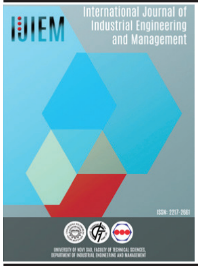
ISSN 2683-345X

journal homepage: http://ijiemjournal.uns.ac.rs/

International Journal of Industrial Engineering and Management

Volume 15 / No 4 / December 2024 / 304 - 319



Original research article

A framework for unsupervised learning and predictive maintenance in Industry 4.0

G. Nota[a,*] 0000-0002-4829-9967, F. D. Nota[b] 0000-0002-3380-0544,
A. Toro[c] 0000-0001-7593-8026, M. Nastasia[d] 0000-0002-9670-5464

[a] *University of Salerno, Department of Business Sciences—Management & Innovation Systems (DISA-MIS), 00165 Fisciano, Italy;*
[b] *Defense Analysis & Research Insitute, Center for Higher Defense, Piazza della Rovere, 00165 Rome, Italy;*
[c] *Catholic University of Pereira, Department of Basic Sciences and Engineering, 660001 Pereira, Colombia;*
[d] *STAMEC s.r.l., 83030 Montefredane (AV), Italy*

ABSTRACT

In recent decades, the economic importance of maintaining machines, equipment, and production facilities has prompted many scholars to examine various aspects of the maintenance of physical assets. However, the industry continues to face the recurring problem of improving product and equipment maintenance processes. New opportunities for improving these processes arise from Industry 4.0 technologies because they make it possible to realize better solutions to the problem of predictive maintenance. Starting from a Big Data and Internet of Things (IoT) architecture as a reference, this paper proposes an abstract framework for predictive maintenance using unsupervised learning models to support decision-making in maintenance programs. From the abstract framework, a predictive maintenance system was developed to enable effective just-in-time maintenance strategies. An unsupervised machine learning algorithm, based on the Gauxian mixtures model, allows us to study the influence on a machine's behavior of a single variable, a group of variables of the same type, and combined variables of different types. The algorithm provides experts with information on which part of the machine they need to focus on to find potential causes of future failures. The case study conducted for an Italian automotive company shows preliminary results on the effectiveness of the approach.

ARTICLE INFO

Article history:

Received December 6, 2023
Revised July 10, 2024
Accepted September 4, 2024
Published online October 11, 2024

Keywords:
Predictive maintenance;
Industry 4.0;
Internet of things;
Unsupervised machine learning;
Gaussian mixtures

*Corresponding author:
Giancarlo Nota
nota@unisa.it

1. Introduction

Industrial machines and equipment are products whose operation performance inevitably deteriorates over time. They operate continuously for extended periods under specific stress, load, or extreme conditions in the manufacturing environment [1]. This deterioration, also called degradation, often leads to failures in machine components, entire machines, or even production lines. Underperformed components or subsystems might fail when reaching a critical degradation degree and risk system safety [2].

According to Chan et al. [3], "in most cases, there is a measurable degradation process before a machine fails." In fact, throughout its useful life, the machine can continuously degrade until it reaches a condition in which it is possible to observe a drop in its performance level, the moment at which incipient defects or initial failures can occur due to the deg-

Published by the University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia.

https://doi.org/10.24867/IJIEM-2024-4-365

radation progress. The continuous proliferation of these defects gradually increases their severity, causing the machine not to perform its function correctly.

Therefore, maintenance has been introduced as an efficient way to assure a satisfactory level of reliability during the useful life of a physical asset [4]. In the manufacturing industry, maintenance consists of carrying out all the necessary actions to restore the durable equipment or keep it in specific operating conditions [5] since the equipment is intended to last a certain amount of time and must, therefore, be maintained.

With the introduction of Industry 4.0 technology, new maintenance opportunities arise in networked factories with the availability of massive data from processes, machines, and systems. This development allows operators, or even intelligent scheduling systems, to monitor the machinery conditions instead of their faults, anticipating possible failures and optimizing the assets utilization [6]. Furthermore, the Predictive Maintenance (PdM) approach makes it possible thanks to the evolution of preventive maintenance by enabling just-in-time work strategies [7].

PdM has acquired great relevance for industrial scenarios as a maintenance strategy for diagnosing and prognosing a machine based on its condition. Compared with other maintenance strategies, the PdM strategy has the advantage of lowering maintenance costs and time [8].

As argued by Busse [9], to use PdM, a condition monitoring system must provide information on the current machine condition (Diagnosis) and, depending on the system's maturity, predict the future condition (Prognosis). Machine Learning (ML) is one of the trend methods used to make predictions and estimations using real-world datasets [10], with critical applications in many industrial areas [5], [11], [12].

Key challenges in industrial machinery maintenance have been highlighted in several papers. These include the need for advanced health diagnosis systems in complex industrial setups [13], the implementation of deep learning-based predictive maintenance in the Industrial Internet of Things (IIoT) [14], and obstacles in developing generalized data-driven predictive maintenance systems [15]. Common themes across the studies include the importance of addressing skill shortages, improving data acquisition and analysis, enhancing service quality, and transitioning from time-based to condition-based maintenance. The papers also emphasize the potential of Artificial Intelligence (AI) and ML in overcoming these challenges, particularly in fault diagnosis and prognosis. However, issues such as data quality, real-time processing, and the development of global rather than equipment-specific approaches remain significant hurdles in the field of industrial machinery maintenance.

Recent research highlights the potential benefits and challenges of implementing PdM in industrial settings. While PdM offers significant advantages, practical implementation can be difficult due to uncertain benefits, advanced IT requirements, and lack of failure data [16]. The integration of PdM with Digital Twins (DT) and IIoT technologies presents opportunities for improving maintenance processes across various industries [17], [18].

ML techniques, including supervised learning and anomaly detection, are being explored to develop accurate prediction models for industrial machinery maintenance [19]. However, challenges persist, such as the analysis of unsupervised data from different sources and managing large volumes of time series data. Innovative approaches to managing the large-scale storage needs of time series data are being developed to address these issues and enhance the effectiveness of PdM in industrial applications [18].

The literature on predictive maintenance (PdM) in industrial settings reveals several critical gaps, some of which our research aims to address. Firstly, there is a pressing need to transition from traditional time-based maintenance strategies to more efficient condition-based approaches. This transition can improve resource utilization and reduce downtime but involves challenges such as technology integration and initial costs.

On the other hand, while AI and ML offer significant potential for fault diagnosis and prognosis, challenges remain, particularly concerning the use of unsupervised ML models for PdM. However, their real-time processing capabilities are limited, requiring improvements to handle rapid data analysis and decision-making in industrial settings.

This research focuses on the proposition of an abstract framework for supporting PdM in cases when only unsupervised learning algorithms can be used. The proposed framework can be used for predicting equipment failures given the indication of abnormal behaviors of industrial machines in a manufacturing context, supporting decision-makers in planning maintenance activities.

As a main contribution, we introduce an abstract framework for PdM and unsupervised learning to support maintenance teams to make informed decisions, propose better maintenance scheduling, and establish effective maintenance policies based on the condition of factory equipment. The abstract framework was implemented in a case study scenario where a Gaussian Mixture Model (GMM) was used for the diagnosis and prognosis of machinery failures.

The paper is structured as follows. After the literature review on the maintenance of industrial machinery and the PdM approach in an unsupervised context, section 3 resumes the research methodology and establishes research questions. In section 4, an Industry 4.0 Big Data and IoT architecture is designed to meet the needs of unlabeled data-driven analytics systems focused on predictive machinery maintenance. As a complement to this architecture, an abstract framework for predictive maintenance using unsupervised learning is presented in section 5, together with an application that uses the Gaussian mixture algorithm. Section 6 shows a real-world system that implements the proposed approach through a case study in an Italian automotive industry company. Finally, after discussing this work's findings, and the comparison with the state-of-the-art, the research results are summarized, discussing benefits, limitations, and future developments.

2. Literature review

Maintenance is crucial for manufacturing organizations, with the advent of Industry 4.0 driving new approaches and models for Maintenance 4.0 [20]. Proper maintenance practices are essential for maximizing operational efficiency, preventing costly breakdowns, and extending the lifespan of machinery [21]. Best maintenance practices encompass various aspects, including preventive maintenance, safety, and specific techniques to ensure maintaining machine reliability [22].

The integration of Industry 4.0 techniques enables the development of predictive maintenance systems, such as those using ML to estimate and predict machinery degradation. These systems allow for informed decision-making regarding maintenance operations [6]. Overall, adopting advanced maintenance strategies is crucial for improving industrial productivity and efficiency in the context of Industry 4.0.

