

# Intelligent Choice of Machine Learning Methods for Predictive Maintenance of Intelligent Machines

Marius Becherer\*, Michael Zipperle† and Prof. Dr. Achim Karduck‡

*Furtwangen University*

---

Machines are serviced too often or only when they fail. This can result in high costs for maintenance and machine failure. The trend of Industry 4.0 and the networking of machines opens up new possibilities for maintenance. Intelligent machines provide data that can be used to predict the ideal time of maintenance. There are different approaches to create a forecast. Depending on the method used, appropriate conditions must be created to improve the forecast. In this paper, results are compiled to give a state of the art of predictive maintenance. First, the different types of maintenance and economic relationships are explained. Then factors for the forecast are explained. Requirements for the data are collected and algorithms for machine learning are presented. Based on the relationships found, a process model is presented that shows a fast implementation of the predictive maintenance for machines.

Keywords: Predictive Maintenance, Machine Learning, Artificial Intelligence, Smart Machines, Industry, Process Model

---

## 1. INTRODUCTION

Maintenance is an integral part of the machine life cycle, checking the machine status and repairing defects. In this way, service life can be extended, and maintenance repair costs reduced. Especially with large machines, high costs for extensive maintenance work and machine failures can be the result. For this reason, needs-based maintenance is essential to minimize production downtime and avoid high maintenance costs due to machine component failures. The company Trenitalia invests approx. 1.3 billion euros for the maintenance of its trains [1]. At Airbus, there are different levels of maintenance for the aircraft, but the maintenance of an aircraft can take several weeks and thus cause high costs [2]. Predictive maintenance is used to reduce costs and maintain efficiency.

The predictive maintenance should determine the ideal time for maintenance and thus increase the availability of the aircraft.

Data is needed to create a prediction. Therefore, an infrastructure that collects and stores data of a machine by sensors is needed. Specific data requirements arise during the prediction generation. For example, these must correspond to the context of the application case and, on the other hand, useless or incorrect data should be excluded. The data then serves as the basis for a selected algorithm, which provides the forecast for the maintenance time of a machine. There are different algorithms for a forecast. For predictive maintenance, the main algorithms used are those of machine learning. These have different characteristics and requirements. Which algorithm is most suitable depends on the application.

This article will present the possibilities of predictive maintenance and show how this process is structured and can be implemented. For this purpose, section II first references related works, which reflect the importance of predictive maintenance. Then, in section III, the theoretical basics of maintenance are

---

\*marius.baech@gmail.com

†michael@zipperle.de

‡karduck@hs-furtwangen.de

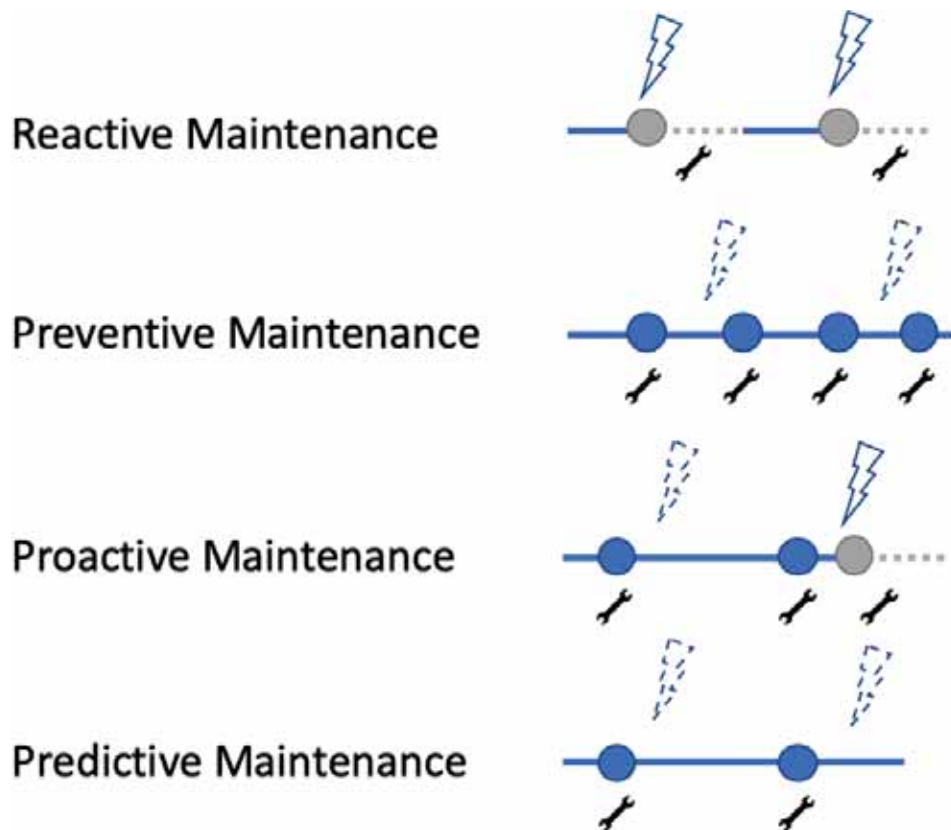


Figure 1 Maintenance times of the different stages.

discussed, introducing the different levels of maintenance and their economic impact. Afterwards, section IV explains the process of predictive maintenance in more detail. The sub-processes of data collection, data pre-processing, algorithms and evaluation are examined in detail. Finally, in section V, the connection between data, algorithm and use case is explained. Based on these relationships, a procedural model is presented, which shows the process for a quick implementation of predictive maintenance. At the end of section VI, the gained knowledge is summarized, and an outlook on further development is given.

