PrEvelOp:

Die Wirtschaftlichkeit von KMU steigern

Dank datenbasierter Ähnlichkeitsanalyse die Varianz in Fertigungsprogrammen reduzieren

Im Forschungsprojekt PrEvelOp werden datengetriebene Ansätze entwickelt, um kleine und mittlere Unternehmen (KMU) im Maschinen- und Anlagenbau bei der Reduktion ihrer Fertigungsprogrammvarianz zu unterstützen. Im Mittelpunkt steht die Verringerung der Varianten in der Produktion, um dadurch Effizienzpotenziale zu nutzen und die Transparenz in den Abläufen zu verbessern. Mithilfe moderner Datenverarbeitungstechnologien und maschineller Lernverfahren wird das Projekt im Ergebnis innovative Lösungen bieten, die KMU einen niederschwelligen Zugang zur Digitalisierung ermöglichen. >

PrEvelOp:

Increase the profitability of SMEs

Reducing Production Program Variance Through Data-Based Similarity Analysis

The PrEvelOp research project develops data-driven approaches to support small and medium-sized enterprises (SMEs) in mechanical and plant engineering in reducing variance in their production programs. The project focuses on reducing production variants to unlock efficiency potential and improve process transparency. Leveraging advanced data processing technologies and machine learning methods, the project provides SMEs with practical and accessible pathways to digitalization. >

Die zunehmende Produktvielfalt stellt kleine und mittlere Unternehmen (KMU) im Maschinen- und Anlagenbau vor wachsende Herausforderungen[1]. Durch die Steigerung der Produktvielfalt sinken Losgrößen und die Komplexität der Produktion steigt, da sich Unternehmen vermehrt für flexible, aber nur bedingt ökonomisch effiziente Werkstatt- und Inselfertigungen entscheiden[2]. In Summe führt die durch eine hohe Produktvielfalt entstehende Artikel- und Fertigungsprozessvielfalt, im Folgenden zusammengefasst unter dem Begriff der Fertigungsprogrammvarianz (siehe zur Abgrenzung der Begriffe auch Figure 1), somit zu steigenden Herstellkosten pro Stück und sinkender Transparenz über deren Entstehung[3]. Vor diesem Hintergrund adressiert das PrEvelOp-Projektteam diese Problematik, indem es innovative datengetriebene Ansätze zur Reduktion der Fertigungsprogrammvarianz entwickelt.

Im Rahmen des Projekts werden Funktionen entwickelt, die in Auftragsdaten (bestehend aus Artikel- sowie Fertigungsprozessdaten) Ähnlichkeiten mithilfe von Methoden des unüberwachten Lernens (eng. *unsupervised learning*) identifizieren. Die Funktionen ermöglichen die Gruppierungen von Aufträgen nach artikelspezifischen und/oder prozessspezifischen Merkmalen. Je nach Gruppierungen werden durch eine Entscheidungsunterstützung passende Maßnahmen zur Reduzierung der Fertigungsprogrammvarianz vorgeschlagen. Das Projekt wird federführend vom *FIR e. V. an der RWTH Aachen in* Kooperation mit dem *Werkzeugmaschinenlabor (WZL) der RWTH Aachen* durchgeführt.

Vorgehen zur Datenvorbereitung

Ein wesentlicher Schritt im Projekt PrEvelOp ist die strukturierte Aufbereitung der Daten, die sowohl CAD-Bauteildaten als auch Prozessdaten umfassen (siehe Bild 2). Die Bauteildaten werden aus CAD-Dateien extrahiert und in ein maschinenlesbares, strukturiertes Format überführt. Parallel dazu werden die Prozessdaten aufbereitet, um einheitliche und konsistente Datensätze zu erzeugen. Das Ziel dieser Maßnahmen besteht darin, beide Datentypen über einen gemeinsamen Nenner – in der Regel die Bauteilnummer – zusammenzuführen und zu konsolidieren. Das Ergebnis dieses Prozesses ist eine bereinigte und standardisierte Tabelle, die sowohl CAD- als auch Prozessdaten integriert und damit die Grundlage für weiterführende Analysen bildet.