2.1 Industrial machinery maintenance

The maintenance of industrial assets is a critical aspect of companies' efficiency and product quality [11]. In general, maintenance is defined as all the technical and managerial actions taken during the period of use to maintain or restore the required functionality of a product or resource [12].

In the manufacturing industry, maintenance consists of carrying out all necessary actions to restore durable equipment or keep it in specific operating conditions [23]. The purpose is to maximize the effectiveness of the production system, preserving its functionality while controlling the cost induced by maintenance activities.

Equipment maintenance is a task entailing significant factories' resources (budgets, operators, time, etc.) [24]. For this reason, the simplest maintenance option is emergency maintenance, which essentially consists of repairing a machine when a failure is detected.

However, unexpected breakdowns strongly impact companies' costs [25] since repair expenditures are much higher than scheduled maintenance costs [26]. Thus, maintenance plans and strategies are designed to reduce or eliminate the number of failures and related costs. Selecting a successful maintenance strategy requires a good knowledge of maintenance management principles and practices and specific facility performance [27]. Therefore, the main recognized maintenance approaches for an industrial context, such as those presented in [28], should be considered to better understand these strategies.

2.2 PdM in an unsupervised context

Discovering imminent faults (prognosis) can be seen as a particular case of outlier detection since it is an observation deviating from other observations that can arouse the suspicion that it was generated by anomalous equipment behavior [29]. Consequently, supervised, semi-supervised, and unsupervised learning methods are employed to solve this problem [30].

While supervised learning provides an effective approach to building ML models, in practice, labeled data in manufacturing is not usually available. Therefore, a valuable alternative is unsupervised learning, which aims to build the representation of a given dataset without any label-based feedback mechanism [31].

In the literature, studies oriented to PdM using the unsupervised learning approach for the maintenance of industrial assets can be found. Amruthnath and Gupta suggested a methodology for early fault detection and fault class prediction [24]. The methodology has been applied using density estimation via GMM Clustering and the K-means algorithm starting from the vibration data of an exhaust fan [32].

An unsupervised learning system based on Cluster Analysis [33] has also been applied to the predictive maintenance of aircraft engines.

Some research papers promote the idea of frameworks for unsupervised learning to solve problems related to equipment maintenance. For example, Farbiz et al. [34] describe a framework based on cognitive analytics with unsupervised learning for machine health monitoring, anomaly detection, and predictive maintenance. The experimental results on an industrial robot demonstrate the effectiveness of their approach.

The predictive maintenance framework proposed by Kim et al. [35] is based on unsupervised learning and consists of constructing a Health Index and Remaining Useful Life (RUL) predictions. The usefulness and applicability of the proposal were conducted through two different real-life cases: monitoring the condition of a pump in a manufacturing plant and a robotic arm in a production line of automobiles.

The technical or organizational problems encountered when implementing unsupervised ML solutions relate to the management of big data and the need to involve application domain experts. In particular, the acquisition of relevant manufacturing data is a widespread problem for ML applications, as the availability, quality, and composition of the manufacturing data at hand strongly influence the performance of ML algorithms [36]. Furthermore, preprocessing data also has a critical impact on the accuracy of the results, and a great deal of time is spent preparing the data and extracting information, the reason for which performance problems (low speed, low accuracy, high memory complexity [34]) may occur.

According to Wuest [37], in some cases, there might be no expert feedback about the results, making it difficult to interpret them. Considering a preventive maintenance strategy, a period of time or number of cycle intervals is defined, and an expert evaluates the condition at the end of this interval [38]. Expert knowledge is also relevant when interpreting data [39].

Our approach focuses on the unsupervised context. The goal is to contribute to filling some gaps in the literature related to the acquisition and analysis of manufacturing machine data through a Big Data and IoT architecture. In particular, we propose an abstract framework for implementing PdM in unsupervised learning contexts, which is formed of several steps independent of the algorithms used and, therefore, provides an abstract structure for testing different unsupervised learning solutions.

3. Research Methodology

The main steps of the empirical research methodology adopted in this work to guide the investigation, as illustrated in **Figure 1**, are: *1. Literature review*, *2. System design*, *3. System implementation*, *4. Data analysis*, and *5. Evaluation*. The literature review revealed a lack of in-depth studies for predictive maintenance of machinery in a scenario where historical data on machinery failures and their causes do not exist or are scarce.

This consideration led to the formulation of the following two research questions:

RQ1: how effective are unsupervised learning models in identifying potential causes of machine failures and predicting future failures in industrial environments within an Industry 4.0 context?

RQ2: how can a predictive maintenance system that uses unsupervised learning models supported by Big Data and IoT architecture improve decision-making processes in maintenance programs?

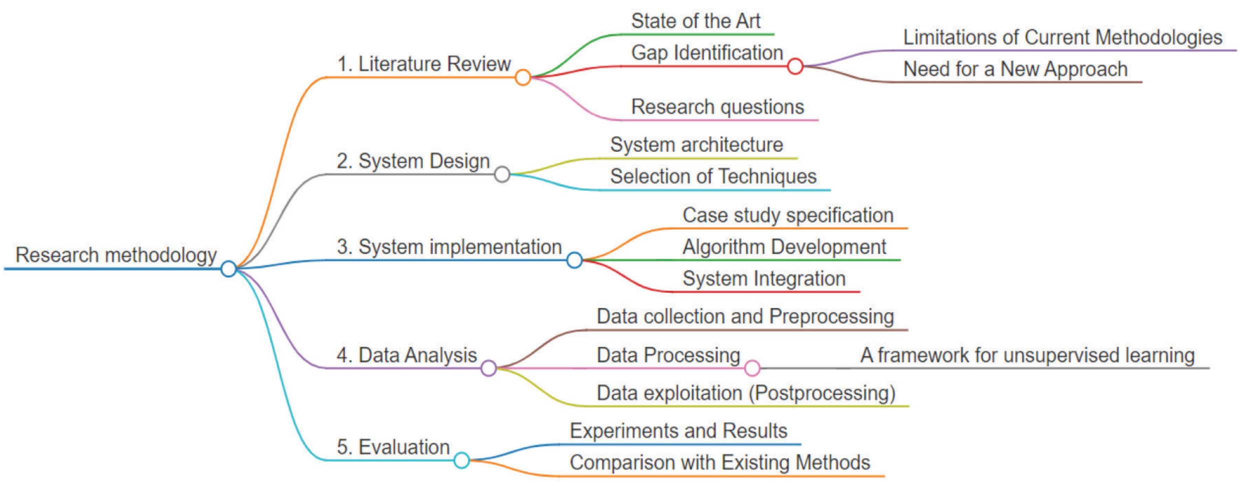


Figure 1. The research methodology

In the following, the methodological steps of the research will be covered as follows: *System Design* in sections 4 and 5, where a system architecture design and selection of ML techniques for prediction are performed; *System implementation* and *Data analysis* will be developed in section 6, where a case study was carried out in a mechanical company within the automotive sector to verify the efficiency of the developed PdM approach. In this section, a PdM system model is designed and an abstract framework for PdM which utilizes unsupervised ML algorithms is proposed; Finally, the approach's Evaluation is carried out in section 7, based on the results presented in section 6.3.

4. System architecture

Advanced diagnostic and prognostic algorithms require the definition of an appropriate architecture, such as the one presented in **Figure 2**. This PdM system architecture designed for the implementation of our project is principally based on IoT and Big Data Analysis (BDA) technologies. This set of technologies is completed with the Internet of Services (IoS), which takes the processed information from Big Data tools and deploys it at the right place and in the right form [40].

Generally, in the life cycle of Big Data in a data processing environment, it is always necessary to acquire data, store it temporarily or permanently, analyze it, and produce outputs/results [41]. Therefore, the phases for extracting and analyzing large amounts of data and the development of an analytical model of industrial big data, that can be used for preventive and predictive maintenance, is typically carried out in three major stages: *pre-processing (I)*, *processing (II)*, and *postprocessing (III)*. These stages contain the four phases of the Big Data life cycle: *(1) data acquisition, (2) data storage, (3) data analysis, and (4) data exploitation*.

I. Pre-processing: this stage is related to acquiring and storing data from various sources of information. It involves the implementation of the following phases:

1) Data acquisition: in this phase, the data is collected from several sources such as CNC machines (computerized numerical control)/PLC (programmable logic controller) and sensors connected to the machines (e.g., energy and vibration sensors). After this, filtering and cleaning data processes are performed before storage.

2) Data extraction, transformation, and storage: once content from the mentioned data sources is retrieved, it is stored in a centralized server through a software component developed as an automatic storage mechanism, which allows for managing large-scale datasets.

