

54th CIRP Conference on Manufacturing Systems

Predictive analytics in quality assurance for assembly processes: lessons learned from a case study at an industry 4.0 demonstration cell

Peter Burggräf^a, Johannes Wagner^a, Benjamin Heinbach^a, Fabian Steinberg^a, Alejandro R. Pérez M.^{*a}, Lennart Schmallenbach^a, Jochen Garcke^{b,c,d}, Daniela Steffes-lai^b, Moritz Wolter^{b,c,e}

^aChair of International Production Engineering and Management (IPEM), Universität Siegen, Paul-Bonatz-Straße 9-11, 57076 Siegen, Germany

^bFraunhofer Institute for Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven 1, 53757 Sankt Augustin, Germany

^cFraunhofer Center for Machine Learning, Schloss Birlinghoven 1, 53757 Sankt Augustin, Germany

^dInstitute for Numerical Simulation, Universität Bonn, Endenicher Allee 19b, 53115 Bonn, Germany

^eInstitute for Computer Science, Universität Bonn, Endenicher Allee 19a, 53115 Bonn, Germany

* Corresponding author. Tel.: +49-271-740-2630; fax: +49-271-740-12628. E-mail address: alejandro.perez@uni-siegen.de

Abstract

Quality assurance (QA) is an important task in manufacturing to assess whether products meet their specifications. However, QA might be expensive, time-consuming, or incomplete. This paper presents a solution for predictive analytics in QA based on machine sensor values during production while employing specialized machine-learning models for classification in a controlled environment. Furthermore, we present lessons learned while implementing this model, which helps to reduce complexity in further industrial applications. The paper's outcome proves that the developed model was able to predict product quality, as well as to identify the correlation between machine-status and faulty product occurrence.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the 54th CIRP Conference on Manufacturing System.

Keywords: Machine-Learning; Predictive Quality; Production; Quality Assurance; Classification

1. Introduction

For companies aiming at quality leadership, the final quality of any goods produced is vital for market success. Thus, companies need to ensure that every product meets the expected customer and legal requirements. The quality control (QC) of production processes and products is one fundamental part of general quality management. It contains the measures and activities needed to fulfill specific quality requirements [1].

Over the years, diverse methods to ensure the quality of products have been developed [1]. Among the most used methods, we have Auditing, Total Quality Control (TCQ) [2], Poka-Yoke Techniques [3], the 100% control of each produced product [4], and statistical methods like the Statistical Process Control (SPC) [5]. Independently of the method used, the goal is the same, to avoid the production of faulty products or to identify and remove them before delivering those to the customer.[1]. However, the relationship between cost and efficiency varies from one method to another. This is noticeable, e.g., at the cost

of highly specialized equipment for automated inspection, or the cost of employees performing a manual review of the produced products versus the accuracy of detecting all faulty products [6].

In general, the quality characteristics of a product can be classified as attributes or variables [1]. *Attributes* are qualitative measurable characteristics, i.e., color or texture. On the other hand, *variables* are precise and measurable characteristics, which are shown as numbers, i.e., length or width. The advantage of quantitative measurements is that in the event of a defective intermediate or end product, precise measured values are available, which enable controlled adjustment of the process parameters, as well as the possibility of observing trends in the collected data. This allows the employment of predictive models to determine the quality of the product before the QC process takes place [7, 8].

With the evolution of machine learning (ML) applications [7], approaches combining QC and predictive models are becoming more relevant. The applications range from the defect analysis of the produced piece [9, 10], to process-oriented QC [11, 8, 12, 13], see Section 2. Independently of the research

area previously stated, within the scope of the related works, there is no evidence of an evaluation framework for diverse ML techniques applied in QC process.

