Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

ANOMALY DETECTION IN IOT-BASED MANUFACTURING USING MACHINE LEARNING AND PROCESS MINING: AN INTEGRATIVE TECHNIQUE (FRAMEWORK)

[1]Muhammad Umair Warsi, [2]Dr. Shafi Ullah, [3]Dr. Mohammad Imran

[1]Department of Computer Science, Faculty of Information and Communication Technology (FICT), Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan. ORCID: https://orcid.org/0009-0006-9077-9485

[2]Department of Computer Science, Faculty of Information and Communication Technology (FICT), Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan

[3]Department of Computer Science, Faculty of Information and Communication Technology (FICT), Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan

[1]uwarsi124@gmail.com

DOI: https://doi.org/10.5281/zenodo.21431930

Keywords:
IoT-enabled manufacturing; Industry 4.0; process mining; machine learning; hybrid anomaly detection; Tennessee Eastman Process; conformance checking.

Article History
Received on: 14 Feb, 2026
Accepted on: 28 March, 2026
Published on 31 March, 2026

Copyright @Author
Corresponding Author:

Abstract

This study proposes and empirically evaluates a hybrid approach of Machine Learning (ML) and Process Mining (PM) to solve the problem of anomaly detection in IoT-enabled manufacturing systems in a way that optimises simultaneously the accuracy of the fault diagnosis, its interpretability, and its operational reliability. The framework combines data-driven statistical intelligence and process-aware behavioural reasoning to find deviations in real-time industrial workflows by using a dataset from the Tennessee Eastman Process (TEP) and a Kaggle oil-and-gas sensor proxy. The experimental pipeline was built in Python 3.11, with the use of scikit-learn, TensorFlow and PM4Py on a Google Colab GPU environment. The five ML models were trained and validated with the stratified 5-fold cross validation technique, namely, Support Vector Machine (RBF kernel), Random Forest, Multilayer Perceptron (MLP), Autoencoder and Isolation Forest. Partitioning models have significantly lagged behind the performance of Random Forest and MLP which classified 100% (Precision = Recall = F1 = 1.000) of the results, thus confirming their robustness in critical fault identification. Process mining, with Directly-Follows Graphs and conformance checking, provided contextual knowledge about the deviations in the process: 76.9% of process traces contained micro anomalies in otherwise normal process behavior. A hybrid fusion model that added the ML and PM outputs by a conservative logical AND rule had perfect precision (1.000) and no false positive but at the expense of lower recall (0.133) which is a trade-off suitable for safety-critical applications where false alarms have a high cost. The system's robustness and resilience were tested under the assumption of sensor noise following a Gaussian distribution and simulated data losses, and the results demonstrated resilience (F1 ≥ 0.95) for up to 10% missing data and end-to-end latency of inferences was kept under millisecond to sub-second, supporting the real time deployment of the system. The study presents an interpretable, precision optimised architecture that integrates the predictive power of ML with the behaviour transparency of PM, providing a methodological basis for digital twins and predictive maintenance in Industry 4.0 manufacturing systems that can be scaled.

INTRODUCTION

Background and Problem Statement

With the advent of the Internet of Things (IoT), manufacturing has been changed as it is now possible to monitor continuously with high resolution across various devices and systems, allowing the operation to be the backbone of Industry 4.0 (Soori et al., 2023). While this connectivity can lead to significant improvements in productivity, quality control, and prediction, it also introduces a new challenge to an old nemesis: the ability to accurately recognize anomalies.This connectivity can bring a lot of benefits in terms of productivity, quality control and prediction, but it also comes with a new challenge to an age-old nemesis: accurately identifying anomalies or deviations from normal operation that could indicate inefficiency, error or imminent failure (Pillai, 2022). Traditional rule-based monitoring methods are not really able to cope with the volume, velocity and dimensionality of IoT data, and are often too prescriptive to discern between normal variation and real fault states.

Both Machine Learning (ML) and Process Mining (PM) have proven to be useful in this area. On the one hand, ML models are capable of scaling to a high number of dimensions when dealing with sensor data for pattern recognition and outlier detection (Al-Amri et al., 2021), and on the other hand, PM techniques make use of event information to uncover inefficiencies in the structure and deviations in behaviour of business and operational processes (Maeyens et al., 2020). In the field of industrial anomaly detection, however, the two methods have rarely been used hand-in-hand, as they each complement one another with respect to their strengths and weaknesses: ML, at the one end, is statistically precise and has been the focus of a broad range of studies, while, at the other end, PM is behaviourally rich and less sensitive in providing predictive power. The Tennessee Eastman Process (TEP), a benchmark simulated chemical plant with known non-linear dynamics and multiple multivariate sensor streams and a known library of fault conditions (Xu et al., 2022), is an ideal testbed for an integrated ML–PM solution.

In this study, to bridge this gap, we designed, implemented, and evaluated a hybrid framework of ML and PM for an IoT-based manufacturing system, with the help of a TEP inspired dataset. Industrial data is often complex, noisy and time-synchronous, and an event that is purely statistical or purely rule-based fails to work properly, resulting in too many false alarms or missing early-stage faults (Pillai, 2022; Anton et al., 2019). The framework that is proposed will aim to cross-validate statistical anomaly scores with deviations in process behaviour, thus providing detections which are at the same time accurate and explainable.

