Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 92 (2020) 63–68

www.elsevier.com/locate/procedia

16th CIRP Conference on Computer Aided Tolerancing (CIRP CAT 2020)

Virtual In-line Inspection for Function Verification in Serial Production by means of Artificial Intelligence

Raphael Wagner*, Jakob Fischer, Daniel Gauder, Benjamin Haefner, Gisela Lanza

wbk Institute of Production Science, Karlsruhe Institute of Technology (KIT), Kaiserstr. 12, 76131 Karlsruhe, Germany

* Corresponding author. Tel.: +49 1523-950-2627; fax: +49-721-608-44005. *E-mail address:* Raphael.Wagner@kit.edu

Abstract

In high-tech production, companies often deal with the manufacture of assemblies with quality requirements close to the technological limits of manufacturing processes. The article shows an approach of a virtual in-line inspection, predicting the products functionality. An artificial neural network (ANN) fed with product characteristics and process data as well as the resulting functional fulfillment of the product is trained for virtual function prognosis. Through the preventive identification of defective products before the final assembly step, components can be recovered and returned to serial production. By optimizing the parameters of the ANN, incorrect classifications are reduced and the practical applicability is ensured. The approach is demonstrated in an industrial application in the automotive industry.

Keywords: Quality Control; Artificial Intelligence; Predictive Model

1. Motivation

The functional requirements for new products in many industries are becoming increasingly demanding. For example, combustion engines are required to consume less fuel and emit fewer emissions without compromising performance. Precisely manufactured fuel injectors enable it to meet these requirements, resulting in more precise injection quantities and thereby optimized combustions.

Due to the rising functional requirements parts must be manufactured closer to their target specifications. This results in tighter tolerances for the components close to the current technological limits [1,2]. Technologically immanent process deviations make it impossible to meet these high quality requirements eventually leading to scrap parts.

If holistic functional fulfillment of a product is examined, interdependencies between components with production deviations can nevertheless enable the functionality of the product in some cases. By utilizing product knowledge, these deviating components can possibly still be used and the yield of good parts increased if the deviations of components are compensated by other subsystems.

While selective assembly is successful with product characteristic interdependencies, it is not always possible for complex products with multiple complex interdependencies between influencing factors of different production stations [1]. In the manufacture of high-precision products, a high degree of non-linearity between these characteristics often occurs due to different influencing factors from parallel manufacturing machines, assembly and testing stations. These different influencing factors can only be represented with difficulty, if at all, by common statistical methods. Currently the functionality deviations of such products for which selective assembly is not feasible often only become apparent during the final functional test after assembly and thus lead to scrap parts and rising production cost.

To verify the product function of such complex products extensive quality testing is required. In-line verification of quality features in real-time is one approach to detect errors early in production and prevent unnecessary value creation of defective components [3–5]. A virtual function verification

10.1016/j.procir.2020.03.126

based on the production data and artificial intelligence approaches seems to be an effective measure to detect non-conforming component combinations earlier in the process then at the final inspection [6,7].

The article is structured as follows: After a description of the state of the art in chapter 2, the development of the virtual function verification is methodically explained in chapter 3 and demonstrated in chapter 4 with an exemplary application. Chapter 5 concludes with a summary and outlook.

2. State of the Art

In-line quality control provides direct feedback on deviations in production systems and thus contributes to efficient and reliable production (see Fig. 1). In the production line integrated in-line measurement technology is used to determine the current production quality and then react accordingly. [8]

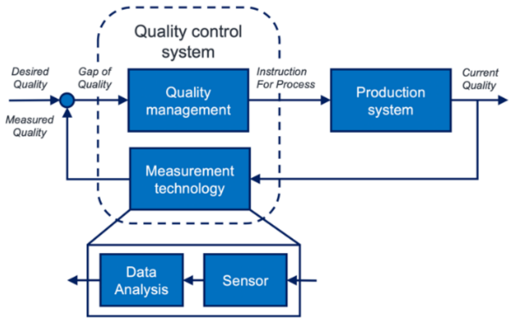

Figure 1: In-line quality control [8]

Artificial intelligence is the discipline of creating intelligent problem-solving machines. Machine learning which refers to artificial learning from experience is a subset thereof.

