Journal of Manufacturing Systems 63 (2022) 550–562

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

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

Review

Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods

Carlos Ferreira [a,*], Gil Gonçalves [b]

[a] University of Porto, Faculty of Engineering, Department Industrial Engineering and Management, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
[b] University of Porto, Faculty of Engineering, Department of Informatic Engineering, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal

ARTICLE INFO

Keywords:
Remaining Useful Life
Artificial intelligence
Data-driven models
Machine learning
Predictive maintenance
Prediction Process Framework

ABSTRACT

Approaches such as Cyber-Physical Systems (CPS), Internet of Things (IoT), Internet of Services (IoS), and Data Analytics have built a new paradigm called Industry 4.0. It has improved manufacturing efficiency and helped industries to face economic, social, and environmental challenges successfully. Condition-Based Maintenance (CBM) performs machines and components' maintenance routines based on their needs, and Prognostics and Health Management (PHM) monitors components' wear evolution using indicators. PHM is a proactive way of implementing CBM by predicting the Remaining Useful Life (RUL), one of the most important indicators to detect a component's failure before it effectively occurs. RUL can be predicted by historical data or direct data extraction by adopting model-based, data-driven, or hybrid methodologies. Model-based methods are challenging, expensive, and time-consuming to develop in complex equipment due to the need for a lot of prior system knowledge. Data-driven methods have primarily used Machine Learning (ML) approaches. They require little historical data, are less complex and expensive, and are more applicable, providing a trade-off between complexity, cost, precision, and applicability. However, despite the increased use of data-driven methods, several studies have pointed out different challenges to RUL prediction. Some works have proposed solutions from individuals and unconnected work structures to overcome these challenges, and there is still a lack of an explicit framework for general process analysis. Moreover, none of them have correlated the different challenges with each micro-step of the RUL prediction process. This work describes the structures, systems and components approached, and datasets used. Next, it proposes a compact framework for the RUL prediction process. Also, it maps the main challenges of this process and the advantages and drawbacks of the most relevant ML methods. Further, it discusses the operational datasets, the accuracy concern, the use of file log systems in the RUL prediction, and approaches the ML Interpretability (MLI) issue. Finally, it concludes with some future research directions.

1. Introduction

Several approaches such as Cyber-Physical Systems (CPS), Internet of Things (IoT), Internet of Services (IoS), and Data Analytics have built a new paradigm called *Industry 4.0*, which has improved the manufacturing efficiency, and helped industries to successful face economic, social and environmental challenges [1]. The Internet of Things (IoT) connects several devices through the internet (wireless and wired), allowing the collection of a large amount of data from machines and components. On the other hand, Data Analytics enables the processing of this large amount of data and efficiently mining insightful knowledge [2]. In this scenario, this increased availability of data and processing tools has favored, leveraged, and more easily applied approaches such as Zero-Defect Manufacturing (ZDM) [3] and Condition-Based Maintenance (CBM) [4]. The ZDM refers to ensuring process and product quality by reducing defects through corrective, preventive, and predictive maintenance techniques under the Industry 4.0 paradigm to improve manufacturing sustainability [5].

The CBM refers to performing systems' maintenance and repair routines (machines and components) based on their needs. CBM allows to schedule these routines in advance, generating a more effective maintenance and logistics plan, improving safety by preventing performance degradation and failures, creating higher operation reliability, increasing production, and minimizing costs due to the reduced

* Corresponding author.
E-mail addresses: Ferreira.carlos@fe.up.pt (C. Ferreira), gil@fe.up.pt (G. Gonçalves).

https://doi.org/10.1016/j.jmsy.2022.05.010
Received 31 March 2022; Received in revised form 15 May 2022; Accepted 16 May 2022
Available online 23 May 2022

downtime [6–9], and even avoiding catastrophic situations [10]. Prognostics and Health Management (PHM), which refers to monitoring component wear evolution (not only the defects) through detection, diagnosis, and prognostics [11], is a proactive way of implementing CBM [6] by predicting the Remaining Useful Life (RUL) from historical and actual data operating conditions [8].

The RUL is currently one of the most important parameters to predict the failure of a component before it effectively occurs [10,12]. It is predicted by detecting the components' conditions through data gathering (e.g., vibration signals) from monitoring sensors [13] and involves adopting model-based (or physics-based), data-driven, or hybrid methodologies [14]. The predicted RUL is expressed according to the primary system measurement, for instance, kilometers for automobiles and cycles for commercial aircraft [15]. CoC is referida a turbofan engines

Model-based methods consist of developing mathematical or physical models based on historical data to determine trends in the health state of a component [10] and may have statistics and computational intelligence approaches [14]. Also, they can be split into micro-level models (or damage propagation models) and macro-level models, which represent a system in a simplified way, defining the relations between the various types of variables (input, state, and output) of that system [16]. On the other hand, data-driven methods mainly consist of developing a model to the historical data and then determining patterns in the health state of a component [10]. They include but are not limited to Machine Learning (ML) approaches [14]. Besides that, data-driven methods can predict the RUL indirectly, estimating the degradation until the failure of the component (degradation-based) and directly from condition monitoring of the equipment's signals [17]. In general, there are three steps related to data-driven methods: 1 data acquisition, 2 degradation or health indicator construction, and health prognosis or 3 RUL prediction [10,18]. However, more complex approaches can be found in the literature [15].

Due to its nature, the model-based methods are challenging [17], expensive, and time-consuming to develop to complex equipment [19] due to the need for a lot of prior system knowledge [10]. Hence, the necessary model's assumptions and simplifications can limit this approach [20], despite the high precision prognostic they can provide [21]. On the other hand, although data-driven methods require large amounts of historical data to model development [22], they are less complex (and precise) and more applicable [21]. Further, they are less dependent on users' expertise to gather and select these data than model-based methods [10], providing, in this sense, a trade-off between complexity, cost, precision, and applicability [21]. On these grounds, and also due to newer technologies regarding sensors and big data [23], leading to advantages related to condition monitoring measurement [24], and the increased flexibility and intelligence, the use of data-driven methods has increased compared to model-based methods in terms of RUL prediction [25].

However, despite the increased use of data-driven models, some challenges need to be overcome when using these models for RUL prediction in PHM. In this context, some works have proposed solutions from individuals and unconnected work structures to overcome these challenges, and there is still a lack of an explicit framework for general process analysis. Moreover, none of them have correlated the different challenges with the steps of the RUL prediction process [26]. Therefore, this work aims to extensively survey the literature to achieve four objectives concerning applying AI-based data-driven models, namely ML approaches, in RUL prediction. First, understanding what Structure, Systems, or Components (SSC) and datasets have been usually used in this research area. Then, describe, in a general way, the macro and micro-steps of the RUL prediction process. Next, describe the main challenges of applying ML methods and correlate them with the macro-steps in the RUL prediction process and the methods and tools proposed to fix them. Finally, capturing advantages and drawbacks of the most relevant ML methods identified in this work.

The remainder of this paper is as follows. Section 2 describes the data gathering by using the PRISMA methodology. Section 3 describes the SSC approach and datasets used. Next, it presents a framework for the RUL prediction process. Also, a mapping of the main challenges of this process and the advantages and drawbacks of the most relevant ML methods is detailed. Section 4 discusses operational datasets in RUL prediction, the weight of RUL prediction concern, its apparent effect in the ensembling of methods, system log files as a data source to predict RUL, and MLI. Finally, Section 5 formally concludes this work and gives directions for future works.

Table 1
Search results.

Search local	Search results	Range	Types of document
ACM	11	2016–2021	• 2 Proceedings • 1 Book • 2 Chapters • 6 Conference Papers
DBLP	29	2008–2021	• 17 Journal Articles • 12 Conference and Workshops Papers
IEEE Xplore	92	2009–2021	• 58 Conference Papers • 33 Journal Articles • 1 Early Access Article
INSPEC	138	2012–2021	• 73 Journal Articles • 64 Conference Articles • 1 Dissertation
ScienceDirect	401	1988–2022	• 64 Review Articles • 307 Research Articles • Encyclopedia • 12 Book Chapters • Conference Abstracts • 2 Editorials • 1 Mini Review • 2 Short Communication • 7 Other
SCOPUS	35	2009–2021	• 18 Conference Papers • 16 Journal Articles • 1 Review
Web of Science	44	2012–2021	• 30 Journal Articles • 12 Proceedings Papers • 2 Reviews

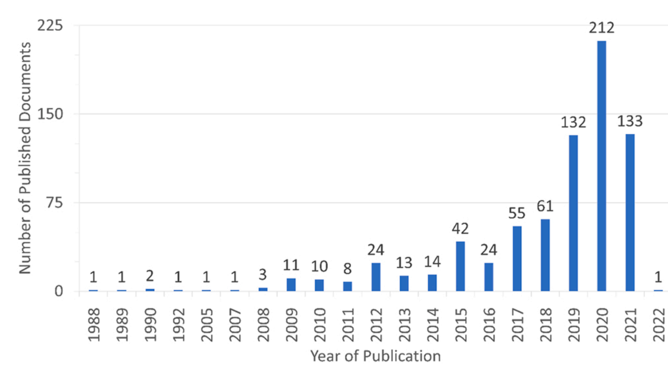

Fig. 1. Published documents per year.

2. Methodology

This study's primary goal is to exploit the main challenges regarding applying ML methods to the RUL prediction and how these challenges have been overcome. Our research questions match the four objectives described in the previous section. To achieve this goal, we carried out a systematic Literature Review (LR) following the four steps described in the PRISMA guidelines [27], which are summarized below.

First, we identified a set of potential documents through DBLP, ACM, IEEE Xplore, INSPEC, ScienceDirect, SCOPUS, and Web of Science. To do this, we have used the following base search expression: ((diagnostics

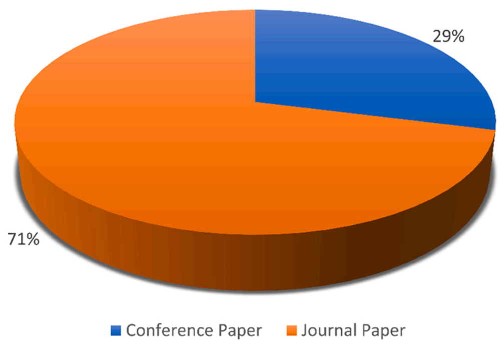

Fig. 2. Types of papers.

Table 2
The most influencing source of papers.

