 27/3: 200–214. https://doi.org/10.1504/IJSNET.2018.093126.
[39] Ehsani-Besheli, F., Zarandi, H.R., 2018, Context-aware Anomaly Detection in Embedded Systems. Vol. 582 Advances in Intelligent Systems and Computing Springer, Cham: 151–165. https://doi.org/10.1007/978-3-319-59415-6_15. Vol. 582.
[40] M.A. Hayes, M.A. Capretz, Contextual anomaly detection framework for big sensor data, in: Proc. of the 3rd Int. Congress on Big Data (IEEE BigData 2014), Vol. 2, 2014, pp. 64–71.10.1186/s40537–014-0011-y.
[41] N.Q.V. Hung, D.T. Anh, An improvement of PAA for dimensionality reduction in large time series databases, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 5351 LNAI, Springer, Berlin, Heidelberg, 2008, pp.698–707. 10.1007/978–3-540–89197-0_64.
[42] Laxhammar, R., Falkman, G., 2012, Online Detection of Anomalous Sub-trajectories: A Sliding Window Approach Based on Conformal Anomaly Detection and Local Outlier Factor. Vol. 382 AICT IFIP Advances in Information and Communication Technology Springer, Berlin, Heidelberg: 192–202. https://doi.org/10.1007/978-3-642-33412-2_20. Vol. 382 AICT.
[43] Kirchgässner, G., Wolters, J., 2007, Univariate Stationary Processes. no. 1 Introduction to Modern Time Series Analysis Springer Berlin Heidelberg, Berlin, Heidelberg: 27–91. https://doi.org/10.1007/978-3-540-73291-4_2. no. 1.
[44] Rousseeuw, P.J., Driessen, K.V., 1999, A Fast Algorithm for the Minimum Covariance Determinant Estimator. Technometrics, 41/3: 212–223. https://doi.org/10.1080/00401706.1999.10485670.
[45] F.T. Liu, K.M. Ting, Z.-H. Zhou, Isolation Forest, in: 2008 Eighth IEEE International Conference on Data Mining, IEEE, 2008, pp. 413–422.10.1109/ICDM.2008.17.
[46] M.M. Breuniq, H.P. Kriegel, R.T. Ng, J. Sander, LOF: Identifying density-based local outliers, in: SIGMOD Record (ACM Special Interest Group on Management of Data), Vol. 29, 2000, pp. 93–104.10.1145/335191.335388.
[47] Franzese, M., Iuliano, A., 2019, Correlation Analysis. Vol. 1-3 Encyclopedia of Bioinformatics and Computational Biology Elsevier: 706–721. https://doi.org/10.1016/B978-0-12-809633-8.20358-0. Vol. 1-3.
[48] Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010, Boruta - A System for Feature Selection. Fundamenta Informaticae, 101/4: 271–285. https://doi.org/10.3233/FI-2010-288.
[49] Lee, I., Yang, J., 2009, Common Clustering Algorithms. Vol. 2 Comprehensive Chemometrics Elsevier: 577–618. https://doi.org/10.1016/B978-044452701-1.00064-8. Vol. 2.
[50] Kar, A.K., 2016, Bio Inspired Computing - A Review of Algorithms and Scope of Applications. Expert Systems with Applications, 59:20–32. https://doi.org/10.1016/j.eswa.2016.04.018.
[51] Shaughnessy, D.O., 1988, Linear Predictive Coding. IEEE Potentials, 7/1: 29–32. https://doi.org/10.1109/45.1890.
