Computers in Industry 123 (2020) 103298

Contents lists available at ScienceDirect

Computers in Industry

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

Review

Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges

Jovani Dalzochio[a], Rafael Kunst[a,*], Edison Pignaton[b], Alecio Binotto[c], Srijnan Sanyal[c], Jose Favilla[d], Jorge Barbosa[a]

[a] University of Vale do Rio dos Sinos (UNISINOS), Av. Unisinos, 950 São Leopoldo, RS, Brazil
[b] Federal University of Rio Grande do Sul (UFRGS), Av. Bento Goncalves, 9500 Porto Alegre, RS, Brazil
[c] IBM Watson IoT Center, Mies-Van-Der-Rohe-Strasse 6 Muenchen, 80807, DE
[d] IBM, 1177 S Belt Line Rd Coppell, TX 75019-4642, USA

ARTICLE INFO

Article history:
Received 6 March 2020
Received in revised form 26 June 2020
Accepted 28 July 2020

Keywords:
Industry 4.0
Internet of Things
Artificial intelligence
Systematic literature review
Predictive maintenance
Ontology

ABSTRACT

In recent years, the fourth industrial revolution has attracted attention worldwide. Several concepts were born in conjunction with this new revolution, such as predictive maintenance. This study aims to investigate academic advances in failure prediction. The prediction of failures takes into account concepts as a predictive maintenance decision support system and a design support system. We focus on frameworks that use machine learning and reasoning for predictive maintenance in Industry 4.0. More specifically, we consider the challenges in the application of machine learning techniques and ontologies in the context of predictive maintenance. We conduct a systematic review of the literature (SLR) to analyze academic articles that were published online from 2015 until the beginning of June 2020. The screening process resulted in a final population of 38 studies of a total of 562 analyzed. We removed papers not directly related to predictive maintenance, machine learning, as well as researches classified as surveys or reviews. We discuss the proposals and results of these papers, considering three research questions. This article contributes to the field of predictive maintenance to highlight the challenges faced in the area, both for implementation and use-case. We conclude by pointing out that predictive maintenance is a hot topic in the context of Industry 4.0 but with several challenges to be better investigated in the area of machine learning and the application of reasoning.

Contents

1. Introduction 2
2. Business challenges involving PdM in Industry 4.0 2
3. Research methodology 3
   3.1. Research questions 4
   3.2. Search process 4
   3.3. Papers selection process 5
   3.4. Quality assessment 5
4. Search results 5
   4.1. Exclusion of papers from the initial corpora 5
   4.2. Performing the quality assessment to select relevant papers 7
5. Answer to the research questions and discussion 7
   5.1. What are the challenges and open questions regarding machine learning and reasoning for predictive maintenance in Industry 4.0? 7

* Corresponding author.
E-mail addresses: jovanidalzochio@edu.unisinos.br (J. Dalzochio), rafaelkunst@unisinos.br (R. Kunst), edison.pignaton@ufrgs.br (E. Pignaton), alecio.binotto@ibm.com (A. Binotto), srijnan.sanyal1@ibm.com (S. Sanyal), jfavilla@us.ibm.com (J. Favilla), barbosa@unisinos.br (J. Barbosa).

https://doi.org/10.1016/j.compind.2020.103298

5.2. Which machine learning techniques are usually in the context of predictive maintenance? 10
5.3. What are the contexts in the use of ontology in predictive maintenance? 12
6. Conclusion 13
References 13

1. Introduction

Industry 4.0 introduces several changes to the original approach of industrial automation. Internet of Things (IoT) and Cyberphysical System (CPS) technologies play roles in this context introducing cognitive automation and consequently implementing the concept of intelligent production, leading to smart products and services [1]. This novel approach leads companies to face challenges of a much more dynamic environment. Many of these companies are not ready to deal with this new scenario where the existence of a large amount does not always collaborate to increase productivity [2].

This large amount of data derives from one of Industry 4.0 principles, which transforms traditional manufacturing into intelligent sensor-equipped factories where technology is ubiquitous. An application of this concept regards the usage of data analytics to design a Decision Support Systems that can be helpful to provide a more efficient decision making, which can allow a faster failure recovery [3,4]. In Industry 4.0, originated from the Decision Support System, we have the Design Support System that provides design assistance through the use of machine learning algorithms, proposing, for example, new versions of products using characteristics of a product [5]. Another example of a large amount of data use is PdM, where CPSs can provide self-awareness and self-maintenance intelligence, this approach allows the industry to predict product performance degradation, and autonomously manage and optimize product service needs [6,7].

Applying predictive maintenance in production environments brings several benefits and also involves overcoming several challenges. On one side, benefits of PdM include productivity improvement, reduction on system faults [8], minimization of unplanned downtimes [9], increased efficiency in the use of financial and human resources [10], and the optimization in planning the maintenance interventions [11]. The use of Machine Learning (ML) is capable of fulfilling the task of prognostics and prediction of failures, for example, estimating the lifetime of a machine using a large amount of data to train an ML algorithm [12,9], in addition to being used to diagnose failures [13,14].

On the other hand, overcoming challenges include the need to integrate data from various sources and systems within a facility what is important to gather accurate information to create prediction models [15–18,3,19–21]. Also, a large amount of data involved in predictive maintenance and the need for real-time monitoring demand dealing with latency, scalability, and network bandwidth-related issues [22,14,23]. Another important aspect regards the use of artificial intelligence, rising other challenges like (I) obtaining training data [24,13]; (II) dealing with dynamic operating environments [24]; (III) selecting the ML algorithm that better fits to a given scenario [13]; and (IV) the necessity of context-aware information [25], such as operational conditions, and production environment [26].

To the best of our knowledge, no related work presents and discusses these challenges covering the use of machine learning, reasoning, ontology in the context of Industry 4.0, and proposing a taxonomy in the same way we present in this paper. Therefore, in this article, we apply a systematic literature review (SLR) methodology [27] to identify relevant frameworks, architectures, and tools in the area of predictive maintenance. Also, we discuss the challenges and investigate the main contributions in the field of research over the last five years. We chose five years to consider the most recent publications in our SLR. Moreover, preliminary searches in the databases returned only a few papers published before the considered period. Our study obtained an initial corpus of 562 publications that were filtered and classified into four groups, considering the approach of each research: (I) integration issues, (II) big data analysis, (III) machine learning approaches, and (IV) reasoning and ontologies. We discuss the top-rated papers in detail and use this corpus to answer four research questions that help one to understand the state-of-the-art and the main challenges of predictive maintenance.

The organization of the remainder of this article is as follows. Section 2 discusses business aspects and the related challenges for the implementation of PdM in Industry 4.0 scenarios. After, in Section 3, we present and discuss the systematic literature review methodology, including the definition of research questions, the search process, the selection and filtering of papers from the initial corpus, and the quality assessment of the papers. We present the results and discussions regarding the analysis of the initial corpus in Section 4. We answer the research questions in Section 5 and present conclusions in Section 6.

2. Business challenges involving PdM in Industry 4.0

The previous industrial revolution focused mainly on improving the physical manufacturing processes, expanding human power with additional power sources (machinery, steam power), establishing a process for mass production through the introduction of assembly lines, and introducing electronics and automation. The 4th industrial revolution, also known as Industry 4.0, focuses primarily on creating a digital representation of the physical processes to get better insights on what is going on with the physical processes. For example, production equipment may have some early signs that something is going wrong and that a breakdown may happen soon. These signs may be detected by predictive models that indicate the deviation from normal operating conditions. So, the digital model can provide early insights about the status of the equipment, allowing the maintenance personnel to determine the best time to repair it, moving from a reactive to planned repair.

Industry 4.0 has a very ambitious scope, aiming to create digital factories, i.e., a digital representation of the physical operations, sometimes called cyber-physical models or digital twins. It aims at integrating processes from the top floor to the shop floor and from suppliers to the end clients, creating vertical and horizontal integration across the value chain. Another goal is to reduce the product design life cycle by creating a digital thread that integrates key processes to design, build, operate, and maintain the equipment. It is also relevant to establish a feedback loop from operation to product engineering to create a piece of fully connected equipment. The connected equipment, regardless of its location in the factory, provides the basis for predictive maintenance. The main idea is to collect a variety of on-line and off-line signals from the equipment to feed models that can detect an early indication of an anomaly or fault.

Fig. 1 presents an overview of a cyber-physical systems architecture designed to provide an early failure detection system. The architecture is composed of two layers. The first one represents the

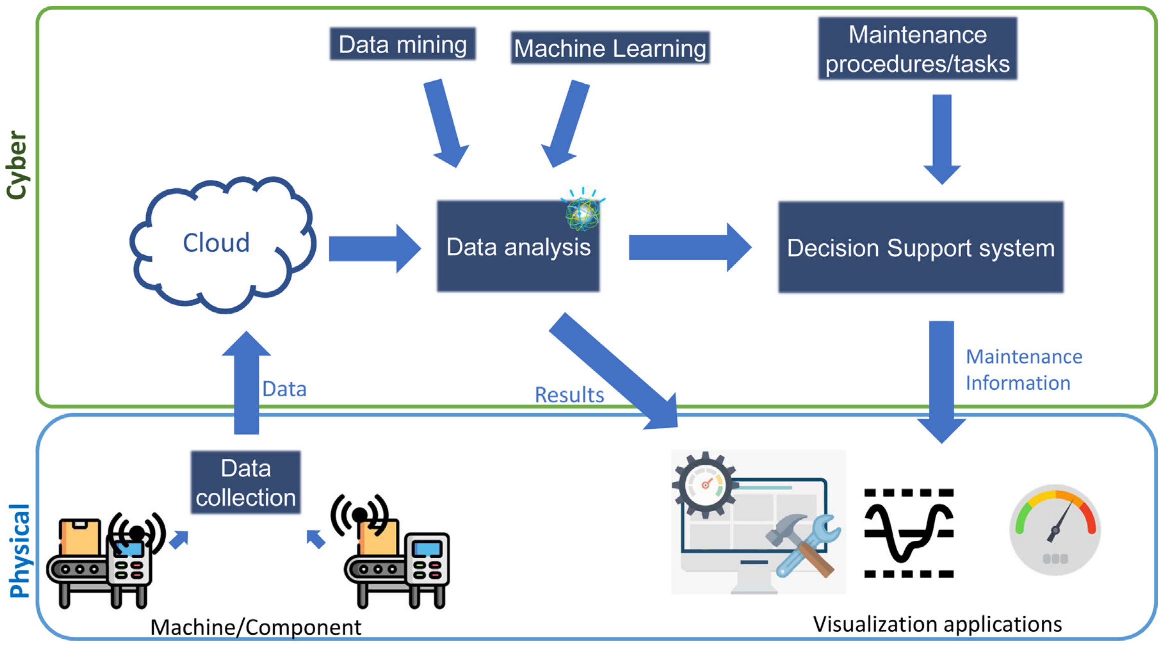

Fig. 1. CPS architecture for maintenance and PdM (adapted from [3]).

physical systems, where sensors monitor the behavior of machines and components. The data collected by these sensors can eventually pass through a pre-processing step before being stored in the Cyber layer of the architecture. This layer is responsible for storing data to feed data mining techniques and ML models training. The Physical layer receives reports regarding the current condition of the machine or component and is capable of providing prognosis-related information, such as the remaining useful life estimates. Besides, in the Cyber layer, there may be a decision support system, which depending on the results generated by the data analysis, can schedule future maintenance and suggest maintenance routes.

