

KI-gestützte Prognose von Durchlauf- und Lieferzeiten in der Einzel- und Kleinserienfertigung

Ein holistischer Ansatz

Janek Bender*,
Martin Trat und
Jivka Ovtcharova

Die Prognose von Durchlauf- und Lieferzeiten ist herausfordernd, besonders in der Einzel- und Kleinserienfertigung. Diese Arbeit stellt einen holistischen Ansatz zur KI-gestützten Prognose dieser Zeiten vor, welcher Techniken aus den Bereichen Automated Machine Learning und Lifelong Learning kombiniert, um leichtgewichtig auch langfristig robuste Prognosemodelle hervorzubringen. Der Ansatz wurde in zwei Fallstudien in der Metallverarbeitung und der Möbelbranche erfolgreich erprobt.

Akkurate Durchlauf- und Lieferzeiten für eine reibungslose Auftragssteuerung

Eine wichtige Eingangsgröße für die unternehmensinterne Auftragsplanung und -steuerung ist die erwartete Durchlaufzeit. Unternehmensextern ist dagegen die erwartete Lieferzeit interessant, welche neben Transportzeiten auch die Durchlaufzeiten der vorgelagerten Wertschöpfungskette miteinschließt (Bild 1).

Eine möglichst genaue Prognose dieser Zeiten ist entscheidend, um die Abläufe in der Produktion optimal steuern zu können und damit die Termintreue hochzuhalten, was letztlich die Kundenzufrie-

denheit sichert. Doch besonders in der Produktion variantenreicher Produkte, wie in der Einzel- und Kleinserienfertigung, ist die akkurate Prognose dieser Durchlauf- und Lieferzeiten komplex. Sie ist abhängig von variierenden Produkteigenschaften, den zugehörigen vielfältigen Produktionsprozessen sowie der Verfügbarkeit und Auslastung von Produktionsressourcen [1, S. 9–11]. Die zunehmende Verfügbarmachung dieser dezentral im Unternehmen entstehenden Produkt-Prozess-Ressourcen-Datenströme sowie der rapide Fortschritt im Bereich Künstlicher Intelligenz (KI) bieten jedoch auch eine Chance, die komplexe Problematik der Durchlauf- und Lieferzeitenprognose datengetrieben zu lösen. Hierzu existiert

bereits eine Menge an Arbeiten, welche KI-Modelle auf Basis von Algorithmen des Maschinellen Lernens für diese Problemklasse entwickelten und evaluierten [2–4]. Nur bei wenigen Arbeiten findet allerdings der produktive Einsatz der entwickelten KI-Modelle Berücksichtigung. Besonders hervorzuheben ist hier die Notwendigkeit, KI-Modelle robust gegenüber Veränderungen in den zugrundeliegenden Datenströmen zu gestalten und damit Performanzverlusten im langfristigen Produktiveinsatz durch Concept Drift [5] vorzubeugen. Die vorliegende Arbeit stellt daher eine Ende-zu-Ende-Methode zur KI-gestützten Prognose von Durchlauf- und Lieferzeiten vor. Im Folgenden werden zunächst zwei Kerntechnologien des Ansatzes vorgestellt, bevor dieser im Detail erläutert und anhand von Beispielen aus der Praxis erprobt wird.

Maschinelles Lernen als Kerntechnologie

Die Fortschritte der letzten Jahre im Bereich von KI-Technologien sind gigantisch und eröffnen ganz neue Lösungswege. Zwei dieser Technologien sind entschei-

* Korrespondenzautor

Dr.-Ing. Janek Bender; FZI Forschungszentrum Informatik; Haid-und-Neu-Straße 10-14, 76131 Karlsruhe; Tel.: +49 (0) 721 9654-501, E-Mail: bender@fzi.de

Weitere Autor:innen

Martin Trat, M. Sc.; FZI Forschungszentrum Informatik
Prof. Dr. Dr.-Ing. Dr. h. c. Jivka Ovtcharova; FZI Forschungszentrum Informatik

Hinweis

Bei diesem Beitrag handelt es sich um einen von den Advisory-Board -Mitgliedern des ZWF-Sonderheftes wissenschaftlich begutachteten Fachaufsatz (Peer-Review).