Thus, machine learning methods provide the possibility to model relationships from data for the purposes of classification and regression [9]. Classification assigns pre-specified classes to an object based on one or more input properties [10]. Therefor the algorithm utilizes input and output data to learn about the functional relationship in form of supervised learning. Potential classification methods for supervised learning are e.g. artificial neural networks or support vector machines [11]. Artificial neural networks are inspired by biological neural networks and are capable of modelling complex relationships between inputs and outputs (see Fig. 2) [12].

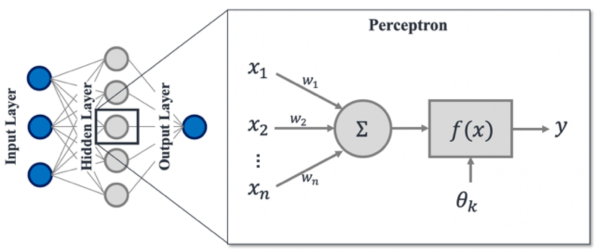

Figure 2: Perceptron and ANN with one hidden layer, cp. [13]

Thus they can be suitable to model the functional relation between deviations especially if a large data base (e.g. from the production ramp-up or previous product generations) is available for training of the model [14]. Artificial neural networks have been successfully applied to model different complex interrelations [15–18].

Artificial neural networks use multiple nodes, so called perceptrons, to calculate an output y from n inputs x_i and the activation function f. The weights w_i and biases θ_k of the perceptrons are determined by learning from existing training data sets (see Eq. 1). [12]

$$y = f\left(\sum_{i=1}^{n} (w_i x_i) + \theta_k\right) \quad (1)$$

After training and testing the ANN model can be applied and predict the output for unknown input data.

Data in which certain classes are underrepresented can influence a classification in favor of the frequently represented class [19]. In the case of such an imbalance, the classification tends to adapt excessively to the majority class. The problem of unbalanced data is widespread in areas such as risk management, fraud detection and medical diagnosis [20–22].

Cluster based under-sampling of the data is a possibility to avoid the bias of the classifier by reducing the incidents of the majority class in the training data in a representative way through clusters [23].

For classification models the confusion matrix summarizes the classification performance. The true class of an object is plotted against the class assigned by the classifier (see Fig. 3). In the case of classifications with two classes, one class is often referred to as positive and the other as negative. Depending on the classification, one also speaks of true positive (TP), false positive (FP), false negative (FN) and true negative (TN). [24]

To determine how accurate the classification recognizes the individual classes, precision and recall can be used for evaluation (see Eq. 2 and 3) [25].

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Often there is a dependency between precision and recall allowing to optimize one parameter at the expense of the other. Therefore the F_β-Score is a weighted key figure [26], where a positive real number β determines how much the recall is weighted against the precision (see Eq. 4).

$$F_\beta = \frac{(1 + \beta^2) * Precision * Recall}{\beta^2 * Precision * Recall} \quad (4)$$

As a conclusion of the state of the art, currently there is no research approach enabling a virtual in-line function verification in serial production before final testing of products with complex interdependencies where selective assembly is not feasible. However, machine learning approaches such as ANN, provide the potential to model these complex interrelations based on in-line production data.

3. Research Approach

In the following, a research approach for a virtual function verification is presented. Using an artificial neuronal network, a virtual function verification is developed from the production data. The model classifies the injectors into conforming and non-conforming products regarding their function fulfillment. The general data mining approach is roughly based on the Knowledge Discovery in Databases (KDD) process [27].

The approach requires in-line data of the relevant features and properties for the classification of products as well as a data base of suitable production and programmable logic control data to train the model. It furthermore requires an understanding of the production process and product itself with its function-determining features that must be incorporated into the model.

Based on the understanding of the product and the production process a performance indicator for model quality is required. While the F_β-Score is a weighted average, the cost of false classifications in quality assurance also known as α-error and β-error can be quantified [5].

Products correctly identified as non-conform by the classifier are true positives (see Fig. 3).

