Reliability Engineering and System Safety 215 (2021) 107864

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry

Andreas Theissler [a], Judith Pérez-Velázquez [b,c,*], Marcel Kettelgerdes [b], Gordon Elger [b,d]

[a] Aalen University of Applied Sciences, 73430 Aalen, Germany
[b] Institute of Innovative Mobility (IIMo), University of Applied Sciences Ingolstadt, 85049 Ingolstadt, Germany
[c] Mathematics Department, Technical University of Munich, 85748 Garching, Germany
[d] Fraunhofer Institute for Transportation and Infrastructure Systems IVI, Fraunhofer-Application Center Networked Mobility and Infrastructure VMI, 85051 Ingolstadt, Germany

ARTICLE INFO

Keywords:
Predictive maintenance
Artificial intelligence
Machine learning
Deep learning
Vehicle
Automotive
Reliability
Lifetime prediction
Condition monitoring

ABSTRACT

Recent developments in maintenance modelling fuelled by data-based approaches such as machine learning (ML), have enabled a broad range of applications. In the automotive industry, ensuring the functional safety over the product life cycle while limiting maintenance costs has become a major challenge. One crucial approach to achieve this, is predictive maintenance (PdM). Since modern vehicles come with an enormous amount of operating data, ML is an ideal candidate for PdM. While PdM and ML for automotive systems have both been covered in numerous review papers, there is no current survey on ML-based PdM for automotive systems. The number of publications in this field is increasing — underlining the need for such a survey. Consequently, we survey and categorize papers and analyse them from an application and ML perspective. Following that, we identify open challenges and discuss possible research directions. We conclude that (a) publicly available data would lead to a boost in research activities, (b) the majority of papers rely on supervised methods requiring labelled data, (c) combining multiple data sources can improve accuracies, (d) the use of deep learning methods will further increase but requires efficient and interpretable methods and the availability of large amounts of (labelled) data.

1. Introduction

The use of data-driven methods like machine learning (ML) is increasingly becoming a norm in manufacturing and mobility solutions — from predictive maintenance (PdM) to predictive quality, including safety analytics, warranty analytics, and plant facilities monitoring [1, 2]. A number of terms such as E-maintenance, Prognostics and Health Management (PHM), Maintenance 4.0 or Smart Maintenance are used to refer to the development of approaches ensuring the integrity of components, products and systems by analysing, prognosticating or predicting problems caused by performance deficiencies which may cause adverse effects on safety [3–6]. The influx of data and the emergence of the industrial internet of things have led to ML-based approaches playing a major role in this context, taking traditional maintenance modelling methods to unprecedented levels.

A prime example of how machine learning (ML) has revolutionized an industrial sector is the automotive industry, fuelled by the transformation of the vehicle into an increasingly complex system [7]. Especially, with regard to current developments towards automated driving and the transformation of the drive-train, there is a strongly increasing demand for cost-efficient technical solutions to ensure the vehicles' functional safety and reliability over lifetime [8–11]. In order to exploit the vehicles' data-richness while handling the high system complexity, ML-based PdM of safety- and cost-relevant components constitute a major solution approach with increasing attention in research. This is backed by the observed increase of publications in the field, which will be shown in Section 5.5.

While the topic of predictive maintenance (PdM) itself as well as machine learning (ML) for automotive systems have both been covered in various, separate review papers, there is a research gap: there is no current survey on the growing field of ML-based PdM for automotive systems. However, the increasing number of publications in this field emphasizes the need for such a survey. Our main contribution is twofold: First, we survey and categorize papers on ML-based PdM for automotive systems and in addition analyse them from a use case- and machine learning-perspective. Second, we identify open challenges and discuss possible research directions aiming to contribute to the development of the field and to inspire research questions. In that context,

* Corresponding author at: Mathematics Department, Technical University of Munich, 85748 Garching, Germany.
E-mail address: Judith.Cerit@thi.de (J. Pérez-Velázquez).

https://doi.org/10.1016/j.ress.2021.107864
Received 20 November 2020; Received in revised form 30 March 2021; Accepted 12 June 2021
Available online 24 June 2021
0951-8320/

we focus on maintenance related to the vehicle in use, not during manufacturing, i.e. we focus on predictive maintenance of automotive systems.

The aim of this paper is to give a broad range of readers an overview on ML-enabled PdM in automotive applications. Besides scholars and students, these are (1) maintenance specialists which have so far used classical approaches and are interested in the added value of data-driven methods, (2) ML professionals interested in key use cases with such potential as maintenance represents to the automotive industry, and (3) automotive engineers interested in machine learning for the enhancement of safety and reliability for automotive systems.

The main contributions of this paper are:

1. We introduce the machine learning subfields most relevant for predictive maintenance. This aims to make the research field of ML-based PdM accessible for experts with a background in either maintenance or machine learning, aiming to initiate fruitful collaborations.
2. We systematically survey and categorize papers in the field of ML-based PdM for automotive systems from a use case - and a machine learning-perspective.
3. From the surveyed papers we identify the most frequent use cases, frequently used ML methods and most active authors.
4. As a major contribution, we identify open challenges in the field and discuss possible research directions. This may serve readers to identify open research questions.

The paper is structured as follows: In the related work section (Section 2) we discuss literature reviews on all three topics of this survey's concern, namely maintenance, ML and automotive applications. Based on that we identify the research gap, motivating our work: in contrast to the related work, our work combines the named three topics into ML-based PdM for automotive systems. In Section 3 we give a general introduction on maintenance modelling and its subfields including PdM. Following that, in Section 4, we introduce the broad field of ML focusing on the tasks most relevant for PdM. The main contributions are to be found in Section 5, where we systematically survey and categorize ML-based PdM for automotive systems and in Section 6, where we identify challenges and discuss future research directions which may be used by readers to identify open research questions. Finally, Sections 7 and 8 contain the discussion and conclusion, respectively. Commonly used abbreviations are listed in the appendix in Table A.7.

2. Related work and research gap

Three factors can be mentioned as the drivers for the astounding development of machine learning (ML): data availability, breakthroughs in algorithm development and the advancements in computational power. The type of ML-methods used in maintenance modelling is dictated by the application and ultimately, the data available.

The pace at which this development takes place is fast and has resulted in a growing number of publications — as will be shown in Section 5.5. There is a number of reviews discussing the findings presented in research papers in the main three topics which concern our survey: PdM models, ML, and automotive applications. They comprise reviews which surveyed AI or ML methods for maintenance but also reviews on maintenance in general in which ML is treated as one possible approach. Also, reviews can be found which survey the use of ML in the automotive industry. Lastly, also reviews considering maintenance, data-driven approaches and automotive applications can be found, but cover only a very narrow field of automotive use cases. Table 1 gives an overview of related work and in what follows we elaborate further.

With regards to related work, our starting point were reviews on maintenance modelling approaches. Wu [17] and Werbińska-Wojciechowska [16] both presented a review on preventive maintenance models, the latter focusing on technical systems. Ran et al. [13] considered ML but as one of many approaches for PdM. Sutharssan et al. [14] also reviewed publications in prognostics and health monitoring (PHM) and categorized them into model-driven, data-driven and hybrid methods, subdividing the data-driven methods into statistical and ML approaches. In a similar manner, Peng et al. [15] categorized condition-based maintenance approaches, but added knowledge-based methodologies, including fuzzy logic and expert systems, as a further category.

More specifically on data-driven modelling approaches, Tsui et al. [26] made a review of the use of data-driven methods in PHM and illustrated it with practical examples. Schwabacher and Goebel [25] looked specifically at AI-methods used in prognostics and structured the revised models by the following types of methods: Physics based, classical AI, numerical methods, and ML. Wu et al. [27], on the other hand, focused on data-driven models for PHM. Carvalho et al. [21] surveyed ML for predictive maintenance in general and assessed the performance of state-of-the-art methods.

Ali [18] focuses on recent research and developments in the field of acoustic emission signal analysis through AI in machine condition monitoring and fault diagnosis. Deep Learning (DL) approaches for machine health monitoring and for system health management were reviewed by Zhao et al. [29] and Khan and Yairi [23], respectively. Bhargava [20] presented a collection of papers from various authors with a broad range of reliability prediction for electronic components comprising a wide range of AI methods.

Li et al. [33] surveyed AI applications in vehicles, tending to focus on applications in autonomous driving. In the work of Nowakowski et al. [12], the authors give a brief survey of maintenance of technical systems. Although one of their case studies is from the automotive domain, the survey itself is not automotive-specific. Falcini et al. [34] discuss deep learning (DL) for automotive systems, however not with a focus on predictive maintenance. Also, in the context of automotive applications, Singh and Arat [35] give an overview of the advances and challenges in DL.

Although it does not focus on automotive applications, the classification of recent literature from Wu et al. [28] can be seen as an important premise to our contribution. In a similar fashion to our approach, Fink et al. [22] and Lei et al. [24] evaluate current developments and discuss potential research trends but they do it in the fields of deep learning applied to PHM and machine fault diagnosis, respectively.

Finally, the closest work to our survey are two reviews combining PdM, data-driven approaches and automotive applications: Ahsan et al. [30] and Sankavaram et al. [32]. However, our survey is significantly different to these in following aspects: (1) both are brief reviews with a limited selection of papers, (2) Ahsan et al. [30] have a different focus, reviewing the subfield of automotive electronics in the context of data-driven methods — ML methods covered as one of subfield that, (3) Sankavaram et al. [32] reviewed work up to the year of 2009.

A point worth mentioning is, that earlier reviews already pointed towards the importance of data-driven approaches in PHM, see Mesgarpour et al. [31].

In contrast to the named reviews, we (a) focus specifically on ML-based PdM for automotive systems, (b) survey a variety of applications and categorize them into PdM sub-fields, and (c) as a key contribution, identify open challenges and research directions in the field. Our paper closes the research gap of a case study survey of ML-based PdM for automotive systems with a focus on vehicles in operation. To the best of our knowledge, there is no current survey of that kind, combining ML, PdM and automotive applications.

Table 1
Overview of related work.

Reference	Topic
Maintenance Modelling:	
Nowakowski et al. [12]	Brief survey of predictive maintenance
Ran et al. [13]	Predictive Maintenance
Sutharssan et al. [14]	PHM methods
Peng et al. [15]	Prognostics in condition-based maintenance
Werbińska-Wojciechowska [16]	Preventive Maintenance for Technical Systems
Wu [17]	Preventive Maintenance Models
PdM & ML	
Ali [18]	AI-based signal analysis in CM and fault diagnosis
Alsina et al. [19]	ML for reliability prediction of components
Bhargava [20]	AI for reliability prediction of electronic components
Carvalho et al. [21]	ML for predictive maintenance
Fink et al. [22]	DL in PHM
Khan and Yairi [23]	DL in system health management
Lei et al. [24]	ML in machine fault diagnosis
Schwabacher and Goebel [25]	AI in prognostics
Tsui et al. [26]	Data-driven approaches in PHM
Wu et al. [27]	PHM and AI
Wu et al. [28]	ML in reliability and maintenance
Zhao et al. [29]	DL for machine health monitoring
Data-based models & PdM & Automotive	
Ahsan et al. [30]	Prognostics of Automotive Electronics
Mesgarpour et al. [31]	Telematics PHM for Commercial Vehicles
Sankavaram et al. [32]	Prognosis of Automotive and Electronic Systems
Automotive & ML:	
Li et al. [33]	AI for vehicles
Falcini et al. [34]	DL in automotive software
Singh and Arat [35]	DL in the automotive industry

3. Maintenance modelling — terminology and taxonomy

In this section, the terminology and categories of maintenance modelling as used in this survey are introduced. Maintenance comprises measures to maintain a system in its specified operation mode either by repairing failures or by taking actions in order to avoid them. In the European standard PN-EN 13306 [36], maintenance is defined as "a combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it, or restore it to a state, in which it can perform the required function".

Maintenance strategies can be subdivided in various ways, a commonly used categorization is [12,37,38]:

1. corrective maintenance: Corrective maintenance, also called reactive maintenance, fix-upon-failure or run-to-failure, aims to repair a system or its components *after a failure occurred*.
2. preventive maintenance: Preventive maintenance (see e.g. [16]) bases on pre-scheduled maintenance intervals mainly using fixed time intervals, sometimes in addition utilizing a system's usage (e.g. the mileage of vehicles). Preventive maintenance aims to repair a system *before a failure occurs*, not taking into account the system's actual health status.
3. predictive maintenance (PdM): PdM aims to predict the optimal time point for maintenance actions, taking into account information about the system's health state and/or historical maintenance data. It tries to *avoid the premature and costly repair of a system*, while at the same aiming to *ensure a timely repair prior to a failure*. Advanced methods aim to predict the expected time of a failure, thereby estimating the remaining useful life (RUL), see e.g. [39,40].

While condition-based maintenance (CBM) is often used as a synonym for predictive maintenance, in [41] CBM is viewed as a subcategory of PdM, subdividing PdM into:

1. statistical PdM: based on data not directly connected with the state of an individual vehicle, e.g. historical maintenance data or data from a vehicle fleet or entire population
2. condition-based PdM: monitoring the system health using real-time data to determine maintenance decisions

In this survey, we adopt the categorization of [41], i.e. statistical PdM and condition-based PdM are viewed as the two subcategories of PdM.

An alternative categorization of maintenance is the use of the terminology from "Industry 4.0". Doing so, maintenance can be categorized according to the level of maturity from maintenance 1.0 (corrective maintenance) to maintenance 4.0 comprising advanced data-driven methods, and methods estimating the probability of failures and their effects (reliability centred maintenance) [12].

4. Machine learning for predictive maintenance

Since this paper discusses machine learning (ML) for predictive maintenance, in this section, the ML fundamentals relevant for PdM are reviewed and ML is related to PdM.

ML is a subfield of artificial intelligence (AI) that focuses on teaching computers how to learn without the need to be programmed for specific tasks. ML approaches can be subdivided into unsupervised, semi-supervised, supervised and reinforcement learning (see Fig. 1) — see e.g. [42]. In unsupervised learning, the data is not labelled. The ML model aims to discover unknown patterns in the data, e.g. by means of similarities between the data points. Algorithms are therefore formulated such that they can find patterns and structures in the data on their own. In semi-supervised learning the input data is a mixture of labelled and unlabelled data points. In supervised learning, the ML model uses labelled training data. That is, it is given labels with the correct output and aims to learn a mapping of inputs to outputs, often adjusting the model in an iterative way. This process is repeated until the model achieves a desired level of accuracy on the training data and can correctly predict the outputs for new instances.

Finally, reinforcement learning uses trial and error in an exploration vs. exploitation manner to discover the actions that yield the greatest rewards. Regarding applications to the automotive industry, reinforcement learning has been pivotal to enable autonomous driving. RL has

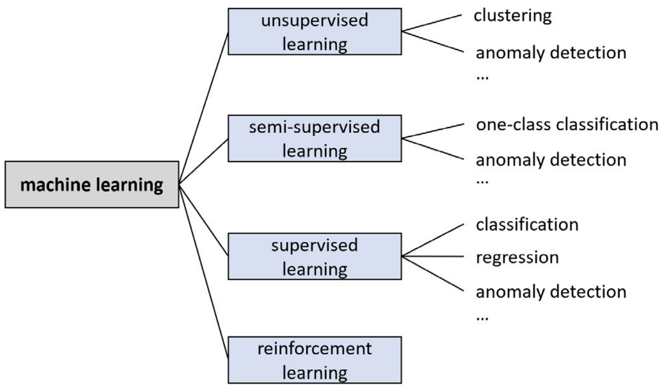

Fig. 1. Common categorization of machine learning with the ML tasks relevant for this work. For reinforcement learning no further details are shown, since none of the surveyed papers used reinforcement learning.

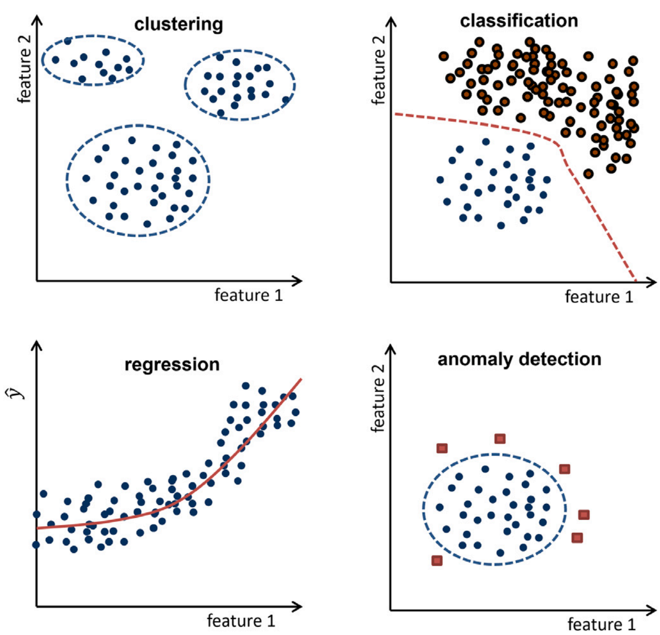

Fig. 2. Machine learning tasks most relevant for PdM.

been applied to PdM, see for example [43] in the context of power grids and for a general approach for maintenance planning in [44]. Also, in one of the reviews mentioned in Section 2, Ran et al. [13] discussed several works on PdM and deep reinforcement learning, a combination of reinforcement learning (RL) with deep learning. However, based on our search criteria for ML-based predictive maintenance for automotive systems (see Table 2 in Section 5.1), none of the surveyed papers used reinforcement learning. Hence, it is not further detailed here. The interested reader is referred to [45,46] for surveys of the field and e.g. to [47,48] for (non-automotive-specific) application.