[52] P. Bansal, S.A. Imam, R. Bharti, Speaker recognition using MFCC, shifted MFCC with vector quantization and fuzzy, in: 2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI), IEEE, 2015, pp. 41–44.10.1109/ICSCTI.2015.7489535.
[53] Valero, X., Alias, F., 2012, Gammatone Cepstral Coefficients: Biologically Inspired Features for Non-Speech Audio Classification. IEEE Transactions on Multimedia, 14/6: 1684–1689. https://doi.org/10.1109/TMM.2012.2199972.
[54] H. Shin, J.K. Lee, J. Kim, J. Kim, Continual learning with deep generative replay, in: Advances in Neural Information Processing Systems, Vol. 2017-Decem, 2017, 2991–3000.arXiv:1705.08690.
[55] McClelland, J.L., Rumelhart, D.E., 1987, Parallel Distributed Processing, Computational Models of Cognition and Perception. The MIT Press, San Diego, California. https://doi.org/10.7551/mitpress/5237.001.0001.
[56] Z. Xie, T. Quirino, M.L. Shyu, S.C. Chen, L.W. Chang, UNPCC: A novel unsupervised classification scheme for network intrusion detection, in: Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, IEEE, 2006, pp.743–750.10.1109/ICTAI.2006.115.
[57] M.A. Rassam, A. Zainal, M.A. Maarof, One-Class Principal Component Classifier for anomaly detection in wireless sensor network, in: 2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN), IEEE, 2012, pp. 271–276.10.1109/CASoN.2012.6412414.
[58] Li, D., Wang, Y., Wang, J., Wang, C., Duan, Y., 2020, Recent Advances in Sensor Fault Diagnosis: A Review. Sensors and Actuators, A: Physical, 309:111990 https://doi.org/10.1016/j.sna.2020.111990.
[59] Suh, J.H., Kumara, S.R., Mysore, S.P., 1999, Machinery Fault Diagnosis and Prognosis: Application of Advanced Signal Processing Techniques. CIRP Annals - Manufacturing Technology, 48/1: 317–320. https://doi.org/10.1016/S0007-8506(07)63192-8.
[60] Gao, R., Wang, L., Teti, R., Dornfeld, D., Kumara, S., Mori, M., Helu, M., 2015, Cloud-enabled Prognosis for Manufacturing. CIRP Annals, 64/2: 749–772. https://doi.org/10.1016/J.CIRP.2015.05.011.
[61] Peng, Y., Dong, M., Zuo, M.J., 2010, Current Status of Machine Prognostics in Condition-based Maintenance: A Review. The International Journal of Advanced Manufacturing Technology, 50/1–4: 297–313. https://doi.org/10.1007/s00170-009-2482-0.
[62] Montero Jimenez, J.J., Schwartz, S., Vingerhoeds, R., Grabot, B., Salaün, M., 2020, Towards Multi-model Approaches to Predictive Maintenance: A Systematic Literature Survey on Diagnostics and Prognostics. Journal of Manufacturing Systems, 56/July: 539–557. https://doi.org/10.1016/j.jmsy.2020.07.008.
[63] B. Freyermuth, Knowledge based Incipient Fault Diagnosis of Industrial Robots, IFAC Proceedings Volumes 24(6) (1991)369–375.10.1016/s1474–6670(17)51169–6.
