

# PREDICTION OF WORKPIECE QUALITY: AN APPLICATION OF MACHINE LEARNING IN MANUFACTURING INDUSTRY

Günther Schuh<sup>1</sup>, Paul Scholz<sup>2</sup>, Sebastian Schorr<sup>3</sup>, Durmus Harman<sup>2</sup>,  
Matthias Möller<sup>4</sup>, Jörg Heib<sup>4</sup>, Dirk Bähre<sup>3</sup>

<sup>1</sup>Laboratory for Machine Tools & Production Engineering (WZL), RWTH Aachen University Aachen, Germany

<sup>2</sup>Fraunhofer Institute for Production Technology, RWTH Aachen University, Aachen, Germany

<sup>3</sup>Institute of Production Engineering, Saarland University, Saarbrücken, Germany

<sup>4</sup>Bosch Rexroth AG, Bexbacher Street, Homburg, Germany

## ABSTRACT

*A significant amount of data is generated and could be utilized in order to improve quality, time, and cost related performance characteristics of the production process. Machine Learning (ML) is considered as a particularly effective method of data processing with the aim of generating usable knowledge from data and therefore becomes increasingly relevant in manufacturing. In this research paper, a technology framework is created that supports solution providers in the development and deployment process of ML applications. This framework is subsequently successfully employed in the development of an ML application for quality prediction in a machining process of Bosch Rexroth AG. For this purpose the 50 most relevant features were extracted out of time series data and used to determine the best ML operation. Extra Tree Regressor (XT) is found to achieve precise predictions with a coefficient of determination ( $R^2$ ) of constantly over 91% for the considered quality characteristics of a bore of hydraulic valves.*

## KEYWORDS

*Technology Management Framework, Quality Prediction, Machine Learning, Manufacturing, Workpiece Quality*

## 1. INTRODUCTION

Due to the digitalization of manufacturing companies, an increasing number of machine tools are connected – the manufacturing process is mapped digitally [1, 2]. The analysis of the recorded process data allows for measures to improve quality, time, and costs [3]. Therefore, a considerable amount of data is generated, on basis of which decisions often have to be made correctly and in real time [4]. However, large amounts of data and different data sources significantly increase the complexity of the evaluation. Commonly programmed software encounters limitations when processing large amounts of data while facing the high complexity of the application environment. For some problems, however, no algorithm can be used, since an infinite number of scenarios is conceivable and not all variations can be covered rule-based [5, 6]. Given the currently high level of capacity utilization and a thoroughly optimized

production, further efficiency increases are likely to be possible only through the introduction of completely new technologies.

ML is known as the field of study that gives computers the ability to learn without being explicitly programmed and is considered as particularly effective method of data processing with the aim of generating usable knowledge from data [7]. ML systems have the potential to capture complex correlations instantaneously from unstructured data, constantly improve analyses and dynamically adapt to external environmental conditions, which is why they are regarded as particularly promising to further optimize the production process [7-9]. ML applications therefore bear the potential to further increase efficiency in the manufacturing industry and thus ensure the attractiveness of production locations in high-wage countries [4]. ML as a result is becoming increasingly relevant and is being applied to the manufacturing process.

However, there is the challenge that ML as a still rather novel technology in the manufacturing industry is relatively unknown and untested [9]. It is therefore difficult for potential solution providers to fully assess the technology potential, identify relevant applications in the manufacturing industry and thus develop functional solutions to relevant problems. Hence, a framework that supports solution providers in the development process and deployment of ML applications is needed. The first goal of this research paper is to develop a multi-layered structure that can be used to develop ML applications. The second goal is to develop an application for quality forecasting on basis of the multi-layered structure.

Following on from the introduction, Chapter 2. discusses in depth the subject area of ML in manufacturing. In this section, among other things, a framework for applying ML in manufacturing is developed and a use case from Bosch Rexroth AG is described in order to finally select an appropriate path in the framework regarding the specific technology demands of the use case investigated. Chapter 3. deals with the data collection process, the feature extraction, importance, and selection as well as the specific ML operations, their performance results, and evaluation. Chapter 4. concludes this research work and refers to special challenges and future research-related potentials.

## **2. MACHINE LEARNING IN MANUFACTURING**

### **2.1. Framework for Applying ML in Manufacturing**

According to the opinion of SCHUH & SCHOLZ, ML represents the most important among many working areas of artificial intelligence such as e.g. Image Processing and Vision or Natural Language Understanding [9-11]. At this point in time, ML can be interpreted as a rather diverse bundle of technologies. The subject area is still much disorganized. A rather unstructured and partial discussion revolves around different aspects of the technology at different levels [9]. The necessity of the development of such a framework has already been widely demonstrated. Also the different technology layers required for the description of ML applications are derived. Within this research paper the framework will be enriched with more detailed technology information regarding the various decision alternatives per layer as well as a decision sequence along the different layers during the development of ML applications [9]. The framework is designed to provide an easily accessible theoretical overview of the subject area, to accompany researchers as well as solution providers during the implementation process and to support the development of solutions. The framework was developed in vertical reading direction and claims to classify problems from the ML domain into these layers. By the defined processing of the individual layers the selection options are narrowed down more and more, whereby only certain classes and decision alternatives of the following layer are selectable.