Regarding the use of ML models in QC, several challenges arise that have not yet been sufficiently addressed, e.g., what to consider while evaluating the output from ML models for industrial applications? (model accuracy vs. applicability at industrial environments); what steps should be followed to prepare the data for predictive QC methods?; and what findings and lessons are worth considering for further applications in the area?. For this, we cover the methods used for quality checks in production, as well as the bases of predictive analytics in production (Section 2). We describe a suitable industrial platform to implement the ML models (Section 3). Then, we explain the three machine learning methods implemented in this work: Multilayer perceptrons, Support Vector Machines, and a Decision Tree (Section 4), followed by the description of the experiment process from the data acquisition and preprocessing, to the classifier optimization and testing (Section 5). Finally, we present the findings achieved after the finalization of this project phase (Section 6), as well as present a list of lessons learned for the implementation of the project, which might be helpful for further implementations (Section 7).

2. Related Work

The digital transformation of the manufacturing industry [14, 15, 16] allows the application of machine learning-powered predictive analytics techniques. According to Delen and Damirkan, predictive analytics (PA) consists of unraveling the inherent relationships (if any) between input and output by using data, and mathematical techniques [17]. PA offers a wide range of applications and is usable as long as sufficient data are available [7]. According to Krauß, there is a wide range of applications for PA in the industrial and production environment. Some industrial applications of PA focus machines and assets (predictive maintenance, anomaly detection, self learning-machines) [18], and process (scheduling, process design, predictive process control) [8].

An explorative literature review on this matter reveals that publications that combine PA and QC became more relevant in the last six years. This could be due to computer power limitations in earlier research and the recent increase of ML applications in this area. The related work mostly belongs to two groups, defect analysis of produced pieces and process-oriented QC. Due to the recent advances in ML approaches specialized in image and pattern recognition [7] and the advent of standardized data collection [14], a great number of publications focus on product analysis based on images from cameras during the production process, e.g., Escobar employed pattern recognition for defect detection in a binary classification problem based on a l_1 -regularized logistic regression [9]; Sohnius focussed on defect prediction within printed layers by employing supervised models based on features extracted directly from the machine code, and the scan of the layers surface [10]. Other authors focussed their research on process-oriented QC, e.g.,

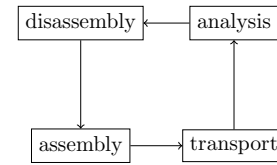


Fig. 1: Idealized flow-chart of our simulated manufacturing process loop. All process elements can be disturbed at any time.

Krauß focussed on describing the ML pipeline to implement automated machine learning for predictive quality in production [11], and product quality prediction in a process chain [8]; Aumi focussed on the model development for predictive quality control of batch processes [12]; Ritter studied the ways of process modeling on quality prediction, and assurance of chipboards [13].

Applications that combine PA with QC might improve the current QC processes. However, based on the previous findings, there are no unambiguous references to the impact of implementing ML models based just on model accuracy, together with industrial factors of the QC process. By employing machine learning (ML) techniques (Section 4), together with the QC traditional methods, it would be possible to improve the relationship between inspection costs and the probability of detecting faulty products during the QC process, as well as decrease the risk of labeling NOK pieces as OK and deliver them to costumers. However, before implementing these techniques, it is important to define the platform where these predictive analytics models will be implemented in a controlled and transparent industrial environment.

3. The case: industry 4.0 demonstration cell

Our case study is an abstraction of an assembly process of the industry. This case study is a highly flexible and transparent platform regarding software and hardware access and manipulation. The industry 4.0 demonstration cell is composed of three independent conveyor belts, a robotic assembly arm (UR3), a laser scanner used for quality control as well as a wide range of sensors, all orchestrated by a SIEMENS PLC S7-1200. We show an idealized process flow-chart in Figure 1 and an illustration of the cell components in Figure 2. The executed abstracted assembly process consists of stacking two disks of different sizes on top of each other. The corresponding pseudo product quality is later determined using the laser scanner and evaluating the concentricity of both disks. If the disks' concentricity is within a tolerance of 1.5mm, the piece is classified as OK, if not, it is classified as NOK. By considering that malfunctions can occur during industrial production processes, our demonstration cell can simulate failures in diverse areas, such as:

- Assembly errors due to the robotic arm: incorrectly positioned disks.
- Bearing damage on the conveyor belts: vibration changes in the conveyor belt motors.

