

DISS. ETH NO. 28031

**ARTIFICIAL INTELLIGENCE IN MANUFACTURING:
AUGMENTING HUMANS AT WORK**

A thesis submitted to attain the degree
DOCTOR OF SCIENCES OF ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

JULIAN SENONER
MSc ETH MTEC, ETH Zurich

born on 08.07.1992
citizen of Austria

accepted on the recommendation of

Prof. Dr. Torbjørn Netland, examiner
Prof. Dr. Stefan Feuerriegel, co-examiner
Prof. Dr. Volodymyr Babich, co-examiner

Abstract

This dissertation presents three scientific studies with a common objective: exploring the capabilities of data-driven decision-making in manufacturing by combining operations management theory with the latest methodological developments in artificial intelligence (AI). All three studies are conducted with industrial collaborators covering different applications of AI in manufacturing. Leveraging this empirical research design enables validating the potential of state-of-the-art AI methods while reporting actual operational improvements made in real-world factories. Thereby, this dissertation provides theoretical contributions to the literature on operations management and important implications for practice.

The first study of this dissertation develops a data-driven decision model to improve process quality in manufacturing. A challenge for traditional methods in quality management is to handle high-dimensional manufacturing data with nonlinear relationships. This challenge is addressed by using nonlinear modeling with SHAP values to infer how a set of production parameters and the process quality of a manufacturing system are related. Thereby, this dissertation contributes a measure of process importance based on which manufacturers can prioritize processes for quality improvement. Grounded in quality management theory, the proposed decision model selects improvement actions that target the sources of quality variation. The decision model was validated at a semiconductor factory of Hitachi ABB. Seeking to improve production yield, the decision model was applied to select improvement actions for a transistor chip product. The effectiveness of the chosen improvement actions was confirmed in a field experiment. Compared to the average yield in the sample, the experiment returned a reduction in yield loss of 21.7%. Overall, the first study demonstrates the operational value of explainable AI by showing that critical drivers of process quality can go undiscovered by the use of traditional methods.

The second study analyzes whether reducing the opacity of AI algorithms improves the task performance of production workers in a real-world inspection task. To test this hypothesis, two preregistered experiments were conducted in which participants were supported by either (a) black-box AI or (b) explainable AI. The first experiment was conducted online with $N = 288$ participants recruited from Amazon Mechanical Turk. The second experiment was conducted in the field with $N = 48$ production workers from an electronics factory of Siemens. The results of the study show that participants supported by explainable AI outperformed those supported by black-box AI because they were more likely to follow AI predictions when they were accurate and overrule them when they were wrong. Overall, explainable AI improved the task performance by 6.3 percentage points in the online experiment and by 7.7 percentage points in the field experiment compared to black-box AI.

The third study explores how predictive models can be used to improve scheduling accuracy in manufacturing settings with customized products. Customization often leads to shifts in the probability distribution of operational data (so-called distributional shifts). This can harm the performance of predictive models when deployed to future customer orders with new specifications. To improve scheduling accuracy in the presence of distributional shifts, this dissertation proposes

the use of adversarial learning to predict order throughput times while simultaneously accounting for distributional shifts between different customer orders. The proposed adversarial learning model was empirically validated in a real-world job shop production at Aker Solutions. Based on operational data from two consecutive customer order settings, the study confirms that adversarial learning outperforms several data-driven baseline models in predicting order throughput times. Moreover, it is established that the improved prediction performance from adversarial learning can reduce the cost of planning errors by 12.6% to 18.3% relative to an expert baseline.

This dissertation makes eight main contributions. First, it proposes and validates a theory-informed measure of process importance based on which manufacturers can prioritize processes for quality improvement. Second, it gives behavioral insights into how explainable AI can improve task performance in human-AI collaboration. Third, it demonstrates how operations managers can address distributional shifts in operational data. Fourth, it presents experimental evidence from real-world factories demonstrating that AI is effective for productivity improvement. Fifth, it offers insights into how AI can augment human work in manufacturing. Sixth, it contributes a general overview of how manufacturers can use AI to improve their operational performance. Seventh, it discusses the challenges of implementing AI in manufacturing. Eighth, it provides recommendations for manufacturing practitioners.

