

Abstract

Leveraging data mining for a more intelligent manufacturing process is a top priority in today's industry to get ahead in the transformation process towards industry 4.0. Because of this, large amounts of data is generated, but perfectly fitting algorithms for the given data structures and wanted results are a huge challenge and often not yet ready for application in industrial engineering. This thesis addresses two highly relevant use cases which involve data mining on real data from an industrial manufacturing process. The first use case, which is to measure the influence of components to localize reasons for production problems with methods revolving around Principal Component Analysis (PCA). The second use case is to research the application of the Hidden Markov Model (HMM), to predict quality scores on production data with sequential character. It is shown that PCA is a useful tool to analyze features of production data, whereas the Hidden Markov Model (HMM) can be used to predict on production data, but only under difficult circumstances, such that the model is not suitable for this use case. Consequently, this work is concluded with a research proposal for a new model and method for the second use case.

Keywords: Industry · Data Mining · Sensor Data · Graph Data · Graph Models · Dimension Reduction · Feature Selection · Scoring · Prediction · Principal Component Analysis (PCA) · Hidden Markov Model (HMM) · Gaussian Mixture Model (GMM) · Artificial Neural Network (ANN) · Graph Neural Network (GNN)

Zusammenfassung

Die Nutzung von Data Mining für einen intelligenteren Fertigungsprozess hat in der heutigen Industrie höchste Priorität, um im Transformationsprozess in Richtung Industrie 4.0 voranzukommen. Aus diesem Grund werden große Datenmengen erzeugt, aber perfekt passende Algorithmen für die gegebenen Datenstrukturen und die gewünschten Ergebnisse sind eine große Herausforderung und oft noch nicht bereit für die Anwendung in der Industrietechnik. Diese Arbeit befasst sich mit zwei hoch relevanten Anwendungsfällen, bei denen es um Data Mining auf realen Daten aus einem industriellen Fertigungsprozess geht. Der erste Anwendungsfall ist die Messung des Einflusses von Komponenten zur Lokalisierung von Gründen für Produktionsprobleme mit Methoden, die sich um Principal Component Analysis (PCA) drehen. Der zweite Anwendungsfall ist die Untersuchung der Anwendung des Hidden Markov Model (HMM) zur Vorhersage von Qualitätskennzahlen auf Produktionsdaten mit sequentiellem Charakter. Es wird gezeigt, dass PCA ein nützliches Werkzeug zur Analyse von Merkmalen von Produktionsdaten ist, während Hidden Markov Model (HMM) zur Vorhersage auf Produktionsdaten verwendet werden kann, aber nur unter schwierigen Umständen, sodass das Modell für diesen Anwendungsfall nicht geeignet ist. Folglich wird diese Arbeit mit einem Forschungsvorschlag für ein neues Modell und eine neue Methode für den zweiten Anwendungsfall abgeschlossen.

Schlüsselwörter: Industrie · Data Mining · Sensor Daten · Graph Daten · Graph Modelle · Dimensionsreduktion · Merkmalsselektion · Bewertung · Prognose · Principal Component Analysis (PCA) · Hidden Markov Model (HMM) · Gaussian Mixture Model (GMM) · Artificial Neural Network (ANN) · Graph Neural Network (GNN)