II. Processing: this stage is related to analyzing the data obtained in the previous phases and developing the Big Data analytical model for predictive maintenance. Hence, the data analysis phase reported in **Figure 2** is performed as follows:

3) Data analysis: the implementation of the ML

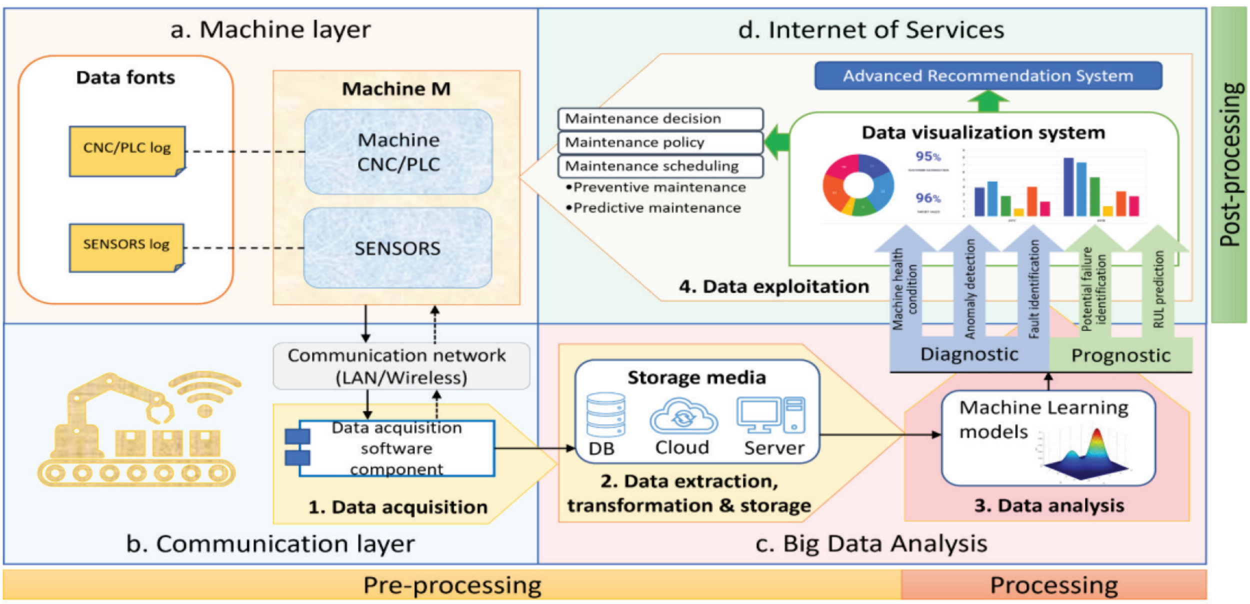


Figure 2. A predictive maintenance system architecture is based on integrating the IoT, BDA, and IoS concepts mentioned in [6]

analytical model for predictive maintenance is first performed. An algorithm is used afterward to process and analyze the collected data, aiming at diagnosing anomalous behaviors on machines and prognosing maintenance needs. Details of this implementation will be presented in the section describing Gaussian Mixtures (GM) as an indication of anomalous machine behavior.

III. Post-processing: this stage concerns the exploitation of the results of the performed analysis employing the following process:

4) Data exploitation: it includes the interpretation of the discovered patterns by the implemented algorithm. The knowledge acquired is organized and presented through data communication/visualization mechanisms so that the client can use it. To achieve this, thresholds are created and used in the visualization tool to communicate instantly with all stakeholders and thus help them identify patterns, trends, and correlations. Finally, the results are evaluated to verify the approach's effectiveness and to find improvement opportunities, allowing for new and better functionalities to be subsequently integrated into the feedback mechanisms.

As presented in **Figure 2**, the execution of the life cycle of Big data analysis within the technological architecture occurs initially with the integration of systems of sensors, IoT, and communication devices with the machine (quadrant *a*, upper left), turning the machine into a structure capable of providing production data, through the communication layer, serving phase 1 of the Big data analytics process: data acquisition using a specialized software component (quadrant *b*, lower left). Then, this information is received and sent to a central server, database, cloud system, data warehouse, or any other storage solution as part of phase 2: *data extraction, transformation, and storage*. These two phases constitute *stage I* of the life cycle of Big Data: preprocessing.

Once the data has been stored, Big Data analytics processes (quadrant *c*, bottom right) can be run to diagnose machine health conditions, predict potential failures, or make RUL predictions. These steps correspond to phase 3: data analysis, performed in *stage II*: processing.

Finally, in the IoS (quadrant *d*, upper right), data visualization, as part of phase 4: data exploitation enables the maintenance responsible to making informed decisions, establishing maintenance policies, and planning preventive maintenance activities. In this way, *stage III*: *post-processing* is carried out with the help of specialized data visualization tools.

5. Machine Learning for predictive analysis

ML refers to the process of teaching a computer system by exploring patterns and discovering inferences among data without explicitly programmed instructions [42], using an algorithm capable of learning with minimum or no additional support. In maintenance, ML can be used to predict potential faults and future equipment conditions (prognostics).

ML algorithms are commonly divided into supervised and unsupervised models [42]. The supervised models predict future events by learning models trained using labeled data points [43], while unsupervised models are trained on all data points and are used primarily for data clustering. In this research, given the unlabeled data coming from industrial machines, the main focus has been on using unsupervised learning as the foundation of an unsupervised predictive maintenance approach. The problem of PdM using unsupervised learning was approached with the definition of a framework with several steps, which aims to provide a structure for testing different unsupervised learning solutions. **Figure 3** shows the proposed framework at an abstract level; the steps are independent of the decision of the algorithm and the evaluation metrics to be used, and therefore different techniques could be used and tested in other contexts.

Based on the framework and according to the characteristics and behavior of the variables used in this research, corresponding to energy, acceleration, and velocity data, we have used clustering to group points based on a defined algorithmic criterion. In the case study of Section 6, we have chosen the GMMs and trained them to detect anomalous data points.

5.1 Gaussian Mixtures as an indicator of anomalous machine behavior

The GMM is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in several systems [44]. It is one of the most popular data clustering methods where each cluster obeys Gaussian distribution. The task of clustering is to group observations into different components by estimating each cluster's parameters [45]. The parameters are estimated from the training data using the iterative

