Real-Time Machine Learning Automation Applied to Failure Prediction in Automakers Supplier Manufacturing System

Arthur Beltrame Canciglierie, Tainá da Rocha, Anderson L. Szejka, Leandro Dos Santos Coelho, Osiris Junior Canciglieri Canciglieri

Real-time machine learning automation applied to failure prediction in automakers supplier manufacturing system

Arthur Beltrame Canciglierie[1,2], Tainá da Rocha[1,2], Anderson L. Szejka[1], Leandro dos Santos Coelho[1,3] and Osiris Canciglieri Junior[1]

[1] Industrial and Systems Engineering Graduate Program (PPGEPS), Pontifical Catholic University of Parana (PUCPR), Curitiba, Brazil
[2] Robert Bosch, Ind. Plant - Curitiba, Brazil
[3] Department of Electrical Engineering (PPGEE), Federal University of Parana (UFPR), Polytechnic Center, Curitiba, Brazil
{anderson.szejka,osiris.canciglieri}@pucpr.br
arthur.canciglieri2@br.bosch.com

Abstract. The Industry 4.0 smart factories allow both optimization and integration of internal processes, utilizing the predictability of failure elements/components in a manufacturing process to prevent reprovals at the end of the process for quality control. The Supervised Machine Learning (ML) methods could be useful to detect anomalies and gain even more value throughout the entire supply chain. The ML approaches face barriers since it demands a changing in the production plant mindset to a more digital production and in the organization's structure for a more advanced data security. The paper aims to propose a smart inconsistency and fail prediction system for manufacturing systems of an automakers supplier assembly process based on the applications of ML techniques. The data provided for the training showed significant deviations and non-linearity allied to only 5 attributes as input variables, which is considered a small number of features for similar problems in the literature. The trained model was then applied to the assembly line with unobserved data of new products, with its result compared with similar previous productions. The results of the tests showed that the proposed stacking model lessens the possibility of rework in the next stages of assembly and creates a more precise process control for the supervisor. The implementation's results pointed out the potential of the stacking model proposed to be a useful tool in the context of Industry 4.0 since the reductions mean greater availability of production time and lower costs with quality control.

Keywords: Industry 4.0; Machine Learning; Quality Control; Failure Prediction; Smart manufacturing; Automaker Supplier.

1 Introduction

Nowadays, the increased use of traceability and connectivity technologies by the automotive industry and its suppliers has made it possible to collect and study manufacturing data focusing on process control and its optimization. There is a collective need from global industries to digitize products and processes in the context of the fourth industrial revolution, thus guaranteeing product quality and digital

excellence [1]. The heart of Industry 4.0 is the rise of smart factories, which includes smart networking between industry units, processes, and operations. It focuses on flexibility and interoperability of processes, integrating the customer's and supplier's information and requirements with the adoption of new business models [2].
requirements with the adoption of new business models [2].

Implementing the new paradigm of Industry 4.0 follows three characteristics: i) The horizontal integration across the entire value network; ii) End-to-end engineering across the entire product lifecycle management, and iii) Vertical integration and networked manufacturing systems [3]. The vertical integration and networked manufacturing systems allow both optimization and integration of internal processes, which leads to a higher understanding of the whole manufacturing process and quality control of the product features.

For quality control, the predictability of failure elements/components in a manufacturing process can help prevent reprovals at the end of the process, and thus ensuring higher quality for the end-user. In this way, AI techniques come to assist the process of failure prediction to make the whole process more robust and less subjective. It integrates technologies/machinery, workers, and information, creating an agile and responsive process, and process control systems [4][5].

The paper aims to propose a smart inconsistency and fail prediction system for manufacturing systems based on the applications of ML techniques. The validation of the system is in a real environment of assembly of parts for the automotive sector to carry out the control quality, generating a powerful tool for the production line' supervisors.

2 Related Works

Over recent years, industries and research centers have made significant efforts and progress into implementing smart manufacturing and smart industries techniques in the Industry 4.0 era. According to [6][7][8][9], in the last decade, the manufacturing industries are having more demand for customized products without losing quality and production indicator and remaining competitive. The companies need to analyze and remodel their processes in this new context of production.