## 2. RELATED WORK

There is an excellent potential for optimization in the maintenance of machines. By using predictive maintenance, the availability of a machine can be increased by up to 15% and costs can be reduced by up to 25% [3]. The report “Predictive Maintenance - predict the unpredictable” shows that only a small percentage of them predict the time of maintenance. The report also provides introductory information on predictive maintenance [4]. First of all, different maintenance levels and areas of application are presented. This includes a survey result regarding predictive maintenance. Besides, two use cases in the area of mobility and chemistry are described.

Further current figures on economic efficiency are contained in a report by BHGE [5]. This shows for the oil recovery application that predictive maintenance reduces the failure rate and saves costs compared to other maintenance levels. Although predictive maintenance has the lowest failure rate, it is not necessarily the best alternative. There are advantages and disadvantages to the

different maintenance levels [6]. Depending on the use case, the arguments can be weighted more or less heavily.

Various factors are essential for the implementation of predictive maintenance, such as IoT and significant data Architecture for predictive maintenance [7], the data and the selected algorithm [8]. For the algorithms, supervised and unsupervised learning is used. These also have different strengths and weaknesses [9]. Because of the different data and use cases, different algorithms from the field of machine learning can be used [10]. To implement predictive maintenance with less data is a challenge. The research deals with the transfer of use cases [11].

Thereby algorithms of already existing use cases are used, which are used in the same or similar contexts.

## 3. FUNDAMENTALS

For a prediction, data are required. Intelligent machines use sensors to record current characteristic values that describe the condition of a machine. Often the term digital twin is used, which represents the machine in a digital image. As shown in Fig. 1, maintenance can be divided into different maintenance levels. The purest form is reactive maintenance. Here, no planning or further action takes place, but rather a reaction is made when the machine fails. This leads to unplanned breakdowns and can cause significant damage because smaller defects are not repaired early enough.

The next stage is preventive maintenance. Maintenance takes place at fixed time intervals. Resource planning can be carried out at an early stage, but maintenance costs increase because

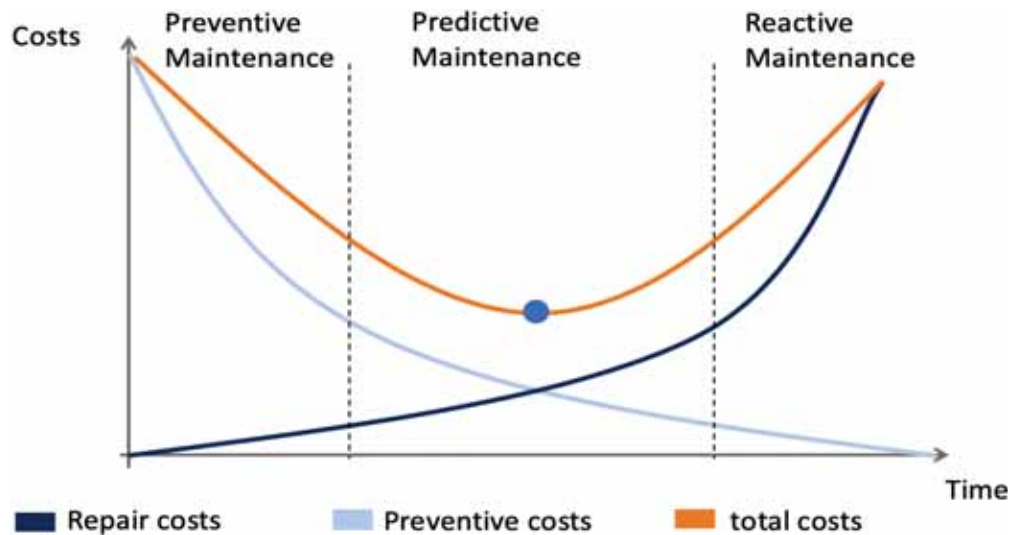


Figure 2 The optimum maintenance time.

maintenance can take place too often. However, maintenance can also take place too rarely, which increases repair costs. This means that preventive maintenance does not allow for needs-based maintenance. With proactive maintenance, the machine is monitored in real-time. Various events can be triggered when specific values are exceeded. This principle allows small errors to be fixed, thus preventing critical errors. Here, too, unplanned breakdowns can occur, since even small errors lead to machine downtimes and these are not detected in advance. To be able to carry out proactive maintenance, sensors must be installed on the machine to be inspected so that the current status can be recorded. The highest level of maintenance is predictive maintenance. A forecast is created based on a data basis. The forecast is intended to provide timely information about necessary maintenance before the machine breaks down. Furthermore, the forecast can be used to optimize resource planning. For example, replacement machines can be provided, or spare parts can be ordered in advance.

