Combining Random Forest and SHAP for Interpretability and Insights in Predictive Maintenance

MSc Research Project
Data Analytics

Bruna Rafaela Alves Garcia
Student ID: x21245835

School of Computing
National College of Ireland

Supervisor: Arjun Chikkankod

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Bruna Rafaela Alves Garcia
Student ID:	x21245835
Programme:	Data Analytics
Year:	2025
Module:	MSc Research Project
Supervisor:	Arjun Chikkankod
Submission Due Date:	03/01/2025
Project Title:	Combining Random Forest and SHAP for Interpretability and Insights in Predictive Maintenance
Word Count:	6689
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Bruna Garcia.
Date:	3rd January 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Combining Random Forest and SHAP for Interpretability and Insights in Predictive Maintenance

Bruna Rafaela Alves Garcia
x21245835

Abstract

This study investigates the integration of Random Forest models with SHAP (SHapley Additive exPlanations) to address key challenges in Predictive Maintenance (PdM), particularly within resource-constrained environments. The study employs the MetroPT3 dataset, comprising time-series sensor data from metro air compressor unit, to identify machine failure patterns and potentially inform maintenance decisions. Given the dataset's highly imbalanced nature and the challenges associated with preserving temporal dependencies, the methodology employs data preprocessing techniques, such as rolling statistical, feature engineering, class weighting and correlation-based feature selection, were employed to prepare the dataset for analysis. The initial Random Forest model demonstrated a good predictive capability, but the integration of SHAP provided feature-level insights that enhanced transparency and informed further refinements. Threshold tuning further optimised the model's performance, achieving a precision of 0.94, recall of 0.74, and F1-score of 0.83 for the minority class. SHAP analysis highlighted key predictors, such as pressure-related and motor current features, offering actionable insights for maintenance decisions. This work underscores the importance of balancing predictive accuracy with interpretability, addressing both technical and practical challenges in Predictive Maintenance. The findings demonstrate the potential of combining machine learning with Explainable AI to deliver transparent and actionable insights. Limitations related to class imbalance and dataset specificity are discussed, along with proposals for future research.

1 Introduction

The increasingly dependence on Machine Learning (ML) models in industrial settings needs strategies for making these models more interpretable, particularly in applications like Predictive Maintenance (PdM). Which has considerable benefits in terms of minimising unplanned downtime and lowering operating costs. However, the complexity of ML models can sometimes limit transparency, posing challenges to trust and adoption.

Unlike traditional maintenance strategies, such as reactive or preventive maintenance, Predictive Maintenance leverages advanced data-driven methodologies, to anticipate issues before they arise. Integrating explainable AI (XAI) techniques into Predictive Maintenance systems further addresses critical concerns such as transparency, accountability and trustworthiness. These enhancements give maintenance teams important insights that help them to understand the rationale behind AI-driven decisions, fostering collaboration and improving decision-making in industrial environments (Gawde et al., 2024).

Furthermore, XAI not only supports the practical application of Predictive Maintenance but also offers opportunities to refine and improve the underlying predictive models. This study explores the potential of SHAP in tackling these challenges, delivering action- able insights and improving the effectiveness of these models.

1.1 Background and Motivation

Many Predictive Maintenance models operate as so called "black boxes", providing limited visibility into their decision-making processes. This opacity is particularly problematic in critical applications, where understanding why a prediction is made is as important as the prediction itself.

By combining real-time data and ML algorithms, Predictive Maintenance significantly reduces operational disruptions and maintenance costs while improving equipment reliability. According to Madhu and Nagaraju (2024), AI-driven Predictive Maintenance solutions implemented in industries like wind energy and aviation have achieved up to a 30% reduction in maintenance costs and a 40% decrease in unplanned downtimes. The ability to predict equipment failures enables industries to optimise resource allocation, extend machinery lifespan and ensure operational safety. Furthermore, the authors highlight that flexibility and interpretability of ML models, such as Random Forests, make them ideal for processing high-dimensional sensor data typical in industrial settings. These models provide frameworks for anomaly detection and lifecycle predictions.

The MetroPT3 dataset, used in this study, presents additional challenges as it is both highly imbalanced and structured as a time-series. Traditional class imbalance solutions, such as SMOTE, cannot be applied without disrupting temporal dependencies, necessitating alternative approaches like class weighting. In this context, Random Forest models were selected for their robustness and interpretability, while SHAP was employed to provide detailed feature-level insights into model predictions. This research aims to balance the dual objectives of achieving actionable predictive insights and maintaining model transparency, even in resource-constrained environments.