Industries experience a complex manufacturing environment based on multicriteria decisions, and the forecast for manufacturing processes influence the quality and efficiency indicators significantly [10][11]. Machine Learning core technologies work well with the complexes problems generated inside de industry and can be applied to reduce fixed costs, rework indicators, and improve key production indicators and speed [12].

The use of artificial intelligence and machine learning in the quality control of companies is becoming increasingly frequent. In the production floor, one of the most challenging problem for quality and process control is the product measurement variability in processes like assembly and machining and can occur in industries such as the automotive, since it can be influenced by various factors [13][14]. According to [15], Machine Learning applications can help Lean Manufacturing lines to have more quality and efficiency, especially in the automotive supplier industry, where the connectivity from machines and production lines data is vital for efficient production.

The application of machine learning models can be facilitated by a smart manufacturing environment since they refer to the new paradigm of production where manufacturing machines are fully connected, monitored, and controlled via smart systems, improving productivity, sustainability and reducing costs [16][17][18].

Finally, adapting machine learning concepts to a domain of knowledge within a real industrial process can present several challenges. One of the challenges is that ML applications cannot impact machine cycle times, and often must be optimized to not significantly impact process times. In addition, the processing power of machines on the production line is generally limited to the production processes themselves, requiring the design of shared processing in the cloud to achieve robust ML model training and satisfactory results [19]. Therefore, the quality control in real time in automotive supplier processes by ML models are still very limited to the complexity and training time required by the chosen models, being often simplified to obtain results that do not negatively impact other production indicators.

3 Problem Statement

This case study was carried out between two workstations located in an assembly line of an Automaker's Supplier Plant, and all names of variables and parts were adapted due to the Company's data security policies. The assembly process focused in this paper is shown in Figure 1.

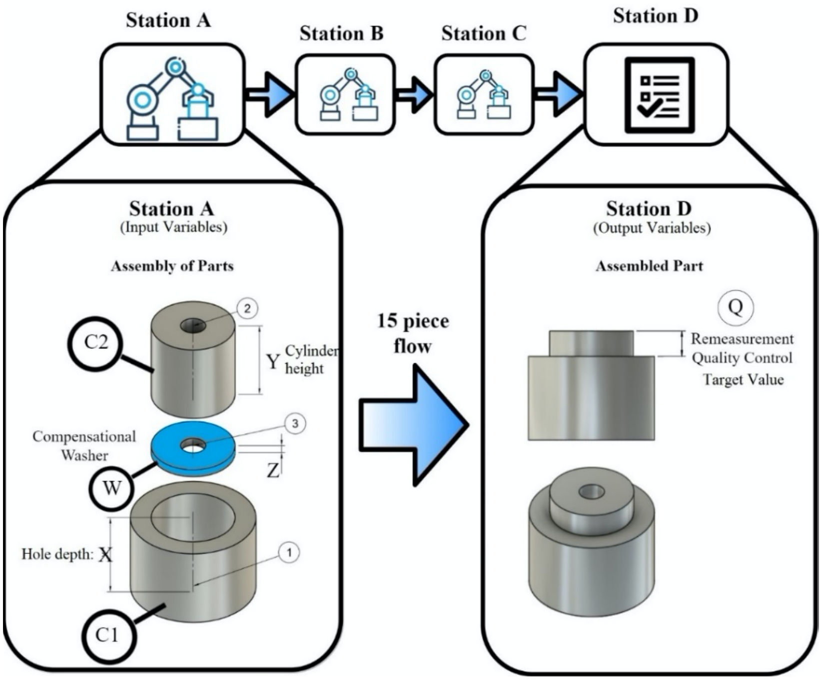

Fig. 1. Diagram of the automaker's supplier assembly process.

In this assembly process, the first station (Figure 1 - station "A") receives two cylindrical parts (Figure 1 - Detail "C1" and "C2") from previous machining processes and measures the internal height of the first part (Figure 1 – dimension "X") and the external height from the second part ("Y" of Figure 1). With these both measures, it calculates a compensational washer (Figure 1 - Detail "W") using a Setup variable from the configuration of the product, and an Operator adjusts to adequate the manufacturing assembly into product requirements. The Setup Variable comes from customers' requirements, and the Operators adjustment is a readjustment to deal with deviations and inaccuracies, whether they are from previous machining processes or for general deviations from the assembly line.