Before applying PdM, it is necessary to consider how critical the equipment is in the context of its operation. For example, if the collapse of the piece of equipment does not impact the plant operation, a strategy can be to operate it to failure and fix it when the failure occurs. On the other hand, a piece of equipment can be critical. In this case, a failure could lead to a significant impact on plant operation. If that is the case, companies should consider using predictive models that can provide an early indication of any potential anomaly. Some companies, however, use redundant equipment that operates in case of the primary hardware fails, allowing the company to keep the operations running while it repairs the problematic unit. In this case, the critically is reduced by eliminating the single point of failure.

Predictive models are usually data-driven models that require a variety of data streams provided by multiple real-time and offline sources. Data also comes from Computerized Maintenance Management Systems (CMMS). Failure-related data is also necessary to build and test the predictive models. As equipment becomes more reliable, one of the challenges data scientists face when creating these models is the lack of failure-related data. One approach to overcome this challenge is to consider a cohort of similar equipment and generate a meta-model that reflects the collective learning from the gathering and then applies the meta-model to the specific piece of equipment.

Predicting equipment failure is just one step in the traditional process to perform maintenance. Fig. 2 describes several steps performed to fix a piece of equipment, whenever an anomaly or failure is detected. First, the CMMS system receives a repair request. Thus, a technician assesses the anomaly to determine the critically and the potential root causes. The next step consists on the elaboration of a maintenance plan, considering pre-conditions, such as how long one should wait to dissipate flammable gases in the environment, or which are the step by step processes to diagnose it. The plan also cares for post-conditions, including tests required to ensure the resolution of the problem. Following the establishment of the maintenance plan, the process continues by allocating an expert technician, the required spare parts, and the tools needed to fix the problem. This process usually involved several people and can be quite labor-intensive.

One of the concepts of Industry 4.0 is that the cyber-physical systems (or digital twins) can take actions and communicate with each other, executing a given end-to-end process autonomously. Fig. 3 illustrates an example of this process. Cognitive maintenance is a cyber-physical system that combines a predictive model with a cognitive system that can determine the severity of the anomaly and the potential root causes. This cyber model automatically opens a repair order in the CMMS that feeds a scheduling model, which aims to minimize the impact of the repair in the plant's production. Moreover, this model cares for the availability of qualified maintenance technicians, tools, and spare parts. At the moment of the repair, the cyber-physical system guides the maintenance technician to execute the repair and automatically collects the information required to update the CMMS records. After the completion of the repair, this information feeds another cyber-physical system that controls the status of the parts inventory, optimizing the spare parts inventory for a given service level.

3. Research methodology

An SLR is an approach used to identify, evaluate, and interpret the papers published in a given field of research. This approach enables the identification of existing gaps and points out new research opportunities [28]. In this article, we follow the SLR approach proposed by Kitchenham et al. [27]. We applied the methodological steps as follows.

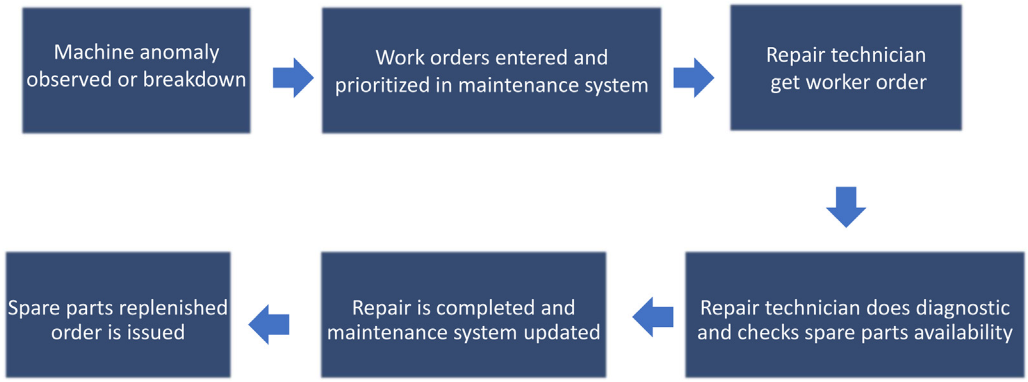

Fig. 2. Traditional process to perform maintenance.

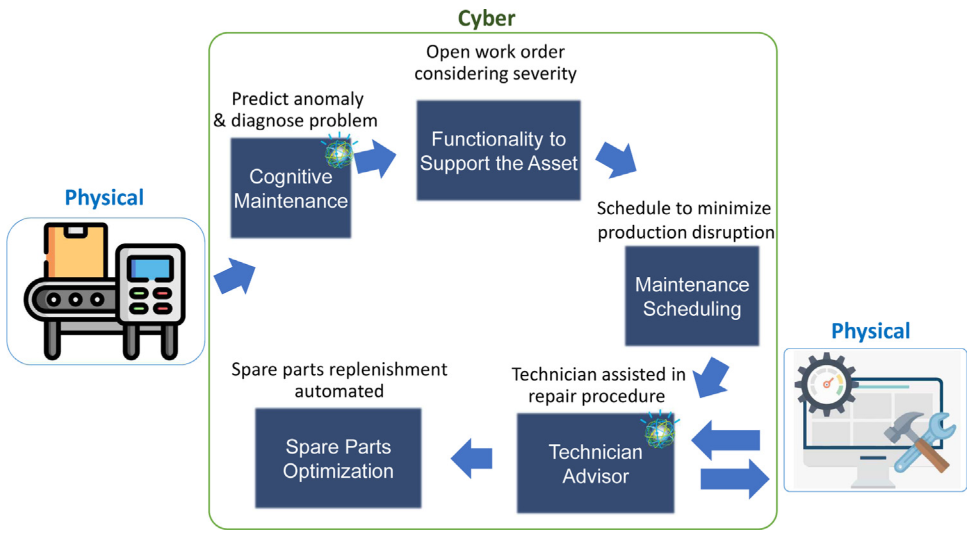

Fig. 3. Cyber-physical process to perform maintenance.

1. **Definition of research questions:** it guides the elaboration of research questions to be used to search for relevant papers in the literature;
2. **Search process:** it presents the research strategy and the scientific sources used in the search for relevant papers;
3. **Studies selection:** definition of the criteria applied to select relevant papers;
4. **Quality assessment:** quantitative analysis of the quality of the selected studies;

We discuss each of the steps in the following subsections.

3.1. Research questions

A crucial part of an SLR is the elaboration of research questions [28]. In this article, the research questions should provide the means to understand the use of ML along with ontologies in the context of PdM in Industry 4.0 scenarios.

To elaborate the research questions, we carried out preliminary research and the analysis of the resulting papers. Besides, the authors also used their experience to create the research questions.

We specify a general question to guide the search for challenges in the field of research. Based on the central question, we establish specific ones to emphasize existing solutions and to identify gaps and directions for future research.

With the general research question in mind, we created the following specific questions (SQ).

What are the challenges and open questions regarding machine learning and reasoning for predictive maintenance in Industry 4.0?

- SQ1: What are the challenges of applying machine learning towards predictive maintenance?
- SQ2: Which machine learning techniques are usually in the context of predictive maintenance?
- SQ3: In which contexts are ontologies used for predictive maintenance?

3.2. Search process

The purpose of SQ1 is, meeting the general question elaborated, to identify the technical challenges that researchers encounter when proposing solutions for PdM using ML. Moving forward on this issue, in SQ2, we seek to identify which ML algorithms are employed to verify if there is any consensus in the scientific community. Finally, SQ3 aims to understand whether ontologies or reasoning have any space in the area of PdM, and what role it plays.

We perform two steps to conduct the search process. The first one is the creation of a search string, while the second involves the selection of the sources. The design of the search string demands a preliminary reading of selected papers related to the field of interest. Boolean operators are applied to improve search string performance. These operators contemplate terms that are synonymous with the keywords already defined.

(''Industry 4.0'' AND (''Machine Learning'' OR ''Deep Learning'') AND ''Predictive Maintenance'' AND ('Àrchitecture'' OR ''Framework'') AND ('Òntology'' OR ''Reasoning''))

In this article, all the results obtained came from electronic sources. The choice of the databases aimed at analyzing papers published in journals and conferences that cover the concepts of Industry 4.0. The selected electronic databases were IEEE,[1] Google Scholar,[2] Springer,[3] ACM Digital Library[4] and ScienceDirect.[5]

3.3. Papers selection process

Following the definition of the search string and gathering articles from the selected electronic databases, it is necessary to remove all studies that are not relevant to the goals of this article. To remove these papers, the following exclusion criteria (EC) applies.

- EC 1: papers not directly related to PdM.
- EC 2: papers not directly related to ML.
- EC 3: papers that presented results of surveys or reviews.
- EC 4: papers published before the year 2015.

Two researchers conduct the filtering process following the steps presented below. The results are analyzed by a third one whenever a discrepancy occurred.

1. **Removal of duplicates:** in some situations, the same paper is available in different sources, like in IEEE and Google Scholar, for example. In this case, we remove the duplicates;
2. **Title and abstract analysis:** the researcher reads the title and the abstract of the paper and judges whether or not it is sufficient to assess the importance of the paper to the study;
3. **Entire text analysis:** it applies in situations where the title and the abstract are not very clear about the proposed solution. Nevertheless, the presented ideas look promising for the goals of this literature review.

As part of the SLR methodology, we exclude from the corpora papers published before 2015 and those classified as surveys or reviews. Besides, we disregarded any work with no scientific character, such as a blog post or magazine article. We also remove the duplicates of papers that appear in more than one database. After the conclusion of this step of the SLR, the remaining papers pass to the phase of quantitative evaluation.