Journal	# of papers
IEEE Access	6
Measurement	6
Reliability Engineering and System Safety	6
Journal of Manufacturing Systems	5
IEEE International Conference on Prognostics and Systems Health Management (PHM)	5
Engineering Applications of Artificial Intelligence	4
Mechanical Systems and Signal Processing	4
IEEE International Conference on Prognostics and Health Management (ICPHM)	4
Applied Sciences	3
IEEE Transactions on Industrial Electronics	3
Neurocomputing	3
Advanced Engineering Informatics	2
Applied Soft Computing Journal	2
IEEE Transactions on Cybernetics	2
IEEE Transactions on Industrial Informatics	2
IEEE Transactions on Instrumentation and Measurement	2
IEEE/ASME Transactions on Mechatronics	2
International Journal of Prognostics and Health Management	2
ISA Transactions	2
IEEE Data Driven Control and Learning System Conference	2

OR prognostics OR "failure detection" OR "condition monitoring" OR "predictive maintenance" OR "condition-based maintenance") AND (("Remaining Useful Life" OR RUL) AND ("Data-driven" AND "artificial intelligence"))). Depending on the local search, some changes were made in the above expression to comply with specific rules. The search was performed between July and August 2021. In all, 750 documents were found with a range time from 1998 to 2022. Table 1 presents the search results, and Fig. 1 shows the number of published documents per year.

Then, after removing the duplicates, we screened the title and abstract of 562 documents to select a more specific group related to AI data-driven models applied to RUL prediction. The inclusion criterion was the existence of one or more keywords used in the search expression in the document's title. In this step, 302 documents were excluded, and the remaining 260 documents have moved forward to the next step. Finally, we analyzed those 260 documents in full text. The inclusion criterion in this step was the existence of a detailed description of ML methods application to RUL prediction. We chose a group of 107 journal and conference papers from 60 different sources, with index SJC 2020 varying from 0.18 to 3.11, to include at the end of this step since they matched the established criteria. Fig. 2 presents the share of each type of paper. Table 2 presents the most influencing sources of the included papers (only journals with more than one paper). Fig. 3 presents our search's complete PRISMA flow diagram.

3. Results

In this section, we intend to describe the findings of our work. We divided these findings into four different pillars such that we can explain them in a better way. First, we have mapped the different SSC and datasets used in RUL prediction when ML methods have been applied (Section 3.1). Second, we have found different processes to achieve RUL prediction, even in treating the same problem or using the same dataset. Based on our findings, we proposed and described a compact framework for this process to standardize all the critical and necessary steps (Section 3.2). Third, we have mapped the RUL prediction process challenges using ML methods (Section 3.3). The mapping of these challenges allowed us to systematize the approaches described in the literature to overcome each one of these challenges (Appendix A). Fourth, we have pointed advantages and drawbacks of some ML approaches (Section 3.4). All these results are described in detail in the following subsections.

3.1. *The SSC approached and used datasets*

The studies evaluated have described 25 different SSC and 15 different datasets used for RUL prediction. Regarding the systems, we can highlight Rolling Elements (37,07%) and Aircraft Turbofan Engines (29,13%), which accounted for almost 70% of all studies evaluated. Rolling elements are key parts of different industrial machines, and their integrity is extremely related to the machine's safety [10], often causing operating failures [28]. For example, they account for more than 50% of induction motors' failures [29], resulting in unexpected stops, increasing maintenance costs, and poor productivity [13]. On aircraft, attempted maintenance is essential to ensure operation safety [12], besides increasing economic efficiency [30]. According to the International Air Transport Association (IATA), maintenance costs of the major aviation companies reached $15.57 billion between 2012 and 2016, which represented a growth of 3% [31]. Turbofan engines, specifically, are responsible for about 30% of the failures in an aircraft, and in great-proportion accidents, these systems have been the root cause in 40% of the cases. Besides, the maintenance cost of propulsion devices shares about 40% of the whole aircraft maintenance costs [32]. Fig. 4 presents all SSC used for RUL prediction and the share of included papers.

Regarding the datasets, we can highlight the share of datasets gathered from the Prognostics Center of Excellence – PCoE, from the National Aeronautics and Aerospace Administration (NASA) [33], including turbofan engines, rolling bearings, and batteries; datasets gathered through the use of the PRONOSTIA platform, specific for rolling bearings, and operating datasets, which included different systems and components. They accounted for about 36%, 25%, and 24% of the share of studies evaluated. Fig. 5 presents all datasets used for RUL prediction and the share of included papers. Fig. 6 presents operating datasets used for RUL prediction and the share of included papers.

3.2. *RUL prediction process*

We have found different processes to achieve RUL prediction, even in treating the same problem or using the same dataset. About 50% of the studies analyzed described the process as a "step-by-step" in applying the methodology to predict RUL. Nevertheless, a few of them could generalize the process of RUL prediction, and none of them correlated the steps of the process with the challenges involved. In this sense, we adapted the framework for PHM proposed in [34] by adding and removing some micro-steps described in that work. After we have clustered the micro-steps into four different macro-steps, such that we can correlate each of these macro-steps with the challenges inherent to the steps inside them. In this subsection, we will describe the proposed RUL prediction process, and in the following subsection, we will describe the challenges and correlate them with each cluster. The proposed RUL prediction process comprises ten micro-steps, divided into

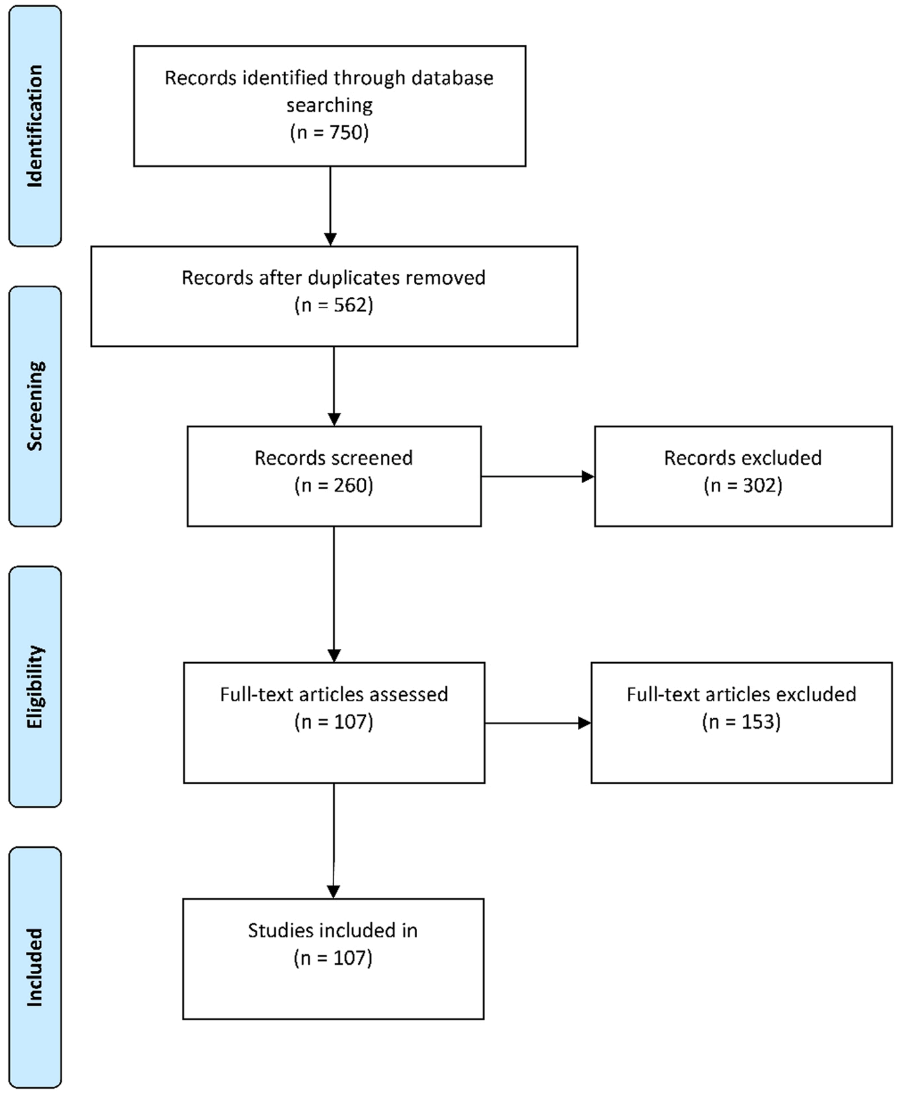

Fig. 3. Search PRISMA flow.

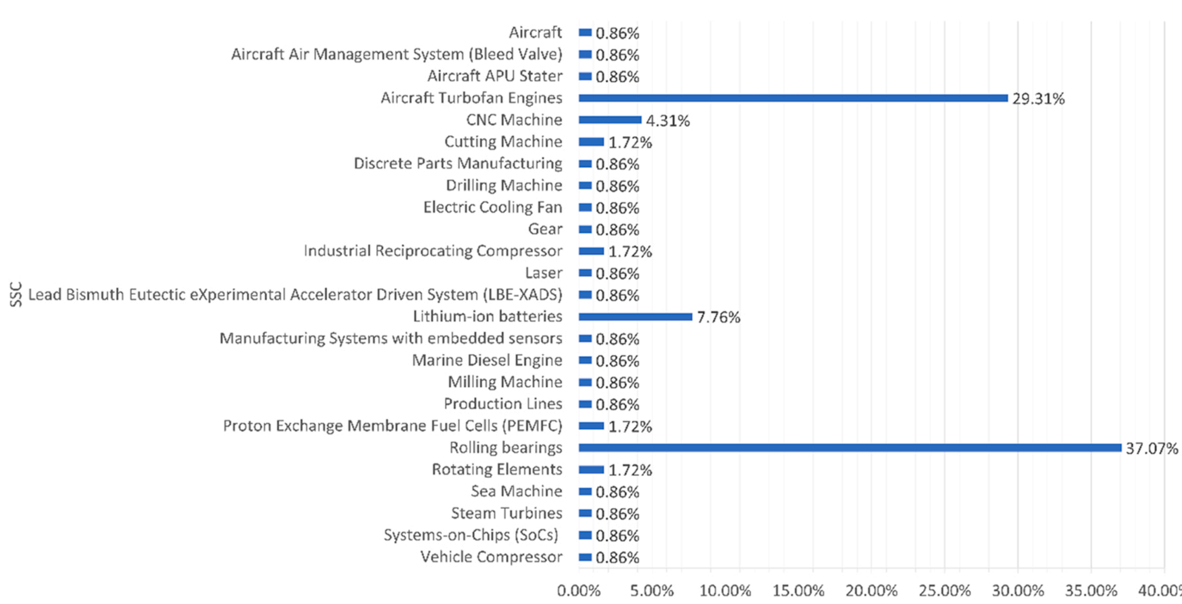

Fig. 4. SSC used for RUL prediction.