In the following, the ML tasks from Fig. 1, clustering, classification, regression and anomaly detection, are briefly introduced. Fig. 2 shows the four tasks for contrived two-dimensional data sets. A survey of ML-based PdM use cases is then given in Section 5.

4.1. Clustering

The most common unsupervised learning method is cluster analysis (clustering) which is used for exploratory data analysis to find hidden patterns or groupings in data (see top left in Fig. 2). Examples of clustering methods are k-means, fuzzy c-means [49], hierarchical cluster analysis, DBSCAN [50], and HDBSCAN [51].

Definition 1 (*Clustering*). *Clustering* is the partitioning of data points x_i of a data set X with feature space F into k groups $C_j = \{c_1, \dots, c_k\}$ by a model M_{clust}.

4.2. Classification

Supervised learning can be further subdivided into classification and regression. Classification uses labelled data to learn a mapping from inputs to class labels aiming to learn some decision function (see top right in Fig. 2). Some common classifiers are k-nearest neighbours (k-NN), naïve Bayes classifiers, support vector machines (SVMs) [52], (deep) artificial neural networks [53], decision trees, random forests [54] and XGBoost [55].

Definition 2 (*Classification*). *Classification* is the assignment of data points x_i to any or multiple of k pre-defined classes $Y = \{y_1, \dots, y_k\}$ by a model M_{class}, where M_{class} was trained on a train set X_{tr} with labels Y_{tr} and feature space F.

4.3. Regression

Regression comprises supervised methods that use pairs of inputs and outputs to learn to predict continuous outputs for new inputs (see bottom left in Fig. 2). Some common ML methods for regression are support vector regression [52], (deep) artificial neural networks [53], random forests [54] and XGBoost [55].

Definition 3 (*Regression*). *Regression* is the prediction of (typically continuous) outputs $\hat{Y}$ for input data X by a model M_{reg}, where M_{reg} was trained on a train set X_{tr} with outputs Y_{tr} and feature space F.

4.4. Anomaly detection

The task of predictive maintenance (PdM) is closely related to modelling a system's normal behaviour and detect deviations, so-called anomalies, which may point to present or evolving failures (see bottom right in Fig. 2). This is known as anomaly detection and has been used in a variety of domains [56–61], not limited to the automotive industry. The importance of anomaly detection is due to the fact that anomalies translate to significant information about a system's health status. Due to its high relevance for PdM, we review anomaly detection in more detail.

While anomaly detection can be achieved by classification (one-class [62,63], two-class or multi-class classification) or clustering (e.g. outlier detection with DBSCAN [50] or HDBSCAN [51]), we decided to have anomaly detection as an own subcategory spanning methods from all three levels of supervision in accordance with [64] (see Fig. 1).

A general definition of anomaly detection is given in Definition 4, where x_i may refer to some subset of the data e.g. an individual data point, a group of data points, a subsequence of a time series, or a region of an image:

Definition 4 (*Anomaly Detection*). Let M_{ad} be some function or model that, applied on a subset x_i of given input data X utilizing some feature space F, returns 0 if x_i is normal and 1 if x_i is an anomaly.

In the automotive safety-related standard ISO-26262 [65], an anomaly is defined as a "condition that deviates from expectations, based, for example, on requirements, specifications, design documents, user documents, standards, or on experience". In a more general context, in [64] an anomaly is defined as a deviation from expected behaviour.

Table 2
Criteria which were combined into one search query on Scopus. The rows are combined with AND operators.

Criterion	Search terms
Automotive	"automotive" OR "vehicle" OR "car" OR "truck"
PdM	"predictive maintenance" OR (detect* W/2 fault) OR "condition-based maintenance" OR "condition monitoring" OR "prognostic and health management" OR "PHM"
ML	"machine learning" OR "deep learning" OR "artificial intelligence" OR "classification" OR "clustering" OR "regression" OR "anomaly detection" OR "reinforcement learning"
Exclude	NOT ("rail" OR "railway" OR "aerial" OR "UAV" OR "underwater" OR "aircraft" OR "road surface" OR "road condition" OR "ship" OR "trains" OR "EOL" OR "end-of-line" OR "review" OR "survey")
Date	≥2010
Type	journal article OR conference paper
Language	English

Anomaly detection is a common approach for fault detection. Within the taxonomy of error, fault, and failure, an anomaly can be considered as a potential error, where an error is caused by a fault and may in turn cause a failure. Hence, an anomaly detection may point to a fault [7] and can therefore be used for *condition-based PdM*.

In *statistical PdM*, anomaly detection can be applied on data from vehicle fleets or historical maintenance data. Detected anomalies may point to issues for an entire vehicle model or to problems of an individual vehicle. Possible data sources can be diagnostic trouble codes, e.g. from connected vehicles or from repair shop visits (see e.g. [66]) or the repair history. In *condition-based PdM*, anomalies can be detected by physical modelling of the normal behaviour or data-driven ML approaches, where data from vehicles or from simulations can be used as the training data set. For the estimation of the *remaining useful life* (RUL), anomalies may be an indication for degradation. The following anomaly detection approaches can be distinguished:

1. **Physical models:** The physical modelling of the normal behaviour based on underlying specifications allows to detect deviations (residuals) as anomalies. In addition or alternatively, implausible situations and known faults can be modelled. Pure physical models are not contained in this paper but there are hybrid models combining physical models with ML, as will be discussed in point 5 (see e.g. [67] for a Deep Learning-based method or [68] for a review).
2. **Unsupervised anomaly detection:** In an unsupervised setting, anomalies can be detected by determining whether a subset x_i is anomalous w.r.t. to the remaining data X [64]. No class labels are used, i.e. X contains no reference data labelled as normal or anomaly. Examples of unsupervised methods use distance or density measures [69], clustering [50,51,70], tree-based methods [71,72], or neural network-based autoencoders [73,74]. Those approaches can, for example, be used to model the normal behaviour and for new data to test how well the models describe the data.
3. **Semi-supervised anomaly detection:** While unsupervised methods can be applied in the absence of labels – which are often very costly to obtain – this lack of information is compensated by the models' assumptions about anomalies. For cases where it is possible to obtain normal data, e.g. using simulations or real systems that are not likely to show abnormal behaviour during the time of data acquisition, semi-supervised anomaly detection [64] can be used. ML models can be trained on normal data to model the system's normal behaviour. During operation, deviations are reported as anomalies. Common methods are one-class classifiers, e.g. support vector machines [62,63] or one-class deep learning methods [75,76].
4. **Supervised anomaly detection:** If both, normal data and representative data with anomalies, are available, the setting can be viewed as a two-class or multi-class classification problem. The presence of a representative set of anomalies is, however, a strong assumption. In such a setting, supervised anomaly detection [64] can be utilized. Common ML models can be trained on both, normal data and anomalies, classifying new instances as either normal or anomaly. For this task, standard classifiers can be used, however, either the data or the classifiers need to be adapted to the high class imbalance [77] — anomalies are usually rare. As an alternative to classification, regression methods can be used. Common ML methods that can be used are support vector machines [52], random forests [54], XGBoost [55], or deep learning models [53].
5. **Hybrid approaches:** Different types of hybrid approaches address the drawbacks of the named individual methods by combining data-driven and physical models.

4.5. Further methods

Further unsupervised methods are dimensionality reduction/ projection methods like principal component analysis (PCA) and t-SNE [78]. These can be used in a pre-processing step to reduce the dimensionality of the data, but PCA can also be used to model the normal behaviour.

A promising approach in the absence of labels are generative models, which offer a way of learning data distributions using unsupervised learning. The aim of learning the true data distribution of the training set is to generate new data points with some variations. Two of the most commonly used approaches are variational autoencoders and generative adversarial networks (GAN), see [79,80] for variants thereof for anomaly detection.

A further promising field is transfer learning which aims to transfer knowledge learned in one setting (the source domain) to another setting (the target domain) in order to leverage available training data and apply the models to deviating scenarios with new conditions or faults. Methods for transfer learning were proposed for bearing fault diagnosis in [81], a weighted domain adaptation network was proposed in [82] and demonstrated for gear box data, and a RUL use case was presented in [83].

5. ML-based PdM for automotive systems: a survey

In this section we survey and structure papers on machine learning-based predictive maintenance for automotive systems. In total we surveyed 62 papers. The selection of the papers is described in Section 5.1. To structure the field, we categorized PdM along the dimensions of "maintenance benefit" and "complexity" into the three subfields *statistical PdM*, *condition-based PdM* and *remaining useful life* (see Fig. 3). The surveyed papers are categorized into these subfields based on their main use case. An overview of the papers is given in Table 4. For readers without ML background, the predominantly used ML methods are briefly introduced in the appendix in Table A.8.

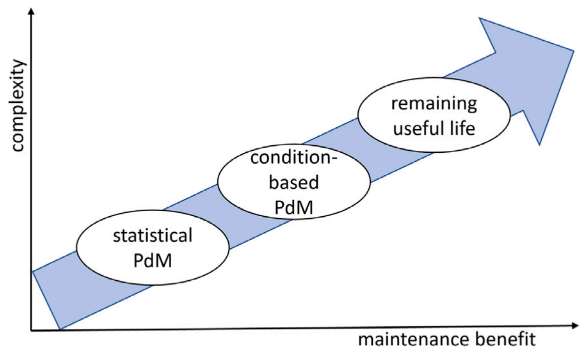

Fig. 3. PdM subfields of selected papers (own categorization).

Table 3
Exclusion criteria to refine the initial result set of the search query.

Exclusion criteria:
pure physical models or no ML included
road, traffic or infrastructure monitoring
manufacturing, logistics, cyber security
driver observation
trains, ships, agricultural vehicles
aircraft, space vehicles
literature reviews, surveys, overview papers

5.1. Research methodology

To allow the reader to understand our selection of papers, our research methodology is described in the following (see Fig. 4 for an overview):

1. We defined search criteria for papers that cover predictive maintenance for automotive systems using machine learning (see Table 2) and searched for Scopus-indexed papers.[1] This resulted in 395 papers.
2. We manually filtered the list of papers in a reproducible way with the exclusion criteria in Table 3. As an additional criterion, we only selected papers with > 5 citations. This resulted in 39 papers, we refer to that list as A.
3. While the number of citations is an acknowledged criterion to reduce the number of papers [143], it introduces a bias towards older papers: Recently published papers are likely to have less citations. Therefore, we additionally selected 5 papers with publication date ≥ 2020, ignoring the number of citations. Instead of that, the inclusion criterion was agreement among this paper's authors that the selected paper is likely to be influential. We denote these papers as B and marked them with[B] in Table 4.
4. We analysed the result set A and grouped them by their main use case. In order to further illuminate these use cases, we conducted an individual targeted search. From this search we manually selected 18 papers, which are marked with[C] in Table 4.

5.2. Statistical predictive maintenance

Statistical PdM, as [41] termed it, uses data not directly connected with the state of an individual vehicle, but rather data from a number of vehicles in some common backend. In some papers this is also referred to as Big Data approaches. Examples are historical maintenance data, vehicle properties like age, mileage and model, and feedback data from vehicle fleets.

Phillips et al. [84] conducted a real-world case study to classify oil samples from diesel engines. The data was obtained from a fleet of 60 off-highway trucks in the mining industry and classified into four classes — from healthy to clear contamination. The authors selected features relying on expert knowledge from domain specialists. Their main goal was an interpretable model, hence they used logistic regression and compared it to an SVM and an ANN.

Zhao et al. [11] applied unsupervised ML methods on data from a vehicle fleet to detect the location of faulty cells within electric vehicle batteries. They focus on potential design problems not addressing random errors that could occur during production or in the field. Their underlying assumption is, that the location of faulty cells within the battery pack is constant in the case of design flaws and follows a Gaussian distribution. The terminal voltages of the cells in the batteries of the vehicle fleet are used. In a first step, the data distribution is approximated with an ANN. Following that, outliers are detected in a statistical way, applying a stepwise 3-σ threshold. In the absence of labels, they compared their approach with the unsupervised local outlier factor (LOF) algorithm [144] and a clustering-based outlier detector. While LOF showed better performance in the case of low fault frequencies, their proposed approach was most robust w.r.t. differing frequencies of faults.

Sankavaram et al. [85] address the classification of known and unknown fault types using incremental learning. They used an ensemble that allows to adapt to new fault types by not re-training the entire model on the previous fault classes. The ensemble's base classifiers were methods like k-NN, SVM and ANN. In a real-world case study, they evaluated their approach on the electronic throttle control. Data from a wide range of vehicles of different ages were used and successful adaption to new fault types was shown.

Nowakowski et al. [12] investigated the benefit of statistical PdM using the example of a vehicle fleet of a transportation company. In a case study they show that simple statistical PdM is indeed superior to the currently used preventive maintenance strategy. They showed that the brand and age of the vehicles can be predictive factors for selected failures.

In [86], Byttner et al. introduced COSMO (Consensus self-organized models), an approach that aims to build up knowledge over time by an explorative search of internal local signals and comparing them with equivalent signals from a group of vehicles that perform similar tasks. COSMO was used for predictive maintenance of vehicle fleets, where the majority of the vehicles is assumed to be healthy and deviations from the majority can be considered as potentially faulty. Fan et al. [87] used COSMO to detect compressor failures in a fleet of city buses. For this purpose, available sensor signals were tracked and compared to the rest of the fleet to detect deviations.

Killeen et al. [88] proposed an IoT approach based on COSMO that detects faulty buses deviating from the rest of the fleet. It differs from the original COSMO approach by proposing a semi-supervised approach for improved sensor feature selection. The IoT infrastructure contains a vehicle node and a gateway which performs sensor data acquisition, aggregation, and lightweight data analytics. A root node is responsible for managing the entire fleet system.

Gardner et al. [89] presented a novel algorithm, PRISM, for automating multivariate sequential data analyses using tensor decomposition. Much of the reason their work is significant comes from the extensive data source: a municipal vehicle fleet of 2500 vehicles operated by the city of Detroit with e.g. police cars and ambulances. The aim was three-fold: to discover sequential patterns in the multivariate maintenance data (using the proposed unsupervised PRISM approach), to perform predictive maintenance (using an LSTM), and to predict vehicle- and fleet-level costs (using an ARIMA time series model). To this end, historical maintenance and purchase data were utilized. They are able to accurately predict both future maintenance jobs and the average future expenses. They state that one advantage of their method is that it is composed of interpretable predictive models, providing

[1] Search query executed on 18th Feb 2021 at Scopus.

Table 4
Overview of surveyed papers, with main use case, ML task (clust = clustering, class = classification, reg = regression, AD = anomaly detection, fcast = forecasting), and ML method (the predominantly used methods are briefly explained in the appendix Table A.8).