1.2 Research Question

This study focuses on investigating the following research questions:

Question 1: What are the practical challenges and potential benefits of applying SHAP for interpreting Random Forest models in Predictive Maintenance?

Question 2: How effectively can Random Forest and SHAP be utilised to deliver actionable maintenance insights in environments with computational and resource constraints?

The study is organised into sections as follows: Part 2, "Related Work", reviews previous research in the fields of Predictive Maintenance, time-series analysis and Explainable AI (XAI). Highlighting the research gap this thesis addresses and the contributions it makes. Part 3, "Methodology", explains the dataset used, the preprocessing techniques applied and the modelling approach chosen for this study. Part 4, "Design Specification", details the framework and techniques employed in the implementation, alongside the requirements for achieving the research objectives. Part 5, "Implementation", describes the final preparation of the dataset, the training of the Random Forest model, and the application of SHAP for interpretability. Part 6, "Evaluation", assesses the effectiveness of the methodology and interpretability techniques, presenting results and their implications. Finally, Part 7, "Conclusion and Future Work", summarises the key findings and research questions addressed, and outlines potential directions for further exploration in Predictive Maintenance and XAI.

2 Related Work

This section presents a comprehensive review of the existing literature on data-driven Predictive Maintenance, with an emphasis on the application of Machine Learning techniques to real-world systems and Explainable AI. The aim is to establish a foundation for the subsequent discussions in this thesis. The review examines the strengths and limitations of various methodologies employed to predict and prevent system failures, detect anomalies, and evaluate their potential for practical implementation in operational settings.

2.1 Machine Learning

Machine learning (ML) has emerged as a groundbreaking force, redefining how computers process data and solve complex problems. As summarised by Rai et al. (2024), ML replaces manually coded rules with algorithms capable of learning from data and adapting over time. This paradigm enables systems to extract insights, recognise patterns and improve their performance independently, shaping industries and everyday life profoundly. Its success relies on leveraging diverse datasets, from structured numerical data to unstructured formats like images and text, underscoring the importance of data quality and volume in building robust models. ML has two primary paradigms, which are: supervised and unsupervised learning. Supervised learning uses labelled datasets to train algorithms, enabling accurate predictions and classifications for new data points. On the other hand, unsupervised learning handles unlabelled datasets, uncovering hidden patterns and relationships, proving particularly valuable for clustering and segmentation tasks.

The large availability of data produced every day is what makes ML evolve so rapidly. As it focusses on creating algorithms based on the machine's prior experiences. The purpose of ML is to detect patterns/trends in data and then use those to make meaningful inferences. Delivering more assertive outcomes through increasingly precise predictions. Machine learning approaches rely largely on computational resources (Jabbar et al., 2018) In simple terms: ML demonstrates the ability to deliver consistently accurate pre- dictions by learning from training data. Its primary goal is to automate knowledge engineering, reducing manual effort through techniques that enhance efficiency by uncovering patterns in data. The effectiveness of machine learning is evaluated experimentally, focusing on its capacity to outperform traditional methods in real-world scenarios (Singh et al., 2006).

Ratha et al. (2020) provided a comprehensive overview of ML applications, emphasising the use of supervised algorithms for analysing trends and autonomous decision-making in domains such as: Healthcare and smart cities. Ra and Souza (2019) discussed the evolution of ML, highlighting supervised, unsupervised, and reinforcement learning algorithms, and illustrating their applications in predictive modelling and decision-making tasks. A work by Singh et al. (2016), explored the comparative performance of supervised learning techniques, such as Decision Trees, Random Forests and Support Vector Machines, demonstrating their adaptability in solving classification and regression problems across diverse domains. More recently, Yu et al. (2023) investigated advancements in ML, presenting neural networks and boosting algorithms as key contributors to break-throughs in areas like defect detection and earthquake prediction. Whereas Gupta et al. (2022) focused on supervised learning's applicability, showcasing regression and classification models in use cases ranging from stock price prediction to academic performance analysis.

2.1.1 Random Forest

Random Forest algorithm has demonstrated great performance in Predictive Maintenance applications across different domains. Aji et al. (2020) implemented Random Forest to develop a Predictive Maintenance model for magnetic sensors. The model achieved a high RF score of 0.98 and a Mean Absolute Error (MAE) of 0.83, demonstrating its effectiveness in predicting maintenance time. The study emphasised the algorithm's ability to aggregate predictions from multiple decision trees, enhancing accuracy and robustness in scenarios requiring continuous real-time monitoring.