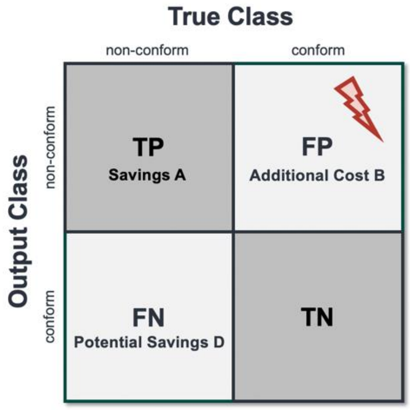

Figure 3: Confusion matrix with cost consideration

Each true positive (TP) therefore saves costs. False positives (FP) in contrast are conform products declared as non-conform by the classifier and thus generate additional costs through unneeded reworking or scrap. False negatives (FN) on the other hand are non-conform parts misclassified as conform. These do not represent a deterioration in the status quo but are potential savings that have not been identified. In order to take this into account in the evaluation of model quality and model selection, these are evaluated in terms of opportunity costs as penalties in the amount of the lost potential savings. True negatives (TN) are products whose function has been correctly classified as conform.

The evaluation of costs and savings can be used to assess the economic performance of an in-line classification model (see Eq. 5).

$$C = \frac{A * TP - B * FP - D * FN}{TP + FP + TN + FN} * N \tag{5}$$

Whereas A is the saving from a correctly detected non-conform part, B the additional cost generated by a conform part incorrectly being rejected and D the unused cost saving potential of not detected scrap parts.

N is the annual production quantity and TP, FP and FN being the number of parts classified accordingly. C is a theoretical cost saving value to economically evaluate the performance of the classification model in usage, which can turn negative if there is a large number of FN, even though the model effectively generates savings. In order to assess the actual impact of the classification model C_{eff} can be calculated without opportunity cost consideration (see Eq. 6).

$$C_{eff} = \frac{A * TP - B * FP}{TP + FP + FN + FP} * N \tag{6}$$

While this value indicates the effective costs and savings of the classifier, it does not take into account the non-identified scrap parts.

3.1. Data Preprocessing

In the first step, the features relevant for functionality are determined by product experts and the required production data extracted from the quality control database.

Before the production data can be used to train the classification model, it must be preprocessed, i.e. missing and erroneous data must be handled accordingly. Therefore, incorrect values in the in-line production data are removed while general noise is not, as it will also occur, e.g. measurement uncertainty, when the model is applied.

To train the artificial neural network in form of supervised learning the training data needs to be labeled. The target values, therefore, must be converted from scalar values to nominal classes. For the classification it is possible to choose two or three target classes (see Fig. 4). In the case of two target classes, only a distinction is made between conform and non-conform parts. With three classes, a differentiation is made whether the tolerance limit has been exceeded or fallen short of.

For aggregated characteristics consisting of the results of several operation states the entire product is also classified accordingly. In the few cases in which the various operation states of the product has contrary results, i.e. one measured value below and another one above the allowed tolerance limit, the product is not classified as a whole and the corresponding entry in the data remains empty.

Figure 4: Target class labelling

Due to measuring times that exceed the production cycle time statistical process control is applied when possible. Therefore, some features are not available for every component. This leads to a large percentage of the use case data tuples being incomplete. The affected features are examined whether the missing values can be interpolated to increase the data basis. To interpolate these features, the production time of the component is used as reference variable for linear interpolation.

Subsequently, the data is standardized so that the artificial neural network can process the data more efficiently and faster [28]. Nominal features such as station numbers and the target feature, however, are excluded from the standardization. To prevent large nominal numbers from distorting the classification results, these values are transformed to smaller numbers.

3.2. Under-Sampling

In production data the share between conform and non-conform parts in most cases is unbalanced. This class imbalance is an intrinsic problem, since the goal is always to achieve the highest possible yield of conform parts. In order to minimize the bias of the classification model due to the class imbalance, cluster based under-sampling proves to be suitable [29].

To ensure that the under-sampling does not distort the test results, a test data set must be separated from the data before the under-sampling of the training data (see Fig. 5). For the test data set random data tuples are selected to prevent systematic influences. The test data set thus still approximates the real class distribution independently of the under-sampling.