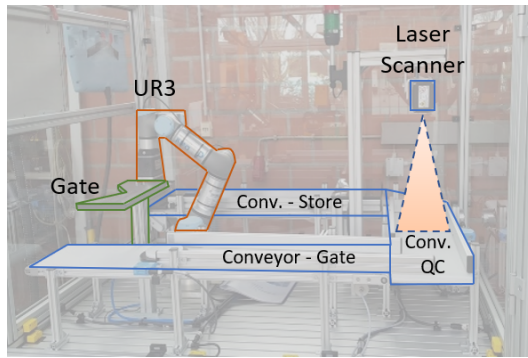


Fig. 2: The industry 4.0 demonstration cell at the University of Siegen

- Resistance on the conveyor belts: temperature changes in the conveyor belt motors.
- Malfunction of the gate door due to leakage in the compressed air system: cylinder stroke at the assembly barrier is slower.
- Faster operating point: disks shift position due to the high belt speeds.
- Missing material at the end of the warehouse: production stop due to lack of material.

Those simulated failures have an impact on the data collected by the sensors. Anomalies can be observed at the vibration sensors, temperature rise, fluctuations in the air pressure system, noticeable shifting on pieces' position as well as the stop of the process. The continuous data collection and the clear correlation between the machine parameters and the quality of the product, is the ideal play-ground to implement a sandbox for machine learning implementations. With this in mind, our goal is to predict the quality of the product (concentricity of the assembled parts), before the actual quality check done by the laser scanner.

4. Machine Learning Methods

To study which implementation would be more suitable for industrial applications, we compare Multilayer Perceptrons (MLPs), Support Vector Machines (SVM), as well as Decision Trees on the task of quality prediction. All three will be introduced next.

4.1. Multilayer perceptrons

These feedforward neural networks typically combine multiple layers and activation functions [7]. A layer contains a large weight matrix and a bias vector. To evaluate the layer the input vector must be multiplied with the weight matrix before the bias vector is added. Each layer is typically followed by an activation function which adds non-linearity to the network graph. Rectified linear units (ReLUs) are zero for negative inputs and leave positive inputs unchanged. ReLUs are the recommended activation function for modern neural networks [19], we use these here as well. Since the linear parts of the ReLU and the

linear matrix multiplication are differentiable we can employ stochastic gradient descent to train our classifier.

4.2. Support Vector Machines

Support Vector Machines (SVM) attempt to separate data along hyperplanes [20]. SVM offer a convex alternative to MLPs [21]. A convex classifier is guaranteed to converge to a global solution regardless of its initialization. Non-separable data can become separable if projected to a different possibly high dimensional space through a Kernel function. The Kernel trick allows SVM to solve some non-linear classification problems [21]. Radial basis functions (RBF) are commonly chosen as kernels [21], therefore, we choose to follow this practice as well.

4.3. Decision Tree

Decision Trees are essentially binary trees, which work with the input features at the roots, decisions are made by moving up the tree to top leaves. At every branch in the tree, a decision is made, until one arrives at the decision [22]. For numerical input features, each decision partitions the data [20], for example by comparison to a threshold value. To construct the tree these cut-off values have to be chosen at every branch. The construction problem turns into an optimization problem when a split criterion is introduced. Common choices are cross-entropy of Gini-coefficient [20]. During the construction splits minimizing, the selected functions are chosen.