Two types of costs can be distinguished for maintenance: preventive and repair costs, which are shown in Fig. 2. Preventive costs are high at the beginning, as the machine is not operated during permanent maintenance. If there is no maintenance, the preventive costs are reduced. On the other hand, there are repair costs. These are low at the beginning because the material wear of machine parts takes place over time. The longer the machine is in operation, the greater the probability of major breakdowns and increased costs. Finally, it is important to find the minimum maintenance costs. By adding the preventive and repair costs, the total cost curve can be established. The ideal maintenance time cannot be determined by time or events but must be forecast.

Many companies maintain reactively and preventively, as shown in Fig. 3. Proactive and predictive maintenance requires an infrastructure that captures behaviour and current data from machine sensors. Depending on the machine, reactive and preventive maintenance is more economical than investing in infrastructure. Ultimately, for an assessment, the additional downtime costs must be compared to the investment costs.

An example that predictive maintenance leads to a reduction in costs and failure rate is shown in Fig. 4. These developments

are documented in the oil and gas industry. The unplanned outages were reduced by half with the predictive maintenance compared to reactive and preventive maintenance. Furthermore, the costs for maintenance were halved since the prediction could determine the ideal time for maintenance. This not only reduced unplanned outages but also saved costs.

## 4. PROCESS

In the following, the process chain for implementing the predicted maintenance of a machine is explained. Predictive maintenance is a complex process that can be divided into the following four sub-processes:

- 1) Data collection
- 2) Data preprocessing
- 3) Algorithm
- 4) Evaluation

These sub-processes build on each other so that the result of one process step influences the quality of the next process step. If, for example, incorrect data is captured, this affects the result of the algorithm. Furthermore, the design of the individual sub-processes depends on the respective application. This determines which data can be collected and which results the forecast should deliver. It is, therefore, essential that the maintenance requirements are defined first. These requirements can then be used to determine individual sub-processes. In the following sections, individual sub-processes are described in more detail.

### 4.1 Data Collection

This sub-process forms the basis of the predictive maintenance because no meaningful forecast can be made without data from a machine. Data can be collected from different sources,

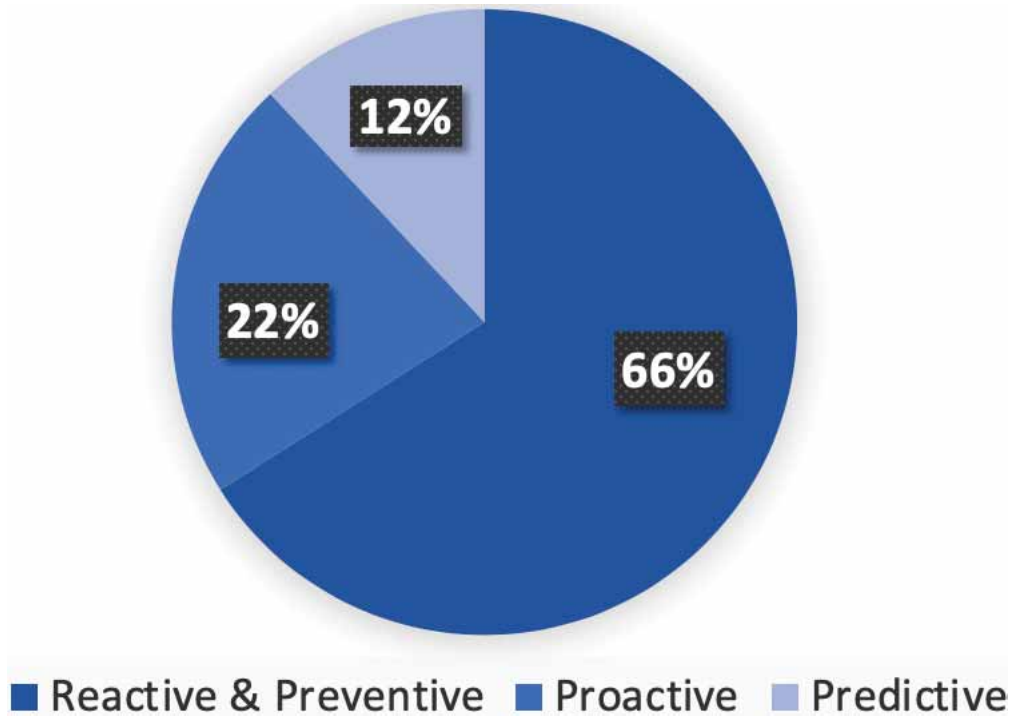


Figure 3 Used maintenance methods [4].