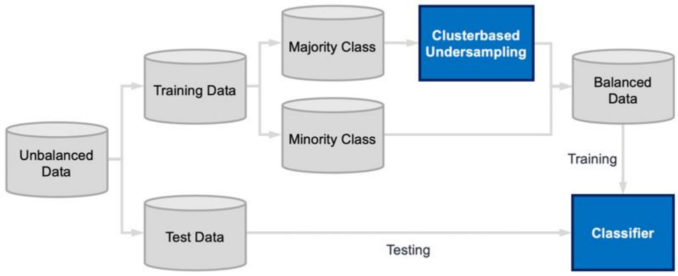

Figure 5: Cluster based under-sampling process [29]

For the under-sampling the majority class is grouped into k clusters with k-means++ clustering. Each of those k clusters are reduced to one data point. There are two possible methods for the reduction, either the cluster center is used as a representative artificial data point or the next real data point is determined by the Euclidean distance [29].

So k determines the reduced number of data tuples of the majority class. This minimizes the class imbalance, leading to the share of the majority class v_{new} in the under-sampled data (see Eq. 7).

$$v_{new} = \frac{k}{k+m} \tag{7}$$

With m being the quantity of data tuples of the minority class in the training data.

3.3. Artificial Neural Network

A two-layer feed-forward network, with sigmoid hidden and softmax output neurons, is then trained with supervised learning of the preprocessed data using the MATLAB® Neural Net Pattern Recognition toolbox. For testing the performance of the artificial neural network hold-out validation is performed. In order to minimize random influences, the mean value of the confusion matrix entries is determined from five training and corresponding test iterations.

In order to optimize the artificial neural network parameters a sensitivity analysis is performed in the form of a statistical design of experiments. Thereby, the influence of several parameters such as the number of hidden neurons, number of epochs, learning rate and number of target classes are examined with a Box-Behnken design. The target function of the analysis is to maximize C (see Eq. 8).

$$\max_{TP,FP,FN\,\in\,\mathbb{N}} C\,(TP, FP, FN) \tag{8}$$

For the training function that updates the weights w_{ki} and biases values θ_k of the perceptrons Bayesian regularization backpropagation, Levenberg-Marquardt backpropagation and Scaled conjugate gradient backpropagation are considered.

For evaluation of the analysis the averaged confusion matrices are used and for each classification model and the corresponding C and C_{eff} calculated. This way the influence of the individual parameters is approximated.

While C_{eff} indicates the actual impact on production, the theoretical value C is used to determine the best classification model so that the number of unrecognized scrap parts is also taken into account.

4. Case Study

Together with an industrial partner, the research approach is applied to an exemplarily use case regarding the production of a piezo controlled common-rail injector for combustion engines in passenger cars and light weight utility vehicles. The product consists of several high-precision subsystems from various disciplines such as mechanics, mechatronics and fluid mechanics (see Fig. 6).

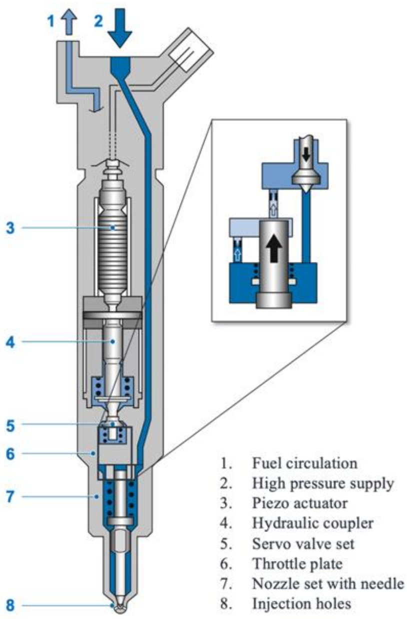

Figure 6: Product principle of a high-pressure injector, cp. [30]

All subsystems of the product, such as piezo actuator, servo valve, throttle plate and nozzle set, affect the functionality of the injector significantly. Precise injection quantities with tolerance less than a milliliter are defined as the functional quality requirements for several operating states simulating different driving situations. The product not only has to meet high precision in volume portioning but also high-dynamic control with up to five injections per cycle at high pressure to meet performance and efficiency requirements [30].