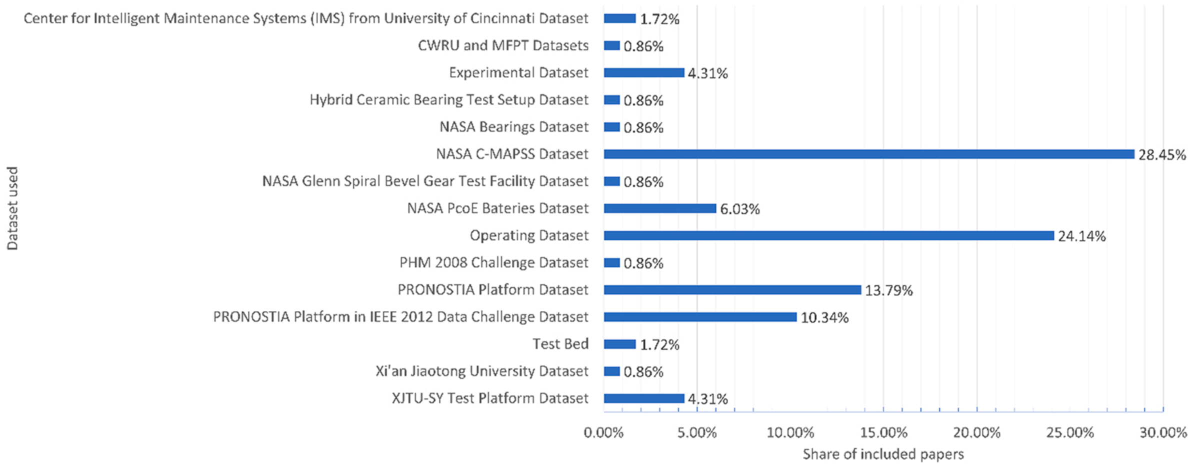

Fig. 5. Datasets used for RUL prediction.

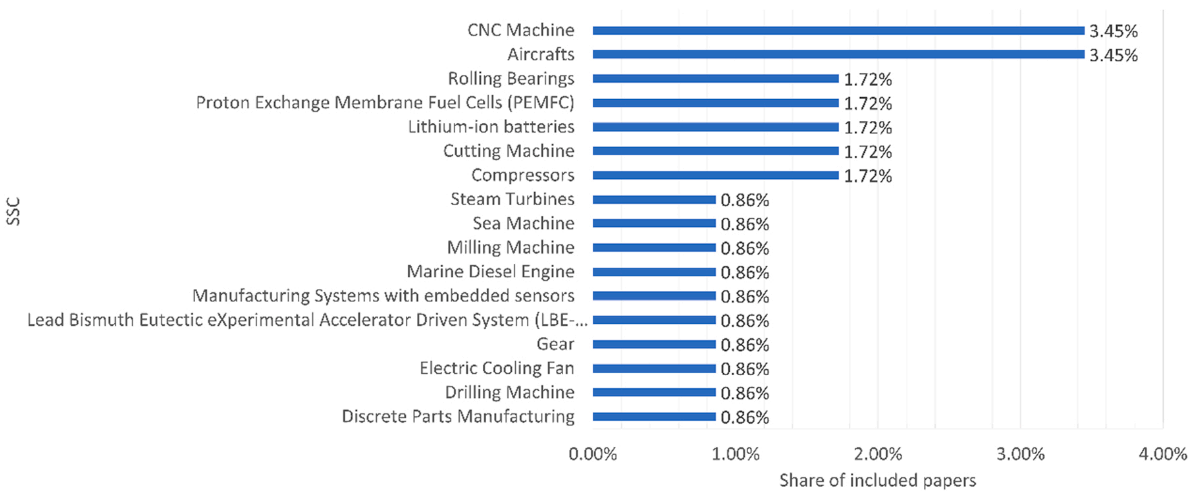

Fig. 6. Operating datasets used for RUL prediction.

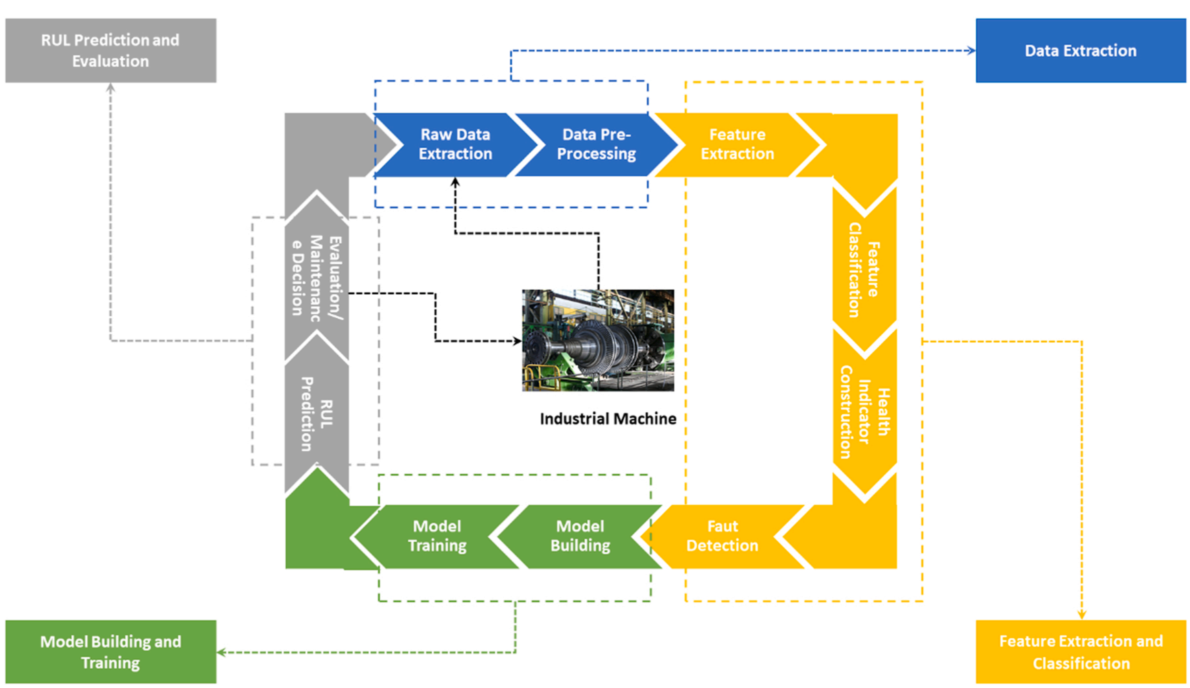

Fig. 7. RUL prediction process.

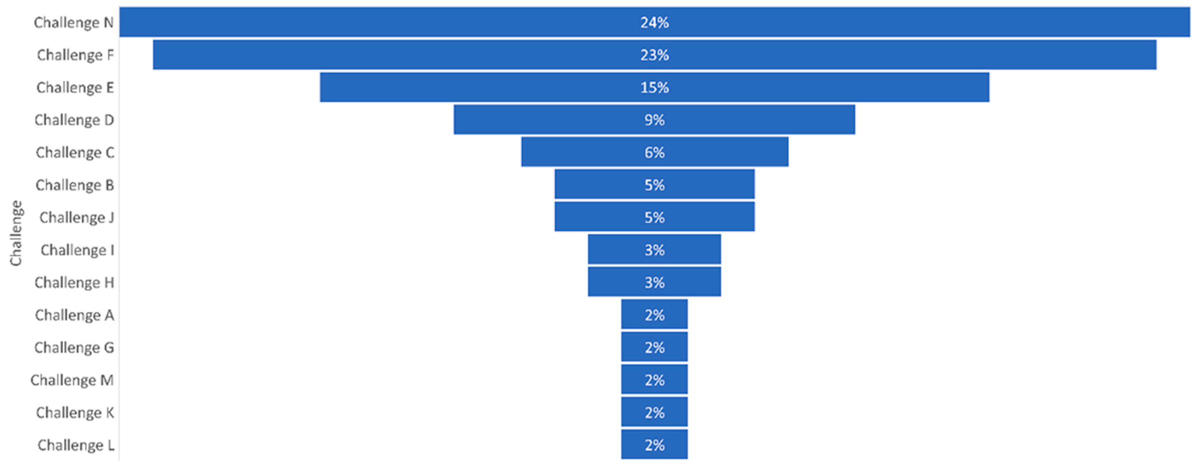

Fig. 8. Share of each challenge in the studies analyzed.

four macro-steps. Fig. 7 presents an overview of the entire process.

3.2.1. Data Extraction

a) *Raw Data Extraction* – Consists in acquiring different types of data (e.g., vibrations or acoustic signals) through the various sensors installed in the System [35].
b) *Data Pre-Processing* – Consists in exploiting the features associated with the degradation of the SSC, which are hidden in the raw signal collected through the sensors [36].

3.2.2. Feature Extraction and Classification

a) *Feature Extraction* – Consists in transforming vibration data (raw data) collected from running SSC into relevant information about the running status of that SSC [37].
b) *Feature Classification* – Consists of determining which extracted features are more sensitive to detecting the degradation path in SSC [37].
c) *Health Indicator Construction* – Consists, in general, of fusing multi-dimensional features extracted from the time domain, frequency domain, and time-frequency domain into only one health indicator through dimension reduction approaches [10].
d) *Fault Detection* – Consists of fault location determination, fault type identification, and fault degree estimation [6].

3.2.3. Model Building and Training

a) *Model Building* – Consists of defining which type of model (or algorithm) to use to predict RUL (e.g., Random Forest, Convolutional Neural Networks). Besides, selected features are divided into two sub-datasets (training and testing) used to train the model and predict the RUL [38].
b) *Model Training* – Consists in creating a relationship between the real health state of the SSC and the measured condition monitoring by using the selected features – training dataset [39].

3.2.4. RUL Prediction and Evaluation

a) *RUL Prediction* – Consists in determining the time to failure before the failure effectively occurs [10] using the selected features, which the model has never seen before – testing dataset.
b) *Evaluation/Maintenance Decision* – This evaluates the prediction result based on specific metrics (e.g., RMSE, MAE, Score Function) and then decides on the applied maintenance approach [34].