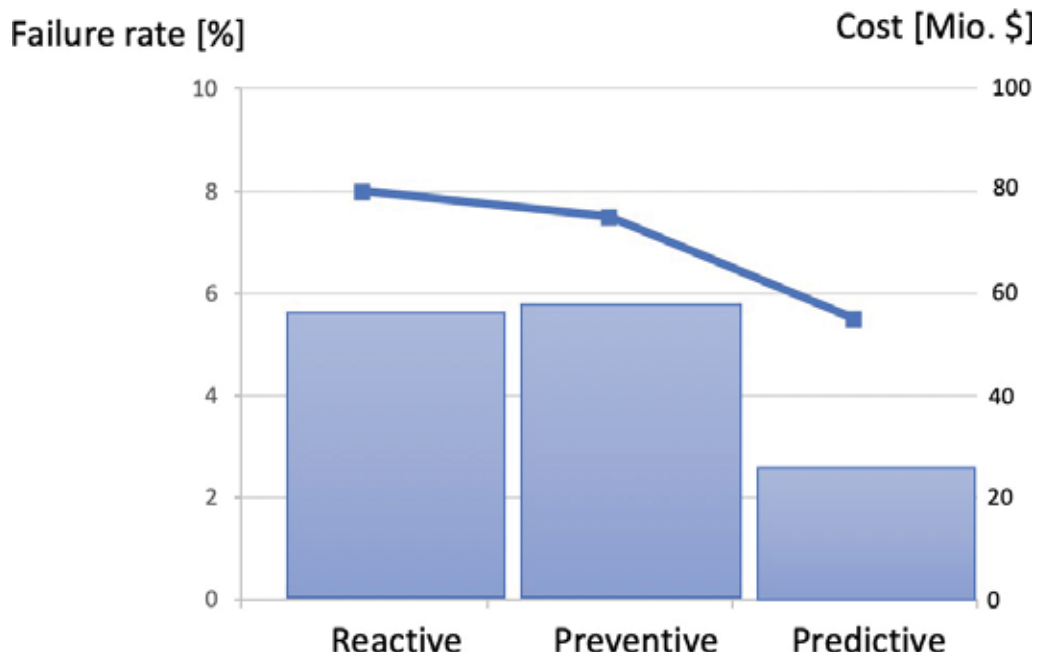


Figure 4 Maintenance methods used [5].

which may differ depending on the application. The sources provide either temporary data, which serves to test and compare the machine condition or static data, which form the basis for comparison. The following sources can be used for data collection:

- Temporary data sources:
  - Operating status of the machine: The sensors installed in a machine provide metrics about the current status. This metrics such as the temperature or speed of an engine can be shown on display. Machines contain

various sensors that are used to monitor the functioning of the machine. A key assumption in predicted maintenance is that the condition of a machine will deteriorate over time in routine operation. The data contains time-dependent features that detect this aging pattern and anomalies that lead to a deterioration in machine condition. The time aspect of the data is necessary for the algorithm to predict errors [8].

- Monitoring sensors: These types of sensors are installed subsequently and are used for machine

monitoring. For example, acoustic sensors can be used to analyze the noise development of a machine. If the noise development does not correspond to the normal behaviour, a possible fault condition of the machine may exist [4].

- Static data sources:
  - Historical maintenance logs: Maintenance logs exist for machines that were previously maintained manually by an inspector. These maintenance logs document the machine status at a specific point in time and contain details of replaced components and repair activities performed. Maintenance logs provide much information, especially about the behaviour of a machine, and are therefore an important source of data for predictive maintenance [4].
  - Fault logs: Machines provide fault logs in which occurred faults are documented. These can be used to determine hardware or software errors.
  - Machine metadata: The metadata of a machine includes the manufacturer, model, date of manufacture, date of commissioning, location of the system and other technical specifications such as throughput and errors per day [8].

Thus there are several possibilities to collect machine data. However, not all data is the same. The accuracy of a forecast for maintenance depends on the relevance, quantity, and quality of the data. This results in the following additional requirements for them:

- **Relevance:** The data must be relevant to the problem to be solved. This means that a complex machine cannot be considered as a whole system. It is, therefore, necessary to divide a machine into components and to examine these individually. A domain expert can divide a system into components. For example, in a motor vehicle, the engine temperature cannot be used for the predictive maintenance of the transmission [8].
- **Quality:** There are numerous sources of error in data collection, which can ultimately lead to insufficient quality of the data for predictive maintenance. It is therefore important to avoid these sources of error; methods for doing so are explained in the next section. Possible sources of errors are listed below:
  - **Error in data entry:** In many cases, data is still entered by humans or transmitted by voice. This can lead to typing errors or misunderstandings during communication [12].
  - **Measurement error:** Measured values are manually measured and read by humans. Errors may occur when applying the measurement method or reading the measured value. However, these errors can be avoided by using sensors and automated data storage. Nevertheless, measurement errors can also occur with sensors, for example, if several sensors interfere with each other or are influenced by the environment [12].

- **Distillation error:** In many machines, data from sensors are processed and combined before being stored in a database. This reduces the complexity of the data and the noise of the sensor data. However, this process can also result in errors or the aggregation of data that is particularly important for predicted maintenance [12].