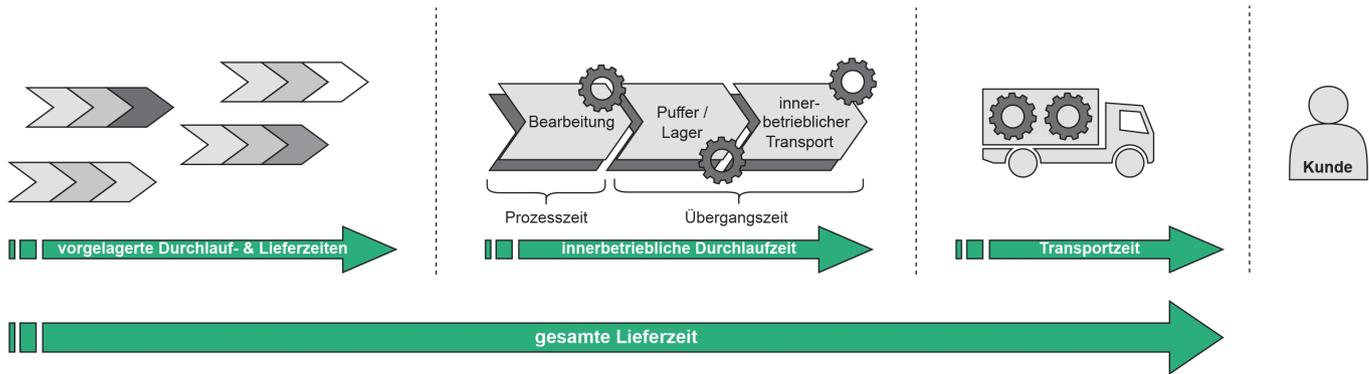


Bild 1. Durchlauf- und Lieferzeiten im Gesamtkontext

dend für die Überlegenheit des Ansatzes gegenüber herkömmlichen Verfahren.

Mit Automated Machine Learning zur unternehmensinternen KI-Entwicklung

Eine technische Herausforderung für Unternehmen liegt in der Entwicklung konkreter KI-Modelle. Die zugrundeliegenden Technologien erfordern tiefgehendes Fachwissen, über welches hochspezialisierte Data Scientists verfügen, welche unter hohem Aufwand komplexe KI-Pipelines entwickeln [6]. Aus dieser Motivation heraus wird zunehmend versucht, notwendige Schritte zur Entwicklung von KI-Modellen zu automatisieren. Dazu eignen sich Techniken des Automated Machine Learning (AutoML) [7]. Im Kern wird das Suchproblem nach der optimalen Kombination aus Algorithmus und Hyperparametern adressiert und dessen Lösung weitgehend automatisiert [8]. Eine zentrale AutoML-Methode für die Anwendung auf Datenströme ist das Sequential Halving [9]. Es handelt sich dabei um einen populationsbasierten Ansatz, bei dem eine Menge von Algorithmus-Hyperparameter-Kombinationen iterativ trainiert und anhand der gewählten Performanzmetrik, bspw. dem Mean Absolute Error (MAE), evaluiert wird. Zu definierten Iterationen, den Halvings, werden die 50 Prozent der Kombinationen verworfen, die sich im unteren Bereich des Performanzrankings befinden. Dies geschieht so lange, bis nur noch eine einzige Kombination übrig bleibt. Neben der konkreten Modellentwicklung werden vorgelagerte Entwicklungsprozesse, wie bspw. Datenvorverarbeitung oder Feature Engineering, sowie nachgelagert Betrieb und Wartung der KI-Modelle jedoch

nur bedingt durch AutoML unterstützt [1, S. 44–45, 10]. Dennoch befähigt der Einsatz von AutoML Unternehmen, selbst maßgeschneiderte KI-Modelle zur Durchlauf- und Lieferzeitenprädiktion zu entwickeln, sofern diese Techniken in eine holistische Methode eingebettet sind. Damit bildet AutoML eine der Kerntechnologien der vorliegenden Arbeit.