The findings from this research are likely to be generalizable across different manufacturing industry segments and are applicable beyond manufacturing. Overall, this dissertation provides a basis for future empirical research on AI in manufacturing and guides manufacturers in their AI journey.

Zusammenfassung

Diese Dissertation präsentiert drei wissenschaftliche Studien mit einem gemeinsamen Ziel: die Erforschung von datengetriebener Entscheidungsfindung in der Produktion durch die Kombination der Operations Management Theorie mit den neuesten methodischen Entwicklungen im Bereich der künstlichen Intelligenz (KI). Alle drei Studien decken verschiedene Produktionsanwendungen von KI ab und wurden mit verschiedenen Industriepartnern durchgeführt. Dieses empirische Forschungsdesign ermöglicht es, das Potenzial modernster KI-Methoden zu validieren und gleichzeitig Verbesserungseffekte in realen Fabriken zu messen. Damit liefert diese Dissertation theoretische Beiträge zur Operations Management Literatur und wichtige Implikationen für die Praxis.

Die erste Studie dieser Dissertation entwickelt ein datengetriebenes Entscheidungsmodell zur Verbesserung der Prozessqualität in Fertigungssystemen. Eine bestehende Herausforderung für traditionelle Methoden im Qualitätsmanagement ist der Umgang mit hochdimensionalen Fertigungsdaten mit nichtlinearen Zusammenhängen. Diese Herausforderung wird durch die Verwendung nichtlinearer Modellierung mit SHAP-Werten adressiert, um abzuleiten, wie gewisse Produktionsparameter und die Prozessqualität eines Fertigungssystems zusammenhängen. Dabei schlägt diese Dissertation ein neues Maß für Prozessrelevanz vor, anhand dessen Produktionsunternehmen Prozesse zur Qualitätsverbesserung priorisieren können. Basierend auf der Qualitätsmanagementtheorie ermittelt das vorgeschlagene Entscheidungsmodell Verbesserungsmaßnahmen, die auf die Ursprünge von Qualitätsschwankungen abzielen. Das Entscheidungsmodell wurde in einer Halbleiterfabrik von Hitachi ABB validiert. Um die Produktionsausbeute für ein Transistorchipprodukt zu verbessern, wurde das Entscheidungsmodell angewendet, um Verbesserungsmaßnahmen zu ermitteln. Die Wirksamkeit der ermittelten Verbesserungsmaßnahmen wurde in einem Feldexperiment bestätigt. Verglichen mit der durchschnittlichen Ausbeute in der ursprünglichen Datengrundlage ergab das Experiment eine Verringerung des Ausbeuteverlusts von 21.7%. Insgesamt demonstriert die erste Studie den operativen Wert von erklärbarer KI, indem sie zeigt, dass kritische Prozessqualitätstreiber durch den Einsatz traditioneller Methoden unentdeckt bleiben können.

Die zweite Studie analysiert, ob die Reduktion der Intransparenz von KI-Algorithmen die Leistung von Produktionsmitarbeitern in einer realen Inspektionsaufgabe verbessern kann. Um diese Hypothese zu testen, wurden zwei vorregistrierte Experimente durchgeführt, bei denen die Teilnehmer entweder durch (a) Black-Box-KI oder (b) erklärbare KI unterstützt wurden. Das erste Experiment wurde online mit $N = 288$ Teilnehmern durchgeführt, welche über Amazon Mechanical Turk rekrutiert wurden. Das zweite Experiment wurde im Feld mit $N = 48$ Produktionsmitarbeitern einer Elektronikfabrik von Siemens durchgeführt. Die Ergebnisse der Studie zeigen, dass Teilnehmer, die von erklärbarer KI unterstützt wurden, bessere Leistungen erbrachten als diejenigen, die von Black-Box-KI unterstützt wurden. Der Grund dafür ist, dass Teilnehmer, die von erklärbarer KI unterstützt wurden, mit größerer Wahrscheinlichkeit KI-Vorhersagen befolgten, wenn sie richtig waren, und sie überstimmten, wenn sie falsch waren. Insgesamt verbesserte