Research Objectives and Questions

The five objectives of this study are: (1) the design and implementation of an integrative ML–PM anomalydetection framework for manufacturing enabled with IoT; (2) comparison of the performance of SVM, Random Forest, MLP, Autoencoder and Isolation Forest models on the TEP-inspired dataset; (3) application of process-mining techniques — event-log construction, conformance checking and deviation analysis — for identifying behavioural fault patterns; (4) comparative evaluation of the accuracy, false-positive rate, interpretability, and real-time applicability of the standalone ML, standalone PM and hybrid ML–PM configurations; and (5) outlining a reporting pipeline for translating detection results into actionable dashboards, alerts and maintenance triggers.

The study thus poses several questions: (i) how to effectively combine ML and PM for anomaly detection in IoT-based manufacturing; (ii) which of the ML models is best suited to perform anomaly detection on the TEP with respect to accuracy, efficiency, and robustness; (iii) which behaviour insight does PM add beyond the statistical detection; (iv) how does this combined framework perform compared to baselines using only ML, only PM, and rules; and (v) how to operationalise the detection output for maintenance and safety decision-making.

Significance and Scope

The study joins a literature in which ML and PM have mostly been studied separately (Fernandes et al., 2019), providing a systematic, and empirically approved integration. The framework is designed for real-life deployment in industry, rather than just academic benchmarking, by focusing on low false positive rate, real-time latency and interpretability, and by using the TEP as an example case study which enables reproducibility and comparison with previous studies. The scope of the study is intentionally limited: it focuses on the anomaly detection problem (and not the more general problems of predictive maintenance or quality control), the evaluation is performed on a controlled, historical TEP-inspired dataset and not in a live industrial setting, and the process-mining part is limited to the type of analyses that are possible given the event logs. Findings

should thus be understood as indicative of feasibility and relative performance and not as a field deployment validation.

LITERATURE REVIEW

As the increasing amount and complexity of IoT generated manufacturing data renders many conventional statistical process-control (SPC) methods inadequate under nonstationary, nonlinear conditions typical of cyber-physical systems, this work has spurred a growing body of literature on hybrid ML–PM frameworks (Shkarupylo et al., 2024; Lakshmi et al., 2023). Presciuttini et al. (2024) point out that while ML shows great promise as a predictor, its lack of explainability could hinder trust in critical manufacturing applications, and Cannizzaro et al. (2025) report that such ML–PM integration is still largely underexplored on an industrial-grade benchmark like the TEP, which is relevant for anomaly-detection research.

Architectural Integration and Data Fusion

The ability to accommodate flexibility in architecture is a common design concern when dealing with a multitude of data streams within the industry. Iuhasz et al. (2025) introduce SCAMP-ML, a modular platform for processing real-time IIoT data that utilizes configurable ML pipelines and pipelines, and Wang et al. (2023) propose a distributed stream-processing architecture based on containers that uses LSTM-based prediction and dynamic windowing for real-time IIoT anomaly detection. In the context of missing values, drift, and conflicting readings, the authors of Desikan et al. (2025) and Moldovan and Iantovics (2025) highlight the importance of multisensor trust-based fusion and signal denoising at the edge, which can significantly enhance the reliability of downstream ML. However, Diro et al. (2021) warn that there are still a number of insecure or fragile ML-based IoT anomaly-detection systems, and that there will always be a trade-off between system flexibility and reliability.

ML Benchmarking on the Tennessee Eastman Process

The TEP continues to be one of the key criteria to assess multidimensional fault-detection studies (Ahang et al., 2024). Generally, comparative studies prefer deep and ensemble learners: Ahang et al. (2024) observe that the feature extraction approach based on autoencoders (AE) is more effective than the conventional algorithms in detecting faults in TEP, and Ji et al. (2023) propose a genetic algorithm optimised backpropagation network (GA-BPNN) that enhances the convergence stability compared to the standalone neural network. Rezgui et al. (2023) also report ensemble regression models to be more robust when compared to single learners under various TEP operating conditions, but warn about generalizing about deep-learning superiority without comparing against simpler and well structured baseline models. Reinartz and Enevoldsen (2022) note that the practical use of TEP has been hampered by simulator rigidity and present a new package pyTEP, which makes it possible to create more realistic and dynamic simulation scenarios.

Interpretability, Robustness, and Adversarial Risk

Throughout, a common theme is the tension between predictive performance and explainability. However, recently, Allen et al. (2024) have introduced the Knowledge-Enhanced Spatiotemporal Analysis (KESA) model, that combines chemical-engineering domain knowledge with graph-based deep-learning method, to explain, rather than just detect, the faults, thus directly addressing the operator-trust gap that black-box models leave unanswered (Visconti et al., 2024; Moldovan & Iantovics, 2025). However, Pozdnyakov et al. (2024) show that highaccuracy deep learners on the TEP are able to be fooled by adversarialsensor perturbations and suggest adversarial training and input quantisation as a defence. Altogether, these results suggest that the detection of TEP anomalies can be seen as a three-objective optimisation problem, as opposed to a single-objective optimisation problem, because of the trade-off between accuracy, interpretability and robustness.

