Advanced Engineering Informatics 45 (2020) 101101

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing

Jacqueline Schmitt[a,*], Jochen Bönig[b], Thorbjörn Borggräfe[a], Gunter Beitinger[b], Jochen Deuse[a,c]

[a] Institute of Production Systems, TU Dortmund University, Leonhard-Euler-Str. 5, 44227 Dortmund, Germany
[b] Siemens AG, Werner-von-Siemens-Str. 50, 92224 Amberg, Germany
[c] Advanced Manufacturing, School of Mechanical and Mechatronic Engineering, University of Technology Sydney, Sydney, Australia

ARTICLE INFO

Keywords:
Edge Cloud Computing
Machine Learning
Quality inspection
Quality prediction
Manufacturing

ABSTRACT

The supply of defect-free, high-quality products is an important success factor for the long-term competitiveness of manufacturing companies. Despite the increasing challenges of rising product variety and complexity and the necessity of economic manufacturing, a comprehensive and reliable quality inspection is often indispensable. In consequence, high inspection volumes turn inspection processes into manufacturing bottlenecks.

In this contribution, we investigate a new integrated solution of predictive model-based quality inspection in industrial manufacturing by utilizing Machine Learning techniques and Edge Cloud Computing technology. In contrast to state-of-the-art contributions, we propose a holistic approach comprising the target-oriented data acquisition and processing, modelling and model deployment as well as the technological implementation in the existing IT plant infrastructure. A real industrial use case in SMT manufacturing is presented to underline the procedure and benefits of the proposed method. The results show that by employing the proposed method, inspection volumes can be reduced significantly and thus economic advantages can be generated.

1. Introduction

As a result of increasing competitive pressure, the supply of high-quality products continues to evolve as an important competitive factor to secure the long-term success of a company. In order to guarantee the delivery and transfer of zero-defect products, it is essential to ensure a constantly high quality for all products. Additionally, in the ever-growing personalization paradigm, the number of variants and thus the complexity of inspection planning and operation increase tremendously. The design of inspection processes is therefore an extremely important and economically critical procedure, which requires the application of the latest and most sophisticated technologies.

In the era of Industry 4.0, appliances are enabled to bring benefits, including personalization, prediction, energy savings, defect reductions, and quality improvement [1]. Industry 4.0 denotes the trend towards automation and data exchange in manufacturing technologies and processes, including Cyber-Physical Systems (CPS), the Internet of Things (IoT), cloud computing, and Artificial Intelligence (AI) [2]. CPS constitute a new generation of systems with integrated computational and physical capabilities that enable the interaction with humans through new modalities [3]. The IoT is designated a key enabler for the next generation of advanced manufacturing [4], describing the technologies of a global infrastructure which allows to connect physical and virtual objects through information and communication technologies (ICT). Cloud computing is dominating today's computation as it enables on-demand and convenient access to a large pool of scalable and configurable computing resources [5]. AI has numerous applications in manufacturing, such as predictive analytics, quality inspection, intelligent automation and sensors, etc., which are based on different AI technologies [6]. One of the most relevant AI technologies is Machine Learning (ML), which offers great potential for the development and integration of strategies for optimizing products and manufacturing processes [7]. Applying statistical methods to structured and unstructured databases allows to extract previously unknown patterns and laws to generate new knowledge [8,9]. This enables the formation of prediction models for data-based and computer-aided prediction of future events [10].

In this contradictory field between constantly growing requirements and new technological possibilities, the contribution of this paper is the development of an integrated solution for predictive model-based quality inspection in industrial manufacturing. The core element is a prediction model based on supervised ML algorithms that allows to predict the final product quality on the basis of recorded process parameters. Additionally, the solution comprises a data preprocessing

* Corresponding author.
E-mail address: jacqueline.schmitt@ips.tu-dortmund.de (J. Schmitt).

https://doi.org/10.1016/j.aei.2020.101101
Received 12 May 2019; Received in revised form 9 April 2020; Accepted 15 April 2020
Available online 20 May 2020

module as well as a model evaluation and deployment module, which allows to interpret the prediction result and enables quality-based process-integrated decision support. The solution can be integrated into the IoT-architecture of the manufacturing plant and connected to other databases and solutions via web services. The results of a case study in surface mount technology (SMT) assembly in electronics manufacturing are presented to highlight the potential of the proposed methodology.

The paper is organized as follows. Section 2 introduces the theoretical background of the proposed solution. The solution framework is proposed in Section 3. Section 4 displays a case study in a real industrial setting. Section 5 concludes the paper.

2. Theoretical background

The present work is based upon two main fields of research, Machine Learning and Edge Cloud Computing, and builds on existing work in the field of model-based quality inspection in manufacturing. Consequently, a comprehensive overview of these fields is put forward in this section.

2.1. Machine learning

In recent years, ML has provided advantages in various fields of application, where the success can be credited to the invention of more sophisticated ML models [11,12], the availability of large data sets [13,14], and the development of software platforms [15,16] that allow easy employment of vast computational resources for training ML models on large data sets [17].

ML is a subfield of Artificial Intelligence that enables information technology (IT) systems to recognize patterns and laws on the basis of existing data and algorithms and develop solutions autonomously [18]. Hence, it is a collective term for the artificial generation of knowledge from experience. The knowledge gained from data can then be generalized and used to solve new problems and analyze previously unknown data. A central role in ML are algorithms, which are responsible for the recognition of patterns and generation of solutions. They can be categorized according to different learning paradigms into [19,20]:

- supervised learning,
- unsupervised learning,
- semi-supervised learning,
- reinforcement learning, and
- active learning.

Supervised learning refers to training models based on labeled training data. This entails the training of models by taking the expected outcome into account, e.g. the classification group. In unsupervised learning, on the other hand, the model groups are formed automatically on the basis of independently recognized patterns [21]. Semi-supervised learning is located between supervised and unsupervised learning. It has gained increasing importance recently, as fully labeled data sets are often not available or can only be generated with high costs. The method of reinforcement learning uses rewards and penalties to improve model performance. Active learning aims at finding useful rather than merely statistical findings. Thereby, instead of using statistical evaluations, the supervising user is asked to provide feedback on a question from which the algorithm should learn in a targeted manner [22]. Despite their different approaches, all learning tasks require algorithms to solve the anticipated problem.

2.2. Edge cloud computing

Computation today is dominated by the two trends of cloud computing and mobile computing. Even though mobile devices improve in storage and processing power [23] in accordance with Moore's Law [24], large data volumes and sophisticated ML models evolve at the same rate. Therefore, high level applications continue to have a cloud-based backend.

Cloud computing is a new computing model enabled by the rapid development of processing and storage technologies and the success of the Internet [25]. In cloud computing resources are provided as general utilities that can be internally or externally hosted, leased and released on-demand by users via the Internet [25]. The key advantages of cloud computing are [26,27]:

- cost-effective and dynamic access to large amounts of computing power,
- almost immediate access to hardware resources without upfront capital investments,
- lower barriers to innovation,
- easy dynamic scaling of enterprise services, and
- enabling of new classes of applications and services such as
  - location-, environment- and context-aware mobile interactive applications with real-time response,
  - parallel batch processing,
  - resource-intensive business analytics, and
  - extensions of compute-intensive desktop applications.