3.4. Quality assessment

According to the methodology [27], in this step, we define the criteria for qualitative evaluation of the selected papers. The evaluation takes into account the following points: (I) the purpose of the research; (II) whether the authors contemplate a research methodology or propose an architecture or a framework; (III) the results accomplished; and (IV) whether the selected work uses ontologies. The following questions apply to select papers that meet the quality requirements.

- Is the purpose of the research presented?
- Is there an architecture/framework proposal or a research methodology?
- Are the research results presented and discussed?
- Does the paper use reasoning or an ontology?

Based on Kitchenham's methodology [27], we define three possible answers, each one receiving a grade: Yes = 1, Partial = 0.5, and No = 0. After two researchers have graded the papers, we dealt with the discrepancies in a discussion meeting. In a few cases, where the two researchers did not reach consensus, a third researcher read the papers to resolve the discrepancies. At the end of such a meeting, to decide whether each study should be kept or excluded from the original corpora, we then applied criteria for excluding articles according to the grade. More formally, such criteria are:

- Articles mainly organized as comments or personal opinions are excluded from the set since they usually do not present a validation methodology;
- Articles that are graded below 2.5 by the researchers, since at least 2 of the questions received 'no' or 'partial' as the answer, indicating a not very relevant publication for this SLR;

After applying the mentioned criteria onto the original set of papers, we read the remaining ones to answer the research question. We discuss the results in Section 4.

4. Search results

This section discusses the result of the search process, the selection process, and the qualitative analysis of the selected papers. We summarize the results in Fig. 4. The description of each step and the number of remaining papers are also in the figure.

We detail the application of the SLR methodology as follows. Section 4.1 discusses relevant papers that are applied to the context of Industry 4.0 but do not meet all the criteria to be part of this SLR. After analyzing the exclusion criteria, details on the quality assessment of the papers are presented in Section 4.2.

4.1. Exclusion of papers from the initial corpora

Fig. 4 shows the number of articles obtained in each of the databases selected in the *Initial Search* stage. We group these articles in the step called *Database Joint*, resulting in a total of 562 papers. The filter of *Impurities Removal* removed duplicates eventually found in more than one database, and surveys, reviews, book chapters, or non-scientific papers like magazine articles, resulting in a total of 288 out of 562 papers. The next step comprised the application of the *Exclusion Criteria* mentioned in Section 3.3. The number of papers considered relevant for these reviews reduced down to 155 in this phase.

We removed some papers because, despite clearly addressing issues related to PdM, they do not reflect the use of ML models. For example, Stojanovic et al. [29] created an architecture that applies the concepts of big data in the context of self-healing manufacturing. The solution, named PREMIuM, briefly mentions ML models as part of a prediction architecture. Between the several layers of PREMIuM, the cloud layer is responsible for analyzing data, using a strategy that involves at least two ML methods. Despite that, the paper does not provide details on the ML models or the training process.

[1] https://ieeexplore.ieee.org/.
[2] https://scholar.google.com/.
[3] https://link.springer.com/.
[4] https://dl.acm.org/.
[5] https://www.sciencedirect.com/.

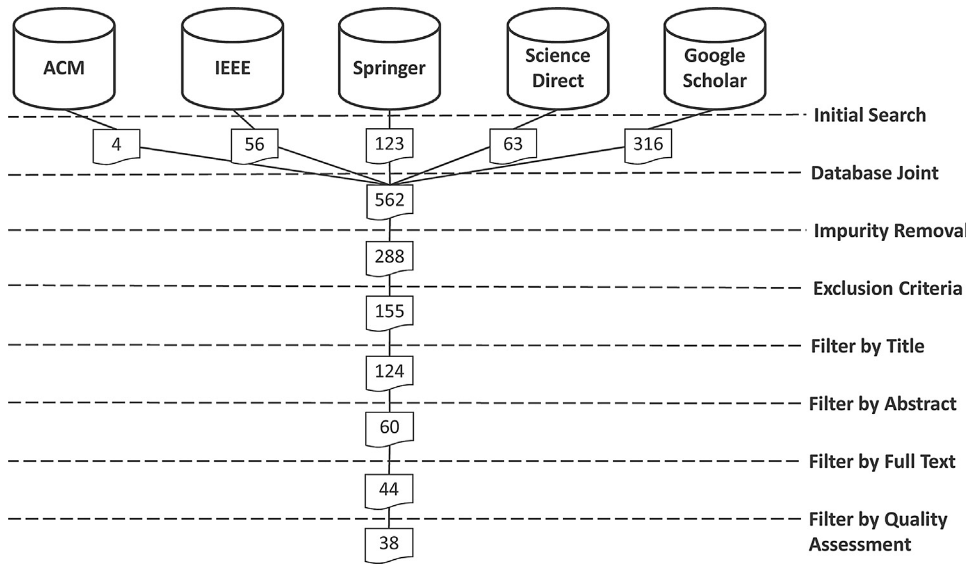

Fig. 4. Filter process of selected articles by database.

The exclusion criteria also removed a solution proposed by Zenisek et al. [30]. This solution generates the first data streams as a way of helping time series researchers to simulate scenarios with realistic monitoring conditions. We also removed the intelligent maintenance support framework proposed by Bumblauskas et al. [31]. In this case, the authors proposed an algorithm to minimize costs across the supply chain and a methodology for establishing predictive maintenance plans. Both approaches fail to apply ML models, so we excluded them from the corpora.

May et al. [32] proposed a novel approach to prevent failures, composed of a set of strategies, called Z-strategy, intended to extend the life of production systems. Z-strategy is capable of predicting failures at the component, machine, or system level. This platform matches the subject of PdM but is not related to ML. Golightly et al. [33] concentrated on human factors, such as data interpretation and visualization. The study identified factors to help the implementation of predictive asset management. Through interviews with experts, the authors identified the organizational problems associated with the development and adoption of PdM systems. The results are recommendations on how to mitigate these problems.

The solutions show several cases where the application of PdM goes beyond the use of ML models. These solutions are generally related to the elaboration of frameworks that introduce the concepts of PdM with a more holistic view. We can also find opposite approaches in the literature. These solutions, although addressing ML in the context of Industry 4.0, do not apply for PdM. These works appeared in the initial corpora because they mention PdM-related terms in sections like related work or references.

In this context, Syafrudin et al. [34] designed a real-time system to improve decision making. This system helps to prevent losses due to unplanned manufacturing failures. The system's workflow considers three steps. The first one is the selection of IoT devices to monitor an automotive manufacturing environment. The second one involves processing large amounts of generated data. The final step is the creation of a hybrid model for fault detection. The hybrid model uses density-based spatial clustering of applications for anomaly detection, and Random Forest to classify events as normal or abnormal.

Romeo et al. [5] provides an ML-enabled framework to help designers and laboratory technicians to make the best operating description of a machine. The solution relies on Decision Trees, k-Nearest Neighbors, and Neighborhood Component Features Selection algorithms to obtain the recommendations. The resulting solution predicts whether the machine specifications, e.g., the number of blades, speed, and shaft size, match the operating parameters, like torque, flow, pressure, and gate. The authors claim that the solution provides easier decision making, conserving company knowledge, saving working hours, and increasing computational speed and accuracy.

Zhang et al. [35] propose an approach to apply ML techniques in the industry. The solution consists of a framework to provide a reference to plan, design, and implement industrial artificial intelligence in different areas of manufacturing. The proposal is theoretical and broadly covers seven dimensions related to industrial artificial intelligence: objects, domain, application stages, application requirements, intelligent technology, intelligent function, and solutions. The framework was evaluated considering five industrial fields and showed to be effective in helping industries to plan how to use artificial intelligence-related solutions.

Ali et al. [4] designed a software middleware to collect and analyze data from different applications in a real-time fashion. The authors evaluate the solution in a scenario where the middleware provided information to ensure optimal production forecasts. Even though not dealing with PdM, the use of ML for production forecasting presents some challenges that look like those of PdM. These challenges include the difficulty of obtaining relevant data without lacks or gaps and the need for domain knowledge for the selection of variables and construction of models. According to the nature of the collected data, Multiple Linear Regression, Support Vector Regression, Decision Tree, and Random Forest algorithms were applied by the middleware to make predictions.

Malek et al. [36] proposed a failure prediction methodology to provide reliability in different scenarios. The authors evaluated the performance considering two scenarios: (I) malware detection and (II) computer failure detection. To do that, the authors applied Naive Bayes, Logistic Regression, and J48 Decision Tree algorithms. The paper briefly mentions concepts related to Industry 4.0 and PdM in the literature review.

A framework combining data collection, pre-processing, and ML models training to identify behaviors that may influence manufacturing is the proposal of Carbery et al. [37]. The solution uses artificial intelligence to assist engineers in increasing machine performance and supporting decision making. It results in a four-

Table 1
Answers and grades.

Answer	Description	Grade
Y (Yes)	The paper explicitly answers the question	1.0
P (Partial)	The paper answers to part of the question	0.5
N (No)	The paper does not mention the topic	0.0

stage workflow: (I) data collection, (II) pre-processing, (III) training data generation, and (IV) artificial intelligence model creation. The authors focused on challenges and solutions for data pre-processing and feature selection.

To demonstrate the role of cloud computing and the use of artificial intelligence to improve factory performance, Wan et al. [38] proposed a vertically integrated, four-tier cloud-assisted smart factory architecture. The layers that compose the architecture are: (I) Smart Device Layer, (II) Network Layer, (III) Cloud Layer, and (IV) Application Layer. ML models are implemented in the Network Layer to perform tasks involving network optimization. These models are also present in the Application Layer, where the use of ML aims to perform failure detection, but not predictive maintenance.

Costa et al. [39] introduced a framework for knowledge representation that transforms unstructured data such as logs or machine documentation into highly-structured data. The approach uses an ontology to assist in structuring and enriching information. It also applies ML techniques to process natural language. Although the solution uses reasoning, its purpose is not related to predictive maintenance.

Finally, we can cite the work developed by Sala et al. [40]. The solution applies a data-driven strategy to predict temperature and chemical concentration in the Basic Oxygen Furnace Steelmaking process. To do that, the authors consider different machine learning models, like Ridge Regression, Random Forest, and Gradient Boosted Regression Trees. As in several other works, the authors only mention PdM-related terms.