3.3. The challenges involving the RUL prediction process by ML methods

Of the papers analyzed, 66 (60,00%) focused on overcoming one or more challenges in applying ML methods for RUL prediction. The remaining 42 studies (40,00%) focused on testing some new approach or did not explicitly reveal the challenge they tried to overcome. By analyzing those 66 works carefully, we could identify 14 different challenges throughout the four macro-steps of the RUL prediction process described in the previous sub-section. Furthermore, we also could identify some approaches used to overcome these challenges. The challenges are described as follows, and Fig. 8 presents the share of each challenge in the studies analyzed. Table A.1 shows some approaches described in the literature to overcome these challenges.

3.3.1. Data Extraction

- Data extraction does not provide a meaningful representation of time-varying sensor raw measurements – Challenge A.

3.3.2. Feature Extraction and Classification

- Models assume that data used in training and testing phases are drawn from the same distribution (similar conditions) – Challenge B.
- Difficult to deal with data noise, high level of uncertainty, the variance of local features, data sparsity, and censored data – Challenge C.
- Health Indicator Construction Complexity – Challenge D.
- Deal with multiple operating conditions and diversified degradation patterns – Challenge E.
- Feature selection and efficient feature comprehension relies heavily on labor and requires extensive domain knowledge (prior knowledge of the entire system) – Challenge F.

3.3.3. Model building and training

- The correlations of different sensor data are not explicitly considered in representation learning – Challenge G.
- Model Uncertainty – Challenge H.
- Deal with chronological order and temporal correlation of the health monitoring data (time-varying dynamics) – Challenge I.
- Access to labeled failure data is scarce due to the rarity of failures – Challenge J.

3.3.4. RUL prediction and evaluation

- Prediction Uncertainty – Challenge K.

LSTM

- DNN models suffer from single-path and top-down propagation – Challenge L.
- Interpretability of the ML Results – Challenge M.
- RUL Prediction Accuracy – Challenge N.

3.4. *Advantages and drawbacks of some ML approaches*

Data-driven methods have great merits, such as high accuracy, fast response, and easy implementation. Nevertheless, these methods are data-dependent, which means they need large amounts of data with high quality to ensure those merits [40]. Besides, three critical issues regarding data-driven models are still open: (i) How to ensure that a model will face input variations concerning those learned? (ii) How to ensure that the learned model faces unknown data? (iii) How to ensure convergence of the algorithms? [41]. For example, traditional ML methods such as Kalman Filter (KF), Extreme Learning Machine (ELM), Support Vector Machine (SVM), Echo State Network (ESN), and Convolutional Neural Network (CNN) do not consider the relevance of time-series signals that reflect the micro-change of the health state [42]. Besides, they often rely on manually extracted features from raw sensor data followed by an estimation of the health indicator, degradation states, and prediction of RUL using failure threshold [43].

No Good

Through the analysis of Table A.1, it is possible to identify the prominent use of some methods such as Neural Networks (NN) and Deep Learning (DL) Methods, especially Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM), and some variations of them.

Neural network architectures allow the model to store inferred pieces of knowledge, which can be accessed and manipulated through an external memory [19]. Another advantage is the capacity of these architectures to deal with non-linear complexity by training multilayer neural networks, improving the prediction accuracy [44]. However, some drawbacks are also part of these architectures, such as the lack of ability to process the input data in different upgrade modes based on its importance degree [45] and the high computational costs to optimize the models' weights [46].

DL methods have been widely used to overcome the challenges treated in this work, as well as other ones, and improve the accuracy of RUL prediction [38,39,47]. They have emerged and achieved outstanding results in different areas, mainly due to their strong capacity to map the relationship between degradation path and measured data and ability to learn features representation automatically, such that it is not necessary to design features manually, eliminating the need for previous knowledge of the system [47]. Finally, DL methods have a high capacity to deal with complex data [48]. Nonetheless, the literature reports some drawbacks, such as the data deficit issue, especially considering the varying operation conditions and the degradation mode of the components in practical industrial applications [49]. Also, the definition of DL in the literature is still not standardized, as described by Zhang et al. [50]. In this work, two DL methods and their variations were identified as the most prominent in RUL prediction: RNN and LSTM.

RNN and variants such as Echo State Network (ECN) have a strong capacity, due to their structure (weight sharing and feedback structure), for extracting and updating correlations of sequential data by augmenting recurrent links to hidden neurons [51] and capitalizing the historical information [12], besides having a great potential of modeling dynamic systems [52]. Notwithstanding some identified drawbacks concerns their training phase, which is very time-consuming and can restrict their applications in some practical industry fields [19]. Moreover, although RNN can exploit the sequential nature of the data, they cannot deal with long time series properly [53]. In addition, RNN cannot process which information is valuable to remember and which to forget [54]. For instance, applying traditional RNN makes it challenging to deal with long-term dependence in practice due to the problem of gradient vanishing and exploding [44].

LSTM corrects two drawbacks of the conventional RNN architecture by adding memory control, allowing correctly deciding when to select meaningful information for prediction and discard redundant information [54], and adequately dealing with the long-term dependency, eliminating the gradient problem vanish and exploding [44]. Besides, LSTM can fully use sensor sequence information [12]. Nevertheless, some drawbacks of applying LSTM are that it does not consider multi-sensor data [39], consumes substantial computation resources in actual deployment, and is not conducive to parallel computing [55].

4. Discussion

This section aims to provide a concise discussion centered on four topics we believe deserve to be explored from our findings. First, we will raise the issue concerning operational datasets (or at least the small number) used for RUL prediction. Then, we will approach the weight of RUL prediction concern and its apparent effect in the ensembling of methods. Next, we will focus on the few studies that have considered predicting RUL using system log files as a data source, highlighting this gap. Finally, we will describe some insights from the literature regarding MLI.

4.1. *The operational datasets for RUL prediction*

Since ML methods need a significant amount of data for training and predicting RUL, different types of datasets are expected to be available. These datasets are essential for testing the model (training/test phase) and validating the results against real situations. From our findings, we can observe from Figs. 4 and 5 that a significant variety of SSC and datasets were used to test RUL prediction models. However, only a tiny part of these datasets (24.14%) concerns operational datasets, which means data from machines running under real conditions. Indeed, we could capture that most of the datasets were retrieved from the different sources from NASA (36.02%) and the PRONOSTIA Platform (24.13%). Although these datasets are widely used and ensure the testability of the models, they are simulated datasets. Thus, they do not reproduce real operating conditions.

In this context, there are still some challenges for the companies, such as dealing with a massive amount of data produced by machines and production lines. Also, due to data privacy issues, many companies do not share their datasets publicly. The final effect is that researchers cannot effectively validate new models with more datasets [56].

Therefore, it is imperative to discuss closer relationships between universities and companies to mitigate this effect. The goal is that companies can trust the operational data of their machines to be used for a more consistent validation of the models developed by universities, always respecting ethics and data privacy and ensuring a better development of predictive maintenance methods and policies. Finally, the benefits of this integration would be felt by companies, universities, and customers, who would have higher quality products and services at lower prices.

4.2. *RUL prediction accuracy concern*

The importance of accurately predicting the RUL in different SSCs can reflect in several aspects, such as maintenance cost reduction by preventing unwanted failures and machine downtime and enhancing operational safety [18,24], being, therefore, essential for determining the machine health status and what maintenance strategy to apply [13]. In this work, from Fig. 8, it is possible to observe this worry about RUL prediction accuracy.

About 24% of the studies that explicitly tried overcoming any challenges have pointed to the improvement of RUL prediction accuracy as a direct target. However, by analyzing the other challenges described, they can indirectly affect the accuracy. For example, if we improve the process of feature selection and comprehension, our prediction's accuracy will probably be enhanced. The same occurs if we can eliminate (or

RNN Problema nel tempo del training!

at least mitigate) the model uncertainty. In this context, using an isolated model or method often may not bring the expected result, and this worry may be causing an ensembling effect. This effect can be observed by analyzing Table A.1, where only 13 (20%) of the 66 studies described have applied an isolated model or method. The rest of the studies have used ensembled methods, which means they have applied one, two, or even three methods, besides one or two prediction models, to overcome the targeted challenge.

Nevertheless, even though these studies have majority demonstrated an improvement in accuracy, they do not offer sufficient information about the real cost of this improvement in terms of computational needs, time-consuming effort, or even model development time (study and programming). Hence, more studies focusing on these issues would be valuable to improve the ensembled model already developed and help develop other ones, reflecting in the RUL accuracy improvement.

4.3. The use of system log files in RUL prediction

In Section 3.1, we have analyzed the SSC and the datasets used in RUL prediction studies. The most used dataset was the NASA C-MAPSS Dataset (28,45% corresponding 33 papers), and the second one was operating datasets (24,14% corresponding 28 papers), which means, data gathered from operating machines. There are two ways of collecting data from operating devices in the second case: sensors and system log files.

System log files are generated from diagnostics systems embedded in the SSC, providing messages or codes that identify abnormal events or deviations from optimal operation [57]. Few publications considered system log files in the context of RUL prediction [58], and in this work, only three of the papers included have used this data source. The first paper has considered using aircraft fault messages [57], the second one has predicted the need for vehicle compressors repair [58], and finally, the third has approached predictive maintenance in discrete parts manufacturing [59]. The first two works' models were designed to process messages from embedded systems without additional sensors. In the last one, ML algorithms were used based on real-world datasets, including machine log messages, event logs, and operational information, without using additional sensors.

Some advantages of using system log messages are: (i) the possibility of monitoring many different components with the same stream (instead monitor single equipment equipped with sensors) [59], (ii) they are asynchronously generated (are not triggered), (iii) they generate categorical variables (instead of continuous), and (iv) they are, in general, concise and rich in information and can be a valuable data source alternative [57].

Despite these advantages, and because these types of files were not designed for data mining [58], some challenges in using these files lie in (i) their interpretation [57] and (ii) the need for the development of techniques to extract relevant features [59].

Therefore, more in-depth studies aiming to raise the development of models which can use system log files could enhance the RUL prediction, besides running as a redundancy mechanism, for instance, in aircraft and vehicles, when coupled with traditional sensor data gathering methods.