Gradually, the decision alternatives of the framework become more granular, which makes the classification of the problem more thorough. A classification according to ML learning strategies is employed because all applications learn in a certain way and specific data types are related to these learning strategies such as e.g. supervised learning, unsupervised learning and reinforcement learning [12]. ML learning strategies are based on the natural learning mechanism of humans [13]. Supervised learning involves learning from examples. Each set of input (features) is labeled with a specific output. After each processed sample, the internal parameters are adjusted (e.g. Neural Networks (NN): weights, Decision Trees (DT): branches, Support Vector Machines (SVM): hyperplane etc.) to minimize a certain error function, often in the sense of the mean square error. In unsupervised learning, only the available data are used to recognize correlations and patterns without defined target classes [13]. Reinforcement learning can shortly be described as the mapping from situations to actions to maximize a scalar reward. The learner is not told which action to take, but instead must discover which action yields to the highest reward by testing [14]. ML tasks such as e.g. classification, regression, and clustering refer to the ML learning strategies and take the specific data formats of the problem into account [12]. Supervised learning can be divided into regression and classification tasks, while unsupervised learning is mainly used for clustering tasks. In the case of regression problems, intervalscaled data are predicted, whereas in the case of classification problems, discrete characteristics such as good or bad can be predicted. The clustering task can be compared with that of the classification, but the target classes are derived from data and are not defined beforehand [13]. ML operations such as e.g. Random Forest (RF) [15], (DT) [15], and (NN) [16] represent the definite procedures which are used for the analysis. They can be described by their internal model structure (e.g. algorithm, hyper parameters, loss function and other general properties). Each ML operation thereby has an individual structure. ML implementation procedures determine which methodical procedure (e.g. KDD (knowledge discovery in databases), Crisp-DM (cross-industry standard process for data mining), and SEMMA (acronym describing the individual steps of the method)) which programming language (e.g. R and Python) and library (e.g. Scikit-Learn, TensorFlow, Keras, and Theano) should be used [17-19]. The development of methodologies, software tools and languages serve as the standardized approach for industrial applications of data processing [20]. Methodical procedures support the systematic implementation of real-life ML applications. The selection of the contents from this layer depends on the use case. For the development of a ML application the layers are processed and defined in the sequence presented here. The well founded examination of the fundamental literature and implemented ML applications allows for an empirical inductive conclusion on the following features per layer and suggests an integration of the layers into the following decision sequence. The wide variety of literature sources presented above support this classification [9-20].