From this, the assembled part is processed by operations "B" and "C" and performs a quality check at station "D". In this station (Figure 1 - Station "D"), a remeasurement process, more accurate, checks whether the product complies with the specifications and whether the assembly at station "A" was correct (Figure 1 - Detail "Q") and generates the Target Value in this study.

All variables are collected through sensors, which generate a text file at stations "A" (input) and "D" (target) containing all the information collected by the machines of each part. Every new part that passes through the stations then generates a text file being identified by a unique code of the part and that contains its variables which is transferred to a backup server to carry out quality and traceability analyses.

4 Conceptual Proposal

The proposed solution is to use a machine learning architecture for quality control that reads the text file generated by the machine before being sent to the backup server, since the creation of this file is instantaneous and presents all the available variables, and then makes prediction and an automatic adjustment at station "A" piece-by-piece based in historic production data. For the training of the ML models and analysis of variable behavior, it was used the historical data stored on the backup server of the production line, in which there is the text files with the individual data of all the parts that were processed in the last 3 years. This process can be seen in Figure 2.

Five ML models were chosen for the architecture: i) *K*-Nearest Neighbor (*KNN)*, ii) Random Forest (*RF)*, iii) Gradient Boosting Machine (*GBM)*, iv) Light GBM and v) XGBOOST. These ML models are classified into 3 groups of learning: association, interaction, and ensemble learning. These ML methods were chosen for their effectiveness in other similar case studies in the literature [20][21] and for their simplified implementation on an assembly line since they are white-box methods, and it is possible to understand the importance of the variables.

The research focus is on the application of a machine learning model architecture and on overcoming the challenges faced in real-time implementation on a production line, where there are several factors to be considered. Therefore, models already used widely in the literature and with easy access to programming libraries were chosen, since this research does not propose a comparison of performance between individual models, being used a range of different algorithms for better overall performance. Finally, several algorithms like those proposed were discarded due to the difficulty of implementing them on computers within the production line, where access to information and programming libraries is restricted by secrecy and security.

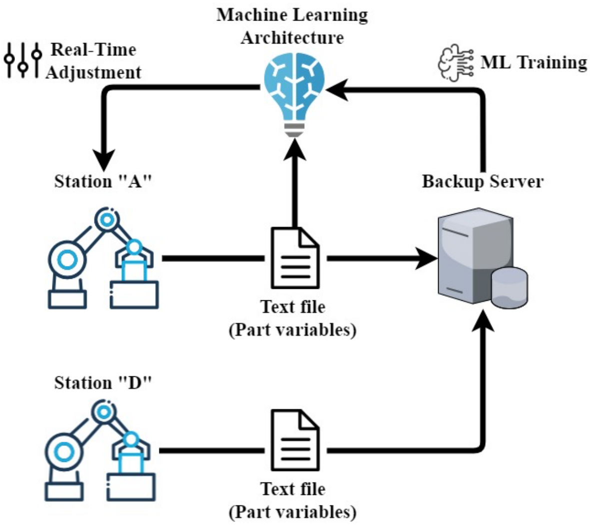

Fig. 2. Data Flow Diagram

After choosing the models to be used, a Staking structure of the machine learning models was designed, where KNN, RF, GBM and light GBM would be used as primary estimators since they present different types of learning methods between themselves and can gain knowledge in a different and optimized way, and XGBOOST as a meta-estimator. The XGBOOST model was chosen as a meta-estimator for its performance in similar benchmarks problems [22][23] and the full use of available hardware.

The models were trained using the six variables (variable "*X*", "*Y*", "*Z*", "Setup variable", "Operator Adjustment" and "Remeasurement-Target Value"), mentioned in Figure 1. All estimators were trained with historical production data of 120.000 product database and the grid-search of the hyperparameter was done for the tuning and optimization of each single model.

5 Results

The conceptual proposal provided the necessary knowledge for the functional test on the assembly line. The models used hyperparameters with the best result of the grid-search, and all attributes for better accuracy. The functional test was carried out on the assembly line during regular operation, indicating the adjustment of the washer in real-time by the created stacking model.