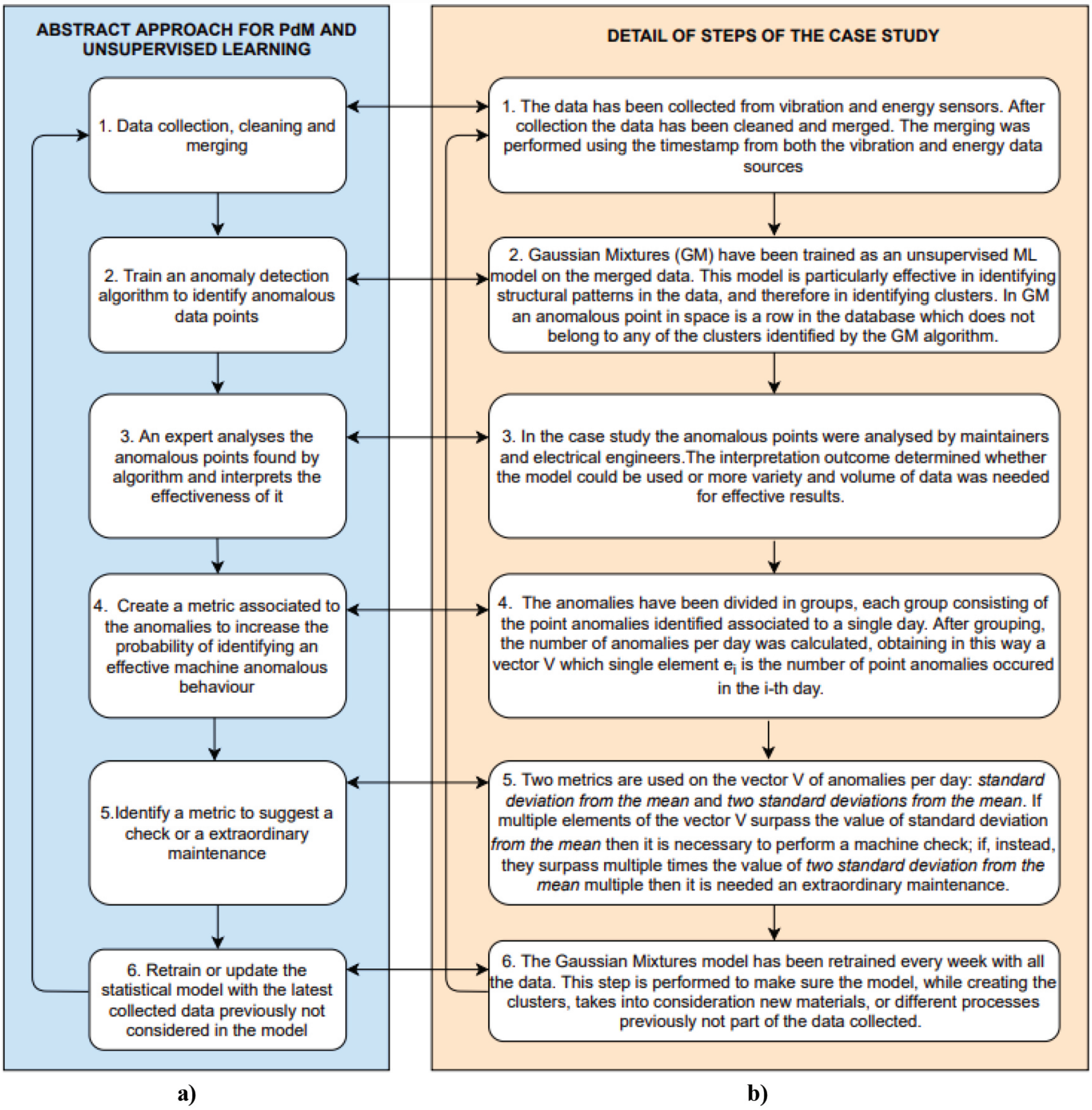


Figure 3. a) A framework for unsupervised learning; b) A framework instance that uses Gaussian Mixtures for predictive maintenance

Expectation-Maximization (EM) algorithm [44], a general method of finding the Maximum Likelihood Estimation (MLE) [46], which is used to estimate the parameters of an assumed probability distribution, given some observed data.

GMM can be used as an anomaly detection algorithm. In particular, an anomaly is a point in space that does not belong to any cluster. The anomalous points can be further processed to obtain a metric for the effective and probable anomalous behavior of an industrial machine.

The discussion of how GMM is used for predictive maintenance machine tools is further explained in the next section.

6. Case study

A case study was conducted in a mechanical company in the automotive sector to validate the effectiveness of the developed approach to predictive maintenance. Its implementation in GMM attempts to improve facility maintenance management activities, fitting an ML model able to support the prediction of degradation of a milling machine. A milling machine is used to remove metal from the workpiece with the help of a revolving cutter called a milling cutter. Due to its operating characteristics, we decided to analyze the energy, acceleration, and velocity variables, using sensors, since they are associated with some of the main reasons for failure in this type of machine.

The developed software stores and analyzes the data collected from the milling machine and visualizes potential energy and vibration anomalies, the latter as a possible consequence of alterations in acceleration or velocity. The method, models, and technology can be applied to other milling machines or lathes with little or no variation.

6.1 A predictive maintenance system model

This section describes the PdM system model designed to implement our approach. The prototype created as a PdM system is based on KITAMURA milling machine with CNC. Energy and vibration sensors have been applied to these machines for energy and vibration monitoring and analysis because they take care of measuring parameters that change according to modifications and disturbances in the mill process.

These sensors are connected to a local network, which becomes part of a CPS as shown in the quadrant a of the architecture in **Figure 2**. In turn, IoT techniques are used to retrieve the information collected from the sensors and send it securely to a centralized server.

The preventive maintenance application is served by a software component directly accessing the data server. This way, the data produced by the sensors (Sensor data logs) are first read and then processed to analyze the machine's performance. The Big Data Node-Red framework has been used to manage this networked data in daily and sensor-wise synchronized CSV files (see quadrant *b* of the architecture).

Once the massive data produced in the factory is rendered into a more concise and informative form, it is used to train and fit an anomaly detection algorithm, which identifies anomalies in the data. For the specific case of this study, the GMM was implemented according to the abstract steps of the framework of PdM in unsupervised learning. The data processing phase, presented in the next section, describes a particular instance of the abstract framework focused on the data analysis process.

Finally, the anomalies extrapolated from the data are represented graphically to the maintainer responsible for carrying out informed maintenance scheduling activities. In case of excessive acceleration, velocity, or energetical alterations during machine operation, the application triggers appropriate alerts as necessary which can be related to preventive checks or extraordinary maintenance.

6.2 PdM approach implementation

This section presents an application of the theoretical concepts discussed above, following the three major stages indicated in the Big data and IoT architecture of **Figure 2**: preprocessing (I), processing (II), and post-processing (III). In addition, it includes the abstract framework, which supports the data analysis process as part of stage II, where the GMM was used as the anomaly detection algorithm.

I. Data collection and preprocessing

As part of the data collection phase, industrial machine data in this research was collected from the CNC machine using energy and vibrational sensors. Among the numerous variables returned by these sensors, for this work, we have selected those described in **Figure 4**. The energy sensors collected the tension (Volts), current (Amperes), and power (Watts) data every 200 milliseconds, while the vibration sensor collected acceleration, velocity, and displacement data every 100 milliseconds.

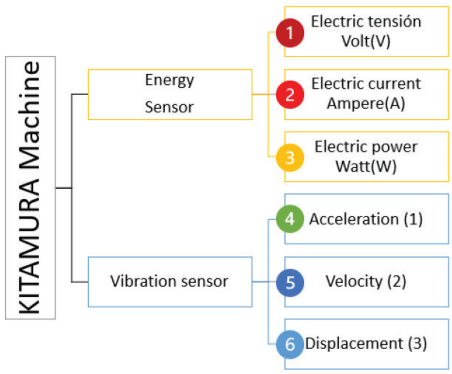


Figure 4. Data collected from a milling machine

Voltage and current values correspond to data from a three-phase power installation. We also have the value of average voltage, average current, and total power, respectively (Vsis, Asis, and PAtsis), and the timestamp. **Figure 5** shows an extract of the energy data log.

KI2-SENTOT-2021-04-08.txt

variable_type	phase	value	time
Volt;	V1N;	109.41230010986328;	1617860515395
Volt;	V2N;	109.64459991455078;	1617860515395
Volt;	V3N;	109.59760284423828;	1617860515395
Volt;	V12;	189.4940948486328;	1617860515395
Volt;	V23;	189.93141174316406;	1617860515395
Volt;	V31;	189.8134002685547;	1617860515395
Volt;	Vsis;	189.74876403808594;	1617860515395
Current;	A1;	16.780000686645508;	1617860515395
Current;	A2;	18.440000534057617;	1617860515395
Current;	A3;	16.200000762939453;	1617860515395
Current;	AN;	0;	1617860515395
Current;	Asis;	17.13999389648438;	1617860515395
Power;	PAt1;	1511.81005859375;	1617860515395
Power;	PAt2;	1546.239990234375;	1617860515395
Power;	PAt3;	1306.18994140625;	1617860515395
Power;	PAtsis;	4364.240234375;	1617860515395

Figure 5. Energy data log

On the other hand, **Figure 6** shows the data contained inside the acceleration, velocity, and displacement data log, where the machine state indicates whether the machine is running (1) or is stopped or turned off (0). As an example of the cleaning process, data not offering a significant contribution to the analysis are removed or not taken into account, such as those corresponding to machine state 0, since they do not generate any variation in the machine acceleration and velocity measurement.

The data has been obtained as a summary of the original signal for each of the vibration sensor variables (variable type: acceleration-1, velocity-2, and displacement-3). The original signal was divided into intervals, and each interval was summarized with the following values: maximum value, average of the highest values, minimum value, average of lower values, and timestamp.

The data collected by the energy and vibration sensors were sent to the data server on the local network. The data server is responsible for filtering and cleaning out incoming data from sensors while the machine is not running and organizing the data in CSV files related to a single day.

II. Data processing

To carry out the data analysis phase established by this stage, we first derived the abstract framework shown in the left side of **Figure 3** where the steps are independent of the decision of the algorithm and the evaluation metrics to be used. Then, the instance shown on the right side uses GMM and ML as one of the possible solutions. The choice of using GMM in our case study was given by the necessity of a fast algorithm able to perform accurately in a context characterized by high volumes of data. In particular, GM models represent data more accurately, compared to other clustering algorithms (such as K-means), as they incorporate information about the covariance structure of the data.

Once data was collected from different sources and stored on a centralized server as part of the data collection and preprocessing phase, it was organized into five directories: *Current, Power, Tension, Acceleration, velocity (Vibration), ElectricalAndVibrational*. Each directory represents a variable group (e.g., the set of variables V1N, V2N, V3N, V12, V23, V31 of a three-phase system) while, in the case of the ElectricalAndVibrational directory, it contains temporally synchronized data from power, acceleration, and velocity variable groups.