4.4. ML Interpretability (MLI)

ML methods have become increasingly popular for understanding and predicting patterns and trends in data gathered from different domains. They have become more complex and challenging for users to evaluate and trust their results [60]. The literature also categorizes the modeling of phenomena into two basic categories: black-box (empirical-based) and white-box (principle-based) [61]. In this sense, ML methods fall in the black-box category, such that their forecast loses the necessary interpretability. This interpretability refers to the transparent decision logic of the model itself [31].

Why has MLI been so important? For instance, MLI has been an essential issue concerning information visualization, with interpretable models leading to enhanced predictions and more confident results [60]. It has not been different in RUL prediction, and there are reasons why interpretable models have been essential. For example, MLI allows potential insights into the decision-making process by comparing predictive models with experts' knowledge to validate them. Also, it can reduce errors in identifying possible causes for reduced machine lifetime. Finally, some industries have regulatory requirements concerning RUL prediction (e.g., aircraft and railroad) [31].

Therefore, regardless of the field, MLI has been a crucial concern because it can provide transparency, ensure the algorithms' expected performance and impartiality, and highlight potential issues in the training data (e.g., bias), although these interpretations are not yet systematically assessed nor neither standardized [62]. Fortunately, this is a burning discussion. In addition to highlighting this issue, our contribution in this sense will be to briefly describe three more works that may be relevant in this area besides those already cited above. First, Carvalho et al. [63] review the research field on machine learning interpretability, focusing on developed methods and metrics and the impacts on society. Then, Linardatos et al. [64] focused on machine learning interpretability methods by conducting an LR and presenting a taxonomy of these methods and the links to their programming implementations. Finally, Krishnan [65] has challenged the broad consensus regarding the ML black-box problem's existence and importance.

5. Conclusion

The study of data-driven models (particularly ML methods) for RUL prediction has grown over the last ten years and has gained more momentum from 2019 onwards (Fig. 1). The quality of these studies has been verified by the quality of the journals in which they have been published. However, the literature lacks papers that directly address the RUL prediction process and the advantages and drawbacks of the available methods in a general way. These are fundamental aspects that can contribute significantly to improving the RUL prediction.

In this context, in this work, we hope to have contributed to this area in distinct ways. First, we explored the SSC and datasets that have been used in the study of RUL prediction using ML methods (Section 3.1). Then, we presented and described an RUL prediction process consisting of four macro-steps and ten micro-steps, which we hope can be used independently of the SSC (Section 3.2). Furthermore, we mapped 14 challenges in using ML methods for RUL prediction (Section 3.3) and pointed out some approaches that have been used in the literature to overcome these challenges (Table A.1). Afterward, we highlighted some ML approaches' advantages and drawbacks that must be considered when choosing the prediction model. We hope to contribute to a better understanding of this issue (Section 3.4). Finally, we have discussed the importance of operational datasets to validate the RUL prediction models (Section 4.1). Next, we have pointed out the worry demonstrated in the literature about the RUL prediction accuracy and its apparent effect on ensembling model development (Section 4.2). We also pointed out the scarcity of papers approaching system log files as a data source to predict RUL and exploited some advantages and drawbacks in this field (Section 4.3). Finally, we have approached some considerations about MLI (Section 4.4).

Future works should focus on different fronts. The first would be to propose and test new combinations and compare the results to state-of-the-art ones due to the increasing use of ensembled models. The second would be to improve those already developed methods, whether ensembled or not, by integrating specialists' knowledge-based methodologies, for example [66]. The third would be to improve the processes of data-driven model maintenance such that drifts (change in data distribution) and the obsolescence of these models can be avoided, and their accuracy does not decrease over time [67]. On these three fronts, the main objective is to obtain models capable of predicting RUL with the lowest possible error and computational effort and effectively improve RUL prediction accuracy, which was the main challenge observed in our LR (see Fig. 8). The fourth would be to increase the development of strategies that steadily increase the RUL prediction by using system log files, jointly or separately, with data gathered from sensors. These files are a relevant source of information regarding the systems and can also contribute to minimizing error and computational effort. Nevertheless, another essential point concerning file logs is that their fair use can also be a possible source of redundancy for some systems, contributing more expressively to improving safety.

Last but not least, some efforts should be considered to develop more complex frameworks to predict RUL accurately and use the RUL predicted to activate other sequential production systems. For example, these frameworks should consider technological components capable of dealing with the large amount of data produced by the equipment. Besides, they should take into account some way to integrate production and maintenance systems, triggering these systems through the RUL predicted. The objective is to reduce downtimes, improve maintenance plans and production schedules under the scope of CBM and PHM and reduce the number of production defects as far as possible under the scope of ZDM, besides leading to increased profits for the companies and higher levels of satisfaction to the customers.

Finally, we expect the results and insights provided in this work to strengthen the body of knowledge in this field, improve the process analysis, algorithms comparison, and model development for RUL prediction, besides guide professionals and academics in the future works that can improve the RUL determination within the industrial context.

Funding

This work was financially supported by: INDTECH 4.0 – POCI-01-0247-FEDER-026653 and DEZMPP – NORTE-01-0247-FEDER-39781, both projects co-funded by the European Regional Development Fund (FEDER), through the Portuguese Competitiveness and Internationalization Operational Program (POCI).

Declaration of Competing Interest

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.

Appendix A

See Table A.1.

Table A.1
Methods and tools described the literature to overcome ML challenges in RUL prediction.

Step of the RUL Prediction Process	Challenge	Proposed Solution to Overcome the Challenge	Ref.
Data Extraction	Data extraction does not provide a meaningful representation of time-varying sensor raw measurements	Denoising Autoencoder (DAE)	[68]
		Updated Selection Strategy (USS)	
		Double Dynamic Forgetting Factors (DDFF)	
		Online Sequential Extreme Learning Machine (OS-ELM)	
Feature Extraction and Classification	Health Indicator Construction Complexity	XGBoost	[69]
		Dynamic Time Warping (DTW)	
		Hilbert-Huang Transform	[46]
		e-Support Vector Regression (e-SVR)	
		Continuous Wavelet Transform (CWT)	[13]
		Convolutional Neural Network (CNN)	
		Polynomial Regression	[48]
		Convolutional Neural Networks (CNN)	
		Long Short-Term Memory (LSTM)	
		Regulated Amplitude Signal	[70]
		Accumulative Feature	
		Support Vector Regression (SVR)	
		Isometric Feature Mapping (ISOMAP)	[71]
		Support Vector Regression (SVR)	
	Deal with multiple operating conditions and diversified degradation patterns.	Bidirectional Long Short-Term Memory (Bi-LSTM)	[7]
		Just in Time Learning (JITL)	[72]
		Gradient Boosting Decision Tree (GBDT)	
		Online Transfer Learning Method	[73]
		Simulation-based Digital Twin (DGT)	[49]
		Domain Adversarial Neural Network (DNN)	
		Bidirectional Long-Short Term Memory (Bi-LSTM)	
		Deep Belief Network (DBN)	[74]
		Fitting Curve Derivative Method of Maximum Power Spectrum Density (FDMPD)	[75]
		Kernel Extreme Learning Machine (KELM)	
		Waiting Application to Failure Time (WAFT)	
		t-Distributed Stochastic Neighbor Embedding (t-SNE)	[76]
		Deep Multiscale Convolutional Neural Network (DMSCNN)	
		Dynamic Time Wrapping (DTW)	[77]
		Wasserstein Distance (WD)	
		Maximal Overlap Discrete Wavelet Transform (MODWT)	
		Transfer-Bidirectional Gated Recurrent Unit (TBGRU)	
		Bidirectional Long Short-Term Memory (Bi-LSTM)	[78]
		Variational Mode Decomposition (VMD)	[38]
		1D Convolutional and Long Short-Term Memory (LSTM)	
		Attention Mechanism	

(continued on next page)

Table A.1 (*continued*)

Step of the RUL Prediction Process	Challenge	Proposed Solution to Overcome the Challenge	Ref.
	Difficult to deal with data noise, high level of uncertainty, the variance of local features, data sparsity, and censored data.	Enhanced Multivariate Degradation Modeling (EMDM) Subtractive Maximum Entropy Fuzzy Clustering (S-MEFC) Summation Wavelet Extreme Learning Machine (SW-ELM)	[15]
		Complete Ensemble Empirical Mode Decomposition with Adaptative Noise (CEEMDAN) Linear Combination of Monotonicity and Correlation Criteria Gated Recurrent Unit (GRU)	[79]
		Basic Characteristics based Ensemble Empirical Mode Decomposition with Adaptative Noise (BC-CEEMDAN) Gated Recurrent Unit (GRU)	[80]
		Complete Ensemble Empirical Mode Decomposition (CEEMD) Reconstruction Procedure Conditional Neural Process (CNP) Long Short-Term Memory (LSTM)	[54]
	Feature selection and efficient feature comprehension relies heavily on labor and requires extensive domain knowledge (prior knowledge of the entire system)	Time Window (TW) Extreme Learning Machine (ELM)	[12]
		Stacked Sparse Autoencoder (SSA)	[30]
		DNN-based Denoising Autoencoder Regression models based on Shallow Neural Networks (SNN)	[29]
		Correlative and Monotonic Metrics Long Short-Term Memory (LSTM) Gaussian Mixture Model (GMM) Categorical Distribution	[81]
		Principal Component Analysis (PCA) Pseudo Nearest Neighbor Method (PNNM) Particle Swarm Optimization (PSO) Least-Squares Support Vector Machine (LS-SVM)	[37]
		Time-Frequency Representation (TRF) Bilinear Interpolation Multiscale Convolutional Neural Network (MSCNN)	[47]
		Deep Convolutional Neural Network (DCNN) Bidirectional Long-Short Term Memory (Bi-LSTM) Supervised Attention Mechanism (SAM)	[39]
		Multiscale Deep Bidirectional Gated Recurrent Neural Network (MDBGRU)	[82]
		Autoencoder (Unsupervised Learning Technique) Pearson's Correlation Coefficient Least-Square Support Vector Machine Transfer Learning Method	[83]
		Autoencoder (Unsupervised Learning Technique) Cox Proportional Hazard Deep Learning (CoxPHDL) Long Short-Term Memory (LSTM)	[84]
		Autoencoder (Unsupervised Learning Technique) Deep Neural Network (DNN)	[85]
		Pearson's Correlation Analysis Random Vector Functional Link Network (RVFL) Non-Linear Autoregressive with Exogenous Set of Single RVFL	[86]
		Convolutional Layer for Feature Extraction Long Short-Term Memory (LSTM)	[43]
		Deep Belief Network (DBN) Feed Forward Neural Network (FFNN)	[87]
		Long Short-Term Memory (LSTM) Domain Adversarial Neural Network	[88]
		Convolutional Neural Network (CNN)	[89]
	Models assume that data used in training and testing phases are drawn from the same distribution (similar conditions)	Bidirectional Long-Short Term Memory (Bi-LSTM) RUL Predictor Network Unsupervised Domain Adaptation	[52]
		Generative Adversarial Neural Network (GA) Adversarial Training (AT)	[40]
		Transfer Learning Method using Deep Representation Regularization (data alignment)	[90]
Model Building and Training	Model Uncertainty	Bayesian Neural Networks (BNN) Bayesian Logistic Regression Markov Chain Monte Carlo – MH algorithm (MCMC(MH))	[6]
		Autoencoder (Unsupervised Learning Technique) Convolutional Neural Networks (CNN) Long Short-Term Memory (LSTM)	[25]
	Deal with chronological order and temporal correlation of the health monitoring data (time-varying dynamics)	Gated Recurrent Unit (GRU)	[30]
		Error-based Evolving Takagi-Sugeno Fuzzy Model (EBeTS)	[91]