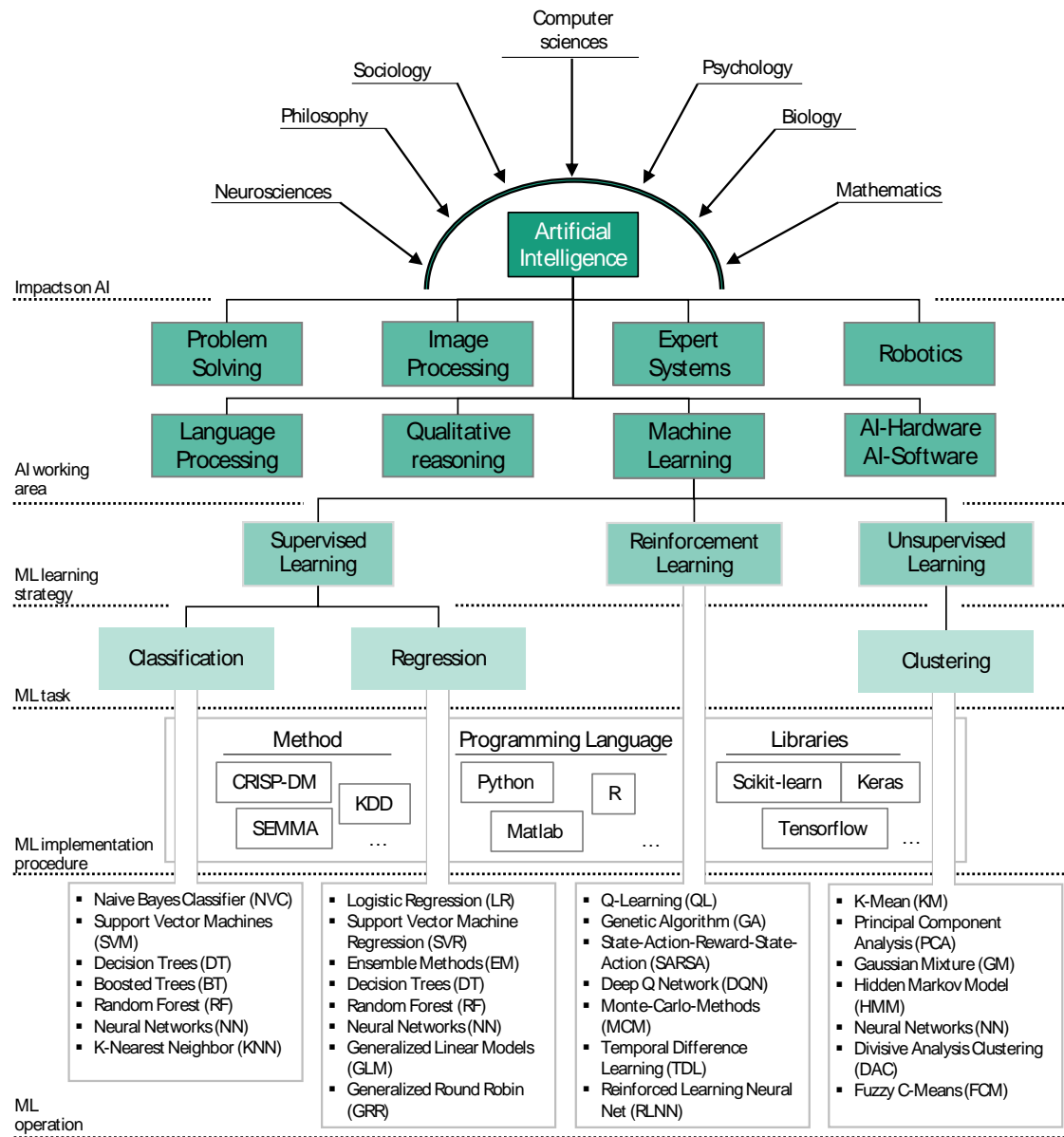


Figure 1. Framework for applying ML in manufacturing

The specific benefit of the framework for the user is that each analysis attempt using ML represents a unique path in the framework. Established paths in the framework can serve as a blueprint for a potential future application design. Thus, less development time may be needed to solve certain problems. In addition, there are more validated and proven solution options to problems, which ultimately leads to lower overall opportunity costs for solution providers.

## 2.2. Description of the Use Case

The manufacturing process considered within this research is an integral part for the serial production of hydraulic valves at the Bosch Rexroth AG. Hydraulic valves are characterized by bores with tolerances of only a few microns to enable seal-less fits and to prevent oil leakage. A valve is machined first by a milling machine, assembled and finally tested, see Fig. 2. Slight anomalies in the production process can lead to unacceptable quality deviations causing high scrap rates and financial losses. To guarantee the required quality of bores of a valve, sampling

inspection after the machining process (approx. 1% of a batch) and end of line testing (100% of a batch) are applied in industry. Both quality control methods lead to indirect costs and a high latency between the machining of a valve and its measurement results. Therefore, feasible and affordable in-process quality control methods are desired from manufacturing companies. The so gained increased transparency over the machining process is used to make adjustments as soon as the required quality is not reached. Process data from the machining process and ML operations are used to predict the quality of one of the most quality critical bores of a valve.

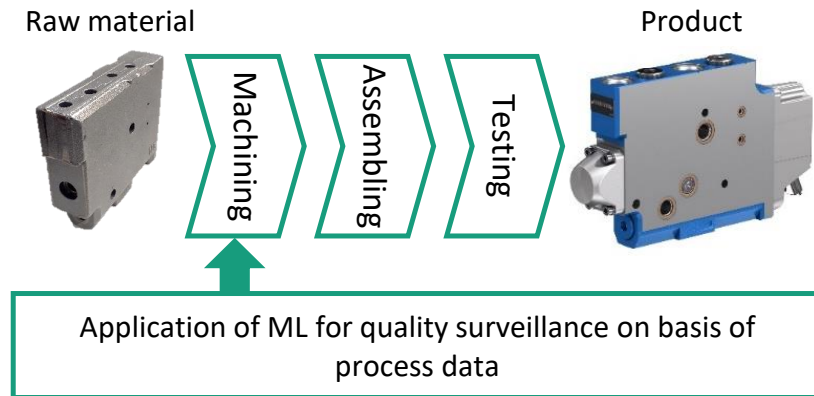


Figure 2. Manufacturing process of a hydraulic valve with quality surveillance

The quality of a bore is determined by dimensional (diameter) and locational (concentricity) quality characteristics. Quality deviations are caused by wear on cutting tools that are used to drill and to ream bores. Tool wear leads to increased cutting forces and torque, which can be measured directly from the drives of a milling machine in form of the motor current and torque [20, 21]. This indirect measurement approach together with ML operations is the only economical technique to obtain quality predictions with a latency close to zero when facing industrial conditions. Such a quality surveillance detects quality deviations in an early manufacturing state and enables cost and resource savings.

### 2.3. Selection of the Right Path in the Framework Suitable to the Use Case

ML is considered as one working area of AI that can take the manufacturing industry to a next level. In order to cope with the use case considered in this research, technologies from the context of ML should therefore be used. First the learning strategy supervised learning is chosen because the quality of the considered bore of each valve is available and less training data are required to build a model. This leads to a less time-consuming data collection process. In addition, the diameter and concentricity of the bore have to be expressed as numerical values, which defines the ML task as a regression task. For the ML implementation procedure the method CRISP-DM is chosen because an understanding of the business and the data already exists and data preparation is required in form of feature extraction and selection. Due to modelling and evaluation the best ML operation suitable for the use case is determined and finally deployed. The lists of ML operations helps to choose the right category of algorithm regarding a regression ML task. For the implementation the programming language Python and the library Scikit-learn in combination with the distribution Anaconda are used. These choices determine the path in the framework to accomplish the quality prediction of machined workpieces.