Use case	ML task	ML method	Reference
Statistical PdM:			
SoH engines	class	log. regression, SVM, ANN	Phillips et al. [84]
Faults in EV batteries	AD	fuzzy c-means clustering, FDA, PCA	Zhao et al. [11]
Full vehicle fault detection	class	ensemble (incremental learning)	Sankavaram et al. [85]
Maintenance in vehicle fleets or populations	n/a	Statistics	Nowakowski et al. [12][C]
	AD	COSMO	Byttner et al. [86][C]
	AD	COSMO	Fan et al. [87]
	AD	IoT & COSMO	Killeen et al. [88][C]
	seq. pattern mining, fcast	PRISM, LSTM, ARIMA	Gardner et al. [89][C]
	reg	gcForest	Chen et al. [41][C]
	reg, class, fcast	linear regression, gradient boosting	Khoshkangini et al. [90][B]
Condition-based PdM:			
Faults in engines	class	ensemble of ELMs	Wong et al. [91]
	class	ensemble of ELMs	Zhong et al. [92]
	clust, class	ANN variant	Wang et al. [93]
	class	ANN	Zabihihesari et al. [94]
	AD, class	residual selection, logistic regression	Jung and Sundström [95]
	AD, class	LSTM and CNN, log. regression, SVM, random forest	Wolf et al. [96]
	AD, class	one class-SVM	Jung [97][C]
	class	ANN	Wang et al. [98][B]
SoH EV batteries	reg	ELM vs. ANN	Pan et al. [99]
	fcast	LSTM	You et al. [100][C]
Faults in EV batteries	class	random forest	Yang et al. [101]
	reg	polyn. regression, SVR, ANN	Quintián et al. [102]
Faults in EV powertrains	AD, class	SVM, k-NN, ANN variant	Sankavaram et al. [103]
Full vehicle fault detection	AD	ensemble of one- and two-class classifiers	Theissler [104]
	class	dec. trees, SVM, k-NN, random forest	Shafi et al. [105]
	class	autoencoder variant	Tagawa et al. [106]
	clust, AD, class	PCA, ICA, own clustering method	Routray et al. [107]
Faults in air pressure system	class	ANN, LSTM, CNN	Rengasamy et al. [108]
	class	boosted decision trees	Cerqueira et al. [109]
	fcast	relaxed prediction horizon algorithm	Nowaczyk et al. [110]
Faults in gearboxes	class	ANN	Heidari Bafroui and Ohadi [111]
	class	k-NN, Gaussian mixture models	Gharavian et al. [112]
	class	hybrid deep belief networks	Zhang et al. [113]
Faults in suspension systems	AD, clust, class	fuzzy c-means clustering variant, PCA, FDA	Yin and Huang [114]
	AD, clust, class	fuzzy c-means clustering variant, PCA, FDA	Wang and Yin [115]
	class	CNN, ANN	Zehelein et al. [116][B]
	reg	NARX ANN	Capriglione et al. [117][C]
	reg	NARX ANN	Capriglione et al. [118][C]
	AD	SVM	Jeong et al. [119][C]
Faults in brake systems	class	clonal selection classification algorithm	Jegadeeshwaran and Sugumaran [120]
	class	ANN, SVM, best first trees, Hoeffding trees	Alamelu Manghai and Jegadeeshwaran [121]
Faults in steering systems	reg	SVR	Ghimire et al. [122]
	AD, class	rough set theory, dec. trees, SVM, k-NN, ANN variant	Ghimire et al. [123]
Sensor fault detection	AD	ELM-based autoencoders and ANN	Fang et al. [124][C]
Tyre monitoring	class	dec. trees, PCA	Siegel et al. [125]
	class	CNN	Siegel et al. [126][C]
Fuel cell vehicles	class	ANN	Mohammadi et al. [127]
	reg	attention-based LSTM and gated recurrent unit (GRU)	Zuo et al. [128][B]
Faults in generators	class	neuro-fuzzy inference system	Wu and Kuo [129]
Engine starter system	class	ensemble of multinomial regression models	Peters et al. [130][C]

(*continued on next page*)

actionable insights. Their work also highlighted the need to improve the accuracy and granularity of existing data and collecting additional data, including vehicle drivers, time, location, and the total time a vehicle was in use.

Table 4 (*continued*).

Use case	ML task	ML method	Reference
Faults in electric motors	reg, class class	ANN ensemble of random forest and ANN variant	Şimşir et al. [131] Seera et al. [132][C]
SoH autonomous vehicles	class	IoT & MLP	Jeong et al. [133][C]
SoH automated vehicles	AD	CNN	Van Wyk et al. [134]
Remaining useful life (RUL):			
EV batteries	reg class, fcast fcast fcast	ANN, L-PEM multi-target probability estimation LSTM CNN, MLP	Rezvani et al. [135][C] Last et al. [136] Wu et al. [137][C] Wang et al. [138][C]
Air pressure systems	reg, fcast	random forest	Prytz et al. [139]
Gearboxes	class	least squares-SVM, k-NN	Taie et al. [140]
Electric motor bearings	reg	SVR, least squares	Lee et al. [141]
Electric vehicle capacitors	clust, reg	fuzzy c-means clustering and Markov model	Al-Dahidi et al. [142][C]

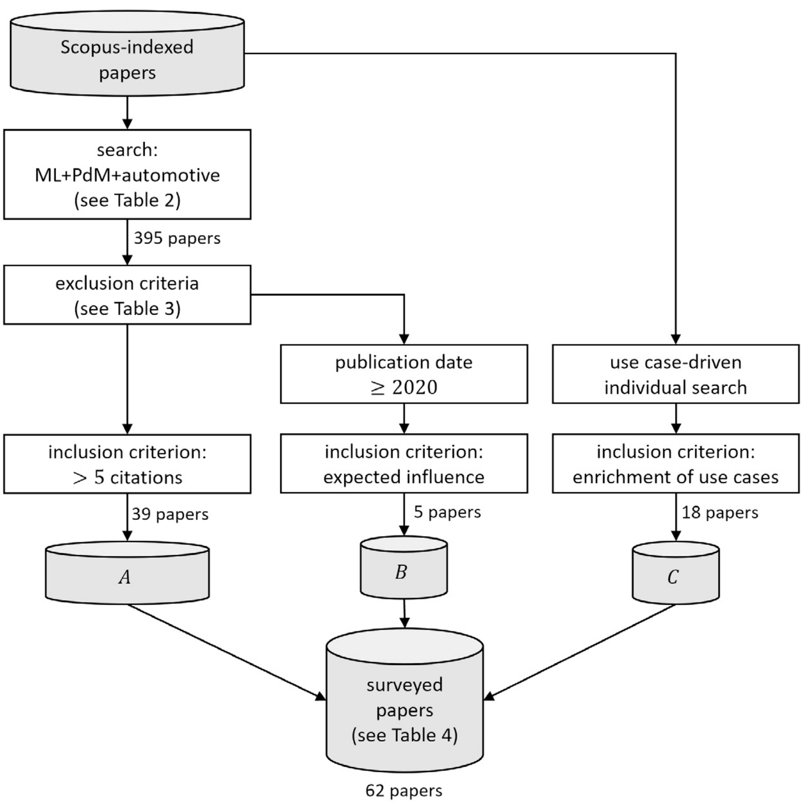

Fig. 4. Overview of research methodology for the selection of papers for the survey.

In [41], Chen et al. proposed an approach to predict the time between failures of a vehicle, with the aim to optimize the maintenance strategy of a fleet management company with repair shops across the UK. The approach combines the repair shops' historical maintenance data with geographical information around the repair shops (weather, terrain and traffic). In a supervised regression setting, an ANN, SVR, random forest and the deep tree-based model gcForest [145] were used, with gcForest reported to yield the best results.

Khoshkangini et al. [90] used a massive data set from a manufacturer's entire vehicle population with the aim to forecast the ratio of failures per month over the population. Two general approaches were compared: The first one uses claim data about parts and components that were repaired or changed in workshops. With linear regression, a forecast of the failure ratio is conducted for the entire population. The second approach combines the claim data with logged vehicle data (configuration and usage of vehicle) into a sequence per vehicle. Then gradient boosting is used to classify whether the specific vehicle will have a failure in a given month. The classifications are then combined to obtain the forecast for the entire population. From their experiments they concluded, that initially the claim-data based regression approach performs better up to a point when the vehicles have been in service for a longer time. Then the combined approach with gradient boosting outperforms the first one.

5.3. Condition-based predictive maintenance

As opposed to statistical PdM, condition-based PdM uses operating data from individual vehicles in order to derive either the state of the overall system, or of one or more components. Based on this, a component-related maintenance decision can be made.

A fundamental approach to predict failures is the detection of faults. Early detection of a fault can prevent its propagation, hence actions can be taken prior to breakdowns. Thereby, condition-based PdM is often achieved by means of anomaly detection or classification. In general, unsupervised, semi-supervised and supervised learning can be utilized, depending on the availability of data and labels. Due to the high number of papers in this subfield, the papers are grouped into subsections relating to different use cases, or more specifically vehicle key components on which condition-based PdM was applied.

5.3.1. Faults in engines

Wong et al. [91] propose a method for the detection of faults in engines focusing on simultaneous faults, i.e. multiple single faults occurring concurrently. They used an ensemble of Bayesian extreme learning machines (ELM) in a supervised classification setup. While the base classifiers were solely trained on different single faults, their experiments on data from a vehicle show that the ensemble is capable of detecting single as well as simultaneous faults. One novelty is, that they set the output of a base classifier to zero if it was not trained for a specific fault type. Zhong et al. [92] addressed a similar problem as Wong et al. [91]. While both research groups used similar classifiers – ensembles of Bayesian ELMs – they differ in the way the signals are decomposed and features are extracted. In addition, Zhong et al. [92] used weighting of the base classifiers w.r.t. their performance. The approach is evaluated with real data from a vehicle utilizing the signals of the engine sound, air-ratio, and ignition.

Wang et al. [93] researched the detection of engine faults from vibration signals. They propose a novel variant of an ANN for that purpose. They refer to their method as extension neural network type-1, and state that it is computationally more efficient than a standard ANN. Their method measures the distances between new data points and cluster centres of fault types. The approach is compared with a standard ANN and evaluated with faults in the spark plug and fuel injection of one vehicle.

Zabihihesari et al. [94] presented a hybrid approach for combustion fault detection of a 12-cylinder 588 kW diesel engine. Their combined method is based on vibration signature analysis using fast Fourier transform, discrete wavelet transform, and ANNs. Their results reliably distinguish between normal and faulty conditions.

Jung and Sundström [95] proposed a residual selection algorithm to monitor the air path through an internal combustion engine. In that context, a residual generator $r(z)$ of a system is a function of a sensor or other actuator data z where a fault-free system implies that the residual output $r(z) = 0$. Jung and Sundström [95] used a hybrid approach to detect faults and isolate, i.e. classify, them. Using a mathematical model, they generated the residuals. From these residuals a subset is selected in a feature selection step. Following that, a logistic regression classifier is trained on the selected residuals. In the fault isolation step they used t-SNE to project the data to a two-dimensional space. The underlying data are sensor signals such as pressure before throttle, pressure in intake manifold, temperature before throttle, engine speed and throttle position. In essence, their proposed approach finds residual generator candidates with specific fault properties and thereby allows to establish a direct link between faults and residuals.

Wolf et al. [96] on the other hand, focused solely on pre-ignition faults in turbocharged petrol engines. They used data extracted from electric control units (ECU) of a vehicle fleet. The authors introduced a deep learning architecture with four CNN layers, two LSTM layers and one softmax layer. In addition, different subsets of features were selected. The proposed approach identified faults with an F1-score of 0.9 and was superior to stand-alone CNNs and LSTMs as well as to linear SVMs, logistic regression and a random forests.

Jung [97] used one-class SVMs enhanced by Weibull calibration to model fault classes individually. The goal was to detect both, known and unknown fault classes — the latter being faults not present in the training data. The approach was evaluated with data from a combustion engine and shown to detect seven known fault types as well as previously unseen fault types.

Wang et al. [98] classified faults in engines based on sound measurements. They used wavelets to extract discriminative features which they used to train an ANN. With their approach, they could successfully classify data from an engine on a test bed into fault-free and any of 8 faults.

5.3.2. Faults and state-of-health for EV batteries

Pan et al. [99] addressed state-of-health (SoH) estimation, i.e. capacity degradation, for batteries of electric vehicles. The crucial step is the identification – in ML wording this would be feature extraction and selection – of two health indicators from the battery's internal resistance using correlation analysis. In a second step, they trained an extreme learning machine (ELM) on the identified features. To train and test the approach, they generated a data set with batteries with different levels of capacity loss. They compared the ELM with a standard ANN and concluded that for their application, ELM is easier to use, faster to train and yields better accuracy.

You et al. [100] offered a solution to the problem of diagnosing battery states in real-world driving patterns. Their data-driven approach uses measurable data from electric vehicles, such as current (I) and voltage (V), and can thereby monitor the SoH. They simulated the operation in vehicles for more than 70 battery cells: The data was obtained with a battery cycler, which is used to charge and discharge batteries by imposing a varying current. While other data-driven approaches divide the entire available region of I/V into multiple subregions and judge the ageing based on the varying density distributions, they analyse I/V instantiation patterns over short periods. They used a LSTM as the estimation model, which can handle time-variant data such as the charge sequence and also memorize long-term information.

Yang et al. [101] proposed an approach to detect external short circuits in batteries, which is a safety-relevant fault in electric vehicles. In a first step they compared two physical models, estimating their parameters with a genetic algorithm. Following that, they used their domain knowledge to identify features that allow to identify the fault. As the final step, they trained a random forest classifier in a supervised manner. The approach was tested with real batteries in a lab environment.

Quintián et al. [102] presented a hybrid model for fault detection of a LFP (Lithium Iron Phosphate — LiFePO4) power cell type, used in electric vehicles. A k-means clustering algorithm was used to identify groups of data with the same behaviour. Then, three different regression techniques were tested for each group: polynomial regression, ANN and SVR. Their approach was able to classify all the fault situations as a battery fault; however, they report that their model cannot distinguish between erroneous measurements.

5.3.3. Faults in electric vehicle powertrain

Sankavaram et al. [103] investigated data-driven fault detection and diagnosis for regenerative braking systems of hybrid electric vehicles. Therefore, The authors modelled the overall powertrain of a two motor series–parallel hybrid vehicle mathematically and conducted fault injection experiments in order to simulate the most common system faults. During the simulations 25 system state variables, including for example wheel speed, engine torque and battery state of charge, were monitored. To minimize the computational costs, the state space was reduced using multi-way principal component analysis and multi-way partial least squares. Based on the reduced data, the authors implemented algorithms including SVMs, a probabilistic ANN, partial least squares and k-NN to classify the faults into 12 classes, e.g. "Battery SOC Fault", "Wheel Inertia Fault" or "Motor1 Current Sensor Fault" reaching an accuracy of up to 100% with SVM and k-NN.

5.3.4. Full vehicle fault detection

Using data from the in-vehicle network, Theissler [104] proposed an ensemble of one-class and two-class classifiers for fault detection in recordings from road trials. The one-class classifiers are trained on recordings from normal operation mode, the two-class classifiers are trained on normal data and faults. This semi-supervised setting allows to detect known and unknown fault types, while a standard classification approach would be limited to only detect known fault types. As one crucial base classifier, support vector data description (SVDD) [62] is utilized and enhanced by a heuristic to tune the hyperparameters solely from normal data (see also [7,60]). The approach is evaluated on data from multiple vehicles.

Shafi et al. [105] propose an architecture and a procedure for fault detection in a vehicle fleet. They acquired signals from the in-vehicle network via the OBD interface, i.e. without additional sensors. This data is transferred to a common backend, where a knowledge base of the entire fleet is built. Fault detection for different vehicle subsystems is achieved with decision trees, SVM, k-NN, and random forests. They state that faults identified in one vehicle will also be used to warn drivers of vehicles with similar conditions. The approach was evaluated with data from 70 vehicles.

Tagawa et al. [106] proposed a method they refer to as structured denoising autoencoders, which is a variant of denoising autoencoders. As a main contribution, their method allows to incorporate partial knowledge about relations between variables and about faults. In a driving scenario, their approach outperformed methods such as one class-SVM, vanilla denoising autoencoders, and LOF. As a limitation, they did not use real faults: they recorded different driving scenarios assuming one of them to represent the normal condition and scenarios like going down a slope to simulate a fault.

Routray et al. [107] introduced a data-driven fault detection framework utilizing unsupervised independent component analysis (ICA) and principal component analysis (PCA) for data reduction and clustering based on feature distance and mutual information. While their approach does not focus on a specific component, they demonstrated it for anomaly detection in automotive MAF sensors.

5.3.5. Faults in air pressure systems

Rengasamy et al. [108] demonstrated the impact of weighted loss functions used in neural network architectures to increase the fault detection accuracy in air pressure systems of heavy trucks. They worked on a truck manufacturer's (Scania) data set and evaluated ANNs, CNNs, LSTMs and gated recurrent units (GRU). With that said, the authors were able to reach a prediction accuracy of 98.8% with the evaluated CNN and an adapted loss function. Also working on a Scania truck air pressure system data set, Cerqueira et al. [109] implemented boosted trees to successfully classify air pressure faults. Both authors did not directly refer to the source of the data set, but most probably used the open-source data set published by Lindgren and Biteus [146].

In another work, Nowaczyk et al. [110] introduced a fuzzy rule-based algorithm with relaxed prediction horizon to classify air pressure failures. Unlike the previous works, the authors worked on a manufacturer's internal data set (Volvo). They compared their approach to classical approaches like k-NN, decision trees and random forests. The authors found that the fuzzy rule-based algorithm performed particularly well on unbalanced data, meaning that the amount of available training data is not equally distributed over the classification labels.

