Proactive Error Prevention in Manufacturing Based on an Adaptable Machine Learning Environment

Holger Ziekow[1], Ulf Schreier[1], Alaa Saleh[1], Christof Rudolph[2], Katharina Ketterer[2], Daniel Grozinger[2], Alexander Gerling[1]

[1] Furtwangen University
{holger.ziekow, ulf.schreier, alaa.saleh, gera}@hs-furtwangen.de
[2] Sick AG
{Christof.Rudolph, Katharina.Ketterer, Daniel.Grozinger}@sick.de

Abstract. This paper gives an introduction into a concept for proactive error prevention in manufacturing. The challenges for research include the accuracy of predictions, the automated adaptation of prediction models to heterogeneous and varied processes, and the management of prediction models in changing processes. The solution is a combination of machine learning technology, big data approaches and knowledge modelling.

Keywords: Proactive Error Preventing, Machine Learning, Manufacturing.

1 Introduction

Technological developments such as the Internet of Things and Industry 4.0 hold the potential for innovation in production processes. The costs are going to be lowered and quality of the production processes will be increased [1] [2]. These developments coincide with significant advances in the processing of mass data and machine learning methods [3]. Currently many tools with different merits and weaknesses are available for the processing of mass data [4], which can be combined with concepts such as the lambda architecture [5]. We are at a point where the current technologies are ready to capture high-resolution production data and quickly evaluate those with complex methods. The challenge for production companies is to exploit the technological potential for specific innovations in production processes and to scale the use of technology in practice.

2 Adaptable Prediction Models

Production companies especially in high-wage countries face the challenge of producing high-quality products at competitive prices. With the project PREFERML we want to develop prediction methods which can predict and proactively avoid production errors in current and future processes. These methods, promise to greatly production costs and waste can be greatly reduced. For practical use, especially in heterogeneous and highly dynamic production processes, it is important to increase the degree of automation in the creation and use of prediction methods. Currently, the creation and maintenance of prediction models is highly demanding in terms of manual effort and time of human experts. Production companies carry out many production processes which are changing over time. To be able to create and use numerous models for different processes, it is necessary to increase the degree of automation in modelling and maintenance. The aim of this project is to achieve this and thereby to enabling the widespread use of proactive error prevention in production. The interdisciplinary combination of expertise in data analysis, software architecture of business applications and domain knowledge from production is indispensable for the investigation of practicable prognosis models in the production process.

3 Related Work

Using production data and big data technologies for prediction in production is the subject of current research [6] [7]. The challenges for research include (1) the accuracy of predictions, (2) the automated adaptation of prediction models to heterogeneous and varied processes, and (3) the management of prediction models in changing processes.

The prediction accuracy is significantly influenced by the appropriate selection and pre-processing of the data which use feature extraction and selection, as well as the parameterisation and selection of the prediction models [8] [9]. This often happens with the help of domain knowledge and as part of a manual process. However, there are approaches in research to automate these processes and the individual adaptation of models [9] [10] [11]. Recent work has shown the use of semantic models for the modelling of background knowledge in industrial applications [11] [12]. This work [13] already considers their application for the automated selection of model building features for model building. Here it is necessary to tailor and expand such approaches to the domain of error prediction in production processes.

Regarding the administration of analytical models and their scalable use, a research gap has been identified and initial approaches developed [14] [15]. A key challenge is to identify and address changes in model quality and, if necessary, causal changes in the deployment environment. This requires continuous monitoring of performance parameters of the models and of relevant conditions of the deployment environment. Technologies of complex event processing are suitable as a basis to detect certain situations in the available sensor data streams and to initiate necessary reactions [16] [17]. Previous works in the field of complex event processing for the evaluation of data streams in the production already exist [18] [19]. However, the definition of the relevant situations is not set by Complex Event Processing but is to be described for the respective task via event patterns. For monitoring of predictive models, it is necessary to adjust the use of complex event processing technology and to automate the definition of relevant event patterns.

4 Solution Approach

The proposed solution is a combination of machine learning technology, big data approaches and knowledge modelling. The combination of these technologies and their specific future development should solve the main challenges at the creation and scalable use of the methods for proactive error prevention. Especially the following problems should be solved: (1) high quality prediction for production errors, (2) automated creation of many predictive models, (3) automated maintenance of the models in the production. **Fig. 1** show the complete concept of the approach to solve the mentioned challenges. The basic components are explained below.