Besides several important benefits, cloud computing also entails some disadvantages. Because cloud services, especially hosted cloud services, are usually remote, they can suffer from latency- and bandwidth-related issues [28]. Additionally, hosted cloud services are used by multiple customers, implicating various issues for customers sharing the same hardware. Further, data access to third parties such as cloud service providers can cause security, compliance, and regulatory issues [28]. Relocating all computing task to the cloud has been efficient to process data because of the enormous computing power. However, despite data-processing speeds having increased rapidly, the bandwidth of the networks has not advanced sufficiently. Because of increasing amounts of generated data, the network thus becomes the bottleneck of cloud computing [29]. The ability to execute computations at the network edge near the data sources is enabled by the technology of edge computing [29]. An edge device can be any computing or networking resource, located between data sources and cloud-based data storages. The devices not only consume and produce data, but also handle computing tasks such as processing, storage, caching, and load balancing and exchange data with the cloud [29].

The main demand for preparation and deployment of ML-applications on an edge-device is calculating power. Calculations in IT-environments are executed on central processing units (CPUs) of PCs. In order to handle not only current but also more complex future assignments, a system should be more than satisfactorily capable regarding computational power. A distinct understanding of the required calculating power for a specific analytics task can also facilitate the implementation of more task-specific hardware [30]. After the first introduction of executing matrix-multiplications on a graphics processing unit (GPU) in 2001, the release of NVIDIA CUDA as a high-language to execute operations on a GPU in 2006 made GPUs available for broad usage in non-graphical use-cases, such as Finite Element Methods (FEM) and ML. These tasks highly benefit from the parallelized architecture of GPUs [31]. For large scale applications, Field Programmable Gate Arrays (FPGA) as well as application-specific integrated circuits (ASICs) for analytics and ML (e.g. Google TPU) provide superior performance compared to CPUs [32], can vastly reduce energy consumption and boost the overall model performance [30].

To exploit the advantages and overcome the disadvantages of both technologies, they are combined in Edge Cloud Computing that allows to collect and analyze data in real time while avoiding excessive data transfer and response delays [33]. The computation power of the cloud is beneficial for predictive model-based quality inspection to train sophisticated models on large historic data sets and store models. Handling of online process data and the model application, however, frequently has to take place in (near) real time to yield gainful inspection decisions. This circumstance restricts cloud-based model-execution because of network latency, additionally colliding with proposed long-term visions of several connected data sources that would further strain network band-width and reliability. To avoid latency- and bandwidth-problems, data processing and model application performed on the edge endows shorter response times, more efficient processing and less pressure on the network.

2.3. Model-based quality inspection

The product quality is essential for the long-term success of a producing company [34] and the economic realization of a comprehensive, reliable quality inspection is therefore of great interest. Manufacturing metrology, conventionally used in quality control, progressively reaches its limits due to the increasing requirements for speed, accuracy, safety, and flexibility [35]. Advanced manufacturing technologies such as data-driven approaches are therefore highly favored to overcome given limitations and to meet recent requirements.

While the recent state of research contains some literature reviews on general applications of ML in manufacturing, e.g. [36–38], specific reviews with focus on quality-related applications are rarely found, e.g. [39]. According to Köksal et al. [39] and Rostami et al. [40], different quality tasks for the application of ML in manufacturing can be distinguished:

- Description of product/process quality,
- Classification of quality,
- Quality prediction, and
- Parameter optimization.

The description of the product or process quality is usually the first step in respective quality-related projects, especially in complex, highly dynamic systems with multi-factorial and non-linear interactions. In the next step, a predictive model maps the available quality-related input information and data, e.g. master data, operating states or process parameters, to the resulting product quality [41]. This model can subsequently be used to predict quality feature values from a given set of input parameter values which allows a variety of measures to be applied in order to achieve an economic manufacturing [39,42,43]:

- Reduction of scrap through early control interventions,
- Optimization of process parameter settings and product quality,
- Stabilization of processes,
- Dynamization of inspection plans, and
- Design of model-based inspection processes.

Most theoretical contributions only focus on the development of new methods and algorithms without reference to a specific application case. Only a few authors propose new methods to solve a specific industrial problem, e.g. Wan et al. [44] propose a new classification method, called Soft Competitive Learning Fuzzy Adaptive Resonance Theory (SFART), to diagnose bearing faults. The recent state of research, however, increasingly comprises industrial application cases which utilize existing methods and algorithms to directly address actual problems and questions in manufacturing from an engineering point of view. These applications are not limited to a specific industrial sector, but can be equally found in different industries [45], e.g. electronics [46–55], metal [56–61], and process industries [62–65]. Likewise, the chosen ML methods are not restricted to a certain type of algorithms but include amongst others Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Decision Trees (DTs).

ANNs “are massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections” [66]. They have found extensive acceptance for modeling complex real-world problems in various disciplines [67]. They have been applied for a wide range of quality-related applications across different industries. In the electronics industry, Yang et al. [47] applied a ANN-based prediction model to solve a solder-paste stencil-printing quality problem. Liukkonen et al. [57] applied ANNs to identify the most important factors affecting the number of detected defects in a soldering process. Shi et al. [52] used ANNs to improve the quality in different industrial processes. Further, ANNs have been applied to predict quality related features based on process parameters for various applications in metal and process industries, e.g. the prediction of the molten steel temperature in a ladle furnace [61] or the estimation of ester formation during beer fermentation [64].

SVMs represent a more sophisticated extension of linear classification, which allows the implementation of nonlinear class boundaries through linear models by transforming the instance space [68]. Due to their advantages of high generalization capabilities, fast classification and the ability to process high dimensional data sets, they can be found in a variety of quality-related industrial applications. Kim et al. [46] and Kang et al. [49] applied SVMs in semiconductor manufacturing for faulty wafer detection and the development of a virtual metrology system respectively. Gola et al. [58] used SVMs in metal industry to distinguish between different microstructures of two-phase steels. Lieber [42] applied SVMs to establish a predictive quality control mechanism in steel bar manufacturing.

DTs are hierarchical models that can be used for regression as well as classification tasks. A hierarchical set of rules is generated that divides the feature space into sections parallel to the axis by sequential checking of the feature values [69]. DTs are highly interpretable and provide a high accuracy which makes them a good choice for applications that require insights into the procedure [70] such as the detection of novel defects [46] and the identification of key quality drivers [51].

In addition to the application of existing methods, there are also some contributions to new and further development of algorithms that address recent quality issues in industry. Wang and Liu [71] develop a new soft sensor modeling approach based on a deep learning network and apply the proposed model for estimating the rotor deformation of air preheaters in a thermal power plant boiler. Wang et al. [72] propose a smart surface inspection system, using faster regions with convolutional neural networks (R-CNN) in a cloud-edge computing environment to identify potential defects in part images. To address quality problem solving in the automobile industry, Xu et al. [73] developed a novel knowledge-driven intelligent quality problem-solving system (IQPSS). In electronics industry, Wentin et al. [74] propose an integrated framework of solder joint defects in the context of Automatic Optical Inspection (AOI) of Printed Circuit Boards (PCBs).

While the models usually show promising results, a specific integration into the manufacturing process, especially with focus on the design of ML-based quality inspection processes, has not been sufficiently addressed so far. In some application cases the models are only used for retrospective failure analysis, e.g. [57], so that integration is generally not planned. Some authors, e.g. [46,47], mention integration as a component of future research work in the outlook of their contribution. However, the description of explicit implementation and integration strategies can only be found sporadically, e.g. [49,75–77]. However, a holistic view, i.e. model results achieved under real-life conditions of use, the analysis of specific effects on quality assurance and the determination of economic savings, does not take place. Additionally, resource constraints in terms of real-time execution, hardware restrictions or energy consumption have not been considered at all or only to a limited extent with specific focus on the individual application, e.g. in [42,49,75,78].