Unlabeled data in the directories has been used to create five different GM models whose goal is to indicate an electrical problem (current, voltage, or power anomaly), a mechanical problem (acceleration or velocity anomalies), or a problem related to an anomalous relationship between mechanical and electrical variables.

It is possible to identify data anomalies starting from the clusters of the five models given as output from the GMMs. A point is considered anomalous if it does not belong to any cluster. Specifically, a point *P* does not belong to any cluster when the probability of *P* belonging to each cluster is less than a threshold T. For example, if the threshold *T* is 40% and there are four clusters, then a point P is considered anomalous if, for each cluster, the probability of *P* belonging to it is less than 40%.

In most cases, an anomalous point does not define a faulty behavior of the machine or a failure; it only defines an anomalous instantaneous condition of the machine. For this reason, we propose an approach that exploits anomalies for predictive maintenance by using this simple assumption:

"The greater the number of anomalies of a variable group is present in a specific time interval, the greater the probability of a real failure occurring".

log-110-2021-04-08.txt

machine_state	time	variable_type	max_value	min_value	max_value_media	min_value_media
0;	1617857920;	1;	0,00;	0,00;	0,00;	0,00
1;	1617857954;	3;	2,00;	0,00;	2,00;	0,00
0;	1617857954;	0;	0,00;	0,00;	0,00;	0,00
1;	1617857964;	2;	4,00;	0,00;	4,00;	0,00
1;	1617857964;	3;	6,00;	0,00;	6,00;	0,00
0;	1617857964;	0;	1,00;	0,00;	1,00;	0,00
1;	1617857967;	3;	2,00;	0,00;	2,00;	0,00
0;	1617857967;	0;	1,00;	0,00;	1,00;	0,00
1;	1617857981;	3;	2,00;	0,00;	2,00;	0,00
0;	1617857981;	0;	1,00;	0,00;	1,00;	0,00
1;	1617857981;	2;	2,00;	0,00;	2,00;	0,00

Figure 6. Acceleration, velocity, and displacement data log

The data processing steps can be summarized as follows:

a) Data is stored in CSV format and synchronized temporally.
b) Five GM Models elaborate the clusters over five different variable groups (1. current, 2. power, 3. tension, 4. acceleration and velocity, and 5. power, acceleration, and velocity). Each model follows a training process having 25 mixture components (each component with its general covariance matrix), 200 iterations of the expectation-maximization algorithm performed and randomly initialized weights.
c) For each GM model, the anomalous points are identified at about a threshold T (in the test case, generally close to 40%), with T calculated heuristically as the threshold guaranteeing few anomalies in the data per day.
d) For each GM Model, the anomalies identified are summed over a specific time interval (e.g., one day).
e) Every week the models are retrained to ensure more reliable results and to be able to identify previously not encountered anomalies.

The fourth step represents the assumption and is necessary to create the appropriate graphical tools to support predictive maintenance. These graphical tools have been developed using the software described in the next section.

III. Data exploitation

As part of the PdM system, a software application was developed for data visualization, providing the maintainer with a graphical tool allowing a quick, clear understanding of the information. Thanks to graphic representation, we summarised the multivariate analysis of industrial machine data understandably and coherently, which helps to comprehend the information and establish preventive maintenance decisions.

The system can show information about anomalous events ordered and summed temporally in different time intervals (1 day, 3 days, 1 week, 2 weeks, and 1 month) and for different variable groups (current; power; tension; acceleration, and velocity; and power, acceleration, and velocity). As an example, **Figure 7a** shows the potential anomalies in a 1-day interval for the current variable group, while **Figure 7b** includes the combination of power, acceleration, and velocity; each point of the line drawn in the graph represents the number of anomalous data points, derived from the output of the GMM models, in a specific day. Therefore, the more the trend grows in the graph, the more it is necessary to consider the possibility of extraordinary maintenance or plan one shortly.

The user interface considers a simplified representation of the multimodal analysis of the GMMs, avoiding showing tables of n-dimensional data. This choice is coherent with the visual management scenario where visual stimuli communicate important information about a phenomenon at a glance, helping to convey relevant, easy-to-understand information [47].

The horizontal yellow line represents the maximum level that the number of anomalies can exceed before requiring a preventive check, while the red line indicates the need for extraordinary mainte-

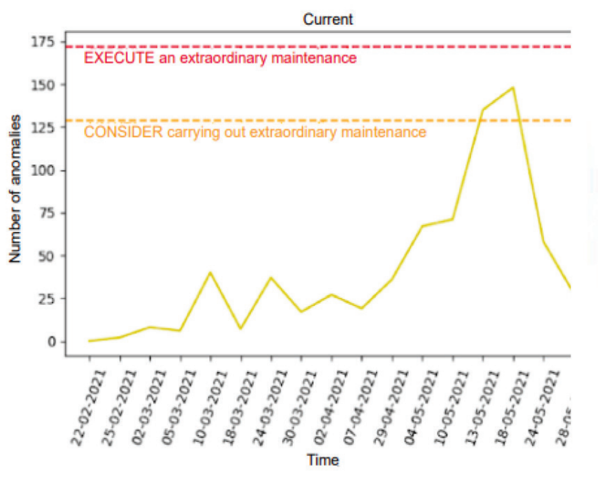


Figure 7a. Tension variable group resulting chart

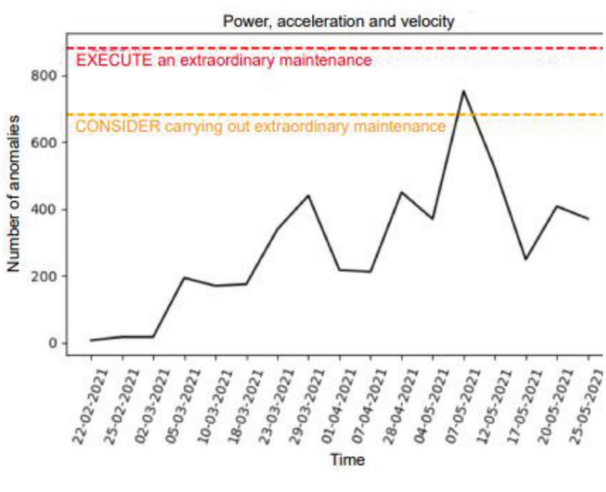


Figure 7b. Power, acceleration, and velocity combination resulting chart

nance. The position of the yellow line is given by one standard deviation from the mean, while that of the red line is by two standard deviations from the mean.

Standard deviations are a simple anomaly detection tool in single variable distributions and are used in this case for detecting a higher-than-normal number of anomalies that occurred in a single time interval (e.g., one day) compared to other time intervals.

To summarize, an excessive number of anomalies in a time interval is a probabilistic indication of a possible failure of the machine. Domain experts analyzed those patterns after the data visualization tool reported a condition worthy of attention from a potential machine failure standpoint. As the GMM–based system can identify unseen anomalous behaviors of the machine, the role of domain experts is to confirm that the anomaly identified by the system corresponds to a potential failure of the device. In other words, the system works as an early alerting system to prevent future costly damage. In addition, because a machine tool consists of mechanical, electrical, and electronic subsystems, the signal of probable impending failure obtained through GMs must be interpreted to identify the subsystem and the cause of the signal.

These tools have been used to suggest when to do maintenance, depending on which chart shows the yellow or red line surpassed, if the problem is electrical, mechanical, or if it is a problem caused by the interaction of the mechanical and electrical phenomena.

6.3 Preliminary results

The establishment of Condition-based maintenance (CbM) policies, as part of an effective preventive maintenance program, was performed based on the technological architecture implemented in the model shown in **Figure 2.** At the time of writing, the development concerns the two KITAMURA pilot machines object of experimentation. Still, shortly, all machine tools in the production department will be equipped to be part of the CPS. Experimental data have been collected from the start of February 2021 to the end of December 2022.

The scenarios we considered to use our framework refer to the different types of production: batch, mass, and engineer-to-order. In the case of mass and batch production, a model training in as little as a few weeks of data can produce an effective analysis. On the other hand, the situation is more complex in the engineer-to-order production described in the case study. In fact, in this type of production, the setup, tools, and raw materials can change even during the same day, which also leads to significant variations in the energy input power and the degree of vibration of the machine tool. In this case, to deal with the wide variety of scenarios, the unsupervised learning models were trained periodically to auto-adapt to the new and diverse data values.

Our approach allowed us to study the influence on the machine's behaviors of 1) a single variable, 2) a group of variables of the same type representing a machine behavior (e.g., the three current variables in a three-phase system), and 3) combined variables of different types (such as power and acceleration). This procedure is possible because the ML algorithm can be instructed to learn from a single variable type or a combination of variables. Moreover, in the case of potential machine anomalies, this feature provides selected information to an expert, which helps orient the analysis of possible causes of an imminent fault.