### 3. METHODOLOGY

Fig. 3 summarizes the methodology of this research to obtain a quality prediction for each workpiece from the data of the related machining process. Each process step is described in detail in the following subsections.



Figure 3. Approach to obtain a quality prediction from machining data

#### 3.1. Data Collection Process

The process data were collected during the serial production of hydraulic valves from a milling-machine. To operate a milling-machine a numerical control unit (NC) is used, which processes data from sensors integrated into the machine and the drives. The measured actual values obtained by sensors give a direct feedback on the cutting conditions and are used to control the machine. These machine-internal data were directly collected with a frequency of 1,000Hz from the NC and are the input data for the feature extraction. Selected features are used to make predictions with ML operations. The stored process data were the actual torque values of the drive of the z-axis as well as the spindle, the actual position values of the x-, y-, and z-axis and also the actual speed value of the spindle. Process data from a total of 160 workpieces were collected during the machining. Fig. 4 exemplarily depicts the actual torque value of the spindle for the second tool, which are used to machine the considered bore.

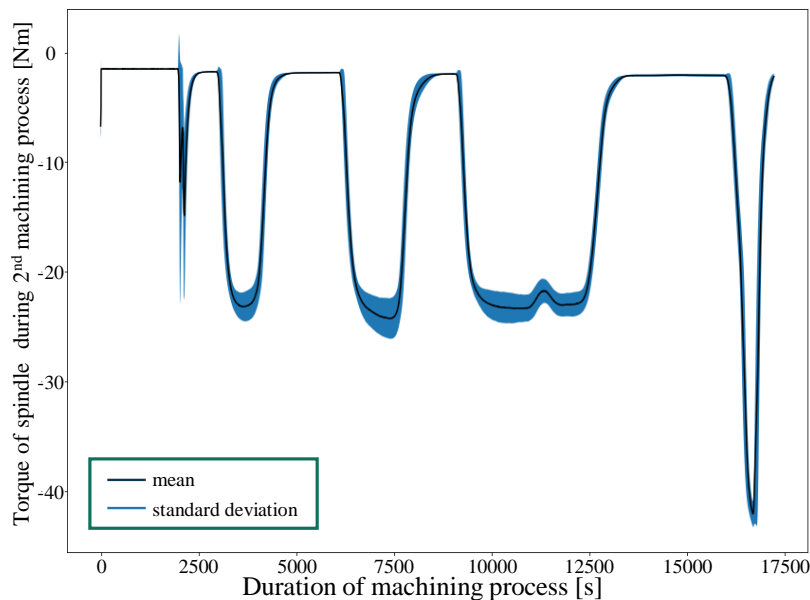


Figure 4. Visualization of process data of the 2nd machining process

In addition, for each workpiece the dimensional and locational quality characteristics were measured representing the target and output parameter for the ML operations. The measured characteristics are the diameter and concentricity of a bore.

### 3.2. Feature Extraction

To reduce the complexity and volume of the collected time series data as well as to accelerate the training and test process of the ML operations a feature extraction is applied. The aim is to extract characteristics in order to accordingly represent the time series with a reduced set of features. General features (e.g. the shape of the signal) and statistical features (e.g. mean or standard deviation) are very common to represent time series data. However, both forms of features often do not contain all the information needed to expose dynamical time series data. Therefore, VUNUNU ET AL. also analyzed the frequency domain to identify additional patterns in the data [22]. In this paper, 63 methods are used to calculate 794 features for each machining operation with the tsfresh library. For each process parameter a total number of 2382 features is obtained due to the three subprocesses [23]. The extracted feature can be divided into three groups as shown in Fig. 5.

