SEMI-SUPERVISED MACHINE LEARNING TECHNOLOGY IN FAULT DIAGNOSIS AND PROGNOSIS: A SYSTEMATIC LITERATURE REVIEW

By: Iwueze Chidinma, Okonkwo Ikechukwu and Amalagu Chikezie

ABSTRACT

This study undertakes a comprehensive exploration of recent advancements in the domain of mechanical fault diagnosis and prognosis, employing semi-supervised machine learning methodologies. Conducting a systematic review, we scrutinized prominent databases such as Web of Science, ACM Digital Library, Science Direct, Wiley Online Library, and IEEE Xplore spanning the timeframe from January 2015 to October 2023. Among the 50 records initially identified, 28 primary studies were meticulously chosen for inclusion, provided they utilized semi-supervised machine learning algorithms for mechanical fault detection or prognosis, and offered empirical results derived from industrial case studies.

In the selected studies, fault diagnosis and prognosis were executed through diverse techniques such as artificial neural networks, decision tree methods, hybrid models, or latent variable models, with some studies incorporating two distinct techniques independently. The remaining studies showcased a spectrum of machine learning approaches, spanning from rule-based models to partition-based algorithms, with only two studies adopting online learning methods. Key advantages of these methodologies includes: the utilization of readily available unlabelled data, high performance, the capacity to unveil intricate nonlinear relationships, and computational efficiency. However, a notable limitation surfaces in the form of reduced model performance when faced with concept drift. This review underscores the growing trend of studies in the field of fault detection and prognosis in recent years, yet emphasizes the need for additional research to tackle challenges inherent in real-world scenarios.

INTRODUCTION

The significance of machine maintenance in influencing machine downtime and production costs directly correlates with a manufacturing company's competitiveness concerning cost, quality, and performance [28, 27]. Asides repairing malfunctioning equipment, maintenance serves the crucial purpose of preserving machinery functionality and minimizing breakdowns. Predictive maintenance, as the term implies, revolves around early problem detection. In this proactive approach, maintenance activities are based on monitoring the actual condition of machinery and addressing components as soon as signs of deterioration is detected, thus avoiding reactive fault repairs [2]. This strategy surpasses reactive and preventive maintenance in several aspects [19, 20], including the

- prevention of catastrophic failures,
- extension of equipment life,
- optimization of preventive maintenance tasks,
- enhanced maintenance inventory management,
- optimization of equipment availability,
- and improved productivity.

By averting severe failures, curbing unexpected faults, and maximizing mean time between failures (MTBF), predictive maintenance contributes to reduced workplace accidents and their severity, diminished repair instances and mean time to repair (MTTR), and elongated equipment life. Consequently, these outcomes translate into increased earnings, lowered maintenance and production costs, and a more sustainable manufacturing approach [19, 8]. According to Sullivan et al. [20], the successful implementation of predictive maintenance programs can yield an average reduction in maintenance costs ranging from 25% to 30%, accompanied by a remarkable return on investment (ROI) of 1000%.

Predictive maintenance aligns with condition-based maintenance [19], in condition-based maintenance, decision-making leans on diagnostics and prognostics techniques [23]. Diagnostics, entailing fault detection and identification (FDI), is typically executed through hardware redundancy or analytical redundancy methods. Hardware redundancy involves measuring the same parameters using multiple sensors and comparing duplicate signals through various techniques, such as signal processing methods [26]. Analytical redundancy methods are grounded in mathematical models, further categorized into quantitative (model-based) and qualitative (data-driven) methods [26]. While both approaches compare predicted or estimated parameters to real measured values, model-based methods estimate parameters based on mathematical models under normal operating conditions, whereas data-driven methods leverage historical data and artificial intelligence algorithms for parameter prediction or anomaly detection.
While diagnostics focuses on detecting, isolating, and identifying faults, prognostics aims to predict faults in the monitored system before they occur [23]. Prognostics techniques, particularly those estimating the remaining useful life (RUL), are employed to predict how soon and how likely a fault is to occur, with a predominant emphasis on RUL estimation in the literature [23]. RUL estimation methods, which may also be data-driven, strive to predict the duration a machine will function before a fault arises or if it will fail within a specified time interval [23]. Analytical redundancy methodologies, not requiring additional hardware, are more cost-effective to implement than hardware redundancy methods [26]. The advent of Internet of Things (IoT) technologies in industrial settings facilitates real-time digital representation of production processes and equipment status, leading to a substantial increase in industrial data volume [1]. Data-driven methods, especially machine learning and data mining techniques, adeptly extract knowledge from this wealth of data and have proven successful in predictive maintenance contexts [4].