- **Error in data integration:** The collected data comes from different sources, which are stored in a database at different times using different methods. Furthermore, in practice, it may be necessary to combine data from different sources. When merging, inconsistencies have to be resolved, and errors may occur. Individual data sets may differ in their units and measurement period [12].
- **Quantity:** There is no general answer to the question of how much data is needed for the predicted maintenance. The quantity required depends on the application case or the context of the problem to be solved. For predictive maintenance, it is important to have data sets where the machines are in unexpected operating states. If, for example, only data records of a machine in ideal condition (without errors) are available, predictive maintenance cannot be performed based on this data. Therefore, data records where a machine goes into a fault condition are particularly important [8].

This section described the data sources, the requirements for these and possible sources of error. The next section describes methods for preparing this data so that an algorithm can use it for predictive maintenance.

## 4.2 Data Preparation

The aim of this sub-process is the pre-processing of the data, whereby the data sets are optimized and prepared for the algorithm used. The input format of the data is determined by the algorithm used. Besides, the sources of error presented in section IV-A must be analyzed and excluded.

- **Temporary data:** Temporary data reflects the current machine status. Each data set includes the time and a detailed description of the behaviour and the associated sensor values. This machine data is recorded periodically, but may be recorded too often or rarely. If they are recorded too often, two consecutive measured values may be identical. In this case, these measured values can be removed to reduce data complexity. In the other case, a mathematical procedure can be used to supplement missing measured values. For example, this can be done by averaging two consecutive measured values to obtain an average value.
- **Static data:** Static data can contain data records that are not relevant to the problem to be solved. These data records must be filtered out and removed. Also, data sets can be combined to form a new data set, and this is called feature development. Furthermore, the data sets can be

**Table 1** Labeled data sets for supervised learning.

Time stamp	Feature 1	Feature 2	Lifetime	Lifetime <1 day
01.01.2018	0,1432	8,532	4	false
02.01.2018	0,1858	7,942	3	false
03.01.2018	0,2145	6,606	2	false
04.01.2018	0,3243	4,348	1	true

categorized, for example, according to repair measures. Thus, repair actions can be assigned to records that trigger them.

### 4.3 Algorithms

In the previous section, the sub-processes data collection and data preprocessing, which form the basis for the algorithm, were examined in more detail. This section shows possible algorithms that can use the collected and preprocessed data to predict the ideal time for maintenance. Which type of algorithm is best suited for a particular application case is explained later in the context of the procedure model in section V.

Algorithms from the field of machine learning are used for maintenance prediction. The algorithms learn from example data and recognize patterns in them. These patterns are stored as mathematical functions in a model. A model always refers to a specific system and can be used on any number of such systems. In the case of similar systems, an existing model can be adapted so that it can also be used there. This means that no new pattern recognition is necessary, and time can be saved since the previous process steps no longer need to be run through, and the learning process does not have to be carried out. Furthermore, the algorithms of machine learning can be divided into two categories, supervised and unsupervised learning. Ultimately, the goal of the two categories is identical, namely to learn a mathematical function  $f: X \rightarrow Y$ . However, the sets X (input data) and Y (output data) differ in their nature. In the following, supervised and unsupervised learning, as well as sample algorithms for these, will be examined in more detail [13].

1) *Supervised Learning*: In supervised learning, an algorithm has a sufficiently large amount of input and output data available. The data records are already selected; that is, an input X is assigned an output Y. The task of the algorithm is to learn a mathematical function with which the output Y can be generated based on the input X [14]. The marking of the data is done by a domain expert. For this purpose, the domain experts must have sufficient knowledge about the system behaviour. A significant advantage of supervised learning is that the data sets can be divided into training and test data sets. The training records are then used for supervised learning. The resulting model can then be evaluated with the test data records. With regard to predicted maintenance, data sources such as historical maintenance logs and error logs are particularly important for this. The result of the monitored learning is either classification or regression of the set X.

- **Classification**: In classification, quantity Y is discrete. This means that the elements of this set are clearly isolated from each other and can be clearly separated. Thus, an element

of set X can be assigned by the classification to an element of set Y in the form of a type or category [14]. Table I shows labeled records of a specific system component. For each record, there is a timestamp and values for certain characteristics. The last two columns represent values to be predicted by supervised learning. The value for the "life days" describes how many days this component will be without maintenance. Whereas the value for "lifetime <one day" describes whether the component will fail in the next 24 hours. Both values can be predicted by classifying the data sets. These two values are not only an example of the classification but also represent the added value in the area of predicted maintenance. By stating these values, a system component can be maintained at the right time, and system failures can be avoided [15]. Decision trees are suitable for the predicted maintenance. Fig. 5 shows the decision tree, which can be obtained from the records of table I. Furthermore, Long Short-Term Memory (LSTM) algorithms, which belong to the neural networks, can be used [10]. LSTM algorithms are often used to predict rare events in time series data. Since machine errors are rare, LSTM algorithms are suitable for predicting maintenance [15].