(*continued on next page*)

Table A.1 (*continued*)

Step of the RUL Prediction Process	Challenge	Proposed Solution to Overcome the Challenge	Ref.
	Access to labeled failure data is scarce due to the rarity of failures	Long Short-Term Memory – Ordinal Regression (LSTM-OR)	[92]
		Empirical Standard Deviation (ESD)	
		Convolutional Generative Adversarial Network (CGAN)	[93]
		Deep Gated Recurrent Unit (DGRU)	
		Feed-Forward Neural Network (FFNN)	[94]
		Radial Basis Network (RBN)	
	The correlations of different sensor data are not explicitly considered in representation learning	Deep Separable Convolutional Network (DSCN)	[35]
RUL Prediction and Evaluation	Prediction Uncertainty	Long Term-Short Memory (LSTM)	[51]
		Gaussian Process Regression (GPR)	
	Single-path and top-down propagation (DNN models)	Time Window (TW)	[16]
		Convolutional Neural Network (CNN)	
		Long Short-Term Memory (LSTM)	
	RUL Prediction Accuracy	Long Term-Short Memory (LSTM)	[19]
		Neural Turing Machine (NTM)	
		Complete Ensemble Empirical Mode Decomposition with Adaptative Noise (CEEMDAN)	[44]
		Particle Swarm Optimization (PSO)	
		Attention Mechanism (AM)	
		Long Term-Short Memory (LSTM)	
		Hidden Markov Model (HMM)	[24]
		Multiple Layer Perceptron (MLP)	
		Long Short-Term Memory (LSTM)	[2]
		Deep Belief Network (DBN)	
		Back Propagation Neural Network (BP)	
		Adaptative Moment Estimation Algorithm (ADAM)	[42]
		Deep Long Short-Term Memory (LSTM)	
		Extreme Learning Machine (ELM)	[95]
		Sliding-Windows (SW)	[96]
		Multi-Scale Convolutional Neural Networks (MSCNN)	
		Bidirectional Long-Short Term Memory (Bi-LSTM)	
		Kernel Principal Component Analysis (KPCA)	[97]
		Gated Recurrent Unit (GRU)	
		Support Vector Machine (SVM)	[98]
		Isometric Map Algorithm (ISOMAP)	[99]
		Long Short-Term Memory with Weight Amplification (LSTMP-A)	
		Attention Mechanism	
		Classification System	[100]
		Clustering	
		Support Vector Machine (SVM)	
		Covariance Matrix	[101]
		Maximum Mean Discrepancy	
		Meta Convolutional Neural Network (Meta-CNN)	
		Meta Gated Recurrent Unit (Meta-GRU)	
		Ant Algorithm (ANTA)	[102]
		Adaptative-Neuro Fuzzy Inference System (ANFIS)	
		Neo-Fuzzy Neuron (NFN)	
		Spectrum Principal Energy Vector (SPEV)	[103]
		Deep Convolutional Neural Network (DCNN)	
		Smoothing Method	
		Wavelet Package	[104]
		Deep Perceptron Neural Network (DPNN)	
		Wavelet Filter	[105]
		Mahalanobis Distance (MD)	
		Cumulative Sum (CUMSUM)	
		Neural Network with Exogenous Inputs (NARX-NN)	
	Interpretability of the ML Results	Structured-Effect Neural Networks (SENN)	[31]

References

[1] Ruiz-Sarmiento JR, Monroy J, Moreno FA, Galindo C, Bonelo JM, Gonzalez-Jimenez J. A predictive model for the maintenance of industrial machinery in the context of industry 4.0. Eng Appl Artif Intell 2020;87:103289. https://doi.org/10.1016/j.engappai.2019.103289. January 2019.

[2] Ren L, Liu Y, Wang X, Lu J, Deen MJ. Cloud-edge based lightweight temporal convolutional networks for remaining useful life prediction in IIoT. IEEE Internet Things J 2020;XX(XX). https://doi.org/10.1109/JIOT.2020.3008170 [pp. 1–1].

[3] Psarommatis F, May G, Dreyfus PA, Kiritsis D. Zero defect manufacturing: state-of-the-art review, shortcomings and future directions in research. Int J Prod Res 2020;58(1):1–17. https://doi.org/10.1080/00207543.2019.1605228.

[4] Aydemir G, Acar B. Anomaly monitoring improves remaining useful life estimation of industrial machinery. J Manuf Syst 2020;56(June):463–9. https://doi.org/10.1016/j.jmsy.2020.06.014.

[5] Psarommatis F, Sousa J, Mendonça JP, Kiritsis D. Zero-defect manufacturing the approach for higher manufacturing sustainability in the era of industry 4.0: a position paper. Int J Prod Res 2022;60(1):73–91. https://doi.org/10.1080/00207543.2021.1987551.

[6] Wang R, Chen H, Guan C. A Bayesian inference-based approach for performance prognostics towards uncertainty quantification and its applications on the marine diesel engine. ISA Trans 2021;(xxxx). https://doi.org/10.1016/j.isatra.2021.02.024.

[7] Huang CG, Huang HZ, Li YF. A bidirectional LSTM prognostics method under multiple operational conditions. IEEE Trans Ind Electron 2019;66(11):8792–802. https://doi.org/10.1109/TIE.2019.2891463.

[8] Wu D, Jennings C, Terpenny J, Gao RX, Kumara S. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. J Manuf Sci Eng 2017;139(7):1–9. https://doi.org/10.1115/1.4036350.

[9] Kan MS, Tan ACC, Mathew J. A review on prognostic techniques for non-stationary and non-linear rotating systems. Mech Syst Signal Process 2015;62–63:1–20. https://doi.org/10.1016/j.ymssp.2015.02.016.

[10] Cheng Y, Hu K, Wu J, Zhu H, Shao X. A convolutional neural network based degradation indicator construction and health prognosis using bidirectional long short-term memory network for rolling bearings. Adv Eng Inform 2021;48(June 2020):101247. https://doi.org/10.1016/j.aei.2021.101247.

[11] El-Thalji I, Jantunen E. A summary of fault modelling and predictive health monitoring of rolling element bearings. Mech Syst Signal Process 2015;60:252–72. https://doi.org/10.1016/j.ymssp.2015.02.008.

[12] Zheng C, et al., A data-driven approach for remaining useful life prediction of aircraft engines. In: IEEE conf. intell. transp. syst. proceedings, ITSC, vol. 2018-Novem; 2018, p. 184–9. ⟨https://doi.org/10.1109/ITSC.2018.8569915⟩.

[13] Yoo Y, Baek JG. A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network. Appl Sci 2018;8(7). https://doi.org/10.3390/app8071102.

[14] Mrugalska B. Remaining useful life as prognostic approach: a review. Adv Intell Syst Comput 2019;876:689–95.

[15] Javed K, Gouriveau R, Zerhouni N. A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. IEEE Trans Cybern 2015;45(12):2626–39. https://doi.org/10.1109/TCYB.2014.2378056.

[16] Li J, Li X, He D. A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction. IEEE Access 2019;7:75464–75. https://doi.org/10.1109/ACCESS.2019.2919566.

[17] Yang Z, Baraldi P, Zio E. A comparison between extreme learning machine and artificial neural network for remaining useful life prediction. In: Proceedings of the 2016 prognostics and system health management conference (PHM-Chengdu), vol. Oct. 2016(no. 201506280015). p. 1–7. ⟨https://doi.org/10.1109/PHM.2016.7819794⟩.

[18] Sadoughi M, Lu H, Hu C. A deep learning approach for failure prognostics of rolling element bearings. In: Proceedings of the 2019 IEEE int. conf. progn. heal. manag. ICPHM 2019; 2019, p. 1–7. ⟨https://doi.org/10.1109/ICPHM.2019.8819442⟩.

[19] Falcon A, D'Agostino G, Serra G, Brajnik G, Tasso C. A neural turing machine-based approach to remaining useful life estimation. In: Proceedings of the Annu. Conf. Progn. Heal. Manag. Soc. PHM, vol. 2020-June; 2020, p. 1–8. ⟨https://doi.org/10.1109/ICPHM49022.2020.9187043⟩.

[20] Zio E, Maio FDi. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system. Reliab Eng Syst Saf 2010;95(1):49–57. https://doi.org/10.1016/j.ress.2009.08.001.

[21] Tobon-Mejia DA, Medjaher K, Zerhouni N, Tripot G. A data-driven failure prognostics method based on mixture of gaussians hidden markov models. IEEE Trans Reliab 2012;61(2):491–503. https://doi.org/10.1109/TR.2012.2194177.

[22] Liewald M, et al. Perspectives on data-driven models and its potentials in metal forming and blanking technologies. Prod Eng 2022;(0123456789). https://doi.org/10.1007/s11740-022-01115-0.

[23] Trinh HC, Kwon YK. A data-independent genetic algorithm framework for fault-type classification and remaining useful life prediction. Appl Sci 2020;10(1). https://doi.org/10.3390/app10010368.

[24] Zhu J, Chen N, Shen C. A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions. Mech Syst Signal Process 2020;139:106602. https://doi.org/10.1016/j.ymssp.2019.106602.

[25] Ren L, Dong J, Wang X, Meng Z, Zhao L, Deen MJ. A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life. IEEE Trans Ind Inform 2021;17(5):3478–87. https://doi.org/10.1109/TII.2020.3008223.