Process Mining for Behavioural Deviation Detection

Process mining provides an additional structure-aware view. Tadi (2024) incorporates GNNs and reinforcement learning into a multi-layer mining pipeline that can deal with very dynamic event logs, and Marin-Castro and Tello-Leal (2021) report that the quality of event-log preprocessing directly influences the reliability of conformance checking and trace clustering, a phenomenon also noted by Vitale et al. (2023), who document that variable trace lengths, missing timestamps, and noise systematically bias downstream models. Where labelled anomalies are limited, unsupervised and weakly supervised PM approaches are available (Vitale et al., 2023; Elaziz et al., 2023) but are scalable.

However, Wang et al. (2024) and Bhanage et al. (2021) warn, unsupervised PM alone often comes with a higher false-positive rate and is less effective in distinguishing between deviations that are of

functional significance and those that are merely variations of the process.

Theoretical Framework and Research Gap

The proposed integration is inspired by four complementary theories: Cyber-Physical Systems (CPS) theory (Lakshmi et al., 2023; Desikan et al., 2025), which considers manufacturing to be an interconnected feedback loop between physical processes and digital controllers; Control Systems theory, which underpins the use of adaptive, non-parametric ML to translate classical SPC into non-linear, high dimensional scenarios (Ji et al., 2023; Ahang et al., 2024); Process Mining theory, which treats event logs as behavioural traces from which structural anomalies can be reconstructed (Vitale et al., 2023; Tadi, 2024); and Explainable AI (XAI), which provides the causal-reasoning layer necessary for operator trust (Allen et al., 2024). Even though these are mature literatures individually, there is still a gap that distributes the systematic integration of PM-ML hybrids, especially on stored historical industrial datasets, where a lack of preprocessing of eventlogs, semantic differences between activities and their labels, and the absence of adaptive learning mechanisms still hinder the transferability of PM-ML hybrids from datasets and over time (MarinCastro & Tello-Leal, 2021; Bhanage et al., 2021). This study aims to fill this gap directly by building and empirically testing an ML–PM pipeline based on fusion on a dataset inspired by TEP.

METHODOLOGY

Research Design

The study is positivist and quantitative and experimental in nature as it is appropriate for the data collected from the sensors and objective and replicable data evaluation measures. A two-pipeline approach was used: a multi-variate sensor features driven ML pipeline to classify each observation window as healthy or anomalous, and a process-aware PM pipeline to analyse the event log for conformance and behaviour deviation. Both outputs are then merged together, to bring the benefits of statistical detection power together with process level transparency.

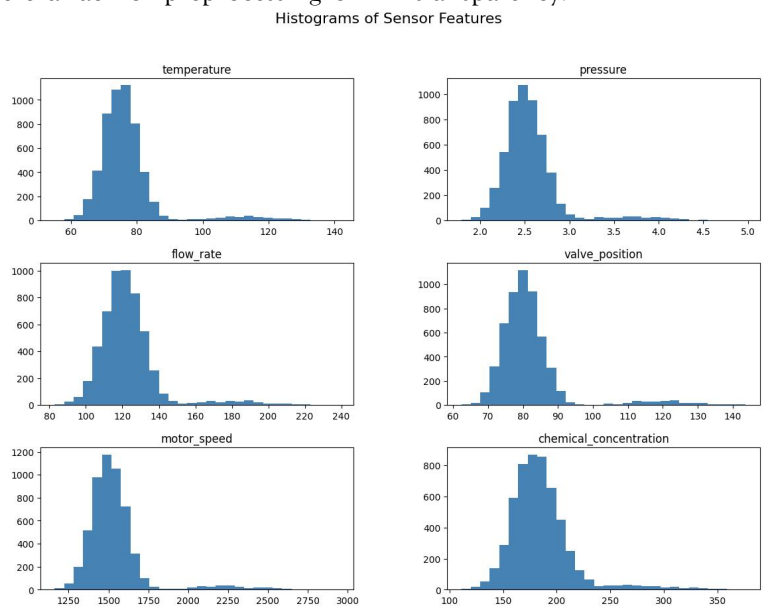

Figure 1. Schematic of the Tennessee Eastman Process (TEP), the benchmark chemical-plant simulation underlying the study's sensor dataset (adapted from Ahang et al., 2024)

Dataset

The dataset is a Kaggle-hosted Oil & Gas Chemical Plant Anomaly Detection proxy for the TEP, and consists of multivariate, high-resolution (~1 Hz) time-series readings from industrial sensors – temperature, pressure, valve position, flow rate, motor speed and chemical concentration – each timestamped and labelled as normal (0) or anomalous (1). The raw data set consisted of 5432 observations and 7 continuous variables, one of which was statistically irrelevant (low-variance) and was removed from the data set, labeled "vibration_level. Table 1 shows the attributes of the sensors.