Maashi et al. (2020) proposed a Predictive Maintenance framework using Random Forest for industrial duct fans. Their model was able to perform binary classification to identify equipment abnormalities and regression to estimate Remaining Useful Life (RUL). With an impressive accuracy of 99% and a regression Root Mean Squared Error (RMSE) of 80, the study underscored the model's ability to handle complex datasets efficiently. The authors highlighted Random Forest's ensemble learning approach, which combines low bias and high variance decision trees to deliver superior performance in both predictive and regression tasks. Similarly, Vinh et al. (2022) also employed Random Forest to estimate RUL, but their approach was different. Their system integrated data from diverse sensors through LoRa nodes, enabling real-time monitoring and prediction. The Random Forest model demonstrated robust performance in this case, achieving a normalised Root Mean Square Error (NRMSE) of 0.1427, highlighting its capability to predict RUL accurately in varied machine environments. The authors emphasized the model's utility in improving operational efficiency and proactively preventing equipment failures.

Random Forest was also explored for predicting relative humidity in a smart factory environment. By using an Industrial Internet of Things (IIoT) platform, real-time data on temperature, particulate concentration and relative humidity were collected to train their model. The study achieved really good accuracy of 82.49%, showing RF's ability to handle high-dimensional data with reliable performance. This Predictive Maintenance approach not only reduced maintenance costs but also mitigated risks associated with static charge and photolithographic degradation, which are critical for cleanroom operations (Prihatno et al. (2021).

These studies illustrate the flexibility and efficacy of Random Forest in Predictive Maintenance applications. Its strengths in handling imbalanced data, generating reliable predictions and providing interpretability make it an ideal choice for scenarios involving critical industrial systems and sensor-based monitoring.

2.1.2 Class Imbalance

Class imbalance occurs in many real-world applications, it is when the distribution of classes in a dataset is uneven. Where one or more classes are significantly underrepresented compared to others. This imbalance can severely affect the performance of ML models.

The classification of imbalanced datasets is a well-known challenge when it comes to handling data. A key issue is the high misclassification rate of the minority class, as classifiers tend to prioritise the majority class. To address this, proposed solutions in the literature can be broadly categorised into approaches such as: Data Sampling: This involves adjusting the sampling of training instances to create a more balanced dataset. Cost-Sensitive Learning: This approach assigns higher misclassification costs to minority class instances (e.g., false positives) and lower costs to the majority class (e.g., false negatives), encouraging the model to treat both classes more equally (Katrakazas et al., 2019).

The former, cannot be used in time-series data due to the temporal dependency, so different solutions need to be addressed.

According to Ahmadzadeh et al. (2019), class imbalance is a prevalent issue with potential solutions, many of which are widely recognised but still prone to misuse. This is especially evident when the primary goal is not centred on Machine Learning itself but on evaluating and examining domain-specific theories. The complexity of the problem, coupled with a lack of expertise in data handling, often leads to an underestimation of the required level of precision. This can result in analyses that are impractical, unreliable, and of limited real-world applicability.

Various studies have been proposed to address this issue. Johnson and Khoshgoftaar (2019) conducted an extensive review of deep learning techniques for imbalanced data, categorising solutions into data-level, algorithm-level, and hybrid methods. They highlighted the effectiveness of traditional resampling techniques, such as SMOTE, and algorithmic adjustments, such as cost-sensitive learning, to improve performance in skewed datasets. Chen et al. (2024) further explored ensemble learning methods, including bagging and boosting, which enhance the robustness of classifiers against imbalanced data distributions by reweighing or resampling techniques. A work by Niaz et al. (2022) discussed hybrid techniques combining data balancing with model-specific enhancements, identifying Easy Ensemble and Balance Cascade as effective solutions for binary and multi-class imbalance problems. These methods integrate the benefits of both resampling and algorithmic tuning, achieving improved predictive accuracy. Finally, Jafarigola and Trafalis (2023) reviewed emerging trends in imbalanced learning, such as long-tail learning and imbalance regression, noting their application in real-world problems like fraud detection and medical diagnosis, where skewed datasets are prevalent.

2.1.3 Time-series Classification

Time-series classification is a challenging ML domain that aims to predict class labels for temporally ordered data. Several approaches have been developed to deal with this complexity. Geurts (2001) suggested a pattern extraction approach that identifies and combines local features in time-series. This approach uses decision trees to generate interpretable classification rules, producing competitive results in artificial and real-world datasets by focusing on the identification of discriminatory temporal patterns.