Furthermore, while model-based methods can yield favorable results with precise system models, constructing an accurate mathematical model for complex systems is a challenging endeavor, making model-based methods less viable [23, 26]. Recent review papers [4, 3] focusing on machine learning techniques for predictive maintenance highlight commonly used data-driven methods such as artificial neural networks, support vector machines, decision trees (including ensemble methods), k-means, and logistic regression, among others. Predicting and detecting faults in industrial equipment pose challenging tasks necessitating the selection of appropriate techniques for accurate results. This study conducts a systematic literature review of machine learning methods employed for detecting mechanical faults and prognosing faults in equipment within real-world scenarios. Serving as a foundation for predictive maintenance system implementation, this review aims to identify future research opportunities. Notably, it focuses solely on real-world industrial cases, acknowledging the challenges specific to fault diagnosis and prognosis in practical settings often overlooked in studies relying on controlled experiments or simulations. Manufacturing systems' inherent complexity, non-stationary processes, noise, and other disturbances underscore the importance of suitable machine learning methods, especially in the absence of historical fault data, common in industrial settings, limiting the learning task to unsupervised and semi-supervised methods. The study presents an overview of the current landscape of fault diagnosis and prognosis in real-world scenarios using machine learning techniques, guided by two research questions exploring the state-of-the-art semi-supervised machine learning methods for fault detection and prognosis in manufacturing equipment, their strengths and weaknesses, and their application in the context of data stream learning. The document is organized into sections covering the review protocol, results, discussion, concluding remarks, and future work directions.

METHODS

This study adheres to the PRISMA statement [10], which provides a checklist and flow diagram for reporting systematic reviews. However, recognizing the healthcare-oriented focus of PRISMA, this review, centered on engineering and computer science themes, also incorporates guidelines from [25]. These guidelines adapt medical review standards to the nuances of software engineering and other scientific fields. The study follows three main phases: planning, conducting, and reporting the review, along with related activities. The identification of the need for this review emerged during the literature search on semi-supervised machine learning methods for fault detection and prognosis. To our knowledge, no existing systematic literature review specifically addresses semi-supervised machine learning methods for mechanical fault detection and prognosis in real-world scenarios within manufacturing equipment.

Before initiating the research, a review protocol was formulated, outlining research questions, the search strategy, study selection criteria, and the data extraction process.

RESEARCH QUESTIONS

The research questions are detailed as follows:

RQ ONE: What semi-supervised machine learning algorithms and methods are currently employed for mechanical fault detection and fault prognosis in manufacturing equipment? (It delves into aspects such as the most frequently used machine learning algorithms, types of learning tasks addressed, and the utilization of hybrid and ensemble methods.)

RQ TWO: What limitations and advantages do these algorithms and methods present? (addressing factors like why specific algorithms are chosen, their weaknesses, and the overall rationale behind their application).

The protocol development began with formulating meaningful research questions to guide the state-of-the-art review, emphasizing a comprehensive exploration of semi-supervised machine learning methods in the context of mechanical fault detection and prognosis in manufacturing equipment.

THE SEARCH STRATEGY

The search strategy, designed to identify recently published research on semi-supervised machine learning methods for fault detection and prognosis in real-world scenarios, is detailed in the subsequent sections.

INFORMATION SOURCES

The study selected academic databases, aligning with systems appropriate for systematic reviews [7] and compatible with the study's subject matter. While IEEE Xplore may not be the ideal platform for systematic reviews, its significance in engineering, electronics, and computer science warranted its inclusion to complement results from the other four databases [7].

SEARCH STRING

For the search string, a combination of terms using Boolean operators OR and AND was employed to construct a comprehensive query. To identify relevant studies involving machine learning, terms such as "mining," "machine learning”, “learning”, “semi-supervised," and "knowledge discovery" were included. The decision to incorporate terms beyond "semi-supervised" and "machine learning" was driven by the substantial overlap between machine learning and data mining, often used interchangeably. Broad terms like "mining" and "learning" aimed to encompass related concepts like "pattern mining" or "data stream learning." Additionally, "knowledge discovery" was included due to its frequent use of machine learning techniques and relevance to fault detection and prognosis. To capture research on mechanical fault detection and prognosis, terms like "fault detection," "fault prediction," "fault prognosis," and "predictive maintenance" were included in the search string.

STUDY SELECTION CRITERIA

The study implemented a set of inclusion and exclusion criteria to discern relevant studies from the search results. Inclusion criteria focused on publications centering on the utilization of semi-supervised machine learning algorithms and methods specifically for fault detection and prognosis. Only studies aligning with one or more inclusion criteria were deemed relevant for the review's purpose. On the other hand, exclusion criteria were employed to filter out research lacking essential characteristics. Duplicate publications and those not in English were eliminated, and only full-length articles published since 2015 were considered for inclusion.

.