5. Experiments

5.1. Data acquisition

Industrial projects require to work with a variety of sensors and controllers from different manufacturers. Thus, the risk of incompatibility between technologies increases. To avoid these problems, universal protocols are often implemented in industrial projects. Figure 3 describes a simplified version of the real connection diagram of the demonstration cell. In our case, the data is acquired independently from two OPC-UA servers. This represents the interaction between multiple servers in real production environments. Before collecting the data, we needed to find the variables of interest and map their source, as well as to determine the use cases of interest. Our study focussed on the data described in Table 1 and the six use cases described in Table 2. The data was manually collected by using the software "UAExpert - v1.5.1". Approximately 15 hours of data were collected in different sessions, while equally sampling each use case during the data acquisition stage.

5.2. Data preprocessing

Once the data was collected, the next step was to clean the raw data and organize it accordingly [7]. In our case, that means to merge the data collected from the two OPC-UA servers, and

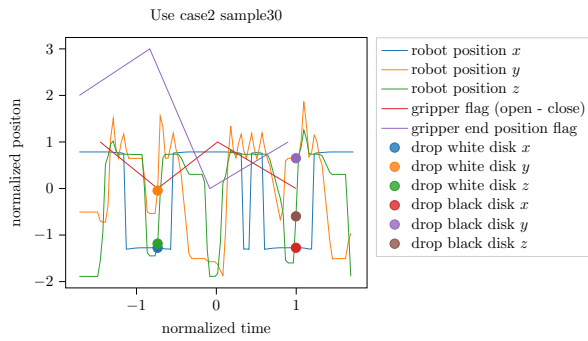


Fig. 4: Three dimensional robot arm movement patterns.

Approach	accuracy
Naïve-Baseline	64 %
Support Vector Machine	82 %
Multilayer Perceptron	88 %
Decision Tree	98 %

Table 3: Comparison of Support Vector Machine (SVM), Multilayer Perceptron and Decision Tree classification on our anomaly detection task.

was used as the training target. It can be OK or NOK which we encoded as zero and one. We worked with a total of 528 measurements. Each containing arm and belt data logs of individual cell run. 50 samples were set aside at random for testing purposes, leaving 478 training samples. The random number generator seed was set to one to ensure the train and test set splits are identical for all experiments. We compared a total of three different classifier architectures on the data, a Support Vector Machine (SVM), a Multilayer Perceptron (MLP) and a Decision Tree structure. In this case, we chose to train all three models using exemplary default hyperparameters, which we found to work well.¹ Results are shown in table 3. For 32 of the total 50 test samples quality measurements indicated a problem. The first row shows the performance of naively predicting a broken piece every time. This accuracy sets the baseline over the entire data set, which would be obtained by simply labeling all samples as faulty. Therefore, we expected any naive classifier to produce at least 64% accuracy, which we had to surpass. In Table 3, we observe that this was indeed the case for all three approaches evaluated here. The MLP prediction results presented on Table 4b show that even though the prediction model was not 100% accurate, the percentage of false-positive predictions was 0%. The false-negative results represented 12% of the test sample and the true predictions constituted 88% of the test data. The outcome of the models showed that the Decision Tree performed best followed by the Multilayer Perceptron and the Support Vector Machine. However, by reviewing the confusion matrices presented in Table 4, Multilayer Perceptron would be the preferred solution to be implemented in industrial applications due to the 0% ratio in false-positive results and high accuracy of 88%.

¹ To allow exact reproduction of our results source code is available at <https://github.com/USi-IPEM/Demo-Cell-Classification>

		True Values		Total
		OK	NOK	
Predicted Values	OK	18	0	18 + 0 = 18
	NOK	8	24	8 + 24 = 32
Total		18 + 8 = 26	0 + 24 = 24	50

(a) Support Vector Machine.

		True Values		Total
		OK	NOK	
Predicted Values	OK	18	0	18 + 0 = 18
	NOK	6	26	6 + 26 = 32
Total		18 + 6 = 24	0 + 26 = 26	50

(b) Multilayer Perceptron.

		True Values		Total
		OK	NOK	
Predicted Values	OK	17	1	17 + 1 = 18
	NOK	0	32	0 + 32 = 32
Total		17 + 0 = 17	1 + 32 = 33	50

(c) Decision Tree.

