HyPredictor: Hybrid Failure Prognosis Approach Combining Data-Driven and Knowledge-Based Methods

Miguel Almeida[1][a], Eliseu Pereira[1,2][b] and Gil Gonçalves[1,2][c]

[1]*Faculty of Engineering, University of Porto, Porto, Portugal*

[2]*SYSTEC - ARISE, Faculty of Enginnering of the University of Porto, Porto, Portugal*

Keywords: Failure Prediction, Hybrid Approaches, Knowledge-Based Methods, Data-Driven Methods, Explainable Artificial Intelligence.

Abstract: In modern manufacturing, marked by an unprecedented surge in data generation, utilising this wealth of information to enhance company performance has become essential. Within the industrial landscape, one of the significant challenges is equipment failures, which can result in substantial financial losses and wasted time and resources. This work presents the HyPredictor framework, a comprehensive failure prediction and reporting system designed to enhance the reliability and efficiency of industrial operations by leveraging advanced machine learning techniques and domain knowledge. Six machine learning algorithms were evaluated for failure prediction. The predictions from the algorithms are then refined using rule-based adjustments derived from domain knowledge. Additionally, Explainable Artificial Intelligence (XAI) techniques were incorporated, as well as the capability of users to customise the system with their own rules and submit failure reports, prompting model retraining and continuous improvement. Integrating domain-specific rules improved the performance by up to 28 percentage points in the F1 Score metric in some prediction models, with the best hybrid approach achieving an F1 Score of 90% and a Recall of 92% in failure prediction. This adaptive, hybrid approach improves prediction accuracy and fosters proactive maintenance, significantly reducing downtime and operational costs.

1 INTRODUCTION

In industrial operations, equipment failures cause significant downtime and financial losses. Predictive maintenance is a proactive strategy that anticipates faults before they occur, minimising disruptions. In the era of Industry 4.0, technologies like the Internet of Things, Artificial Intelligence (AI), and Machine Learning enhance productivity and operational intelligence by leveraging real-time data and combining data-driven and knowledge-driven methods, such as analysing sensor data and incorporating expert insights.

Integrating these advanced technologies and data-driven insights promises to revolutionise industrial operations but also presents challenges. The vast quantity and complexity of data requires sophisticated tools to avoid information overload and extract relevant insights. Traditional machine learning methods may struggle with this complexity, and static models may fail to capture the dynamic nature of industrial processes. Relying solely on data-driven or knowledge-driven approaches has inherent limitations. Combining both methodologies offers a more comprehensive understanding of equipment health and enhances decision-making processes. Addressing these challenges is crucial for effectively utilising the abundant data available.

By integrating data-driven and knowledge-driven methodologies, this work aims to address these challenges and provide a comprehensive understanding of the equipment's health. By developing accurate predictive models, this work also seeks to minimise downtime, reduce maintenance costs, and promote sustainable industrial practices by mitigating the environmental impact of equipment failures.

Data-driven methods process various sensor data, extracting valuable insights from the information generated by industrial equipment. Concurrently, expert knowledge will be integrated through knowledge-

[a] https://orcid.org/0009-0003-6559-6271

[b] https://orcid.org/0000-0003-3893-3845

[c] https://orcid.org/0000-0001-7757-7308

Almeida, M., Pereira, E. and Gonçalves, G.
HyPredictor: Hybrid Failure Prognosis Approach Combining Data-Driven and Knowledge-Based Methods.
DOI: 10.5220/0012924300003822
In *Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO 2024) - Volume 1*, pages 245-252
ISBN: 978-989-758-717-7; ISSN: 2184-2809

specific rules, enriching the analytical framework with insights and domain expertise. Finally, XAI techniques will be deployed to provide transparent explanations of the model's predictions. This integrated approach of data-driven methods, knowledge-driven rules, and XAI forms a robust analytical framework that leverages the unique advantages of each method, aligning with the principles of Industry 4.0.

This paper is organised as follows: Section 2 presents the literature review, Section 3 describes the dataset, Section 4 outlines the implementation, Section 5 presents the results, and Section 6 concludes and discusses potential future work.

2 LITERATURE REVIEW

This literature review section discusses predictive maintenance and explores data-driven methodologies, knowledge-driven strategies, and XAI methods.