2.4. Contribution

Within the context of model-based quality inspection, most of the existing research works only focus on training the ML-based predictive models on historic data. However, the targeted selection of data with respect to potential quality-relevant cause and effect relationships and the integration of the developed models into the quality planning and assurance do not take place.

On the other hand, technical implementation concepts and IT architectures demonstrate the general capability of today's manufacturing plant IT infrastructures for such novel solutions. In this paper, a holistic approach is proposed for model-based quality inspection of semi-finished and finished products in industrial manufacturing. The approach covers all steps for the realization of model-based quality inspection, ranging from the systematic data selection to cover existing quality control loops in the model representation, through the training of predictive models, as well as to the technical implementation and integration. Thereby, in contrast to existing approaches, the underlying business case and existing expert knowledge are particularly emphasized and incorporated.

3. Predictive model-based quality inspection framework

To facilitate the collection, processing, and analyzing of recorded process data, training and deployment of predictive models, as well as their technical implementation and integration, the proposed framework consists of four main elements:

(1) data collection and processing,
(2) model training and scoring,
(3) model deployment, and
(4) technical implementation.

Additionally, the integration of the predictive model-based quality inspection requires the fifth step of the technical integration into the existing IT-infrastructure, which, however, is too individualized to be described in a generally valid and applicable methodology and therefore not part of this contribution. The layout of the proposed framework is shown in Fig. 1.

3.1. Data collection and processing

The first step is the identification and selection of relevant process and quality data, called horizontal data selection. Based on expert knowledge and results of conducted manufacturing experiments, data points that correlate with the main quality-relevant influencing factors and features of the product are selected. The second step is the vertical data selection in which a representative sample of historic data sets has to be chosen and extracted from data holding systems. Thereby, non-representative data sets, which, for example, were recorded under obsolete process configurations or during manufacturing trials, have to be eliminated in order to allow the model to only learn regularly occurring patterns and dependencies. Next, the data quality is evaluated and respective measures of necessary data pre-processing, e.g. treating missing values, redundancies, or inconsistencies or eliminating special cause outliers, are taken. The result is a prepared and cleaned training data set for subsequent modeling, including a unique identifier, all relevant features as well as the quality label, which can be continuous or discrete depending on the applied measurement method.

3.2. Model training and scoring

The model training and scoring process is designed to find the best performing model for a given set of data. The process can be subdivided into training, testing, and model comparison and selection.

The training of the models takes place in a nested structure of inner and outer cross validation and hyperparameter optimization. At the beginning of the model building process, different supervised learning algorithms have to be trained and parameterized in a coarse parameter optimization to allow a comparison of their performances in order to select the best performing model. As the a-priori selection of adequate algorithms is not achievable in a generalized way [79], different learning methods and algorithms have to be tested and evaluated for each individual application [70]. The pre-selection must be made on the basis of selected criteria, e.g. complexity, interpretability, and speed, or the expertise of the respective data scientist and the insights and results from previous projects. For the considered use case of model-based quality inspection, the prediction time as well as the potential precision, which is associated with model complexity, are of greater interest. However, algorithm performance is also affected by factors such as data volume, variety, and velocity [70]. For prediction tasks, which will be within the focus of this contribution, it is thereby recommended to take the mathematical character of the algorithms, e.g. tree-, probability-, or distance-based and ensemble methods, into account and choose a distributed selection of algorithms in order to explore their performance for the given data set [69]. Popular supervised ML algorithms include k-Nearest Neighbors (kNN), Naïve Bayes classifiers (NB), Decision Trees (DT), Logistic Regression (LR), Support Vector Machines (SVM), Random Forests (RF), and Artificial Neural Networks (ANN) [70].

For the evaluation and comparison of model performances, different statistical performance metrics can be applied. For binary classifications, the metrics can be calculated based on the entries of a confusion matrix, as shown in Table 1.

The comparison of the predicted class with the true class allows to distinguish between correctly positive or negative classified examples (true positive, true negative) and incorrectly classified examples (false

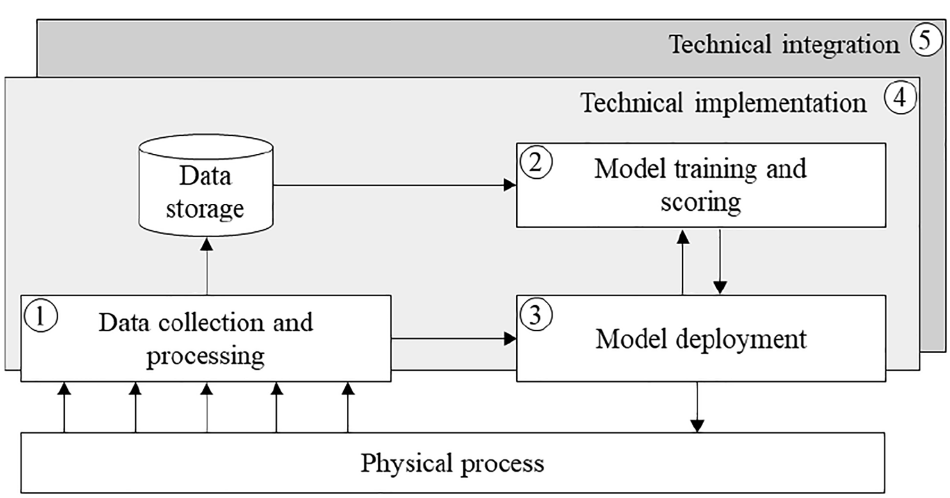

Fig. 1. Layout of the proposed predictive model-based quality inspection framework.

Table 1
Confusion matrix for binary classification according to [42,80,81].

		Reference data		Row sum
		True Class 1 *(Positive)*	True Class 2 *(Negative)*	Σ
Classification	Pred. Class 1 *(Positive)*	a *(True Positive)*	b *(False Positive)*	a + b
	Pred. Class 2 *(Negative)*	c *(False Negative)*	d *(True Negative)*	c + d
Column sum	Σ	a + c	b + d	

positive, false negative). This distinction in turn enables the calculation of various statistical quality measures. The standard performance measure for classification models is accuracy $\left(acc = \frac{a+d}{a+b+c+d}\right)$, which is the percentage of correctly classified examples [82]. However, this measure is inappropriate in applications with unbalanced representations of the classes [82] which is common for quality-related industrial applications, as the assignment of all examples to the overrepresented class would lead to high accuracy but no added value for the differentiation between the classes. Therefore, a more appropriate model evaluation is based on the measures true-positive-rate $\left(TPR = \frac{a}{a+c}\right)$ and false-positive-rate $\left(FPR = \frac{b}{b+d}\right)$. The graphical representation of the trade-off between TPR and FPR, accounting for the trade-off between slack and pseudo defects in quality engineering terminology, is captured by receiver operating characteristics (ROC) curves.

In the configuration of inspection processes, the proportion of false negatives is determined by the inspection severity where an increase in severity implies a lower proportion of false negatives with a higher proportion of good parts falsely marked as defective (false positives) as a trade-off. In ROC-curves the model whose curve is positioned closest to the upper left corner of the plot is considered to bear the best trade-off between the considered quality measures.

