Trustworthy anomaly detection for smart manufacturing

Li, Z.

Abstract

The widespread integration of artificial intelligence (AI) technology into manufacturing processes has fostered the evolution of what is now termed *smart manufacturing*. Smart manufacturing leverages data from various sources, including sensors, event logs, and more, to optimize operations through the use of advanced machine learning techniques. A critical component of smart manufacturing is anomaly detection, which aims to identify abnormal patterns in data that may indicate machine malfunctions or system faults, facilitating predictive maintenance. However, the application of anomaly detection to real-world manufacturing environments presents challenges, such as complex data types, high-dimensionality, explainability, generalizability, and automatability. This dissertation addresses these challenges by focusing on *trustworthy anomaly detection*, approached from both *data-centric AI* and *model-centric AI* perspectives.

Data-centric AI emphasizes the importance of data quality, consistency, and richness, while model-centric AI focuses on developing effective algorithms with existing, often fixed, datasets. This dissertation argues that these two paradigms are complementary. While recent trends have shifted towards data-centric AI, we contend that both approaches are crucial for advancing smart manufacturing. Solving the challenges associated with anomaly detection requires understanding both the algorithms used and the insights extracted from complex, large datasets. Therefore, this work adopts a holistic approach, addressing the challenges from both perspectives to enhance the reliability and efficiency of smart manufacturing systems.

This dissertation is structured around two primary research questions, each targeting different aspects of anomaly detection in smart manufacturing. The first research question addresses how to develop and improve anomaly detection from a data-centric AI perspective. In real-world manufacturing, data can take various forms, such as event logs, time series, and graphs. To address this complexity, the dissertation pro-

poses a novel graph-based anomaly detection method that converts event logs into directed, weighted graphs. This method enables unsupervised anomaly detection and provides explanations for each detected anomaly (Chapter 2). Furthermore, a feature selection method is developed to handle high-dimensional data, improving the accuracy of traditional log anomaly detection methods by identifying and reducing redundant log events (Chapter 3).

The second research question examines anomaly detection from a model-centric AI perspective, addressing issues related to explainability, generalizability, and automatability. The dissertation introduces a novel approach using Quantile Regression Forests for inherently interpretable contextual anomaly detection (Chapter 4). It also explores the vulnerability of post-hoc explanation methods for graph neural networks (GNNs) to adversarial attacks, developing an optimization-based adversarial attack method to test the robustness of post-hoc GNN explanations (Chapter 5). Additionally, a cross-domain anomaly detection method is proposed to improve the generalizability of models across different datasets when encountering concept drift (Chapter 6), and a novel strategy for automated hyperparameter tuning in unsupervised anomaly detection systems is presented to enhance their reliability and performance without requiring labeled data (Chapter 7).

Through these contributions, the dissertation provides a comprehensive framework for advancing trustworthy anomaly detection in smart manufacturing. By integrating data-centric and model-centric AI approaches, it offers novel solutions that improve anomaly detection accuracy, interpretability, robustness of explanations to adversarial attacks, generalizability across domains, and automatability in deployment. These advancements have the potential to substantially enhance the reliability and efficiency of smart manufacturing systems, while the insights gained from this research are also applicable to other domains where anomaly detection plays a critical role.

In conclusion, this dissertation advances the state-of-the-art in trustworthy anomaly detection for smart manufacturing by addressing key challenges through a dual focus on data-centric and model-centric AI. The proposed methods contribute to improved accuracy, explainability, robustness of post-hoc explanations, and automatability of anomaly detection systems, ensuring their applicability in complex, high-dimensional, and dynamic real-world manufacturing environments.