RESULTS

The search, selection, and exclusion protocols culminated in the identification of 28 primary studies from an initial pool of 50. These records underwent screening based on publication details, including language, as well as an assessment of the information provided by the title and abstract

On addressing Research Question One (RQ ONE): "What semi-supervised machine learning algorithms and methods are currently employed for mechanical fault detection and fault prognosis in manufacturing equipment?",

The selected primary studies showcased a diverse array of semi-supervised machine learning algorithms and methods utilized for mechanical fault detection and prognosis, often incorporating combinations of different algorithms. Many studies conducted comparative analyses among various semi-supervised machine learning algorithms, either to demonstrate the efficacy of the proposed method or to identify the most suitable algorithm. In instances where a specific algorithm was chosen as the best-performing, only those selected algorithms are detailed in the review.

The selected studies encompassed various learning tasks tailored to the specific problems at hand and the available data. In certain studies [18, 17, 16, 15, 21, 14, 13, 12, 11] (with optional inclusion of 5, 22, 9, 74, 24, 18), labeled data was accessible but primarily utilized for validating the semi-supervised learning models. The ensuing subsections present detailed summaries, outlining the advantages and limitations of the machine learning algorithms and methods employed in these studies. Following this section, a more in-depth description of these algorithms and methods is provided.

On addressing Research Question 2 (RQ TWO): “What limitations and advantages do these algorithms and methods present?”, the majority of selected studies provided insight into the rationale behind their choice of specific machine learning algorithms or combinations. Many studies emphasized the inherent strengths of the selected algorithms or considered their specific advantages in the context of fault detection, prediction, and implementation in industrial settings. Additionally, several studies reported the benefits offered by their proposed approaches. However, only a limited number of studies presented the limitations of the employed algorithms or the proposed methodologies.

For instance, deep learning models were recognized for their ability to learn features from raw data when the training and test data shared the same distribution and feature space. Nevertheless, in real industrial settings characterized by time-varying conditions, this condition often does not hold. To address this issue, a method proposed in [6] surpassed simple pattern recognition and classification by employing semi-supervised deep learning to identify the dynamic properties of a machine tool. This approach enabled early fault detection and health status diagnosis under time-varying operation conditions. Similarly, the framework presented in [21] tackled the challenge of time-varying operations by utilizing two conditional variational autoencoder (CVAE) models to estimate the health index of a machining center and predict its future condition for a given operating regime. The authors highlighted the CVAE's ability to remove noise from sensor data, automatically extract meaningful features, and handle complex conditional probability distributions. These characteristics facilitated the development of a semi-supervised learning method capable of estimating a machine's health under time-varying operations, even in scenarios with limited labelled data, common in real industrial settings.

Some of the selected studies showcased hybrid machine learning models tailored to address tasks that a single algorithm or type of algorithms may not effectively handle. An analysis conducted in [16] revealed that supervised learning algorithms, while suitable for classification tasks, require labelled data and may misclassify unknown faults. To overcome these limitations, semi-supervised algorithms were introduced, but they faced challenges in distinguishing between fault types. Here, unsupervised clustering algorithms proved useful. The combination of these diverse algorithms resulted in the creation of a predictive maintenance system adept at detecting and classifying various mechanical faults from unlabelled data.

In the study outlined in [22], clustering techniques and a recurrent neural network (RNN) were employed to address the issue of missing labels. The weighted pair-group method using centroid (WPGMC) was chosen for its capacity to create homogeneous groups that are more easily interpreted. Simultaneously, the RNN, equipped with internal memory, was selected for its ability to capture complex, non-linear relationships in time series data. This capability is crucial for uncovering wear and tear patterns in industrial equipment over time.

Furthermore, the methodology proposed in [15] involved constructing a health index by leveraging an autoencoder's ability to learn the relationship between input data variables. Subsequently, simple linear regression was applied to predict future values of the health index and calculate the Remaining Useful Life (RUL). While the methodology demonstrated its capacity to learn from unlabelled data and showcased applicability across different domains, the absence of run-to-failure data necessitated the somewhat arbitrary definition of the anomaly threshold. Additionally, it was acknowledged that the prediction accuracy could be enhanced by employing more sophisticated algorithms than simple linear regression.

In the methodology presented in [12], the approach integrates multiple learning models for fault prediction in a press module. Notably, interpretability was a key concern, leading to the adoption of an association rules approach. This choice ensures that the fault prediction method remains interpretable, and the approach demands minimal parameter tuning, exhibiting generic applicability to various types of sensor data.

The study detailed in [17] utilized a Hidden Markov Model (HMM) due to its suitability for time series data, capable of detecting long-term degradation and handling dynamic features in a semi-supervised manner. The proposed approach, which combines HMM with sliding windows and a genetic algorithm, accommodates data with asynchronous sampling rates and does not require extensive domain knowledge. However, the model's performance may decrease when the contamination probability distribution of a production cycle substantially differs from the distributions used to train the HMM. Addressing this, the study suggests considering expert advice when selecting production cycles for HMM training.