[64] Majstorović, V.D., Milačić, V.R., 1990, Expert Systems for Maintenance in the CIM Concept. Computers in Industry, 15/1–2: 83–93. https://doi.org/10.1016/0166-3615(90)90086-5.
[65] Vingerhoeds, R.A., Janssens, P., Netten, B.D., Fernández-Montesinos, M. Aznar, 1995, Enhancing Off-line and On-line Condition Monitoring and Fault Diagnosis. Control Engineering Practice, 3/11: 1515–1528. https://doi.org/10.1016/0967-0661(95)00162-N.
[66] Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M., 2019, Fuzzy Early Warning Systems for Condition Based Maintenance. Computers & Industrial Engineering, 128/December 2018: 736–746. https://doi.org/10.1016/j.cie.2018.12.056.
[67] Cao, Q., Samet, A., Zanni-Merk, C., De Beuvron, F.D.B., Reich, C., 2019, An Ontology-based Approach for Failure Classification in Predictive Maintenance Using Fuzzy c-means and SWRL Rules. Procedia Computer Science, 159:630–639. https://doi.org/10.1016/j.procs.2019.09.218.
[68] M. Baban, C.F. Baban, B. Moisi, A. FuzzyLogic-Based Approach for Predictive Maintenance of Grinding Wheels of Automated Grinding Lines, 2018 23rd International Conference on Methods and Models in Automation and Robotics, MMAR 2018(2018)483–486.10.1109/MMAR.2018.8486144.
[69] Berredjem, T., Benidir, M., 2018, Bearing Faults Diagnosis Using Fuzzy Expert System Relying on an Improved Range Overlaps and Similarity method. Expert Systems with Applications, 108:134–142. https://doi.org/10.1016/j.eswa.2018.04.025.
[70] Tang, X., Xiao, M., Liang, Y., Zhu, H., Li, J., 2019, Online Updating Belief-rule-base Using Bayesian Estimation. Knowledge-Based Systems, 171:93–105. https://doi.org/10.1016/j.knosys.2019.02.007.
[71] Boral, S., Chaturvedi, S.K., Naikan, V.N., 2019, A Case-based Reasoning System For Fault Detection and Isolation: A Case Study on Complex Gearboxes. Journal of Quality in Maintenance Engineering, 25/2: 213–235. https://doi.org/10.1108/JQME-05-2018-0039.
[72] Kolodner, J.L., 1992, An Introduction to Case-Based Reasoning*xxxxxx. Artificial Intelligence Review, 6:3–34.
[73] Nasution, F.P., Sævik, S., Gjøsteen, J.K., 2012, Fatigue Analysis of Copper Conductor for Offshore Wind Turbines by Experimental and FE Method. Vol. 24 Energy Procedia Elsevier Ltd: 271–280. https://doi.org/10.1016/j.egypro.2012.06.109. Vol. 24.
[74] Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A., Wu, B., 2007, Intelligent Fault Diagnosis and Prognosis for Engineering Systems. John Wiley and Sons. https://doi.org/10.1002/9780470117842.
[75] Qiu, J., Seth, B.B., Liang, S.Y., Zhang, C., 2002, Damage Mechanics Approach for Bearing Lifetime Prognostics. Mechanical Systems and Signal Processing, 16/5: 817–829. https://doi.org/10.1006/mssp.2002.1483.