Batal et al. (2009) developed the STF-Mine method, which produces temporal abstraction patterns for multivariate time-series classification. This technique divides raw time-series into qualitative states, extracts common patterns with an expanded A-priori algorithm, and chooses discriminatory features using statistical metrics. Their experimental results revealed the effectiveness of STF-Mine in enhancing classifier accuracy, particularly in medical datasets where temporal patterns are crucial.

These developments highlight continuing attempts to increase the accuracy and interpretability of TSC approaches by emphasising temporal feature extraction and pattern identification.

2.1.4 Performance Metrics

The assessment of predictive models in scenarios with class imbalance often calls for metrics that are used to skewed distributions of data. The F1-score, recognised as the harmonic mean of precision and recall, is particularly effective when true negatives are not as critical as true positives and false positives (Puthiya Parambath et al., 2014). This metric is preferred in binary classification tasks involving imbalanced data, where it provides a balanced perspective on precision and recall. Unlike traditional accuracy measures, the F1-score does not overemphasise majority class performance, making it suitable for applications like Predictive Maintenance where failure events are rare but consequential.

Flach and Kull (2015) further emphasised the advantages of precision-recall curves over receiver operating characteristic (ROC) curves in imbalanced datasets. They highlighted that precision-recall curves focus solely on positive class performance, a critical consideration for Predictive Maintenance. They proposed Precision-Recall-Gain curves as a refinement, which calibrate precision and recall metrics to address incoherencies in scale and better align them with real-world decision thresholds. These refinements underscore the importance of precision-recall metrics in scenarios like this research, where understanding model trade-offs is critical.

Lipton et al. (2014) explored the optimisation of the F1 score in binary and multilabeled classification tasks. The study also discussed the challenges associated with determining optimal decision-making thresholds to maximize the F1 score, especially in the presence of unbalanced datasets. It was noted that while F1 provides an effective measure for balancing precision and recall, its nonlinear properties and asymmetric treatment of positive and negative classes can introduce complexities in optimization, particularly for rare labels (Lipton et al., 2014).

By adopting the F1-score and related metrics, this research aligns with the best practices for evaluating predictive models in imbalanced datasets. The focus on precision and recall ensures a targeted assessment of the model’s ability to predict failure events accurately, avoiding metrics that prioritise majority class performance. This methodological choice supports the study’s objectives of ensuring reliable and interpretable predictions in Predictive Maintenance contexts.

2.2 Predictive Maintenance

Predictive Maintenance (PdM) is an important method for reducing unexpected downtime and operating expenses by predicting equipment problems. It employs Machine Learning (ML) to assess sensor and operational data, allowing for prompt actions. Amer et al. (2023) investigated the performance of various supervised machine learning algorithms, including Random Forest, XGBoost and Support Vector Machines. They discovered that XGBoost was especially successful for big datasets. Their findings highlight the relevance of algorithm selection based on data characteristics and indicate that Predictive Maintenance systems can dramatically minimise production stoppages.

Paolanti et al. (2018) developed a Predictive Maintenance architecture targeting Industry 4.0 contexts, using Random Forests to categorise machine states with excellent accuracy (95%). This strategy used real-time data collected by IoT sensors and analysed using a cloud-based system, emphasising the importance of data-driven approaches in supporting maintenance decisions and improving operational efficiency.

In an earlier study, Susto et al. (2015) proposed a Multiple Classifier Predictive Maintenance (PdM) approach for dealing with imbalanced datasets, which are typical in maintenance tasks. The strategy avoided unexpected failures and optimised maintenance schedules by using classifiers with different prediction horizons. Their work in semiconductor production demonstrated the efficiency of this strategy in balancing operational costs with performance.

2.2.1 Maintenance Costs

Maintenance costs represent a significant aspect of industrial operations, and it would not be different for the metro train industries. Maintenance costs can reach as high as 70% of production expenses. This high proportion is driven by automation, the integration of Industry 4.0 technologies, and the complexity of systems, which increases labour and repair expenses (Lemes and Hvam, 2019). The authors state that proper maintenance cost categorisation, such as: downtime, wages, spare parts and degradation, helps industries improve cost visibility and decision-making. However, lack of standardisation and detailed quantification, with Predictive Maintenance offering a promising alternative for cost reduction, since it generally cost less than corrective or preventive actions. Models that integrate predictive strategies with traditional maintenance frameworks enhance the operational and financial efficiency of industries.