As shown in Fig. 4, after applying the *Exclusion Criteria*, we screened the 155 remaining papers considering three filters mentioned in Section 3.3: (I) *Filter by Title*, (II) *Filter by Abstract*, and (III) *Filter by Full Text*. After the application of these three filters, we classified the 44 remaining papers according to quality parameters. In this case, we score each paper according to a methodology and eliminate those graded below a pre-defined threshold.

4.2. Performing the quality assessment to select relevant papers

This section follows the quality criteria defined in Section 3.4 to conduct the qualitative analysis of the papers. Researchers answered to each question according to Kitchenham's methodology [27]. We present the possible answers and the respective grades in Table 1. Relevant papers for this SLR are those that received 2.5 points or more. We selected these threshold values to guarantee that at least one of the questions receives maximum evaluation in the worst-case. Also, we identified that all papers that obtained less than two points do not present results based on a use-case. Besides, all papers graded less than 2.5 do not present the concepts of reasoning or propose an ontology.

To grade the papers, the researchers first applied a filter by the title, reducing the number of studies from 155 to 124. These 124 papers went through the abstract filtering process, which selected 60 relevant ones. At the final filtering process, the remaining papers went through a complete analysis of the text. This phase excluded 16 studies and resulted in the 44 papers considered the most relevant for this SLR.

The research team answered the following questions to assess the quality of the 44 remaining papers.

- Is the purpose of the research presented?
- Is there an architecture/framework proposal or a research methodology?
- Are research results presented and discussed?
- Does the paper implement the concepts of reasoning or proposes an ontology?

We present the answers to the questions and the resulting scores in Table 2, in descending order. References [16,41,42] received an asterisk (*) because, according to the filters, these papers should be out of the corpora. However, we kept them because they propose the use of ontologies to implement reasoning in the context of PdM, which is a topic of interest for this SLR.

After the quality assessment performed over 44 papers, 6 references [63,64,61,62,11,10] were removed due to the score metric. The 38 remaining papers are classified by year and database in Fig. 5, where the *x*-axis represents the range of years considered in this research process, from 2015 to June 2020. Discussions regarding these papers are presented in Section 5.

Fig. 6 presents an illustration where we group the articles according to the total score obtained. The shape of the icon identifies the publication source of each article. The results show that only four papers fully answered all quality assessment questions. In the second column of the graph, one can see that more than half of the papers scored 3 points. These results allow us to conclude that only relevant papers remained in the corpora after the application of the filtering process.

Fig. 7 shows the importance of each question in the selection of papers. The image presents the grade received by the papers grouped by question. It is possible to conclude that SQ1 and SQ2 are very important to the results of this SLR. The scores indicate that all papers satisfied SQ1, and at most, two papers did not satisfy SQ2. This behavior shows that questions SQ3 and SQ4 were the main reasons for the exclusion of papers from the initial corpora. Another conclusion is that the majority of the papers focus on practical aspects. One can conclude this because a positive evaluation of SQ3 means that the paper presents the results of a use-case. Finally, a close look at the evaluation of SQ4 allows us to deduce that reasoning and ontologies did not receive much focus in the analyzed period. Only five papers fully answered the question or at most ten partially answered the question.

There is growing interest in the application of ML and reasoning for predictive maintenance. This field of research did not receive attention in 2015 and started growing in 2016 to a peak of 14 publications in 2018. Splitting the period into two parts, we can see that the first three years of research represent only 24% of the total number of selected publications. On the other hand, the last three years of the period concentrate 76% of the total number of contributions.

Among the five sources considered to carry out the search process, four had papers selected in this SLR. Google Scholar collaborated with six contributions. In SpringerLink, we found eight publications, while nine manuscripts came from IEEE Xplore. The only source that did not publish research considered in this SLR was the ACM digital library.

5. Answer to the research questions and discussion

This section analyzes the contributions of the most relevant papers selected in this SLR. To do that, in each subsection, one of the research questions defined in Section 3.1 is answered, considering the contributions of the papers found in the literature.

Table 2
Quality assessment scores.

Ref.	Year	Authors	SQ1	SQ2	SQ3	SQ4	Score
[43]	2019	Cao et al.	Y	Y	Y	Y	4.0
[44]	2019	Ansari, Glawar and Nemeth	Y	Y	Y	Y	4.0
[41]*	2018	Nuñez and Borsato	Y	Y	Y	Y	4.0
[42]*	2016	Schmidt, Wang and Galar	Y	Y	Y	Y	4.0
[45]	2020	Cheng et al.	Y	Y	Y	N	3.0
[46]	2020	Calabrese et al.	Y	Y	Y	N	3.0
[47]	2020	Daniyan et al.	Y	Y	Y	N	3.0
[48]	2020	Hoffmann et al.	Y	Y	Y	N	3.0
[49]	2020	De Vita et al.	Y	Y	Y	N	3.0
[50]	2020	Chen et al.	Y	Y	Y	N	3.0
[51]	2019	Huang et al.	Y	Y	Y	N	3.0
[52]	2019	Rivas et al.	Y	Y	Y	N	3.0
[24]	2019	Xu et al.	Y	Y	Y	N	3.0
[53]	2019	Cerquitelli et al.	Y	Y	Y	N	3.0
[54]	2018	Ding et al.	Y	Y	Y	N	3.0
[8]	2018	Carbery, Woods and Marshall	Y	Y	Y	N	3.0
[55]	2018	Yuan et al.	Y	Y	Y	N	3.0
[26]	2018	Schmidt et al.	Y	Y	Y	N	3.0
[56]	2018	Strauss et al.	Y	Y	Y	N	3.0
[14]	2018	Zhou et al.	Y	Y	Y	N	3.0
[57]	2018	Peres et al.	Y	Y	Y	N	3.0
[22]	2018	Liu et al.	Y	Y	Y	N	3.0
[13]	2018	Adhikari et al.	Y	Y	Y	N	3.0
[12]	2018	Schmidt et al.	Y	Y	Y	N	3.0
[20]	2018	Kiangala et al.	Y	Y	Y	N	3.0
[18]	2018	Hegedus et al.	Y	Y	N	Y	3.0
[21]	2018	Kaur et al.	Y	Y	Y	N	3.0
[58]	2017	Diez-Olivan et al.	Y	Y	Y	N	3.0
[59]	2017	Li et al.	Y	Y	Y	N	3.0
[19]	2017	Ferreira et al.	Y	Y	Y	N	3.0
[23]	2017	Crespo et al.	Y	Y	Y	N	3.0
[60]	2016	Gatica et al.	Y	Y	Y	N	3.0
[9]	2016	Chukwuekwe et al.	Y	Y	Y	N	3.0
[16]*	2020	Ansari et al.	Y	Y	P	N	2.5
[17]	2019	Sarazin et al.	Y	Y	P	N	2.5
[15]	2019	Bousdekis et al.	Y	Y	P	N	2.5
[3]	2018	Cachada et al.	Y	Y	P	N	2.5
[32]	2018	May et al.	Y	Y	P	N	2.5
[61]	2019	Glawar et al.	Y	Y	N	N	2.0
[62]	2019	Talamo et al.	Y	Y	N	N	2.0
[11]	2018	Balogh et al.	Y	Y	N	N	2.0
[63]	2018	Issam et al.	Y	Y	N	N	2.0
[64]	2015	Gao et al.	Y	Y	N	N	2.0
[10]	2017	Wang et al.	Y	N	N	N	1.0

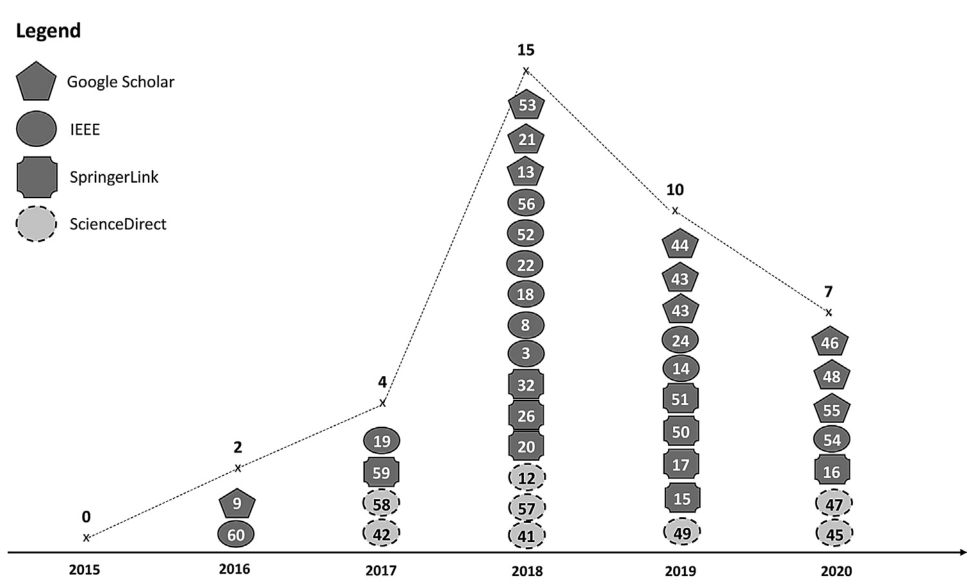

Fig. 5. Publication year of selected articles by database.

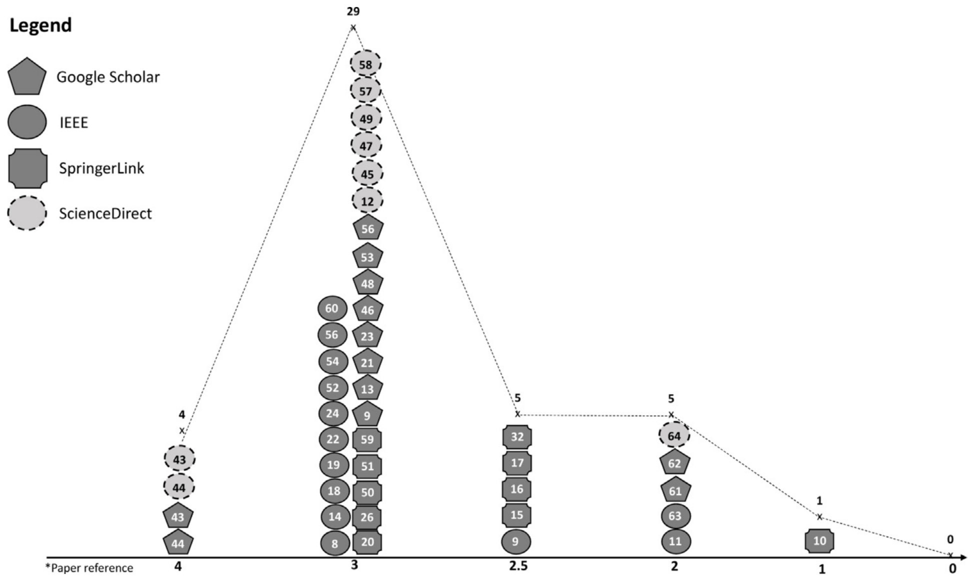

Fig. 6. Publication score of selected articles by database.