5.3.6. Faults in the gearbox

Heidari Bafroui and Ohadi [111] used a supervised ML setting to classify data of gearboxes into healthy, and three types of faults: chip, worn 10%, worn 5%. A continuous wavelet transform is applied on the vibration signals and statistical features are extracted. Following that, the energy and Shannon entropy is used for feature reduction and the resulting features are classified with an ANN. In their experiments, their proposed feature reduction was superior in accuracy and training time, compared to classifying the non-reduced feature space. For a similar problem setting, Gharavian et al. [112] classify the vibration data of a gearbox. Their focus is on the comparison of feature extraction methods. A continuous wavelet transform is applied on the signals, followed by feature extraction with PCA and Fisher discriminant analysis (FDA). They classified the resulting feature vectors with k-nearest neighbours (k-NN) and Gaussian mixture models. In their experiments with a gearbox operating at constant speed, feature extraction with FDA was superior to PCA.

Also focusing on automotive gearboxes, Zhang et al. [113] introduced a hybrid deep belief network (DBN) architecture in order to classify common faults in planetary gearboxes, like (partially) broken teeth. Therefore, an experimental setup consisting of a healthy and several fault injected gearboxes, a BLDC motor, braker and several sensors was built up. The tracked sensor data included, among others, motor current and voltage, torque, vibration and rotational speed. After preprocessing and segmentation of the measurements, the hybrid DBN was trained and compared to methods like a classic deep belief network (DBN), CNN, SVM, an autoencoder and a LSTM. In their results, the authors showed that their hybrid DBN performed superior compared to the other classification strategies.

5.3.7. Faults in suspension systems

In Wang and Yin [115] and Yin and Huang [114] a research group investigated fault detection in vehicle suspension systems. More precisely, faults in springs are detected using acceleration sensors at the four corners of the vehicle body. In Wang and Yin [115], they proposed a semi-supervised approach, starting with an initial cluster of normal data. Newly occurring faults are then detected in an online-manner using possibilistic c-means clustering (a modification of fuzzy c-means clustering). In Yin and Huang [114] they enhanced their work and proposed an unsupervised approach. From a data set with a majority of normal data and potentially some faults, the number of potential clusters is manually identified using a PCA transformation. With a variant of fuzzy c-means, the data is then subdivided into clusters. Both papers used Fisher discriminant analysis to isolate the faults in a final step. Their methods avoid the classification of underlying faults with different intensities into different fault types, rather classifying them into the same fault type using so-called fault lines. In both papers the approach was evaluated with simulation data from a full car suspension model.

Capriglione et al. [117] approximated the stroke sensor of the rear suspension in motorcycles by a soft sensor. In a supervised setup, the true sensor values were used as the target variables and an autoregressive model was trained on the time series data. They used a combination of "Nonlinear Auto-Regressive with exogenous inputs" (NARX) and an artificial neural network (ANN). While their main focus was to approximate the true sensor, they propose to use the soft sensor to detect faults in the actual sensor. This work was later enhanced in Capriglione et al. [118] focusing on fault detection.

In [119], Jeong et al. used a combination of a physical model and an ML model to detect faults in suspension systems. They first extract residuals based on the physical model and then apply a SVM to detect faults.

Zehelein et al. [116] injected faults into a vehicle's damping system by manipulating the respective damper currents and therefore reducing damping force. Based on that, the authors collected wheel speed, yaw rate as well as longitudinal and lateral acceleration data from the vehicle's electronic stability control during typical driving scenarios with varying trunk weight and tyre type (winter and summer). After pre-processing the data with selected methods, like fast Fourier transformation and short time Fourier transformation, they trained a CNN in order to classify the suspension system into one intact and three fault classes, each relating to the position of the defect dampers. The results were compared to the classification results of a classical ANN, yielding a classification accuracy of up to 92.22% for the CNN, compared 87.27% for the ANN.

5.3.8. Faults in the brake system

Jegadeeshwaran and Sugumaran [120] propose an approach for the detection of faults in the brake system. Vibration signals are acquired from a brake test rig with an additional accelerometer. From these signals, statistical features are extracted and a subset is algorithmically selected. Classification is then conducted with a clonal selection classification algorithm.

In [121], the research group around Jegadeeshwaran preprocessed the vibration data from the brake test rig via wavelet analysis. Based on that they evaluated a set of classification approaches, namely best first trees, Hoeffding trees, SVMs and an ANN, to detect and classify ten different hydraulic brake states, ranging from "Good" over "Drum Brake Pad Wear", "Air in Brake Fluid", up to failure states like "Brake Oil Spill" and "Reservoir Leak". The best classification results were achieved with the Hoeffding trees.

5.3.9. Faults in the electric power steering

Ghimire et al. [122] developed a physics-based model of an electric power steering system. Through a series of fault injection experiments they were able to derive fault-sensor measurement dependencies to isolate the faults. They used support vector regression (SVR) to estimate the severity of faults. In the later work Ghimire et al. [123], the same main author enhanced the previous work. In a first step, a physical model was built and simulations were conducted under different fault conditions. In addition to physical models, the authors evaluated data-driven approaches, namely k-NN, a probabilistic ANN, SVM, decision trees as well as rough set-theory. According to the authors, the rough-set theory method performed superior on sparse data compared to the other methods.

5.3.10. Sensor fault detection

Although this application mainly concerns autonomous vehicles, we categorize it here since it focuses on sensors. The methods of Fang et al. [124], based on sensor monitor cluster, allows not only sensor fault detection but also fault location. They used extreme learning machine based autoencoder applied to anomaly detection. Their framework has three parts which allow sensor monitor cluster, anomaly detection and fault location, respectively. The monitoring part is basically dedicated to determine normality of the data, this is carried out through discrete wavelet transform. It allows de-noising and extracting features from the sensor data to analyse the health conditions of each sensor. They then performed anomaly detection using an autoencoder. Finally to carry out the fault location, ANNs are used for actuator fault tests. To this end, they used a fuzzy PID controller. They validate their method with both experiments and simulations.

5.3.11. Tyre monitoring

In two papers, the same main author addressed tyre monitoring. The research group took an approach that – in addition to the ML part – implemented the entire mobile communication infrastructure. Siegel et al. [125] utilized GPS and acceleration data from a mobile phone, which was mounted in a vehicle, in order to classify a 20% increase or decrease of the vehicles' tyre pressure. Therefore, the authors used PCA to reduce the dimensionality of the data and decision trees as a classifier, reaching a classification accuracy of 80%.

In the later work [126], Siegel et al. proposed an approach for visual tyre inspection. The main idea relies on pattern matching based on features present in tyres with cracked sidewalls. CNNs were chosen for being particular suitable for textures [147]. Using mobile phones, pictures of the tyres are sent to a cloud-based system, where the ML model evaluates the pictures. The ML model was trained with a variety of images of tyres both in good and bad conditions. Human assistance was used to verify the labelling for training and to help calculating the baseline accuracy. They used two CNN models: a baseline model and one constructed by densely connected blocks.

5.3.12. Fuel cell vehicles

Mohammadi et al. [127] implemented an ANN in order to conduct water management fault classification for automotive fuel cells. They trained the neural network based on a mathematical model, which was again validated by experimental measurement data. The ANN successfully classified and localized drying and flooding faults among nine modelled cells.

Zuo et al. [128] utilized LSTMs and gated recurrent units (GRUs) to predict the terminal voltage of a fuel cell depending on degradation and load current. In order to do so the authors conducted long-term dynamic load durability tests with a proton exchange membrane fuel cell (PEMFC). In the subsequential test phase, the best prediction results could be achieved by the attention-based LSTM with a coefficient of determination of up to 0.89.

5.3.13. Faults in electric motors, generators and starter

Wu and Kuo [129] utilized an adaptive neuro-fuzzy inference systems (ANFIS) – a combination of an ANN with fuzzy logic – to classify faults of automotive generators. In order to create the data set, they used an experimental setup with an engine and a generator as core components. They ingested synthetic failures in the generator and applied a discrete wavelet transformation (DWT) to its output voltage signal. Extracting the DWT coefficients as features, the ANFIS was trained to identify different fault classes under different engine speeds, resulting in an classification accuracy of 98.8%.

Şimşir et al. [131] performed real-time monitoring and diagnosis of the faults of a hub motor. They measured main system parameters and trained an ANN. Different faults were diagnosed: coil open circuit, coil short circuit, fault in hall effect sensor, short circuit between coils and damaged bearing faults. The model was embedded into an Arduino Due microcontroller card to allow mobile real-time monitoring and fault diagnosis.

Peters et al. [130] addressed the classification of multiple, interacting fault modes and the determination of the corresponding health states. Following a feature extraction step, an ensemble of multinomial regression models was trained. The approach was evaluated for the engine starter system, which comprises the battery and the starter motor. The authors state that while the methodology is applicable to common industrial systems, knowledge about the diagnosable faults and about system health indicators is necessary. The approach was evaluated with data from an engine test-rig.

In electric vehicles, induction motors are widely used. In this context, Seera et al. [132] developed a hybrid condition monitoring model that consists of an ensemble of a fuzzy min–max ANN and a random forest. They examined the efficacy of their model in monitoring multiple incipient faults from induction motors using information from only one source (i.e., stator currents) in both noise-free and noisy environments. An important contribution is, that they take interpretability of the model into account and extract a decision tree in order to explain the model predictions to domain users. An experiment was conducted, to monitor and predict three different induction motor conditions: fault-free, stator winding faults, and eccentricity problems.

5.3.14. Health state of autonomous or automated vehicles

Jeong et al. [133] presented a self-diagnosis system for autonomous vehicles utilizing an IoT infrastructure and deep learning models. In their work they propose a detailed communication network to read data from the vehicle and to transfer data to a backend in a cloud. An ANN, more precisely a multilayer perceptron (MLP), is used in a supervised learning scenario where the number of nodes and layers is dynamically adapted. The condition of vehicle components is classified as "normality", "inspection", or "danger" using this information to warn the driver in the case of a risk.

In [134], van Wyk et al. combined a CNN and well-established anomaly detection methods to detect and identify anomalous behaviour in connected and automated vehicles. They showed, using real data that a combined approach of feature extraction and classification outperforms the use of either of these two alone. The proposed system can be applied to any motor data.

5.4. Remaining useful life

The remaining useful life (RUL) is the continuous estimation of the future time span in which a component is still considered operational — in addition to the aforementioned approaches which yield a crisp decision whether a maintenance action is necessary or not. This is precious to avoid premature costly repairs, e.g. in the natural case of signs of wear, a repair is not necessarily instantaneously required.

Electric vehicle batteries have been the subject of a number of studies aiming to quantify the SoH and predict their RUL. Rezvani et al. [135] compared two methods for a publicly available data set: an adaptive version of an ANN and the linear prediction error method (L-PEM) were trained in a supervised manner on time series obtained over the battery cycles. They report that the ANN yielded a higher accuracy in capacity estimation but L-PEM showed a better performance in RUL prediction. Last et al. [136] proposed an own tree-based algorithm, namely multi-target Info-Fuzzy Network, which was utilized to identify rules in order to predict probability distributions of battery failure and remaining useful lifetime based on features like the open-circuit voltage, state of charge or battery temperature. The algorithm was tested on synthetically generated data and was able to outperform classic Weibull statistics. More recently, Wu et al. [137] used LSTMs to determine the SoH of lithium-ion batteries in electric vehicles. They state that the extraction of healthy features prior to the use of ML methods were crucial for their application.

Wang et al. [138] developed a prediction method for voltage and lifetime of lead–acid batteries in electric vehicles. They used a CNN and a standard ANN (a MLP). The data was recorded from 10 lead–acid batteries over a time span of 155 weeks. Different voltages were used as features and the CNN and MLP were compared, where the CNN achieved a higher accuracy.

Unlike the aforementioned works that addressed batteries, Prytz et al. [139] estimated the RUL of air compressor systems. In a large-scale study on real-world data from trucks and buses, they aimed to schedule repair shop visits. To address this, they estimated the RUL and compared it with the time of the next planned service visit. The key is the combination of two data sources: (a) the vehicles' usage patterns, which are read-out during repair shop visits, and (b) the repair shops' service records. They focused on four failures of the air compressor which were detected using a random forest classifier in a supervised learning setting. Prytz et al. state that the use of feature selection methods were crucial.

Taie et al. [140] proposed a general vehicle remote diagnosis platform analysing vehicle sensory data with a least squares-SVM and demonstrated it on gearbox data. Expert knowledge was utilized to label the gearbox condition based on engine speed, wheel speed, and gearbox temperature as input features. Based on that, the least squares-SVM was trained to classify a gearbox into "NOK", "10% RUL", "40% RUL" and "OK", reaching a classification accuracy of 93%, which was more accurate than k-NN reaching 82%. This use case can be considered as borderline case between condition-based PdM and RUL prediction, since the RUL prediction is solely based to two crisp classification labels.

Lee et al. [141] predicted the RUL of an electric motor's rotary bearing by training and evaluating ordinary least squares, feasible generalized least squares and support vector regression (SVR), based on experimentally generated vibration data. While the SVR was the computationally most expensive algorithm, it outperformed the other methods.

Lastly, in a promising approach, Al-Dahidi et al. [142] estimated the RUL from a "fleet of equipment", rather than solely relying on data from an individual unit. In a first step, the approach identifies the degradation levels with unsupervised ensemble clustering. These are then used as the states for a homogeneous discrete-time finite-state semi-Markov model. The approach is evaluated with two case studies, where one is from the automotive domain: In an artificial case study, the RUL of capacitors in a fleet of electric vehicles was estimated.

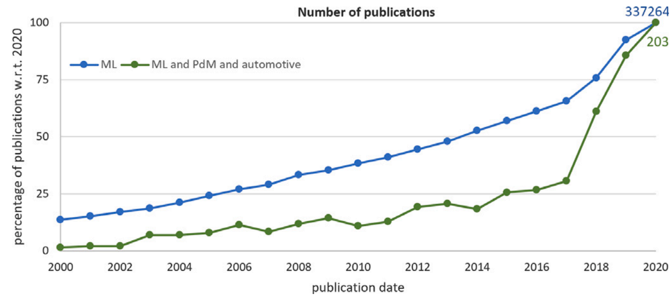

Fig. 5. Number of publications on ML in general (blue) and on ML-based predictive maintenance for automotive systems (green). To allow comparison, the publications are scaled as percentages w.r.t. 2020, the absolute numbers for 2020 are given on the top right. While the field of ML keeps increasing, ML-enabled PdM for automotive systems has experienced a boost starting in 2017.

5.5. Bibliometric analysis

A brief bibliometric analysis was conducted to deduce further insights. The number of Scopus-indexed publications was observed for general machine learning topics (search terms see Table 2, criterion "ML") and for ML-based PdM for automotive systems (search terms see Table 2, criteria "automotive" AND "PdM" AND "ML", prior to application of exclusion criteria). Interestingly, while the field of ML keeps increasing, ML-enabled PdM for automotive systems has experienced an additional boost starting in 2017 (see Fig. 5), underlining the increasing importance of the field.

In addition, from the paper list *A* (see Fig. 4), the Top-5 influential papers (measured by number of citations on Scopus), the authors with the most contributions, the most frequently addressed use cases, and the most frequently used ML methods were extracted — this analysis is shown in Table 5. From the author's contributions (second column), two research groups can be identified as most active: the group around Sankavaram and the group of Nowaczyk, Rönvaldson, Prytz, Byttner and others. The most investigated uses cases are concerned with the engine as well as electric vehicle batteries, followed by full vehicle use cases.

The – by far – most frequently used ML methods are ANNs, where we categorized standard neural networks like MLPs and their variants as ANN. In addition, an increasing number of papers uses neural network architectures like CNNs, LSTMs, ELMs and autoencoders. Following neural networks, SVMs have been widely used. This can be partially explained by the available reformulation of standard SVMs for anomaly detection, i.e. one-class SVMs. Furthermore, the wide use of ensembles should be noted.

6. Challenges and research directions

In addition to surveying a broad range of automotive use cases, this paper contributes by identifying challenges, open questions, and research gaps. The aim is to inspire researchers by identifying potential research directions. Based on the use case survey and the authors' experience the following challenges were identified:

6.1. Non-availability of public real-word data sets

As opposed to data set repositories for general ML problems (e.g. [148] for computer vision or [149,150] for time series), public real-world data sets for automotive data are very rare. One main reason is, that the data is considered as highly confidential by the automotive industry — understandably since for predictive maintenance the data contains faults or wear in their products. This has severe consequences on the research in this field, shown by following observations:

Table 5
Top-5 influential papers (measured by number of citations), Top-5 most active authors, most frequently addressed use cases, and most frequently used ML methods. Extracted from the paper list A (see Fig. 4).

