Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 83–98

Application of Basic Machine-Learning Classifiers for Automatic Anomaly Detection in Shewhart Control Charts

Aleksander Woźniak*, Klaudia Krawiec**, Roger Książek***

Abstract. In today's dynamic technological environment, innovation plays a crucial role – especially for manufacturing enterprises that constantly strive to improve the quality of their products. This article examines the quality-management issue in a company producing car rims. It was identified that real-time quality control can sometimes be unreliable due to controller fatigue, leading to erroneous data interpretation or delayed responses to deviations in the production process. The study aimed to investigate the possibility of eliminating or significantly reducing these errors by employing a tool that is based on artificial intelligence. The article covers the preparation of training data, the training of classifiers, and the evaluation of their effectiveness in analyzing control charts in real time. The adopted hypothesis assumes that machine-learning classifiers can be effective methods of support for quality controllers. The research began with collecting measurement data from the machine and dividing it into training and test sets. The obtained results were evaluated using standard quality measures for machine-learning models. The results showed that the use of artificial intelligence can bring significant benefits in improving quality supervision in the production process of car rims.

Keywords: Machine Learning, Artificial Intelligence, AI, Statistical Process Control, SPC, Quality Control, Classifiers, Quality Metrics, Python Programming, Car Wheel, Quality Issues

Mathematics Subject Classification: 68T10

JEL Classification: C44

Submitted: June 5, 2024

Revised: October 28, 2024

* AGH University of Krakow, Faculty of Energy and Fuels, Krakow, Poland, e-mail: alekwozniak@student.agh.edu.pl
** AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: klakrawiec@student.agh.edu.pl
*** AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: roger@agh.edu.pl

DOI: https://doi.org/10.7494/dmms.2024.18.6345

1. INTRODUCTION

The origins of Shewhart control charts can be traced back to the 1920s, when they were first developed for quality control. These represent a statistical tool that monitors process stability by tracking sequence samples and identifying patterns that deviate from expected norms. By distinguishing between common-cause and special-cause variations, this method enables the early detection of shifts in production, thus ensuring consistent quality control over time. Shewhart control charts have been employed for qualitative analysis, becoming a key tool for collecting data during production and later analyzing it (Shewhart, 1926).

It needs to be stressed that Shewhart control charts remain in use even today; however, their roles have evolved with modern digital tools. This shift has introduced several important considerations regarding the role of AI in quality control. It remains uncertain whether AI-based systems can fully replace human operators or if they will function more effectively as supplementary tools. The potential of AI to improve the accuracy and precision of measurements in rim production is significant; however, challenges and limitations come with its implementation. The effectiveness of different machine-learning algorithms in identifying and eliminating measurement errors must also be evaluated, along with the necessary data for adequately training AI models. Addressing these uncertainties is crucial for determining how AI can be optimally integrated into modern manufacturing processes (Malindzakova et al., 2023; Tran et al., 2022).

In the rim-manufacturing company to which this paper refers, their rims are sold on a global scale (both branded and for retail purposes). Quality control operators in the company were responsible for manual production measurements and the use of a special machine during the MD860 testing phase, which measured the necessary dimensions of each produced rim. Despite the use of Shewhart control charts, a key issue persisted in continuous quality control: the system relied heavily on human cognitive capabilities. Employee fatigue became a significant concern, as the constant manual supervision that was required for monitoring rims through the measurement device led to decreased efficiency and a heightened risk of errors. This human-centered limitation underscored the need for more-automated and -reliable solutions in the quality control process. This became the motivation for using the data sets from this process for research into the potential advantages of using simple AI classifiers over decision-making algorithms.

This paper showcases the process of selecting the most suitable machine-learning classifiers for supporting the automated detection and recognition of the patterns that were recorded in the Shewhart control charts that were used in the quality control in the above-mentioned rim-manufacturing process. Data sets that detailed the features of the rim-manufacturing process were utilized to train and evaluate selected machine-learning classifiers to detect and recognize classic seven-sample sequences that signal process degradation. Based on the performed literature review (which is presented in more detail in Section 2, showing the scarcity of works on the application of basic machine-learning classifiers for automatic analyses of seven-sample sequences in Shewhart control charts), data from the rim-manufacturing process was used to conduct the research. The study focused on basic ML classifiers due to their simplicity, thus allowing the presented experiment to be repeated on data sets from other processes without requiring significant hardware resources nor the performance of tedious advanced-programming work. The preparation of simple classifiers for pattern analysis is less time-consuming and less engaging than, for example, the training of an artificial neural network; so, the classifiers can be used to identify potentially promising research areas such as an n-sample sequence that precedes a seven-sample sequence that is indicative of process degradation. As this was a pilot study, the main research question (RQ) that was posed was as follows: which basic ML classifiers achieve acceptable performance (F1-score) in recognizing seven-sample patterns that are indicative of process degradation?

This paper is structured as follows. In Section 2, the research gap of insufficient research on the application of machine learning for enhancing statistical-process control via the automated analysis of Shewhart control charts is identified based on a systematic literature review. In Section 3, the control sequences that were selected for the experiments are presented, while in Section 4, the approach for training and evaluating the performance of the selected ML classifiers is explained. Section 5 reports and discusses the obtained results, and the main conclusions are presented in Section 6.