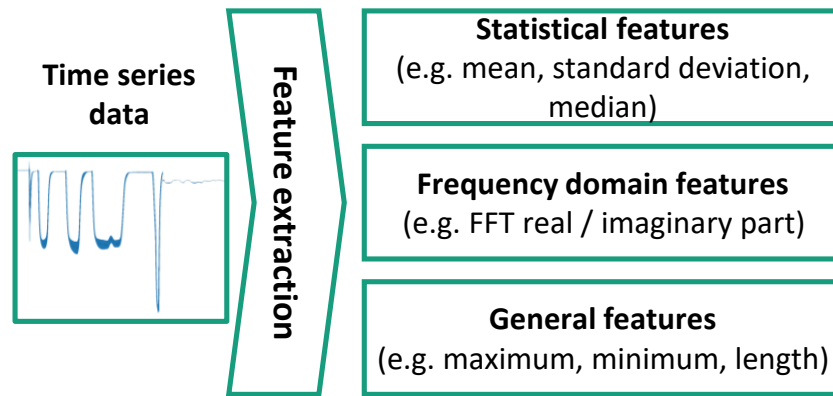


Figure 5. Visualization of process data of the 2nd machining process

### 3.3. Feature Importance and Selection

Irrelevant features can lead to weak prediction performances and increased computational costs. In a huge set of features the probability of containing irrelevant features is high and therefore a feature selection is recommended. The determined importance of each feature increases the interpretability of a feature in the related use case and enables to select the features that contribute most to an accurate and efficient prediction [24]. In this paper, the RF technique is used to determine the feature's importance. The RF algorithm uses subsets of the training data to construct each of the tree configurations. Therefore, an independent test set is not required for the evaluation of the feature ranking. The feature importance is obtained by comparing the prediction errors after permuting the feature's values of all examples [25]. Feature importance of the RF is well suited for datasets with a small sample and big feature size [26]. The importance of all features is calculated and the features with the lowest importance values are excluded from the data. The features are ranked in descending importance order as shown in Fig. 6. The three most important features are the low frequency fast Fourier transformation (FFT) coefficient of tool 3 (8.97%), the approximate entropy of tool 2 (7.93%) and the high frequency FFT coefficient of tool 2 (6.79%).

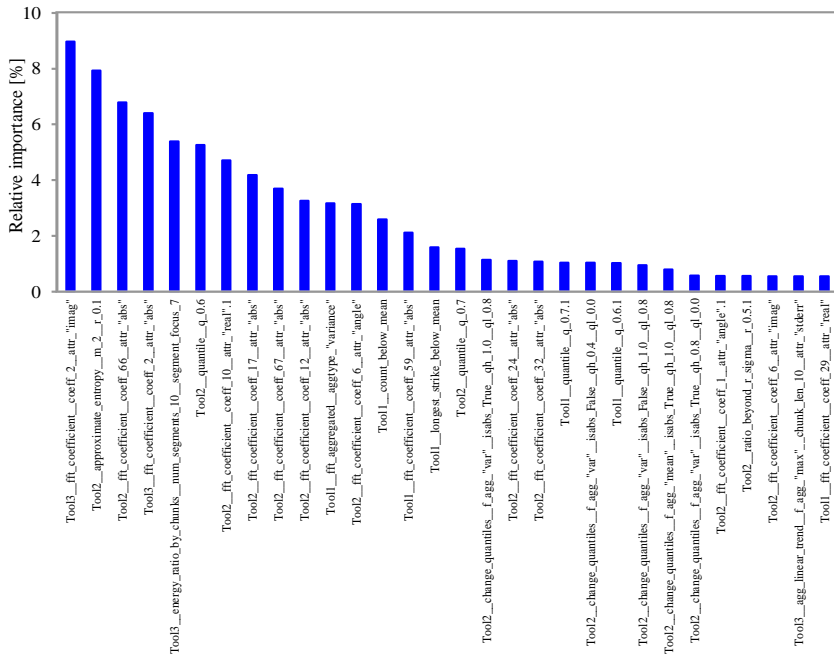


Figure 6. The 30 most important features

To decrease the model complexity and the computational costs the 50 most important features are selected. For the selection of the most relevant features a systematic approach is considered. This procedure answers the question of how many features are needed to maintain the initial accuracy. The initial accuracy is determined by a grid-searched operation based on all features. The individual features ranked by importance are made incrementally available to the operation as a training set. First, the operation is trained with the highest ranked feature. Then, the next important feature is inserted into the feature table and the procedure is repeated. For each feature subset the  $R^2$  is evaluated with a five-fold cross validation. These steps are performed for every feature in the dataset. The mean and the standard deviation of the results for the first 50 important features are shown in Fig. 7.

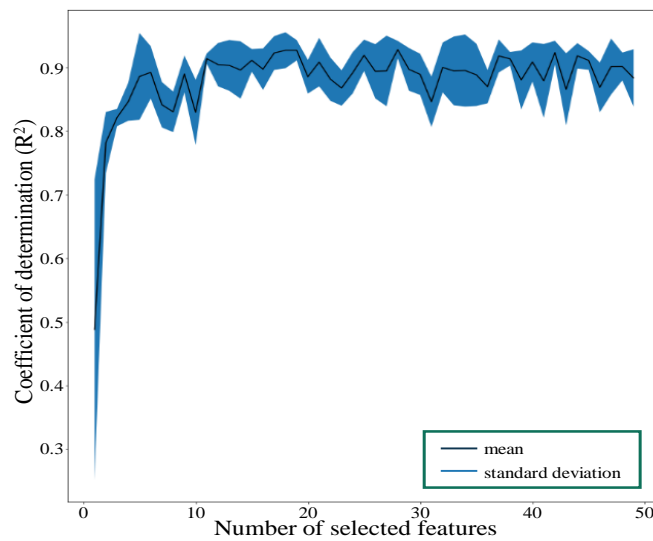


Figure 7.  $R^2$  plotted over the number of selected features



It can be stated that already with the ten most important features a  $R^2$  of 91.39% is reached. Additional features do not lead to an extraordinary increase in accuracy or a saturation behavior, which is possibly due to the small number of samples. By adding more features, the accuracy decreases (not shown here). Thus, the number of features which serve as inputs to the ML operations is set to 50.