[76] Rajkumar, R., Lee, I., Sha, L., Stankovic, J., 2010, Cyber-physical Systems: The Next Computing Revolution. in: Proceedings - Design Automation Conference ACM Press, New York, New York, USA: 731–736. https://doi.org/10.1145/1837274.1837461. (pp).
[77] Rabiner, L., 1989, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. in: Proceedings of the IEEE, 77/2: 257–286. https://doi.org/10.1109/5.18626.
[78] Imam Mujahidin Iqbal, N. Aziz, Comparison of various Wiener model identification approach in modelling nonlinear process, in: 2011 3rd Conference on Data Mining and Optimization (DMO), IEEE, 2011, pp.134–140.10.1109/DMO.2011.5976517.
[79] Singpurwalla, N., Youngren, M., 1993, Multivariate distributions induced by dynamic environments. Scandinavian Journal of Statistics, 20/3: 251–261.
[80] Leemis, L.M., 2010, Variate Generation in Reliability. Vol. 36 Springer Series in Reliability Engineering Springer, London: 85–103. https://doi.org/10.1007/978-1-84882-213-9_4. Vol. 36.
[81] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 2008, Time Series Analysis, Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/9781118619193.
[82] Namuduri, S., Narayanan, B.N., Davuluru, V.S.P., Burton, L., Bhansali, S., 2020, Review—Deep Learning Methods for Sensor Based Predictive Maintenance and Future Perspectives for Electrochemical Sensors. Journal of The Electrochemical Society, 167/3:037552 https://doi.org/10.1149/1945-7111/ab67a8.
[83] Schmidhuber, J., 2015, Deep Learning in Neural Networks: An Overview. Neural Networks, 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003. arXiv:1404.7828.
[84] Hsu, Chih-Wei, Lin, Chih-Jen, 2002, A Comparison of Methods for Multiclass Support Vector Machines. IEEE Transactions on Neural Networks, 13/2: 415–425. https://doi.org/10.1109/72.991427.
[85] Breiman, L., 2001, Random Forests. Machine Learning, 45/1: 5–32. https://doi.org/10.1023/A:1010933404324.
[86] T. Chen, C. Guestrin, XGBoost, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Vol. 13–17-Augu, ACM, New York, NY, USA, 2016, pp. 785–794.arXiv:1603.02754, 10.1145/2939672.2939785.
[87] Kohonen, T., 1990, The Self-organizing Map. in: Proceedings of the IEEE, 78/9: 1464–1480. https://doi.org/10.1109/5.58325.
[88] Kaiser, K.A., Gebraeel, N.Z., 2009, Sensor-Based Degradation Models. IEEE Transactions on Systems, Man, and Cybernetics, 39/4: 840–849.
[89] You, M.Y., Liu, F., Wang, W., Meng, G., 2010, Statistically Planned and Individually Improved Predictive Maintenance Management for Continuously Monitored Degrading Systems. IEEE Transactions on Reliability, 59/4: 744–753. https://doi.org/10.1109/TR.2010.2085572.
[90] Lei, Y., Li, N., Gontarz, S., Lin, J., Radkowski, S., Dybala, J., 2016, A Model-Based Method for Remaining Useful Life Prediction of Machinery. IEEE Transactions on Reliability, 65/3: 1314–1326. https://doi.org/10.1109/TR.2016.2570568.
[91] Li, N., Lei, Y., Yan, T., Li, N., Han, T., 2019, A Wiener-process-model-based Method for Remaining Useful Life Prediction Considering Unit-to-unit Variability. IEEE Transactions on Industrial Electronics, 66/3: 2092–2101. https://doi.org/10.1109/TIE.2018.2838078.
[92] Gianoglio, C., Ragusa, E., Bruzzone, A., Gastaldo, P., Zunino, R., Guastavino, F., 2020, Unsupervised Monitoring System for Predictive Maintenance of High Voltage Apparatus. Energies, 13/5: 1109. https://doi.org/10.3390/en13051109.
[93] Hu, J., Chen, P., 2020, Predictive Maintenance of Systems Subject to Hard Failure Based on Proportional Hazards Model. Reliability Engineering & System Safety, 196:106707 https://doi.org/10.1016/J.RESS.2019.106707.
[94] Garcia, M.C., Sanz-Bobi, M.A., del Pico, J., 2006, SIMAP: Intelligent System for Predictive Maintenance. Application to the Health Condition Monitoring of a Windturbine Gearbox. Computers in Industry, 57/6: 552–568. https://doi.org/10.1016/j.compind.2006.02.011.