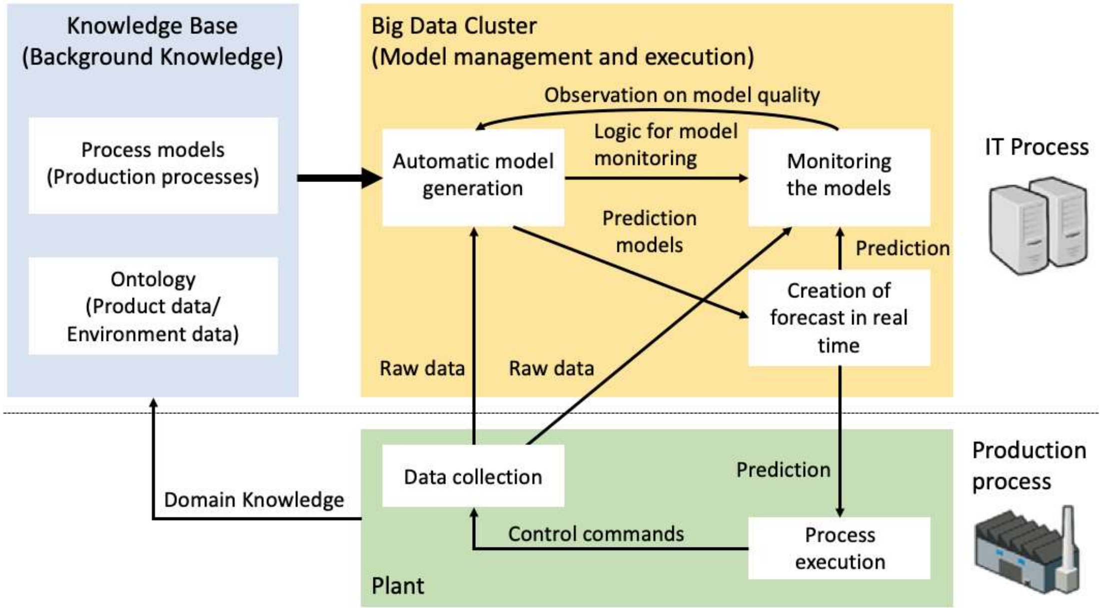

Fig. 1. Conception overview over the approach

A fundamental component of the solution is the modelling and use of background knowledge about the production processes and collected data. A domain expert with concrete process knowledge can often name relevant relationships. This background knowledge will get used to support the automatic creation and maintenance of prediction models. For example, when selecting data for a prediction model it is critical to select only relevant and non-redundant sources related to the predicted value. This selection can be done by heuristics but is faced with challenges, in particular for high dimensional data. Our approach is to create a modelling method that can be used to capture the required domain knowledge and make it available for automated data processing. Existing technologies such as ontologies are candidates for modelling language for production data and may be adapted to suit the requirements.

Another key element in the concept is the automatic generation of prediction models. Therefore, statistical models and machine learning models are going to be used. With these models, error predictions can be captured both as classification and as regression problems. The occurrence of certain error classes in current and future processes should be detected using classification and be avoided in advance. Regression will be used to predict the properties of products. If the expected properties differ from the objective, the parameters of the production process can be changed during operation and during production breaks. The corresponding product property can thus be corrected proactively.

With the help of big data technologies for parallel and highly scalable data processing, the solution spectrum for the model building will be investigated to find a case specific optimised solution. The parallelisation possibilities of established and emerging big data technologies should be used and if necessary adapted to achieve an efficient solution of the highly complex calculations.

The creation of models is not related to real-time requirements and can be implemented through batch-processing technologies. In comparison, the execution of prediction models must be continuous and - for some cases - in real time. The implementation of the prediction with technologies of stream processing and complex event processing should be investigated and suitable solutions should be identified. Technologies of stream- and complex event processing should be analysed to identify a suitable solution

Due to changes in the production process or in the data collection, so called concept drift may occur, which may result in a loss of accuracy for the prediction models. Therefore, methods should be explored, which detect crucial changes in an early stage. The necessary logic is to be generated in the context of the automated modelling process. The objective is to automatically adapt the monitoring logic to the specific models and relevant inputs.