According to He et al. (2022), in the metro train sector, maintenance represents 40% of total costs, directly influencing operational efficiency, energy savings and system reliability. The research highlights that a reliability-centered maintenance (RCM) approach, combined with strategies like "opportunistic maintenance", allows for simultaneous optimisation of costs and train availability. The study demonstrates that arranging maintenance tasks based on component reliability thresholds significantly reduces shut- down frequency, leading to cost savings and increased operational active time. The use of mathematical optimisation models to schedule maintenance tasks under constraints like failure probability and cost provides a structured methodology for managing these complex systems. By reducing downtime and aligning maintenance schedules with operational needs, such frameworks are pivotal in achieving economic and technical feasibility in metro systems.

ML-based approaches can further reduce costs by improving the precision of failure predictions, allowing for just-in-time interventions. However, as noted by Leukel et al. (2021), the selection of appropriate ML algorithms and features is crucial, as imbalanced data can skew predictions and lead to inefficient maintenance scheduling.

2.3 Explainable AI

Explainable Artificial Intelligence (XAI) is becoming more important for improving the transparency and trustworthiness of machine learning (ML) models, especially in sensitive areas such as healthcare, criminal justice and Predictive Maintenance. Karlsson and Bengtsson (2022) evaluated Integrated Gradients (IG), a prominent XAI method, across sequential and non-sequential datasets. Their results revealed that IG effectively interprets deep learning models such as RNN, LSTM, and GRU but is influenced by baseline selection, particularly for non-sequential data. On a different study, Samek et al. (2017) discussed the challenges of interpreting the decisions of complex deep learning models, often regarded as "black boxes" due to their non-linear structures. The study introduced methods like Sensitivity Analysis (SA) and Layer-wise Relevance Propagation (LRP) to explain predictions, highlighting their effectiveness in visualising model decisions in tasks such as image classification, text analysis and human action recognition. LRP, in particular, demonstrated superior performance over SA by assigning relevance scores to input features, enabling a more accurate decomposition of predictions. Additionally, they also highlighted the role of XAI in identifying biases in datasets and improving model performance through a better understanding of decision-making processes. These advances underscore the potential of XAI to bridge the gap between complex ML models and actionable insights in various real-world scenarios.

Cummins et al. (2024) categorised XAI methods for Predictive Maintenance, highlighting challenges in achieving interpretability while preserving performance. They emphasised the importance of local explanations for actionable insights and proposed systematic evaluation frameworks to enhance user trust in Predictive Maintenance systems (Cummins et al., 2024). Rudin (2019) argued for prioritising inherently interpretable models over post-hoc explanations, criticising the reliability of black-box models in critical decision-making contexts. The study advocates for model simplicity to achieve transparent predictions without compromising accuracy.

2.3.1 SHAP

SHapley Additive exPlanations (SHAP) is a XAI approach based on game theory that aims to explain the predictions of any machine learning model. It uses Shapley values, which is a concept from game theory, by computing the contribution of each feature to the prediction. In a recent work (Van den Broeck et al., 2020) the computational complexity of SHAP explanations was analysed. They demonstrated that while SHAP ensures desirable properties such as local accuracy and consistency, its computational tractability is limited, especially for models like logistic regression and neural networks. The study emphasised that fully-factorised distributions simplify computations but remain computationally intensive in certain cases.

Kedar and Mhatre (2024) compared SHAP with other XAI methods, including LIME, Anchors and Permutation Importance. According to their findings, SHAP demonstrated certain superiority in providing detailed feature-level insights, crucial for building trust in AI systems. In addition, they noted that while SHAP excels in global interpretability, challenges remain in computational efficiency and feature dependency, particularly for high-dimensional datasets.

3 Methodology

3.1 Research Methodology

This study adopts a mixed-methods approach, combining quantitative analysis with interpretive techniques, to examine the potential of Machine Learning for Predictive Maintenance in a time-series dataset. The primary objective is to investigate how SHAP (SHapley Additive exPlanations) improves interpretability and supports maintenance decision-making with a particular focus on applying Random Forest models to the MetroPT3 dataset.