	Influential papers		Authors		Use cases		ML method
1	Yin and Huang [114]	1	Pattipati, K. (4 papers)	1	Engine	1	ANN
2	Heidari Bafroui and Ohadi [111]	2	Sankavaram C. (3 papers)	2	EV battery	2	SVM
3	Zhao et al. [11]	3	Nowaczyk S. (3 papers)	3	Full vehicle	3	Ensemble
4	Pan et al. [99]	4	Rögnvaldsson, T. (3 papers)	4	Air pressure system	4	Decision tree
5	Theissler [104]	5	12 authors with 2 papers	5	Gearbox	5	ELM, fuzzy c-means, k-NN

Table 6
Categorization of surveyed approaches by PdM and machine learning subcategories. Note that some papers were assigned to multiple levels of supervision, since they evaluated different approaches.

	Unsupervised ML	Semi-supervised ML	Supervised ML
Statistical PdM	[11,12,86,87,89]	[85,88]	[41,84,89,90]
Condition-based PdM	[95,106,107]	[97,104,114,115,132,134]	[91–94,96,98–103,105,108–113,116–123,125–131,133]
RUL	[139,142]		[135–138,140,141]

- **Lack of continuous lines of research:** Due to the non-availability of public real-world data, research is typically conducted by academia in collaboration with automotive manufacturers or suppliers, or within the automotive companies themselves. When a research project is finished, the researchers – often Ph.D. students – take on different positions in academia or industry. As a consequence, access to the underlying data is no longer given, hence, that line of research is not continued. This observation is backed by researching publications by various researchers publishing frequently for 3-5 years in a certain field of the automotive domain, before moving on to publish different topics or not publishing any further.
- **Non-efficiency in follow-up research:** Follow-up research by further researchers building on data from prior research is usually not possible. The costly phase of data acquisition needs to be repeated by new researchers entering the field.
- **Lack of reproducible research:** The described non-accessibility of data prevents published research from being reproduced and possibly built on.
- **Limitation in research rigour:** Proposed approaches are usually presented with results on own data sets (e.g. [97,104,139]). Quantitatively comparing a new approach with state-of-the-art results is not possible due to the non-availability of the data sets previous approaches were tested on.
- **High cost for data acquisition:** Researchers of data-driven topics in the automotive industry typically face the problem of having to acquire data themselves or getting granted access to confidential data. While the latter is more of a political issue, acquiring data is an enormous effort. Measurements from different operation modes are required in order to have a representative data set. In order to train models, or at the minimum in order to test models, recordings of systems with different types of faults or signs of wear are required. Due to the enormous effort, some researchers need to rely on simulated data.
  One option to get access to a large amount of data is to use recordings from vehicles that are tested during pre-series or end-of-line tests. In some cases – especially for commercial vehicles – data is also recorded from customer vehicles during operation in the field.
  In the vast majority of the surveyed papers, the data had to be acquired by the researchers, e.g. in Pan et al. [99], Heidari Bafroui and Ohadi [111], Gharavian et al. [112], Wong et al. [91], Wu et al. [137], Theissler [104], You et al. [100]. Other researchers, e.g. Wang and Yin [115], Last et al. [136], Al-Dahidi et al. [142], had to rely on simulated data. Tagawa et al. [106] had to rely on assumed faults simulated by different driving scenarios, e.g. going down a slope or driving in the neutral gear. Zhong et al. [92] and Wong et al. [91] explicitly mention the high cost to obtain the data for their problem setting of simultaneous engine faults. Zhong et al. [92] had to take an enormous effort to generate data: 10 types of single faults and four types of simultaneous faults were generated in the test vehicle. Recording for each of these faults were repeated 200 times for each of two operation modes. Next to Cerqueira et al. [109] and Rengasamy et al. [108], who most probably worked on a publicly available data set of air pressure system faults [146], Rezvani et al. [135] could use a publicly available data set for their research on batteries.

Open question: Need for public real-world data sets. As a consequence, we state that there is a need for more public real-word data sets from the automotive domain. This would be a boost in the research of data-driven methods in the automotive industry and would allow for continuous lines of research exploiting previous results. In addition, it would also yield more accurate, robust and general solutions, being able to evaluate methods on a variety of different data sets.

6.2. Lack of labelled data

Working with real-world data offers the chance of being able to evaluate developed methods in practise. On the downside, real-world data is often not or only partially labelled, since annotating the large amount of data is time-consuming and requires expert knowledge. While there are unsupervised methods not requiring labels (see Table 6), semi-supervised or supervised methods usually yield more robust results. As a minimum, labelled data is required to test models, even if having trained them on unlabelled data.

The vast majority of the surveyed use cases in Section 5 relied on labelled data, e.g. Pan et al. [99], Heidari Bafroui and Ohadi [111], Capriglione et al. [118], Jeong et al. [119], Chen et al. [41]. Zhao et al. [11] proposed an unsupervised approach. In the absence of labels, they had to evaluate their approach against other unsupervised methods in contrast to an evaluation w.r.t. ground truth.

Research direction: Efficient labelling processes. While there are simple statistical methods to determine outliers in the statistical sense, only a human expert can decide whether some subset of the data is to be considered a true anomaly, i.e. a fault or a sign of wear. For that reason it appears to be natural to incorporate expert knowledge in the labelling process, with the aim to move beyond the prohibitively costly way of manually labelling the entire data set. Visual interactive labelling (VIAL) [151–154] is a promising approach, where Active Learning [155,156] is used in combination with expert knowledge to interactively select subsets of data to be labelled. For anomaly detection problems, VIAL-AD was proposed in [157], discussing a number of open research questions specific to anomaly detection.

6.3. Complexity of problem setting

Several challenges can be traced back to the complex problem setting caused by the fact, that the automotive industry supplies long-term products for a wide range of application areas.

- **High number and variety of vehicles:** Due to the enormous number and variety of vehicles (e.g. discussed in [7]), it is a challenge to acquire a representative data set for ML models to be trained on. In addition, vehicles have a long life time, e.g. a commercial vehicle manufactured today, can be assumed to still be in operation in 20–30 years, adding to the variability of data seen during operation.
  Research direction: enhancing the representativeness of the training set.
  One option to improve the representativeness of the training set is to – in addition to real-world data – use simulated data from hardware- or software-in-the-loop systems. Another option is to artificially create data that resembles the original data. This is a common approach in deep learning termed data augmentation [158] and can lead to a boost in accuracies e.g. in computer vision applications by creating images with different rotations or zoom-ins. For PdM, care must be taken to ensure the data is not falsified in such a way that incorrect classification occurs.
- **Rarity of faults and wear:** During the operation of vehicles, faults are very rare and wear occurs only after a long period of time. Hence, obtaining data from vehicles yields a data set that is highly imbalanced and is not likely to contain all faults [97,104] and signs of wear.
  Research direction: Methods for learning from imbalanced data.
  The imbalanced data requires the use of ML models that are robust to class imbalances, cost functions that do not optimize towards the majority class, and resampling methods for the data [77,159].
- **Concept drift:** In ML, there is an underlying assumption, that the data a model was trained on is representative for the data the model will see during operation. Despite enormous efforts during quality assurance, this assumption is not fulfilled in the automotive industry. Reasons are the high number and variety of vehicles, the manifold locations the vehicles are used in (e.g. different weather conditions and territories), and the different usage of vehicles. Further reasons are the long life time and the resulting wear, replacement of parts, and software updates. Hence, it is highly likely that an ML model will encounter situations during operation that it was initially not trained on. This phenomenon is termed concept drift [160,161]. There are different taxonomies in literature, one categorization divides it into virtual and real concept drift:
  - A virtual concept drift, also referred to as covariate shift, refers to a change of the data distribution, e.g. caused by the vehicle being operated in a different way, at different locations or in different applications. Formally it can be described as
    $$P(X_{op}) \neq P(X_{tr}) \tag{1}$$
    where $P(X_{tr})$ denotes the distribution of the training data and $P(X_{op})$ the distribution of the data encountered during operation.
  - With a real concept drift, situations that were initially viewed as anomaly may become normal during the life time of a vehicle, or vice versa. Formally a real concept drift is expressed as:
    $$\exists x_i \rightarrow y_i \,:\, P(Y_{op}|X_{op}) \neq P(Y_{tr}|X_{tr}) \tag{2}$$
    where x_i refers to a subset of the data, y_i to the corresponding label, Y_{tr} and Y_{op} to the labels, X_{tr} to the training set and X_{op} to the data seen during operation. In other words, the correct assignment of $x_i \rightarrow y_i$ may change w.r.t. time. This problem setting was for example addressed with domain adaptation for RUL in [83].

  Research direction: Addressing concept drift.
  There appears to be no silver bullet to cope with concept drift in automotive PdM. A virtual concept drift was implicitly addressed by [97,104], by assuming that the distribution of faults in the training set is not representative for the unseen data. For real concept drift, an obvious option is to incorporate the driver or an expert into the decision process, in case of uncertainty. This requires, however, that the model is aware of its uncertainty about a decision. For ML in general, this has been addressed e.g. in [162]. A further option is to continuously update the models using the knowledge acquired from the entire vehicle population, thereby adapting to changes that are not specific to an individual vehicle or to unforeseen situations that appear in the field. The research field of federated learning [163] is a promising direction in that context.

6.4. Acceptance of ML-based maintenance

In general, the acceptance of ML-based methods for an application area that is used to the predominant use of physical models or models defined by domain experts is strongly connected with the experts' and users' trust in these models. Trust in turn is related to the reliability and interpretability of models [164], as well as the understanding how one's personal data is used by the models:

- **Non-interpretability of complex ML models:** Manufacturers, repair shops and customers demand explanations for the replacement of parts. With interval-based maintenance the explanation is trivial, however, with the use of sophisticated ML models, e.g. deep learning models [53], explanation and interpretation of the models and their decisions become a challenge. From the surveyed papers, Phillips et al. [84] emphasized the importance of understanding the ML models in order for industry experts to trust the models. They favoured an interpretable model opting for logistic regression. In addition, only Seera et al. [132] and Gardner et al. [89] have explicitly addressed the interpretability of the models.
  Research directions: explainable AI and interpretable models.
  The field of explainable AI (XAI) has emerged in recent years [165,166], e.g. with methods to explain the inner workings of models or to explain a black box model with an interpretable model. As opposed to that, there is also some research advocating not to use black box models for "high stake decisions" [167].
- **Data privacy:** Vehicles are moving sensors, measuring data from the vehicle, the driving behaviour, the environment, and the current location. On the one hand, from a technical point of view the ideal solution would be to transfer the data of all vehicles to a common backend, in order to build, improve, and adapt ML models. Each vehicle would then benefit from the knowledge extracted from all vehicles. For example ML models could adapt to unforeseen breakdowns occurring in the field.
  On the other hand, the aforementioned approach allows to determine how, where, and when a vehicle was operated, which is not desired by many vehicle owners.[2] As opposed to a common backend, the other extreme is to have the models work locally solely considering the vehicle and the knowledge that was built in during manufacturing. However, due to considering only a single vehicle, the added value is not as high.
  Research direction: addressing data privacy.
  In general, anonymization can help to build trust, but prevents to make potentially important links between the same vehicle or vehicles within the same region. Pseudonymization is

[2] Note, that data privacy is viewed differently in different regions of the world.

a practical compromise. However, it has been shown that allegedly anonymized or pseudonymized data can in some cases be used to uncover individuals, e.g. using location information (see e.g. [168]), unique usage patterns or data fusion. A promising research direction is federated learning, where not the data is transferred in a distributed environment, but rather the model parameters. In [163] a secure federated learning framework is proposed, with the aim to preserve data privacy.

6.5. Increasing complexity due to transformations in the automotive industry

- **Maintenance of ML-based automotive systems:** While traditionally, mechanical and electrical parts were the target of predictive maintenance (e.g. vee-belt or suspension system), for ML-based advanced driver assistance systems and autonomous driving systems, the ML models might become additional items for predictive maintenance — although not likely on the vehicle level, rather on the level of all vehicles with an ML model of a certain version. Reasons for ML models to be updated or replaced are improved accuracies of new model generations, higher efficiency, lower latency, or improved robustness against adversarial attacks [169–171]. In the use case survey, Jeong et al. [133] addressed PdM for autonomous vehicles. In addition, in [172] an ANN is used to detect data injection attacks on the cooperative adaptive cruise control of connected vehicles.
- **Transformation of the drive train:** The advent of alternative drive trains brings along new vehicle components requiring maintenance. Specifically for full electric vehicles, hybrid vehicles as well as fuel cell vehicles, the monitoring of batteries is crucial. A lot of research has recently been conducted in this field, as shown by the surveyed use cases You et al. [100], Wu et al. [137], Wang et al. [138], Rezvani et al. [135] and others. Due to the ongoing transformation of the drive train, it is highly likely that this research direction will become increasingly important.

7. Discussion

In this section, we draw overriding conclusions from the findings of the literature survey and the identified open challenges:

ML has become a pivotal approach for the field of maintenance modelling. ML has earned its place among the most used and useful methods currently employed within maintenance modelling. The availability of data as well as developments in computing power have the potential to keep boosting the field in years to come. Addressing open questions and challenges will enrich this research area.

ML-enabled PdM accompanies the transformation of the drive train. From the surveyed uses cases the use of PdM for components of the drive train stands out (see Tables 4 and 5). It is no surprise that the engine as one of the vehicle's core components in terms of cost and complexity has been researched in many papers. We find it noteworthy that the transformation of the drive train brings along new challenges visible by the number of publications in that field. A high number of papers address batteries of electric vehicles due to the fact that the battery is safety-relevant and an enormous cost driver for electric vehicles. In addition, the electric motor and fuel cell vehicles were investigated in some papers.

The combination of condition-based pdm and the available big data in statistical PdM solutions offers enormous potential. Another finding is, that statistical PdM appears to be the key approach when using data from vehicle fleets or entire populations. We believe that data fusion from different sources can lead to a boost in the accuracy and pave the way for new applications. Examples could be the combination of statistical PdM using data like historical maintenance data and trends in a vehicle population in connection with condition-based PdM using individual vehicle related data, e.g. its current state and usage patterns. In Section 5 some promising approaches using data fusion were discussed: Prytz et al. [139] and Khoshkangini et al. [90] combined usage patterns with service records and Chen et al. [41] combined historical maintenance data with geographical information like weather, terrain and traffic. In addition, Shafi et al. [105] used data from within single vehicles and also from the entire fleet.

Artificial neural networks dominate the field, however, currently not with massively deep architectures. Analysing the use of ML methods, neural networks were by far used most frequently. In Table 4 we categorized standard neural networks like MLPs and their variants as "ANN". In addition to these, an increasing number of papers use neural network architectures like CNNs, LSTMs, ELMs and autoencoders. This is consistent to the increased use of neural networks in general machine learning, i.e. in deep learning. Since in many of the papers the mentioned neural network architectures were used with a small number of layers – i.e. one hidden layer in some of the papers – they should not automatically all be assigned to deep learning. By definition deep learning comprises "deep" neural networks. However, the transition is not sharp and we expect to see an increase in the use of deep networks applied to PdM. While the accuracy achieved by deep learning is often superior to other methods, really deep networks were only used in a small number of the surveyed PdM use cases. One reason might be, that deep learning methods require enormous computing power due to a wide range of hyperparameters and a tremendous number of tuneable weights. In addition, deep learning requires large amounts of data. Hence, the engineering effort of a deep learning solution might be higher, as discussed in [173]. Furthermore, the use of a trained model in operation (in the inference step), requires powerful hardware, where specific hardware is nowadays offered by various manufacturers. Besides neural networks, different variants of support vector machines and ensembles are used in a number of publications.

Most approaches use black box models, calling for research in explainability and interpretability of the models. With the observed wide use of neural networks in the surveyed papers and the expected increase in the use of deeper networks, we argue that model interpretability will become a branch of the research field of ML-enabled PdM: A (deep) neural network is not interpretable by itself, they are often considered as black boxes. The fields of explainable AI (XAI) and interpretable ML address this shortcoming, as discussed in Section 6.4.

An overwhelming majority of papers rely on a fully labelled data set, making the availability of labelled data a major bottleneck. In Table 6 the surveyed use cases are categorized by their level of supervision, as discussed in Section 4, and by their categorization within PdM as it was shown in Fig. 3. It is noticeable that most papers used methods that rely on fully or partially labelled data. While the results are usually more reliable with (semi-)supervised learning, these methods require the labels to be available. In Section 6.2 some approaches were named that are promising for the generation of labelled data in an efficient way [151–156]. Note that, while for data recorded from a test bed labelled data can be obtained by injecting e.g. faults or wear, this is not possible when working on data recorded in the field. Hence, the necessity of a fully labelled data set is a major bottleneck when applying ML for predictive maintenance.