The R4RE algorithm, an enhancement of the QARMA framework introduced in [24], stands out for its fully distributed nature and the assurance that resulting rules meet user-defined interestingness criteria. Overcoming limitations of QARMA, R4RE allows quantification of the consequent item in closed intervals and incorporates online pruning of generated rules within the search process. The study demonstrates that these developments lead to improved Remaining Useful Life (RUL) estimates and reduced error rates in the test set.

The rule-based method XCS, as described in [11], features a "covering" mechanism facilitating rule set recalibration for unseen data without re-training the entire model, making it suitable for online learning. Recognizing dependencies between variables and identifying different failure patterns, the rules generated by XCS offer valuable insights for failure origin identification.

In [14], the MCOD algorithm was selected for its low-memory and processing requirements, making it ideal for processing streaming data and producing easily understandable results. However, it is noted that MCOD is sensitive to input parameters, which can influence the number of outlier reports.

DISCUSSION

In this section, the focus is on discussing the results in more depth, aiming to identify trends and ideas. Additionally, it provides an overview of the challenges encountered in using machine learning methods for detecting mechanical faults and predicting faults in real manufacturing scenarios, while considering potential directions for future research.

Concerning RQ ONE, which investigates the machine learning algorithms and methods employed for mechanical fault detection and prognosis in manufacturing equipment, specific attention is given to the diverse set of algorithms discussed earlier. In the study from 2019 [13], the authors leveraged plant floor automation and information system (PFS) data from a real-world automotive manufacturing line. They introduced a manufacturing system-wide balanced random forest (MBRSF) model, utilizing a random survival forest to estimate a hazard function from balanced system-wide data. This approach aimed to quantify the likelihood of breakdown events over time. Notably, experiments on 20 machines demonstrated that the MBRSF outperformed other survival models by approximately 90%, as measured by the integrated Brier's score. This underscores the effectiveness of the proposed model in capturing complex and dynamic machine breakdown patterns.

In the 2019 study [9], an LSTM network was developed to predict faults in industrial ovens using sensor data and log events. Trained on consecutive time series, the network predicted five subsequent future events, reaching up to 25 minutes into the future. Despite dealing with strongly imbalanced data, the model's performance was evaluated using metrics such as the Matthews correlation coefficient (0.691), recall (0.790), and F1-score (0.803). The evaluation results indicated that while the performance slightly decreased for predictions further into the future, the network's overall performance was deemed acceptable for all predictions.

Another 2019 study [5] proposed a fault diagnosis method leveraging digital twin technology. A high-fidelity dynamic virtual model of a car body-side production line, representing the entire product life cycle, was created as the digital twin. This virtual model was used to build a diagnosis model incorporating an SSAE layer for feature extraction from unsupervised data and a softmax classifier for assigning probabilities to class labels. Deep transfer learning facilitated the transfer of knowledge gained in the virtual space to a new fault diagnosis model in the physical space. The cooperation between the virtual and physical entities of the digital twin resulted in accurate fault predictions (average accuracy = 97.96%) and adaptability to new working conditions.

In a subsequent 2021 study [21], a health model for a predictive maintenance system was introduced. This system considered time-varying operational conditions and allowed for the scheduling of maintenance and production. Utilizing condition monitoring sensor data, production data, and future production orders, the proposed framework employed two CVAE models: HA-CVAE and DS-CVAE. HA-CVAE assessed the machine's condition under time-varying operational conditions, while DS-CVAE served as a data simulator to generate realistic sensor data. The method was capable of estimating the machine's health under different operating conditions, even in scenarios with scarce labelled data. Furthermore, it demonstrated the ability to predict the machine's future health and degradation condition.

Hybrid Models

In the 2018 study by Strauß et al. [16], a predictive maintenance approach was proposed for a heavy lift EMS at the BMW Group. The study integrated semi-supervised, unsupervised, and supervised learning techniques to detect and classify mechanical faults. Recognizing the scarcity of fault data, a semi-supervised method for anomaly detection was initially employed. Three semi-supervised models were built using normal data exclusively and evaluated on a dataset containing both normal and fault data. Unsupervised models were then used to cluster the fault data, allowing the creation of a dataset containing normal data and instances of three different types of failures. Eight supervised models were trained and evaluated on this dataset, with four achieving an F1-score of over 90%. The final selected models for the predictive maintenance system were one-class SVM, K-means, and random forest, considering both performance and computational requirements.