2. PAPER POSITIONING

Since the early 1990s, there has been growing interest in applying machine learning (ML) to recognize patterns on Shewhart control charts. These studies have typically approached the subject from the perspective of enhancing the detection accuracy and efficiency of the deviations from the process norms and anomalies. Techniques like support vector machines (SVMs) have emerged as particularly effective in automating the analysis of data from control charts – often outperforming traditional statistical methods in many use cases. However, much of the existing research has tended to overlook the crucial aspect of interpreting the results generated by ML-based systems – especially in the context of more dynamic and complex production environments.. This gap in the literature is significant because interpreting ability is key to ensuring that these systems can be effectively implemented in real-world scenarios; therefore, this article seeks to fill this gap by conducting a systematic review of the literature, focusing on the integration of ML with Shewhart control charts. The aim is to identify any recent developments that address the interpret ability of ML outcomes and explore how these approaches can be applied in actual production processes (Hwarng, 1992; Shewhart, 1992)

For this study, scientific articles from the SCOPUS database that covered the period from 2010 through 2024 were selected. The search criteria included the use of key phrases that were related to ‘Machine learning,’ ‘Machine learning + SPC,’ and ‘Machine learning + automotive.’ The study was conducted in September 2024 using the results of database queries from the day of September 4, 2024; the data filtering followed these steps:

1. **Machine learning** – included articles where the phrase ‘Machine learning’ was present. The number of results increased year by year (from 46,162 publications in 2010 to 780,574 in 2024).

2. **Machine learning + SPC** – included articles that contained both 'Machine learning' and 'SPC'. The number of results grew from 232 in 2010 to 2279 in 2024.
3. **Machine learning + automotive** – included articles that contained both 'Machine learning' and 'automotive'. The number of results rose from 125 publications in 2010 to 3060 in 2024.
4. **Machine learning + automotive + SPC** – included articles where the phrase 'Machine learning' and 'automotive' and 'SPC' was present. In total, 37 articles were identified that met all of the above criteria. Most of these were rejected due to not fitting within the thematic scope of the case at hand. A few were selected for analysis.

In those articles that focused on the utilization of machine-learning classifiers for automatic anomaly detection on Shewhart control charts, the data was analyzed and utilized to recognize irregularities (in a similar fashion as with other studies). However, Staněk et al. (2023) adopted the real-time recognition and classification of wheels and rims as the primary goal, while Rameshkumar et al. (2021) focused on using acoustic-emission features to predict the conditions of grinding wheels; the article concentrated on using Shewhart control charts to detect any anomalies in the processes. Compared to Lee et al. (2023), which examined the recognition of wheel-component conditions using machine learning, the research that this article focused on was anomaly detection in general (rather than specific wheel components). Additionally, there are articles such as Fang et al. (2023) and Krummenacher et al. (2017), which also employed machine-learning techniques to improve efficiency and safety in the railway industry. Despite the differences in the application and research area, all of these articles harnessed the potential of machine learning to enhance the effectiveness and quality of production processes and identify any abnormalities.

Moser et al. (2021) presented a machine-learning-based emission model for diesel engines that supported on-board diagnostics. This application of ML in automotive systems contributes to improving the control of emissions and is relevant to the automotive industry's performance and environmental sustainability. In Oh et al. (2019), the authors introduced a real-time quality-assessment system for an automotive-production process that used support vector machines (SVMs). The research was highly relevant to automotive safety and control (SPC), as it leveraged machine learning to ensure manufacturing precision in the assembly of sunroofs. Sharmin et al. (2022) explored a machine-learning approach to intrusion detection in automotive CAN networks. This research focused on cybersecurity for vehicle networks, making it highly relevant to automotive SPC by ensuring that secure data flows in in-vehicle systems. In Staf and McKelvey (2018), the authors proposed predictive models for brake performance in vehicles. While this paper did not heavily focus on ML, the predictive nature of the models ties into vehicle control systems and contributes to performance optimization, making it somewhat relevant to SPC. Lampe and Meng (2024) discussed a curated data set for improving automotive cybersecurity via machine learning. The relevance to SPC was clear, as it provided a foundational data set for developing ML models that enhance automotive network security. In Dettu et al. (2024), the focus was on integrating physical and virtual models for optimizing vehicle dynamics. While machine learning was not a core aspect, the approach was relevant to SPC through its emphasis on improving vehicle performance and control.

True et al. (2021) explored THz technology for semiconductor testing; this was not directly related to automotive SPC or machine learning, so this paper was less relevant to the automotive field. Lokman et al. (2019) provided a comprehensive review of intrusion-detection systems for automotive CAN networks, focusing on the use of ML for network security. It was highly relevant to automotive SPC – particularly in the context of enhancing vehicle cybersecurity that uses machine-learning techniques.