Table 4: Confusion matrices result of the prediction of 50 test data-points. We present the result of the predictions given by all three algorithms in detail based on the comparison between true values vs predictions. The tables read from the upper-left cell as: true-positive, false-positive, false-negative, true-negative.

6. Conclusion

Based on the explorative literature review, we observed that the hundred percent quality check method proves to be the most efficient method to guarantee the final product quality, however, the high implementation cost is its biggest draw-back. On the other hand, we have the statistical methods, which are known to be fast and efficient, but the lack of reviewing each product until the next batch review might imply a considerable loss of the produced products whenever NOK products are found. This is based on the uncertainty of when the error occurred and how many products are affected. Implementing ML models based on just the initial conditions of the process and capable to predict the quality of the product, would be a great asset for the industrial processes. Within the limits of our work, we predicted with 98% certainty the product quality by implementing a Decision Tree model with a 6-variables vector as input. However, our decision for industrial applications would be the MLP with 88% accuracy, to guarantee that all passed parts are OK and avoid labeling NOK pieces as OK. For further work, we expect to increase the accuracy of the prediction model by collecting more data-samples and experimenting with other, more complex, ML models. It is important to state that this prediction model is not yet ready to fully replace the statistical methods, nor the hundred percent quality check method, since this research is still in its early stages and needs further development. However, a combination of SPC along prediction models could work well together, since the prediction model will evaluate each produced piece, and this might compensate for disadvantages of the statistical methods.

7. Lesson Learned

It is worth to mention the most important points learned during the project implementation. This might help as building blocks for further research in the area. Difficulties at implementing ML models in production are also described by others authors [23].

Incompatibility between technologies; devices configured on the same network presented troubles to share data due to network restrictions, as well as the limitation to access and modify restricted program-code given by the machine manufacture. To overcome the situation stated above, it was necessary to reprogram big sections of the machine control system.

Data synchronization; two OPC-UA servers were used on this implementation. This led to a series of data synchronization problems due to differences in the internal clock of both servers. To solve the issue, it was needed to write a time-stamp synchronization script that was deployed during the data preprocessing phase. Without harmonious time-stamps, determining the production cycle would have been impossible and the input data for the ML model would have been useless, and the model inaccurate.

Data interpretation; this process is considered complicated, time-consuming, and a fundamental step for any ML applications. It is fundamental to understand the assembly process, the timing from the machine movements, the boundaries of the machine's sensors, as well a basic correlation within the presented data.

Missing values; to minimize the risk of false results delivered by the ML model, it would be compulsory to carefully review the ML input data. Missing values can have a great impact on the ML prediction, that is the reason why we deliberately review all the input data and validate it with the expected output from the diverse sensors. Whenever there was missing data, and as long as we did not fabricate data, we employed statistical methods to fill the data gap.

Adapting the industry 4.0 demonstration cell; the initial state of the demonstrator was very limited. Even though the process was well known, the lack of labels or flags whenever certain actions occur, made it complicated to have accurate data sampling from each assembly cycle. To overcome this uncertainty, we modified the machine program to raise specific flags whenever certain actions were executed.

Machine Learning; careful measurement and meaningful preprocessing of our cell-data was key to the successful application of all three algorithms. Without the labels, the problem could not have been framed as a supervised classification problem.

Dataset size; as the starting point we collected as many data points like the ones found in the iris Dataset [24]. The next step was to collect batches of data of around 30 minutes per use case. This process was repeated until accumulate around 15 hours of data.

Author Contributions Pérez M. and Wolter writing - original draft, conceptualization, review, and editing; investigation; software and data curation. Schmallenbach writing - review, editing. Burggräf, Wagner, Heinbach, Steinberg writing - conceptualization, review. Garcke, and Steffes-lai writing - review.

Acknowledgments This work was funded by the European Regional Development Fund (ERDF) within the project "ManuBrain".