Table 1: *Sensor attributes in the dataset.*

Sensor	Type	Unit	Description
T1	Temperature	°C	Reactor temperature
P2	Pressure	bar	Separator outlet pressure
V3	Valve status	Boolean	Valve open/closed
F4	Flow rate	L/min	Pump circulation rate

Preprocessing, Feature Engineering, and Event-Log Construction

Linear interpolation was used to fill missing values and redundant and near zero variance sensors were removed. Continuous features were scaled (MinMax or z-score, depending on the model) and time series were segmented into 60-second intervals with 30-second stride and each interval was labelled by a majority-vote rule. New rolling statistics (mean and standard deviation) were added to the dataset, as were gradients, Shannon entropy and inter-sensor correlation matrices, bringing the total number of records to 10,862 in 53 features. At the same time, sensors' state values were mapped into symbolic activities (e.g., Temp_High, Pressure_Spike) and an event log was created in the PM4Py format, with fields Case ID, Activity, Timestamp, and Outcome, which allowed to discover and check conformance of the process.

Machine-Learning Model Implementation

The five models used in supervised and unsupervised learning in scikit-learn and Tensorflow 2.x were Support Vector Machine (RBF), Random Forest, Multilayer Perceptron (MLP), Autoencoder and Isolation Forest. Some important hyperparameters are given in Table 2. The data was partitioned 80/20 (Train/Test) and hyperparameter tuning was done using Stratified 5 fold cross validation while SMOTE oversampling and class weight were used to address class imbalance.

Table 2: *Machine-learning model hyperparameters.*

Model	Key parameters	Value(s)
SVM	Kernel, C, Gamma	RBF, 1.0, scale
Random Forest	n_estimators, max_depth	100, None
MLP	Hidden layers, activation, optimizer	(64, 32), ReLU, Adam
Autoencoder	Bottleneck size, epochs	16, 50

Precision, Recall, F1-score, ROC-AUC and False Positive Rate (FPR) were used for evaluation, which were chosen specifically because of the significant class imbalance of the dataset (around 95% normal / 5% anomalous) for which the use of Accuracy alone is uninformative.

Process-Mining Technique

The Process Discovery was carried out using the Heuristic Miner (which allows to deal with noisy and incomplete logs) and the Inductive Miner (which produces sound and complete Petri-net models). Conformance checking by token replay and alignment was performed to check the observed traces against the discovered reference model to find missing or additional activities, loops and skipped steps. Process-performance analysis also uncovered any process bottlenecks, timing inconsistencies, and sequence violations. Completely the PM pipeline was adopted in PM4Py.

Hybrid Fusion Architecture

The hybrid approach fuses the ML based anomaly probabilities with the deviation scores obtained from the PM using a conservative logical-AND rule: a window is classified as a high-confidence anomaly if the ML based classifier assesses the window as an anomaly, and the PM conformance assessment also classifies the window as a deviation. This design approach is deliberately based on sacrificing recall for precision, since it is easier to inspect or shutdown the machine, than to recall it, in a safety critical manufacturing application.

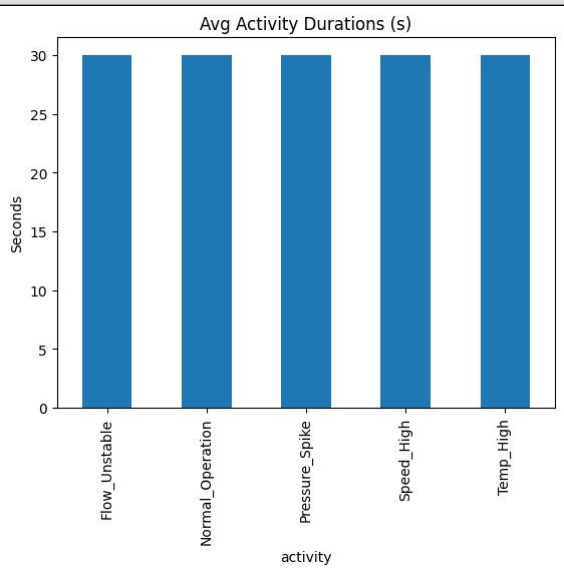

Figure 2. Hybrid Machine Learning and Process Mining architecture, showing the parallel data-driven and process-aware pipelines integrated through the fusion layer.

Experimental Validation and Ethics

The hybrid performance was compared to the performance of three baselines — ML-only, PM-only, and a rulebased threshold detector — in multi-fault test scenarios, both in terms of classification accuracy and generalisation, and in terms of latency. The robustness was further tested by adding Gaussian sensor noise (σ = 0.0–1.0), using different window sizes, and introducing random noise on the sensor. All experiments were performed using Python 3.11 on Intel i7 (3.4 GHz) processor with 16 GB RAM and using GPUs on Google Colab. The study only uses a publicly available, anonymised dataset and no proprietary nor human-subject data; code, parameters and logs are planned to be released openly under an MIT license for the sake of reproducibility.