2.1 Predictive Maintenance

In the realm of maintenance, as outlined by (Zonta et al., 2020), four primary categories have been identified: corrective, preventive, predictive, and prescriptive strategies. Predictive maintenance represents a transformative shift in industrial asset management, employing advanced data analytics, machine learning, and sensor technologies to anticipate equipment failures proactively. By utilising extensive datasets from sensors and other industrial sources, data-driven methodologies uncover intricate patterns and anomalies, forming the foundation of predictive maintenance. Model-driven approaches utilise mathematical or computational models to simulate equipment behaviour and optimise maintenance strategies. In contrast, knowledge-driven methods integrate human expertise, providing qualitative insights and contextual understanding to enhance predictive accuracy, especially in scenarios requiring careful considerations and predictions of rare events.

2.2 Data-Driven Methods

Data-driven methods in predictive maintenance utilise the abundance of data generated in industrial environments to identify patterns, correlations, and anomalies, facilitating the prediction of equipment health and potential failures. This approach encompasses both supervised and unsupervised learning techniques, with supervised methods like classification and regression being particularly prevalent (Angelopoulos et al., 2020). Notably, machine learning algorithms such as Random Forest (RF), Support Vector Machines (SVMs), Gradient Boosting (GB), Artificial Neural Networks (ANNs), and Logistic Regression (LR) are among the most widely utilised in predictive maintenance (Leukel et al., 2021).

Unsupervised methods operate without labelled data, making them particularly valuable in scenarios where the outcomes are not well-defined or unknown. Clustering and anomaly detection are prevalent unsupervised techniques in predictive maintenance, with Principal Component Analysis (PCA) and K-Means clustering being among the most widely used methods. Moreover, PCA was utilised to reduce the dimensionality of features (Canizo et al., 2017; Li et al., 2014). In the study by (Bekar et al., 2020), a K-Means clustering technique is employed to obtain diagnostic information, which can be utilised for labelling the data and supporting practitioners in predictive maintenance decision-making.

2.3 Knowledge-Drive Methods

Knowledge-driven methods in artificial intelligence and machine learning play a pivotal role in decision-making by utilising expert knowledge and domain expertise. These methods contrast data-driven approaches, as they involve incorporating explicit rules, ontologies, and logical reasoning into the decision-making process.

One popular example of knowledge-based methods is the development of Rule-Based Systems, where predefined rules guide decision-making (Sun and Ge, 2021). These rules can be derived from expert knowledge, logical reasoning, or established guidelines within a specific domain.

Ontologies, structured representations of knowledge defining relationships and entities within a domain, are frequently used in knowledge-based methods (Chi et al., 2022). Expert Systems are a specialised category of knowledge-based methods designed to emulate the decision-making capabilities of human experts. However, traditional knowledge representations built upon expert systems demand a specific data structure design, and most of these systems possess intricate architectures, restricting the ease of knowledge sharing and reuse (Chi et al., 2022).

2.4 Explainable Artificial Intelligence

XAI plays a crucial role in enhancing the transparency and trustworthiness of failure prediction systems in industrial settings. XAI techniques provide insights into how and why specific predictions are made, enabling humans to understand the underlying factors influencing the model's decisions (Ahmed et al., 2022; Barredo Arrieta et al., 2020).

Techniques such as Local Interpretable Model-agnostic Explanations (LIME) and Shapley Additive Explanations (SHAP) are commonly used to break down model predictions into understandable components, thus closing the gap between complex machine learning algorithms and practical, applicable insights. Integrating XAI into failure prediction systems not only enhances their reliability but also empowers users to leverage these systems more effectively, encouraging a proactive approach to maintenance and operational efficiency (Barredo Arrieta et al., 2020).

2.5 Gap Analysis

Current predictive maintenance methodologies have notable gaps. While data-driven approaches are prevalent, they often overlook the valuable insights domain experts provide. Moreover, the lack of transparency in machine learning models poses a significant challenge, hindering trust and interpretability. Additionally, many existing systems lack mechanisms for continuous improvement based on real-time feedback, resulting in models that struggle to adapt to changing operational environments. These gaps highlight the need for innovative hybrid solutions that integrate data-driven methods with domain expertise while also prioritising transparency and adaptability.