During this period, the system has shown an increasing trend of data anomalies provided as an output of the methods used on the KITAMURA machine four times.

<u>Case 1</u>: an anomaly has been signaled because of an anomalous trend of variables A1, A2, and A3, representative of current data. As shown in step 3 of the proposed framework (**Figure 3**), the expert's analysis has shown incorrect electrical wiring that has been fixed with a corrective maintenance operation.

<u>Case 2 and 3</u>: An increasing trend of the number of anomalies detected by the model analyzing power and acceleration variables was observed, indicating a potential failure in the imminent future. In both cases, it was decided to continue production and monitor the machine's status. In case 2, the machine KITAMURA1 has produced a not negligible number of defective pieces. Consequently, a mechanical engineer identified the cause of a spindle bearing about to break. In case 3, concerning the machine KITAMURA2, the expert identified tool wear as the cause of the anomalous behavior.

<u>Case 4</u>: In this case, the anomalous behavior signaled by the model has not been associated with any cause.

The results first indicate the system's effectiveness in detecting preventively anomalous machine behaviors and effectively transmitting this information to maintainers and production managers, who can decide to start a maintenance operation based on the system output if necessary.

7. Discussion

The main contribution of this work concerns the introduction of a framework for predictive maintenance based on unsupervised ML, a topic not sufficiently covered in the scientific literature. The data analysis was positively impacted by a PdM system developed for data acquisition, analysis, and visualization, which uses unsupervised ML models for early failure diagnosis and prognosis in situations with unlabeled data. Thus, the maintenance manager has a graphical tool that helps to better understand the behavior of the machines through the information collected by energy and vibrational sensors. This tool provides support for more effective preventive maintenance decisions.

Similar works have also studied predictive maintenance in unsupervised conditions. For example, Farbiz et al. [34] studied the correlation between robot controller data, energy data, and vibration data sources to detect machine anomalies. Our approach is different in that it allows the analysis of single or correlated variables with the advantage of providing information on the kind of failure. From Kim et al. [35], the equipment RUL prediction is based on the Health Index degradation curve of the equipment and a threshold when there is no historical data.

In additional works such as that of Del-Campo et al. [39], Zhao and Mata [48], Alaoui-Belghiti et al. [49], and Amruthnath and Gupta [24], dictionary learning methods to generate a set of anomaly scores, together with unsupervised learning algorithms are used to discover patterns and rules from historical data. Our work enriches an area, the PdM of machine tools in Industry 4.0 using unsupervised learning, not sufficiently covered in the literature.

In **Table 1**, the differences between our framework and other frameworks proposed in the literature are listed. The framework features are the following: A) unsupervised learning for predictive maintenance; B) temporally continuous strategies of auto-adapting the unsupervised learning model (e.g., via retraining strategies); C) machine-human feedback loop aimed at improving predictive maintenance for specific industrial contexts; D) specialized machine fault identification capabilities by using multiple unsupervised learning models per data source (e.g., one model for energy data, one for vibrational data and one including both sources); E) scalable retraining solutions dealing with Big Data issues that are typical of manufacturing industries.

This work was developed as part of a multi-year research and innovation programme aimed at achieving several objectives for the transition to Industry 4.0, including:

- The realisation of a total productive maintenance system by integration of Industry 4.0 into Total Productive Maintenance [50], with particular reference to the innovation proposed in this work.
- The realisation of a system for sustainability in manufacturing [51]
- The realisation of a software system, currently under development, for the intelligent and sustainable management of a supply chain in line with the results described in [52] for manufacturing supply chains. We also are considering the challenges entailed by the creation of new management systems in the Industry 5.0 scenario [53].

In this holistic view, the issue of machinery maintenance, in addition to the main benefits of decreasing the cost of keeping machinery in good shape and decreasing production downtime, also brings smaller but equally significant benefits in relation to production sustainability. Additional benefits concern reduction of waste and scrap due to malfunctioning machinery and energy savings as production stops caused by breakdowns result in unnecessary energy consumption in absence of production.

From a management point of view, the proposed solution has the following implications. Firstly, one

Table 1. Differences between our proposal and other frameworks existing in the literature

Work	Feature A	Feature B	Feature C	Feature D	Feature E
Our work	**x**	**x**	**x**	**x**	**x**
Farbiz et al. [34]	x	x		x	
Kim et al. [35]	x		x	x	
Del-Campo et al. [39]	x	x			
Zhao and Mata [48]	x				
Alaoui-Belghiti et al. [49]	x				
Amruthnath and Gupta [24]	x		x		

issue to be addressed when implementing predictive maintenance solutions using Industry 4.0 technologies concerns the cost of the solution. The manager must evaluate if it's worth the risk because the cost of the technology can be important. In most cases, there is already a cyber-physical system in place in the factory for planning, monitoring and controlling production, so the investment comes down to the software component alone.

A second aspect concerns the decision on how to allocate resources (hardware, software, people, time) to Total Productive Maintenance systems involving a mix of approaches to machine maintenance. It was found, during the development of the case study, that fewer resources were allocated to processes involving periodic machinery maintenance in favour of those involving predictive maintenance. This is a consequence of the advantages that the latter entails (e.g. it is not necessary to replace a machine component periodically if it is in good condition). The maintenance planner has therefore revisited the mix of approaches to Total Productive Maintenance in order to make the most of the advantages of using predictive maintenance.

Although this work is empirical, it raises some theoretical reflections. Referring to the picture shown in **Figure 3a**, the second step involves the use of an ML algorithm to identify anomalous states in the collected data. In **Figure 3b**, a GMM-type algorithm is suggested because it allows good flexibility in clustering and the treatment of multimodal data. These advantages make GMMs a powerful and flexible tool for anomaly detection. However, the framework is general and does not limit the choice of ML algorithm.

Algorithms that compete with GMM (e.g. K-Means Clustering) include various clustering, density estimation and unsupervised learning methods. The choice of the best algorithm often depends on the specific nature of the data and the problem to be solved. In the case of predictive maintenance, further studies are needed which can indicate in the various possible scenarios (applications on machine tools, in handling systems and in supply chains) which are the best choices with respect to the problem to be solved.

8. Conclusions

Diagnosing and prognostics are two crucial aspects of preventive maintenance programs. Prognostics are usually applied to achieve zero-downtime performance through prediction, whereas diagnostics are required when faults occur. To improve the failure prediction capabilities of machine tools, the research achieved the following findings:

1. An architectural model designed specifically for preventive and predictive maintenance, to optimize the maintenance of industrial machinery.
2. A new abstract framework for predictive maintenance, which uses an unsupervised machine learning algorithm to improve the efficiency and accuracy of maintenance processes.

The abstract framework gives a general perspective allowing the use of different unsupervised ML for anomaly detection, which means that the framework does not depend on a specific algorithm. Additionally, our model is automatically retrained to improve the accuracy of the results.

To validate the framework in a real-world application, a cyber-physical architecture was implemented together with a framework instance that uses ML and a GMM–based system for the PdM of machine tools. The application is operational in an Italian automotive manufacturing industry company with encouraging results from the point of view of the effectiveness of the implemented solution.

Following the experimentation carried out over a two-year period in which data were collected, analyzed, and evaluated, it was possible to make the following observations on the research questions:

Implementing unsupervised machine learning algorithms in the project has proven effective within the experimental timeframe. These algorithms successfully identified anomalies in the machinery, leading to necessary emergency maintenance activities, as demonstrated in the case study. In this sense, the results obtained empirically allow us to conclude positively for RQ1, that unsupervised learning models effectively identify potential causes of machine failures and predict future failures in industrial environments.

Similarly, in response to RQ2, the decision-making processes in maintenance programs have significantly improved due to the visual management enabled by the developed PdM system. The results generated by this system are useful for various roles in the maintenance process, including operators, maintenance planners, and the maintenance team. The damage hypotheses provided by the algorithm facilitate more autonomous and timely decision-making. This is because the system, supported by Big

Data and IoT architecture, enables the identification of anomalies according to the visual management paradigm, allowing for effective informed decisions.

A limitation of this work certainly stems from the fact that expert advice is needed to validate the automatic reporting and interpret the data to identify which machine tool component is likely to fail in the near future. However, unsupervised learning solves the problem of predictive maintenance in all those scenarios, which is widespread in the manufacturing environment, where historical data on past failures have not been collected that could enable supervised learning and a consequent reduction in the effort of the maintenance expert.