[95] Liao, L., Jin, W., Pavel, R., 2016, Prognosability Regularization for Prognostics and Health Assessment. IEEE Transactions on Industrial Electronics, 63/11: 7076–7083.
[96] Li, Z., Wu, D., Hu, C., Terpenny, J., 2019, An Ensemble Learning-based Prognostic Approach with Degradation-dependent Weights for Remaining Useful Life Prediction. Reliability Engineering & System Safety, 184:110–122. https://doi.org/10.1016/j.ress.2017.12.016.
[97] Cheng, J.C., Chen, W., Chen, K., Wang, Q., 2020, Data-driven Predictive Maintenance Planning Framework for MEP Components Based on BIM and IoT Using Machine Learning Algorithms. Automation in Construction, 112:103087 https://doi.org/10.1016/J.AUTCON.2020.103087.
[98] Ruiz-Sarmiento, J.R., Monroy, J., Moreno, F.A., Galindo, C., Bonelo, J.M., Gonzalez-Jimenez, J., 2020, A Predictive Model for the Maintenance of Industrial Machinery in the Context of Industry 4.0. Engineering Applications of Artificial Intelligence, 87/October 2019:103289 https://doi.org/10.1016/j.engappai.2019.103289.
[99] Ayvaz, S., Alpay, K., 2021, Predictive Maintenance System for Production Lines in Manufacturing: A Machine Learning Approach Using IoT Data in Real-time. Expert Systems with Applications, 173/September 2020:114598 https://doi.org/10.1016/j.eswa.2021.114598.
[100] Gebraeel, N.Z., Lawley, M.A., Li, R., Ryan, J.K., 2005, Residual-life Distributions From Component Degradation Signals: A Bayesian Approach. IIE Transactions (Institute of Industrial Engineers), 37/6: 543–557. https://doi.org/10.1080/07408170590929018.
[101] Simon, D., 2006, The Particle Filter. Optimal State Estimation John Wiley & Sons, Inc., Hoboken, NJ, USA: 461–483. https://doi.org/10.1002/0470045345.ch15. (pp).
[102] Man, J., Zhou, Q., 2018, Prediction of Hard Failures with Stochastic Degradation Signals Using Wiener Process and Proportional Hazards Model. Computers & Industrial Engineering, 125:480–489. https://doi.org/10.1016/j.cie.2018.09.015.
[103] Inglis, J., Meeker, W.Q., Escobar, L.A., 2000, Statistical Methods for Reliability Data. Journal of the American Statistical Association, 95/449: 340. https://doi.org/10.2307/2669573.
[104] Qiu, H., Lee, J., Lin, J., Yu, G., 2003, Robust Performance Degradation Assessment Methods for Enhanced Rolling Element Bearing Prognostics. Advanced Engineering Informatics, 17/3–4: 127–140. https://doi.org/10.1016/j.aei.2004.08.001.
[105] Paris, P., Erdogan, F., 1963, A Critical Analysis of Crack Propagation Laws. Journal of Basic Engineering, 85/4: 528–533. https://doi.org/10.1115/1.3656900.
[106] Guastavino, F., Dardano, A., Torello, E., 2008, Measuring Partial Discharges Under Pulsed Voltage Conditions. IEEE Transactions on Dielectrics and Electrical Insulation, 15/6: 1640–1648. https://doi.org/10.1109/TDEI.2008.4712668.
[107] T. Wang, J. Yu, D. Siegel, J. Lee, A similarity-based prognostics approach for remaining useful life estimation of engineered systems, 2008 International Conference on Prognostics and Health Management, PHM 2008(2008). 10.1109/PHM.2008.4711421.
[108] Sarkka, S., 2013, Bayesian Filtering and Smoothing. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139344203.
[109] Fox, D., Hightower, J., Liao, L., Schulz, D., Bordello, G., 2003, Bayesian Filtering for Location Estimation. IEEE Pervasive Computing, 2/3: 24–33. https://doi.org/10.1109/MPRV.2003.1228524.
[110] Lee, W.J., Wu, H., Yun, H., Kim, H., Jun, M.B., Sutherland, J.W., 2019, Predictive Maintenance of Machine Tool Systems Using Artificial Intelligence Techniques Applied to Machine Condition Data, Procedia CIRP, 80:506–511. https://doi.org/10.1016/j.procir.2018.12.019.