In Xu et al. (2024), the focus was on digital twin technologies for materials science. While interesting, it did not directly apply to automotive SPC or ML, making it less relevant to this domain. Zhou et al. (2023) reviewed monitoring techniques for selective laser melting in manufacturing. Though relevant to advanced manufacturing processes, it did not have a direct connection to ML or automotive SPC, thus making it less applicable to the field. The article by Mjimer et al. (2023) discussed the role of ML in industrial continuous improvement, but it was not focused on automotive applications; therefore, it was not strongly relevant to automotive SPC. Gong et al. (2023) reviewed machine learning in the context of laser-based manufacturing. Although it touched on ML, the focus was more on manufacturing processes outside the automotive industry, making it less relevant to automotive SPC. The paper by Benzaza et al. (2023) discussed improving quality-management systems in the automotive industry, but it did not focus heavily on machine learning. While relevant to automotive SPC in terms of quality control, the lack of a strong ML focus made it less applicable. Tsenev and Ivanova (2022) provided insights into using ML to evaluate control systems in automotive production lines; this was directly relevant to both automotive SPC and machine-learning applications. The paper by Lestyán et al. (2019) applied machine learning to identify drivers based on CAN network data. It was highly relevant to automotive SPC and ML – particularly in the context of vehicle data security and driver identification. Minawi et al. (2020) discussed a machine-learning-based system for detecting intrusions in automotive CAN networks. This work was highly relevant to automotive SPC, focusing on cybersecurity. In Habibullah et al. (2024), the authors explored software engineering for automotive perception systems. Although relevant to automotive systems, the paper did not focus on ML, making it only partially aligned with automotive SPC. Park and Baek (2023) presented a cyber-attack detection system using decision trees for automotive cyber-physical systems. This paper was highly relevant to both automotive SPC and ML due to its focus on vehicle network security. Lampe and Meng (2023) introduced a data set for automotive intrusion detection, leveraging ML for enhanced vehicle security; this was directly applicable to both automotive SPC and ML. Shi et al. (2016) focused on validating ML systems for autonomous driving, thus ensuring reliability in automotive applications. This paper was highly relevant to automotive SPC and ML – particularly in the field of autonomous vehicles. Escobar et al. (2021) discussed how AI and ML could enhance automotive manufacturing quality control, thus making it relevant to both automotive SPC and ML. In Kidmose and Meng (2024), the authors examined and evaluated data sets for intrusion-detection in automotive systems using machine learning. This was directly applicable to both ML and automotive SPC.

Cui et al. (2021) applied machine learning for predictive modeling in materials science, but it was not specifically focused on automotive SPC. In Robinson et al. (2020), the authors discussed testing techniques for automotive component reliability; however, there was little connection to machine learning, thus making it less relevant to automotive SCP. Hofmann et al. (2024) focused on ML applications in additive manufacturing, but it was not directly relevant to automotive SCP. Kalyanasundaram et al. (2018) applied machine learning for detecting denial of service (DoS) attacks on automotive CAN networks, thus making it highly relevant to both automotive SCP and ML. The review by Paturi et al. (2023) focused on machine-learning applications in manufacturing, but it was not directly related to automotive SCP (making it less relevant). Finally, Alfardus and Rawat (2023) proposed a hybrid machine-learning method for detecting cyberattacks in automotive networks, thus making it highly relevant to both ML and automotive SCP.

The reason for choosing simple classifiers is that they can analyze production data and detect subtle patterns that might indicate any potential quality issues that a human controller might overlook. These minor deviations can be easily identified by an algorithm, but they may be missed by manual inspections. The manual inspection of each rim in mass production is time-consuming and prone to errors, whereas ML can quickly analyze vast amounts of production data – especially when combined with large data sets; this increases process efficiency. Additionally, machine-learning models can be trained on various data sets, allowing them to adapt to changing production conditions or new rim specifications, providing greater flexibility as compared to static quality control methods. Following this approach, concise explanations of confusion matrixes and quality metrics such as accuracy and precision are included to ensure that those readers without programming knowledge can better understand the text.

3. SEQUENCE DETECTION IN STATISTICAL PROCESS CONTROL

To conduct the experiment, specific software that is capable of calculating and simulating certain results for analysis is needed (such as Python or Excel). Without the appropriate methodology, however, these tools will be useless. To adequately conduct the experiment, Statistical Process Control (SPC) was utilized; this is a tool that is used for monitoring and controlling production processes in order to ensure that products meet specified quality standards. SPC relies on the collection and analysis of process data, enabling quick responses to deviations from the expected quality of a product.

The main quality issues that SPC helps address include process deviations, process instability, and variability control. Deviations can arise from factors such as changes in raw materials, machinery, or environmental conditions. SPC allows for the early detection of these deviations, enabling corrective actions to be taken before defective products leave the production line.

Process instability refers to fluctuations that can affect product quality. SPC enables the monitoring of these fluctuations and the identifications of their causes, allowing for appropriate adjustments to be made. Control of process variability is crucial, as excessive variations can lead to unpredictable quality outcomes. SPC helps control and reduce this variability, resulting in more stable product quality. SPC can also be used to monitor processes that are influenced by multiple variables; this allows for identifying significant variables and controlling them in order to ensure stable quality. In Figure 1, the use of one of the SPC methods is illustrated; namely, having seven or more points on one side of the nominal line without crossing it.

The triangles in Figure 1 highlight sequences of seven or more consecutive points that exceed the nominal line, thus signaling deviations from the expected process behavior. The first such deviation begins at Sample 5, which has a value of 6. Following this, the next six samples (from Sample 6 through Sample 11) also show values that are above the nominal line (with Sample 11 reaching a value of 7). As per the rule in place, when seven or more points consecutively exceed the nominal line, they are classified as being erroneous; this means that all of the measurements from Samples 5 through 11 are considered to be incorrect and are assigned a classification of 1. The sequence concludes at Sample 14 (with a value of 6); this is the last measurement above the nominal line, making the previous six measurements also part of the error pattern.