As an additional requirement from the technical point of view, the scoring time of the model may not exceed the required response time and should allow time for appropriate control measures. However, the scoring time does not only depend on the algorithm used but also on the hard- and software it is deployed on. In most cases, however, the response time allowed is sufficiently large such that the scoring time constraint is mostly not delimiting the model selection process.

3.3. Model deployment

The model deployment takes place through the organizational integration into the inspection planning process. Thereby, the inspection strategy defines the role of the ML model in the context of the inspection planning and design. While an inspection, exclusively based on the prediction model, requires high confidence in the model and extremely high model accuracy to reach or exceed the level of conventional inspection principles, hybrid approaches seem promising for the current state of development. Thereby, the inspection reliability is given by the combination of quality prediction and conventional inspection. The introduction of quality prediction in quality assurance can facilitate the generation of additional added value by reducing physical inspection volume without sacrificing inspection reliability. To leverage this potential, the prediction can be integrated upstream of the conventional inspection process in order to dynamically adjust the inspection volume according to the prediction result. Two different strategies can be deduced depending on the trustworthiness of the model. Either only those parts are subjected to a physical inspection whose prediction result was OK (not defective), or vice versa, whose result was NOK (defective). In conjunction with the selected strategy, the algorithm must be tuned with respect to zero false positives or zero false negatives accordingly. The savings in inspection volume result from the share of the alternative prediction class. As the class imbalance of data sets in the quality context is usually quite high, selecting only NOK predicted parts to undergo the physical inspection offers vastly superior potential savings.

3.4. Technical implementation

The design of the technical implementation is hardly driven by the actual given requirements and resource constraints and therefore cannot be specified in a generally valid way. However, a generalized framework can serve as an orientation for the individual configuration.

The amount of data available is rapidly increasing. The access of data independent of time and location is enabled by networked devices and sensors. However, this development leads to two major challenges: High dimensional data and large data volumes on the one hand versus highly distributed data, accessible on devices with limited processing capability on the other hand. Generally, hardware requirements depend on the demand for speed and parallelization, accuracy and reliability of the hardware. Further, the 3 V's of Big Data, velocity, variety and volume [83], impact the selection of storage. Accordingly, it is important to investigate the given requirements and resource constraints before the implementation and integration is started. The main challenges when deploying ML models in the manufacturing environment are:

- Limited processing capabilities,
- High dimensionality and large data volumes,
- Energy and memory constraints, and
- Real-time constraints for execution times.

The specification and effects of these constraints cannot be estimated generically. Instead, considering given requirements and arising interactions, a project should strive for an optimal interaction between the individual components of the deployment.

The real-time constraints are given by the takt time of the manufacturing line, which requires the data preprocessing and model application to perform correspondingly fast on the hardware used. The dimensionality of the data set is significantly determined by the number of parameters and processes taken into account. The evaluation of data volume must be performed separately for model training and application. While the data volume for training and optimization depends on the amount of historical data sets available, the volume in model application is defined by the number of classifications executed simultaneously. The process capabilities as well as energy and memory constraints are given by the respective hardware.

4. Case study

The case study was conducted in the electronics industry on the example of SMT manufacturing at Siemens electronics plant in Amberg, Germany. The electronics industry is characterized by almost fully automated manufacturing of large quantities, decreasing size of parts and components as well as high quality requirements. Since the foundation in 1989, programmable logic controllers (PLC) of the Simatic type have been produced in the Siemens electronics manufacturing plant in Amberg. These PLCs serve to automate machines and plants in order to save time and money while increasing the product quality. Due to the high manufacturing volume of the plant with an output of one device per second [84], the speed and robustness of the applied manufacturing processes and a continuous and reliable quality assurance are thoroughly important.

The application of the developed predictive model-based quality inspection methodology for printed circuit boards (PCBs) was motivated by the high capacity utilization of the existing X-ray inspection system and the pending investment decision regarding the purchase of an additional X-ray system due to increasing product demand.

At the beginning of the considered process chain of PCB manufacturing, the initial component, an unassembled PCB, is transported via conveyor belts to a printer, where lead-free solder paste is applied by means of stencil printing. Directly after the printing, the solder paste inspection (SPI), a visual inspection station, checks the quality of the solder paste position. Subsequently, the PCBs are fed into the assembler, where individual components such as resistors, capacitors or microchips are mounted by several assembling heads. The conveyor belt then leads into the furnace, where the applied solder paste is cured in several heat zones. Depending on the product variant, an automatic optical inspection (AOI) or X-ray checks the correct position of soldered components and connection of solder pins after completion of the soldering process. While the AOI can take place in takt time within the manufacturing line, some variants with soldered joints underneath rule out the AOI and therefore require X-ray inspection, which, due to the extensive inspection time, has to be conducted in a separate batch process.

For the case study the focus has been narrowed down to one product variant, its respective manufacturing line and the data sources of SPI and X-ray. The selected product is a connector PCB of a distributed I/O which runs through the SMT line as a panel consisting of 48 boards that are separated afterwards. The selection of SPI as the initial data source of input parameters is justified by the large impact of the solder paste position on the overall quality of the PCBs, which can be confirmed by process experts and found in domain-specific literature [85–88].

The considered data set consists of numeric SPI features (see Table 2) and a binary X-ray label on the aggregation level of solder pins. Historic data sets from SPI and X-ray were matched from different manufacturing data bases via a unique identifier.

Historical data sets for a period of five production months were used for the case study. In total, 1,461,037,321 data points were parsed, of which ∼ 0.0008% were not OK, inducing an extraordinary high class imbalance. Barcodes with too many not OK pins were removed during data exploration and cleansing to eliminate unrepresentative data sets, e.g. from manufacturing trials and not-representative manufacturing conditions. Due to the stability and reliability of the measurement technology used and the high degree of data quality maturity, no further preprocessing steps were needed.

4.1. Modeling results

At the beginning of the model building process, the supervised learning algorithms DT, NB, LR, SVM, and GBT were trained and parameterized in a coarse parameter optimization on a smaller balanced data sample with 4,000 data points and validated with a 5-fold cross validation. The achieved results were compared in terms of accuracy, standard deviation, recall, precision, training time and scoring time (see Table 3). As the absolute times are not yet relevant, the description of hardware used is omitted at this point.

Comparing the results based on the statistical performance measures, SVM and GBT performed best. Comparing the scoring time of both models, however, the GBT scoring time was eight times faster than the one of the SVM. With future scaling and parallel classification of multiple solder joints at the same time in mind, GBT was selected over SVM.

According to the hybrid inspection strategy that implies relying on the prediction model and only inspect those parts whose prediction has been NOK to reduce the X-Ray inspection volume, further optimization of the GBT model parameters aimed at reducing false negatives. As the proportion of false positives strongly correlates with the amount of savings in inspection volumes, different levels of conservativeness of the models were investigated. The high conservative model highly penalizes false negatives, resulting in zero false negatives, but a high proportion of false positives, limiting the savings in inspection effort. On the other hand, the low conservative model allows a small proportion of false negatives for the benefit of higher X-ray savings.

Table 4 shows the classification results of a highly conservative GBT model, trained and tested with a 5-fold cross validation and optimized hyperparameters.

Table 2
Considered SPI features and units.

SPI feature	Unit
Height	%
Shape 2D	%
Shape 3D	%
Surface	%
Volume	%
Offset X	µm
Offset Y	µm

4.2. Deployment strategy and results

After achieving promising results in the modelling phase that showed a good correlation between SPI parameter values and the X-ray result, selecting and arranging a technically and economically feasible deployment and inspection strategy was the main focus of the subsequent phase of the case study.