Machine-Learning Performance

All five models are summarized in terms of their test-set performance in Table 3. Random Forest and MLP reached the top performance with Precision, Recall, F1 and ROC-AUC equal to 1.000 respectively, FPR equal to 0.000. SVM had similar performance (F1 = 0.992), but did slightly worse for recall due to the effects of kernel-margin. Performance of the unsupervised models was significantly lower: the Autoencoder showed full recall and a low precision of 0.361 - often compensating for this by generating too many false positive alerts due to normal variation; and Isolation Forest had poor performance on precision (0.054) and recall (0.050) - indicating that this model has difficulty identifying the nonlinear, correlated relationships present in this dataset.

RESULTS

Table 3: *Machine-learning model performance metrics.*

Model	Precision	Recall	F1	ROC-AUC	FPR
SVM (RBF)	1.000	0.983	0.992	1.000	0.000
Random Forest	1.000	1.000	1.000	1.000	0.000
MLP	1.000	1.000	1.000	1.000	0.000
Autoencoder	0.361	1.000	0.531	0.999	0.050
Isolation Forest	0.054	0.050	0.052	0.954	0.025

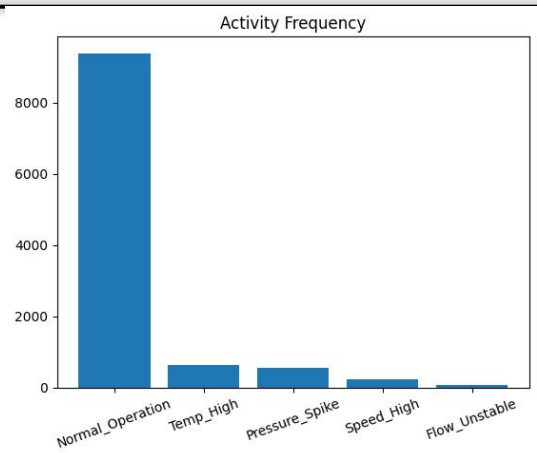

Figure 3. ROC curves for all five models; SVM, Random Forest, and MLP achieve AUC ≈ 1.0.

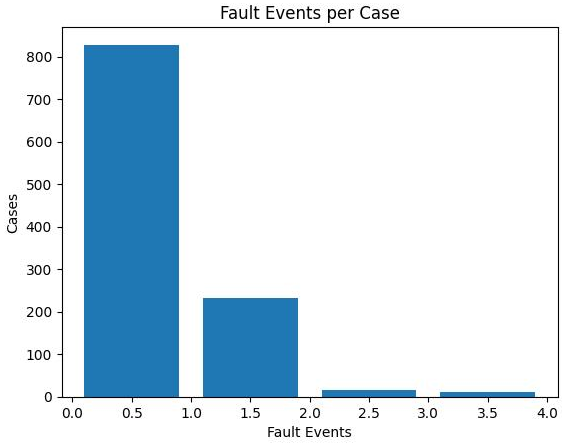

Figure 4. Precision–Recall curves, remaining near unity for the top three supervised classifiers even under class imbalance.

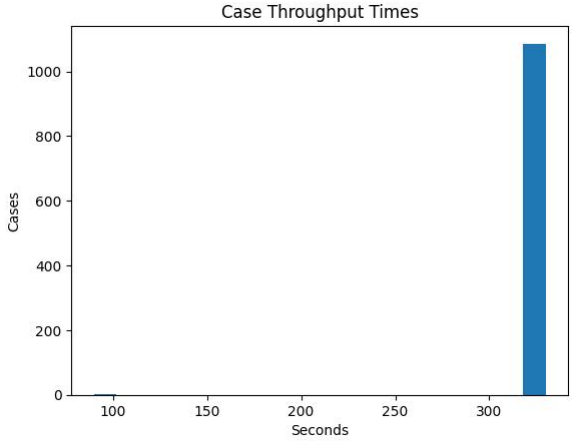

Figure 5. Comparative bar chart of Precision, Recall, and F1 across all five models.

Process-Mining Results

Ten thousand eight hundred and sixty two (10,862) windowed observations were analyzed and revealed that 84.3% of all activities belonged to Normal_Operation, 5.6% to Temp_High, 4.9% to Pressure_Spike, 2.2% to Speed_High and 0.6% to Flow_Unstable revealing that temperature and pressure variation are the most frequent fault triggers, in accordance with the thermal–pressure coupling of chemicalprocess dynamics. Comparing the fuzzily discovered Directly-Follows Graph to the behavior of the generator found that 76.9% of traces deviated from the reference (Table 4); most of these deviations were minor "micro-anomalies", not necessarily critical faults, but indicate inefficiency or drift in an early stage of the generator that may be useful for predictive maintenance.