erklärbare KI die Leistung im Online-Experiment um 6.3 Prozentpunkte und im Feldexperiment um 7.7 Prozentpunkte im Vergleich zur Black-Box-KI.

Die dritte Studie untersucht, wie KI-Modelle verwendet werden können, um die Planungsgenauigkeit in Fertigungssystemen mit kundenspezifischen Produkten zu verbessern. Kundenspezifikationen führen häufig zu Verschiebungen in der Wahrscheinlichkeitsverteilung von operativen Daten (sog. Verteilungsverschiebungen). Dies kann die Vorhersagegenauigkeit von KI-Modellen beeinträchtigen, wenn für zukünftige Kundenaufträge mit neuen Spezifikationen Vorhersagen getroffen werden müssen. Um die Planungsgenauigkeit bei Verschiebungen in der Wahrscheinlichkeitsverteilung zu verbessern, schlägt diese Dissertation die Verwendung von Adversarial Learning vor, um Auftragsdurchlaufzeiten vorherzusagen, während gleichzeitig Verteilungsverschiebungen zwischen verschiedenen Kundenaufträgen berücksichtigt werden. Das vorgeschlagene Adversarial Learning Modell wurde empirisch in einer realen Werkstattfertigung bei Aker Solutions validiert. Basierend auf operativen Daten von zwei aufeinanderfolgenden Kundenaufträgen bestätigt die Studie, dass das vorgeschlagene Adversarial Learning Modell im Vergleich zu mehreren datengetriebenen Vergleichsmodellen bessere Vorhersagen für Auftragsdurchlaufzeiten liefert. Darüber hinaus wurde festgestellt, dass die verbesserte Vorhersageleistung durch Adversarial Learning die Kosten von Planungsfehlern um 12.6% bis 18.3% im Vergleich zu Expertenvorhersagen reduzieren kann.

Diese Dissertation leistet acht Hauptbeiträge. Erstens wird ein theoretisches Maß für Prozessrelevanz vorgeschlagen und validiert, auf dessen Grundlage Produktionsunternehmen Prozesse zur Qualitätsverbesserung priorisieren können. Zweitens gibt diese Dissertation Verhaltenseinblicke darüber, wie erklärbare KI die Leistung in der Mensch-KI-Kollaboration verbessern kann. Drittens wird demonstriert, wie Operations Manager Verteilungsverschiebungen in operativen Daten adressieren können. Viertens präsentiert diese Dissertation experimentelle Ergebnisse aus realen Fabriken, die belegen, dass KI zur Produktivitätssteigerung effektiv ist. Fünftens bietet die Dissertation Einblicke, wie KI die menschliche Arbeit in der Fertigung augmentieren kann. Sechstens bietet diese Dissertation einen allgemeinen Überblick darüber, wie Produktionsunternehmen KI nutzen können, um ihre Betriebsleistung zu verbessern. Siebtens werden die Herausforderungen bei der Implementierung von KI in der Produktion erörtert. Achters gibt diese Dissertation Empfehlungen für Praktiker, die in der Produktion tätig sind.

Die Ergebnisse dieser Forschung sind auf verschiedene Segmente der Produktionsindustrie übertragbar und sind über die Produktion hinaus anwendbar. Insgesamt bietet diese Dissertation eine Grundlage für zukünftige empirische Forschung zu KI in der Produktion und unterstützt Produktionsunternehmen bei der Implementierung von KI.