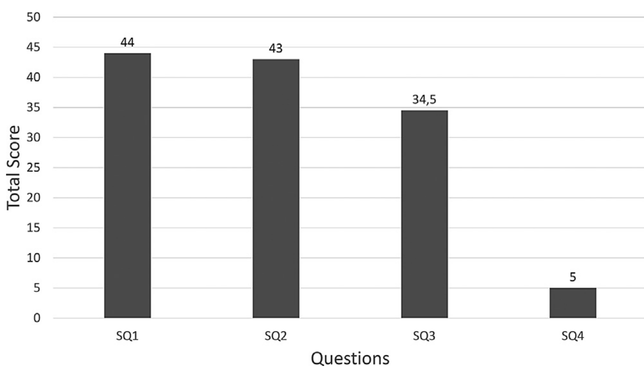

Fig. 7. Total score by question.

5.1. What are the challenges and open questions regarding machine learning and reasoning for predictive maintenance in Industry 4.0?

This research question contributes to the scientific community by identifying and classifying the current challenges and open issues regarding ML and reasoning in the context of PdM. To do that, we propose the taxonomy presented in Fig. 8.

We designed the taxonomy considering two types of elements. The blue boxes represent broad fields related to predictive maintenance that can have different kinds of challenges related to them. On the other hand, the green boxes are specific challenges or open issues that we identified based on the results of the SLR. The numbers under each green box are citations to papers that propose solutions to tackle that specific challenge.

The first element of the taxonomy refers to the general field of *Predictive Maintenance*. The analysis of the literature showed that this field usually poses challenges related to the reduction of maintenance-related costs or aims at improving production efficiency by predicting necessary maintenance. Solutions to deal with these generic challenges can be divided into three groups: (I) *Big Data* analytics, (II) *Machine Learning* models, and (III) *Ontology* and reasoning-related proposals. Each of these three areas has specific challenges in the context of PdM. We present and discuss these challenges as follows.

The first specific challenge in the taxonomy is directly related to predictive maintenance and concerns to *integration* issues. This kind of problem typically affects the company as a whole. The literature presents several solutions to build integrated solutions and methods that are capable of handling different processes related to PdM. The most common approach of these works is to propose architectures, strategies, principles, and tools that seek to unify each step of the process for deploying systems to enable predictive maintenance [15–18,3,19–21,45,47,48,50]. These tools are generally capable of integrating processes such as sensor data collection on machines, data processing, creation and training of machine learning models, as well as the integration with other information sources, e.g., ERP systems to send PdM-related alerts in a user-friendly interface.

Big data is one of the more challenging areas in the context of PdM. Some issues concern the need for real-time monitoring and consequently processing a large amount of data generated by the sensors. In this context, guaranteeing good values on metrics like *Latency*, *Scalability*, and network *Bandwidth* is a problem as some predicted events require immediate action to prevent failures. Works like Liu et al. [22] and Zhou et al. [14] propose the use of edge computing to bring the processing closer to the data collection point and delegate to the cloud only tasks that do not require immediate action. On the other hand, Crespo et al. [23] proposed a framework that implements the concept of distributed computing, so when the system capacity reaches high usage values, the system manages the resulting overhead by distributing the processing.

Still in the context of *Big Data*, another open challenge concerns *Data acquisition*. The open issues include the difficulty of obtaining quality data and interpreting it. A considerable portion of the collected data has missing values, is poorly structured, or has no annotations. To deal with this issue, Hegedhus et al. [18] proposed a solution to preprocess data and turn it usable for predictive maintenance. Another approach to deal with data acquisition is the framework designed by Strauss et al. [56] that enables the monitoring and data acquisition in legacy machinery. Although some pieces of equipment do not have data from condition monitoring sensors, such as vibration and temperature, some alternatives can be put into practice as demonstrated by Calabrese et al. [46], which makes use of event log data provided by diagnostic systems already attached to the machine, creating a low-cost alternative to predict failures.

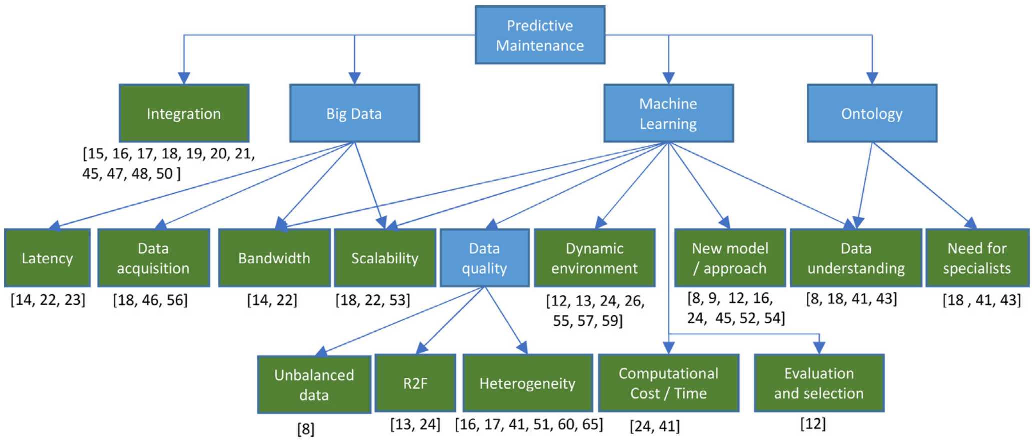

Fig. 8. A taxonomy to classify challenges and open issues in PdM.

Another major challenge in predictive maintenance is to establish the grounds for applying *Machine Learning* models. Like in the context of *Big Data*, the need for real-time decision making requires high levels of *Scalability* and network *Bandwidth*. In this sense, training learning models in the edge of the networks is a solution that has been proposed by Liu et al. [22] and by Cerquitelli et al. [53].

Imbalanced data appears as a challenge related to data quality. Carbery et al. propose a framework using multiple stages of pre-processing, feature selection, and combining factor levels to deal with datasets that show small amounts of failures compared to the quantity of operational data. Another challenge in the context of ML is to obtain data that shows the tendency of normal state behavior to failure, called run to fail (*R2F*). Xu et al. [24] and Adhikari et al. [13] deeply explore R2F. This kind of data is important to identify problems because, in this case, it is necessary to train the models with annotated failure-related datasets. In this sense, Gatica et al. [60] proposed a top-down strategy consisting of first understanding machine operation and then taking action to deal with the problem.

Besides, the *heterogeneity* of the datasets also poses as an issue that deserves attention from the scientific community. Both the lack and the excess of heterogeneity harm the ML models. The lack of heterogeneity was explored by Schmidt et al. [65] and by Nunez et al. [41]. In both approaches, the authors conclude that this characteristic turns training the ML models more difficult. The authors also found out that in several cases, only one data source is available, e.g., all data come from one specific machine. Excessive heterogeneity also impacts negatively on the ML model training. In this direction, Sarazin et al. [17] and Gatica et al. [60] analyzed the behavior of ML models that receive large amounts of training datasets from many different sources. Another problem related to this topic was investigated by Ansari et al. [16]. The work considers the challenges related to processing data with multiple structures of maintenance data. These structures include sensors information, maintenance text reports, and multimodal data where a machine sensor signal can provide more than one information and can result in inappropriate planning, monitoring, or controlling that decreases the remaining useful life. This large amount of data is addressed by Huang et al. [51], since using the original dataset for diagnosis and fault prediction is difficult, so the authors proposed an algorithm for data fusion processing.

The general conclusion we can reach on this topic is that both lack and excess of *heterogeneity* can impact on the predictability of the algorithms. Proposals to mitigate this problem are available. Different authors found out that one of the causes of data heterogeneity is the fact that manufacturing plants are dynamic environments. Schmidt et al. [12] claim it is not recommended to use data obtained only in the laboratory. Moreover, Li et al. [59] showed that it is also not feasible to train ML models using data provided by only one model of equipment. Another aspect emphasized in related work is that the process executed by a given machine can also change dynamically [24]. Alternatives to deal with this challenge include training multiple learning models [13], utilizing data produced by the pieces of equipment that have operated in comparable conditions [26], applying data mining techniques to generate context information [57], or apply strategies using deep learning algorithms [55].

Another challenge related to ML concerns the lack of a universal model that applies to multiple scenarios. In this context, Ansari et al. [16] discussed the possibility of proposing new models to deal with data heterogeneity issues. Xu et al. [24] use deep transfer learning to extract a high-level representation from a large amount of data from a specific domain and transfer that knowledge to a target domain. Transfer learning makes it possible to use models trained in a given domain to perform tasks in other scenarios that are related to the original one. Other approaches advocate in favor of applying existing models to new scenarios [52,8,9]. Generally, these approaches test different models to evaluate which one fits better for a specific situation. For example, Schmidt et al. [12] propose a classification based on vibration limit values to predict failure, and Cheng et al. [45] compare two algorithms to predict the HVAC system condition. Ding et al. [54] propose Knowledge-based reasoning to predict steel bridge performance deterioration. The computational cost for training these ML models is also a challenge to be addressed, according to Xu et al. [24] and to Nunes et al. [41].

The last area identified in this SLR refers to the application of ontologies for predictive maintenance. This scenario typically associates ontologies with the need to understand the data, as an example, we have the Measurement Ontology, presented by Schmidt et al. [42], which relates measurements to other data such as date of collection, the machine that generated the data, and the process executed at that time. It can also store information about the environment, such as temperature and physical location [42]. This field of research is still incipient, but some works already investigated the topic (Hegedus et al. [18], Cao et al. [43], Nunez et al. [41]).

5.2. *Which machine learning techniques are usually in the context of predictive maintenance?*

The literature covers a wide variety of ML techniques, each with specific characteristics and applications. The focus of this section is the application of these models for predictive maintenance. We intend to highlight the commonly used ML techniques and the reasons for selecting specific techniques.