[26] T. R. G, et al. A deep neural networks based model for uninterrupted marine environment monitoring. Comput Commun 2020;157(April):64–75. https://doi.org/10.1016/j.comcom.2020.04.004.

[27] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339(jul21 1). https://doi.org/10.1136/bmj.b2535 [pp. b2535–b2535].

[28] Li X, Zhang W, Ding Q. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. Reliab Eng Syst Saf 2019;182(July 2018):208–18. https://doi.org/10.1016/j.ress.2018.11.011.

[29] Xia M, Li T, Shu T, Wan J, De Silva CW, Wang Z. A two-stage approach for the remaining useful life prediction of bearings using deep neural networks. IEEE Trans Ind Inform 2019;15(6):3703–11. https://doi.org/10.1109/TII.2018.2868687.

[30] Deng K, et al., A remaining useful life prediction method with automatic feature extraction for aircraft engines. In: Proceedings of the 2019 18th IEEE international conference on trust, security and privacy in computing and communications/13th IEEE international conference on big data science and engineering (TrustCom/BigDataSE), vol. 2019-Janua; 2019, p. 686–92, ⟨https://doi.org/10.1109/TrustCom/BigDataSE.2019.00097⟩.

[31] Kraus M, Feuerriegel S. Forecasting remaining useful life: interpretable deep learning approach via variational Bayesian inferences. Decis Support Syst 2019;125(April):113100. https://doi.org/10.1016/j.dss.2019.113100.

[32] Tang R, Fang G, Liu G, Wang H. Propulsion life prediction based on support vector machine. IOP Conf Ser Earth Environ Sci, vol. 687(no. 1); 2021. ⟨https://doi.org/10.1088/1755-1315/687/1/012082⟩.

[33] National Aeronautics and Aerospace Administration (NASA). Prognostics Center of Excellence – PCoE. ⟨https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/⟩, [Accessed 28 December 2021].

[34] Wang Y, Deng L, Zheng L, Gao RX. Temporal convolutional network with soft thresholding and attention mechanism for machinery prognostics. J Manuf Syst 2021;60(June):512–26. https://doi.org/10.1016/j.jmsy.2021.07.008.

[35] Wang B, Lei Y, Li N, Yan T. Deep separable convolutional network for remaining useful life prediction of machinery. Mech Syst Signal Process 2019;134:106330. https://doi.org/10.1016/j.ymssp.2019.106330.

[36] Zuo L, Zhang L, Zhang ZH, Luo XL, Liu Y. A spiking neural network-based approach to bearing fault diagnosis. J Manuf Syst 2020;(2006). https://doi.org/10.1016/j.jmsy.2020.07.003.

[37] Dong S, Luo T. Bearing degradation process prediction based on the PCA and optimized LS-SVM model. Meas J Int Meas Confed 2013;46(9):3143–52. https://doi.org/10.1016/j.measurement.2013.06.038.

[38] Liu C, Zhang L, Niu J, Yao R, Wu C. Intelligent prognostics of machining tools based on adaptive variational mode decomposition and deep learning method with attention mechanism. Neurocomputing 2020;417:239–54. https://doi.org/10.1016/j.neucom.2020.06.116.

[39] Li X, Jia X, Wang Y, Yang S, Zhao H, Lee J. Industrial remaining useful life prediction by partial observation using deep learning with supervised attention. IEEE/ASME Trans Mechatron 2020;25(5):2241–51. https://doi.org/10.1109/TMECH.2020.2992331.

[40] Li X, Zhang W, Ma H, Luo Z, Li X. Data alignments in machinery remaining useful life prediction using deep adversarial neural networks. Knowl-Based Syst 2020;197. https://doi.org/10.1016/j.knosys.2020.105843.

[41] Javed K, Gouriveau R, Zerhouni N, Zemouri R, Li X. Robust, reliable and applicable tool wear monitoring and prognostic: approach based on an improved-extreme learning machine. In: Proceedings of thePHM 2012–2012 IEEE int. conf. on progn. heal. manag. enhancing safety, effic. availability, eff. syst. through PHM technol. appl. conf. progr.; 2012. ⟨https://doi.org/10.1109/ICPHM.2012.6299516⟩.

[42] Wu J, Hu K, Cheng Y, Zhu H, Shao X, Wang Y. Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network. ISA Trans 2020;97:241–50. https://doi.org/10.1016/j.isatra.2019.07.004.

[43] Hinchi AZ, Tkiouat M. Rolling element bearing remaining useful life estimation based on a convolutional long-short-Term memory network. Procedia Comput Sci 2018;127:123–32. https://doi.org/10.1016/j.procs.2018.01.106.

[44] Qu J, Liu F, Ma Y, Fan J. A neural-network-based method for RUL prediction and SOH monitoring of lithium-ion battery. IEEE Access 2019;7:87178–91. https://doi.org/10.1109/ACCESS.2019.2925468.

[45] Xiang S, Qin Y, Luo J, Pu H, Tang B. Multicellular LSTM-based deep learning model for aero-engine remaining useful life prediction. Reliab Eng Syst Saf 2021;216(June):107927. https://doi.org/10.1016/j.ress.2021.107927.

[46] Cheng C, et al. A deep learning-based remaining useful life prediction approach for bearings. IEEE/ASME Trans Mechatron 2020;25(3):1243–54. https://doi.org/10.1109/TMECH.2020.2971503.

[47] Zhu J, Chen N, Peng W. Estimation of bearing remaining useful life based on multiscale convolutional neural network. IEEE Trans Ind Electron 2019;66(4):3208–16. https://doi.org/10.1109/TIE.2018.2844856.

[48] Kong Z, Cui Y, Xia Z, Lv H. Convolution and long short-term memory hybrid deep neural networks for remaining useful life prognostics. Appl Sci 2019;9(19). https://doi.org/10.3390/app9194156.

[49] Liu C, Mauricio A, Qi J, Peng D, Gryllias K. Domain adaptation digital twin for rolling element bearing prognostics. In: Proceedings of the annu. conf. progn. heal. manag. soc. PHM, vol. 12(no. 1); 2020, p. 1–10. ⟨https://doi.org/10.36001/phmconf.2020.v12i1.1294⟩.

[50] Zhang WJ, Yang G, Lin Y, Ji C, Gupta MM. On definition of deep learning. In: World autom. congr. proc., vol. 2018-June; 2018, p. 232–6. ⟨https://doi.org/10.23919/WAC.2018.8430387⟩.

[51] Liu K, Shang Y, Ouyang Q, Widanage WD. A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. IEEE Trans Ind Electron 2021;68(4):3170–80. https://doi.org/10.1109/TIE.2020.2973876.

[52] Ragab M, Chen Z, Wu M, Kwoh CK, Li X. Adversarial transfer learning for machine remaining useful life prediction. In: Proceedings of the annu. conf. progn. heal. manag. soc. PHM, vol. 2020-June; 2020. ⟨https://doi.org/10.1109/ICPHM49022.2020.9187053⟩.

[53] Zheng S, Ristovski K, Farahat A, Gupta C. Long short-term memory network for remaining useful life estimation. In: Proceedings of the 2017 IEEE int. conf. progn. heal. manag. ICPHM 2017; 2017, p. 88–95. ⟨https://doi.org/10.1109/ICPHM.2017.7998311⟩.

[54] Zhao S, Zhang Y, Wang S, Zhou B, Cheng C. A recurrent neural network approach for remaining useful life prediction utilizing a novel trend features construction method. Meas J Int Meas Confed 2019;146:279–88. https://doi.org/10.1016/j.measurement.2019.06.004.

[55] Liu Y, Wang X. Deep attention: a self-attention based neural network for remaining useful lifetime predictions. In: Proceedings of the 2021 7th int. conf. mechatronics robot. eng. ICMRE 2021; 2021, p. 98–105. ⟨https://doi.org/10.1109/ICMRE51691.2021.9384841⟩.

[56] Kang Z, Catal C, Tekinerdogan B. Remaining useful life (Rul) prediction of equipment in production lines using artificial neural networks. Sensors 2021;21(3):1–20. https://doi.org/10.3390/s21030932.

[57] Baptista M, de Medeiros IP, Malere JP, Nascimento C, Prendinger H, Henriques EMP. Comparative case study of life usage and data-driven prognostics techniques using aircraft fault messages. Comput Ind 2017;86:1–14. https://doi.org/10.1016/j.compind.2016.12.008.