[1] s. Koether u. Meier 2020, S. 69
[2] s. Esmaeilian et al. 2016, S. 86
[3] s. Krause u. Gebhardt 2018, S. 43

Increasing product diversity poses growing challenges for medium-sized enterprises (SMEs) in the mechanical and plant engineering sector[1]. As product diversity expands, batch sizes shrink, and production complexity rises, with companies increasingly relying on flexible, but often economically inefficient workshop production and cellular manufacturing (or production island) methods.[2] Overall, the variety of articles and production processes resulting from high product diversity, collectively referred to as production program variance (see Fig. 1 for terminology definitions), drives up per-unit manufacturing costs and reduces transparency regarding cost origins.[3] To address these challenges, the PrEvelOp research projecters develop innovative, data-driven approaches to reduce production program variance.

As part of the project, functions are being developed that identify similarities in order data (comprising article and production process information) using unsupervised learning methods. These functions enable orders to be grouped according to item- and/or process-specific characteristics. Based on these groupings, suitable measures for reducing production program variance are proposed through decision support tools. The project is led by *FIR e. V. an der RWTH Aachen* in collaboration with the *Laboratory for Machine Tools and Production Engineering (WZL) at RWTH Aachen University*.

Data Preparation Procedure

A key step in the PrEvelOp project is the structured preparation and integration of data, including both CAD component data and process data (see Fig. 2). Component data is extracted from CAD files and converted into a machine-readable, structured format. At the same time, process data is prepared to generate uniform and consistent data records. The aim of these steps is to merge and consolidate both data types using a common reference – typically the component number. This process results in a cleaned and standardized table that integrates both

[1] see Koether and Meier 2020, p. 69
[2] see Esmaeilian et al. 2016, p. 86
[3] see Krause and Gebhardt 2018, p. 43

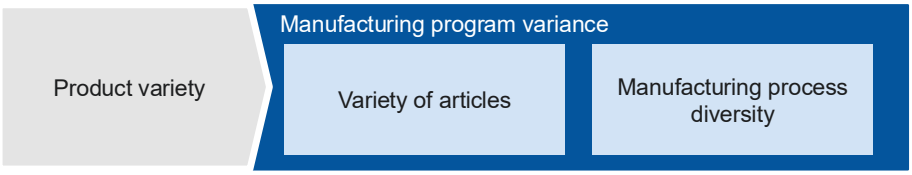

Figure 1: Terminology Definition

Neben der Zusammenführung der beiden Datensätze zu einer gemeinsamen Tabelle müssen die Daten weiter aufbereitet werden, um eine konsistente und qualitativ hochwertige Datenbasis zu schaffen. Im Projekt PrEvelOp wurden hierfür verschiedene Funktionen zur Datenvorverarbeitung und -analyse entwickelt. Die explorative Datenanalyse umfasst dabei die Analyse der vorhandenen Datenstruktur. Hierzu gehören die Identifikation von Datentypen und potenziellen Ausreißern, beispielsweise durch die Berechnung des Interquartilsabstands. Eine Prüfung der Datenqualität wird durchgeführt, indem Fehl- und Leereinträge erkannt werden, die häufig in Auftrags- und Artikeldaten auftreten.

Für die automatisierte Datenvorverarbeitung wurden Funktionen wie Lowercasing und Encoding implementiert. Dabei werden Zeichenketten wie Artikelnummern oder Artikelbezeichnungen normalisiert und vereinheitlicht. So wird beispielsweise aus Varianten gleicher Bedeutung wie „Artikel“, „artikel“ und „ARTIKEL“ eine einheitliche Form erstellt. Dieses Verfahren stellt sicher, dass beim anschließenden Encoding unterschiedliche Schreibweisen nicht fälschlicherweise als unterschiedliche Attribute verarbeitet werden. Diese Maßnahmen sind essenziell, da maschinelle Lernverfahren Zeichenketten nicht direkt verarbeiten können[4].