Table 4: *Conformance-checking results.*

Trace type	Proportion
Deviant	0.769
Conformant	0.231

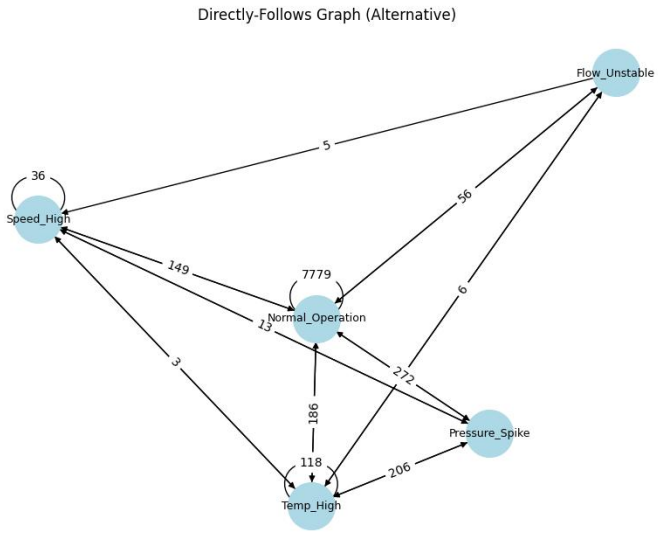

Figure 6. Directly-Follows Graph discovered from the process event log, showing Normal_Operation as the dominant hub with sporadic transitions to fault-related states.

Hybrid Integration Results

The comparison between the hybrid (AND) fusion model and three other models which use only ML or PM or rules is shown in Table 5. Precision (1.000) and correct negatives (no false positives) was achieved with the hybrid model which was designed to be conservative, however at a significantly reduced recall (0.133) compared to ML only detection. Performance of the PM-only detector was poor, failing to achieve an F1 score of > 0.074 confirming that behavioural deviation alone is not enough to achieve a good anomaly signal; the high recall (1.000) but poor precision (0.189) of the rule-based detector would also prove an undesirable balance in safety-critical environments.

Table 5: *Hybrid model performance versus baseline configurations.*

Model	Precision	Recall	F1	ROC-AUC	FPR
ML-only (MLP)	1.000	0.983	0.992	1.000	0.000
PM-only (DFG)	0.049	0.150	0.074	0.505	0.083
Hybrid (AND)	1.000	0.133	0.235	1.000	0.000
Rule-based	0.130	1.000	0.231	0.905	0.189

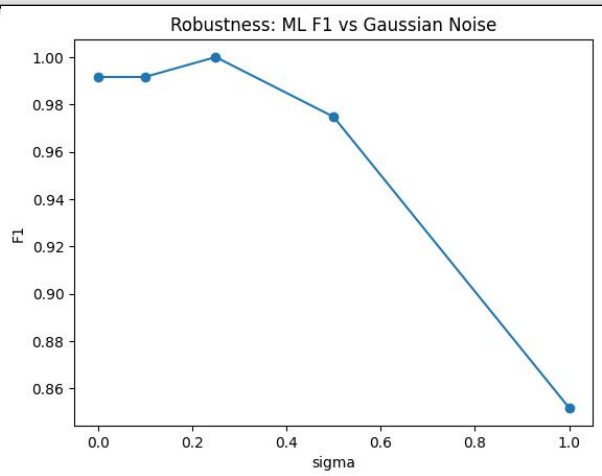

Figure 7. Hybrid model confusion matrix, showing zero false positives across the test set.

Robustness and Latency

Under injected Gaussian sensor noise, F1 remained above 0.975 up to σ = 0.5, degrading to 0.855 at σ = 1.0 (Table 6). With simulated missing data, F1 maintained a good performance of > 0.95 until 10% of the data was missing, > 0.89 at 20% and > 0.71 at 30% due to the graceful degradation under missing data caused by the redundancy of correlated sensor features. The end-to-end latency (including millisecond time-scale ML inference and its conformance alignment to the PM around the clock in tens-of-milliseconds) did not exceed one second per case, which is expected for the framework to be used in real time supervisory monitoring.

Table 6: *Robustness under Gaussian noise and missing data.*

Condition	Level	F1
Gaussian noise (σ)	0.0	0.995
Gaussian noise (σ)	0.5	0.975
Gaussian noise (σ)	1.0	0.855
Missing data	10%	0.958
Missing data	20%	0.892
Missing data	30%	0.710

DISCUSSION

There are three conclusions that can be drawn that are all related. Firstly, the performance of supervised ML models on this well structured and feature engineered dataset significantly outcompeted unsupervised methods: The Fully Connected (FC) and Ensemble (Random Forest) approaches achieved a perfect discrimination, indicating that approaches based on distance or reconstruction of data do not work well with this kind of high dimensional, non linear sensor data, as previously reported in TEP benchmarking (Ahang et al., 2024; Ji et al., 2023). The feature-importance ranking by Random Forest reveals high relevance of temperature variance, flow/valve correlation and drift in chemical-concentrations in the model, which is physically and well known in the thermal/ hydraulic processes and gives the model physical 7or not purely statistical) credibility.

Second, process mining served its purpose which was not replaced by the traditional use of the accuracy measure: it did not identify where an anomaly might have occurred as an ultimate detector per se (the standalone F1 of 0.074 statistically proves that process mining is not competitive as a classifier) but it explains why, and where, an anomaly flagged through the evidence by a traditional way was occurring within the process workflow. This aligns with the general literature which denotes the complementary role of PM with ML as it is not a replacement (Diro et al., 2021, Vitale et al., 2023).