3 DATASET

This work utilised a public dataset from a metro train of Porto in an operational context (Davari et al., 2021). The dataset includes 15 signals, such as pressure readings, temperature, motor current, and air intake valves, collected from a compressor's Air Production Unit between February and August 2020. The data was logged at a frequency of 1Hz by an onboard embedded device, resulting in 1,516,948 instances.

4 IMPLEMENTATION

The HyPredictor methodology encompasses several critical stages, including data reception and pre-processing, model development, rule-based adjustments, explainability through XAI, user-implemented rules, and failure reporting with model retraining, as can be seen in Figure 1. Each stage ensures robust, accurate, and transparent predictions, promoting continuous improvement and user engagement.[1]

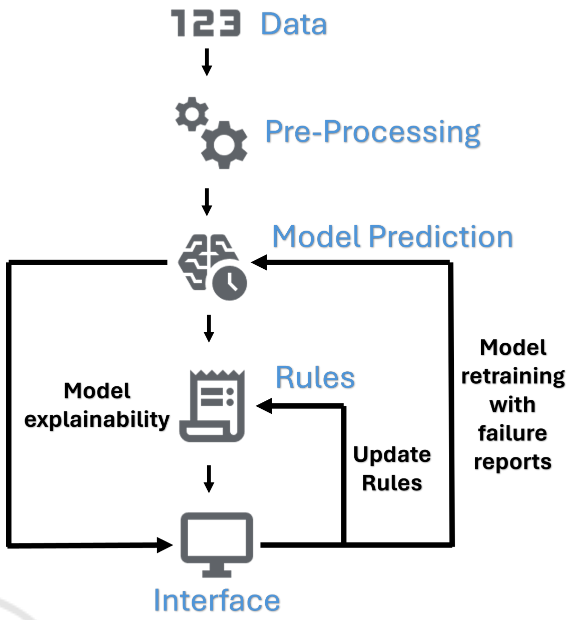

Figure 1: Proposed implementation structure.

Given that the dataset was initially unlabeled, the process began by labelling it using the failure reports provided by the company. This involved correlating the collected sensor data with documented failure events to create a labelled dataset. To allow operators to address issues before an actual failure occurred, labels were assigned to indicate a failure beginning two hours before the reported failure time. This preemptive labelling strategy aimed to provide a sufficient lead time for operators to intervene and potentially prevent failure, thereby enhancing the system's effectiveness in facilitating proactive maintenance. To enhance the model's predictive capabilities, feature engineering was performed. This process involved aggregating the data into 15-minute intervals, computing the median value within each interval, and integrating these median values into the original dataset.

Moreover, a thorough analysis was conducted to enhance the model's predictive capabilities for failure prediction. Initially, the target variable was identified as 'Failure', and correlation coefficients between this variable and all other features in the dataset were calculated. The features most strongly associated with failure occurrences were determined by ordering these correlation coefficients by their absolute values. Subsequently, a correlation threshold of $|0.2|$ was applied to filter out features with low cor-

[1] To access the code for this project, please visit the repository at: https://github.com/miguelalmeida8/HyPredictor-Framework

relation to the target variable. These low-correlation features were then dropped from the dataset to optimise the model's training process. Additionally, to mitigate multicollinearity issues, the correlation matrix of the remaining features was visualised through a heatmap, as can be seen in Figure 2, and pairs of features with high correlation coefficients were identified using a threshold of $|0.8|$ to denote high correlation. Finally, the features that exhibit a high correlation with others were removed. This feature selection and multicollinearity mitigation process aimed to optimise the model's predictive performance and ensure the robustness of the failure prediction system for industrial applications.

This step involved dividing the dataset into training and testing sets using a temporal split. The testing data, comprising 638,486 samples from June 4, 2020, onward, represented approximately 42% of the dataset. Six machine learning algorithms were evaluated to determine the most suitable method for accurate failure prediction: Random Forest, XGBoost, CatBoost, Gradient Boosting Machine (GBM), LightGBM and a voting algorithm. The voting algorithm was used to combine the strengths of multiple models, aiming to enhance overall prediction reliability. Due to their strong overall performance, the four algorithms chosen for the ensemble were XGBoost, CatBoost, GBM, and LightGBM. The models were trained using the pre-processed data, and k-fold cross-validation was implemented to ensure that the model generalises well to unseen data. Additionally, grid search was employed to optimise the model parameters, enhancing its performance. Recall was used as the primary metric to evaluate the model's effectiveness in accurately identifying failures, ensuring a focus on minimising false negatives in the predictions.