### 3.4. Machine Learning Operation

An artificial neural network (ANN) is a computational model consisting of a collection of artificial neural neurons inspired by the brain of living beings. All those neurons are linked via interconnections to form a large network. Information are insert into the network via input neurons and are conveyed within the network to provide results to the output neurons [27, 28]. The output of each neuron  $y$  is determined according to equation (1) by its activation function  $\phi$  that activates the neuron if the weighted sum of  $n$  inputs  $x_i$  multiplied with the related weights  $w_i$  is above a particular threshold  $u$  [28].

$$y = \phi \left( \sum_{i=1}^n w_i x_i - u \right) \quad (1)$$

Ensemble methods combine several ML algorithms to achieve a higher prediction accuracy. Often these methods are tree-based like Random Forest (RF) and Extra Trees (XT). Random Forest, developed by BREIMAN, is a ML operation for classification and regression problems [15]. An ensemble of decision trees is used to obtain the final prediction  $y$  by averaging the predictions  $y_t$  of all trees  $T$  for a given dataset  $x$  as denoted in equation (2). To build a single tree bagging is applied that describes the random selection of a sample from the original training dataset. RF do not overfit due to bootstrap aggregation, can reach highly accurate predictions, and are fast to train [29, 30]. The ML operation XT differs from RF in two main points to grow the trees. First, XT uses the entire training dataset instead of bagging. Second, splitting-points are chosen fully randomly to split a node [31]. In this paper the RF and XT, as regression tree approaches, are used to predict numerical values for the quality characteristics.

$$y(x) = \frac{1}{T} \sum_{t=1}^T y_t(x) \quad (2)$$

Support Vector Machine (SVM) is a ML operation for classification and regression tasks. Input data are mapped to a high dimensional feature space where the prediction task gets linearly separable. The support vector kernel together with the number of support vectors and its parameter determine the decision function of SVM [32, 33]. Further explanation can be obtained from STEINWART & CHRISTMANN as well as VAPNIK [33, 34].

For the considered ML operations ANN, RF, XT, and SVM the optimal hyperparameters have to be determined. For this purpose, a parameterspace is set up in which different variations of hyperparameters are examined in a grid and the  $R^2$  of each combination is measured. For the individual operations, both the absolute and the repetitive accuracy are calculated. Thus, the best model configuration obtained from the grid search can be determined from the results shown in Table 1.

Table 1. Results of Gridsearch

ML operation	R <sup>2</sup> [%] (absolute)	R <sup>2</sup> [%] (repetitive)
Artificial Neural Network (ANN)	55.13 ± 5.30	51.63 ± 3.20
Random Forest Regressor (RF)	91.28 ± 2.81	89.59 ± 0.82
Extra Trees Regressor (XT)	93.64 ± 0.65	92.81 ± 0.18
Support Vector Machine Regressor (SVM)	<0	<0

The highest R<sup>2</sup> can be obtained for the XT Regressor. The associated standard deviation is the lowest of all ML operations, which proves the suitability for this task. The hyperparameters and other properties for the best operation are listed in Table 2.

Table 2. Extra Trees Regressor summary

Parameter	Value
Splitting criterion	MSE
Max. features to consider best split	50
Min. number of samples to split	3
Number of trees	10
Mean fit time	0.115 sec
Mean score time	0.103 sec
R <sup>2</sup> of training data	99.95 ± 0.02%
R <sup>2</sup> of test data	93.64 ± 0.65%

### 3.5. Results and Evaluation

The prediction accuracy reached with the described ML operation (XT) is evaluated using the R<sup>2</sup>, the mean squared error (MSE), and the mean absolute error (MAE). Table 3. summarizes the results for the two predicted quality characteristics diameter and concentricity. The diameter is predicted with a higher precision compared to the concentricity. The MAE is with 0.26 μm very little and the MSE is close to zero. For the concentricity the MAE and MSE are much higher than the obtained results for the diameter but the predicted values are still precise enough for the utilization in the described use case.

Table 3. Extra Trees Regressor summary

	Diameter	Concentricity
R <sup>2</sup> [%]	93.75	91.66
MSE [mm <sup>2</sup> ]	1.2x10 <sup>-7</sup>	0.008
MAE [μm]	0.26	22.91

That fact that the desired quality characteristics can be predicted very reliably with XT can be retrieved from the high R<sup>2</sup> (Table 3.) as well as in the graphical comparison of the predicted and the measured values in Fig. 8.

