Findings from Production Management Research

RESEARCH

Fabian Steinberg

# Machine Learning-based Prediction of Missing Parts for Assembly

Springer Vieweg

# Findings from Production Management Research

**Series Editor**

Peter Burggräf, International Production Engineering & Management, Universität Siegen, Kreuztal, Nordrhein-Westfalen, Germany

In the dynamic realm of industrial operations, production management remains the core for organizational efficiency. This series is a collection of doctoral thesis delving into the latest insights and innovations shaping production management, offering a comprehensive exploration of emerging trends and best practices. From the integration of advanced digital technologies to the effects of sustainability initiatives, this series covers key facets driving the evolution of modern production management. Drawing upon cutting-edge research and practical case studies, this collection is an indispensable resource for industry professionals, academics and students seeking to navigate and capitalize on the dynamic landscape of modern production management.

Fabian Steinberg

# Machine Learning-based Prediction of Missing Parts for Assembly

Springer Vieweg

Fabian Steinberg
Universität Siegen
Siegen, Germany

Dissertation zur Erlangung des Grades eines Doktors der Ingenieurswissenschaften vorgelegt von Fabian Steinberg, M. Sc. eingereicht bei der Naturwissenschaftlich-Technischen Fakultät der Universität Siegen, Siegen 2024
Betreuer und erster Gutachter: Univ.-Prof. Dr.-Ing. Peter Burggräf, Universität Siegen
Zweiter Gutachter: Univ.-Prof. Dr. rer. nat. Jochen Garcke, Universität Bonn
Tag der mündlichen Prüfung: 09. Februar 2024

ISSN 3005-1649 ISSN 3005-1657 (electronic)
Findings from Production Management Research
ISBN 978-3-658-45032-8 ISBN 978-3-658-45033-5 (eBook)
https://doi.org/10.1007/978-3-658-45033-5



This Springer Vieweg imprint is published by the registered company Springer Fachmedien Wiesbaden GmbH, part of Springer Nature.
The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

# Acknowledgements

This thesis on the prediction of missing parts for assembly was written during my time as a research assistant at the Chair of International Production Engineering and Management (IPEM) at the University of Siegen, Germany. I would like to thank the people who contributed to the success of this work.

Special thanks go to Prof. Dr.-Ing. Peter Burggräf, holder of the IPEM chair, for the opportunity to do my doctorate and for his support, encouragement and trust in me as a person. The environment he created, the first-class education, the wealth of ideas and the collegial management style were ideal conditions for this work and for my personal development.

I would also like to thank Univ.-Prof. Dr. rer. nat. Jochen Garcke, Research Group Leader at the Institute for Numerical Simulation at the University of Bonn and Head of the Business Area Numerical Data-Driven Prediction at the Fraunhofer Institute for Algorithms and Scientific Computing SCAI, for his willingness to serve as the second expert witness. Dr.-Ing. Robert Brandt, holder of the Chair of Materials Systems for Lightweight Vehicle Construction at the University of Siegen, for assuming the chair of the doctoral committee, and Univ.-Prof. Dr.-Ing. Axel van Heel, holder of the Chair of Materials Science and Materials Testing, for being a member of the doctoral committee.

Many of my former and active colleagues at IPEM have had a decisive influence on the content of this work, and I am deeply grateful to them. The mutual helpfulness, the high motivation, the creativity and the friendship have always given me new motivation. I would like to emphasize the cooperation with Dr.-Ing. Johannes Wagner, Benjamin Heinbach, Till Saßmannshausen, Rene Sauer, Alexander Becher, Philipp Nettesheim and Maximilian Schütz. Furthermore, I would like to thank my student assistants and the students who wrote their bachelor and master theses under my supervision, and continuously supported me in my work at IPEM. I would also like to thank Mrs. Mailin Klaas for the first-class administrative organization and for her sympathetic support throughout my time at the chair.