Langfristig robuste KI dank Concept-Drift-Adaption

In realen Anwendungsfällen sind Daten häufig Änderungen unterworfen, was als Concept Drift bezeichnet wird [11]. Dieser kann gravierend destruktive Auswirkungen auf KI-Modelle haben, weswegen dedizierte Maßnahmen zum Umgang mit diesem erforderlich sind. Paradoxaerweise wird dies in der Praxis oft nicht berücksichtigt und es mangelt stark an entsprechenden Best Practices [12]. Der Forschungsbereich Lifelong Learning widmet sich dieser Herausforderung und bringt Methoden hervor, die häufig auf dem Nachtrainieren von KI zu geeigneten Zeitpunkten basieren [5]. Die vorliegende Arbeit sieht hierfür insbesondere die Kombination mehrerer KI-Modelle als Mitglieder eines Ensembles vor, was in diversen anderen Anwendungsfällen bereits gute Performanz gezeigt hat [13]. Wann ein Mitglied hinzugefügt wird, entscheidet dabei ein naturinspirierter Algorithmus. Dieser erkennt auftretende Concept Drifts mittels eines statistischen Detektionsverfahrens, fügt als Folge eines solchen ein neues Mitglied hinzu, beginnt sodann dessen Training auf aktuellen Daten und lässt dessen Vorhersagen mit stetig wachsendem Gewicht in die ensembleweiten Vorhersagen einfließen [14]. Auf diese Weise

kann eine laufende Anpassung der eingesetzten KI auf sich verändernde Bedingungen sichergestellt werden.

Methode zur KI-gestützten Prognose von Durchlauf- und Lieferzeiten

Die hier vorgestellte Methode in kompakter Form basiert auf [1], gewichtet jedoch die Entwicklung robuster KI deutlich stärker. Angelehnt an etablierte Vorgehensmodelle wie CRISP-DM [15] ist sie iterativ und inkrementell angelegt und besteht aus drei aufeinanderfolgenden Schritten (Bild 2).

Zu Beginn erfolgt die Datenvorbereitung. Hier geht es zunächst um die Erschließung geeigneter Datenquellen. Besonders eignen sich Produktdatenmanagement-, Enterprise Resource Planning (ERP) und Manufacturing Execution Systems, da sie Produkt-Prozess-Ressourcendatenströme in strukturierter Form bereitstellen können. Wertvoll für die Prognose der Durchlaufzeiten sind Daten zur Produktbeschaffenheit (wie Material)