In 2019, Rousopoulou et al. [9] presented a solution for predictive maintenance of industrial ovens used by a medical devices manufacturer. The study selected an LSTM network to predict faults from condition monitoring sensor data and log events. Additionally, an outlier detection method was combined with a classifier to detect faults in unlabeled acoustic data. Three outlier detection algorithms (DBSCAN, LOF, and MAD) were tested, with DBSCAN yielding the best results. To address the imbalance in the dataset, the synthetic minority oversampling technique (SMOTE) was applied. An SVM trained on this balanced dataset achieved an accuracy of 85% and an F1-score of 0.86, detecting new faults in live audio measurements. The system issued a fault notification if the model detected five consecutive faults.

In the study published in [22], the focus was on developing a prognostic maintenance model in a context where no labeled data existed. Conducted at a German automotive manufacturer, the study aimed to address a scenario where maintenance of a milling tool was subjectively performed by machine operators based on visual inspections, and no labeled data were available.

The proposed method aimed to uncover latent information hidden in historic data, which included maintenance and production records, control data, and sensor data. The approach considered two orthogonally related dimensions: the time dimension, indicating when a tool was replaced, and the condition dimension, referring to information about damaged and undamaged tools inferred from the available data.

To differentiate between correct and incorrect tool replacement decisions, a 4-field matrix was defined based on these dimensions. Clustering techniques were applied along both dimensions. The time dimension clustering utilized the agglomerative hierarchical clustering algorithm named weighted pair-group method using centroid (WPGMC). Cluster one represented tool replacements performed late in the tool’s lifetime, while cluster two represented early replacements.

For the condition dimension, time series’ sequences were clustered into two groups using the mean absolute deviation (MAD) to measure the intensity of the sequences’ oscillations. Sequences with a lower MAD value, indicating weaker oscillations, were assigned to cluster one, while the remaining sequences were assigned to cluster two.

The results obtained from clustering the data were orthogonally related, and based on that, the data observations were assigned to each quadrant of the 4-field matrix. Time series’ sequences of "type 1" in the matrix represented the replacement of an undamaged tool late in its lifetime, reflecting correct decisions by machine operators. This data was used to train and test a recurrent neural network (RNN) to predict the tools’ remaining useful life (RUL). The RNN model was then used to predict the RUL of "type 3" observations (undamaged tools replaced too early), demonstrating that using the prognostic model would have extended the tools’ lifetimes for about one-third of these tool replacements.

In 2021, the authors of [15] proposed a framework based on autoencoders and simple linear regression for the construction of a health index, aiming to predict the Remaining Useful Life (RUL) of industrial equipment, especially in scenarios lacking fault history data. The methodology involves using an autoencoder to learn the normal structure of the data and construct the health index by assessing the difference between the input and reconstructed data through mean absolute error (MAE). Subsequently, simple linear regression is employed to predict the trend of the health index over time. Detection of abnormal behavior is based on an increase in the trend of the health index and a large slope parameter in the regression.

The RUL is defined as the time difference between the current health index and the predicted time of failure occurrence. The proposed approach was applied in practical industrial settings, including cases involving pump equipment and robot arms. For pumps, where a limited amount of failure data was available, the health index accurately signaled an increase before an impending failure, and the methodology successfully predicted faults. In the case of robot arms, vibration sensor data was collected over several months, and the proposed method accurately predicted faults. To verify its reliability, the authors conducted additional experiments comparing the proposed method to an isolation forest for RUL prediction. Results based on metrics such as mean absolute error (MAE) and root mean square error (RMSE) demonstrated superior performance of the proposed method in all experiments involving data from both pumps and robot arms.

In a 2021 study presented in [12], a methodology for fault detection and prediction in cold forming processes at a Phillips factory in the Netherlands is detailed. The study leverages Gaussian Mixture Model (GMM), the FP-Growth algorithm, and CBA-CB for the analysis.

Over a year, information on the normal operating conditions of a press module was collected, encompassing material batch data, maintenance logs, and information from acoustic emission sensors. The authors employed the matrix profile, a data structure for time series analysis, to derive two meta-time series from the acoustic emission data. These meta-time series were utilized for anomaly detection and fault prediction through rule mining.

For anomaly detection, a statistical approach was chosen due to the absence of labeled data guiding the anomaly threshold definition. The matrix profile was also employed in fault prediction, involving the mining of salient subsequences to identify common patterns. Principal Component Analysis (PCA) was then applied to these subsequences, and resulting principal components were clustered using GMM. The acoustic emission data was segmented into non-overlapping time windows, and each pattern within a window was labeled based on the previously discovered clusters. Integrating the acoustic emission data with maintenance logs, FP-Growth was utilized to mine association rules, forming the basis for a classifier with the modified CBA-CB algorithm.

Comparative evaluation against a majority classifier showed promising results. While the majority classifier achieved a high micro F1-score but failed to predict any events, the proposed method demonstrated the ability to predict faults related to three out of the four infrequent maintenance events of interest. The achieved micro F1-score was 0.632, showcasing the effectiveness of the proposed approach in addressing extremely rare events occurring less than 0.05% of the time.