My deepest and most heartfelt thanks go to my wife Kristina Bieker-Steinberg for her loving support, her motivating encouragement and the space she has given me, as well as to my parents Birgit and Andreas Steinberg for their unconditional support on my life's journey and for encouraging me to change my career from industry back to academia. Without their loving support, this thesis would not have been possible.

Olpe Fabian Steinberg
February 2024

# Abstract

Manufacturing companies are faced with the challenge of managing increasing process complexity, while at the same time having to meet ever higher demands in terms of on-time delivery and product costs. Especially at points in the value chain such as assembly, where different material flows converge, it is often not possible to provide the components required for an order in a timely and synchronized manner. Early identification of missing parts at the beginning of assembly can help to take countermeasures to meet the required delivery dates. To achieve this, this thesis develops machine learning based prediction models that can predict potential missing parts at the start of assembly at an early stage in the value chain. The development of the models was carried out as case studies at manufacturing companies in the machine industry. As a basis for the development, an extensive systematic literature search was conducted on existing approaches for the prediction of lead times of production orders. The result was that no approach exists that takes into account the full complexity of manufacturing companies. In particular, with regard to the data used, it became clear that information about the product to be manufactured—so-called material data—has not been used up to now. Based on the systemic review, a model for predicting missing parts from in-house production was implemented. It was shown that classification approaches achieve the best possible model quality for components from in-house production. With the defined modeling approach—classification—it was then verified that material data has a significant influence on the model quality and is therefore relevant for the prediction of missing parts at the start of assembly. Finally, a model for predicting delivery delays in the purchasing process was implemented, which makes it possible to predict potential missing parts from suppliers at the time of ordering. The case studies show that the use of machine learning for the prediction of missing parts in both in-house production and the purchasing process can identify delays in the start of assembly at an early stage. The developed models are therefore suitable as a support system for production planners and controllers as well as purchasing departments to improve material availability at the start of assembly.

# Zusammenfassung

Produzierende Unternehmen stehen vor der Herausforderung, eine stetig wachsende Prozesskomplexität bei gleichzeitig steigenden Anforderungen an Termintreue und Produktkosten zu beherrschen. Insbesondere an Stellen der Wertschöpfungskette wie der Montage, an denen verschiedene Materialflüsse zusammenlaufen, gelingt es häufig nicht, die für einen Auftrag benötigten Komponenten rechtzeitig und synchron bereitzustellen. Das frühzeitige Erkennen von Verzögerungen bis zum Montagebeginn kann helfen, Gegenmaßnahmen einzuleiten, um die geforderten Liefertermine einzuhalten. Um dies zu erreichen, wurden im Rahmen dieser Arbeit maschinelle lernbasierte Prognosemodelle entwickelt, die potentielle Fehlteile zum Montagestart frühzeitig vorhersagen können. Die Entwicklung der Modelle wurde jeweils im Rahmen von Fallstudien bei produzierenden Unternehmen aus dem Maschinen- und Anlagenbau betrachtet. Als Grundlage für die Entwicklung wurde zunächst eine umfassende systematische Literaturrecherche zu bestehenden Ansätzen zur Vorhersage der Durchlaufzeit von Fertigungsaufträgen durchgeführt. Das Ergebnis war, dass bislang kein Ansatz existiert, der die gesamte Komplexität produzierender Unternehmen berücksichtigt. Insbesondere bei den verwendeten Daten zeigte sich, dass Informationen über das zu fertigende Produkt—so genannte Materialdaten—bisher nicht genutzt werden. Auf Basis dieser Untersuchungen wurde zunächst ein Modell zur Vorhersage von Fehlteilen aus der eigenen Produktion implementiert. Dabei zeigte sich, dass in diesem Bereich Klassifikationsansätze die bestmögliche Modellgüte erreichen. Mit der gewählten Modellierungsart—Klassifikation—wurde anschließend ermittelt, dass Materialdaten einen signifikanten Einfluss auf die Modellgüte haben und somit für die Vorhersage von Fehlteilen am Montagestart relevant sind. Schließlich wurde ein Modell zur Vorhersage von Lieferterminverzögerungen im Einkaufsprozess implementiert, mit dessen Hilfe potentielle Fehlteile von Lieferanten bereits zum Zeitpunkt der Bestellung vorhergesagt werden können. Die betrachteten Fallbeispiele zeigen, dass durch den Einsatz von maschinellem Lernen zur Vorhersage von Fehlteilen in der eigenen Fertigung sowie im Einkaufsprozess Ursachen für Verzögerungen des Montagestarts frühzeitig identifiziert werden können. Die entwickelten Modelle eignen sich somit als Assistenzsystem für Produktionsplaner und -steuerer sowie Einkaufsabteilungen, um die Materialverfügbarkeit zum Montagestart zu verbessern.