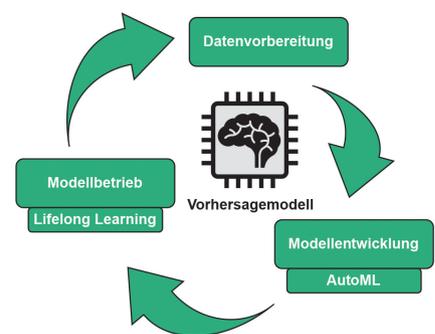


Bild 2. Methode zur KI-gestützten Prognose von Durchlauf- und Lieferzeiten

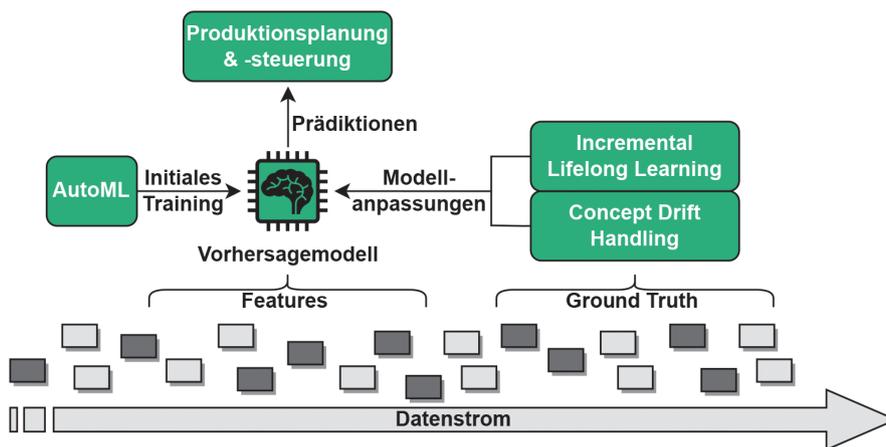


Bild 3. Inferenzprozesssystem im Modellbetrieb

und zugehörige Prozessinformationen (Produktionsschritte, Maschinenparameter). In Bezug auf die Lieferzeiten sind Lokationen und Losgrößen interessant. In der angeschlossenen Transformation werden die Datenströme systematisch aufbereitet und zusammengeführt, um eine qualitativ hochwertige Ausgangsbasis aufzubauen. Dies beinhaltet die Bereinigung von fehlerhaften Daten sowie Ausreißern, Typenkonversionen, Homogenisierung der Zeiteinheiten und Berücksichtigung von Schichtzeiten [16, S. 128], erstreckt sich aber auch auf die explizite Modellierung von zusätzlichen Variablen durch Experten [17]. Der so aufbereitete Datenstrom wird im zweiten Schritt zur Modellentwicklung genutzt. Im Zentrum steht dabei die erwähnte AutoML-Methode Sequential Halving [9]. Für die Modellentwicklung wird zunächst eine Menge von Algorithmus-Hyperparameter-Kombinationen als Ausgangspopulation generiert. Dies geschieht über die Vorauswahl geeigneter Algorithmen sowie obere und untere Schranken für die zugehörigen Hyperparameter. Es wurde gezeigt, dass auf dieser Problemklasse baumbasierte Algorithmen gute Ergebnisse erzielen [1, S. 65–66]. Diese werden entsprechend des beschriebenen Ensemble-Ansatzes kombiniert [14]. Nach dem Training über die Sequential-Halving-Methode [9] bleibt eine Algorithmus-Hyperparameter-Kombination übrig, die das initiale Modell darstellt. An dieser Stelle erfolgt der Übergang in den Modellbetrieb. Um die Performanz des Modells dauerhaft aufrechtzuerhalten,

kommt an dieser Stelle ein Lifelong-Learning-Ansatz zum Tragen. Hierzu inferiert das Modell auf dem eingehenden Datenstrom und prognostiziert entsprechend Durchlauf- oder Lieferzeiten, die in die Produktionsplanung und -steuerung eingehen. Steht zu einem späteren Zeitpunkt nach Ablauf des Vorgangs die tatsächliche Zeit, die Ground Truth, zur Verfügung, greift die beschriebene Concept-Drift-Adaption und das Modell wird bei Bedarf automatisch angepasst [14]. Bild 3 visualisiert dieses Inferenzprozesssystem. Sollte sich nach dem Modellentwicklungsschritt oder auch im Modellbetrieb zeigen, dass die Performanzanforderungen an das Modell grundsätzlich nicht erfüllt werden können, erlaubt der Methodenzyklus einen Neustart, um in weiteren Iterationen sukzessiv Verbesserungen herbeizuführen.

■ Erfolge in der Praxis

Der Ansatz wurde branchenübergreifend in zwei konkreten Anwendungsfällen erfolgreich erprobt. Hierbei wurde jeweils ein Schwerpunkt auf die Entwicklungs- und die Betriebsphase gelegt.

Modellentwicklung in der Metallverarbeitung

Die Methode, und insbesondere der Schritt Modellentwicklung, wurde im Rahmen einer Fallstudie bei einem mittelständischen Einzelfertiger für Schmiedeteile angewendet [1, S. 123–147]. Das Ziel war die Prognose von Prozess- und Übergangszeiten, welche gemeinsam die innerbetrieblichen Durchlaufzeiten bil-

den. Hierzu wurden etwa 100.000 Datenpunkte, welche circa 15.000 Fertigungsaufträge abbilden, aus einem Zeitraum von über zwei Jahren genutzt. Diese lagen zentral in einem ERP-System vor. Mithilfe von Experten aus dem Unternehmen wurden die Daten angereichert, u. a. mit Umformungstemperaturen und spezifischen Gewichten der verwendeten Werkstoffe, um zusätzliche Einflussgrößen auf die Prozessdauern zu modellieren. Ausreißer sowie fehlerhafte oder nicht plausible Werte wurden im Rahmen der Bereinigung entfernt. Auf den so aufbereiteten Datenstrom wurde die oben genannte Sequential-Halving-AutoML-Methode angewendet. Das Modell wurde gegen Expertenschätzungen validiert. Im Ergebnis bei der Prognose der Prozesszeiten konnte der MAE von 57,95 Minuten auf 35,95 Minuten, also um 37,96 Prozent, gesenkt werden. Eine Schätzung der Übergangszeiten war durch das Modell erstmals möglich. Der MAE belief sich hier auf 880,15 Minuten. Hier ist zu beachten, dass die Übergangszeiten deutlich länger sind als die Prozesszeiten, wodurch sich die Unterschiede in der Größenordnung der beiden MAE erklären. Eine Detailanalyse der Fehler auf Prozessartebene hat gezeigt, dass das Modell verschiedene Prozessarten unterschiedlich gut schätzen kann. Beispielsweise lag der MAE bei der Prognose von Prozesszeiten für Sägeprozesse deutlich höher als bei Schmeldeprozessen. Letztendlich ist hier eine breitere Datenbasis und höhere Datenqualität notwendig, um die Modellperformanz noch weiter steigern zu können.