In a similar manner, the circles in Figure 1 represent sequences of seven or more consecutive points that are below the nominal line. These points signal a drop below the expected range (again, indicating process deviations). The first such deviation occurs at Sample 17 (which has a value of 1). The following six measurements (up to Sample 20, with a value of 2) remain below the nominal line, thus fulfilling the rule for classifying the entire sequence as being erroneous. The final erroneous measurement in this sequence occurs at Sample 24 (with a value of 0), as it completes a series of seven or more points that are consecutively below the nominal line. As a result, this group of measurements is also classified as 1 (for incorrect results) (Helmold, 2021).

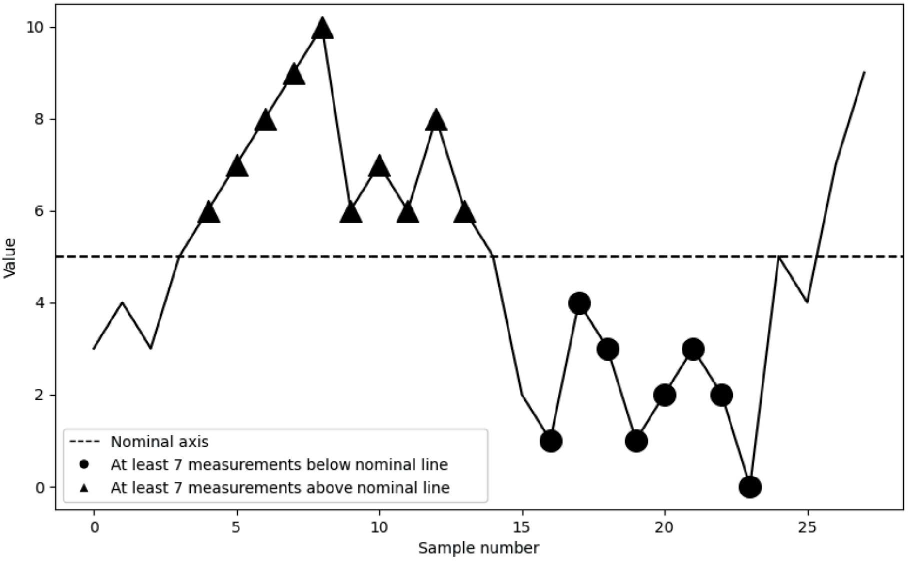

Fig. 1. *Graphical representation of method on chart*

4. MATERIAL AND METHODS

In the context of artificial intelligence and machine learning, classifiers are algorithms or models that assign objects or data to one of several classes or categories. Their main goal is to learn decision rules based on available training data to classify new unknown data. This is useful in many fields, such as image recognition, disease diagnosis, email spam filtering, sentiment analysis in social media, and more. Classifiers represent a crucial component in many artificial-intelligence-based systems. In the research on the presented problem, the following classifiers were utilized:

- AdaBoost (Adaptive Boosting) is a machine-learning algorithm that is primarily used for binary classification (but it can also be adapted for multi-class problems). The idea of AdaBoost involves sequentially training weak classifiers (known as "base classifiers") and assigning greater weights to misclassified examples in order to focus on any difficult-to-classify areas of the data.
- Decision Tree is a machine-learning algorithm that is used for both classification and regression problems. It divides a data set based on features to predict a target value.
- Gradient Boosting is a machine-learning algorithm that combines multiple weak models to create a strong predictive model. It operates sequentially, correcting the errors of previous models.
- Random Forest is an algorithm that is based on the concept of ensemble learning. It constructs multiple decision trees and combines them into one, making decisions by voting.
- Stochastic Gradient Descent (SGD) is an optimization algorithm that is used in machine learning – especially with large data sets.
- Support Vector Classifier (SVC) is a classification algorithm that uses support vector machines (SVM) to find the hyperplane that best separates data from different classes.
- Voting Classifier combines multiple different classification models and uses voting to make the final decision about the predicted class of a new sample (Cichosz, 2000).

4.1. Confusion matrix

In the conducted research, the confusion matrix was used to assess the quality (see Table 1). The confusion matrix is a tabular representation of the classification results, showing the number of correct and incorrect classifications for each class.

Table 1. *Confusion matrix*

	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

The above terms (especially in confusion matrices or performance metrics such as precision, recall, and F1 score) are commonly used to evaluate the performance of classification algorithms:

- TN (True Negatives): the number of observations that are correctly classified as negative. These are cases that are actually negative and have been classified as negative by the classifier.
- TP (True Positives): the number of observations that are correctly classified as positive. These are cases that are actually positive and have been classified as positive by the classifier.
- FP (False Positives): the number of observations that are incorrectly classified as positive. These are cases that are actually negative but have been incorrectly classified as positive by the classifier.
- FN (False Negatives): the number of observations that are incorrectly classified as negative. These are cases that are actually positive but have been incorrectly classified as negative by the classifier (Kurp, 2023).

4.2. Measures of classifier quality

By using the confusion matrix, the quality of the results was strengthened, thus expanding the scope of a further analysis. The prepared data served as modern methods for evaluating binary classifiers based on various indicators, thus allowing for a comprehensive analysis of their performance. Four key measures (precision [P], recall [R], average F1-score, and accuracy [A]) constitute a significant set of tools in this analysis.