As small deviations in a component already influence the flow rate, high quality requirements are assigned to the relevant characteristics to assure the product functionality. In addition, non-linear, process-related variations occur, which can be assumed due to the manufacture of the products at different production stations. However, the technological limits of the manufacturing processes are reached resulting in a comparatively high percentage of scrap, and thereby increased production costs.

In the final assembly step the injectors are screwed together. Up to this stage the injectors can easily be dismantled and the components reused. If a faulty product is detected after this assembly step the dismantling is more costly and due to plastic deformation not all components can be recovered. Following the screwing the functionality of the injectors is tested by measuring the fuel flow at the different operating states. If the flow rate at one of the operating states does not meet the functional tolerances, the injector is scrap and gets dismantled so that individual components can be recovered.

Therefore, a virtual in-line function verification is introduced before the final assembly step (see Fig. 7). The virtual function verification utilizes the in-line production data to classify a product with regard its functionality.

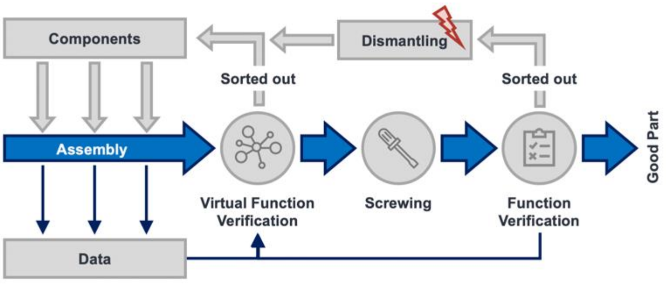

Figure 7: Virtual in-line function verification

Based on the classification result the injectors can be sorted out before screwing and thus costs are reduced (see Fig. 8). Dismantling of unscrewed products is only 27 % of the dismantling cost of non-conforming tested products in the final functional test, after screwing.

Figure 8: Cost structure before and after screwing

For the virtual function verification an artificial neural network is trained and tested with production data from 197,981 injectors with 17 input features each. To evaluate the ANN and the effects of certain parameters the corresponding effective savings C_{eff} and the savings with penalties for unrecognized scrap parts C are calculated.

4.1. Results

The results of the ANN were evaluated with a five times hold-out validation and additionally compared with other methods of machine learning, such as Decision Trees, Support Vector Machines (SVM), k-Nearest-Neighbor (kNN) and ensemble methods, i.e. subspace kNN (see Table 1). All these classification models are combining all operating states which simulate different driving situations in one aggregated feature.

The effect of cluster based under-sampling unfortunately led to mixed results with smaller data sets depending on the examined operating state in pre-analyses. Since the calculations are highly computation-intensive they were not examined on the entire data set. Nevertheless, in further applications, the method could significantly improve the ANN performance.

Table 1: Classification results for various models

Model	Precision	Recall	F_1	C	C_{eff}
no model applied				-500,643	0
ANN (max C)	**0.197**	**0.271**	**0.228**	**-115,700**	**68,212**
Fine Tree	0.786	0.131	0.234	-349,934	54,705
Fine kNN	0.496	0.416	0.480	-73,107	129,046
Quadratic SVM	0.938	0.075	0.138	-396,508	34,000
Subspace kNN	**0.692**	**0.553**	**0.615**	**8,022**	**215,875**

With the design of experiments the configuration of the artificial neural network was optimized. The effect of the parameters such as the number of neurons per hidden layer, epochs, learning rate, ratio of test data, number of target classes and the backpropagation training function on the neural network are examined in the design of experiments (see Fig. 9).