Third, the hybrid AND-fusion approach shows that it is a precision–recall trade-off that is purposefully and defensibly made. This directly mitigates alarm-fatigue and unwarranted-shutdown problems reported in literature for industrial applications of AI (Desikan et al., 2025; Pozdnyakov et al., 2024) coming at an accepted cost of some missed true anomalies when

compared to its ML-only detection. This puts the system in the context of a more limited use case such as in a monitoring application for a reactor or pipeline for instance, in which fewer but more highly reliable alerts may be desired. Smoothing combination approaches (e.g., weighted or fuzzy-logic) can be viewed as a natural extension to apply for use-cases where recall is emphasized.

The combination of robustness (graceful degradation of F1 score in the presence of noisy data and missing data) and sub-second end-to-end latency, also counters one of the main concerns expressed about anomaly detectors heavily reliant upon deeplearning, namely laboratory grade accuracy may not stand up to field data that is often noisy and incomplete (Lakshmi et al., 2023; Iuhasz et al., 2025). However, there are some caveats associated with the above conclusions. The data used is semi-synthetic and TEP-derived, not from a real plant deployment so the results about how much noise and missingness the system can tolerate may not apply to other plant sensor designs, or maintenance policies. Although the AND-fusion rule ensures false positives are suppressed, it is still a rather coarse-grained tool and the conformance checking in process-mining assumes a certain purity of the event logs and consistent labelling of events, which is not guaranteed in legacy industrial systems (Marin-Castro & Tello-Leal, 2021).

CONCLUSION AND FUTURE WORK

In this study, a hybridity framework, based on a dataset inspired by a Tennessee Eastman Process and model combining Machine Learning and Process Mining techniques was designed, prototyped and empirically validated for the case of anomaly detection in IoT manufacturing. Among the five benchmarked ML models the Random Forest and MLP gave a perfect classification performance; adding the interpretable behavioural layer using process mining showed that a significant part of process traces (76.9 %) deviates at a micro-level from the normative process; and the resulting fusion model based on the hybrid had a perfect precision with zero false positives, while still maintaining a deliberately low recall, was robust to quite a moderate sensor noise and loss of information, and had real-time latency budgets. As a whole, these results back up the fundamental argument of the study: that a statistically accurate anomaly detection method based on ML and a behaviourally transparent PM provides a more reliable and practical approach than either of these methods on its own.

There are four areas for future research. First, adaptive/fuzzy-logic fusion technologies might dynamically rebalance the ML/PM based on the uncertainty of the operating context for better recall, but still have the accuracy of the existing framework. Second, extending the framework towards asset-level diagnostics is possible through multimodal data sources such as acoustic, visual and text in maintenance-logs. Third, the framework could be extended to integration with a digital-twin platform to facilitate the shift from detecting issues after the work from the prediction and control using simulations. Fourth and most important, the framework needs to be tested on real industrial data from manufacturing partners, with the impact of real sensor noise and human performance and deployment aspects that cannot be covered by historical data or data from the TEP. The potential of expanding the hybrid ML–PM paradigm to other domains of alike nature — such as energy grids, logistics or healthcare systems — where sensor-instrumented equipment and events play a key role, and where data-driven optimisation holds great promise, also offers an interesting possibility to test the generalisability of the paradigm beyond manufacturing..

REFERENCES

- Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification: A review. Qubahan Academic Journal, 1(2), 81–90.
- Ahang, M., Charter, T., Ogunfowora, O., Khadivi, M., Abbasi, M., & Najjaran, H. (2024). Intelligent condition monitoring of industrial plants: An overview of methodologies and uncertainty management strategies. arXiv preprint arXiv:2401.10266.
- Allen, L., Lu, H., & Cordiner, J. (2024). Knowledge-enhanced spatiotemporal analysis for anomaly detection in process manufacturing. Computers in Industry, 161, 104111.
- Aria, M., Cuccurullo, C., & Gnasso, A. (2021). A comparison among interpretative proposals for Random Forests. Machine Learning with Applications, 6, 100094.
- Bhanage, D. A., et al. (2021). A framework of deep reinforcement learning for weakly supervised anomaly detection in business processes. Journal of Big Data.
- Cannizzaro, D., Antonioni, P., Ponzio, F., Galati, M., Patti, E., & Di Cataldo, S. (2025). Machine learning-enabled real-time anomaly detection for