Neben diesen automatisierten Verfahren wird auch das Domänenwissen der Anwender berücksichtigt. Die manuelle Anpassung bestimmter Vorverarbeitungsschritte, wie die Attributauswahl für die Clusteranalyse, ermöglicht es,

[4] s. Müller U. Guido 2017

CAD and process data, providing a solid foundation for subsequent analyses.

In addition to merging the two data sets into a common table, the data must be further processed to create a consistent, high-quality database. In the PrEvelOp project, various functions for data pre-processing and analysis were developed to achieve this goal. Exploratory data analysis involves analyzing the existing data structure, including identifying data types and detecting potential outliers, for example by calculating the interquartile range. A data quality check is carried out by identifying incorrect or missing entries, which commonly occur in order and item data.

For automated data pre-processing, functions such as lowercasing and encoding have been implemented. Character strings such as article numbers or article descriptions are normalized and standardized. For example, variants with the same meaning, such as “article”, “artikel”, and “ARTIKEL” are converted to a standardized form. This ensures that different spellings are not incorrectly treated as different attributes during subsequent encoding. These steps are crucial, as machine learning methods cannot process character strings in their raw form.[4]

In addition to these automated processes, users' domain knowledge is also considered. Manual adjustments to certain pre-processing steps, such as attribute selection for cluster analysis, allow both item-specific and process-

[4] see Müller and Guido 2017

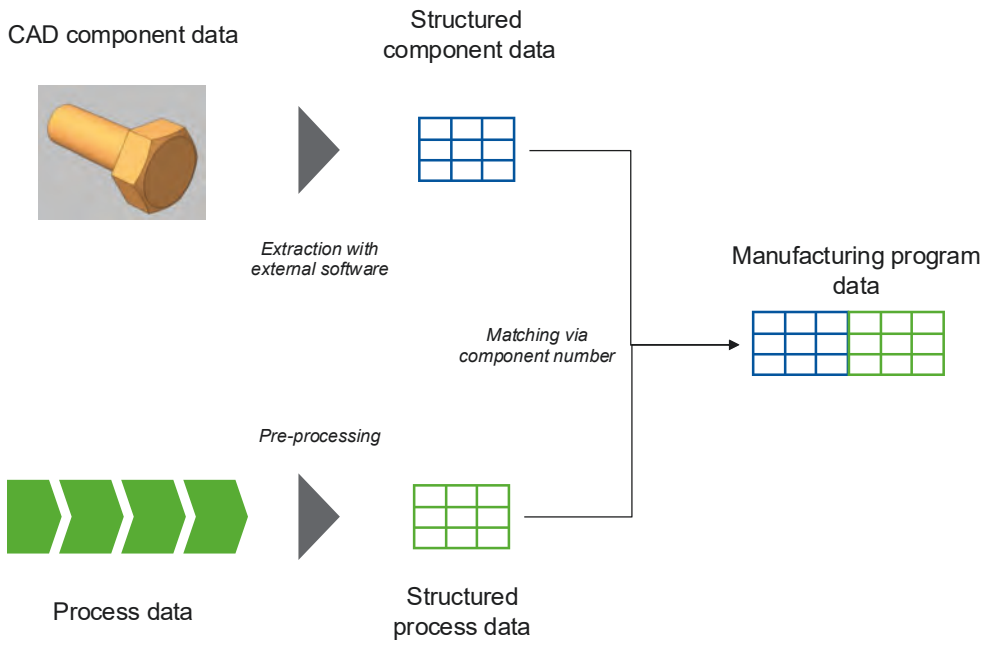

Figure 2: Data Preparation Procedure

sowohl artikelspezifische als auch prozessbezogene Merkmale in die Analyse einzubeziehen. Dies gewährleistet, dass die Clusteranalyse auf einer fachlich fundierten Datenbasis durchgeführt wird und relevante Merkmale für die Gruppierung einbezieht.