For the knowledge-driven approach, rules were defined based on expertise and experience to adjust the model's predictions. These rules are crucial for improving accuracy and handling edge cases, ensuring that rare but critical scenarios are correctly addressed. The system applies these rules post-prediction, refining the initial outputs. The implemented rules are designed to enhance the model's predictive accuracy by incorporating domain-specific insights, particularly concerning the median oil temperature and pressure readings. These rules adjust the model's initial predictions based on predefined conditions indicative of potential failures. The specific rules applied are as follows:

- **Rule 1:** If the initial prediction is 0 (no failure) but the median oil temperature exceeds 83°C, the prediction is adjusted to 1 (failure).
- **Rule 2:** If the initial prediction is 1 (failure) but the median oil temperature is below 67.25°C, the prediction is adjusted to 0 (no failure).
- **Rule 3:** If the initial prediction is 0 (no failure) and both the median oil temperature exceeds 75.65°C and the median differential pressure exceeds -0.02 Bar, the prediction is adjusted to 1 (failure).

Subsequently, XAI methods like LIME reveal the rules and contributions of individual features in predictions, as illustrated in Figure 3.

A user interface (UI) displays these explanations, allowing experts to understand the model's behaviour behind specific predictions. The visual representation of feature contributions provides clear insights into the factors influencing each prediction, making the model's decision-making process transparent and comprehensible.

The system also empowers experts to implement new rules based on their observations and insights from the model explanations. These insights can then be codified into new rules integrated into the system to refine its predictive accuracy. This continuous feedback loop is crucial in enhancing the system's performance. As experts identify and address new scenarios or anomalies, the rules evolve, making the model more robust over time.

This dynamic interaction between the model and the experts promotes a proactive maintenance strategy. The system can predict failures more accurately by preemptively adjusting the model based on real-world observations, reducing downtime and operational costs. This integration of human expertise with machine learning not only optimises the predictive model but also ensures that the system remains aligned with the evolving operational context and complexities of the industrial environment.

The interface also allows experts to submit a failure report whenever they identify a failure. This reporting mechanism is crucial for maintaining the system's accuracy and responsiveness. When a failure report is submitted, the system triggers a retraining process for the model. The model adapts and learns from recent occurrences by incorporating the new failure data, continuously improving its predictive accuracy and reducing future prediction errors. This process not only updates the model with the latest data but also enhances its ability to recognise similar patterns and anomalies in the future. The integration of real-time feedback ensures that the model remains relevant and effective in an ever-changing operational environment. This capability is particularly important in industrial settings, where conditions and failure modes can evolve rapidly.

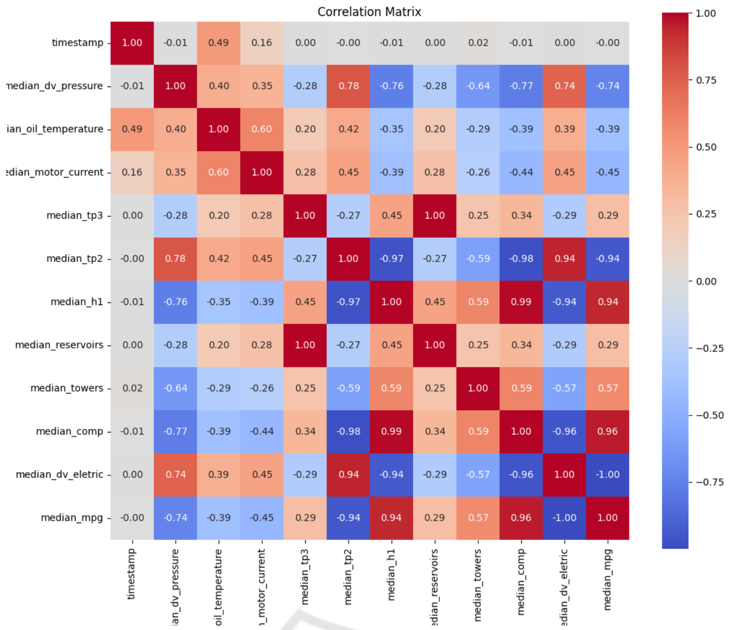

Figure 2: Correlation matrix.