None of the surveyed papers used reinforcement learning, pointing to a potential research gap. As mentioned in Section 4, based on our search terms and exclusion criteria, no works on PdM and reinforcement learning applied to automotive applications are contained in the survey. However, there is research on reinforcement learning for PdM, with a potential to apply it in an automotive context. For example Ding et al. [174] tested their method on rotating machinery, which employs raw vibration signals under different health states and working conditions. More examples are discussed in Ran et al. [13]. These methods are promising, one reason being that they do not require a fully labelled data set.

Table A.7
List of the common abbreviations in machine learning and maintenance used in this paper.

Acronym	Topic
AI	Artificial intelligence
ML	Machine learning
DL	Deep learning
PdM	Predictive maintenance
CBM	Condition-based maintenance
CM	Condition-monitoring
EV	Electrical Vehicle
IoT	Internet of things
RUL	Remaining useful life
SoH	State of health
PHM	Prognostics and Health Management

Machine learning will not entirely replace physical models, but rather accompany them. The advantages of ML-solutions in PdM such as major cost savings, higher predictability, and the increased availability of the systems have attracted more and more researchers' and manufacturers' attention to adopting it. Currently, enormous expectations are set on ML. While it is undeniable that ML is transforming PdM, it is our opinion that it will not entirely replace classical methods. However, even in those cases where physical models are used, ML can be leveraged to build hybrid models, e.g. to model the dynamic behaviour of highly nonlinear components.

8. Conclusion

In this work we surveyed and categorized recent research contributions on ML-enabled PdM for automotive systems. In order to do so, a systematic literature research on Scopus was conducted, using a reproducible research methodology with defined exclusion and inclusion criteria. This yielded 62 papers which were surveyed and categorized with respect to (a) the corresponding use cases (in terms of the examined vehicle components), (b) the applied ML methods, (c) the ML tasks (e.g. classification, regression and clustering), and finally (d) the respective predictive maintenance categories, which where identified

Table A.8
Brief explanation of the machine learning methods predominantly used in the surveyed papers.

Method	Explanation
ANN, MLP	An artificial neural network (ANN) [53] consists of interconnected nodes organized in layers. A node's output is determined by the weighted sum of the input with some (typically non-linear) activation function applied on the sum. During training, the weights are tuned, such that some error function is minimized. There is a variety of different network architectures, making ANNs applicable for different data types and for different problem settings like classification, regression and anomaly detection. A common ANN architecture is the feed-forward multilayer perception (MLP) which consists of several layers where the nodes of layer L_i are solely connected with the nodes of layer L_{i+1}.
Auto-encoder	An autoencoder [74] is an unsupervised ANN that learns to reproduce the input data at the output layer. It thereby learns to capture the characteristics of the data, minimizing the reproduction error. In the test phase, data with different characteristics is reproduced with a high reproduction error, making an autoencoder applicable to anomaly detection.
CNN	A convolutional neural network (CNN) [175,176] is an ANN that exploits the neighbourhood between data points, e.g. in images, spectrograms or time series [177]. In the convolution layers, windows are moved over the data in order to learn filters that capture the characteristic features.
DL (Deep Learning)	Deep learning [53] comprises a variety of methods based on ANNs with a high number of layers, e.g. CNN, LSTM, MLP, Autoencoder. Deep learning has led to a boost in the accuracies of many ML applications.
ELM	Extreme learning machines (ELMs) are a variant of ANNs that are less computationally expensive to tune compared to standard ANNs. This is achieved by randomly assigning constant weights in the lower layers, solely tuning the weights in the upper layers.
ensemble methods	Ensemble methods combine multiple ML models, so-called base models (also called weak learners). The base models' decisions are combined to one overall decision. A famous example is a random forest, combining multiple decision trees.
gcForest	Deep forest algorithm which uses a multi-grain scanning approach for data slicing and a cascade structure of multiple random forests layers (see [145]).
LSTM, RNN	To model sequential data like time series, ANNs with a memory are used. Recurrent neural networks (RNN) model feed a node's output back into the network. Hence, an output depends not only on the current input but also on the previously determined output, which allows to model sequences. A commonly used variant of RNNs are long short-term memory (LSTM) networks [178], which solve issues in the training process of the original RNNs. LSTM can be used for classification [179], forecasting, and anomaly detection [180].
Random forest	A random forest is an ensemble method [54], combining multiple decision trees. The aim is to have diverse trees which is achieved by using different subsets of the training data and the feature space. The overall output is then determined by a majority vote over all trees. Random forests can be used for classification and regression and can cope with high-dimensional feature spaces.
SVM, SVR	For classification problems, support vector machines (SVM) (see e.g. [52]) determine the decision function by finding the hyperplane that maximizes the distance between the classes. SVMs can be enhanced to learn non-linear decision functions by transforming the data to a higher-dimensional space, where the data can be separated by the hyperplane. This is achieved by the so-called kernel trick, where the inner product of the data points is replaced by a kernel function. For regression problems, support vector regression (SVR) can be used, where the optimization problem is reformulated in order not to separate classes, but to minimize the error between the data and the hyperplane.
SVDD, one-class SVM	SVMs were reformulated as one-class classifiers and are frequently used for semi-supervised anomaly detection. There are two variants: ν-SVM [63] and support vector data description (SVDD) [62]. The ν-SVM [63] learns a hyperplane to separate normal data points and anomalies. SVDD finds a hypersphere around the normal data points such that the radius is minimized. Both methods are usually used in their soft-margin variant with kernel functions, allowing for flexible decision functions.

in advance with regard to maintenance benefit and complexity. As a complement to the survey, we conducted a bibliometric analysis to reveal influential papers, active authors, frequently addressed use cases and frequently utilized ML methods.

By addressing researchers and practitioners with a background either in machine learning, maintenance, or automotive engineering the paper aims to create the basis for interdisciplinary collaborations. A key contribution is the identification of open challenges and research directions, which may serve researchers to identify open research questions. A number of overriding conclusions can be drawn: (1) more publicly available data for automotive systems would lead to a boost in research activities, (2) ML-based PdM methods are promising to accompany the transformation of the drive train, (3) combining data from multiple sources can improve accuracies and enable new applications, (4) the use of deep learning methods in PdM is likely to increase further, but this requires tailored methods in terms of efficiency and interpretability as well as the availability of data.

While new insights were brought to light in this paper, the work is not without limitations. One limitation is, that the literature survey did not allow for a quantitative comparison of the results, due to entirely different problem settings and data sets. Consequently, we decided not to include the reported performance metrics (e.g. accuracy) in the comparison in order not to mislead readers to favour a supposedly superior ML method, which might be superior only due to the differences in the data set. A second limitation is, that we did not evaluate the surveyed approaches in terms of implementability, e.g. on constrained hardware resources when used as an on-board solution in a vehicle or regarding bandwidth requirements when used in a connected vehicle setting. A third limitation, also recognized in other reviews, is the variety of terms used to refer to ML approaches, particularly in older literature. This inevitably led to systematic searches overseeing relevant papers. We attempted to target this by extending our criteria, although there is always the chance to have missed some important contributions.

Future research could address the transferability of general PdM achievements to automotive use cases. For example in PdM there are benchmark data sets like CWRU or C-MAPPS, with a variety of publications. In addition, some of the research on PdM in manufacturing has the potential to be applicable for vehicles. As a concrete further step, we plan to use state-of-the art time series models like Inception Time [181], LSTM-FCN [182] and ROCKET [183] for PdM tasks.

Finally, in a more general context, we argue that ML has brought about a call for modification. At the same time that new technologies in the vehicle appear every day, new approaches to maintenance are being developed, from undertaking maintenance through the internet to computing complex calculations in-vehicle, such as sophisticated driving manoeuvres. These approaches are shifting maintenance away from error-prone models by utilizing data generated by the car to identify potential signs that could lead to downtime or failure. The ultimate goal is to conduct effective predictive maintenance before occurring issues impact the system as a whole. The surveyed use cases presented in this work show that ML can indeed effectively predict failures or abnormalities in a wide range of applications. As a final remark, we believe that machine learning has enhanced the set of tools for predictive maintenance and will continue to do so. This does, however, not mean, that ML will fully replace other approaches. In some cases hybrid models or pure physical models will still be the most reasonable choice.

CRediT authorship contribution statement

Andreas Theissler: Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Judith Pérez-Velázquez:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Marcel Kettelgerdes:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Gordon Elger:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Abbreviations and reference of ML methods

The commonly used abbreviations are listed in Table A.7. A brief reference of the machine learning methods that are predominantly used in the surveyed papers is given in Table A.8.

References

[1] The Manufacturer. Annual manufacturing report 2020. Technical report, Hennik Group; 2020.
[2] World Manufacturing Forum. 2019 world manufacturing forum report, skills for the future of manufacturing. Technical report, World Manufacturing Forum; 2019.
[3] Iung B, Levrat E, Marquez A Crespo, Erbe H. E-maintenance: principles, review and conceptual framework. IFAC Proc Vol 2007;40(19):18–29. http://dx.doi.org/10.3182/20071002-MX-4-3906.00005.
[4] Pecht MG, Kang M. Prognostics and health management of electronics: fundamentals, machine learning, and the internet of things. Wiley - IEEE, Wiley; 2018.
[5] Cachada A, Barbosa J, Leitño P, Gcraldcs CAS, Deusdado L, Costa J, et al. Maintenance 4.0: Intelligent and predictive maintenance system architecture. In: 2018 IEEE 23rd international conference on emerging technologies and factory automation, vol. 1. 2018, p. 139–46. http://dx.doi.org/10.1109/ETFA.2018.8502489.
[6] Bokrantz Jon, Skoogh Anders, Berlin Cecilia, Wuest Thorsten, Stahre Johan. Smart Maintenance: an empirically grounded conceptualization. Int J Prod Econ 2020;223:107534. http://dx.doi.org/10.1016/j.ijpe.2019.107534.
[7] Theissler Andreas. Detecting anomalies in multivariate time series from automotive systems [Ph.D. thesis], Brunel University London; 2013.
[8] Becker Jan, Helmle Michael, Pink Oliver. System architecture and safety requirements for automated driving. In: Automated driving: safer and more efficient future driving. Cham: Springer International Publishing; 2017, p. 265–83. http://dx.doi.org/10.1007/978-3-319-31895-0_11.
[9] Reschka Andreas. Safety concept for autonomous vehicles. In: Autonomous driving: technical, legal and social aspects. Berlin, Heidelberg: Springer Berlin Heidelberg; 2016, p. 473–96. http://dx.doi.org/10.1007/978-3-662-48847-8_23.
[10] Kim Junsung, Rajkumar Ragunathan (Raj), Jochim Markus. Towards dependable autonomous driving vehicles: A system-level approach. SIGBED Rev. 2013;10(1):29–32. http://dx.doi.org/10.1145/2492385.2492390.
[11] Zhao Y, Liu P, Wang Z, Zhang L, Hong J. Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods. Appl Energy 2017;207:354–62. http://dx.doi.org/10.1016/j.apenergy.2017.05.139.
[12] Nowakowski Tomasz, Tubis Agnieszka, Werbińska-Wojciechowska Sylwia. Evolution of technical systems maintenance approaches–review and a case study. In: International conference on intelligent systems in production engineering and maintenance. Springer; 2018, p. 161–74.
[13] Ran Yongyi, Zhou Xin, Lin Pengfeng, Wen Yonggang, Deng Ruilong. A survey of predictive maintenance: Systems, purposes and approaches. 2019.
[14] Sutharssan Thamo, Stoyanov Stoyan, Bailey Chris, Yin Chunyan. Prognostic and health management for engineering systems: a review of the data-driven approach and algorithms. J Eng 2015;2015(7):215–22.
[15] Peng Ying, Dong Ming, Zuo Ming. Current status of machine prognostics in condition-based maintenance: A review. Int J Adv Manuf Technol 2010;50:297–313. http://dx.doi.org/10.1007/s00170-009-2482-0.
[16] Werbińska-Wojciechowska Sylwia. Preventive maintenance models for technical systems. In: Technical system maintenance: delay-time-based modelling. Cham: Springer International Publishing; 2019, p. 21–100. http://dx.doi.org/10.1007/978-3-030-10788-8_2.
[17] Wu Shaomin. Preventive maintenance models: A review. In: Tadj Lotfi, Ouali M-Salah, Yacout Soumaya, Ait-Kadi Daoud, editors. Replacement models with minimal repair. London: Springer London; 2011, p. 129–40. http://dx.doi.org/10.1007/978-0-85729-215-5_4.
[18] Ali Yasir Hassan. Artificial intelligence application in machine condition monitoring and fault diagnosis. In: Artificial intelligence. Rijeka: IntechOpen; 2018, http://dx.doi.org/10.5772/intechopen.74932.
[19] Alsina Emanuel F, Chica Manuel, Trawiński Krzysztof, Regattieri Alberto. On the use of machine learning methods to predict component reliability from data-driven industrial case studies. Int J Adv Manuf Technol 2018;94(5):2419–33.
[20] Bhargava Cherry. AI techniques for reliability prediction for electronic components. Engineering Science Reference; 2019.

[21] Carvalho Thyago P, Soares Fabrízzio AAMN, Vita Roberto, da P. Francisco Roberto, Basto João P, Alcalá Symone GS. A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 2019;137:106024. http://dx.doi.org/10.1016/j.cie.2019.106024.

[22] Fink Olga, Wang Qin, Svensén Markus, Dersin Pierre, Lee Wan-Jui, Ducoffe Melanie. Potential, challenges and future directions for deep learning in prognostics and health management applications. Eng Appl Artif Intell 2020;92:103678. http://dx.doi.org/10.1016/j.engappai.2020.103678.

[23] Khan Samir, Yairi Takehisa. A review on the application of deep learning in system health management. Mech Syst Signal Process 2018;107:241–65. http://dx.doi.org/10.1016/j.ymssp.2017.11.024.

[24] Lei Yaguo, Yang Bin, Jiang Xinwei, Jia Feng, Li Naipeng, Nandi Asoke K. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mech Syst Signal Process 2020;138:106587.

[25] Schwabacher Mark, Goebel Kai. A survey of artificial intelligence for prognostics. In: AAAI fall symposium: artificial intelligence for prognostics. 2007, p. 108–15.

[26] Tsui Kwok-Leung, Chen Nan, Zhou Qiang, Hai Yizhen, Wang Wenbin. Prognostics and health management: A review on data driven approaches. Math Probl Eng 2015;2015:1–17.

[27] Wu Shaomin, Tsui Kwok L, Chen Nan, Zhou Qiang, Hai Yizhen, Wang Wenbin. Prognostics and health management: A review on data driven approaches. Math Probl Eng 2015;2015:793161.

[28] Wu Shaomin, Wu D, Peng R. Machine learning approaches in reliability and maintenance: classifications of recent literature. In: 2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling. 2020, p. 1–5.

[29] Zhao Rui, Yan Ruqiang, Chen Zhenghua, Mao Kezhi, Wang Peng, Gao Robert X. Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 2019;115:213–37. http://dx.doi.org/10.1016/j.ymssp.2018.05.050.

[30] Ahsan M, Stoyanov S, Bailey C. Prognostics of automotive electronics with data driven approach: A review. In: 39th international spring seminar on electronics technology. 2016, p. 279–84. http://dx.doi.org/10.1109/ISSE.2016.7563205.

[31] Mesgarpour Mohammad, Landa-Silva Dario, Dickinson Ian. Overview of telematics-based prognostics and health management systems for commercial vehicles. In: Mikulski Jerzy, editor. Activities of transport telematics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013, p. 123–30.

[32] Sankavaram C, Pattipati B, Kodali A, Pattipati K, Azam M, Kumar S, et al. Model-based and data-driven prognosis of automotive and electronic systems. In: 2009 IEEE international conference on automation science and engineering. 2009, p. 96–101. http://dx.doi.org/10.1109/COASE.2009.5234108.

[33] Li Jun, Cheng Hong, Guo Hongliang, Qiu Shaobo. Survey on artificial intelligence for vehicles. Autom Innov 2018;1(1):2–14.

[34] Falcini F, Lami G, Costanza A. Deep learning in automotive software. IEEE Softw 2017;34(03):56–63. http://dx.doi.org/10.1109/MS.2017.79.

[35] Singh Kanwar Bharat, Arat Mustafa Ali. Deep learning in the automotive industry: Recent advances and application examples. 2019.

[36] European Committee for Standardization, Brussels. PN-EN 13306: 2010 maintenance – maintenance terminology. 2010.

[37] Zhang Weiting, Yang Dong, Wang Hongchao. Data-driven methods for predictive maintenance of industrial equipment: A survey. IEEE Syst J 2019;13(3):2213–27.

[38] Yan Rundong, Dunnett Sarah J, Jackson Lisa M. Novel methodology for optimising the design, operation and maintenance of a multi-AGV system. Reliab Eng Syst Saf 2018;178:130–9.