- **Regression**: In regression, however, the quantity Y is continuous. This means that an element of the set X can be assigned a numerical value to an element of the set Y through regression [14]. Example algorithms for this are linear regression and regression trees[10].

2) *Unsupervised Learning*: While in supervised learning the algorithm has sufficiently large amounts of marked data sets available, in unsupervised learning only unmarked input data is available. Unsupervised learning is intended to find hidden structures in the input data in order to assign them to different groups. This process is also called clustering. For example, a group can represent the normal behaviour of a component of a machine. If anomalies occur, an error and necessary maintenance can be inferred. Unsupervised learning is particularly suitable for predictive maintenance of machines that do not have historical maintenance and error logs. A significant advantage of unsupervised learning is that no domain experts are required for data tagging. Thus, the process of unsupervised learning is more automated and requires less manual preparation of data sets than with supervised learning. The significant differences between classification and clustering are, on the one hand, the data and on the other hand, that in classification, the target categories are already given by the data tagging. In clustering, on the other hand, the target categories are created by the algorithm. This has the disadvantage that the learned model is difficult to evaluate. For unsupervised learning, LSTM algorithms can be used. These can be used for supervised and

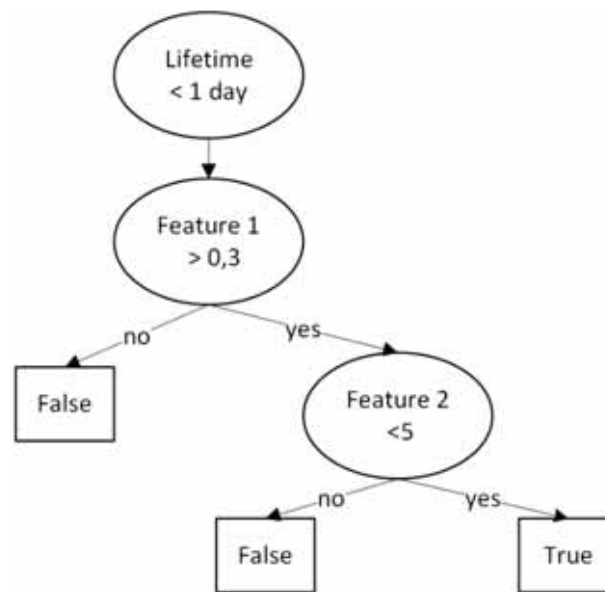


Figure 5 Classification with decision tree.

unsupervised learning, depending on the implementation of the loss function [15]. Furthermore, clustering algorithms such as hierarchical clustering, K-Means, and C-Means can be used to detect anomalies in the data sets. They enable early prediction of errors and thus timely maintenance of the component [16].

In general, supervised learning provides a more accurate prediction because the model can be more easily evaluated by the test data. If the data of a machine component meets the requirements of supervised learning, an algorithm from this area is preferred.

#### 4.4 Evaluation

The evaluation represents the last sub-process of the process chain and deals with the reaction to the result of the underlying sub-processes. By continuously monitoring the components of a machine, continuous condition monitoring can be carried out. With the help of an application directly on the machine or centrally on a server, current statuses can be displayed. The trend of Industry 4.0 and the related networking of machines offer numerous possibilities for central monitoring and predicted maintenance. If an algorithm detects an anomaly, various other processes can be started. On the one hand, automatic resource planning for spare parts can be carried out. If a component threatens to fail, it must be checked whether this component is still available as a spare part in the warehouse. If this is not the case, it must be ordered from the vendor. On the other hand, automatic resource planning can be created for the inspector so that he can maintain the machine in time and replace the affected components. Furthermore, in the event of an impending machine failure that can no longer be prevented, the employees can be scheduled in good time for other work.

### 5. PROCEDURE MODEL

In order to create a concept for predictive maintenance, a concrete use case is important. A use case defines the

general requirements of the predicted maintenance. The general requirements include the type, the possible components and the environment of a machine. On the other hand, the data supplied by the machine and which maintenance tasks are to be forecast. The requirements for the data described earlier must be observed. Based on the general requirements, an algorithm can be selected to obtain the desired result. The conclusion to be drawn from this is that the design of the process chain presented in section IV depends on the respective application of a machine. For example, if the use case changes, other data must be collected, and the selected algorithm must be adapted. In the following, a procedure for the implementation of Predictive Maintenance is presented. This process model is shown in Fig. 6 and is intended to show the process of how a forecast for a use case can be implemented with as little effort as possible.