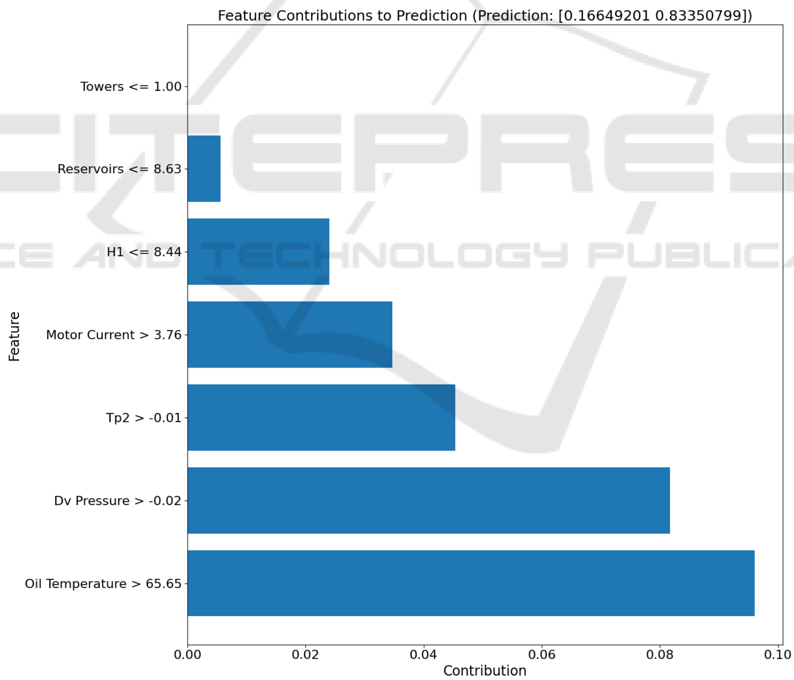

Figure 3: LIME explanation.

The described methodology ensures a comprehensive and adaptative failure prediction and reporting system. The system delivers accurate, transparent, and continuously improving predictions by combining advanced data processing, machine learning, and domain knowledge in the form of rules, XAI, and user interaction.

5 RESULTS

This section presents the performance and results of the failure prediction and reporting system.

5.1 Model Performance Metrics

The model's performance was assessed using several key metrics: Accuracy, Precision, Recall, and F1 Score. The evaluation results for the six different models, Random Forest, XGBoost, LightGBM, GBM, CatBoost and the voting algorithm, are presented in the "Before" column of Table 1 and 2. This comparison highlights the strengths and weaknesses of each model, allowing for an informed selection based on the specific requirements of the predictive maintenance task.

The evaluation of the models showcased significant variations in their performance metrics. While XGBoost, GBM, LGBM, CatBoost and the voting algorithm demonstrated good performance across multiple metrics, with high recall rates and F1 scores, Random Forest fell short in terms of precision and F1 score, indicating a higher incidence of false positives.

5.2 Rule-Based Adjustments

Applying rule-based adjustments significantly enhanced the model's performance in specific scenarios. By integrating domain-specific rules, the system improved its handling of rare but critical failure scenarios. For instance, rules were implemented to address conditions where certain sensor readings, when combined, indicated an impending failure that the model alone did not detect. The combination of machine learning predictions and rule-based adjustments resulted in an overall accuracy improvement.

Table 1 and 2 presents the performance metrics before and after applying the rule-based adjustments. The improvements in Recall, Precision, and F1 Score across different models highlight the added value of expert knowledge in enhancing predictive accuracy.

The decision to deploy GBM was primarily influenced by its highest recall rate after applying the domain knowledge, indicating its effectiveness in capturing a larger proportion of true positive instances. While high recall rates are important for capturing as many true positives as possible, the potential consequences of false alarms must be carefully weighed, especially in industrial settings where false positives could lead to unnecessary maintenance interventions and operational disruptions.

This hybrid approach not only improves the model's predictive accuracy but also ensures that critical, rare failure scenarios are effectively captured, thereby enhancing the reliability and efficiency of the failure prediction system. The system remains adaptive and robust in real-world operational contexts by continuously refining the rules based on new data and expert feedback.