Durch die finale standardisierte und fehlerbereinigte Datenstruktur wird eine solide Grundlage für den Einsatz maschineller Lernverfahren geschaffen.

Einsatz maschineller Lernverfahren zur Analyse der Fertigungsprogrammvarianz

Auf der Grundlage qualitativ hochwertiger Daten ermöglichen maschinelle Lernverfahren die Modellierung und Analyse der Fertigungsprogrammvarianz. Dabei kommen insbesondere Methoden des unüberwachten Lernens, wie k-Means- und hierarchisches Clustering, zum Einsatz. Der k-Means-Algorithmus teilt Datenpunkte iterativ in k Gruppen ein, wobei die Distanz innerhalb der Gruppen minimiert wird. Das hierarchische Clustering organisiert die Daten hingegen in einer Baumstruktur, die die Beziehungen zwischen den Gruppen hierarchisch darstellt und visualisiert.

Gerade das hierarchische Clustering bietet den Vorteil, dass die Ähnlichkeiten und Verbindungen zwischen Bauteilen anschaulich dargestellt werden. In der Baumstruktur werden Bauteile entsprechend ihrer Ähnlichkeit so angeordnet, dass die eng verwandten Elemente nahe beieinander liegen. Je nachdem, an welcher Stelle die Struktur geschnitten wird, entstehen größere oder kleinere Clustergruppen. Diese grafische Darstellung ermöglicht es Anwendern, Gemeinsamkeiten schnell zu erkennen und gezielt für die Analyse sowie Entscheidungsfindung zu nutzen.

Die Ergebnisse dieser Analysen liefern wertvolle Einblicke in die Struktur der Fertigungsprogrammvarianz und dienen als Ausgangspunkt für die Ableitung spezifischer Maßnahmen zur Reduzierung der Fertigungsprogrammvarianz.

Ergebnis und Nutzen für die Praxis

Das Forschungsprojekt PrEvelOp wird von einem projektbegleitenden Ausschuss unterstützt, der sich aus Forschungseinrichtungen und Industrieunternehmen zusammensetzt. Zur Sicherstellung der Praxisnähe und Anwendungsorientierung sind verschiedene Unternehmen aus dem produzierenden Gewerbe beteiligt. Diese Unternehmen unterstützen die Entwicklung und Optimierung der einzelnen Projektbausteine, indem sie nicht nur Daten liefern, sondern auch ihr branchenspezifisches Fachwissen aktiv in die Ent-

related characteristics to be included in the analysis. This ensures that the cluster analysis is carried out on a technically sound database, incorporating relevant characteristics for accurate grouping.

The final standardized and error-corrected data structure forms a solid foundation for applying machine learning methods.

Analyzing Production Program Variance with Machine Learning

With high-quality data, machine learning methods enable the effective modeling and analysis of production program variance. Specifically, unsupervised learning methods such as k-means and hierarchical clustering are employed. The k-means algorithm iteratively partitions data points into k groups, minimizing intra-group distance. Hierarchical clustering organizes the data into a tree structure, visually representing the relationships between the groups.

Hierarchical clustering has the advantage of clearly illustrating the similarities and connections between components. In the tree structure, components are arranged according to their similarity, with closely related items placed close to each other. Depending on where the structure is cut, larger or smaller clusters can be formed. This graphical representation enables users to quickly identify similarities and leverage them for targeted analysis and decision-making.

The results of these analyses provide valuable insights into the structure of production program variance and serve as a starting point for developing specific measures to reduce it.

Results and Benefits for Practice

The PrEvelOp research project is supported by a project committee consisting of research institutions and industrial companies. Various manufacturing industry companies are involved in ensuring practical relevance and application-oriented outcomes. These companies contribute to the development and optimization of the individual project modules not only by providing data, but also by actively sharing their industry-specific expertise in the development and validation of the project solutions. The order and article data provided serve as the foundation for developing the data processing and analysis functions.