Moreover, to test the flexibility and robustness of the proposed model, tests were carried out on 10 different configurations of products with 7400 test parts supplied to different customers. Each consumer company requires different product configurations, such as different levels of quality, product requirements and geometries with different characteristics.

The rework indicators collected at station D were compared using the machine learning architecture proposed with productions of the same configuration of product without its use (normal production), and thus analyzed whether it was possible to reduce rework. It was not possible to evaluate the usage of the proposed model for the same parts since after the assembly processes at station A and subsequent B and C, the assembled part can no longer be disassembled, making it impossible to compare the production with the proposed architecture and the normal production of the same parts. The results obtained in the tests are shown in Table 1, considering that the different configurations mean distinct product specification.

Table 1. Test results of the assembly line.

Product Configuration	Amount Tested	Rework Reduction
Configuration 1	508	0.54%
Configuration 2	726	2.79%
Configuration 3	209	4.65%
Configuration 4	150	2.60%
Configuration 5	719	5.28%
Configuration 6	258	6.03%
Configuration 7	96	8.57%
Configuration 8	584	-2.12%
Configuration 9	2139	1.20%
Configuration 10	1992	1.90%
Total	**7381**	**3.14%**

Table 1 shows that the use of the proposed staking model had a positive result in reducing rework in 9 configurations of products, and only 1 had a negative outcome. The only negative result was motivated by the developed architecture not being able to adapt to configuration 8 and thus accurately predict new parts in real time, in addition to reasons of temperature variation and dirt that can influence the system outcome.

In general, there was a 3.14% reduction in rework that represents a direct decrease in quality control costs and a more stable assembly line. Thus, the tests indicate that the proposed stacking model lessens the possibility of rework and creates a more precise process control for the supervisor.

6 Conclusion

In this article, we proposed and analyzed the process and results of implementing a machine learning stacking model on an automotive parts assembly line. The data provided for the training showed significant deviations and non-linearity allied to only 5 attributes as input variables, which is considered a small number of features for similar problems in the literature. An extensive grid search was performed with cross-

validation to obtain the best hyperparameters for the problem and training the proposed model with a dataset of 2 months of production data.
The trained model was then applied to the assembly line with unobserved data and new products, and its result compared with similar previous productions. Around 7400 parts were tested, obtaining an average rework reduction of 3.14%.

The proposed model is highly robust and can be used in several similar problems, in which an adjustment is made, and a quality control is carried out based on this adjustment. The idea is that the quality control should measure the target value of all parts, and thus create machine learning models to understand the relationship between the input variables (until the part is adjusted) and the quality control variable. The proposed model was tested for mechanical adjustments made with washers, but it is possible to be used for mechanical adjustments of different types, and thus be used in almost any assembly process in the automotive industry as well as various processes.

The next steps for this research are to increase the robustness of the proposed model and increase its accuracy with different and newer machine learning models. Besides that, the proposed model must be tested on production and assembly lines of different products with other characteristics to test its flexibility and robustness with different problems.

References

1. Lu, Yuqian, Chao Liu, Kevin I-Kai Wang, Huiyue Huang, e Xun Xu. 2020. Digital Twin-Driven Smart Manufacturing: Connotation, Reference Model, Applications and Research Issues. Robotics and Computer-Integrated Manufacturing. 2019 Feb;101837.
2. Kamble SS, Gunasekaran A, Sharma R. Analysis of the driving and dependence power of barriers to adopt industry 4.0 in Indian manufacturing industry. Computers in Industry. 2018 Oct 1; 101:107-19.
3. Peres, R.S., Rocha, A.D., Leitao, P. and Barata, J., 2018. IDARTS–Towards intelligent data analysis and real-time supervision for industry 4.0. Computers in Industry, 101, pp.138-146.
4. Raj A, Dwivedi G, Sharma A, de Sousa Jabbour AB, Rajak S. Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective. International Journal of Production Economics. 2020 Jun 1; 224:107546.
5. Szejka, A.L., Canciglieri Jr., O., Loures, E.R., Panetto, H., Aubry, A.: Requirements interoperability method to support integrated product development. In: 45th Computers and Industrial Engineering, vol. 147, pp. 1–8. Metz (2015).
6. Bauters, K., Cottyn, J., Claeys, D., Slembrouck, M., Veelaert, P. and Van Landeghem, H., 2018. Automated work cycle classification and performance measurement for manual workstations. Robotics and Computer-Integrated Manufacturing, 51, pp.139-157.
7. Tvenge, N. and Martinsen, K., 2018. Integration of digital learning in industry 4.0. Procedia manufacturing, 23, pp.261-266.
8. Gattullo, M., Scurati, G.W., Fiorentino, M., Uva, A.E., Ferrise, F. and Bordegoni, M., 2019. Towards augmented reality manuals for industry 4.0: A methodology. Robotics and Computer-Integrated Manufacturing, 56, pp.276-286.
9. Adamczyk, B.S., Szejka, A.L., Canciglieri, O.: Knowledge-based expert system to support the semantic interoperability in smart manufacturing. Computers in Industry. 115, 103161 (2020). https://doi.org/10.1016/j.compind.2019.103161.
10. Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W. and Monostori, L., 2018. Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. *Procedia CIRP*, 72, pp.1051-1056.