5.3 Prediction and Explanation Time

The model's prediction generation time was measured to average 24.8 milliseconds per instance. This rapid prediction time ensures the system can provide timely alerts, allowing for rapid intervention and mitigation measures. Similarly, the time required to generate explanations using LIME averaged 233.8 milliseconds per instance. While explanations took slightly longer than predictions, this timeframe is considered acceptable.

The relatively quick prediction and explanation times contribute to enhanced operational efficiency. Users can act on predictions and understand the underlying reasons without significant delays, which is crucial in time-sensitive environments.

5.4 Interface

Transitioning from discussing the technical performance of the predictive model, it's important to highlight the user interface, which is the primary means through which users interact with the system. Images of the interface, showcasing its various features and capabilities, are provided in Figure 4.

To test the interface, raw data was sent via Message Queuing Telemetry Transport (MQTT), a lightweight messaging protocol ideal for real-time data transmission. The interface efficiently processes this incoming data and displays the system's current status on a user-friendly dashboard. Key features of the interface include:

- **Real-Time Failure Detection.** The dashboard displays whether a failure is predicted based on the incoming data, allowing for immediate user intervention.
- **Sensor Data Visualization.** Alongside the failure prediction, the values from various sensors, such as pressure, temperature, motor current, and air intake valve readings, are shown in real-time. This gives users a comprehensive view of the operational status and helps diagnose potential issues.
- **Failure Report Management.** The UI can display existing failure reports and allows users to create new failure reports when a failure is identified. This helps track and document failure instances systematically.
- **Rule Management.** Users can view the rules in place, organised by order of priority, ensuring that the most critical rules are given top priority. Additionally, the interface allows users to create new rules based on their insights and observations.

Table 1: Comparison of model performance before and after rules (Part 1).

Metrics	RF		XGBoost		LGBM	
	Before	After	Before	After	Before	After
Recall (%)	90.685	90.999	91.991	91.991	88.786	90.748
F1 Score (%)	64.655	93.093	90.104	90.471	92.144	92.232
Precision (%)	50.236	95.285	88.293	89.001	95.767	95.855
Accuracy (%)	96.586	99.535	99.305	99.333	99.479	99.547

Table 2: Comparison of model performance before and after rules (Part 2).

Metrics	CatBoost		GBM		Voting	
	Before	After	Before	After	Before	After
Recall (%)	91.413	91.413	91.003	92.241	91.003	91.413
F1 Score (%)	89.303	90.771	89.576	90.245	91.055	91.278
Precision (%)	87.288	90.137	88.193	88.333	91.107	91.143
Accuracy (%)	99.247	99.547	98.886	98.950	99.059	99.081

- **Model Explanation.** The UI includes functionality to show detailed model explanations, providing insights into the factors influencing each prediction. This transparency helps users understand the model's decision-making process.

The combination of real-time failure alerts, detailed sensor data visualisation, failure report management, rule management, and model explanations ensures that users can quickly understand and react to potential problems.

6 CONCLUSIONS AND FUTURE WORK

In modern manufacturing, equipment failures can cause extensive downtime, resulting in considerable financial losses and wasted resources. This paper introduces a methodology for a failure prediction and reporting system designed to mitigate such risks by proactively reducing instances of equipment failure.

The UI enables real-time failure detection, sensor data visualisation, failure report management, rule management, and model explanations, facilitating rapid understanding and response to potential issues. Additionally, the integration of domain knowledge in the form of rules significantly enhanced the system's performance. Notably, GBM, with the applied rules, emerged as the most effective approach in the evaluation. Overall, the HyPredictor methodology ensures a comprehensive and adaptive failure prediction and reporting system, leveraging advanced data processing, machine learning, and domain knowledge to enhance operational efficiency and facilitate proactive maintenance strategies in industrial settings.

Future work should focus on ensuring the scalability of the failure prediction and reporting system to deploy across various industrial environments.

ACKNOWLEDGEMENTS

This work was partially supported by the HORIZON-CL4-2021-TWIN-TRANSITION-01 openZDM project, under Grant Agreement No. 101058673.

REFERENCES

Ahmed, I., Jeon, G., and Piccialli, F. (2022). From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. *IEEE Transactions on Industrial Informatics*, 18(8):5031–5042.

Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., and Zahariadis, T. (2020). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. (2020). Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58:82–115.