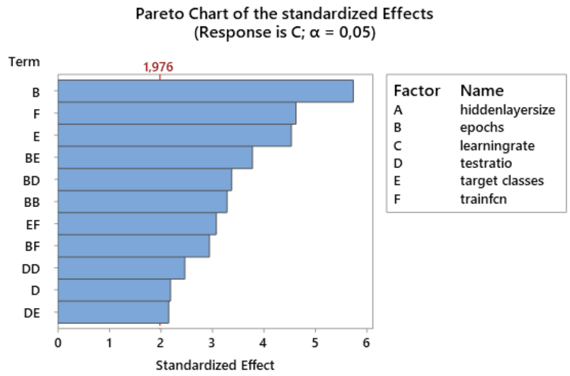

Figure 9: Pareto chart of the standardized effect of the parameters

The optimized ANN was obtained classifying the injector into the three classes achieved slightly better results than two target classes. This could be due to general differences in parts with results above and below the tolerance. Regarding the training function the Bayesian regularization backpropagation showed the best results, exceeding the other functions. The optimal parameters are 250 epochs, a learn rate of 0.136, with a test data ratio of 25% and a hidden layer size of 10.

With these parameters the ANN reaches a precision of 0.197 and a Recall of 0.271, leading to an F_1-Score of 0.228. Even though the theoretical cost savings C value is negative due to the penalties for non-identified defective parts, the artificial neural network effectively generates cost savings C_{eff} of 68,212 €/year.

The number of epochs also has great influence on the results but increases training times considerably. Further training of the ANN with more epochs could therefore additionally improve the results. However, due to nonlinearities, this approach does not necessarily find the global optimum and cannot be used without restriction.

The other machine learning models are obtained with a 10-fold cross-validation with two target classes instead of three. Even without further optimization the subspace k-Nearest-Neighbor ensemble outperforms the other approaches including the ANN, resulting in an effective cost savings value of 215,875 €/year. Overall k-Nearest-Neighbor seems to be a suitable approach with good classification results for the given production data.

5. Summary and Outlook

In this article, an approach for a virtual in-line inspection for function verification during the production process is presented. This contributes to a cost-efficient serial production of complex products with high quality requirements close to the technical production limits through proactive and cost-effective dismantling of non-conforming products. Therefore, in-line production data is utilized to classify the product with an artificial neural network regarding its function fulfillment before the final assembly step. The performance of the artificial neural network was optimized by parameter variation in the training phase using statistical design of experiments. For evaluation of the classification model cost saving assessments are introduced. The approach is demonstrated to the production of a common-rail injector leading to reduced scrap parts and cost savings.

In the future, it will become increasingly necessary to provide users with a guide to the application of artificial intelligence models. Potential further research must therefore be carried out on further methods to optimize model accuracy by varying parameters during the learning phase of the models. Furthermore, research can be carried out on continuously learning approaches of Artificial Intelligence. This ensures that the models continuously learn and take into account time effects and changes in influencing factors in the functional evaluation of products.

Acknowledgements

This article is part of a research project that is funded by the German Federal Ministry of Education and Research (BMBF) within the Framework Concept ''Photonics for flexible, networked production'' (funding number 13N14647) and managed by the Project Management VDI. The authors are responsible for the contents of this publication.