Also known as a positive predictive value, precision focuses on how many predicted positive cases actually belong to the positive class. This is particularly useful in situations where the consequences of false positive errors are significant. In other words, precision measures how many of the instances that are classified as positive by the algorithm actually belong to the given class. Precision can take values from 0 to 1, where 1 indicates perfect precision and 0 indicates no precision. Precision is defined as the ratio of the number of correctly assigned instances of a given class to the sum of all of them:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Also known as sensitivity or the true positive rate, recall evaluates the classifier's ability to detect all actual positive cases. Recall values range from 0 to 1, where 0 indicates the model's inability to detect positive cases and 1 indicates the perfect ability to identify all positive cases. Recall is particularly important in situations where there are high risks that are associated with failing to identify positive cases:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The average F1-score (also known as the harmonic mean) is a measure that combines precision and recall, offering a single comprehensive value for an overall assessment of the balance between these two indicators. This approach is especially important in situations where classes are imbalanced, meaning that one class occurs much more frequently than the other. In practice, a high value of an average F1-score indicates an effective classifier that achieves a balance between minimizing false positives and false negatives:

$$F_1 score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

Accuracy is a general measure of classifier effectiveness that measures the ratio of the correct predictions (both positive and negative) to the total number of cases. High accuracy generally indicates the classifier's overall correct performance; however, it can be misleading in situations where classes are unevenly distributed (the classifier may be "accurate" by assigning most cases to the dominant class). In certain cases (especially when the costs of the errors differ between classes), other measures such as precision and recall may be more informative (Géron, 2020):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

5. RESULTS AND DISCUSSION

At the beginning, the data was collected; it consisted of 1040 measurements that were gathered over a specified period of time. Then, this data was prepared in Excel, where each measurement was assigned an x value (with x representing the values of seven consecutive measurements). The y value was calculated using the SPC method, where y was assigned a value of 1 (if the sequence of seven measurements stayed consistently above or below the nominal line) or 0 (if the results fluctuated above and below the nominal line). The data was then divided into 80% training data and 20% test data.

The next step involved preparing the code in Python. The data was also scaled to check for any impact on the final result for each classifier. NS represented "no scaling," while WS indicated "with scaling"; thus, 16 different classifier models were created. Libraries such as csv (for handling CSV files), NumPy (for numerical operations), AdaBoostClassifier from the scikit-learn library (for AdaBoost classification), and the confusion matrix and classification report (for evaluating the classifiers) were used. The code was developed in such a way that it could be used for each of the classifiers in an identical manner. The necessary libraries were imported, the CSV file that contained the data was opened and read, and each row was displayed. Next, the data was divided into two NumPy arrays: X, and Y. The content of Y, the length of the data, and the sum of the values were displayed to verify the correctness of the data-loading. The data was split into training and test sets, where the first 800 samples were used for the training and the rest for the testing.

Table 2. *Comparison of values for evaluating models using F1-score metric*

No.	Tested models	NS/WS	Results Dataset 1	Results Dataset 2	Results Dataset 3
1	AdaBoostClassifier	NS	0.550	0.715	0.635
2	DecisionTreeClassifier		0.520	0.860	0.935
3	GradientBoostingClassifier		0.550	0.895	0.895
4	**RandomForestClassifier**		**0.590**	**0.900**	**0.980**
5	SGDClassifier		0.445	0.450	0.435
6	SGDClassifierStratifiedShuffleSplit		0.170	0.450	0.435
7	SVC		0.175	0.835	0.930
8	VotingClassifier		0.500	0.840	0.710
9	AdaBoostClassifier	WS	0.550	0.745	0.635
10	DecisionTreeClassifier		0.515	0.875	0.960
11	GradientBoostingClassifier		0.530	0.895	0.895
12	RandomForestClassifier		0.590	0.890	0.980
13	SGDClassifier		0.170	0.450	0.295
14	SGDClassifierStratifiedShuffleSplit		0.210	0.450	0.490
15	SVC		0.175	0.835	0.930
16	VotingClassifier		0.500	0.840	0.725

Prediction was performed on the test set, and then the prediction results and actual labels were displayed in Table 2. The effectiveness of the classifiers clearly depended on the quality of the input data. The data set that was labeled Number 1 exhibited relatively poor results as compared to the other two data sets; this was mainly due to the insufficient number of errors that were identified in this data set. The classifiers did not have enough information to properly learn and assign incorrect results. In contrast, the classifiers for Datasets 2 and 3 achieved very good results, thus suggesting that the numbers of errors that were made and detected were sufficient for effective result assignment. It is worth noting that a slight improvement in results for those classifiers that were applied with data scaling could be observed; this may have been due to the smaller data range, which allowed the classifiers to better adapt to relevant classification information. The best results in all of the conducted simulations were achieved by the RandomForestClassifier. This classifier achieved values that were close to ideal in Datasets 2 and 3; and despite its lower effectiveness in this case, Dataset 1 still performed relatively well. However, it should be remembered that perfection cannot be expected in every experiment. The data that is used for experiments must contain sufficient numbers of errors for the classifiers to effectively learn the operating patterns. Under real production conditions, ensuring adequate quantities and qualities of the data can be crucial for the effective applications of classifiers. The next step after conducting these analyses could be to develop a computer application that utilizes trained classifiers to provide guidance to the controllers. Such a tool would not replace the work of the controllers entirely but could be a valuable aid in their daily work.