This research can be developed in several directions. First of all, the advent of Industry 5.0 poses new questions about the role to be played by humans interacting with automated systems to improve monitoring and control processes. This is considered increasingly important in future research due to the knowledge learned by machine learning cannot win human domain knowledge [54].

This topic needs to be explored in machinery maintenance, as further improvements may result from the harmonious interaction of maintenance systems with human roles involved in maintenance activities (planners, machine workers, maintenance teams).

A second line of research concerns the development of graphical interfaces suitable to support the visual management paradigm [55] for maintenance processes. The usability of such interfaces is essential to enable an operator to promptly and effectively identify and react to potentially critical situations reported by visual software tools. The design and implementation of effective GUIs can benefit from the study proposed in [56].

Finally, future development of this work will consider integrating the designed IoT and Big Data architecture with different ML algorithms, such as those for prescriptive analysis [57]. This way, the maintenance programs could be optimized by developing advanced recommendation systems, which help suggest the actions to undertake during maintenance.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] C. Chen, N. Lu, B. Jiang, and C. Wang, “A Risk-Averse Remaining Useful Life Estimation for Predictive Maintenance,” IEEE CAA J Autom Sin, vol. 8, no. 2, pp. 412–422, 2021, doi: 10.1109/JAS.2021.1003835.

[2] M. Jasiulewicz-Kaczmarek, K. Antosz, P. Żywica, D. Mazurkiewicz, B. Sun, and Y. Ren “Framework of machine criticality assessment with criteria interactions,” Eksploat Niezawodn - Maint Reliab, vol. 23, pp. 207–220, 2021, doi: 10.17531/ein.2021.2.1.

[3] C. K. M. Lee, Y. Cao, and K. H. Ng, “Big Data Analytics for Predictive Maintenance Strategies,” in Supply Chain Management in the Big Data Era, H. K. Chan, N. Subramanian, and M. D-A. Abdulrahman, Eds. New York, NY, USA: IGI Global, 2017, pp. 50-74, doi: 10.4018/978-1-5225-0956-1.

[4] A. K. S. Jardine, D. Lin, and D. Banjevic. “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” Mech Syst Signal Process, vol. 20, no. 7, pp. 1483–1510, 2006, doi: 10.1016/j.ymssp.2005.09.012.

[5] V. L. Trevathan, A Guide to the Automation Body of Knowledge. Research Triangle Park, NC, USA: ISA—The Instrumentation, Systems, and Automation Society, 2006.

[6] J.-R. Ruiz-Sarmiento, J. Monroy, F.-A. Moreno, C. Galindo, J.-M. Bonelo, and J. Gonzalez-Jimenez, “A predictive model for the maintenance of industrial machinery in the context of industry 4.0,” Engineering Applications of Artificial Intelligence, vol. 87, p. 103289, 2020, doi: 10.1016/j.engappai.2019.103289.

[7] E. Gundogar, A. Yilmaz, and B. Erkayman, “A solution approach to a synchronisation problem in a JIT production system,” Prod Plan Control, vol. 25, pp. 990–998, 2014, doi: 10.1080/09537287.2013.794984.

[8] B. Ji, H. Park, K. Jung, S. H. Bang, M. Lee, J. Kim, and H. Cho, “A Component Selection Method for Prioritized Predictive Maintenance,” in Advances in Production Management Systems (APMS 2017), The Path to Intelligent, Collaborative and Sustainable Manufacturing, IFIP Advances in Information and Communication Technology, vol 513, H. Lödding, R. Riedel, K. D. Thoben, G. von Cieminski, and D. Kiritsis, Eds. 2017, pp. 433–440, doi: 10.1007/978-3-319-66923-6_51.

[9] A. Busse, J. Metternich, and E. Abele, “Evaluating the Benefits of Predictive Maintenance in Production: A Holistic Approach for Cost-Benefit-Analysis,” in Advances in Production Research. WGP 2018, R. Schmitt, and G. Schuh, Eds. 2019, pp. 690–704, doi: 10.1007/978-3-030-03451-1_67.

[10] U. Moorthy, and U. D. Gandhi, “A Survey of Big Data Analytics Using Machine Learning Algorithms,” in HCI Challenges and Privacy Preservation in Big Data Security, D. Lopez, and M. A. Saleem Durai, Eds. New York, NY, USA: IGI Global, 2017, pp. 95–123, doi: 10.4018/978-1-5225-2863-0.ch005.

[11] G. Kabir, S. Tesfamariam, J. Loeppky, and R. Sadiq, “Predicting water main failures: A Bayesian model updating approach,” Knowl-Based Syst, vol. 110, pp. 144–156, 2016, doi: 10.1016/j.knosys.2016.07.024.

[12] J-H. Shin, and H-B. Jun, “On condition-based maintenance policy,” J Comput Des Eng, vol. 2, pp. 119–127, 2015, doi: 10.1016/j.jcde.2014.12.006.

[13] H.-Y. Wang, and C.-H. Hung, “Complex Industrial Machinery Health Diagnosis Challenges and Strategies,” in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 557,

