Trustworthy anomaly detection for smart manufacturing

Li, Z.

Stellingen

Behorende bij het proefschrift

Trustworthy Anomaly Detection for Smart Manufacturing

1. Transforming log data into graphs enables more accurate anomaly detection by capturing both quantitative and structural dependencies, improving fault analysis in high-tech systems. (Chapter 2)

2. Selecting a small subset of relevant log events for anomaly detection and prediction improves both accuracy and efficiency, demonstrating that exhaustive analysis of all log events is unnecessary in complex manufacturing systems. (Chapter 3)

3. Contextual anomaly detection, which differentiates between contextual and behavioral features, enhances the identification of meaningful anomalies compared to traditional methods that treat all features equally. (Chapter 4)

4. Post-hoc explanation methods for Graph Neural Networks are highly vulnerable to adversarial perturbations, bringing into question their reliability for decision-critical applications and highlighting the need for robustness in explainable AI. (Chapter 5)

5. Leveraging labeled data from related domains can significantly improve the accuracy of graph-level anomaly detection, mitigating the limitations of purely unsupervised methods in smart manufacturing. (Chapter 6)

6. Truly unsupervised graph anomaly detection is achievable without labeled data by leveraging self-supervised learning techniques and internal evaluation metrics. (Chapter 7)

7. Achieving trustworthy anomaly detection in smart manufacturing requires balancing detection effectiveness, explainability, and robustness, as overemphasizing any single factor can lead to sub-optimal decision-making.

8. The integration of model-centric and data-centric approaches is essential for improving the reliability of anomaly detection in industrial applications, as algorithmic advancements alone cannot compensate for poor data quality.

9. The lack of labeled anomalies in smart manufacturing necessitates the development of unsupervised and semi-supervised anomaly detection methods, yet their effectiveness is highly dependent on domain knowledge and data representation.

10. The interpretability of anomaly detection models is crucial for adoption in industrial settings, as black-box models hinder root cause analysis and limit trust in AI-driven decision-making.

11. Trustworthy AI principles should be mandated for all high-stakes anomaly detection applications, including healthcare and finance, as the cost of untrustworthy AI may far outweigh its benefits.

12. The perception of AI as an infallible decision-maker is flawed; rather than replacing humans, AI should be designed to augment human decision-making, particularly in critical applications such as anomaly detection.

Zhong Li

Leiden, 1 May 2025