The goal is to create an open-source library that provides functions for the machine learning-based reduction of production program variance. This library will be made

wicklung und Validierung der Projektlösungen einbringen. Die bereitgestellten Auftrags- und Artikeldaten dienen als Grundlage für die Entwicklung der Funktionen zur Datenverarbeitung und -analyse.

Als Ergebnis wird eine Open-Source-Bibliothek angestrebt, die Funktionen zur maschinellen, lernbasierten Reduzierung der Fertigungsprogrammvarianz bereitstellt. Diese soll KMU über eine permissive Lizenz sowie zu Test- und Übungszwecken in Form einer rudimentären Demo-Applikation zugänglich gemacht werden.

Eine Minimierung der Fertigungsprogrammvarianz ist eine notwendige Voraussetzung für die Identifikation und Umsetzung varianzreduzierender Maßnahmen. Dies ermöglicht es, varianzinduzierte Komplexitätskostentreiber – wie geringe Losgrößen oder häufige Werkzeugwechsel – zu adressieren. Ziel ist es, das Spannungsfeld zwischen individueller Produktion (*economies of scope*) und Massenproduktion (*economies of scale*) zu entschärfen und dadurch die Wettbewerbsfähigkeit der Unternehmen nachhaltig zu stärken.

Abiraam Kantharajah

available to SMEs under a permissive license, accompanied by a basic demo application for testing and training purposes.

Minimizing production program variance is essential for identifying and implementing measures to reduce it. This makes it possible to address variance-induced complexity cost drivers - such as small batch sizes or frequent tool changes. The aim is to reduce the tension between custom production (economies of scope) and mass production (economies of scale), ultimately strengthening the long-term competitiveness of companies.

Abiraam Kantharajah

Literatur:

Koether, R.; Meier, K.-J.: Lean Production für die variantenreiche Einzelfertigung. Flexibilität wird zum neuen Standard. 2., aktualis. u. überarb. Auflage. Springer Gabler, Wiesbaden 2020.

Esmaeilian, B.; Behdad, S.; Wang, B.: The evolution and future of manufacturing: A review. In: Journal of Manufacturing Systems 39(2016)4, S. 79–100. DOI: 10.1016/j.jmsy.2016.03.001.

Krause, D.; Gebhardt, N.: Methodische Entwicklung modularer Produktfamilien. Hohe Produktvielfalt beherrschbar entwickeln. Springer Vieweg, Berlin [u. a.] 2018.

Müller, A. C.; Guido, S.: Introduction to machine learning with Python. A guide for data scientists. O'Reilly, Sebastopol (CA) 2017.

PREVELOP Production – Development – Optimization

This pre-competitive project was funded by the Federal Ministry for Economic Affairs and Climate Action with IGF funds.

Project Title: PrEvelOp – Production – Development – Optimization: Reduction of manufacturing program variance using data-based similarity analysis to increase the profitability of SMEs

Funding/Promoters: Federal Ministry for Economic Affairs and Climate Action (BMWK), Projektträger im DLR

Funding no.: 01IF22649 N Duration: 01.05.2023 – 31.12.2024

Research Partner: C.A.PICARD®, ESSERTEC GMBH, exprobico, mk Plast GmbH & Co. KG, myOpenFactory Software GmbH, Ortlinghaus-Werke GmbH, SIMUFORM Search Solutions GmbH, simus systems GmbH, STURM® INDUSTRIES, Wulf Zargen GmbH & Co.KG

Website: prevelop.fir.de

Abiraam Kantharajah, M.Sc.
Project Manager
Research Unit Information Management
FIR e. V. an der RWTH Aachen
Phone: +49 163 3152523
Email: Abiraam.Kantharajah@fir.rwth-aachen.de

Bundesministerium für Wirtschaft und Klimaschutz

DLR Projektträger

Open Access: Dieser Artikel wird unter der Creative-Commons-Lizenz „Share Alike 4.0 International – Weitergabe unter gleichen Bedingungen 4.0 International" (CC BY-SA 4.0) veröffentlicht.