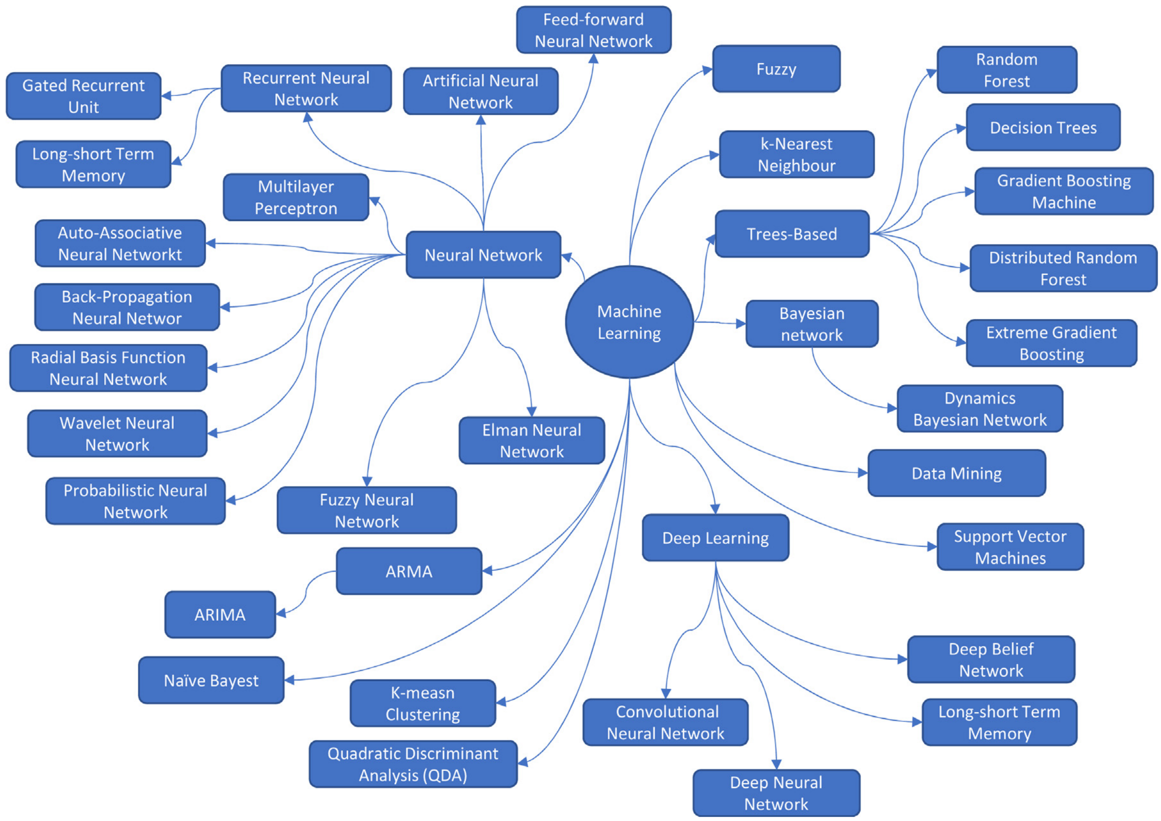

Fig. 9. Taxonomy of machine learning techniques.

ML algorithms apply to several stages of predictive maintenance, such as diagnosis, prognosis, and estimation of useful life. In Fig. 9, we present a taxonomy of these ML algorithms. The taxonomy highlights the connection among classes of algorithms.

Several authors use Artificial Neural Networks (ANN) to tackle problems related to predictive maintenance. Li et al. [59] proposed a framework for fault detection and prediction, which is capable of performing error correction regardless of machine or process type. According to the authors, an ANN model can execute the prediction task through training with backlash errors of the last three weeks to predict the backlash error of the subsequent week. The selection of the ANN technique occurred because it is already widely used to compensate for slack errors in computer-controlled machine centers. Crespo et al. [23] applied ANN to identifying when asset behavior abnormalities can appear in the context of Industry 4.0. The ANN is combined with association rules to determine the operating conditions considered abnormal. The framework proposed by Cheng et al. [45] uses ANN and SVM. The results prove that the framework is capable of predicting the future condition in MEP components and, consequently extend the RUL of the components. Daniyan et al. [47] also present a framework using ANN, but in this case, the model estimates the RUL of a bearing using a data series of temperature.

Other solutions involve the implementation of Recurrent Neural Networks (RNN) that are a type of ANN capable of incorporating memory. Rivas et al. [52] adopted Long Short-Term Memory (LSTM) RNN model for failure prediction. The authors focused on creating an LSTM model to identify a possible future malfunction using two models. The first one classifies whether the engine has more than 100 life cycles (classification problem), while the second one predicts the remaining number of cycles (regression problem). Cachada et al. [3] also used LSTM along with a second technique called Gated Recurrent Unit (GRU) for a similar purpose. Both models were applied because they implement the ability to consider historical data to predict future behavior. Several authors [12,14] assess the performance of ANN and compare it with other techniques like Support Vector Machine (SVM) and Randon Forest (RF). Yet in the area of ANN, some works consider the implementation of Auto-Associative Neural Networks (AANN). Liu et al. [22] proposed a PdM framework in which an AANN identifies irregularities in railways. This information is used to predict failures and suggests the actions to be taken in advance.

As an alternative to using raw sensor data, which can be composed of different types of information coming from various sensors, raw data can pass through a fusion process, conducted by a correlation algorithm. Hung et al. [51] used algorithms based on neural networks, Back-Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), Elman Neural Network (ENN), Probabilistic Neural Network (PNN), Fuzzy Neural Network (FNN) and Wavelet Neural Network (WNN) to implement the fusion process. The authors use the resulting to perform fault prediction and detection.

Table 3 summarizes the works that apply neural networks, specifying their classes, and which task the neural network performs within each paper. As can be seen, an ML algorithm can be used in several stages in the PdM process to perform classification or prediction tasks.

Another common approach identified in the SLR is the application of ML models to validate proposed frameworks. Schmidt et al. [12] evaluated the performance of k-Nearest Neighbor (kNN), Back-propagation Feed-forward Neural Network (FFNN), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayesian (NB) in various scenarios to obtain the best combination of techniques to deal with time-series prediction. Zhou et al. [14] also used kNN and DT along with SVM and Multi-layer

Table 3
Approaches based on NN.

Paper	ML method	Application
Li et al. [59]	ANN	Fault Prediction
Crespo et al. [23]	ANN	Anomaly detection
Rivas et al. [52]	LSTM	RUL estimation
Cachada et al. [3]	LSTM/GRU	Failure prediction
Schmidt et al. [12]	FFNN	Time-series prediction
Zhou et al. [14]	MLP	Fault classification
Liu et al. [22]	ANN	Anomaly detection
Liu et al. [51]	BPNN, ENN, RBFNN, PNN, FNN and WNN	Data fusion
Cheng et al. [45]	ANN	Condition prediction
Daniyan et al. [47]	ANN	RUL estimation

Table 4
Approaches based on ML.

Paper	ML method	Application
Schmidt et al. [12]	kNN, FFNN, DT, RF, NB	Time to fail class classification
Zhou et al. [14]	kNN, FT, SVM	Fault Classification
Carbery et al. [8]	BN	Diagnosing and Predicting faults
Ansari et al. [16]	DBN	Prediction of failure events
Chukwuekwe et al. [9]	ARMA	Fault Prediction
Adhikari et al. [13]	SVM	Anomaly Detection
Adhikari et al. [13]	ARIMA	RUL
Adhikari et al. [13]	DT, SVM, NB, RF	Fault classification of
Calabrese et al. [46]	GBM, DRF, XGBosot	RUL

Perceptron (MLP), to evaluate a framework for diagnostics and prognostics.

Dealing with large amounts of data to predict failures was the goal of Carbery et al. [8]. To do that, the authors used a Bayesian Network (BN). The selection of this technique was justified because BN is known for performing well under uncertainties. Moreover, this technique can decompose complex problems in more manageable ones using conditional probabilities. A special BN, called Dynamics Bayesian Network (DBN), is proposed by Ansari et al. [16]. The proposal is part of a framework designed to predict failures and to measure the impact of such a prediction on the quality of production planning processes and maintenance costs.

Another class of models that is gaining attention from the scientific community is Auto-regressive Moving Average (ARMA). Chukwuekwe et al. [9] presented a solution that considers the vibration date to predict failures in the context of Industry 4.0. Auto-regressive Integrated Moving Average (ARIMA), which is a variation of ARMA, was applied by Adhikari et al. [13] in a predictive maintenance framework to predict the remaining useful life of components. The selection of ARIMA was due to its ability to use historical data to estimate future behavior. In addition to RUL estimation, the framework proposed by Adhikari et al. [13] uses SVM to identify possible anomalies and to classify algorithmic failures.

Applying tree-based algorithms to predict the probability of a failure, Calabrese et al. [46] present an architecture that uses three different algorithms to classify machines with 30 days of less of RUL. The Gradient Boosting Machine (GBM) generated the models that obtained the best results for classification when compared to the Distributed Random Forest (DRF) models and Extreme Gradient Boosting (XGBoost) models.

Table 4 presents a summary of the researches that apply ML without implementing NN or DL. The Application column shows which task the ML performs within each paper. We can see that an ML algorithm can execute more than one process, as well as the execution of that process, may demand more than one ML model.

Deep learning (DL) is an area that is beginning to receive more attention in the context of predictive maintenance, especially for

Table 5
Approaches based on DL.

Paper	ML method	Application
[55]	CNN	RUL
[49]	DNN	Anomaly Detection
[24]	DNN, DTL	Fault Diagnosis
[51]	CNN, LSTM, DBN, DNN	Fault Diagnosis/Prediction

processes like diagnosis, prognosis, and RUL. Among the works identified in this SLR, we can mention the use of a Convolutional Neural Network (CNN) in the task of extracting features from databases without the need for prior knowledge of the data. In this sense, Yuan et al. [55] automate the discovery of new features hidden in the database to perform flaw detection and prediction. Another approach regards anomaly detection using DL techniques. In this context, De Vita et al. [49] apply a deep autoencoder consisting of a Deep Neural Network (DNN) to reduce the dimensionality of the data, generating a new dataset for further training of detection algorithms such as k-means.

The lack of training data or even different data distribution in practical application are challenges that DL tries to solve, as can be seen in the two-phase Digital-twin-assisted Fault Diagnosis using Deep transfer learning (DFDD) proposed by Xu et al. [24]. In the first phase, experts model a virtual Deep Neural Network (DNN), while in the second phase, the authors use Deep Transfer Learning (DTL) to transfer the knowledge previously created to the physical world, overcoming the lack of data. Huang et al. [51] dealt with a use-case that demanded data fusion to allow prognosis and diagnosis. To do that, the authors applied solution as CNN, Deep Belief Network (DBN), DNN, Automatic Encoder (AE), and LSTM.