electron beam powder bed fusion additive manufacturing. Journal of Intelligent Manufacturing, 36(3), 2105–2119.
- Chinthamu, N., Ashish, A., & Rao, D. (2025). Data mining techniques for predictive maintenance in manufacturing industries: A comprehensive review. ITM Web of Conferences, 76, 05012.
- Desikan, J., Singh, S. K., Jayanthiladevi, A., Bhushan, S., Rishiwal, V., & Kumar, M. (2025). Hybrid machine learning-based fault-tolerant sensor data fusion and anomaly detection for fire risk mitigation in IIoT environment. Sensors, 25(7), 2146.
- Diro, A., Chilamkurti, N., Nguyen, V. D., & Heyne, W. (2021). A comprehensive study of anomaly detection schemes in IoT networks using machine learning algorithms. Sensors, 21(24), 8320.
- Elaziz, E. A., Fathalla, R., & Shaheen, M. (2023). Deep reinforcement learning for data-efficient weakly supervised business process anomaly detection. Journal of Big Data, 10(1), 33.
- Hartung, F., Franks, B. J., Michels, T., Wagner, D., Liznerski, P., Reithermann, S., & Kloft, M. (2023). Deep anomaly detection on Tennessee Eastman process data. Chemie Ingenieur Technik, 95(7), 1077–1082.
- Iuhasz, G., Fortiş, T. F., & Panica, S. (2025). Exploring machine learning methods for the identification of production cycles and anomaly detection. Internet of Things, 101508.
- Ji, C., Tao, T., Wang, J., & Sun, W. (2022). Multi-scale process monitoring based on time-frequency analysis and feature fusion. Frontiers in Chemical Engineering, 4, 899964.
- Ji, Q., Soh, A. C., Rahman, R. Z. A., Ahmad, R. K. R., & Ahmad, S. A. (2023). Integrated genetic algorithm and back-propagation neural network approach for fault detection and prediction in the Tennessee Eastman process. Transactions on Science and Technology, 10(4), 181–187.
- Kokossis, A., Georgiadis, M. C., & Pistikopoulos, E. N. (2023). Online monitoring and robust, reliable fault detection of chemical process systems. [Working paper].
- Lakshmi, G. A., Gummadi, A., & Changala, R. (2023). Block chain and machine learning models to evaluate faults in the smart manufacturing system. International Journal of Scientific Research in Science and Technology, 10(5), 247–255.
- Marin-Castro, H. M., & Tello-Leal, E. (2021). Event log preprocessing for process mining: A review. Applied Sciences, 11(22), 10556.
- Mendes, T., Cardoso, P. J., Monteiro, J., & Raposo, J. (2022). Anomaly detection of consumption in hotel units: A case study comparing isolation forest and variational autoencoder algorithms. Applied Sciences, 13(1), 314.
- Moldovan, S. C., & Iantovics, L. B. (2025). Review on information fusion-based data mining for improving complex anomaly detection. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 15(2), e70017.
- Pozdnyakov, V., Kovalenko, A., Makarov, I., Drobyshevskiy, M., & Lukyanov, K. (2024). Adversarial attacks and defences in fault detection and diagnosis: A comprehensive benchmark on the Tennessee Eastman process. IEEE Open Journal of the Industrial Electronics Society.
- Presciuttini, A., Cantini, A., Costa, F., & Portioli-Staudacher, A. (2024). Machine learning applications on IoT data in manufacturing operations and their interpretability implications: A systematic literature review. Journal of Manufacturing Systems, 74, 477–486.
- Qin, Y., Li, C., Shi, X., & Wang, W. (2022). MLP-based regression prediction model for compound bioactivity. Frontiers in Bioengineering and Biotechnology, 10, 946329.
- Reinartz, C., & Enevoldsen, T. T. (2022). PyTEP: A Python package for interactive simulations of the Tennessee Eastman process. SoftwareX, 18, 101053.
- Rezgui, W., Rezki, N., & Kerrouchi, S. (2023). Faults prediction and monitoring of complex processes using an ensemble of machine learning regression models: Application to the Tennessee Eastman process. [Conference paper].
- Selim, G. E. I., Hemdan, E. E. D., Shehata, A. M., & El-Fishawy, N. A. (2021). Anomaly events classification and detection system in critical industrial internet of things infrastructure using machine learning algorithms. Multimedia Tools and Applications, 80(8), 12619–12640.
- Shahab, M., Nagy, Z., & Reklaitis, G. (2024). Deep neural networks-based fault diagnosis

model for process systems. Computer Aided Chemical Engineering, 53, 1963–1968.

- Shkarupylo, V., Alsayaydeh, J. A. J., Yusof, M. F. B., Oliinyk, A., Artemchuk, V., & Herawan, S. G. (2024). Exploring the potential network vulnerabilities in the smart manufacturing process of Industry 5.0 via the use of machine learning methods. IEEE Access.
- Tadi, S. R. C. C. T. (2024). Process mining driven by deep learning for anomaly detection in intelligent automation systems. Journal of Scientific and Engineering Research, 11(1), 317–329.
- Visconti, P., Rausa, G., Del-Valle-Soto, C., Velázquez, R., Cafagna, D., & De Fazio, R. (2024). Machine learning and IoT-based solutions in industrial applications for smart manufacturing: A critical review. Future Internet, 16(11), 394.
- Vitale, F., De Vita, F., Mazzocca, N., & Bruneo, D. (2023). A process mining-based unsupervised anomaly detection technique for the industrial internet of things. Internet of Things, 24, 100993.
- Wang, R., Qiu, H., Cheng, X., & Liu, X. (2023). Anomaly detection with a container-based stream processing framework for industrial internet of things. Journal of Industrial Information Integration, 35, 100507.
- Wang, Y., Chiu, T., & Vasarhelyi, M. A. (2024). Applying deep learning to detect abnormal event log traces: A non-rule-based framework. The International Journal of Digital Accounting Research, 24(30), 119–140.

IEER
Institute for Excellence in Education & Research