Modellbetrieb in der Möbelbranche

Der Schritt Modellbetrieb wurde im Rahmen einer Fallstudie aus der Möbelbranche mit Ziel der Prognose von Gesamtlieferzeiten erprobt. Zur Modellentwicklung wurden Daten verwendet, die zum Zeitpunkt der Produktbestellung verfügbar werden. Die Merkmale dieser umfassen unter anderem die Anzahl der unterschiedlichen Produkte der Bestellung, Preise sowie Lokationen von versendenden und empfangenden Akteuren [18]. Anschließend begann der Produktiveinsatz des Modells mit dem Ausrollen in der Laufzeitumgebung und erstreckte sich über einen Produktivzeitraum von circa 1,5 Jahren, repräsentiert durch etwa 1,5 Millionen

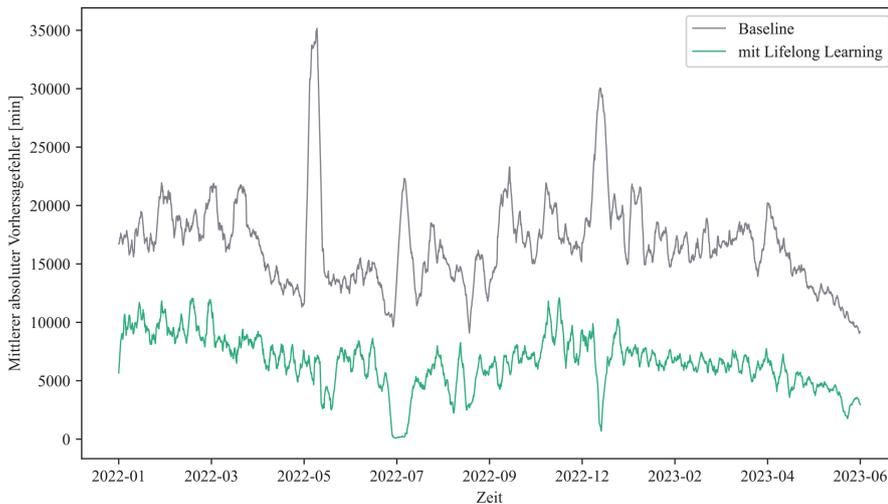


Bild 4. Performanzvergleich von KI ohne (Baseline) und mit Lifelong Learning

Datenpunkte. Als Vergleichsbasis wurde ein einfaches Modell ohne die hier vorgestellten Mechanismen des Lifelong Learning einem komplexen mit diesen gegenübergestellt. Das einfache Modell erreichte einen MAE von 16,70E3 Minuten. Bei genauerer Betrachtung des Fehlerverhaltens fiel auf, dass zahlreiche Phasen existieren, in denen der Fehler häufig um etwa 30,74 Prozent und vereinzelt um bis zu etwa 77,65 Prozent höher ist als zu erwarten wäre. Diese erhebliche Fluktuation ist im hohen Maße Concept-Drift-Ereignissen geschuldet. Die Anwendung des skizzierten Lifelong-Learning-Ansatzes im komplexen Modell führt zu einem MAE von nur 6,61E3 Minuten. Dies entspricht einer Reduktion um ca. 60,42 Prozent gegenüber dem einfachen Modell. Der Fehlerverlauf beider Modelle ist in Bild 4 gezeigt.