The quality prediction takes place on the aggregation level of solder joints, representing the smallest entity of the PCB. However, decisions on dynamic X-ray inspection or alternative routings of the PCBs can only be made on higher aggregation levels. The particular PCB panel variant consists of 48 boards, assembled with one connector on each side - X1 on the top and X2 on the bottom. The X2 connector has 52 solder joints and the X1 connector has 79 solder joints per board, resulting in 2,496 solder joints on the bottom and 3,792 solder joints on the top of each panel. The probability that all solder joints of the panel are uniformly predicted as defect-free is low. Therefore, intermediate aggregation levels for routing and inspection decisions had to be evaluated. In this context, the analysis of the X-ray inspection procedure returned the aggregation level of the field of view (FOV) whereas each FOV is successively inspected in the X-ray system.

For the considered panel variant, the 48 boards are inspected within 8 FOVs (see Fig. 2). The GBT model was re-trained on the FOV-level and validated in a 5-fold cross validation. Table 5 shows the classification results after a hyper parameter optimization.

The X-ray inspection can be divided into 20% handling and 80% inspection time, whereas the inspection of each FOV accounts for an equal proportion of the inspection time. To generate economic benefits from the quality prediction, the prediction results are aggregated on the FOV-level, whereas one FOV is defined as defect-free if all pins are defect-free and defective as soon as one or more pins are predicted as defective. The X-ray inspection is omitted for all defect-free FOVs, decreasing the inspection process time by 10% for each FOV skipped. In case all FOVs of a panel are defect-free, the 20% handling are saved as well, as the panel does not have to go into X-ray at all. Fig. 3 shows the alternative routing options of the PCBs depending on the aggregated prediction result.

To enable the application of the prediction model in real-time manufacturing, the technical realization and integration form the next step to be performed. Thereby, it is important to know the existing boundary conditions and resource constraints and take them into account. Additionally, the selection of an adequate architecture and suitable hard- and software are a key factor for the economic scalability of the solution. However, the anticipated implementation architecture should not only operate independently, but especially in conjunction with existing structures and interfaces.

Table 3
Initial results of prediction models.

Model	Accuracy	Standard Deviation	Recall	Precision	Training Time (1000 rows)	Scoring Time (1000 rows)
Naïve Bayes	83.5%	± 2.7%	94.7%	75.5%	3 ms	9 ms
Decision Tree	88.2%	± 1.5%	91.9%	84.0%	39 ms	6 ms
Logistic Regression	71.9%	± 1.3%	77.0%	66.8%	49 ms	27 ms
Support Vector Machine	92.9%	± 1.3%	96.4%	89.3%	300 ms	360 ms
Gradient Boosted Tree	92.6%	± 1.0%	89.9%	93.1%	2 s	40 ms

Table 4
Classification results of highly conservative GBT model with 5-fold cross validation and optimization of hyper parameters on solder joint level.

		Reference data	
		True Defective	True Defect-free
Classification	Pred. Defective	*36 ± 14*	*91,608 ± 30,271*
	Pred. Defect-free	*246 ± 110*	*7,570,753 ± 81,414*
Class recall		98.8% ± 0.4%	86.4% ± 4.5%

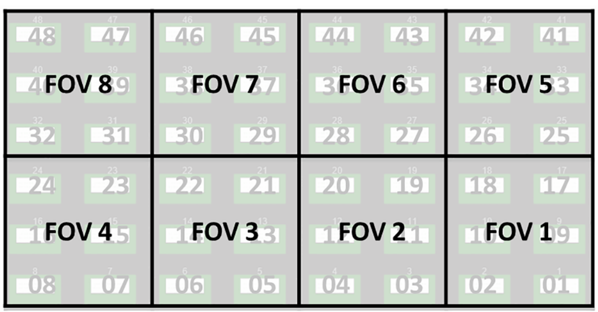

Fig. 2. Field of Views (FOVs) of the selected panel variant.

Table 5
Classification result of highly conservative GBT model with 5-fold cross validation and optimization of hyper parameters on FOV level.

		Reference data		Average volume reduced
		True Defective	True Defect-free	
Classification	Pred. Defective	*41 ± 20*	*13,092 ± 1207*	~29%
	Pred. Defect-free	*4 ± 4*	*5463 ± 1370*	
Class recall		29.4% ± 7.1%	7.6% ± 5.2%	

Within the edge-cloud architecture three main hardware complexes are installed. The edge device is located at the manufacturing line, whereas cloud-based data storage capacities are internally supplied and a company-owned cluster is used for model development, update, and storage. All three components are connected by a network layer.

For the specific use case, model training, fitting, and updating is handled by a company-owned Spark-cluster, handling big data processing with Python libraries. It is horizontally scalable to up to 24 workstations with 16-core-CPUs and 32–64 GB of RAM each.

As the deployment is the only ML-specific task handled by the edge-device, computational requirements for this case are modest. As of 2020, state-of-the-art industrial PCs are equipped with sufficient CPUs for the task, therefore I/O-options and robustness of the solution were prioritized. The edge-device receives, parses, and preprocesses CAMX-xml files of test-wafers via TCP-IP from SPI. Results for current test-sets with seven features are calculated within less than one minute by the chosen solution equipped with an Intel Celeron N2930, leaving headroom for upgraded future models.

The results are transferred to the Quality Measurement Execution (QME) database and data cloud via web-service developed with Amazon Web Services (aws)-SDK. Cloud-Storage is supplied by aws in form of S3 servers for the data lake comprising labeled historic data, extracted from other data-holding systems (like SimaticIT) and the QME data base. Edge device and cloud solutions are connected via ISP, requiring low latency (2–150 ms) and bandwidth of 10 MB per 14 s for the project (approx. 1 Mbit/s permanently per manufacturing line).

4.3. Outlook on future scaling and extension

The deployment and integration is currently limited to one product variant, the respective manufacturing line and the data sources of SPI and X-Ray. The next step will be a multidimensional expansion including the connection of additional data sources, e.g. supplier data, and other process data such as placement data. The expected outcome is an improved model performance in terms of better trade-off between slack and inspection time savings and better coverage of multiple X-Ray defect patterns. The impact on the IT architecture will be increased

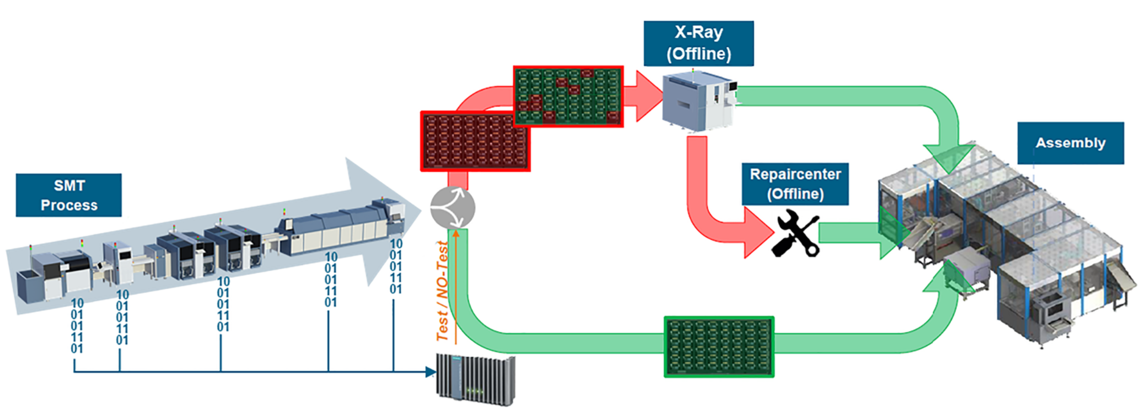

Fig. 3. Alternative routing options of PCBs based on aggregated prediction result.

network structure and complexity requirements, meriting a resource-aware realization and connection.