The main objective of this work was to prepare a training data set as well as train and evaluate the selected classifiers regarding their applicability for real-time control chart analysis. Both of the set objectives were achieved in this work, and the thesis that the known machine-learning classifiers could serve as useful methods for supporting the work of a quality controller was confirmed. During the research, it was proven that the classifiers could effectively identify the defective products on the production line in the analyzed enterprise. The use of simple classifiers along with Shewhart control charts for analyzing real production data opens up new perspectives in the field of quality control. It is also worth emphasizing that the conducted research contributes to the development of the field, where artificial intelligence becomes not only a tool that facilitates work but also an effective mechanism for eliminating human errors. From a practical perspective, defect-detection systems (though generally effective) cannot guarantee that a product is defective with absolute certainty; this is because they operate within predefined static frameworks. As such, these systems serve as valuable support tools for quality control inspectors but not as standalone solutions. However, these systems can surpass human capabilities in terms of measurement accuracy; this is due to their ability to maintain consistent stability in measurement processes.

The primary limitation when adopting these improvements lies in their cost-effectiveness – whether or not the implementation of such technology is truly necessary for a given production line. Additionally, it is crucial to consider that introducing these systems requires specialized software; this raises the question of whether the investment in such technology is justified based on the specific production needs.

The conducted research opens the door for new possibilities in the field of quality control automation, bringing benefits to statistical process control – both in terms of economic efficiency and improvements in accuracy. In detecting those patterns that signal process degradation (especially those that are defined by short seven-sample sequences), the choice of a classifier can significantly influence the predictive accuracy and model reliability. Among the basic machine-learning classifiers, random forest proved to be exceptionally effective, achieving close to 100% accuracy in identifying such degradation patterns in our experiments. This high level of performance arose from two main factors: first, the availability of a sufficient number of error samples provided a rich data set for the classifier from which to learn. When a data set includes enough instances of both normal and degraded process conditions, a random forest model can distinguish patterns that signify the onset of degradation with high precision. Each tree within the random forest ensemble captures nuances in these patterns, thus enabling robust recognition across varied instances. Second, the Random Forest algorithm's inherent structure is well-suited for this type of classification task. By creating multiple decision trees and averaging their outputs, Random Forest effectively captures complex relationships and dependencies within the data; this is especially valuable in process degradation, where subtle changes across samples might otherwise go undetected. The ensemble approach reduces the risk of overfitting on specific anomalies, leading to more generalizable insights into the degradation process. Moreover, Random Forest's ability to handle noise and outliers ensures that sporadic measurement errors do not overshadow true degradation signals. When paired with interpretability in feature importance, this resilience allows researchers and practitioners to better understand which variables most strongly indicate process decline.

6. CONCLUSIONS

All of the proposals that were suggested by this research were realized: the classifiers were well-trained, and it was found that their performance was highly satisfactory when recognizing any data anomalies that were related to process degradation; this is evidence that the implemented machine-learning models can deliver reliable support for finding defects during the production of wheel rims, thus serving as yet another tool for quality controllers.

With such promising results, the obvious next step will be to put this into practice in a real production setting. The practical and scalable application of the trained classification algorithms can be integrated into an application that is specifically designed for this purpose. Such an application would feature the real-time monitoring of production data, thus alerting operators and quality inspectors to potential defects at the times of such occurrences. By incorporating such a model-driven approach directly onto a production line, manufacturers can restrict their reliance on manual inspections to only higher detection accuracy, thereby yielding significant dividends in product quality and operational efficiency.

Further refinement could also be performed on the models in this pilot deployment itself; again, this would provide continuous improvement in anomaly-detection accuracy and adaptability for various production scenarios. This application of machine learning can become a very good avenue for bringing changes in the methods of quality control by applying data-driven insight for improving the reliability and streamlining of manufacturing.

REFERENCES

Alfardus A. & Rawat D.B. (2023). Detection of cyberattacks in controller area network (CAN) bus system using a hybrid ML approach. In: *2023 IEEE 12th International Conference on Cloud Networking (CloudNet)*, pp. 53–60. **DOI**: https://doi.org/10.1109/CloudNet59005.2023.10490074.

Benzaza L., Alfathi N. & Lyhyaoui A. (2023). Selecting key product characteristics to improve the QMS in automotive sector. In: J. Kacprzyk, M. Ezziyyani, V.E. Balas (Eds.), International Conference on Advanced Intelligent Systems for Sustainable Development. *Lecture Notes in Networks and Systems*, **714**, pp. 523–531. **DOI**: https://doi.org/10.1007/978-3-031-35245-4_48.

Cichosz P. (2000). *Systemy uczące się*. Warszawa: WNT.

Cui J., Liu W., Zhang Y., Gao C., Lu Z., Li M. & Wang F. (2021). A novel method for predicting delamination of carbon fiber reinforced plastic (CFRP) based on multi-sensor data. *Mechanical Systems and Signal Processing*, **157**, art. no. 107708. **DOI**: https://doi.org/10.1016/j.ymssp.2021.107708.

Dettú F., Formentin & Savaresi S.M. (2024). The twin-in-the-loop approach for vehicle dynamics control. *IEEE/ASME Transactions on Mechatronics*, **29**(2), pp. 1217–1228. **DOI**: https://doi.org/10.1109/TMECH.2023.3292503.