References

- [1] Amitava Mitra. *Fundamentals of Quality Control and Improvement*. Wiley, 2016.
- [2] B. Illés, P. Tamás, P. Dobos, and R. Skapinyecz. New challenges for quality assurance of manufacturing processes in industry 4.0. *Solid State Phenomena*, 261(481-481), 2017.
- [3] Harry Robinson. Using poka-yoke techniques for early defect detection. In *6th International Conference on Software Testing Analysis and Review*, 1997.
- [4] C Martin Hinckley. Defining the best quality-control systems by design and inspection. *Clinical Chemistry*, 43(5):873–879, 05 1997.
- [5] D. Selvamuthu and D. Das. *Introduction to Statistical Methods, Design of Experiments and Statistical Quality Control*. Springer, 2018.
- [6] Diego A. Sandoval-Chávez and Mario G. Beruvides. Using opportunity costs to determine the cost of quality: A case study in a continuous-process industry. *The Engineering Economist*, 43(2):107–124, 1998.
- [7] Christopher M Bishop. *Pattern recognition and machine learning*. Springer, 2006.
- [8] J. Krauß, J. Dorißen, H. Mende, M. Frye, and R. Schmitt. Machine learning and artificial intelligence in production: Application areas and publicly available data sets. *Production at the leading edge of technology*, 2019.
- [9] Carlos A. Escobar and Ruben Morales-Menendez. Machine learning techniques for quality control in high conformance manufacturing environment. *Advances in Mechanical Engineering*, 10(2), 2018.
- [10] Felix Sohnius, Peter Schlegel, Max Ellerich, and Robert H. Schmitt. Data-driven prediction of surface quality in fused deposition modeling using machine learning. pages 473–481, 2019.
- [11] J. Krauß, B. M. Pacheco, H. M. Zang, and R. Schmitt. Automated machine learning for predictive quality in production. *Procedia CIRP*, 93, 2020.
- [12] S. Aumi, B. Corbett, and Mhaskar P. Model predictive quality control of batch processes. *American Control Conference*, (5646-5651), 2012.
- [13] C. Ritter and F. Schweitzer. Neue wege der prozeßmodellierung für die rechnergestützte qualitätsvorhersage und -sicherung in der spanplattenfertigung. *Holz als Roh- und Werkstoff*, 50(62-66), 1992.
- [14] Dimitris Mourtzis, Nikolaos Milas, and Nikolaos Athinaios. Towards machine shop 4.0: a general machine model for cnc machine-tools through opc-ua. *Procedia CIRP*, 78:301–306, 2018.
- [15] Alessandra Caggiano, Tiziana Segreto, and Roberto Teti. Cloud manufacturing on-demand services for holistic quality assurance of manufactured components. *Procedia CIRP*, 67:144–149, 2018.
- [16] Dimitris Mourtzis, Sophia Fotia, Nikoletta Boli, and Ekaterini Vlachou. Modelling and quantification of industry 4.0 manufacturing complexity based on information theory: a robotics case study. *International Journal of Production Research*, 57(22):6908–6921, 2019.
- [17] D. Delen and H. Demirkan. Data, information and analytics as services. *Decision Support Systems*, 55(359-363), 2013.
- [18] Marcello Colledani and Alessio Angius. Production quality performance of manufacturing systems with in-line product traceability and rework. *CIRP Annals*, 69(1):365–368, 2020.
- [19] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.
- [20] Charu C Aggarwal. *Data classification*. CRC PRESS, 2014.
- [21] Johan AK Suykens, Tony Van Gestel, Joseph De Brabanter, Bart De Moor, and Joos PL Vandewalle. *Least squares support vector machines*. World scientific, 2002.
- [22] Stephen Marsland. *Machine Learning: An Algorithmic Perspective*. CRC Press, 2 edition, 2015.
- [23] Andrei Paleyes, Raoul-Gabriel Urma, and Neil D. Lawrence. Challenges in deploying machine learning: a survey of case studies. page 9, 2020.
- [24] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, 1936.