Another dimension is the expansion to other product variants and types. As each product variant differs in its specific characteristics and defect patterns, new models have to be trained and optimized. This in turn requires further investigation on how to manage numerous models and developing a model management and selection strategy with a certain degree of automation. Additionally, the connection of data-holding systems and the cloud-based data lake will be automated to generate labeled training data for model training and optimization. Another interface between edge and cloud will also be required to exchange classification results in order to observe the online model performance by incorporating recorded X-Ray results.

5. Discussion and conclusion

The ability to inspect product quality comprehensively and reliably is a key success factor for manufacturing companies in today's global competition. A ML-based inspection approach provides economically beneficial inspection strategies. The concept is demonstrated and realized for a case study in electronics industry with a defined scope within the PCB assembly at Siemens electronics plant in Amberg, Germany.

Predictive models are trained based on historic data sets in the cloud and deployed on local edge devices. During the manufacturing process, parameters are recorded and sent to the edge device, on which data processing and model application are handled in near real time. The prediction results are evaluated and aggregated to a processable level allowing a dynamic inspection decision. Depending on the results on the aggregated level, inspection time and/or additional handling time can be saved, generating an attractive business case.

However, more research is required to scale the application of predictive model-based inspection and cope with emerging challenges and research questions. Future research and implementation fields identified during this work are:

- **Addition of further process parameters:** This increases the number of data sources which in turn leads to an expansion and increase in complexity of the solution. However, it may also increase the model performance and allow to increase inspection coverage, e.g. including the distinction of different defect patterns.
- **Roll-out to more than one product variant:** To allow a comprehensive predictive model-based inspection, the coverage of different variants by a single model or the need of additional prediction models must be examined.
- **Automation of data interfaces:** The speed and economic benefit of model building and optimization can be increased by automating the data transfer between data-holding systems and cloud-based applications. In addition, further interfaces will be created to monitor the performance of models deployed on the edge by coupling their results with recorded X-Ray measurements.
- **Investigation of effects on systemic level:** Early knowledge of the expected product quality through predictive model-based inline inspection enables timely and novel control decisions. Bottlenecks may resolve or shift, requiring a detailed consideration of the overall system, including the manufacturing, logistic, and inspection processes as well as the network components of the edge cloud architecture.

To conclude this paper, it can be summarized that, enabled by the ever-rising data availability in manufacturing and the technological advances in computing and respective hard- and software solutions, the predictive model-based quality inspection is a highly promising approach to design inspection processes more economically.

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.

Acknowledgment

Part of the work on this paper has been supported by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center SFB 876 “Providing Information by Resource-Constrained Analysis”, project B3 “Data Mining on Sensor Data of Automated Processes”.

References

[1] S. Singaravel, J. Suykens, P. Geyer, Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction, Adv. Eng. Inform. 38 (2018) 81–90, https://doi.org/10.1016/j.aei.2018.06.004.

[2] Fei Tao, Qinglin Qi, Lihui Wang, A.Y.C. Nee, Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison, Engineering 5 (4) (2019) 653–661 https://linkinghub.elsevier.com/retrieve/pii/S209580991830612Xhttps://doi.org/10.1016/j.eng.2019.01.014.

[3] R. Baheti, H. Gill, Cyber-physical systems, Impact Control Technol. 12 (2011) 161–166.

[4] A.J.C. Trappey, C.V. Trappey, U. Hareesh Govindarajan, A.C. Chuang, J.J. Sun, A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0, Adv. Eng. Inform. 33 (2017) 208–229, https://doi.org/10.1016/j.aei.2016.11.007.

[5] S. Harnal, R.K. Chauhan, Multimedia support from cloud computing: A review, in: International Conference on Microelectronics, Computing and Communication - MicroCom 2016: January 23-25, 2016, Durgapur, India, IEEE, [Piscataway, NJ], 2016, pp. 1–6.

[6] B.-H. Li, B.-C. Hou, W.-T. Yu, X.-B. Lu, C.-W. Yang, Applications of artificial intelligence in intelligent manufacturing: a review, Front. Inf. Technol. Electron. Eng. 18 (2017) 86–96, https://doi.org/10.1631/FITEE.1601885.

[7] Y. Ishino, Y. Jin, An information value based approach to design procedure capture, Adv. Eng. Inf. 20 (2006) 89–107, https://doi.org/10.1016/j.aei.2005.04.002.

[8] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: data mining, inference, and prediction, Springer Series in Statistics, Springer, New York, 2009.

[9] U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy, Advances in knowledge discovery and data mining, 1996.

[10] U.H. Govindarajan, A.J.C. Trappey, C.V. Trappey, Immersive Technology for Human-Centric Cyberphysical Systems in Complex Manufacturing Processes: A Comprehensive Overview of the Global Patent Profile Using Collective Intelligence, Complexity 2018 (2018) 1–17, https://doi.org/10.1155/2018/4283634.

[11] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inform. Process. Syst. (2012) 1097–1105.

[12] V. Mnih, N. Heess, A. Graves, et al., Recurrent models of visual attention, Adv. Neural Inform. Process. Syst. (2014) 2204–2212.

[13] C. Chelba, T. Mikolov, M. Schuster, Q. Ge, T. Brants, P. Koehn, T. Robinson, One billion word benchmark for measuring progress in statistical language modeling, arXiv preprint arXiv:1312.3005, 2013.

[14] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, Imagenet large scale visual recognition challenge, Int. J. Comput. Vision 115 (2015) 211–252.

[15] T. Chilimbi, Y. Suzue, J. Apacible, K. Kalyanaraman, Project adam: Building an efficient and scalable deep learning training system, in: 11th 5USENIX6 Symposium on Operating Systems Design and Implementation (5OSDI6 14), 2014, pp. 571–582.

[16] J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, M. Mao, A. Senior, P. Tucker, K. Yang, Q.V. Le, et al., Large scale distributed deep networks, Adv. Neural Inform. Process. Syst. (2012) 1223–1231.

[17] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: A system for large-scale machine learning, in: 12th 5USENIX6 Symposium on Operating Systems Design and Implementation (5OSDI6 16), 2016, pp. 265–283.

[18] J. Drexl, R. Hilty, F. Beneke, L. Desaunettes, M. Finck, J. Globocnik, B. Gonzalez Otero, J. Hoffmann, L. Hollander, D. Kim, H. Richter, S. Scheuerer, P.R. Slowinski, J. Thonemann, Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Perspective, SSRN J. (2019), https://doi.org/10.2139/ssrn.3465577.

[19] L. Deng, X. Li, Machine learning paradigms for speech recognition: An overview, IEEE Trans. Audio Speech Lang. Process. 21 (2013) 1060–1089.

[20] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, L. Hanzo, Machine learning paradigms for next-generation wireless networks, IEEE Wirel. Commun. 24 (2017) 98–105.

[21] M. Mohammed, E.B.M. Bashier, M.B. Khan, Machine learning: Algorithms and applications, CRC Press, Boca Raton, 2017.