Escobar C.A., Chakraborty D., Arinez J. & Morales-Menendez R. (2021). Augmentation of body-in-white dimensional quality systems through artificial intelligence, In: *2021 IEEE International Conference on Big Data (Big Data)*, pp. 1611–1618. **DOI:** https://doi.org/10.1109/BigData52589.2021.9671610.

Fang X.-Y., Gong J.-E., Zhang F., Zhang H.-N. & Wu J.-H. (2023). Machine learning assisted materials design of high-speed railway wheel with better fatigue performance. *Engineering Fracture Mechanics*, **292**, art. no. 109586. **DOI**: https://doi.org/10.1016/j.engfracmech.2023.109586.

Géron A. (2020). *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*. Wydanie drugie. Gliwice: Helion.

Gong W., Zhao W., Wang X., Li Z., Wang Y., Zhao X., Wang Q., Wang Y., Wang L. & Chen Q. (2023). Machine learning for laser micro/nano manufacturing: Applications and prospects. *Chinese Journal of Lasers*, **50**(20), art. no. 2000001. **DOI**: https://doi.org/10.3788/CJL230827.

Habibullah K.M., Heyn H.-M., Gay G., Horkoff J., Knauss E., Borg M., Knauss A., Sivencrona H. & Li P.J. (2024). Requirements and software engineering for automotive perception systems: an interview study. *Requirements Engineering*, **29**(1), pp. 25–48. **DOI**: https://doi.org/10.1007/s00766-023-00410-1.

Helmold M. (2021). Statistical, quality and resource management tools. *Management for Professionals*, Part F473, pp. 71–79. **DOI**: https://doi.org/10.1007/978-3-030-77661-9_7.

Hofmann J., Li Z., Taphorn K., Herzen J. & Wudy K. (2024). Porosity prediction in laser-based powder bed fusion of polyamide 12 using infrared thermography and machine learning. *Additive Manufacturing*, **85**, art. no. 104176. **DOI**: https://doi.org/10.1016/j.addma.2024.104176.

Hwarng H.B. (1992). Pattern recognition on Shewhart control charts using a neural network approach. Arizona State University [PhD Thesis].

Kalyanasundaram P., Kareti V., Sambranikar M., Ss N.K. & Ranadive P. (2018). Practical approaches for detecting DoS attacks on CAN network. *SAE Technical Papers*, paper no. 0019. **DOI**: https://doi.org/10.4271/2018-01-0019.

Kidmose B. & Meng W. (2024). can-sleuth: Investigating and evaluating automotive intrusion detection datasets. In: *2024 European Interdisciplinary Cybersecurity Conference*, pp. 19–28. **DOI**: https://doi.org/10.1145/3655693.3655696.

Krummenacher G., Ong C.S., Koller S., Kobayashi S. & Buhmann J.M. (2017). Wheel defect detection with machine learning. *IEEE Transactions on Intelligent Transportation Systems*, **19**(4), pp. 1176–1187. **DOI**: https://doi.org/10.1109/TITS.2017.2720721.

Kurp F. (2023). *Sztuczna inteligencja od podstaw*. Gliwice: Helion.

Lampe B. & Meng W. (2023). can-train-and-test: A new CAN intrusion detection dataset. In: *2023 IEEE 98th Vehicular Technology Conference*. **DOI**: https://doi.org/10.1109/VTC2023-Fall60731.2023.10333756.

Lampe B. & Meng W. (2024). can-train-and-test: A curated CAN dataset for automotive intrusion detection. *Computers and Security*, **140**, art. no. 103777. **DOI**: https://doi.org/10.1016/j.cose.2024.103777.

Lee J.-H., Lee J.-H., Yun K.-S., Bae H.B., Kim S.Y., Jeong J.-H. & Kim J.-P. (2023). A study on wheel member condition recognition using machine learning (support vector machine). *Sensors*, **23**(20), art. no. 8455. **DOI**: https://doi.org/10.3390/s23208455.

Lestyán S., Acs G., Biczók G. & Szalay, Z. (2019). Extracting vehicle sensor signals from CAN logs for driver re-identification. In: *Proceedings of the 5th International Conference on Information Systems Security and Privacy*, pp. 136–145. **DOI**: https://doi.org/10.5220/0007389501360145.

Lokman S.-F., Othman A.T. & Abu-Bakar M.-H. (2019). Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review. *EURASIP Journal on Wireless Communications and Networking*, **2019**(1). **DOI**: https://doi.org/10.1186/s13638-019-1484-3.

Malindzakova M., Čulková K. & Trpčevská J. (2023). Shewhart control charts implementation for quality and production management. *Processes*, **11**(4), art. no. 1246. **DOI**: https://doi.org/10.3390/pr11041246.

Minawi O., Whelan J., Almehmadi A. & El-Khatib K. (2020). Machine learning-based intrusion detection system for controller area networks. In: *DIVANet '20: Proceedings of the 10th ACM Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, pp. 41–47. **DOI**: https://doi.org/10.1145/3416014.3424581.

Mjimer I., Aoula E.-S. & Achouyab E.H. (2023). Contribution of machine learning in continuous improvement processes. *Journal of Quality in Maintenance Engineering*, **29**(2), pp. 553–567. **DOI**: https://doi.org/10.1108/JQME-03-2022-0019.