Despite the promising use of DL algorithms in the context of PdM, we found relatively few works that addressed the use of these algorithms to perform predictive maintenance-related tasks. In Table 5, we summarize the proposals that apply DL.

5.3. *What are the contexts in the use of ontology in predictive maintenance?*

Predictive maintenance can rely on ontologies for various purposes. Fig. 10 presents a taxonomy that describes the main applications of ontologies in predictive maintenance. This section discusses these applications.

One of the main goals of an ontology is to provide *context awareness*. Prima framework [44] applies this concept. Prima models real-world objects according to their properties and functions. The solution is capable of gathering and associating sensor data. The framework allows, for example, the association of a motion sensor with a specific machine process and, at the same time, with energy consumption information. This kind of approach *models knowledge, and shares* it for decision-making purposes. Prima also *stores information semantically*. The solution models and stores data in a standardized manner, which allows the framework to access various data sources. This characteristic also copes with another feature, which is to provide *interoperability* among different domains, achieved by implementing the so-called domain ontology concept.

Cao et al. [43] explore the ability to *store information semantically*. The work explores the context of condition-based maintenance. Using the concept of *rules*, the authors propose an algorithm that generates Semantic Web Rule Language (SWRL). This solution implements reasoning to describe events and temporal constraints. The results facilitate decision making through the *prediction* of failures.

A challenging issue, which affects both ML models and ontologies, is the need for *expert users* to analyze the context and provide

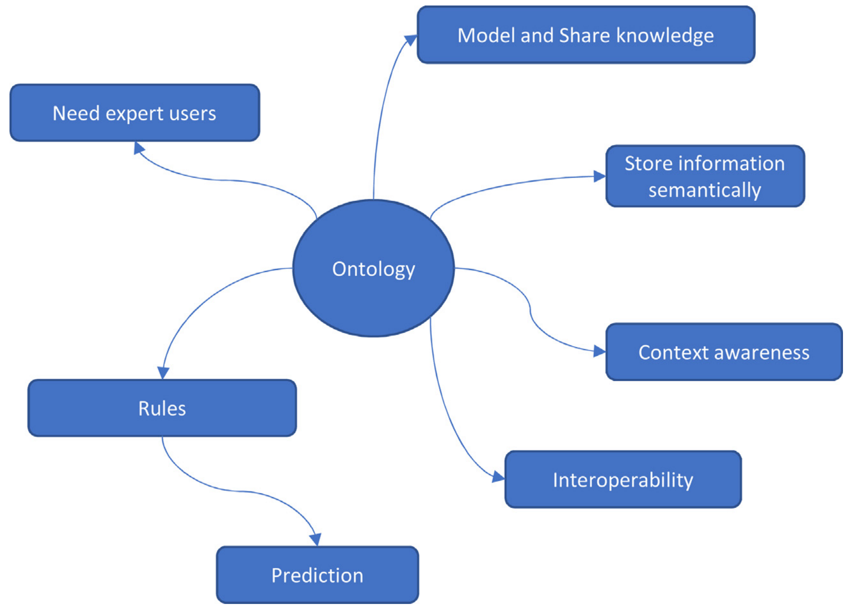

Fig. 10. Taxonomy of ontologies applications in predictive maintenance.

crucial information to set the parameters of the models [66]. In the context of predictive maintenance, this need is discussed by Nunez et al. [41]. In their solution, the authors use expert knowledge to formalize an ontology to perform vibration analysis of machine components. The authors use information stored in the ontology to create SWRL rules for failure prediction and to determine the cause of the potential failure.

6. Conclusion

The development of this systematic literature review aimed to discuss the main issues related to machine learning and reasoning for predictive maintenance in the context of Industry 4.0. We discussed the concepts and technologies applied in this area. We also presented the challenges faced in its application in the real world. This review focused on identifying architectures or frameworks that use reasoning based on ML models or through the adoption of ontologies. The study was limited to predictive maintenance of cyber-physical systems, not including related works that apply predictive maintenance in other contexts, such as predicting software failures.

Three research questions were defined to guide this systematic literature review. The answer to these questions showed that the need for data integration across the company is a topic of interest because it impacts on the overall business performance. Collecting data from a piece of equipment and giving contextual information, semantically improving that data, and providing meaning through the use of information from various sources received attention too. Moreover, using formal methods, although not yet deeply investigated, is also an important matter. For instance, the use of ontologies in the context of predictive maintenance appears as a tool applied for data standardization, aiding in the interoperability of systems, and consequently collaborating with the integration of the company's information as a whole.

ML models applied to PdM also received attention from the scientific community recently. In this sense, we identified that there is a large number of different models being proposed and applied in this field. Nevertheless, the results of the systematic literature review showed that no algorithm is capable of dealing with all existing scenarios in a company. As discussed in Section 2, the maintenance process involves several steps. Among these steps, we can mention the data collection to perform the detection of anomalies, diagnosis, and fault isolation, and the estimation of remain useful life. The construction of ML models in this first stage uses data coming mainly from equipment and machines in the physical world. While anomaly detection and fault isolations are general classifications or clustering problems and prognostics is a regression-related problem. In the context of prognostic techniques, we can classify them in three categories: (i) similarity-based, (ii) extrapolation based, and (iii) model identification and estimation based.

This article showed that predictive maintenance is a hot topic in the context of Industry 4.0. We conclude that because the papers that bring novelty to the field are concentrated in the years 2018 and 2019. Many relevant works are available so far. However, there is still room to deal with several challenges in this field. Taking into account the achieved results, we envision the necessity of implementing the theoretical frameworks found in the literature in real industrial environments. This implementation would allow a more precise evaluation of their effectiveness through metrics such as cost reduction and time spent on the maintenance task. Challenges related to big data are also of interest because predictive maintenance is a field that relies on large amounts of data. Therefore issues like scalability, latency, and data security deserve further investigation.

Declaration of Competing Interest

The authors report no declarations of interest.

References

Kunst, R., Avila, L., Binotto, A., Pignaton, E., Bampi, S., Rochol, J., 2019. Improving devices communication in industry 4.0 wireless networks. Eng. Appl. Artif. Intell. 83, 1–12 http://www.sciencedirect.com/science/article/pii/S0952197619300995.

Lee, J., Kao, H.-A., Yang, S., 2014. Service innovation and smart analytics for industry 4.0 and big data environment. Proc. CIRP 16, 3–8.

Cachada, A., Barbosa, J., Leitño, P., Gcraldcs, C.A., Deusdado, L., Costa, J., Teixeira, C., Teixeira, J., Moreira, A.H., Moreira, P.M., et al., 2018. Maintenance 4.0: intelligent and predictive maintenance system architecture. 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), vol. 1, 139–146.

Ali, M.I., Patel, P., Breslin, J.G., 2019. Middleware for real-time event detection and predictive analytics in smart manufacturing. 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), 370–376.

Romeo, L., Loncarski, J., Paolanti, M., Bocchini, G., Mancini, A., Frontoni, E., 2020. Machine learning-based design support system for the prediction of heterogeneous machine parameters in industry 4.0. Expert Syst. Appl. 140, 112869.

O'Donovan, P., Gallagher, C., Leahy, K., O'Sullivan, D.T., 2019. A comparison of fog and cloud computing cyber-physical interfaces for industry 4.0 real-time embedded machine learning engineering applications. Comput. Ind. 110, 12–35 http://www.sciencedirect.com/science/article/pii/S016636151830366X.

Boyes, H., Hallaq, B., Cunningham, J., Watson, T., 2018. The industrial internet of things (iiot): an analysis framework. Comput. Ind. 101, 1–12 http://www.sciencedirect.com/science/article/pii/S0166361517307285.

Carbery, C.M., Woods, R., Marshall, A.H., 2018. A bayesian network based learning system for modelling faults in large-scale manufacturing. 2018 IEEE International Conference on Industrial Technology (ICIT), 1357–1362.

D. O. Chukwuekwe, T. Glesnes, P. Schjølberg, Condition monitoring for predictive maintenance-towards systems prognosis within the industrial internet of things.

Wang, K., Wang, Y., 2017. How ai affects the future predictive maintenance: a primer of deep learning. In: International Workshop of Advanced Manufacturing and Automation, Springer, pp. 1–9.

Balogh, Z., Gatial, E., Barbosa, J., Leitão, P., Matejka, T., 2018. Reference architecture for a collaborative predictive platform for smart maintenance in manufacturing. In: 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), IEEE, pp. 000299–000304.

Schmidt, B., Wang, L., 2018. Predictive maintenance of machine tool linear axes: a case from manufacturing industry. Proc. Manuf. 17, 118–125.

ADHIKARI, P., RAO, H.G., BUDERATH, D.-I.M., 2018. Machine Learning Based Data Driven Diagnostics & Prognostics Framework for Aircraft Predictive Maintenance.

Zhou, C., Tham, C.-K., 2018. Graphel: a graph-based ensemble learning method for distributed diagnostics and prognostics in the industrial internet of things. In: 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), IEEE, pp. 903–909.

Bousdekis, A., Mentzas, G., Hribernik, K., Lewandowski, M., von Stietencron, M., Thoben, K.-D., 2019. A unified architecture for proactive maintenance in manufacturing enterprises. In: Enterprise Interoperability VIII, Springer, pp. 307–317.

Ansari, F., Glawar, R., Sihn, W., 2020. Prescriptive maintenance of cpps by integrating multimodal data with dynamic bayesian networks. In: Machine Learning for Cyber Physical Systems, Springer, pp. 1–8.

Sarazin, A., Truptil, S., Montarnal, A., Lamothe, J., 2019. Toward information system architecture to support predictive maintenance approach. In: Enterprise Interoperability VIII, Springer, pp. 297–306.

Hegedüs, C., Varga, P., Moldován, I., 2018. The mantis architecture for proactive maintenance. In: 2018 5th International Conference on Control, Decision and Information Technologies (CoDIT), IEEE, pp. 719–724.

Ferreira, L.L., Albano, M., Silva, J., Martinho, D., Marreiros, G., Di Orio, G., Maló, P., Ferreira, H., 2017. A pilot for proactive maintenance in industry 4.0. In: 2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS), IEEE, pp. 1–9.

Kiangala, K.S., Wang, Z., 2018. Initiating predictive maintenance for a conveyor motor in a bottling plant using industry 4.0 concepts. Int. J. Adv. Manuf. Technol. 97 (9–12), 3251–3271.

