KU LEUVEN

FACULTY OF ENGINEERING SCIENCE

Lorenzo Perini

Department	Computer Science
PhD defence	28 March 2024
Supervisor	Jesse Davis
Funding	FWO and FAIR
E-mail	lorenzo.perini@kuleuven.be

Operational, Uncertainty-Aware, and Reliable Anomaly Detection

Blade icing in wind turbines

Introduction / Objective

Anomaly detection automatically identifies instances deviating from normal behavior, like detecting blade icing in wind turbines. Unsupervised machine learning relies on heuristics to assign anomaly scores, but these can be hard to interpret and may introduce uncertainty.
Can we design an anomaly detection system that is *Operational, Uncertainty-Aware, and Reliable*?

Research Methodology

This thesis makes significant contributions to addressing the posed question. First, we make a detector operational by designing approaches to properly threshold the anomaly scores without recurring to labeled data. Second, we propose to quantify a detector's uncertainty in predictions by introducing a Bayesian framework. Finally, we allow the model to abstain (i.e., to return "I don't know") whenever its uncertainty is high without requiring any labeled example.

In a summary, we propose to:

1. *transform the anomaly scores to actionable predictions* by properly setting a **decision threshold** with no labels (Fig. 1);
2. *measure an anomaly detector's uncertainty* in predictions in terms of **stability under data perturbations** (Fig. 2);
3. *allow the anomaly detector to output "**I don't know**"* when its uncertainty is high to increase the user trust (Fig. 3).

Operational

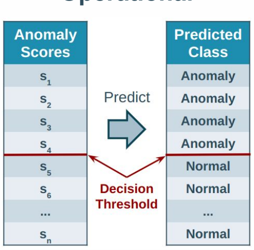

Fig. 1: A decision threshold can transform the scores into actionable predictions.

Uncertainty-Aware

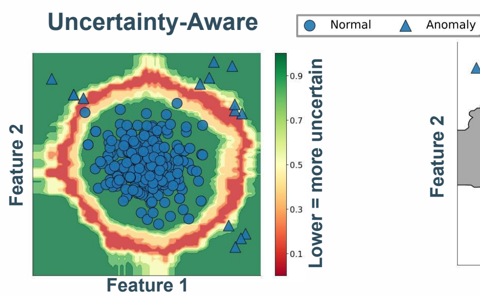

Fig. 2: The red area around the decision boundary is an indication of high uncertainty.

Reliable

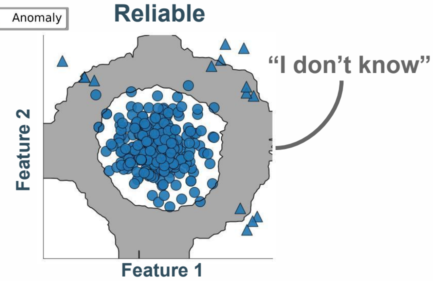

Fig. 3: The grey area around the decision boundary indicates where the model abstains.

Results & Conclusions

Anomaly detection is a relevant task because anomalies are often connected to high costs (monetary, environmental,...). Existing literature has three relevant gaps, namely how to operationalize an anomaly detector, how to capture its uncertainty, and how to increase the user trust in the system. Our proposed approaches fill these gaps, allowing practitioners to use and rely on the chosen anomaly detection model.

Major publications

Perini, L., Burkner, P., Klami, A. Estimating the contamination factor's distribution in unsupervised anomaly detection, ICML 2023.

Perini, L., Vercruyssen, V., and Davis, J. Quantifying the confidence of anomaly detectors in their example-wise predictions, ECML 2020.

Perini, L., and Davis, J. Unsupervised Anomaly Detection with Rejection, NeurIPS 2023.