# Contents

# List of Abbreviations

| | |
|---|---|
| € | Euro |
| AB | Adaptive Boosting |
| AdaBoost | Adaptive Boosting |
| ANN | Artificial Neural Networks |
| APS | Advanced Planning System |
| ASD | Assembly Start Delayer |
| BAG-DT | Bagged Trees |
| CAD | Computer Aided Drawing |
| CO | Combinatorial Optimization |
| CON | Constant |
| CRISP-DM | Cross Industry Standard Process for Data Mining |
| CSV | Comma Separated Value |
| CT | Control Theory |
| DecT | Decision Theory |
| DT | Decision Tree |
| ERP | Enterprise Resource Planning |
| GB | Gradient Boosting |
| H | Heuristics |
| JIQ | Jobs in Queue |
| KNN | K-Nearest-Neighbor |
| LINREG | Linear Regression |
| LOGREG | Logistic Regression |
| LP | Linear Programming |
| LR | Linear Regression |

| | |
|---|---|
| LSA | Latent Semantic Analysis |
| MES | Manufacturing Execution Systems |
| ML | Machine Learning |
| MLP | Multilayer Perceptron |
| NC | Numerical Control |
| NLP | Nonlinear Programming |
| NOP | Number of Operations |
| OP | Operation |
| OR | Operations Research |
| PC | Principal Component |
| PCA | Principal Component Analysis |
| PDM | Precedence Diagram Method |
| PLM | Product Lifecycle Management |
| PPC | Production Planning and Control |
| QC | Quality Criteria |
| QT | Queuing Theory |
| RAN | Random |
| ReLU | Rectified linear unit |
| RF | Random Forest |
| RQ | Research Question |
| SCM | Supply Chain Management |
| SVC | Support Vector Classifier |
| SVM | Support Vector Machine |
| tanh | Hyperbolic Tangent |
| TU | Time Units |
| TWK | Total Work |
| WIQ | Work in Queue |

# List of Symbols

| | |
|---|---|
| ASD | Assembly Start Delayer |
| CD | Completion Date |
| $DD_{finished}$ | Finished Date Deviation |
| DDL | Delivery Date Lateness |
| $DD_{rel}$ | Relative Date Deviation |
| $DD_{start}$ | Start Date Deviation |
| $DOF_{actual}$ | Actual Finished Date of an Order or Operation |
| $DOF_{target}$ | Target Finished Date of an Order or Operation |
| $DOS_{actual}$ | Actual Start Date of an Order or Operation |
| $DOS_{target}$ | Target Start Date of an Order or Operation |
| F | F-score |
| FN | False Positive |
| FP | False Negative |
| $G_{left/right}$ | Impurity of the Left/Right Subset |
| I1 | Inital Inventory |
| I2 | Final Inventory |
| IC | Completed Inventory |
| ID | Disrupted Inventory |
| IP | Inventory in Process |
| IW | Waiting Inventory |
| J(k,t) | Cost Function |
| LT | Lead Time |
| $LT_{actual}$ | Actual Lead Time of an Order or Operation |
| $LT_{target}$ | Target Lead Time of an Order or Operation |