5 Conclusion

In this paper, we shed light over the need for effective predictive maintenance for manufacturing companies and production factories. Also, we gave a short overview of our scheme to use background knowledge and ontologies to prevent costly waste and increase the production output. The desired result is a combination of background knowledge, big data and machine learning technologies. This outcome will be used to create predictive models and automatically adapt them to newly emerging variants in production lines. Among the main objectives of our intended system is to handle arising concept drifts, and continuously adapt itself to the requirements of the production environment.

Acknowledgement

This project was funded by the German Federal Ministry of Education and Research, funding line “Forschung an Fachhochschulen mit Unternehmen (FHProfUnt)“, contract number 13FH249PX6. The responsibility for the content of this publication lies with the authors. Also, we want to thank the company SICK AG for the cooperation and partial funding.

6 References

1. H. Kagermann, Chancen von Industrie 4.0 nutzen. In Industrie 4.0 in Produktion, Automatisierung und Logistik, Springer Fachmedien Wiesbaden, 2014, pp. 603-314.
2. J. Deuse, K. Weisner, A. Hengstebeck and F. Busch, Gestaltung von Produktionssystemen im Kontext von Industrie 4.0. In Zukunft der Arbeit in Industrie 4.0, Heidelberg: Springer Berling, 2015, pp. 99-109.
3. W. Fan and A. Bifet, Mining big data: current status, and forecast to the future. ACM SIGKDD Explorations Newsletter, 2013, pp. 1-5.
4. C. Chen and C. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on Big Data., 2014, pp. 275, 314-347.
5. N. Marz and J. Warren, Big Data: Principles and best practices of scalable realtime data systems. Manning Publications Co., 2015. pp 13-25.
6. J. Krumeich, S. Jacobi, D. Werth and P. Loos, "Big Data Analytics for Predictive Manufacturing Control-A Case Study from Process Industry.," in IEEE International Congress on Big Data, 2014. pp. 16-24
7. T. Wuest, D. Weimer, C. Irgens and K. D. Thoben, Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research,, 2016, pp. 23-45.
8. A. Kumar, R. McCann, J. Naughton, J. M. Patel , T. Barbos, R. Hunt and M. von Klein, A Survey of the Existing Landscape of ML Systems. UW- Madison CS Tech. Rep. TR1827., 2015.
9. C. Thornton, F. Hutter, H. Hoos and K. Leyton-Brown, "Auto- WEKA: Combined selection and hyperparameter optimization of classification algorithms.," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp.847-855.
10. I. Guyon, I. Chaabane, H. Escalante, S. Escalera, D. Jajetic, J. Lloyd and A. Statnikov, "A brief Review of the ChaLearn AutoML Challenge: Any-time Any-dataset Learning without Human Intervention" in International Conference in Machine Learning, 2016.
11. J. Bergstra, R. Bardenet, Y. Bengio and B. Kégl, Algorithms for hyper- parameter optimization. In Advances in Neural Information Processing Systems, 2011, pp. 2546-2554.

12. E. Kharlamov, B. Grau, E. Jiménez-Ruiz, S. Lamparter, G. Mehdi, M. Ringsquandl and I. Horrocks, Capturing Industrial Information Models with Ontologies and Constraints. In International Semantic Web Conference, Springer International Publishing, 2016, pp. 325-343.
13. M. Ringsquandl, S. Lamparter, S. Brandt, T. Hubauer and R. Lepratti, Semantic-Guided Feature Selection for Industrial Automation Systems. In International Semantic Web Conference, Springer International Publishing, 2015, pp. 225-240.
14. D. Crankshaw, P. Bailis, J. Gonzalez, H. Li, Z. Zhang, M. Franklin and M. Jordan, The missing piece in complex analytics: Low latency, scalable model management and serving with velox. arXiv preprint arXiv, 2014, pp. 1409-3809.
15. J. Schleier-Smith, An Architecture for Agile Machine Learning in Real-Time Applications. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015, pp. 2059-2068.
16. D. Luckham, The power of events, 204 ed., 2002.
17. O. Etzion and P. Niblett, Event processing in action, Manning Publications Co, 2010.
18. Y. Zhang, Q. Dai and R. Zhong, An extensible event-driven manufacturing management with complex event processing approach. International Journal of Control and Automation, 2009, pp. 1-12.
19. I. Kellner and L. Fiege, Viewpoints in complex event processing: industrial experience report. In Proceedings of the Third ACM International Conference on Distributed Event-Based Systems, 2009, p. 9.