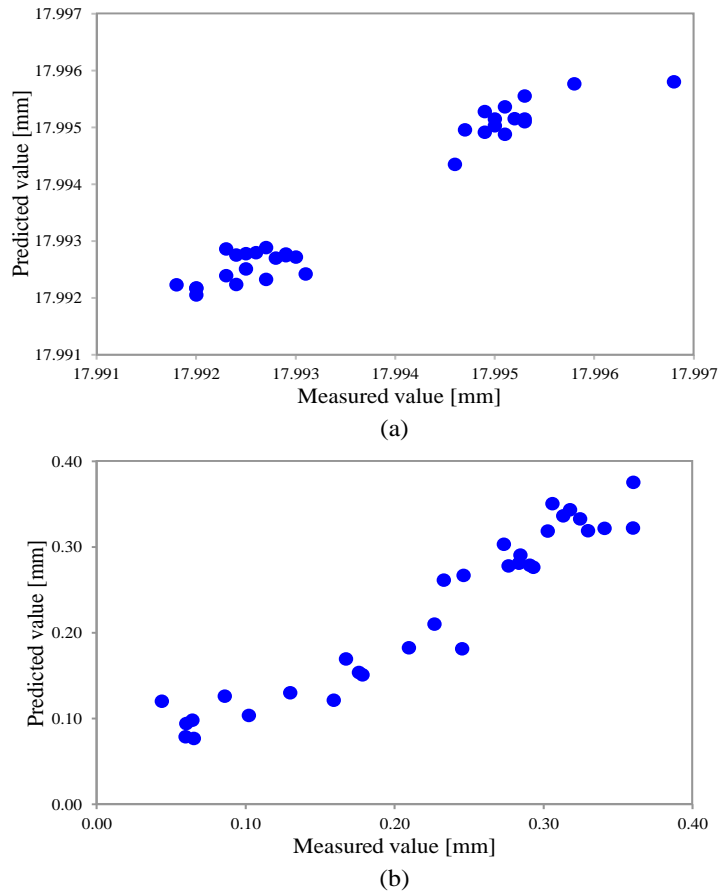


Figure 8. Graphical representation of the predicted and measured values for (a) diameter and (b) concentricity.

To diagnose the systems behavior the learning curve is reviewed. The  $R^2$  is constantly high for the training data, which is typical for an XT. However, a learning effect can still be observed as the  $R^2$  improves with an increasing number of training examples from around 70% to 94%. An  $R^2$  of 100% is hardly achievable, since the learning curve stagnates. In this case a saturation did not occur. Fig. 9 depicts the learning curve. The coefficient of determination ( $R^2$ ) is plotted with an increasing number of training examples. It can be observed that the accuracy increases with an increasing number of training samples, which proves that the XT is becoming continuously accurate.

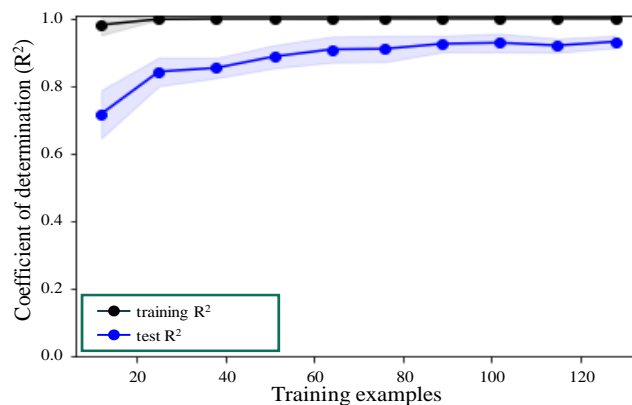


Figure 9. Learning curves of Extra Trees Regressor

## 4. CONCLUSION

Engineers in the manufacturing industry often have only limited knowledge of using and applying AI or ML to solve problems in their business. Vice versa have data scientists only basic expertise concerning manufacturing processes. The framework presented in this paper enables engineers to gain an overview of AI and ML and its related subfields. The framework guides through the ML implementation process and depicts potential tools and approaches for the application of ML in the manufacturing industry. For a specific use case from the Bosch Rexroth AG the engineers applied the framework successfully. All required decisions were taken along a chosen path in the framework which finally led to a ML operation that predicts the quality characteristics of a bore very precisely. A feature extraction and selection is mandatory to reduce the complexity of the time series data as well as the computing time and to increase the prediction accuracy. From the 794 extracted features, 50 features were determined as significant and were considered for the prediction models. The predictions for the diameter and the concentricity of the considered bore were evaluated with the statistical criterions  $R^2$ , MSE, and MAE. With a MAE of only 0.26  $\mu\text{m}$  for the diameter and 22.91  $\mu\text{m}$  for the concentricity the prediction errors were very low. Hence, XT can be used to predict the quality of bores on basis of process data close to real time. This direct feedback regarding the manufacturing process leads to scrap avoidance, resource savings and cost reductions. Consequently, the application of ML is a further contribution to secure production plants in high-wage countries.

In the future, we will gather more data to increase the training samples for the ML operations in order to improve the prediction accuracy. With more training samples ML operations like ANN or SVM will potentially lead to acceptable results, too. Furthermore, we are looking for further use cases where we can apply the described framework to achieve improvements by implementing ML.