Other approaches

In the realm of fault detection and prediction, certain studies explored alternative approaches, employing machine learning algorithms from diverse categories such as instance-based algorithms, rule-based models, and dynamic Bayes networks.

A 2017 study, as presented in [17], introduces a predictive maintenance approach integrating sliding windows, a genetic algorithm, and a hidden Markov model to estimate and predict long-term degradation in a hard masking deposition tool. Handling data sampled at varying rates, the method calculates summary statistics over sliding windows to synchronize data features. Employing a Hidden Markov Model (HMM) to cluster time series data, the algorithm considers past and present states to estimate the tool's degradation. The genetic algorithm collaborates with the HMM to select the most suitable subset of features. Evaluated using historical data and insights from maintenance experts, this unsupervised method demonstrates its effectiveness in estimating and predicting tool degradation.

In 2019, Naskos et al. [14] proposed a real-time method for detecting oil leakages in the large tanks of a BENTELER Automotive factory. Utilizing the micro-cluster continuous outlier detection (MCOD) algorithm on sensor data streams, the authors incorporate domain knowledge of the production cycle to determine machinery's operational status and enhance algorithm performance. Comparative analysis reveals that combining MCOD with domain knowledge yields superior results compared to variants, including the application of MCOD to raw data without prior domain knowledge.

Another study in 2019, presented by Graß et al. [18], outlines an anomaly detection method for identifying faults in the fans of a reflow oven. Faced with challenges such as the absence of labeled data and varying machine configurations due to different items processed in the same production line, the authors devised a multi-step approach. Data clustering based on machine configurations precedes segmenting data for each cluster and extracting relevant features. Utilizing K-Nearest Neighbors (K-NN), an anomaly threshold is defined based on the mean distance between a given segment and its k nearest neighbors. Tested on seven years of historical data, this approach successfully detects fan malfunctions, showcasing its applicability in a complex production environment.

In 2019, a study by the authors of [24] introduced the "Rules 4 Rare Events" (R4RE) algorithm, employing a quantitative association rule mining method to predict the Remaining Useful Life (RUL) of industrial equipment. This algorithm represents an enhancement over a previous method by enabling quantifications of the consequent item in closed intervals and integrating online pruning of the generated rules. Applied to sensor data collected from a real factory, the R4RE algorithm generated approximately 4500 rules estimating the RUL (RUL-time) of monitored machines. Furthermore, the authors expanded the dataset to predict the RUL in terms of produced parts (RUL parts), a more robust measure that considers periods when a machine is idle or turned off. Comparative evaluation against other machine learning models revealed that the R4RE model achieved superior results in predicting RUL-time (RMSE = 34.2; MAE = 28.7; MAPE% = 20.1) and performed competitively in predicting RUL-parts (RMSE = 668.7; MAE = 120.8; MAPE% = 3.76).

Also in 2019, Chen et al. [11] utilized a rule-based method to predict the RUL of machinery, specifically applying a modified version of the eXtended Classifier System (XCS). Focusing on predicting the RUL of a digital radio frequency matching box (RF-MB) used in the semiconductor manufacturing process, the authors employed XCS, a rule-based machine learning method capable of recalibrating its rule set through interaction with the environment. However, as XCS processes binary input data, the authors adapted a continuous-valued input-capable version, XCSR. Framing the estimation of the RF-MB's RUL as a classification problem, the authors applied Fisher discriminant analysis (FDA) to reduce the number of variables before leveraging XCSR, achieving a notable prediction accuracy of 97.3%.

Concerning RQ TWO: on the limitations and advantages of the machine learning algorithms and methods, the reviewed studies highlighted various strengths and drawbacks. Several advantages were commonly identified, emphasizing high performance and the capability to uncover complex nonlinear relationships in the data. This was evident in studies such as [24, 13, 22, 21] where these characteristics were crucial considerations in algorithm selection. Computational efficiency emerged as another significant advantage, with the efficiency of machine learning algorithms like [14] being underscored. However, it was noted that advanced algorithms, particularly artificial neural networks, could be resource-intensive, demanding substantial processing power and memory. The interpretability of models was deemed beneficial, not only for technical aspects like debugging and fine-tuning but also for fostering user trust and aiding in informed decision-making. This interpretability aspect was highlighted in [12], emphasizing its relevance in industrial contexts for identifying the root causes of faults.