Zusammenfassung und Ausblick

Die vorliegende Arbeit skizziert eine Methode zur KI-gestützten Prognose von Durchlauf- und Lieferzeiten. Kerntechnologien sind AutoML zur Modellentwicklung und Lifelong Learning zum Modellbetrieb. Der Einsatz dieser Technologien erlaubt einerseits unternehmensintern eine schnelle Entwicklung von KI-Modellen und sorgt andererseits dafür, dass die negativen Auswirkungen von Concept Drift im Produktivbetrieb minimiert werden. Die Fallstudien in der Metallverarbeitung und der Möbelbranche zeigten,

dass die Methode bei Betrachtung des MAE bessere Prognosen hervorbringt als Experten und diese hohe Performanz der Prognosemodelle auch im langfristigen Betrieb sicherstellt. Zukünftige Arbeiten sollten sich darauf konzentrieren, die Datenqualität in den Unternehmen weiter zu verbessern, AutoML Techniken über das Algorithmus-Hyperparameter-Suchproblem hinaus zu verbreitern und praxisgeeignete Methoden zum Umgang mit Concept Drift aus dem Forschungsstand abzuleiten.

Literatur

- Bender, J.: AutoML-Supported Lead Time Prediction Enabling Smart Job Scheduling in Make-To-Order Production. Dissertation, Karlsruher Institut für Technologie (KIT), Karlsruhe 2024
DOI:10.5445/IR/1000175793
- Gyulai, D.; Pfeiffer, A.; Nick, G. et al.: Lead Time Prediction in a Flow-shop Environment with Analytical and Machine Learning Approaches. IFAC-PapersOnLine 51 (2018) 11, S. 1029–1034
DOI:10.1016/j.ifacol.2018.08.472
- Lingitz, L.; Gallina, V.; Ansari, F. et al.: Lead Time Prediction Using Machine Learning Algorithms: a Case Study by a Semiconductor Manufacturer. Procedia CIRP 72 (2018) 5, S. 1051–1056
DOI:10.1016/j.procir.2018.03.148
- Lim, Z. H.; Yusof, U. K.; Shamsudin, H.: Manufacturing Lead Time Classification Using Support Vector Machine. In: Badioze Zaman, H.; Smeaton, A. F.; Shih, T. K. et al. (Hrsg.): Advances in Visual Informatics. Springer, Cham 2019
DOI:10.1007/978-3-030-34032-2_25
- Gama, J.; Žliobaitė, I.; Bifet, A. et al.: A Survey on Concept Drift Adaptation. ACM Computing Surveys 46 (2014) 4, S. 1–37
DOI:10.1145/2523813
- Zöllner, M.-A.; Huber, M. F.: Benchmark and Survey of Automated Machine Learning Frameworks. Journal of Artificial Intelligence Research 70 (2021), S. 409–472
DOI:10.1613/jair.1.11854
- Feurer, M.; Klein, A.; Eggenberger, K. et al.: Efficient and Robust Automated Machine Learning. In: Proceedings of the 28th International Conference on Neural Information Processing Systems – Bd. 2, NIPS Papers, MIT Press, Cambridge, MA, USA 2015, S. 2755–2763
- Thornton, C.; Hutter, F.; Hoos, H. H.; Leyton-Brown, K.: 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, ACM 2013, S. 847–855
DOI:10.1145/2487575.2487629
- Karnin, Z.; Koren, T.; Somekh, O.: Almost Optimal Exploration in Multi-Armed Bandits. In: Proceedings of the 30th International Conference on Machine Learning, PMLR 28 (2013) 3, S. 1238–1246
- He, X.; Zhao, K.; Chu, X.: AutoML: A Survey of the State-of-the-art. Knowledge-Based Systems 212 (2021)
DOI:10.1016/j.knosys.2020.106622
- Webb, G. I.; Hyde, R.; Cao, H. et al.: Characterizing Concept Drift. Data Mining and Knowledge Discovery 30 (2016) 4, S. 964–994
DOI:10.1007/s10618-015-0448-4
- Vela, D.; Sharp, A.; Zhang, R. et al.: Temporal Quality Degradation in AI Models. Scientific Reports 12 (2022) 1
DOI:10.1038/s41598-022-15245-z
- Krawczyk, B.; Minku, L. L.; Gama, J. et al.: Ensemble Learning for Data Stream Analysis: a Survey. Information Fusion 37 (2017), S. 132–156
DOI:10.1016/j.inffus.2017.02.004
- Trat, M.; Bergmann, P.; Ott, A.; Ovtcharova, J.: A Nature-Inspired Concept Drift Adaptation Method for Industrial Data Stream Regression. In: Wang, Y.-C.; Chan, S. H.; Wang, Z.-H. (Hrsg.): Flexible Automation and Intelligent Manufacturing: Manufacturing Innovation and Preparedness for the Changing World Order. Lecture Notes in Mechanical Engineering, Springer, Cham 2024, S. 3–13
DOI:10.1007/978-3-031-74482-2_1
- Chapman, P.; Clinton, J.; Kerber, R. et al.: CRISP-DM 1.0: Step-by-step Data Mining Guide. SPSS, 2000
- Sauer mann, F.: Datenbasierte Prognose und Planung auftragsspezifischer Übergangszeiten. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2020