[58] Prytz R, Nowaczyk S, Rögnvaldsson T, Byttner S. Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. Eng Appl Artif Intell 2015;41:139–50. https://doi.org/10.1016/j.engappai.2015.02.009.
[59] Gutschi C, Furian N, Suschnigg J, Neubacher D, Voessner S. Log-based predictive maintenance in discrete parts manufacturing. Procedia CIRP 2019;79:528–33. https://doi.org/10.1016/j.procir.2019.02.098.
[60] Chatzimparmpas A, Martins RM, Jusufi I, Kerren A. A survey of surveys on the use of visualization for interpreting machine learning models. Inf Vis 2020;19(3): 207–33. https://doi.org/10.1177/1473871620904671.
[61] Li JW, Zhang WJ, Yang GS, Tu SD, Chen XB. Thermal-error modeling for complex physical systems: the-state-of-arts review. Int J Adv Manuf Technol 2009;42(1–2): 168–79. https://doi.org/10.1007/s00170-008-1570-x.
[62] Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: Proceedings of the 2018 IEEE 5th international conference on data science and advanced analytics (DSAA); 2018, p. 80–9. ⟨https://doi.org/10.1109/DSAA.2018.00018⟩.
[63] Carvalho DV, Pereira EM, Cardoso JS. Machine learning interpretability: a survey on methods and metrics. Electronics 2019;8(8):1–34. https://doi.org/10.3390/electronics8080832.
[64] Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable ai: a review of machine learning interpretability methods. Entropy 2021;23(1):1–45. https://doi.org/10.3390/e23010018.
[65] Krishnan M. Against interpretability: a critical examination of the interpretability problem in machine learning. Philos Technol 2020;33(3):487–502. https://doi.org/10.1007/s13347-019-00372-9.
[66] Modi S, Lin Y, Cheng L, Yang G, Liu L, Zhang WJ. A socially inspired framework for human state inference using expert opinion integration. IEEE/ASME Trans Mechatron 2011;16(5):874–8. https://doi.org/10.1109/TMECH.2011.2161094.
[67] Dreyfus P-A, Pélissier A, Psarommatis F, Kiritsis D. Data-based model maintenance in the era of industry 4.0: a methodology. J Manuf Syst 2022;63 (January):304–16. https://doi.org/10.1016/j.jmsy.2022.03.015.
[68] Berghout T, Mouss LH, Kadri O, Saïdi L, Benbouzid M. Aircraft engines remaining useful life prediction with an adaptive denoising online sequential extreme learning machine. Eng Appl Artif Intell 2020;96(September):103936. https://doi.org/10.1016/j.engappai.2020.103936.
[69] Que Z, Xu Z. A data-driven health prognostics approach for steam turbines based on Xgboost and DTW. IEEE Access 2019;7:93131–8. https://doi.org/10.1109/ACCESS.2019.2927488.
[70] X. Yang, Y. Zhang, W. Yang, Y. Wang, and Y. ZhengYang X, Zhang Y, Yang W, Wang Y, Zheng Y. Long-term RUL prediction of bearings with signal amplitude regulation and accumulative feature. In: Proc. 2020 IEEE 9th data driven control learn. syst. conf. DDCLS 2020; 2020, pp. 633–8. ⟨https://doi.org/10.1109/DDCLS49620.2020.9275213⟩.
[71] Benkedjouh T, Medjaher K, Zerhouni N, Rechak S. Remaining useful life estimation based on nonlinear feature reduction and support vector regression. Eng Appl Artif Intell 2013;26(7):1751–60. https://doi.org/10.1016/j.engappai.2013.02.006.
[72] Li X, Mba D, Loukopoulos P. A just-in-time-learning based data-driven method for valve failure prognostics. In: Proc. - 11th Int. conf. progn. syst. heal. manag. PHM-Jinan 2020; 2020, p. 434–8. ⟨https://doi.org/10.1109/PHM-Jinan48558.2020.00084⟩.
[73] Zeng F, Li Y, Jiang Y, Song G. An online transfer learning-based remaining useful life prediction method of ball bearings. Meas J Int Meas Confed 2021;176 (February):109201. https://doi.org/10.1016/j.measurement.2021.109201.
[74] Li L, Peng Y, Song Y, Liu D. Lithium-ion battery remaining useful life prognostics using data-driven deep learning algorithm. In: Proc. - 2018 progn. syst. heal. manag. conf. PHM-Chongqing 2018; 2019, p. 1094–100. ⟨https://doi.org/10.1109/PHM-Chongqing.2018.00193⟩.
[75] Zhao H, Liu H, Jin Y, Dang X, Deng W. Feature extraction for data-driven remaining useful life prediction of rolling bearings. IEEE Trans Instrum Meas 2021;70. https://doi.org/10.1109/TIM.2021.3059500.
[76] Ge Y, Liu JJ, Ma JX. Remaining useful life prediction using deep multi-scale convolution neural networks. IOP Conf Ser Mater Sci Eng, vol. 1043(no. 3); 2021. ⟨https://doi.org/10.1088/1757-899X/1043/3/032011⟩.
[77] Cao Y, Jia M, Ding P, Ding Y. Transfer learning for remaining useful life prediction of multi-conditions bearings based on bidirectional-GRU network. Meas J Int Meas Confed 2021;178(February):109287. https://doi.org/10.1016/j.measurement.2021.109287.
[78] Zhang A, et al. Transfer learning with deep recurrent neural networks for remaining useful life estimation. Appl Sci 2018;8(12). https://doi.org/10.3390/app8122416.
[79] Xiao L, Liu Z, Zhang Y, Zheng Y. Bearings remaining useful life prediction with combinatorial feature extraction method and gated recurrent unit network. In: Proceedings of the 2020 IEEE 9th data driven control learn. syst. conf. DDCLS 2020; 2020, p. 360–5. ⟨https://doi.org/10.1109/DDCLS49620.2020.9275098⟩.
[80] Xiao L, Liu Z, Zhang Y, Zheng Y, Cheng C. Degradation assessment of bearings with trend-reconstruct-based features selection and gated recurrent unit network. Meas J Int Meas Confed 2020;165:108064. https://doi.org/10.1016/j.measurement.2020.108064.
[81] Huang Y, Tang Y, Vanzwieten J, Liu J, Xiao X. An adversarial learning approach for machine prognostic health management. In: Proceedings of the 2019 int. conf. high perform. big data intell. syst. HPBD IS 2019; 2019, p. 163–8. ⟨https://doi.org/10.1109/HPBDIS.2019.8735480⟩.
[82] Behera S, Misra R, Sillitti A. Multiscale deep bidirectional gated recurrent neural networks based prognostic method for complex non-linear degradation systems. Inf Sci 2021;554:120–44. https://doi.org/10.1016/j.ins.2020.12.032.
[83] Mao W, He J, Zuo MJ. Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning. IEEE Trans Instrum Meas 2020;69(4):1594–608. https://doi.org/10.1109/TIM.2019.2917735.
[84] Chen C, et al. Predictive maintenance using cox proportional hazard deep learning. Adv Eng Inform 2020;44(February):101054. https://doi.org/10.1016/j.aei.2020.101054.
[85] Ren L, Zhao L, Hong S, Zhao S, Wang H, Zhang L. Remaining useful life prediction for lithium-ion battery: a deep learning approach. IEEE Access 2018;6:50587–98. https://doi.org/10.1109/ACCESS.2018.2858856.
[86] Gou B, Xu Y, Fang S, Pratama RA, Liu S. Remaining useful life prediction for lithium-ion battery using ensemble learning method. IEEE Power Energy Soc Gen Meet 2019;2019-Augus. https://doi.org/10.1109/PESGM40551.2019.8973811.
[87] Deutsch J, He D. Using deep learning-based approach to predict remaining useful life of rotating components. IEEE Trans Syst Man Cybern Syst 2018;48(1):11–20. https://doi.org/10.1109/TSMC.2017.2697842.
[88] da Costa PR de O, Akçay A, Zhang Y, Kaymak U. Remaining useful lifetime prediction via deep domain adaptation. Reliab Eng Syst Saf 2020;195(September 2019):106682. https://doi.org/10.1016/j.ress.2019.106682.
[89] Guo Q, Miyamae Y, Wang Z, Taniuchi K, Yang H, Liu Y. Senvis-Net: learning from imbalanced machinery data by transferring visual element detectors. Int J Mach Learn Comput 2018;8(5):416–22. https://doi.org/10.18178/ijmlc.2018.8.5.722.
[90] Zhang W, Li X, Ma H, Luo Z, Li X. Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions. Reliab Eng Syst Saf 2021;211(February). https://doi.org/10.1016/j.ress.2021.107556.
[91] Camargos MO, Bessa I, D'Angelo MFSV, Cosme LB, Palhares RM. Data-driven prognostics of rolling element bearings using a novel Error Based Evolving Takagi–Sugeno Fuzzy Model. Appl Soft Comput J 2020;96:106628. https://doi.org/10.1016/j.asoc.2020.106628.
[92] Diksha VTV, Malhotra P, Vig L, Shroff G. Data-driven prognostics with predictive uncertainty estimation using ensemble of deep ordinal regression models. Int J Progn Heal Manag 2019;(2008). https://doi.org/10.36001/ijphm.2019.v10i4.2612.
[93] Behera S, Misra R. Generative adversarial networks based remaining useful life estimation for IIoT. Comput Electr Eng 2021;92(June 2020):107195. https://doi.org/10.1016/j.compeleceng.2021.107195.
[94] Hu C, Youn BD, Kim T. Semi-supervised learning with co-training for data-driven prognostics, In: Proceedings of the 2012 IEEE conference on prognostics and health management; 2012, p. 1–10. ⟨https://doi.org/10.1109/ICPHM.2012.6299526⟩.
[95] Razavi-Far R, Chakrabarti S, Saif M, Zio E, Palade V. Extreme learning machine based prognostics of battery life. Int J Artif Intell Tools 2018;27(8). https://doi.org/10.1142/S0218213018500367.
[96] Jiang Y, Lyu Y, Wang Y, Wan P. Fusion network combined with bidirectional LSTM network and multiscale CNN for useful life estimation LSTM network and multiscale CNN for useful life estimation. In: Proceedings of the 12th int. conf. adv. comput. intell. ICACI 2020; 2020, p. 620–7. ⟨https://doi.org/10.1109/ICACI49185.2020.9177774⟩.
[97] Chen J, Jing H, Chang Y, Liu Q. Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process. Reliab Eng Syst Saf 2019;185(January):372–82. https://doi.org/10.1016/j.ress.2019.01.006.
[98] Nuhic A, Terzimehic T, Soczka-Guth T, Buchholz M, Dietmayer K. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. J Power Sources 2013;239:680–8. https://doi.org/10.1016/j.jpowsour.2012.11.146.
[99] Xiang S, Qin Y, Zhu C, Wang Y, Chen H. Long short-term memory neural network with weight amplification and its application into gear remaining useful life prediction. Eng Appl Artif Intell 2020;91(February):103587. https://doi.org/10.1016/j.engappai.2020.103587.
[100] Yang C, Letourneau S, Liu J, Cheng Q, Yang Y. Machine learning-based methods for TTF estimation with application to APU prognostics. Appl Intell 2017;46(1): 227–39. https://doi.org/10.1007/s10489-016-0829-4.
[101] Ding P, Jia M, Zhao X. Meta deep learning based rotating machinery health prognostics toward few-shot prognostics. Appl Soft Comput 2021;104:107211. https://doi.org/10.1016/j.asoc.2021.107211.
[102] Soualhi A, Medjaher K, Celrc G, Razik H. Prediction of bearing failures by the analysis of the time series. Mech Syst Signal Process 2020;139:106607. https://doi.org/10.1016/j.ymssp.2019.106607.
[103] Ren L, Sun Y, Wang H, Zhang L. Prediction of bearing remaining useful life with deep convolution neural network. IEEE Access 2018;6:13041–9. https://doi.org/10.1109/ACCESS.2018.2804930.
[104] Hong S, Yin J. Remaining useful life prediction of bearing based on deep perceptron neural networks. In: Proceedings of the 2018 2nd international conference on big data and internet of things – BDIOT 2018, vol. 36(no. 7); 2018, p. 175–79, ⟨https://doi.org/10.1145/3289430.3289438⟩.
[105] Rai A, Upadhyay SH. The use of MD-CUMSUM and NARX neural network for anticipating the remaining useful life of bearings. Meas J Int Meas Confed 2017; 111:397–410. https://doi.org/10.1016/j.measurement.2017.07.030.