## REFERENCES

- [1] X. Yao, J. Zhou, J. Zhang, and C. R. Boer, "From Intelligent Manufacturing to Smart Manufacturing for Industry 4.0 Driven by Next Generation Artificial Intelligence and Further On," in 5th International Conference on Enterprise Systems (ES): IEEE, September 2017, pp. 311–318.
- [2] M. Syafrudin, G. Alfian, N. L. Fitriyani, and J. Rhee, "Performance analysis of IoT-Based sensor, Big Data processing, and Machine Learning model for real-time monitoring system in automotive manufacturing," *Sensors*, vol. 18, 2018.
- [3] E. Alpaydin, *Introduction to Machine Learning*. MIT Press Cambridge, vol. 3, 2014, pp. 1-20.
- [4] McKinsey Global Institute, *Artificial Intelligence. The Next Digital Frontier?*, <https://www.mckinsey.com/~media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx#page=55>, 2017.
- [5] E. Brynjolfsson and T. Mitchell, "What can machine learning do? Workforce implications," *SCIENCE*, vol. 358, December 2017, pp. 1530-1534.
- [6] A. Agrawal, J. Gans, and A. Goldfarb, *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, 2018.
- [7] A. Géron, *Hands-on machine learning with Scikit-Learn and TensorFlow. Concepts, tools, and techniques to build intelligent systems*, Sebastopol, O'Reilly Media, 2017.

- [8] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, p. 886, 2015.
- [9] G. Schuh and P. Scholz, "Development of a framework for the systematic identification of AI application patterns in the manufacturing industry," 2019, unpublished.
- [10] J. Lunze, *Künstliche Intelligenz für Ingenieure. Methoden zur Lösung ingenieurtechnischer Probleme mit Hilfe von Regeln, logischen Formeln und Bayesnetzen*. Berlin, De Gruyter Oldenbourg, 2016.
- [11] N. J. Nilsson, *Principles of Artificial Intelligence*. Palo Alto, Morgan Kaufmann, 2016.
- [12] J. Brownlee, *A Tour of Machine Learning Algorithms*, <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>, 2013.
- [13] S. Raschka, V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow*. Birmingham: Packt Publishing, 2017.
- [14] R. S. Sutton, *Reinforcement Learning*. New York: Springer Science+Business Media, 1992.
- [15] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5-32, 2001.
- [16] G. D. Rey and K. F. Wender, *Neuronale Netze. Eine Einführung in die Grundlagen, Anwendungen und Datenauswertung*. Bern, Huber, 2008.
- [17] U. Shafique and H. Qaiser, "A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA)," *International Journal of Innovation and Scientific Research*, vol. 12, pp. 217–222, 2014.
- [18] G. Piatetsky, Python overtakes R, becomes the leader in Data Science, Machine Learning platforms, <https://www.kdnuggets.com/2017/08/python-overtakes-r-leader-analytics-data-science.html>, 2017.
- [19] S. Shetty, Why TensorFlow always tops machine learning and artificial intelligence tool surveys, <https://hub.packtpub.com/tensorflow-always-tops-machine-learning-artificial-intelligence-tool-surveys/>, 2018.
- [20] F. Klocke, *Manufacturing Processes 1: Cutting*. New York: Springer, 2011.
- [21] M. Eynian, K. Das, A. Wretland, "Effect of tool wear on quality in drilling of titanium alloy Ti6Al4V, Part I: Cutting Forces, Burr Formation, Surface Quality and Defects," *High Speed Machining*, vol. 3, pp. 1-10, 2017.
- [22] C. Vununu, K. Ryong, E. J. Lee, K. S. Moon, S. H. Lee, "Automatic fault diagnosis of drills using artificial neural networks," *16th IEEE International Conference on Machine Learning and Application*, pp. 992-995, 2017.
- [23] M. Christ, N. Braun, J. Neuffer, A. W. Kempa-Liehr, "Time series feature extraction on basis of scalable hypothesis tests (tsfresh – A Python package)," *Neurocomputing*, vol. 307, pp. 72-77, 2018.
- [24] I. Guyon, A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.
- [25] J. Rogers, S. Gunn, "Identifying feature relevance using a random forest," in: *Subspace, Latent Structure and Feature Selection*. Berlin: Springer-Verlag, pp. 173-184, 2006.

- [26] C. Strobl, A. L. Boulesteix, A. Zeileis, T. Hothorn, "Bias in random forest variable importance measures: Illustrations, sources and a solution," *BMC Bioinformatics*, 2007.
- [27] I. N. Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, S. F. R. Alves, *Artificial Neural Networks: A practical course*. Switzerland: Springer International Publishing, 2017.
- [28] H. Rathore, *Mapping Biological Systems to Network Systems*. Switzerland: Springer International Publishing, 2016.
- [29] G. Biau, "Analysis of a random forests model," *Journal of Machine Learning Research*, vol. 13, pp. 1063-1095, 2012.
- [30] A. Cutler, D. R. Cutler, J. R. Stevens, *Random Forests*, in: *Ensemble Machine Learning: Methods and Applications*, New York: Springer Science+Business Media, pp. 157-175, 2012.
- [31] P. Geurts, D. Ernst, L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, p. 3-42, 2006.
- [32] L. Zhang, W. Zhou, L. Jiao, "Wavelet Support Vector Machine," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 34, pp. 34-39, 2004.
- [33] I. Steinwart, A. Christmann, *Support Vector Machines*. New York: Springer Science+Business Media, 2008.
- [34] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.