18. Schuh, G.; Prote, J.-P.; Hünnekes, P. et al.: Impact of Modeling Production Knowledge for a Data Based Prediction of Transition Times. In: Ameri, F.; Stecke, K.E.; Cieminski, G. von; Kiritsis, D. (Hrsg.): Advances in Production Management Systems. Production Management for the Factory of the Future, IFIP Advances in Information and Communication Technology, Bd. 566, Springer, Cham 2019, S. 341–348
DOI:10.1007/978-3-030-30000-5_43
19. Landmesser, A.: KI in der Möbelbranche – Proof of Concept mit IWofurn. (<https://ai.hdm-stuttgart.de/news/2022/ki-in-der-moebelbranche-poc-mit-iwofurn> [Abgerufen am 30.1.2025])

Die Autor:innen dieses Beitrags

Dr.-Ing. Janek Bender, geb. 1990, studierte Wirtschaftsinformatik und arbeitete als Projektingenieur in der Industrie bevor er 2018 an das FZI Forschungszentrum Informatik in Karlsruhe wechselte. Seit 2021 ist er Abteilungsleiter im Forschungsbereich Intelligent Systems and Production Engineering. Seine Forschungsinteressen liegen im Bereich industrieller Künstlicher Intelligenz zur Produktionsoptimierung, in dessen Rahmen er 2024 am Karlsruher Institut für Technologie promoviert wurde.

Martin Trat, M. Sc., geb. 1991, studierte Wirtschaftsingenieurwesen an der Universität Bayreuth und begann 2020 seine Arbeit als wissenschaftlicher Mitarbeiter am FZI Forschungszentrum Informatik in Karlsruhe. Im Forschungsbereich Intelligent Systems and Production Engineering befasst er sich mit der produktiven Anwendung Künstlicher Intelligenz und mit smarten Strategien zum Umgang mit Concept Drift.

Prof. Dr. Dr.-Ing. Dr. h. c. Jivka Ovtcharova, geb. 1957, promovierte in Maschinenbau und Informatik und ist Expertin auf dem Gebiet des Virtual Engineering. Seit 2003 leitet sie das Institut für Informationsmanagement im Ingenieurwesen am Karlsruher Institut für Technologie. Seit 2004 ist sie Direktorin im Forschungsbereich Intelligent Systems and Production Engineering des FZI Forschungszentrum Informatik.

Abstract

AI-supported Forecasting of Throughput and Delivery Times in Make-to-Order and Small-Series Production – A Holistic Approach.

Predicting throughput and delivery times is challenging, especially in make-to-order and small-series production. This work presents a holistic approach to AI-supported forecasting of these times, which combines techniques from

the fields of Automated Machine Learning and Lifelong Learning to produce lightweight and long-term robust forecasting models. The approach was successfully applied in two case studies in the metalworking and furniture industries.

Schlüsselwörter

Durchlaufzeiten, Lieferzeiten, AutoML, Lifelong Learning, Einzelfertigung

Keywords

Throughput Times, Delivery Times, AutoML, Lifelong Learning, Make-to-Order Production

Bibliography

DOI:10.1515/zwf-2024-0162

ZWF 120 (2025) Special Issue; page 292 – 296

 Open Access. © 2025 bei den Autoren,

publiziert von De Gruyter.  Dieses Werk ist lizenziert unter der Creative Commons Namensnennung 4.0 International Lizenz.

ISSN 0947-0085 · e-ISSN 2511-0896