D.-J. Deng and J.-C. Chen, Eds. 2024, pp. 130-140, doi: 10.1007/978-3-031-55976-1_13.
[14] H. Wang, W. Zhang, D. Yang, and Y. Xiang, "Deep-Learning-Enabled Predictive Maintenance in Industrial Internet of Things: Methods, Applications, and Challenges," IEEE Systems Journal, vol. 17, no. 2, pp. 2602-2615, 2023, doi: 10.1109/JSYST.2022.3193200.
[15] P. Nunes, J. Santos, and E. Rocha, "Challenges in predictive maintenance – A review," CIRP Journal of Manufacturing Science and Technology, vol. 40, pp. 53-67, 2023, doi: 10.1016/j.cirpj.2022.11.004.
[16] H. Hviid Hansen, M. Kulahci, and B. Friis Nielsen, "A primer on predictive maintenance: Potential benefits and practical challenges," Quality Engineering, vol. 36, no. 3, pp. 638-649, 2024, doi: 10.1080/08982112.2024.2331140.
[17] N. H. A. Wahab, K. Hasikin, K. W. Lai, K. Xia, L. Bei, K. Huang, and X. Wu, "Systematic review of predictive maintenance and digital twin technologies challenges, opportunities, and best practices," PeerJ Comput Sci, vol. 10, p. e1943, 2024, doi: 10.7717/peerj-cs.1943.
[18] M. Alabadi, A. Habbal, and M. Guizani, "An Innovative Decentralized and Distributed Deep Learning Framework for Predictive Maintenance in the Industrial Internet of Things," IEEE Internet of Things Journal, vol. 11, no. 11, pp. 20271-20286, 2024, doi: 10.1109/JIOT.2024.3372375.
[19] R. Kumar, M. Mishra, S. Suman, and P. Singh Bali, "Predictive Maintenance in Industrial Systems Using Machine Learning," International Journal of Innovative Science and Research Technology, vol. 9, no. 3, pp. 1778-1785, 2024, doi: 10.38124/ijisrt/IJISRT24MAR1367.
[20] P. Poór and J. Basl, "Machinery maintenance model for evaluating and increasing maintenance, repairs and operations within Industry 4.0 concept," IOP Conf. Ser.: Mater. Sci. Eng., vol. 947, no. 1, p. 012004, 2020, doi: 10.1088/1757-899X/947/1/012004.
[21] J-Y. Chen, Y-L. Lin, and B-Y. Lee, "Development of the Adaptive System for Tool Management," Tehnicki Vjesnik, vol. 30, no. 2, pp. 648-654, 2023, doi: 10.17559/TV-20220702054333.
[22] E. Salawu et al., "Impact of Maintenance on Machine Reliability: A Review," E3S Web of Conferences, vol. 430, 2023, doi: 10.1051/e3sconf/202343001226.
[23] R. J. Rabelo, S. P. Zambiasi, and D. Romero, "Softbots 4.0: Supporting Cyber-Physical Social Systems in Smart Production Management," Int J Ind Eng Manag, vol. 14, no. 1, pp. 63 – 93, 2023, doi: 10.24867/IJIEM-2023-1-325.
[24] N. Amruthnath and T. Gupta. "A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance," in 2018 5th International Conference on Industrial Engineering and Applications (ICIEA). Singapore: IEEE, 2018, pp. 355–361, doi: 10.1109/IEA.2018.8387124
[25] J. Lee, E. Lapira, S. Yang, and A. Kao, "Predictive Manufacturing System - Trends of Next-Generation Production Systems," IFAC Proc, vol. 46, no. 7, pp. 150–156, 2013, doi: 10.3182/20130522-3-BR-4036.00107.
[26] H. Li and Q. Zhao. "Maintenance Modeling and Scheduling in Fault Tolerant Control Systems," in Proceedings from the 6th IFAC Symposium, SAFEPROCESS 2006, Beijing, China, 2007, pp. 777-782, doi: 10.1016/B978-008044485-7/50131-7.
[27] M. G. Deighton, "Maintenance Management," in Facility Integrity Management, M. G. Deighton, Ed. Boston: Gulf Professional Publishing, 2016, pp. 87–139, doi: 10.1016/B978-0-12-801764-7.00005-X.
[28] G. Nota, A. Postiglione, and R. Carvello, "Text mining techniques for the management of predictive maintenance," Procedia Computer Science, vol. 200, pp. 778-792, 2022, doi: 10.1016/j.procs.2022.01.276.
[29] R. Langone, C. Alzate, B. De Ketelaere and J. A. K. Suykens, "Kernel spectral clustering for predicting maintenance of industrial machines," 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), Singapore, 2013, pp. 39-45, doi: 10.1109/CIDM.2013.6597215.
[30] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, Least Squares Support Vector Machines. Singapore: World Scientific Publishing, 2002, doi: 10.1142/5089.
[31] D. Rafique and L. Velasco, "Machine learning for network automation: overview, architecture, and applications," J Opt Commun Netw, vol. 10, no. 10, pp. 126–143, 2018, doi: 10.1364/JOCN.10.00D126.
[32] N. Amruthnath and T. Gupta. "Fault class prediction in unsupervised learning using model-based clustering approach," in 2018 International Conference on Information and Computer Technologies (ICICT), DeKalb, IL, USA: IEEE Xplore, 2018, pp. 5–12, doi: 10.1109/INFOCT.2018.8356831.
[33] Y. Bao, G. Rui, and S. Zhang, "A Unsupervised Learning System of Aeroengine Predictive Maintenance Based on Cluster Analysis," in Proceedings of the 2020 International Conference on Aviation Safety and Information Technology. Weihai City, China: ACM, 2020, pp. 187–191, doi: 10.1145/3434581.3434619.
[34] F. Farbiz, Y. Miaolong, and Z. Yu, "A Cognitive Analytics based Approach for Machine Health Monitoring, Anomaly Detection, and Predictive Maintenance," in 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2020, pp. 1104–1109, doi: 10.1109/ICIEA48937.2020.9248409.
[35] D. Kim, S. Lee, and D. Kim, "An Applicable Predictive Maintenance Framework for the Absence of Run-to-Failure Data," Appl Sci, vol. 11, no. 11, p. 5180, 2021, doi: 10.3390/app11115180.
[36] P. Dayan, M. Sahani, and G. Deback. "Adaptation and Unsupervised Learning," in Advances in Neural Information Processing Systems, S. Becker and S. Thrun and K. Obermayer, Eds. 2022, pp. 237-244.
[37] T. Wuest, D. Weimer, C. Irgens, and K-D. Thoben, "Machine learning in manufacturing: advantages, challenges, and applications," Prod Manuf Res, vol. 4, no. 1, pp. 23-45, 2016, doi: 10.1080/21693277.2016.1192517.
[38] S. H. Bang, R. Ak, A. Narayanan, Y. T. Lee, and H. Cho, "A survey on knowledge transfer for manufacturing data analytics," Comput Ind, vol. 104, pp. 116-130, 2019, doi: 10.1016/j.compind.2018.07.001.
[39] S. Martin-del-Campo, and K. Al-Kahwati, "Unsupervised Ranking of Outliers in Wind Turbines via Isolation Forest with Dictionary Learning," PHM Society, vol. 5, no. 1, p. 9, 2020, doi: 10.36001/phme.2020.v5i1.1164.
[40] L. Monostori, "AI and machine learning techniques for managing complexity, changes and uncertainties in manufacturing," Eng Appl Artif Intell, vol. 16, no. 4, pp. 277-291, 2003, doi: 10.1016/S0952-1976(03)00078-2.
[41] J. Lenz, T. Wuest, and E. Westkämper, "Holistic approach to machine tool data analytics," J Manuf Syst, vol. 48, pp. 180-191, 2018, doi: 10.1016/j.jmsy.2018.03.003.
[42] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Comput Netw, vol. 54, no. 15. pp. 2787–2805, 2010, doi: 10.1016/j.comnet.2010.05.010.
[43] R. Casado and M. Younas, "Emerging trends and technologies in big data processing," Concurr Comput Pract Exp, vol. 27, no. 8, pp. 2078–2091, 2014, doi: 10.1002/cpe.3398.
[44] J. Moradi, H. Shahinzadeh, H. Nafisi, M. Marzband, and G. B. Gharehpetian, "Attributes of Big Data Analytics for

Data-Driven Decision Making in Cyber-Physical Power Systems," in 14th International Conference on Protection and Automation of Power Systems (IPAPS), Tehran, Iran, 2019, pp. 83-92, doi: 10.1109/IPAPS49326.2019.9069391.
[45] M. H. ur Rehman, V. Chang, A. Batool, and T. Y. Wah, "Big data reduction framework for value creation in sustainable enterprises," Int J Inf Manag, vol. 36, no. 6, pp. 917–928, 2016, doi: 10.1016/j.ijinfomgt.2016.05.013.
[46] D. Reynolds, "Gaussian Mixture Models," in Encyclopedia of Biometrics, S. Z. Li and A. Jain, Eds. Boston, MA, USA: Springer, 2009, pp. 659-663, doi: 10.1007/978-0-387-73003-5_196.
[47] J. Liu, D. Cai, and X. He, "Gaussian Mixture Model with Local Consistency," in Proceedings of the AAAI Conference on Artificial Intelligence, Atlanta, GA, USA, 2010, doi: 10.1609/aaai.v24i1.7659.
[48] Y. Zhao and G. E. Mata, "Leverage Artificial Intelligence to Learn, Optimize, and Win (LAILOW) for the Marine Maintenance and Supply Complex System," in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2020, pp. 678–684, doi: 10.1109/ASONAM49781.2020.9381319.
[49] A. Alaoui-Belghiti, S. Chevallier, and E. Monacelli, "Unsupervised Anomaly Detection Using Optimal Transport for Predictive Maintenance," in Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series. ICANN 2019, I. V. Tetko, V. Kůrková, P. Karpov, and F. Theis F, Eds. 2019. pp. 686–697, doi: 10.1007/978-3-030-30490-4_54.
[50] G. L. Tortorella, F. S. Fogliatto, P. A. Cauchick-Miguel, S. Kurnia, and D. Jurburg, "Integration of Industry 4.0 technologies into Total Productive Maintenance practices," Int J Prod Econ, vol. 240, p. 108224, 2021, doi: 10.1016/j.ijpe.2021.108224.
[51] G. Nota and A. Toro Lazo, "Leveraging the GQM+ Strategy approach and Industry 4.0 technologies for environmental sustainability in manufacturing," Journal of Smart Environments and Green Computing, vol. 2, no. 3, pp. 143-162, 2022, doi: 10.20517/jsegc.2022.13.
[52] Oubrahim, N. Sefiani, and A. Happonen, "The Influence of Digital Transformation and Supply Chain Integration on Overall Sustainable Supply Chain Performance: An Empirical Analysis from Manufacturing Companies in Morocco," Energies, vol. 16, no. 2, p. 2, 2023, doi: 10.3390/en16021004.
[53] C. L. Karmaker, A.B.M. Mainul Bari, Md. Z. Anam, T. Ahmed, S. M. Ali, D. A. de Jesus Pacheco, and Md. A. Moktadir, "Industry 5.0 challenges for post-pandemic supply chain sustainability in an emerging economy," Int J Prod Econ, vol. 258, p. 108806, 2023, doi: 10.1016/j.ijpe.2023.108806.
[54] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, and L. He, "A survey of human-in-the-loop for machine learning," Future Generation Computer Systems, vol. 135, pp. 364-381, 2022, doi: 10.1016/j.future.2022.05.014.
[55] G. Fenza, V. Loia, and G. Nota, Patterns for Visual Management in Industry 4.0," Sensors, vol. 21, no. 19, p. 6440, 2021, doi: 10.3390/s21196440.
[56] M. Di Gregorio, G. Nota, M. Romano, M. Sebillo, and G. Vitiello, "Designing usable interfaces for the Industry 4.0", in Proceedings of the 2020 International Conference on Advanced Visual Interfaces, AVI '20, New York, NY, USA: Association for Computing Machinery, 2020, pp. 1-9, doi: 10.1145/3399715.3399861.
[57] Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: Literature review and research challenges," Int J Inf Manag, vol. 50, pp. 57–70, 2020, doi: 10.1016/j.ijinfomgt.2019.04.003.