[22] B. Berendt, B. Bringmann, E. Fromont, G. Garriga, P. Miettinen, N. Tatti, V. Tresp, Machine Learning and Knowledge Discovery in Databases, in: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2016.

[23] H. Chang, A. Hari, S. Mukherjee, T.V. Lakshman, Bringing the cloud to the edge, 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2014, pp. 346–351.

[24] R.R. Schaller, Moore's law: past, present and future, IEEE Spectr. 34 (1997) 52–59, https://doi.org/10.1109/6.591665.

[25] M.G. Avram, Advantages and Challenges of Adopting Cloud Computing from an Enterprise Perspective, Procedia Technol. 12 (2014) 529–534, https://doi.org/10.1016/j.protcy.2013.12.525.

[26] W. Dubey A, Delivering software as a service, The McKinsey Quarterly (May 2007) 1–12, 2007.

[27] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, Cloud computing—The business perspective, Decis. Support Syst. 51 (2011) 176–189.

[28] R.L. Grossman, The case for cloud computing, IT Prof. 11 (2009) 23–27.

[29] W. Shi, S. Dustdar, The promise of edge computing, Computer 49 (2016) 78–81.

[30] K. Kambatla, G. Kollias, V. Kumar, A. Grama, Trends in big data analytics, J. Parallel Distrib. Comput. 74 (2014) 2561–2573, https://doi.org/10.1016/j.jpdc.2014.01.003.

[31] M. Saecker, V. Markl, Big data analytics on modern hardware architectures: A technology survey, Eur. Bus. Intell. Summer School (2012) 125–149.

[32] J. Cong, Z. Fang, M. Huang, L. Wang, Di Wu, CPU-FPGA Coscheduling for Big Data Applications, IEEE Des. Test 35 (2018) 16–22, https://doi.org/10.1109/MDAT.2017.2741459.

[33] D. Georgakopoulos, P.P. Jayaraman, M. Fazia, M. Villari, R. Ranjan, Internet of Things and Edge Cloud Computing Roadmap for Manufacturing, IEEE Cloud Comput. 3 (2016) 66–73, https://doi.org/10.1109/MCC.2016.91.

[34] D. Weimer, B. Scholz-Reiter, M. Shpitalni, Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection, CIRP Ann. 65 (2016) 417–420, https://doi.org/10.1016/j.cirp.2016.04.072.

[35] V.D. Majstorovic, N. Durakbasa, Y. Takaya, S. Stojadinovic, Advanced Manufacturing Metrology in Context of Industry 4.0 Model, in: Proceedings of the 12th International Conference on Measurement and Quality Control-Cyber Physical Issue: IMEKO TC 14 2019,[Belgrade, Serbia, from 4 to 7 June 2019], 2019, pp. 1–11.

[36] D. Braha, Data mining for design and manufacturing: Methods and applications, Kluwer Academic, Dordrecht, Boston, 2001.

[37] J.A. Harding, M. Shahbaz, Srinivas, A. Kusiak, Data Mining in Manufacturing: A Review, J. Manuf. Sci. Eng. 128 (2006) 969, https://doi.org/10.1115/1.2194554.

[38] A. Kusiak, Data mining: manufacturing and service applications, Int. J. of Prodn. Res. 44 (2006) 4175–4191, https://doi.org/10.1080/00207540600632216.

[39] G. Köksal, İ. Batmaz, M.C. Testik, A review of data mining applications for quality improvement in manufacturing industry, Expert Syst. Appl. 38 (2011) 13448–13467.

[40] H. Rostami, J.-Y. Dantan, L. Homri, Review of data mining applications for quality assessment in manufacturing industry: support vector machines, Int. J. Metrol. Qual. Eng. 6 (2015) 401, https://doi.org/10.1051/ijmqe/2015023.

[41] M. Kano, Y. Nakagawa, Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry, Comput. Chem. Eng. 32 (2008) 12–24.

[42] D. Lieber, Data mining in quality control at the example of steel bar manufacturing (in German), first ed., Shaker, Herzogenrath, 2018.

[43] C. Gröger, F. Niedermann, B. Mitschang, Data mining-driven manufacturing process optimization, Proceedings of the world congress on engineering, 2012, pp. 1475–1481.

[44] X.-J. Wan, L. Liu, Z. Xu, Z. Xu, Q. Li, F. Xu, Fault diagnosis of rolling bearing based on optimized soft competitive learning Fuzzy ART and similarity evaluation technique, Adv. Eng. Inf. 38 (2018) 91–100, https://doi.org/10.1016/j.aei.2018.06.006.

[45] F. Arif, N. Suryana, B. Hussin, A data mining approach for developing quality prediction model in multi-stage manufacturing, Int. J. Comput. Appl. 69 (2013) 35–40, https://doi.org/10.5120/12106-8375.

[46] D. Kim, P. Kang, S. Cho, H.-J. Lee, S. Doh, Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing, Expert Syst. Appl. 39 (2012) 4075–4083.

[47] T. Yang, T.-N. Tsai, J. Yeh, A neural network-based prediction model for fine pitch stencil-printing quality in surface mount assembly, Eng. Appl. Artif. Intell. 18 (2005) 335–341, https://doi.org/10.1016/j.engappai.2004.09.004.

[48] S. Kang, P. Kang, An intelligent virtual metrology system with adaptive update for semiconductor manufacturing, J. Process Control 52 (2017) 66–74, https://doi.org/10.1016/j.jprocont.2017.02.002.

[49] P. Kang, H.-J. Lee, S. Cho, D. Kim, J. Park, C.-K. Park, S. Doh, A virtual metrology system for semiconductor manufacturing, Expert Syst. Appl. 36 (2009) 12554–12561, https://doi.org/10.1016/j.eswa.2009.05.053.

[50] Y.-J. Chang, Y. Kang, C.-L. Hsu, C.-T. Chang, T.Y. Chan, Virtual Metrology Technique for Semiconductor Manufacturing, in: IEEE International Joint Conference on Neural Networks, Vancouver, BC, Canada, IEEE, Piscataway, New Jersey, 2006, pp. 5289–5293.

[51] J. Schnell, C. Nentwich, F. Endres, A. Kollenda, F. Distel, T. Knoche, G. Reinhart, Data mining in lithium-ion battery cell production, J. Power Sources 413 (2019) 360–366, https://doi.org/10.1016/j.jpowsour.2018.12.062.

[52] X. Shi, P. Schillings, D. Boyd, Applying artificial neural networks and virtual experimental design to quality improvement of two industrial processes, Int. J. Prod. Res. 42 (2004) 101–118.

[53] K.R. Skinner, D.C. Montgomery, G.C. Runger, J.W. Fowler, D.R. McCarville, T.R. Rhoads, J.D. Stanley, Multivariate statistical methods for modeling and analysis of wafer probe test data, IEEE Trans. Semicond. Manuf. 15 (2002) 523–530.

[54] S. Thiede, A. Turetskyy, A. Kwade, S. Kara, C. Herrmann, Data mining in battery production chains towards multi-criterial quality prediction, CIRP Ann. 68 (2019) 463–466, https://doi.org/10.1016/j.cirp.2019.04.066.

[55] C.-F. Chien, K.-H. Chang, W.-C. Wang, An empirical study of design-of-experiment data mining for yield-loss diagnosis for semiconductor manufacturing, J. Intell. Manuf. 25 (2014) 961–972, https://doi.org/10.1007/s10845-013-0791-5.