Moser M., Kipping S., Higuchi K. & Hirayama H. (2021). Machine-learned emission model for diesel exhaust on-board diagnostics and data flow processor as enabler. *SAE Technical Papers*, paper no. 5108. **DOI**: https://doi.org/10.4271/2021-01-5108.

Oh Y., Ransikarbum K., Busogi M., Kwon D. & Kim N. (2019). Adaptive SVM-based real-time quality assessment for primer-sealer dispensing process of sunroof assembly line. *Reliability Engineering and System Safety*, 184, pp. 202–212. **DOI**: https://doi.org/10.1016/j.ress.2018.03.020.

Park S.-H. & Baek Y. (2023). CADS: Cyber-attack detection system using decision tree for automotive cyber-physical systems. In: *2023 International Conference on Electrical, Communication and Computer Engineering*. **DOI**: https://doi.org/10.1109/ICECCE61019.2023.10442984.

Paturi U.M.R., Palakurthy S.T., Cheruku S., Vidhya Darshini B. & Reddy N.S. (2023). Role of machine learning in additive manufacturing of titanium alloys: a review. *Archives of Computational Methods in Engineering*, **30**(8), pp. 5053–5069. **DOI**: https://doi.org/10.1007/s11831-023-09969-y.

Rameshkumar K., Mouli D. & Shivith K. (2021). Machine learning models for predicting grinding wheel conditions using acoustic emission features. *SAE International Journal of Materials and Manufacturing*, **14**(4), pp. 387–406. **DOI**: https://doi.org/10.4271/05-14-04-0026.

Robinson J.C., Sherman K., Price D.W. & Rathert J. (2020). Inline Part Average Testing (I-PAT) for automotive die reliability. In: O. Adan, J.C. Robinson (Eds.), *Metrology, Inspection, and Process Control for Microlitography XXXIV. Proceedings of SPIE*, **11325**, art. no. 113250D. **DOI**: https://doi.org/10.1117/12.2551539.

Sharmin S., Mansor H., Abdul Kadir A.F. & Aziz N.A. (2022). Using streaming data algorithm for intrusion detection on the vehicular controller area network. *Communications in Computer and Information Science*, 1557, pp. 131–144. **DOI**: https://doi.org/10.1007/978-981-19-0468-4_10.

Shewhart M. (1992). Interpreting statistical process control (SPC) charts using machine learning and expert system techniques. In: *Proceedings of the IEEE 1992 National Aerospace and Electronics Conference@m_ NAECON* 1992, Dayton, OH, USA, **3**, pp. 1001–1006. **DOI**: https://doi.org/10.1109/NAECON.1992.220472.

Shewhart W. (1926). Quality control charts. *Bell System Technical Journal*, **5**(4), pp. 593–603. **DOI**: https://doi.org/10.1002/j.1538-7305.1926.tb00125.x.

Shi W., Alawieh M.B., Li X., Yu H., Arechiga N. Tomatsu N. (2016). Efficient statistical validation of machine learning systems for autonomous driving. In: *2016 IEEE/ACM International Conference on Computer-Aided Design,* 7–10 November 2016, Austin, TX, USA. **DOI**: https://doi.org/10.1145/2966986.2980077.

Staf E. & McKelvey T. (2018). Introducing compressed mixture models for predicting long-lasting brake events. *IFAC-PapersOnLine*, **51**(31), pp. 840–845. **DOI**: https://doi.org/10.1016/j.ifacol.2018.10.115.

Staněk R., Kerepecký T., Novozámský A., Šroubek F., Zitová B. & Flusser J. (2023). Real-time wheel detection and rim classification in automotive production. In: *2023 IEEE International Conference on Image Processing (ICIP)*, pp. 1410–1414.

Tran P.H., Ahmadi Nadi A., Nguyen T.H., Tran K.D. & Tran K.P. (2022). Application of machine learning in statistical process control charts: A survey and perspective. In: K.P. Tran (Ed.) *Control Charts and Machine Learning for Anomaly Detection in Manufacturing. Springer Series in Reliability Engineering*, pp. 7–42. **DOI**: https://doi.org/10.1007/978-3-030-83819-5_2.

True J., Xi C., Jessurun N., Ahi K. & Asadizanjani N. (2021). Review of THz-based semiconductor assurance. *Optical Engineering*, **60**(6), art. no. 060901. **DOI**: https://doi.org/10.1117/1.OE.60.6.060901.

Tsenev V. & Ivanova M. (2022). Statistical and machine learning approach for evaluation of control systems for automatic production lines. *Bulletin of Electrical Engineering and Informatics*, **11**(5), pp. 2527–2536. **DOI**: https://doi.org/10.11591/eei.v11i5.3664.

Xu X., Wang G., Xuan S., Shan Y., Yang H. & Yao X. (2024). Digital twin and cross-scale mechanical interaction for fabric rubber composites considering model uncertainties. *Composites Science and Technology*, **248**, art. no. 110431. **DOI**: https://doi.org/10.1016/j.compscitech.2024.110431.

Zhou A., Liu X., Liu Q., Zhu G. & Zhang S. (2023). Progress of in-process monitoring techniques for selective laser melting. *China Surface Engineering*, **36**(4), pp. 36–50. **DOI**: https://doi.org/10.11933/j.issn.1007-9289.20221212001.