Abstract

Modern manufacturing relies on high efficiency, strict product quality, and continuous equipment availability to remain competitive. However, manufacturing environments can be subject to critical problems that significantly impact production. Problems including product defects and equipment failures can occur suddenly without warning and could lead to high production losses. Early identification of these problems is crucial for maintaining production efficiency.

Industry 4.0 technologies enable continuous real-time data collection from production lines. Artificial Intelligence (AI) models can learn from these data to identify signs of problems and prevent them before they occur. However, manufacturing environments generate diverse data types including images, numerical tabular data, and multivariate time series. Various data types usually require different model architectures, making it difficult to design a unified solution. Beyond prediction, industrial operators must be able to understand and trust model decisions in order to take appropriate corrective or preventive actions. Without this transparency, even accurate models are difficult to adopt in practice. The need for Explainable Artificial Intelligence (XAI) in complex AI models has therefore become a significant challenge in industrial environments. The choice of explainability techniques depends on both the data and the model types, with each technique having a different way of representing explanations.

This thesis proposes a unified solution that addresses data diversity by unifying both the processing of different data types and the decision explanation approach. The solution leverages advances in AI models and explainability techniques from the image processing domain and adapts them to numerical and temporal data. We propose the **Heatmap-Based Deep Neural Network (HDNN)** approach, which consists of transforming input from numerical tabular data and multivariate time series into heatmap image representations to feed them into image-based models. The models are trained to detect local and global patterns related to problems in the heatmap input image. We then apply explainability techniques from image processing to highlight the most important regions of the heatmap. By superimposing the saliency map explanations on the raw input data, operators can directly identify the parameters and their values corresponding to these regions. This approach provides practical and efficient visual explanations for operators.

We initially used Convolutional Neural Network (CNN)-based models in the HDNN approach, which demonstrated promising results. Then, we proposed the **Heatmap Vision Tabular Transformer (HViTT)** architecture, an adaptation of the Vision Transformer to industrial tabular and temporal data to better analyze features in heatmap inputs. Finally, we propose an XAI method selection methodology based on a comparative analysis of image-based XAI techniques for a given use case. This contribution introduces **TRansparency Alpha-Mask Faithfulness Evaluation (TRAFE)**, a dedicated method for the HViTT architecture to evaluate the faithfulness of XAI techniques.

The proposed solution has been validated on multiple industrial use cases for classification and regression tasks, including quality control and predictive maintenance applications. The obtained results demonstrate the effectiveness of our approach, showing competitive model performance and practical model decision explanations.