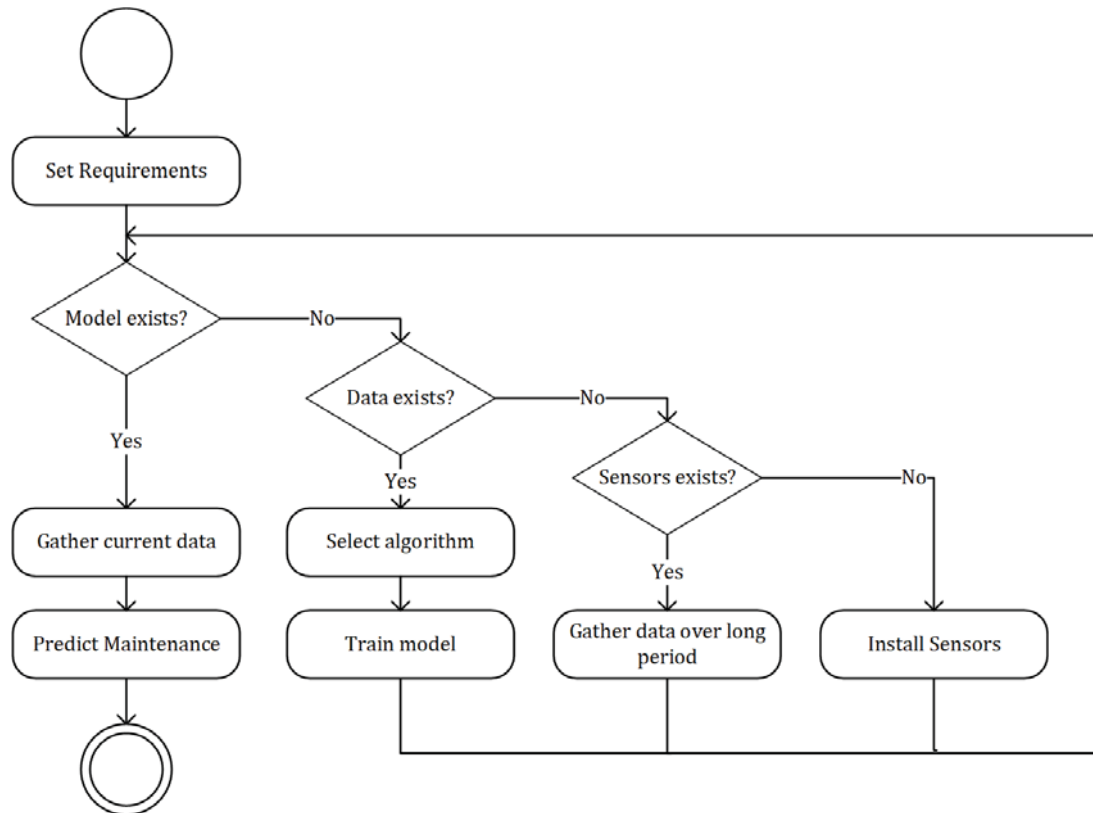
The procedure is cyclical since individual steps have to be optimized depending on the desired result. In the beginning, the use case must be defined with exact requirements and measurable goals. Measurable objectives are essential for the evaluation so that a correct evaluation of the results is possible.

In the first iteration, existing models of supervised learning are analyzed. A model might already exist, which is used in the same or similar context. If such models exist, they can be tested and compared with the requirements and objectives. If the objectives are not met, or no models exist, the second iteration is performed.

The second iteration considers the available data that are necessary for model creation. Based on this data, an algorithm from the area of supervised or unsupervised learning can be selected. If the objectives are not achieved with the selected algorithm, other models or algorithms can be tested. Mainly, the steps described in Section IV can be followed to create a model. If the goals are still not achieved because the data quality and quantity are too low, the third iteration is performed.

In the third iteration, the existing infrastructure is adapted to collect the required data. The data can be collected using built-in hardware such as temperature sensors. However, sometimes this data is not enough, and additional hardware has to be mounted





**Figure 6** Procedure model to set up predictive maintenance.

to the machine. To determine whether additional hardware is required, first, only the built-in hardware is used, and a machine learning model is implemented. If the evaluation of the model shows that predictive maintenance is inaccurate, additional hardware is required. Additional hardware are mostly sensors, which can sense how the machine behave, typical examples are temperature or sound sensors. In order to use the gathered data from the sensors efficiently, the data is stored in a central database. If the machine is already connected to a database, it is recommended to integrate the sensor data into its data structure. Otherwise, the sensors can be connected to a central IoT gateway to store and analyze the data and to add additional business logic there. Since in this iteration, no historical maintenance and error logs are available, a model for unsupervised learning is suitable. The data to train the model can be accessed from the database.

## 6. CONCLUSION

The paper explores the many factors that influence the predictive maintenance process. First of all, it is important to know the different maintenance levels. It turns out that higher maintenance levels can save costs and reduce machine downtime. However, this requires data that can be collected with the appropriate infrastructure. For the adapted infrastructure, the investment costs increase due to the additional hardware. Depending on the application, the possible effects of a machine failure must be compared to the investment costs. Predictive maintenance leads to a reduction in the failure rate and costs.