[39] Li X, Ding Q, Sun J-Q. Remaining useful life estimation in prognostics using deep convolution neural networks. Reliab Eng Syst Saf 2018;172:1–11. http://dx.doi.org/10.1016/j.ress.2017.11.021.

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

[41] Chen Chong, Liu Ying, Sun Xianfang, Di Cairano-Gilfedder Carla, Titmus Scott. Automobile maintenance modelling using gcforest. In: 2020 IEEE 16th international conference on automation science and engineering. IEEE; 2020, p. 600–5.

[42] Russell Stuart, Norvig Peter. Artificial intelligence: A modern approach. 3rd ed. USA: Prentice Hall Press; 2009.

[43] Rocchetta R, Bellani L, Compare M, Zio E, Patelli E. A reinforcement learning framework for optimal operation and maintenance of power grids. Appl Energy 2019;241:291–301. http://dx.doi.org/10.1016/j.apenergy.2019.03.027.

[44] Zhang N, Si W. Deep reinforcement learning for condition-based maintenance planning of multi-component systems under dependent competing risks. Reliab Eng Syst Saf 2020;203. http://dx.doi.org/10.1016/j.ress.2020.107094.

[45] Kaelbling Leslie Pack, Littman Michael L, Moore Andrew W. Reinforcement learning: A survey. J Artificial Intelligence Res 1996;4:237–85.

[46] Arulkumaran Kai, Deisenroth Marc Peter, Brundage Miles, Bharath Anil Anthony. Deep reinforcement learning: A brief survey. IEEE Signal Process Mag 2017;34(6):26–38.

[47] Koprinkova-Hristova P. Reinforcement learning for predictive maintenance of industrial plants. Inf Technol Control 2013;11(1):21–8.

[48] Ong Kevin Shen Hoong, Niyato Dusit, Yuen Chau. Predictive maintenance for edge-based sensor networks: A deep reinforcement learning approach. In: 2020 IEEE 6th world forum on internet of things. IEEE; 2020, p. 1–6.

[49] Bezdek James C, Ehrlich Robert, Full William. FCM: The fuzzy c-means clustering algorithm. Comput Geosci 1984;10(2–3):191–203.

[50] Ester Martin, Kriegel Hans-Peter, Sander Jörg, Xu Xiaowei, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second international conference on knowledge discovery and data mining, vol. 96. 1996, p. 226–31.

[51] Campello Ricardo JGB, Moulavi Davoud, Sander Jörg. Density-based clustering based on hierarchical density estimates. In: Pacific-Asia conference on knowledge discovery and data mining. Springer; 2013, p. 160–72.

[52] Abe Shigeo. Support vector machines for pattern classification (advances in pattern recognition). 2nd ed. Springer-Verlag London Ltd.; 2010.

[53] LeCun Yann, Bengio Y, Hinton Geoffrey. Deep learning. Nature 2015;521:436–44. http://dx.doi.org/10.1038/nature14539.

[54] Breiman Leo. Random forests. Mach Learn 2001;45(1):5–32. http://dx.doi.org/10.1023/a:1010933404324.

[55] Chen Tianqi, Guestrin Carlos. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD int'l conf. on knowledge discovery and data mining. New York, NY, USA: Association for Computing Machinery; 2016, p. 785–94. http://dx.doi.org/10.1145/2939672.2939785.

[56] Tamersoy Acar, Khalil Elias, Xie Bo, Lenkey Stephen L, Routledge Bryan R, Chau Duen Horng, Navathe Shamkant B. Large-scale insider trading analysis: patterns and discoveries. Soc Netw Anal Min 2014;4(1):201.

[57] Garcia-Teodoro Pedro, Diaz-Verdejo Jesus, Maciá-Fernández Gabriel, Vázquez Enrique. Anomaly-based network intrusion detection: Techniques, systems and challenges. Comput Secur 2009;28(1–2):18–28.

[58] Avatefipour Omid, Al-Sumaiti Ameena Saad, El-Sherbeeny Ahmed M, Awwad Emad Mahrous, Elmeligy Mohammed A, Mohamed Mohamed A, et al. An intelligent secured framework for cyberattack detection in electric vehicles' CAN bus using machine learning. IEEE Access 2019;7:127580–92.

[59] Prisacaru Alexandru, Palczynska Alicja, Theissler Andreas, Gromala Przemyslaw, Han Bongtae, Zhang GQ. In situ failure detection of electronic control units using piezoresistive stress sensor. IEEE Trans Compon Packag Manuf Technol 2018;PP:1–14.

[60] Theissler Andreas. Anomaly detection in recordings from in-vehicle networks. In: Big data applications and principles. 2014.

[61] Theissler Andreas, Frey Stephan, Ehlert Jens. OCADaMi: One-class anomaly detection and data mining toolbox. In: Joint European conference on machine learning and knowledge discovery in databases 2019. Springer; 2020, p. 764–8.

[62] Tax David, Duin Robert. Support vector data description. Mach Learn 2004;54(1):45–66.

[63] Schölkopf Bernhard, Platt John C, Shawe-Taylor John C, Smola Alex J, Williamson Robert C. Estimating the support of a high-dimensional distribution. Neural Comput 2001;13:1443–71.

[64] Chandola Varun, Banerjee Arindam, Kumar Vipin. Anomaly detection: A survey. ACM Comput Surv 2009;41(3):15:1–58.

[65] ISO 26262-1. ISO 26262: Road vehicles – functional safety – part 1: Vocabulary. final draft. (ISO 26262-1). Geneva, Switzerland: ISO; 2011.

[66] Theissler Andreas. Multi-class novelty detection in diagnostic trouble codes from repair shops. In: 2017 IEEE 15th international conference on industrial informatics. IEEE; 2017, p. 1043–9. http://dx.doi.org/10.1109/INDIN.2017.8104917.

[67] Raissi Maziar, Perdikaris Paris, Karniadakis George E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707.

[68] Rai R, Sahu CK. Driven by data or derived through physics? A review of hybrid physics guided machine learning techniques with cyber-physical system (CPS) focus. IEEE Access 2020;8:71050–73. http://dx.doi.org/10.1109/ACCESS.2020.2987324.

[69] Breunig Markus M, Kriegel Hans-Peter, Ng Raymond T, Sander Jörg. LOF: Identifying density-based local outliers. In: SIGMOD conference. 2000, p. 93–104.

[70] He Zengyou, Xu Xiaofei, Deng Shengchun. Discovering cluster-based local outliers. Pattern Recognit Lett 2003;24(9–10):1641–50. http://dx.doi.org/10.1016/S0167-8655(03)00003-5.

[71] Liu Fei Tony, Ting Kai Ming, hua Zhou Zhi. Isolation forest. In: Proceedings of the 2008 eighth IEEE international conference on data mining. IEEE Computer Society; 2008, p. 413–22.

[72] Hariri S, Carrasco Kind M, Brunner RJ. Extended isolation forest. IEEE Trans Knowl Data Eng 2019.

[73] Zhou Chong, Paffenroth Randy C. Anomaly detection with robust deep autoencoders. In: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. New York, NY, USA: Association for Computing Machinery; 2017, p. 665–74. http://dx.doi.org/10.1145/3097983.3098052.

[74] Cao VL, Nicolau M, McDermott J. Learning neural representations for network anomaly detection. IEEE Trans Cybern 2019;49(8):3074–87.
[75] Ruff Lukas, Vandermeulen Robert, Goernitz Nico, Deecke Lucas, Siddiqui Shoaib Ahmed, Binder Alexander, et al. Deep one-class classification. In: Proceedings of ICML, vol. 80. Stockholmsmässan, Stockholm Sweden: PMLR; 2018, p. 4393–402.
[76] Sun Jiayu, Shao Jie, He Chengkun. Abnormal event detection for video surveillance using deep one-class learning. Multimedia Tools Appl 2019;78(3):3633–47. http://dx.doi.org/10.1007/s11042-017-5244-2.
[77] Krawczyk Bartosz. Learning from imbalanced data: open challenges and future directions. Prog Artif Intell 2016;5(4):221–32.
[78] Maaten Laurens van der, Hinton Geoffrey. Visualizing data using t-SNE. J Mach Learn Res 2008;9(86):2579–605.
[79] Akçay Samet, Atapour-Abarghouei Amir, Breckon Toby P. GANomaly: Semi-supervised anomaly detection via adversarial training. In: Computer vision. Cham: Springer International Publishing; 2019, p. 622–37.
[80] Akçay Samet, Atapour-Abarghouei Amir, Breckon Toby P. Skip-GANomaly: Skip connected and adversarially trained encoder-decoder anomaly detection. In: 2019 international joint conference on neural networks. 2019, p. 1–8.
[81] Wang X, Shen C, Xia M, Wang D, Zhu J, Zhu Z. Multi-scale deep intra-class transfer learning for bearing fault diagnosis. Reliab Eng Syst Saf 2020;202. http://dx.doi.org/10.1016/j.ress.2020.107050.
[82] Wei Dongdong, Han Te, Chu Fulei, Zuo Ming Jian. Weighted domain adaptation networks for machinery fault diagnosis. Mech Syst Signal Process 2021;158:107744. http://dx.doi.org/10.1016/j.ymssp.2021.107744.
[83] da Costa PRDO, Akçay A, Zhang Y, Kaymak U. Remaining useful lifetime prediction via deep domain adaptation. Reliab Eng Syst Saf 2020;195. http://dx.doi.org/10.1016/j.ress.2019.106682.
[84] Phillips J, Cripps E, Lau JW, Hodkiewicz MR. Classifying machinery condition using oil samples and binary logistic regression. Mech Syst Signal Process 2015;60:316–25. http://dx.doi.org/10.1016/j.ymssp.2014.12.020.
[85] Sankavaram C, Kodali A, Pattipati KR, Singh S. Incremental classifiers for data-driven fault diagnosis applied to automotive systems. IEEE Access 2015;3:407–19. http://dx.doi.org/10.1109/ACCESS.2015.2422833.
[86] Byttner S, Rögnvaldsson T, Svensson M. Consensus self-organized models for fault detection (COSMO). Eng Appl Artif Intell 2011;24(5):833–9. http://dx.doi.org/10.1016/j.engappai.2011.03.002.
[87] Fan Y, Nowaczyk S, Rögnvaldsson T. Incorporating expert knowledge into a self-organized approach for predicting compressor faults in a city bus fleet. Front Artif Intell Appl 2015;278:58–67. http://dx.doi.org/10.3233/978-1-61499-589-0-58.
[88] Killeen Patrick, Ding Bo, Kiringa Iluju, Yeap Tet. IoT-Based predictive maintenance for fleet management. Procedia Comput Sci 2019;151:607–13. http://dx.doi.org/10.1016/j.procs.2019.04.184, The 10th International Conference on Ambient Systems, Networks and Technologies (ANT 2019) / The 2nd International Conference on Emerging Data and Industry 4.0 (EDI40 2019) / Affiliated Workshops.
[89] Gardner Josh, Mroueh Jawad, Jenuwine Natalia, Weaverdyck Noah, Krassenstein Samuel, Farahi Arya, et al. Driving with data in the motor city: Mining and modeling vehicle fleet maintenance data. 2020.
[90] Khoshkangini Reza, Mashhadi Peyman Sheikholharam, Berck Peter, Gholami Shahbandi Saeed, Pashami Sepideh, Nowaczyk Sławomir, et al. Early prediction of quality issues in automotive modern industry. Information 2020;11(7):354.
[91] Wong PK, Zhong J, Yang Z, Vong CM. Sparse Bayesian extreme learning committee machine for engine simultaneous fault diagnosis. Neurocomputing 2016;174:331–43. http://dx.doi.org/10.1016/j.neucom.2015.02.097.
[92] Zhong J-H, Wong PK, Yang Z-X. Fault diagnosis of rotating machinery based on multiple probabilistic classifiers. Mech Syst Signal Process 2018;108:99–114. http://dx.doi.org/10.1016/j.ymssp.2018.02.009.
[93] Wang M-H, Chao K-H, Sung W-T, Huang G-J. Using ENN-1 for fault recognition of automotive engine. Expert Syst Appl 2010;37(4):2943–7. http://dx.doi.org/10.1016/j.eswa.2009.09.041.
[94] Zabihihesari Alireza, Ansari-Rad Saeed, Shirazi Farzad, Ayati Moosa. Fault detection and diagnosis of a 12-cylinder trainset diesel engine based on vibration signature analysis and neural network. Proc Inst Mech Eng C 2018;233:095440621877831. http://dx.doi.org/10.1177/0954406218778313.
[95] Jung D, Sundström C. A combined data-driven and model-based residual selection algorithm for fault detection and isolation. IEEE Trans Control Syst Technol 2019;27(2):616–30. http://dx.doi.org/10.1109/TCST.2017.2773514.
[96] Wolf P, Mrowca A, Nguyen TT, Baker B, Gunnemann S. Pre-ignition detection using deep neural networks: A step towards data-driven automotive diagnostics. In: IEEE conference on intelligent transportation systems, proceedings, ITSC, vol. 2018-November. 2018, p. 176–83. http://dx.doi.org/10.1109/ITSC.2018.8569908.
[97] Jung Daniel. Data-driven open-set fault classification of residual data using Bayesian filtering. IEEE Trans Control Syst Technol 2020;28(5):2045–52.
[98] Wang YS, Liu NN, Guo H, Wang XL. An engine-fault-diagnosis system based on sound intensity analysis and wavelet packet pre-processing neural network. Eng Appl Artif Intell 2020;94:103765.
[99] Pan H, Lü Z, Wang H, Wei H, Chen L. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. Energy 2018;160:466–77. http://dx.doi.org/10.1016/j.energy.2018.06.220.
[100] You G, Park S, Oh D. Diagnosis of electric vehicle batteries using recurrent neural networks. IEEE Trans Ind Electron 2017;64(6):4885–93. http://dx.doi.org/10.1109/TIE.2017.2674593.
[101] Yang R, Xiong R, He H, Chen Z. A fractional-order model-based battery external short circuit fault diagnosis approach for all-climate electric vehicles application. J Cleaner Prod 2018;187:950–9. http://dx.doi.org/10.1016/j.jclepro.2018.03.259.
[102] Quintián H, Casteleiro-Roca J-L, Perez-Castelo FJ, Calvo-Rolle JL, Corchado E. Hybrid intelligent model for fault detection of a Lithium Iron Phosphate power cell used in electric vehicles. Lect Notes Comput Sci 2016;9648:751–62. http://dx.doi.org/10.1007/978-3-319-32034-2_63.
[103] Sankavaram C, Pattipati B, Pattipati K, Zhang Y, Howell M, Salman M. Data-driven fault diagnosis in a hybrid electric vehicle regenerative braking system. In: IEEE aerospace conference proceedings. 2012, http://dx.doi.org/10.1109/AERO.2012.6187368.
[104] Theissler A. Detecting known and unknown faults in automotive systems using ensemble-based anomaly detection. Knowl-Based Syst 2017;123:163–73. http://dx.doi.org/10.1016/j.knosys.2017.02.023.
[105] Shafi U, Safi A, Shahid AR, Ziauddin S, Saleem MQ. Vehicle remote health monitoring and prognostic maintenance system. J Adv Transp 2018;2018. http://dx.doi.org/10.1155/2018/8061514.
[106] Tagawa T, Tadokoro Y, Yairi T. Structured denoising autoencoder for fault detection and analysis. J Mach Learn Res 2014;39(2014):96–111.
[107] Routray A, Rajaguru A, Singh S. Data reduction and clustering techniques for fault detection and diagnosis in automotives. In: 2010 IEEE international conference on automation science and engineering. 2010, p. 326–31. http://dx.doi.org/10.1109/COASE.2010.5584324.
[108] Rengasamy D, Jafari M, Rothwell B, Chen X, Figueredo GP. Deep learning with dynamically weighted loss function for sensor-based prognostics and health management. Sensors (Switzerland) 2020;20(3). http://dx.doi.org/10.3390/s20030723.
[109] Cerqueira V, Pinto F, Sá C, Soares C. Combining boosted trees with metafeature engineering for predictive maintenance. Lect Notes Comput Sci 2016;9897 LNCS:393–7. http://dx.doi.org/10.1007/978-3-319-46349-0_35.
[110] Nowaczyk S, Prytz R, Rögnvaldsson T, Byttner S. Towards a machine learning algorithm for predicting truck compressor failures using logged vehicle data. Front Artif Intell Appl 2013;257:205–14. http://dx.doi.org/10.3233/978-1-61499-330-8-205.
[111] Heidari Bafroui H, Ohadi A. Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. Neurocomputing 2014;133:437–45. http://dx.doi.org/10.1016/j.neucom.2013.12.018.
[112] Gharavian MH, Almas Ganj F, Ohadi AR, Heidari Bafroui H. Comparison of FDA-based and PCA-based features in fault diagnosis of automobile gearboxes. Neurocomputing 2013;121:150–9. http://dx.doi.org/10.1016/j.neucom.2013.04.033.
[113] Zhang T, Li Z, Deng Z, Hu B. Hybrid data fusion DBN for intelligent fault diagnosis of vehicle reducers. Sensors (Switzerland) 2019;19(11). http://dx.doi.org/10.3390/s19112504.
[114] Yin S, Huang Z. Performance monitoring for vehicle suspension system via fuzzy positivistic C-means clustering based on accelerometer measurements. IEEE/ASME Trans Mechatronics 2015;20(5):2613–20. http://dx.doi.org/10.1109/TMECH.2014.2358674.
[115] Wang G, Yin S. Data-driven fault diagnosis for an automobile suspension system by using a clustering based method. J Franklin Inst B 2014;351(6):3231–44. http://dx.doi.org/10.1016/j.jfranklin.2014.03.004.
[116] Zehelein Thomas, Hemmert-Pottmann Thomas, Lienkamp Markus. Diagnosing automotive damper defects using convolutional neural networks and electronic stability control sensor signals. J Sensor Actuator Netw 2020;9:8. http://dx.doi.org/10.3390/jsan9010008.
[117] Capriglione D, Carratù M, Liguori C, Paciello V, Sommella P. A soft stroke sensor for motorcycle rear suspension. Measurement 2017;106:46–52.
[118] Capriglione Domenico, Carratù Marco, Pietrosanto Antonio, Sommella Paolo. NARX ANN-Based instrument fault detection in motorcycle. Measurement 2018;117:304–11.
[119] Jeong Kicheol, Choi Seibum B, Choi Hyungjeen. Sensor fault detection and isolation using a support vector machine for vehicle suspension systems. IEEE Trans Veh Technol 2020;69(4):3852–63.
[120] Jegadeeshwaran R, Sugumaran V. Brake fault diagnosis using clonal selection classification algorithm (CSCA) – A statistical learning approach. Eng Sci Technol Int J 2015;18(1):14–23. http://dx.doi.org/10.1016/j.jestch.2014.08.001.
[121] Alamelu Manghai TM, Jegadeeshwaran R. Vibration based brake health monitoring using wavelet features: A machine learning approach. JVC/J Vib Control 2019;25(18):2534–50. http://dx.doi.org/10.1177/1077546319859704.