References

[1] Lanza G, Haefner B, Kraemer A (2015) Optimization of selective assembly and adaptive manufacturing by means of cyber-physical system based matching. *CIRP Annals* 64(1):399–402.
[2] Peter M, Fleischer J Rotor balancing by optimized magnet positioning during algorithm-controlled assembly process: Selection and assembly of rotor components minimizing the unbalance. *2014 4th International Electric Drives*, pp. 1–4.
[3] Schmitt R, Niggemann C, Isermann M, Laass M, Matuschek N (2010) Cognition-based self-optimisation of an automotive rear-axle-drive production process. *Journal of Machine Engineering* 10(3):68–77.
[4] Moroni G, Petrò S (2013) Inspection Strategies and Multiple Geometric Tolerances. *Procedia CIRP* 10:54–60.
[5] Moroni G, Petrò S, Tolio T (2011) Early cost estimation for tolerance verification. *CIRP Annals* 60(1):195–198.
[6] Haefner B, Lanza G (2017) Function-oriented measurements and uncertainty evaluation of micro-gears for lifetime prognosis. *CIRP Annals* 66(1):475–478.
[7] Weckenmann A, Hartmann W (2013) Function-oriented method for the definition and verification of microstructured surfaces. *Precision Engineering* 37(3):684–693.
[8] Lanza G, Haefner B, Schild L, Berger D, Eschner N, Wagner R, Zaiß M (2019) In-Line Measurement Technology and Quality Control. in Gao W, (Ed.). *Metrology*. Springer Singapore. Singapore, pp. 399–433.
[9] Rosenblatt F (1962) *Principles of Neurodynamics*. Spartan, Washington D.C.
[10] Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods* 43(1):3–31.
[11] Vapnik VN, Vapnik VN (1995 // 1998) *The nature of statistical learning theory*. 2nd ed. Springer, New York.
[12] Bishop CM (2010) *Neural networks for pattern recognition*. Oxford Univ. Press, Oxford.
[13] Jain AK, Mao J, Mohiuddin KM (1996) Artificial neural networks: a tutorial. *Computer* 29(3):31–44.
[14] Wagner R, Haefner B, Lanza G (2018) Function-Oriented Quality Control Strategies for High Precision Products. *Procedia CIRP* 75:57–62.
[15] Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H (2018) State-of-the-art in artificial neural network applications: A survey. *Heliyon* 4(11):e00938.
[16] Tosun E, Aydin K, Bilgili M (2016) Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures. *Alexandria Engineering Journal* 55(4):3081–3089.
[17] Asiltürk İ, Çunkaş M (2011) Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method. *Expert Systems with Applications* 38(5):5826–5832.
[18] Haefner B, Biehler M, Wagner R, Lanza G (2018) Meta-Model Based on Artificial Neural Networks for Tooth Root Stress Analysis of Micro-Gears. *Procedia CIRP* 75:155–160.
[19] Japkowicz N, Stephen S (2002) The class imbalance problem: A systematic study1. *IDA* 6(5):429–449.
[20] Chawla NV, Japkowicz N, Kotcz A (2004) Editorial. *SIGKDD Explor. Newsl.* 6(1):1–6.
[21] Richardson AM, Lidbury BA (2017) Enhancement of hepatitis virus immunoassay outcome predictions in imbalanced routine pathology data by data balancing and feature selection before the application of support vector machines. *BMC medical informatics and decision making* 17(1):121.
[22] Rahman MM, Davis D, Ao SI (2013) Cluster based under-sampling for unbalanced cardiovascular data. *International Conference of Data Mining and Knowledge Engineering* 3:3–5.
[23] Yen S-J, Lee Y-S (2009) Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications* 36(3):5718–5727.
[24] Ting KM, Shultz TR, Fahlman SE, Craw S, Andritsos P, Tsaparas P, Silva R, Drummond C, Ling CX, Sheng VS, Lanzi PL, Gama J, Wiegand RP, Sen P, Namata G, Bilgic M, Getoor L, He J, Jain S, Stephan F, Sammut C, Harries M, Pfahringer B, Case J, Wagstaff KL, Nijssen S, Wirth A, Zhang X, Cancedda N, Renders J-M, Michelucci P, Oblinger D, Keogh E, Mueen A (2010) Confusion Matrix. in Sammut C, Webb GI, (Eds.). *Encyclopedia of Machine Learning*. Springer US. Boston, MA, p. 209.
[25] Vazirgiannis M, Halkidi M, Gunopulos D (2003) *Uncertainty handling and quality assessment in data mining*. Springer, London.
[26] Chinchor N (1992) MUC-4 evaluation metrics. in Unknown, (Ed.). *Proceedings of the 4th conference on Message understanding - MUC4 '92*. Association for Computational Linguistics. Morristown, NJ, USA, p. 22.
[27] Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery in databases. *AI magazine* 17(3):37.
[28] Jayalakshmi T, Santhakumaran A (2011) Statistical normalization and back propagation for classification. *International Journal of Computer Theory and Engineering* 3(1):1793–8201.
[29] Lin W-C, Tsai C-F, Hu Y-H, Jhang J-S (2017) Clustering-based undersampling in class-imbalanced data. *Information Sciences* 409-410:17–26.
[30] Reif K (2020) *Dieselmotor-Management: Systeme, Komponenten, Steuerung und Regelung*. 5th ed. Springer Fachmedien Wiesbaden, Wiesbaden.