[56] N. de Abajo, A.B. Diez, V. Lobato, S.R. Cuesta, ANN quality diagnostic models for packaging manufacturing: An industrial data mining case study, Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 799–804.

[57] M. Liukkonen, T. Hiltunen, E. Havia, H. Leinonen, Y. Hiltunen, Modeling of soldering quality by using artificial neural networks, IEEE Trans. Electron. Packag. Manuf. 32 (2009) 89–96.

[58] J. Gola, D. Britz, T. Staudt, M. Winter, A.S. Schneider, M. Ludovici, F. Mücklich, Advanced microstructure classification by data mining methods, Comput. Mater. Sci. 148 (2018) 324–335, https://doi.org/10.1016/j.commatsci.2018.03.004.

[59] X. Wang, C.X. Feng, Development of Empirical Models for Surface Roughness Prediction in Finish Turning, Int. J. Adv. Manuf. Technol. 20 (2002) 348–356, https://doi.org/10.1007/s001700200162.

[60] H.-Y. Tseng, Welding parameters optimization for economic design using neural approximation and genetic algorithm, Int. J. Adv. Manuf. Technol. 27 (2006) 897–901, https://doi.org/10.1007/s00170-004-2276-3.

[61] X. Wang, M. You, Z. Mao, P. Yuan, Tree-Structure Ensemble General Regression Neural Networks applied to predict the molten steel temperature in Ladle Furnace, Adv. Eng. Inf. 30 (2016) 368–375, https://doi.org/10.1016/j.aei.2016.05.001.

[62] W.-C. Chen, A.H.I. Lee, W.-J. Deng, K.-Y. Liu, The implementation of neural network for semiconductor PECVD process, Expert Syst. Appl. 32 (2007) 1148–1153.

[63] W.-C. Chen, P.-H. Tai, M.-W. Wang, W.-J. Deng, C.-T. Chen, A neural network-based approach for dynamic quality prediction in a plastic injection molding process, Expert Syst. Appl. 35 (2008) 843–849.

[64] C. Riverol, J. Cooney, Estimation of the ester formation during beer fermentation using neural networks, J. Food Eng. 82 (2007) 585–588.

[65] D. Shah, J. Wang, Q.P. He, A feature-based soft sensor for spectroscopic data analysis, J. Process Control 78 (2019) 98–107, https://doi.org/10.1016/j.jprocont.2019.03.016.

[66] A.K. Jain, J. Mao, K.M. Mohiuddin, Artificial neural networks: a tutorial, Computer 29 (1996) 31–44, https://doi.org/10.1109/2.485891.

[67] I.A. Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, J. Microbiol. Methods 43 (2000) 3–31, https://doi.org/10.1016/S0167-7012(00)00201-3.

[68] I.H. Witten, E. Frank, M.A. Hall, C.J. Pal, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.

[69] C.C. Aggarwal, Data mining: the textbook, Springer, 2015.

[70] In Lee, Yong Jae Shin, Machine learning for enterprises: Applications, algorithm selection, and challenges, Bus. Horiz. 63 (2) (2020) 157–170 https://linkinghub.elsevier.com/retrieve/pii/S0007681319301521https://doi.org/10.1016/j.bushor.2019.10.005.

[71] X. Wang, H. Liu, Soft sensor based on stacked auto-encoder deep neural network for air preheater rotor deformation prediction, Adv. Eng. Inf. 36 (2018) 112–119, https://doi.org/10.1016/j.aei.2018.03.003.

[72] Y. Wang, M. Liu, P. Zheng, H. Yang, J. Zou, A smart surface inspection system using faster R-CNN in cloud-edge computing environment, Adv. Eng. Inf. 43 (2020) 101037, , https://doi.org/10.1016/j.aei.2020.101037.

[73] Z. Xu, Y. Dang, P. Munro, Knowledge-driven intelligent quality problem-solving system in the automotive industry, Adv. Eng. Inf. 38 (2018) 441–457, https://doi.org/10.1016/j.aei.2018.08.013.

[74] D. Wenting, M. Abdul, E. Marius, S. Alexei, Soldering defect detection in automatic optical inspection, Adv. Eng. Inf. 43 (2020) 101004, https://doi.org/10.1016/j.aei.2019.101004.

[75] H.-C. Huang, Y.-C. Lin, M.-H. Hung, C.-C. Tu, F.-T. Cheng, Development of cloud-based automatic virtual metrology system for semiconductor industry, Rob. Comput. Integr. Manuf. 34 (2015) 30–43, https://doi.org/10.1016/j.rcim.2015.01.005.

[76] F.-T. Cheng, J.Y.-C. Chang, H.-C. Huang, C.-A. Kao, Y.-L. Chen, J.-L. Peng, Benefit model of virtual metrology and integrating AVM into MES, IEEE Trans. Semicond. Manuf. 24 (2011) 261–272.

[77] J.C. Yung-Cheng, F.-T. Cheng, Application development of virtual metrology in semiconductor industry, in: Proceedings of the 32nd annual conference of IEEE industrial electronics society (IECON 2005). Los Alamitos, CA, USA, 2005.

[78] H.W. Dörmann Osuna, Approach for a process-integrated quality control system for unstable processes (in German), 2009.

[79] L. Kotthoff, Algorithm Selection for Combinatorial Search Problems: A Survey, in: C. Bessiere, L. de Raedt, L. Kotthoff, S. Nijssen, B. O'Sullivan, D. Pedreschi (Eds.), Data Mining and Constraint Programming: Foundations of a Cross-Disciplinary Approach, Springer International Publishing, Cham, 2016, pp. 149–190.

[80] Kohl, Performance Measures in Binary Classification, Int. J. Stats. Med. Res., 2012. https://doi.org/10.6000/1929-6029.2012.01.01.08.

[81] W.M.P. van der Aalst, Process Mining, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[82] M. Kubat, R.C. Holte, S. Matwin, Machine learning 30 (1998) 195–215, https://doi.org/10.1023/A:1007452223027.

[83] S. Sagiroglu, D. Sinanc, Big data: A review, in: 2013 International Conference on Collaboration Technologies and Systems (CTS), San Diego, CA, USA, IEEE, 20.05.

2013–24.05.2013, pp. 42–47.
[84] Siemens, Digital Factory. 99.99885 Percent Quality (in German), 2014. https://www.siemens.com/innovation/de/home/pictures-of-the-future/industrie-und-automatisierung/digitale-fabrik-die-fabrik-von-morgen.html.
[85] A. Lofti, M. Howarth, Industrial application of fuzzy systems: adaptive fuzzy control of solder paste stencil printing, Inf. Sci. 107 (1998) 273–285.
[86] D. Amir, Expert system for SMT assembly, Proceedings of the Surface Mount International Conference and Exposition-Technical Program, 1994, pp. 691–699.
[87] D. He, N.N. Ekere, M.A. Currie, The behavior of solder pastes in stencil printing with vibrating squeegee, IEEE Trans. Comp., Packag., Manufact. Technol. C 21 (1998) 317–324, https://doi.org/10.1109/TCPMC.1998.7102530.
[88] J. Pan, G.L. Tonkay, R.H. Storer, R.M. Sallade, D.J. Leandri, Critical Variables of Solder Paste Stencil Printing for Micro-BGA and Fine-Pitch QFP, IEEE Trans. Electron. Packag. Manuf. 27 (2004) 125–132, https://doi.org/10.1109/TEPM.2004.837965.