[122] Ghimire R, Sankavaram C, Ghahari A, Pattipati K, Ghoneim Y, Howell M, et al. Integrated model-based and data-driven fault detection and diagnosis approach for an automotive electric power steering system. In: AUTOTESTCON (proceedings). 2011, p. 70–7. http://dx.doi.org/10.1109/AUTEST.2011.6058760.

[123] Ghimire R, Zhang C, Pattipati KR. A rough set-theory-based fault-diagnosis method for an electric power-steering system. IEEE/ASME Trans Mechatronics 2018;23(5):2042–53. http://dx.doi.org/10.1109/TMECH.2018.2863119.

[124] Fang Y, Cheng C, Dong Z, Min H, Zhao X. A fault diagnosis framework for autonomous vehicles based on hybrid data analysis methods combined with fuzzy PID control. In: 2020 3rd international conference on unmanned systems. 2020, p. 281–6. http://dx.doi.org/10.1109/ICUS50048.2020.9274856.

[125] Siegel Joshua E, Bhattacharyya Rahul, Sarma Sanjay, Deshpande Ajay. Smartphone-based wheel imbalance detection. In: ASME 2015 dynamic systems and control conference. American Society of Mechanical Engineers Digital Collection; 2015, http://dx.doi.org/10.1115/DSCC2015-9716.

[126] Siegel Joshua E, Sun Yongbin, Sarma Sanjay. Automotive diagnostics as a service: An artificially intelligent mobile application for tire condition assessment. In: Aiello Marco, Yang Yujiu, Zou Yuexian, Zhang Liang-Jie, editors. Artificial intelligence and mobile services. Cham: Springer International Publishing; 2018, p. 172–84.

[127] Mohammadi A, Djerdir A, Steiner NY, Bouquain D, Khaburi D. Diagnosis of PEMFC for automotive application. In: Proceedings: 2015 5th international youth conference on energy. 2015, http://dx.doi.org/10.1109/IYCE.2015.7180793.

[128] Zuo Jian, Lv Hong, Zhou Daming, Xue Qiong, Jin Liming, Zhou Wei, et al. Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. Appl Energy 2021;281:115937. http://dx.doi.org/10.1016/j.apenergy.2020.115937.

[129] Wu J-D, Kuo J-M. Fault conditions classification of automotive generator using an adaptive neuro-fuzzy inference system. Expert Syst Appl 2010;37(12):7901–7. http://dx.doi.org/10.1016/j.eswa.2010.04.046.

[130] Peters B, Yildirim M, Gebraeel N, Paynabar K. Severity-based diagnosis for vehicular electric systems with multiple, interacting fault modes. Reliab Eng Syst Saf 2020;195. http://dx.doi.org/10.1016/j.ress.2019.106605.

[131] Şimşir Mehmet, Bayir Raif, Uyaroğlu Yılmaz. Real-time monitoring and fault diagnosis of a low power hub motor using feedforward neural network. Comput Intell Neurosci 2016;2016:1–13. http://dx.doi.org/10.1155/2016/7129376.

[132] Seera Manjeevan, Lim Chee Peng, Nahavandi Saeid, Loo Chu Kiong. Condition monitoring of induction motors: A review and an application of an ensemble of hybrid intelligent models. Expert Syst Appl 2014;41(10):4891–903. http://dx.doi.org/10.1016/j.eswa.2014.02.028.

[133] Jeong YiNa, Son SuRak, Jeong EunHee, Lee ByungKwan. An integrated self-diagnosis system for an autonomous vehicle based on an IoT gateway and deep learning. Appl Sci 2018;8:1164. http://dx.doi.org/10.3390/app8071164.

[134] Van Wyk F, Wang Y, Khojandi A, Masoud N. Real-time sensor anomaly detection and identification in automated vehicles. IEEE Trans Intell Transp Syst 2020;21(3):1264–76. http://dx.doi.org/10.1109/TITS.2019.2906038.

[135] Rezvani Mohammad, AbuAli Mohamed, Lee Seungchul, Lee Jay, Ni Jun. A comparative analysis of techniques for electric vehicle battery prognostics and health management (PHM). SAE Tech Pap 2011. http://dx.doi.org/10.4271/2011-01-2247.

[136] Last M, Sinaiski A, Subramania HS. Condition-based maintenance with multi-target classification models. New Gener Comput 2011;29(3):245–60. http://dx.doi.org/10.1007/s00354-010-0301-7.

[137] Wu Y, Xue Q, Shen J, Lei Z, Chen Z, Liu Y. State of health estimation for lithium-ion batteries based on healthy features and long short-term memory. IEEE Access 2020;8:28533–47.

[138] Wang Zhi-Hao, Hendrick, Horng Gwo-Jiun, Wu Hsin-Te, Jong Gwo-Jia. A prediction method for voltage and lifetime of lead–acid battery by using machine learning. Energy Explor Exploit 2020;38(1):310–29. http://dx.doi.org/10.1177/0144598719881223.

[139] Prytz Rune, Nowaczyk Sławomir, Rögnvaldsson Thorsteinn, Byttner Stefan. Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. Eng Appl Artif Intell 2015;41:139–50.

[140] Taie MA, Diab M, Elhelw M. Remote prognosis, diagnosis and maintenance for automotive architecture based on least squares support vector machine and multiple classifiers. In: International congress on ultra modern telecommunications and control systems and workshops. 2012, p. 128–34. http://dx.doi.org/10.1109/ICUMT.2012.6459652.

[141] Lee C-Y, Huang T-S, Liu M-K, Lan C-Y. Data science for vibration heteroscedasticity and predictive maintenance of rotary bearings. Energies 2019;12(5). http://dx.doi.org/10.3390/en12050801.

[142] Al-Dahidi S, Di Maio F, Baraldi P, Zio E. Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. Reliab Eng Syst Saf 2016;156:109–24. http://dx.doi.org/10.1016/j.ress.2016.07.019.

[143] Wee Bert Van, Banister David. How to write a literature review paper? Transp Rev 2016;36(2):278–88.

[144] Breunig Markus M, Kriegel Hans-Peter, Ng Raymond T, Sander Jörg. LOF: identifying density-based local outliers. In: Proceedings of the 2000 ACM SIGMOD international conference on management of data. 2000, p. 93–104.

[145] Zhou Zhi-Hua, Feng Ji. Deep forest. Natl Sci Rev 2019;6(1):74–86.

[146] Lindgren Tony, Biteus Jonas. APS failure at scania trucks. Scania CV AB; 2016, URL https://www.kaggle.com/uciml/aps-failure-at-scania-trucks-data-set.

[147] Liu Li, Chen Jie, Fieguth Paul, Zhao Guoying, Chellappa Rama, Pietikäinen Matti. From BoW to CNN: Two decades of texture representation for texture classification. Int J Comput Vis 2019;127(1):74–109.

[148] Deng Jia, Dong Wei, Socher Richard, Li Li-Jia, Li Kai, Fei-Fei Li. Imagenet: A large-scale hierarchical image database. In: IEEE conference on computer vision and pattern recognition. Ieee; 2009, p. 248–55.

[149] Chen Yanping, Keogh Eamonn, Hu Bing, Begum Nurjahan, Bagnall Anthony, Mueen Abdullah, et al. The UCR time series classification archive. 2015, https://www.cs.ucr.edu/~eamonn/time_series_data/.

[150] Bagnall Anthony, Lines Jason, Bostrom Aaron, Large James, Keogh Eamonn. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min Knowl Discov 2017;31(3):606–60.

[151] Bernard Jürgen, Zeppelzauer Matthias, Sedlmair Michael, Aigner Wolfgang. VIAL: a unified process for visual interactive labeling. Vis Comput 2018;34(9):1189–207. http://dx.doi.org/10.1007/s00371-018-1500-3.

[152] Bernard Jurgen, Hutter Marco, Zeppelzauer Matthias, Fellner Dieter, Sedlmair Michael. Comparing visual-interactive labeling with active learning: An experimental study. IEEE Trans Vis Comput Graphics 2018;24(1):298–308. http://dx.doi.org/10.1109/tvcg.2017.2744818.

[153] Grimmeisen Benedikt, Theissler Andreas. The machine learning model as a guide: Pointing users to interesting instances for labeling through visual cues. In: The 13th international symposium on visual information communication and interaction. ACM; 2020, http://dx.doi.org/10.1145/3430036.3430058.

[154] Beil David, Theissler Andreas. Cluster-clean-label: An interactive machine learning approach for labeling high-dimensional data. In: The 13th international symposium on visual information communication and interaction. ACM; 2020, http://dx.doi.org/10.1145/3430036.3430060.

[155] Fu Yifan, Zhu Xingquan, Li Bin. A survey on instance selection for active learning. Knowl Inf Syst 2012;35(2):249–83. http://dx.doi.org/10.1007/s10115-012-0507-8.

[156] Settles Burr. Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences; 2009.

[157] Theissler Andreas, Kraft Anna-Lena, Rudeck Max, Erlenbusch Fabian. VIAL-AD: Visual interactive labelling for anomaly detection – An approach and open research questions. In: 4th international workshop on interactive adaptive learning. CEUR-WS; 2020.

[158] Shorten Connor, Khoshgoftaar Taghi M. A survey on image data augmentation for deep learning. J Big Data 2019;6(1):60.

[159] Chawla Nitesh V, Bowyer Kevin W, Hall Lawrence O, Kegelmeyer W Philip. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57.

[160] Lu Jie, Liu Anjin, Dong Fan, Gu Feng, Gama Joao, Zhang Guangquan. Learning under concept drift: A review. IEEE Trans Knowl Data Eng 2018;31(12):2346–63.

[161] Webb Geoffrey I, Hyde Roy, Cao Hong, Nguyen Hai Long, Petitjean Francois. Characterizing concept drift. Data Min Knowl Discov 2016;30(4):964–94.

[162] Senge Robin, Bösner Stefan, Dembczyński Krzysztof, Haasenritter Jörg, Hirsch Oliver, Donner-Banzhoff Norbert, et al. Reliable classification: Learning classifiers that distinguish aleatoric and epistemic uncertainty. Inf Sci 2014;255:16–29.

[163] Yang Qiang, Liu Yang, Chen Tianjian, Tong Yongxin. Federated machine learning: Concept and applications. ACM Trans Intell Syst Technol (TIST) 2019;10(2):1–19.

[164] Ribeiro Marco Tulio, Singh Sameer, Guestrin Carlos. "Why should I trust you? "Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, p. 1135–44.

[165] Samek Wojciech, Montavon Grégoire, Vedaldi Andrea, Hansen Lars Kai, Müller Klaus-Robert. Explainable AI: interpreting, explaining and visualizing deep learning, vol. 11700. Springer Nature; 2019.

[166] Guidotti Riccardo, Monreale Anna, Ruggieri Salvatore, Turini Franco, Giannotti Fosca, Pedreschi Dino. A survey of methods for explaining black box models. ACM Comput Surv (CSUR) 2018;51(5):1–42.

[167] Rudin Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019;1(5):206–15.

[168] Gambs Sébastien, Killijian Marc-Olivier, del Prado Cortez Miguel Núñez. De-anonymization attack on geolocated data. J Comput Syst Sci 2014;80(8):1597–614.

[169] Akhtar Naveed, Mian Ajmal. Threat of adversarial attacks on deep learning in computer vision: A survey. IEEE Access 2018;6:14410–30.

[170] Cao Yulong, Xiao Chaowei, Cyr Benjamin, Zhou Yimeng, Park Won, Rampazzi Sara, et al. Adversarial sensor attack on lidar-based perception in autonomous driving. In: Proceedings of the 2019 ACM SIGSAC conference on computer and communications security. 2019, p. 2267–81.

[171] Strauss Thilo, Hanselmann Markus, Junginger Andrej, Ulmer Holger. Ensemble methods as a defense to adversarial perturbations against deep neural networks. 2017, ArXiv Preprint ArXiv:1709.03423.

[172] Sargolzaei A, Crane CD, Abbaspour A, Noei S. A machine learning approach for fault detection in vehicular cyber-physical systems. In: 2016 15th IEEE international conference on machine learning and applications. 2016, p. 636–40. http://dx.doi.org/10.1109/ICMLA.2016.0112.

[173] Grüner Tobias, Böllhoff Falco, Meisetschläger Robert, Vydrenko Alexander, Bator Martyna, Dicks Alexander, et al. Evaluation of machine learning for sensorless detection and classification of faults in electromechanical drive systems. Procedia Comput Sci 2020;176:1586–95.

[174] Ding Yu, Ma Liang, Ma Jian, Suo Mingliang, Tao Laifa, Cheng Yujie, et al. Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach. Adv Eng Inf 2019;42:100977. http://dx.doi.org/10.1016/j.aei.2019.100977.

[175] LeCun Yann, Bengio Yoshua. The handbook of brain theory and neural networks. Cambridge, MA, USA: MIT Press; 1998, p. 255–8,

[176] Rawat Waseem, Wang Zenghui. Deep convolutional neural networks for image classification: A comprehensive review. Neural Comput 2017;29:1–98. http://dx.doi.org/10.1162/NECO_a_00990.

[177] Fawaz Hassan Ismail, Forestier Germain, Weber Jonathan, Idoumghar Lhassane, Muller Pierre-Alain. Deep learning for time series classification: a review. Data Min Knowl Discov 2019;33(4):917–63.

[178] Hochreiter Sepp, Schmidhuber Jürgen. Long short-term memory. Neural Comput 1997;9(8):1735–80. http://dx.doi.org/10.1162/neco.1997.9.8.1735.

[179] Sak Hasim, Senior Andrew W, Beaufays Francoise. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: Interspeech. 2014.

[180] Malhotra Pankaj, Vig Lovekesh, Shroff Gautam, Agarwal Puneet. Long short term memory networks for anomaly detection in time series. In: Proceedings, vol. 89. Presses universitaires de Louvain; 2015, p. 89–94.

[181] Fawaz Hassan Ismail, Lucas Benjamin, Forestier Germain, Pelletier Charlotte, Schmidt Daniel F, Weber Jonathan, et al. InceptionTime: Finding AlexNet for time series classification. Data Min Knowl Discov 2020;34(6):1936–62.

[182] Karim Fazle, Majumdar Somshubra, Darabi Houshang, Chen Shun. LSTM Fully convolutional networks for time series classification. IEEE Access 2017;6:1662–9.

[183] Dempster Angus, Petitjean François, Webb Geoffrey I. ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels. Data Min Knowl Discov 2020;34(5):1454–95.