Kaur, K., Selway, M., Grossmann, G., Stumptner, M., Johnston, A., 2018. Towards an open-standards based framework for achieving condition-based predictive maintenance. In: Proceedings of the 8th International Conference on the Internet of Things, ACM, p. 16.

Liu, Z., Jin, C., Jin, W., Lee, J., Zhang, Z., Peng, C., Xu, G., 2018. Industrial ai enabled prognostics for high-speed railway systems. In: 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), IEEE, pp. 1–8.

Crespo Márquez, A., de la Fuente Carmona, A., Antomarioni, S., 2019. A process to implement an artificial neural network and association rules techniques to improve asset performance and energy efficiency. Energies 12 (18), 3454.

Xu, Y., Sun, Y., Liu, X., Zheng, Y., 2019. A digital-twin-assisted fault diagnosis using deep transfer learning. IEEE Access 7, 19990–19999.

da Cunha Mattos, T., Santoro, F.M., Revoredo, K., Nunes, V.T., 2014. A formal representation for context-aware business processes. Comput. Ind. 65 (8), 1193–1214 http://www.sciencedirect.com/science/article/pii/S0166361514001407.

Schmidt, B., Wang, L., 2018. Cloud-enhanced predictive maintenance. Int. J. Adv. Manuf. Technol. 99 (1–4), 5–13.

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O.P., Turner, M., Niazi, M., Linkman, S., 2010. Systematic literature reviews in software engineering – a tertiary study. Inf. Softw. Technol. 52 (8), 792–805.

Kitchenham, B., 2004. Procedures for Performing Systematic Reviews, 33. Keele University, Keele, UK, pp. 1–26.

Stojanovic, L., Stojanovic, N., 2017. Premium: big data platform for enabling self-healing manufacturing. In: 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), IEEE, pp. 1501–1508.

Zenisek, J., Wolfartsberger, J., Sievi, C., Affenzeller, M., 2018. Streaming synthetic time series for simulated condition monitoring. IFAC-PapersOnLine 51 (11), 643–648.

Bumblauskas, D., Gemmill, D., Igou, A., Anzengruber, J., 2017. Smart maintenance decision support systems (smdss) based on corporate big data analytics. Expert Syst. Appl. 90, 303–317.

May, G., Kyriakoulis, N., Apostolou, K., Cho, S., Grevenitis, K., Kokkorikos, S., Milenkovic, J., Kiritsis, D., 2018. Predictive maintenance platform based on integrated strategies for increased operating life of factories. In: IFIP International Conference on Advances in Production Management Systems, Springer, pp. 279–287.

Golightly, D., Kefalidou, G., Sharples, S., 2018. A cross-sector analysis of human and organisational factors in the deployment of data-driven predictive maintenance. Inf. Syst. e-Bus. Manag. 16 (3), 627–648.

Syafrudin, M., Alfian, G., Fitriyani, N., Rhee, J., 2018. Performance analysis of iot-based sensor, big data processing, and machine learning model for real-time monitoring system in automotive manufacturing. Sensors 18 (9), 2946.

Zhang, X., Ming, X., Liu, Z., Yin, D., Chen, Z., Chang, Y., 2019. A reference framework and overall planning of industrial artificial intelligence (i-ai) for new application scenarios. Int. J. Adv. Manuf. Technol. 101 (9–12), 2367–2389.

Malek, M., 2017. Predictive analytics: a shortcut to dependable computing. In: International Workshop on Software Engineering for Resilient Systems, Springer, pp. 3–17.

Carbery, C.M., Woods, R., Marshall, A.H., 2018. A new data analytics framework emphasising pre-processing in learning ai models for complex manufacturing systems. In: Intelligent Computing and Internet of Things, Springer, pp. 169–179.

Wan, J., Yang, J., Wang, Z., Hua, Q., 2018. Artificial intelligence for cloud-assisted smart factory. IEEE Access 6, 55419–55430.

Costa, R., Figueiras, P., Jardim-Gonçalves, R., Ramos-Filho, J., Lima, C., 2017. Semantic enrichment of product data supported by machine learning techniques. In: 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), IEEE, pp. 1472–1479.

Sala, D.A., Jalalvand, A., Van Yperen-De Deyne, A., Mannens, E., 2018. Multivariate time series for data-driven endpoint prediction in the basic oxygen furnace. In: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, pp. 1419–1426.

Nuñez, D.L., Borsato, M., 2018. Ontoprog: an ontology-based model for implementing prognostics health management in mechanical machines. Adv. Eng. Inform. 38, 746–759.

Schmidt, B., Wang, L., Galar, D., 2016. Semantic framework for predictive maintenance in a cloud environment. In: 10th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME'16, Ischia, Italy, 20–22 July 2016, vol. 62, Elsevier, pp. 583–588.

Q. Cao, A. Samet, C. Zanni-Merk, F. d. B. de Beuvron, C. Reich, Combining chronicle mining and semantics for predictive maintenance in manufacturing processes.

Ansari, F., Glawar, R., Nemeth, T., 2019. Prima: a prescriptive maintenance model for cyber-physical production systems. Int. J. Comput. Integr. Manuf., 1–22.

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. Autom. Constr. 112, 103087.

Calabrese, M., Cimmino, M., Fiume, F., Manfrin, M., Romeo, L., Ceccacci, S., Paolanti, M., Toscano, G., Ciandrini, G., Carrotta, A., et al., 2020. Sophia: an event-based iot and machine learning architecture for predictive maintenance in industry 4.0. Information 11 (4), 202.

Daniyan, I., Mpofu, K., Oyesola, M., Ramatsetse, B., Adeodu, A., 2020. Artificial intelligence for predictive maintenance in the railcar learning factories. Proc. Manuf. 45, 13–18.

Hoffmann, M.W., Wildermuth, S., Gitzel, R., Boyaci, A., Gebhardt, J., Kaul, H., Amihai, I., Forg, B., Suriyah, M., Leibfried, T., et al., 2020. Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. Sensors 20 (7), 2099.

De Vita, F., Bruneo, D., Das, S.K., 2020. A novel data collection framework for telemetry and anomaly detection in industrial iot systems. In: 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), IEEE, pp. 245–251.

Chen, G., Wang, P., Feng, B., Li, Y., Liu, D., 2020. The framework design of smart factory in discrete manufacturing industry based on cyber-physical system. Int. J. Comput. Integr. Manuf. 33 (1), 79–101.

Huang, M., Liu, Z., Tao, Y., 2019. Mechanical fault diagnosis and prediction in iot based on multi-source sensing data fusion. Simul. Model. Pract. Theory, 101981.

Rivas, A., Fraile, J.M., Chamoso, P., González-Briones, A., Sittón, I., Corchado, J.M., 2019. A predictive maintenance model using recurrent neural networks. In: International Workshop on Soft Computing Models in Industrial and Environmental Applications, Springer, pp. 261–270.

Cerquitelli, T., Bowden, D., Marguglio, A., Morabito, L., Napione, C., Panicucci, S., Nikolakis, N., Makris, S., Coppo, G., Andolina, S., et al., 2019. A fog computing approach for predictive maintenance. In: International Conference on Advanced Information Systems Engineering, Springer, pp. 139–147.

Ding, K., Shi, H., Hui, J., Liu, Y., Zhu, B., Zhang, F., Cao, W., 2018. Smart steel bridge construction enabled by bim and internet of things in industry 4.0: a framework. In: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), IEEE, pp. 1–5.

Yuan, Y., Ma, G., Cheng, C., Zhou, B., Zhao, H., Zhang, H.-T., Ding, H., 2018. Artificial Intelligent Diagnosis and Monitoring in Manufacturing. arXiv preprint arXiv:1901.02057.

Strauß, P., Schmitz, M., Wöstmann, R., Deuse, J., 2018. Enabling of predictive maintenance in the brownfield through low-cost sensors, an iiot-architecture and machine learning. In: 2018 IEEE International Conference on Big Data (Big Data), IEEE, pp. 1474–1483.

Peres, R.S., Rocha, A.D., Leitao, P., Barata, J., 2018. Idarts-towards intelligent data analysis and real-time supervision for industry 4.0. Comput. Ind. 101, 138–146.

Diez-Olivan, A., Pagan, J.A., Sanz, R., Sierra, B., 2017. Data-driven prognostics using a combination of constrained k-means clustering, fuzzy modeling and lof-based score. Neurocomputing 241, 97–107.

Li, Z., Wang, Y., Wang, K.-S., 2017. Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario. Adv. Manuf. 5 (4), 377–387.

Gatica, C.P., Koester, M., Gaukstern, T., Berlin, E., Meyer, M., 2016. An industrial analytics approach to predictive maintenance for machinery applications. In: 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, pp. 1–4.

Glawar, R., Ansari, F., Kardos, C., Matyas, K., Sihn, W., 2019. Conceptual design of an integrated autonomous production control model in association with a prescriptive maintenance model (prima). Proc. CIRP 80, 482–487.

Talamo, C., Paganin, G., Rota, F., 2019. Industry 4.0 for failure information management within proactive maintenance. In: IOP Conference Series: Earth and Environmental Science, vol. 296, IOP Publishing, p. 012055.

Issam, M., El Majd, B.A., El GHAZI, H., 2018. A new architecture of collaborative vehicles for monitoring fleet health in real-time. In: 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), IEEE, pp. 309–313.

Gao, R., Wang, L., Teti, R., Dornfeld, D., Kumara, S., Mori, M., Helu, M., 2015. Cloud-enabled prognosis for manufacturing. CIRP Ann. 64 (2), 749–772.

Selcuk, S., 2017. Predictive maintenance, its implementation and latest trends. Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf. 231 (9), 1670–1679.

Olivares-Alarcos, A., Beßler, D., Khamis, A., Goncalves, P., Habib, M.K., Bermejo-Alonso, J., Barreto, M., Diab, M., Rosell, J., Quintas, J., Pignaton, E., Olszewska, J., et al., 2019. A review and comparison of ontology-based approaches to robot autonomy. Knowl. Eng. Rev. 34, e29, http://dx.doi.org/10.1017/S0269888919000237.

Rafael Kunst is a professor of Computer Science at the Applied Computing Graduate Program of University of Rio dos Sinos Valley – Universidade do Vale do Rio dos Sinos (UNISINOS), in Brazil, where he is also a member of the Software Innovation Laboratory – SOFTWARELAB. His research interests involve Computer Networks, Internet of Things, and Machine Learning. He holds a PhD and a M.Sc. degree in Computer Science, both received from the Federal University of Rio Grande do Sul (UFRGS).