Bekar, E. T., Nyqvist, P., and Skoogh, A. (2020). An intelligent approach for data pre-processing and analysis in

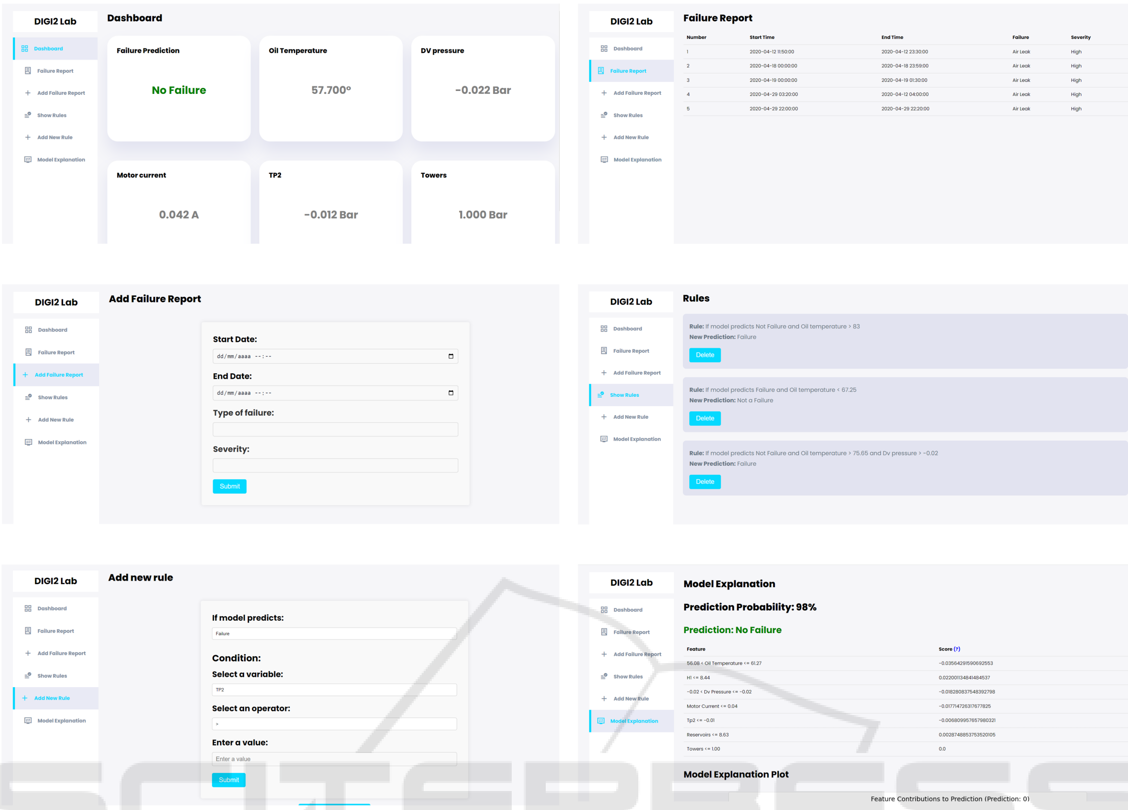

Figure 4: Images of the interface.

predictive maintenance with an industrial case study. *Advances in Mechanical Engineering*, 12.

Canizo, M., Onieva, E., Conde, A., Charramendieta, S., and Trujillo, S. (2017). Real-time predictive maintenance for wind turbines using big data frameworks. In *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pages 70–77.

Chi, Y., Dong, Y., Wang, Z. J., Yu, F. R., and Leung, V. C. (2022). Knowledge-based fault diagnosis in industrial internet of things: A survey. *IEEE Internet of Things Journal*, 9:12886–12900.

Davari, N., Veloso, B., Ribeiro, R. P., Pereira, P. M., and Gama, J. (2021). Predictive maintenance based on anomaly detection using deep learning for air production unit in the railway industry. In *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–10. IEEE.

Leukel, J., González, J., and Riekert, M. (2021). Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review.

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., and Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies*, 45:17–26.

Sun, Q. and Ge, Z. (2021). A survey on deep learning for data-driven soft sensors. *IEEE Transactions on Industrial Informatics*, 17:5853–5866.

Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., and Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers and Industrial Engineering*, 150.