[111] Dugan, J., Bavuso, S., Boyd, M., 1992, Dynamic Fault-tree Models for Fault-tolerant Computer Systems. IEEE Transactions on Reliability, 41/3: 363–377. https://doi.org/10.1109/24.159800.
[112] Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., Liao, H., 2006, Intelligent Prognostics Tools and E-maintenance. Computers in Industry, 57/6: 476–489. https://doi.org/10.1016/J.COMPIND.2006.02.014.
[113] Zhang, Q., Cheng, L., Boutaba, R., 2010, Cloud Computing: State-of-the-art and Research Challenges. Journal of Internet Services and Applications, 1/1: 7–18. https://doi.org/10.1007/s13174-010-0007-6.
[114] Arab, A., Ismail, N., Lee, L.S., 2013, Maintenance Scheduling Incorporating Dynamics of Production System and Real-time Information from Workstations. Journal of Intelligent Manufacturing, 24/4: 695–705. https://doi.org/10.1007/s10845-011-0616-3.
[115] Wang, J., Zhang, L., Duan, L., Gao, R.X., 2017, A New Paradigm of Cloud-based Predictive Maintenance for Intelligent Manufacturing. Journal of Intelligent Manufacturing, 28/5: 1125–1137. https://doi.org/10.1007/s10845-015-1066-0.
[116] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M., 2010, A View of Cloud Computing. Communications of the ACM, 53/4: 50–58. https://doi.org/10.1145/1721654.1721672.
[117] Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L., 2016, Edge Computing: Vision and Challenges. IEEE Internet of Things Journal, 3/5: 637–646. https://doi.org/10.1109/JIOT.2016.2579198.
[118] Bosch Software Innovations GmbH, Edge computing for IoT (2018).
[119] Martin, J.P., Kandasamy, A., Chandrasekaran, K., Joseph, C.T., 2019, Elucidating the Challenges for the Praxis of Fog Computing: An Aspect-based Study. International Journal of Communication Systems, 32/7:e3926 https://doi.org/10.1002/dac.3926.
[120] Peres, R.S., DionisioRocha, A., Leitao, P., Barata, J., 2018, IDARTS - Towards Intelligent Data Analysis and Real-time Supervision for Industry 4.0. Computers in Industry, 101/October 2017: 138–146. https://doi.org/10.1016/j.compind.2018.07.004.
[121] W. Zhang, M. Dong, K. Ota, J. Li, W. Yang, J. Wu, A. Big Data Management Architecture for Standardized IoT Based on Smart Scalable SNMP, 2020 IEEE International Conference on Communications, ICC 2020 2020(2020).
[122] Resende, C., Folgado, D., Oliveira, J., Franco, B., Moreira, W., Oliveira-Jr, A., Cavaleiro, A., Carvalho, R., 2021, TIP4.0: Industrial Internet of Things Platform for Predictive Maintenance. Sensors, 21/14: 4676. https://doi.org/10.3390/s21144676.
[123] Markiewicz, M., Wielgosz, M., Bochenski, M., Tabaczynski, W., Konieczny, T., Kowalczyk, L., 2019, Predictive Maintenance of Induction Motors Using Ultra-low Power Wireless Sensors and Compressed Recurrent Neural Networks. IEEE Access, 7:178891–178902. https://doi.org/10.1109/ACCESS.2019.2953019.
[124] Xu, Y., Nascimento, N.M.M., de Sousa, P.H., Nogueira, F.G., Torrico, B.C., Han, T., Jia, C., Rebouças Filho, P.P., 2021, Multi-sensor Edge Computing Architecture for Identification of Failures Short-circuits in Wind Turbine Generators. Applied Soft Computing, 101:107053 https://doi.org/10.1016/j.asoc.2020.107053.
[125] Cerquitelli, T., Bowden, D., Marguglio, A., Morabito, L., Napione, C., Panicucci, S., Nikolakis, N., Makris, S., Coppo, G., Andolina, S., Macii, A., Macii, E., O'Mahony, N., Becker, P., Jung, S., 2019, A Fog Computing Approach for Predictive Maintenance. Vol. 349 Springer International Publishing. https://doi.org/10.1007/978-3-030-20948-3_13.
[126] J. Díaz-De-Arcaya, R. Miñon, A.I. Torre-Bastida, D. L., K. A., M. R., M. E.M.N., Q. C., S. R., S. P., T. C., W. D., Towards an architecture for big data analytics leveraging edge/fog paradigms, 13th European Conference on Software Architecture, ECSA 2019 2(2019)173–176.

