Contribution to explainability of Anomaly Detection model for Predictive Maintenance in Industry

Le Hoang Nguyen[1, 2,*], Kim Duc Tran[1, 2], Sebastien Thomassey[2], Kim Phuc Tran[2]

[1] *University of Lille, ENSAIT, GEMTEX - Laboratoire de Génie et Matériaux Textiles, F-59000 Lille, France*
[2] *University of Dong A, IAD, Viet Nam*
* *le-hoang.nguyen@ensait.fr*

The objective of this research is to develop and apply advanced technologies in the fields of artificial intelligence (AI) and automation to improve the efficiency of Predictive Maintenance (PdM) in smart factories, in line with the human-centered orientation of the Industrial Revolution 5.0. One of the major challenges in applying AI to predictive maintenance today is the requirement to respond in real time, ensure privacy for sensitive data, and maintain the explainability of deep learning models. Although many studies have achieved promising results in the field of PdM in recent years, there is still a significant gap between research solutions and practical application. In real industrial environments, system behavior may change over time due to factors such as natural wear and tear or system upgrades, leading to models trained on historical data gradually becoming less effective. To address this issue, the study proposes the implementation of online learning, incremental retraining or domain adaptation techniques to maintain the model's adaptability and efficiency over time. In addition, post-hoc explanation methods often reflect only correlations rather than true causes. In the context of predictive maintenance, root cause analysis is key, and post-event explanation methods may miss actual causes that occurred earlier or reside in other subsystems. Therefore, the study proposes applying “Explainable-by-Design” (XbD) approaches, or building hybrid models, such as combining physics-based modeling with AI. Techniques such as gray-box models or Physics-Informed Neural Networks are considered potential solutions to improve explainability and understanding of causal relationships in systems. Finally, to protect data privacy and security during training, the study proposes to apply federated learning as an effective solution, allowing the construction of collaborative models without centralizing sensitive data, thereby enhancing security and meeting the practical requirements of modern industrial systems.

Within the framework of this study, we propose a sequential implementation roadmap to integrate advanced technologies into the predictive maintenance system as follows: applying modern techniques such as Deep Learning and XbD artificial intelligence, combined with Edge Devices in the Federated Learning model. In addition, to support real-time system visualization and monitoring, Digital Twin technology is also integrated into the predictive maintenance process. The goal of this roadmap is to reduce processing latency, ensure data privacy, improve reliability in interpreting results, and enhance system operational readiness. The proposed solution aims to detect anomalies early and accurately predict the remaining useful life of equipment in smart factories.

In the first phase of our research, we built a Test-bed to evaluate the performance of anomaly detection in the context of predictive maintenance. The testbed consists of three Raspberry Pi devices, which act as edge devices, responsible for monitoring potential anomalies or faults arising from sensors mounted on the motors. Each motor is equipped with a sensor system that measures temperature, vibration, and sound at three different locations, providing real-time data for prediction. At the edge devices, machine learning models are trained locally, and then the model parameters are synchronized with the central server through Federated Learning. We evaluate the performance of the models based on the accuracy of anomaly detection, and also check the data security during the training process. In addition, the XbD approach is integrated from the beginning through prototypes, ensuring transparency and explainability of the prediction results. To fully realize the solution, we develop a Digital Twin that allows simulation and monitoring of the entire system operation in real time. This mechanism not only effectively supports the Root Cause Analysis process, but also contributes to optimizing the overall operational performance of the equipment.

References:

[1] D.H.Nguyen and K.P.Tran. Artificial intelligence for smart manufacturing in industry 5.0: Methods, applications, and challenges. Artificial Intelligence for Smart Manufacturing: Methods, Applications, and Challenges, pages 5–33, 2023

[2] Ahn, Jisu and Lee, Younjeong. Federated learning for predictive maintenance and anomaly detection using time series data distribution shifts in manufacturing processes. 2023

[3] Cummins, Logan and Sommers. Explainable predictive maintenance: A survey of current methods, challenges and opportunities, pages 57574—57602, 2024

[4] Van Dinter, Raymon and Tekinerdogan. Predictive maintenance using digital twins: A systematic literature review, page 107008, 2022.

[5] Soares, Eduardo Almeida, Explainable-by-design deep learning. Book, 2022