Furthermore, the ability to learn from unlabeled data, a key feature of unsupervised learning, was recognized as crucial in situations where historical fault data is challenging to obtain. This challenge arises due to infrequent fault occurrences, inadequate logging, or temporal misalignment with condition monitoring data. Several studies, including [18, 17, 16, 15, 21, 14, 9, 13], acknowledged the importance of unsupervised/semi-supervised learning in handling these scenarios. The lack of historical fault data was acknowledged as a prevalent issue, impacting a quarter of the selected studies. However, validating unsupervised/semi-supervised models in the absence of historical data poses challenges, requiring domain knowledge guidance and testing in real production environments, as demonstrated in [12].

On the other hand, only a few studies explicitly identified the limitations of the machine learning algorithms used. Issues such as the decrease in model performance when training and test data do not share the same distribution were recognized as a significant limitation with implications for fault detection and prediction in general. These limitations were addressed in specific studies like [6, 17, 15, 14, 5], shedding light on challenges associated with chosen algorithms and methodologies.

CONCLUSION

This study conducts a systematic literature review on semi-supervised machine learning methods applied to mechanical fault detection and fault prognosis in manufacturing equipment within real-world scenarios. Following PRISMA guidelines and guidelines for software engineering systematic reviews, an initial set of 50 records published between January 2015 and October 2023 were identified. After applying selection criteria, 28 primary studies were included in the review. The selected studies were further examined based on two research questions, focusing on machine learning methods employed, their advantages and limitations, and their application in fault detection and prognostics in diverse manufacturing contexts.

While each study in the review addresses mechanical fault detection or prognosis in real scenarios, differences exist in terms of the context, machinery involved, and data characteristics. These distinctions, typical of industrial case studies, make direct comparisons challenging. The inherent complexity of manufacturing systems and the time-varying properties of production processes pose challenges for fault detection and prognosis. The review emphasizes the need for more research to develop machine learning algorithms capable of handling noisy, non-stationary data and capturing nonlinear patterns in machinery interactions.

A notable challenge is the non-stationarity of manufacturing environments, which can be addressed through online learning or data stream learning. Online learning techniques that adapt incrementally to changes in data distribution caused by non-stationary environments are identified as promising for processing high-speed streams of sensor data. However, the review reveals a deficit in studies dedicated to online learning methods, particularly in the context of mechanical fault detection or prediction.

The absence of labeled data is a common concern in real-world scenarios, limiting the learning task to unsupervised and semi-supervised methods. Nearly half of the encountered but unselected studies in the review employed unsupervised learning techniques (only). The study underscores the necessity not only to demonstrate the effectiveness of semi-supervised models but also to develop new methods capable of learning complex nonlinear relationships without labeled data and while adapting to concept drift.

The study emphasizes the importance of considering factors collectively, such as computational efficiency and interpretability, in the development of predictive maintenance systems. Predictive maintenance offers economic, safety, and environmental benefits, but justifying the investment requires accurate models aligned with business needs. The study advocates for future research to address these challenges and explore opportunities in online learning methods for fault detection and prognosis in manufacturing environments.

REFERENCE

1. Boyes, H., Hallaq, B., Cunningham, J., & Watson, T. (2018). The industrial internet of things (IIoT): An analysis framework. *Computers in Industry, 101*, 1–12. https://doi.org/10.1016/j.compind.2018.04.015

2. Muller, A., Crespo Marquez, A., & Iung, B. (2008). On the concept of e-maintenance: Review and current research. *Reliability Engineering & System Safety, 93*(8), 1165–1187. https://doi.org/10.1016/j.ress.2007.08.006

3. Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering, 137*, 106024. https://doi.org/10.1016/j.cie.2019.106024

4. Zhang, W., Yang, D., & Wang, H. (2019). Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Systems Journal, 13*(3), 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565

5. Xu, Y., Sun, Y., Liu, X., & Zheng, Y. (2019). A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning. *IEEE Access, 7*, 19990–19999. https://doi.org/10.1109/ACCESS.2018.2890566

6. Luo, B., Wang, H., Liu, H., Li, B., & Peng, F. (2018). Early fault detection of machine tools based on deep learning and dynamic identification. *IEEE Transactions on Industrial Electronics, 66*(1), 509–518. https://doi.org/10.1109/TIE.2018.2807414

7. Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar,

PubMed, and 26 other resources. *Research Synthesis Methods, 11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

8. Rødseth, H., & Schjølberg, P. (2016). Data-driven predictive maintenance for green manufacturing. In *Proceedings of the 6th International Workshop of Advanced Manufacturing and Automation*. Advances in Economics, Business and Management Research, Atlantis Press, 36–41.

9. Rousopoulou, V., Nizamis, A., Giugliano, L., Haigh, P., Martins, L., Ioannidis, D., ... & Stirna, J. (2019). Data Analytics Towards Predictive Maintenance for Industrial Ovens: A Case Study Based on Data Analysis of Various Sensors Data. In *Advanced Information Systems Engineering Workshops (CAISE 2019)*, Lecture Notes in Business Information Processing, 349, 83–94.

10. Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology, 62*, e1–e34.

11. Chen, L.-Y., Lee, J.-H., Yang, Y.-L., Yeh, M.-T., & Hsiao, T.-C. (2019). Predicting the Remaining Useful Life of Plasma Equipment through XCSR. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO ’19*, Association for Computing Machinery, New York, 1263–1270. https://doi.org/10.1145/3319619.3326879

12. Avendano, D. N., Caljouw, D., Deschrijver, D., & Van Hoecke, S. (2021). Anomaly detection and event mining in cold forming manufacturing processes. *International Journal of Advanced Manufacturing Technology*, 1–16. https://doi.org/10.1007/s00170-020-06156-2

13. Bukkapatnam, S. T. S., Afrin, K., Dave, D., & Kumara, S. R. T. (2019). Machine learning and AI for long-term fault prognosis in complex manufacturing systems. *CIRP Annals - Manufacturing Technology, 68*(1), 459–462. https://doi.org/10.1016/j.cirp.2019.04.104

14. Naskos, A., Gounaris, A., Metaxa, I., Koechling, D., & Stirna, J. (2019). Detecting anomalous behavior towards predictive maintenance. In *Advanced Information

Systems Engineering Workshops (CAISE 2019)*, Lecture Notes in Business Information Processing, 349, 73–82.

15. Kim, D., Lee, S., & Kim, D. (2021). An applicable predictive maintenance framework for the absence of run-to-failure data. *Applied Sciences, 11*(11), 5180. https://doi.org/10.3390/app11115180

16. 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 *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, IEEE; IEEE Comp Soc; Expedia Grp; Baidu; Squirrel AI Learning; Ankura; Springer, 1474–1483.

https://doi.org/10.1109/BigData.2018.8622076

17. Kinghorst, J., Geramifard, O., Luo, M., Chan, H. L., Yong, K., Folmer, J., ... & Vogel-Heuser, B. (2017). Hidden Markov model-based predictive maintenance in semiconductor manufacturing: A genetic algorithm approach. In *IEEE International Conference on Automation Science and Engineering*, IEEE, 1260–1267.

18. Grass, A., Beecks, C., & Soto, J. A. C. (2019). Unsupervised Anomaly Detection in Production Lines. In *Machine Learning For Cyber Physical Systems, ML4CPS 2018*. Technologien fur die intelligente Automation, 9, 18–25.

19. Mobley, R. K. (2001). Predictive Maintenance. In *Plant Engineer's Handbook*. Elsevier, 867–888.

20. Sullivan, G., Pugh, R., Melendez, A. P., & Hunt, W. D. (2010). Operations & maintenance best practices-a guide to achieving operational efficiency (release 3). Technical Report, Pacific Northwest National Lab. (PNNL), Richland.

21. Zhai, S., Gehring, B., & Reinhart, G. (2021). Enabling predictive maintenance integrated production scheduling by operation-specific health prognostics with generative deep learning. *Journal of Manufacturing Systems*. https://doi.org/10.1016/j.jmsy.2021.02.006

22. Zschech, P., Heinrich, K., Bink, R., & Neufeld, J. S. (2019). Prognostic Model Development with Missing Labels: A Condition-Based Maintenance Approach Using Machine Learning. *Business & Information Systems Engineering, 61*(3), 327–343. https://doi.org/10.1007/s12599-019-00596-1

23. Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing, 20*(7), 1483–1510. https://doi.org/10.1016/J.YMSSP.2005.09.012

24. Christou, I. T. (2019). Avoiding the Hay for the Needle in the Stack: Online Rule Pruning in Rare Events Detection. In *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, 661–665.

25. Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele Univ, 33(2004), 1–26. https://www.researchgate.net/publication/228756057

26. Hwang, I., Kim, S., Kim, Y., & Seah, C. E. (2010). A survey of fault detection, isolation, and reconfiguration methods. *IEEE Transactions on Control Systems Technology, 18*(3), 636–653. https://doi.org/10.1109/TCST.2009.2026285

27. Holmberg, K., Adgar, A., Arnaiz, A., Jantunen, E., & Mascolo, J. (2010). *E-maintenance*. In K. Holmberg, A. Adgar, A. Arnaiz, E. Jantunen, J. Mascolo, & S. Mekid (Eds.), Springer, London.

28. Aboelmaged, M. G. (2014). Predicting e-readiness at firm level: An analysis of technological, organizational, and environmental (TOE) effects on e-maintenance readiness in manufacturing firms. *International Journal of Information Management, 34*(5), 639–651. https://doi.org/10.1016/j.ijinfomgt.2014.05.002

29. Gertler, J. J. (2017). *Fault detection and diagnosis in engineering systems*. CRC Press. https://www.taylorfrancis.com/books/9781351448796