Predictive maintenance can be divided into four process steps: data collection, data pre-processing, algorithms, and

evaluation. There are temporary and statistical data sources for data collection. Temporary data sources provide real-time data. Static data sources provide data over some time, such as maintenance and error logs. Essential requirements for the data are relevance, quality, and quantity. The data must be transformed into a format that the algorithm can use. Furthermore, incorrect and redundant data must be filtered out. Algorithms from the area of supervised and unsupervised learning can be used for prediction. Supervised learning is suitable for predicting the maintenance of machines with sufficient static data. It provides high accuracy because the model can be evaluated using the data markers. Unsupervised learning is suitable for machines with insufficient static data. The algorithm learns from the temporary data and performs clustering. If anomalies occur, an error can be inferred. Based on the forecast, further processes can be started, such as optimizing resource planning for spare parts and personnel. The collected findings show that the individual sub-processes of the predictive maintenance depend on the use case. The process model presented is based on these assumptions and is intended to simplify the forecasting process for machines. The approach here leads from an existing solution to in-house development. Ideally, existing solutions can be used and optimized.

As discussed in the paper, currently, hardly any forecasts are used to determine the lifetime of maintenance. The biggest challenges for companies are data collection or lack of resources to develop their systems for predictive maintenance. One reason for this is that older machines are often not yet networked and therefore no historical data is available. Cloud vendors such as Amazon, Microsoft and IBM offer solutions to easily network machines and collect their data centrally.



They also offer existing models, or new models for predictive maintenance can be created. The presented process model can be implemented with these cloud providers and also enables smaller companies to implement predictive maintenance for their machines.

## REFERENCES

1. SAP, "Creating a dynamic maintenance management system powered", 2015. [Online]. Available: <https://www.sap.com/assetdetail/2015/12/b6caea0d-507c-0010-82c7-eda71af511fa.html> (visited on 11/10/2018).
2. Airlines, "Wartung und Uberholung von flugzeugen folgen strikten regeln", [Online]. Available: <http://www.airliners.de/wartung-ueberholung-flugzeugen-regeln/35275> (visited on 10/20/2017).
3. McKinsey, "Smartening up with artificial intelligence (ai) - what's in it for germany and its industrial sector", 2017. [Online]. Available: <https://www.mckinsey.com/media/McKinsey/Industries/Semiconductors/Our%20Insights/Smartening%20up%20with%20artificial%20intelligence/Smartening-up-with-artificial-intelligence.ashx> (visited on 11/10/2018).
4. PWC, "Predictive maintenance 4.0 - predict the unpredictable", 2017. [Online]. Available: <https://www.pwc.nl/nl/assets/documents/pwc-predictive-maintenance-4-0.pdf> (visited on 11/10/2018).
5. BHGE, "The impact of digital on unplanned downtime", 2016. [Online]. Available: <https://www.bhge.com/sites/default/files/2017-12/impact-of-digital-on-unplanned-downtime-study.pdf> (visited on 11/10/2018).
6. embitel, "Does the future of industrial asset management belong to predictive maintenance?", 2018. [Online]. Available: <https://www.embitel.com/blog/embedded-blog/does-the-future-of-industrial-asset-management-belong-to-predictive-maintenance> (visited on 12/12/2018).
7. W. Yu, T. S. Dillon, F. Mostafa, W. Rahayu, and Y. Liu, "A global manufacturing big data ecosystem for fault detection in predictive maintenance", IEEE Transactions on Industrial Informatics, 2019.
8. M. Azure, "Predictive maintenance", 2018. [Online]. Available: <https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/cortana-analytics-playbook-predictive-maintenance> (visited on 12/14/2018).
9. P. Peeling, "Big data and machine learning for predictive maintenance", 2017. [Online]. Available: <https://www.matlabexpo.com/content/dam/mathworks/mathworks-dot-com/images/events/matlabexpo/uk/2017/big-data-machine-learning-for-predictive-maintenance.pdf> (visited on 12/08/2018).
10. Fraunhofer-Gesellschaft, "Maschinelles lernen. eine analyse zu kompetenzen, forschung und anwendung", 2018. [Online]. Available: <https://www.fraunhofer.de/de/presse/presseinformationen/2018/juni/maschinelles-lernen-kompetenzen-forschung-anwendung.html> (visited on 11/18/2018).
11. Context Awareness in Predictive Maintenance. Berlin: Springer International Publishing, 2016, pp. 197–211.
12. J. M. Hellerstein, "Quantitative data cleaning for large databases", 2018. [Online]. Available: <http://db.cs.berkeley.edu/jmh/papers/cleaning-unece.pdf> (visited on 12/14/2018).
13. Wikipedia. Maschinelles lernen, [Online]. Available: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen) (visited on 01/08/2019).
14. Maschinelles Lernen: Grundlagen und Algorithmen in Python. Berlin: Springer International Publishing, 2018, pp. 197–211, ISBN: 9783446457058.
15. I. Downard, "Conquering the challenges of data preparation for predictive maintenance", 2019. [Online]. Available: <https://www.infoq.com/articles/challenges-of-data-preparation-for-predictive-maintenance> (visited on 12/27/2018).
16. N. Amruthnath and T. Gupta, "A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance", 2018. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8387124&isnumber=8387054> (visited on 12/27/2018).