[127] Liu, C., Su, Z., Xu, X., Lu, Y., 2022, Service-oriented Industrial Internet of Things Gateway for Cloud Manufacturing. Robotics and Computer-Integrated Manufacturing, 73/July 2021:102217 https://doi.org/10.1016/j.rcim.2021.102217.

[128] Foukalas, F., 2020, Cognitive IoT Platform for Fog Computing Industrial Applications. Computers and Electrical Engineering, 87:106770 https://doi.org/10.1016/j.compeleceng.2020.106770.

[129] A. Dionisio Rocha, R. Peres, J. Barata, An agent based monitoring architecture for plug and produce based manufacturing systems, in: 2015 IEEE 13th International Conference on Industrial Informatics (INDIN), IEEE, 2015, pp. 1318–1323.10.1109/INDIN.2015.7281926.

[130] Bellifemine, F., Caire, G., Greenwood, D., 2007, Developing Multi-Agent Systems with JADE, Wiley Series in Agent Technology. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9780470058411.

[131] Wang, G., Koshy, J., Subramanian, S., Paramasivam, K., Zadeh, M., Narkhede, N., Rao, J., Kreps, J., Stein, J., 2015, Building a Replicated Logging System with Apache Kafka. in: Proceedings of the VLDB Endowment, 8/12: 1654–1655. https://doi.org/10.14778/2824032.2824063.

[132] J. Dean, S. Ghemawat, MapReduce: Simplified data processing on large clusters, in: OSDI 2004 - 6th Symposium on Operating Systems Design and Implementation, Vol. 5, 2004, pp.137–149.10.21276/ijre.2018.5.5.4.

[133] H. Sharma, J. Park, N. Suda, L. Lai, B. Chau, V. Chandra, H. Esmaeilzadeh, Bit Fusion: Bit-Level Dynamically Composable Architecture for Accelerating Deep Neural Network, in: 2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA), IEEE, 2018, pp. 764–775.arXiv:1712.01507, 10.1109/ISCA.2018.00069.

[134] Panicucci, S., Nikolakis, N., Cerquitelli, T., Ventura, F., Proto, S., Macii, E., Makris, S., Bowden, D., Becker, P., O'mahony, N., Morabito, L., Napione, C., Marguglio, A., Coppo, G., Andolina, S., 2020, A Cloud-to-edge Approach to Support Predictive Analytics in Robotics Industry. Electronics (Switzerland), 9/3: 492. https://doi.org/10.3390/electronics9030492.

[135] Liu, C., Vengayil, H., Zhong, R.Y., Xu, X., 2018, A Systematic Development Method for Cyber-physical Machine Tools. Journal of Manufacturing Systems, 48:13–24. https://doi.org/10.1016/j.jmsy.2018.02.001.

[136] Teoh, Y.K., Gill, S.S., Parlikad, A.K., 2021, IoT and Fog Computing based Predictive Maintenance Model for Effective Asset Management in Industry 4.0 using Machine Learning. IEEE Internet of Things Journal, 0/c: 0–7. https://doi.org/10.1109/JIOT.2021.3050441.

[137] Alcaraz, J., Maroto, C., 2001, A Robust Genetic Algorithm for Resource Allocation in Project Scheduling. Annals of Operations Research 2001 102:1, 102/1: 83–109. https://doi.org/10.1023/A:1010949931021.

[138] Mantravadi, S., Møller, C., LI, C., Schnyder, R., 2022, Design Choices for Next-generation IIoT-connected MES/MOM: An Empirical Study on Smart Factories. Robotics and Computer-Integrated Manufacturing, 73/July 2021:102225 https://doi.org/10.1016/j.rcim.2021.102225.