11. Sharp, M., Ak, R. and Hedberg Jr, T., 2018. A survey of the advancing use and development of machine learning in smart manufacturing. *Journal of manufacturing systems*, 48, pp.170-179.
12. Murphy, R., Newell, A., Hargaden, V. and Papakostas, N., 2019. Machine learning technologies for order flowtime estimation in manufacturing systems. *Procedia CIRP*, 81, pp.701-706.
13. Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S. and Wrobel, S., 2019. A review of machine learning for the optimisation of production processes. *The International Journal of Advanced Manufacturing Technology*, 104(5-8), pp.1889-1902.
14. Carvalho, T.P., Soares, F.A., Vita, R., Francisco, R.D.P., Basto, J.P. and Alcalá, S.G., 2019. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, p.106024.
15. Şenkayas, H. and Gürsoy, Ö., 2018. Industry 4.0 Applications and Digitalization of Lean Production Lines. *The Annals of the University of Oradea*, p.124.
16. Wang, J., Ma, Y., Zhang, L., Gao, R.X. and Wu, D., 2018. Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems*, 48, pp.144-156.
17. Scurati, Giulia Wally, Michele Gattullo, Michele Fiorentino, Francesco Ferrise, Monica Bordegoni, and Antonio Emmanuele Uva. Converting maintenance actions into standard symbols for Augmented Reality applications in Industry 4.0. *Computers in Industry* 98 (2018): 68-79.
18. O'Donovan, P., Gallagher, C., Leahy, K. and O'Sullivan, D.T., 2019. A comparison of fog and cloud computing cyber-physical interfaces for Industry 4.0 real-time embedded machine learning engineering applications. *Computers in Industry*, 110, pp.12-35.
19. Oh, Y., Busogi, M., Ransikarbum, K., Shin, D., Kwon, D. and Kim, N., 2019. Real-time quality monitoring and control system using an integrated cost-effective support vector machine. *Journal of Mechanical Science and Technology*, 33(12), pp.6009-6020.
20. Peres, R.S., Barata, J., Leitao, P. and Garcia, G., 2019. Multistage quality control using machine learning in the automotive industry. *IEEE Access*, 7, pp.79908-79916.
21. Jia, Y., Jin, S., Savi, P., Gao, Y., Tang, J., Chen, Y. and Li, W., 2019. GNSS-R soil moisture retrieval based on a XGboost machine learning aided method: *Performance and validation. Remote Sensing*, 11(14), p.1655.
22. Park, S., Moon, J., Jung, S., Rho, S., Baik, S.W. and Hwang, E., 2020. A two-stage industrial load forecasting scheme for day-ahead combined cooling, heating and power scheduling. *Energies*, *13*(2), p.443.
23. Gao, K., Chen, H., Zhang, X., Ren, X., Chen, J. and Chen, X., 2019. A novel material removal prediction method based on acoustic sensing and ensemble XGBoost learning algorithm for robotic belt grinding of Inconel 718. *The International Journal of Advanced Manufacturing Technology*, *105*(1-4), pp.217-232.