The research phases were as it follows: 1. Problem Definition: Identify the challenges in Predictive Maintenance, particularly in imbalanced a time-series dataset. And formulate research questions aiming to explore SHAP's applicability to this domain. 2. Data Understanding and Preparation: Investigate the characteristics of the MetroPT3 dataset followed by data preprocessing and class imbalance handling. 3. Model Development: Implement and train Random Forest models, chosen for its interpretability and computational constraints. 4. Interpretability: Integrating SHAP to the analysis to elucidate the global contributions of features to model predictions. 5. Validation and Insights: Evaluate the model's performance and interpret the implications of its predictions for the MetroPT3 dataset.

3.2 Data Analytics Methodology

The CRISP-DM, which stands for Cross-Industry Standard Process for Data Mining, methodology was created by a consortium of companies including Daimler AG, Integral Solutions Ltd (ISL), NCR Corporation and OHRA, an insurance company in 1996. The 1.0 version of CRISP-DM was released and fully documented, providing data miners with a standardized framework and guidelines. The process comprises six well-defined and structured stages or phases and it provides a simple and clear model for data analysis (Shafique & Qaiser, 2014). The standard six-step CRISP-DM model, as shown in Figure 1, was slightly adapted to meet the specifications of this project. A short summary of what will be done is also shown in Figure 1.

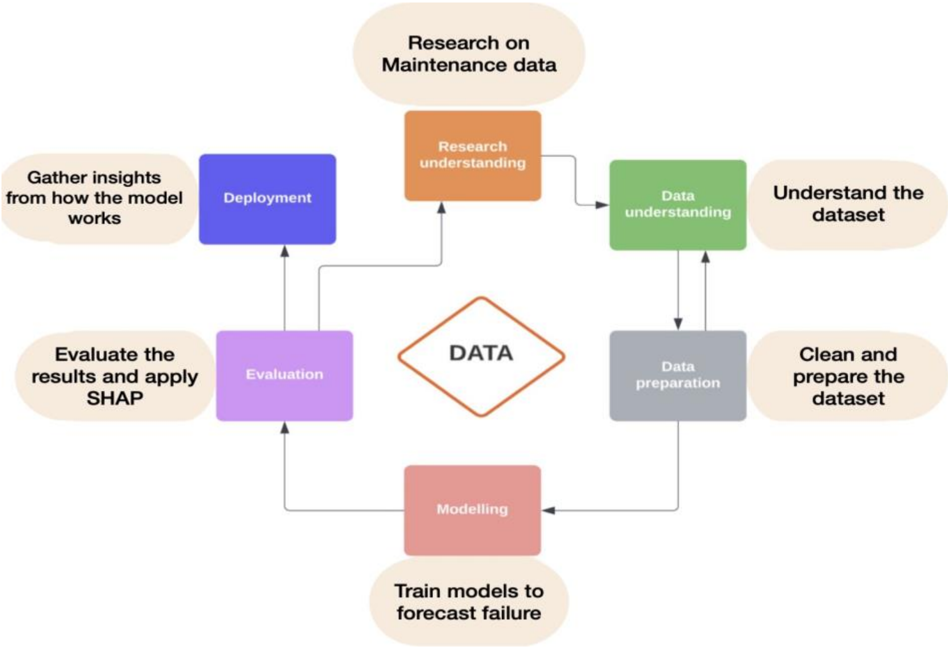

Figure 1: CRISP-DM Process Model (Source: Adapted from Shafique & Qaiser 2014, p.3 and Vijaywargiya 2019, p.8)

3.3 Data Description

The dataset, named MetroPT3, which is available at UC Irvine Machine Learning Repository[1], was collected from the air production unit (APU) of "Metro do Porto" trains in an operational context. It spans February to August 2020 and consists of 15,169,480 data points logged at a frequency of 1 Hz. This multivariate time-series dataset contains readings from 15 sensors, including seven analogue sensors measuring variables such as pressure motor current and oil temperature. Plus, eight digital sensors monitoring electrical signals like air intake valve activity and compressor outlet valve status. The dataset was designed to support Predictive Maintenance tasks, anomaly detection and remaining useful life (RUL) prediction (Davari et al., 2021).

3.4 Data Preprocessing

Data preprocessing is a crucial foundation for conducting credible data analyses. It encompasses a variety of techniques designed to improve the integrity of raw data. These methods include checking for outliers and deciding what to do with them, which may skew the results, and the imputation of missing values (if needed) to ensure completeness. This stage is vital for preparing data for meaningful analysis by ensuring accuracy, consistency, and reliability in the dataset. Expanding on these techniques, preprocessing also often involves normalising data to a common scale and encoding categorical variables to better fit analytical models. By addressing these preliminary needs, data preprocessing enhances the overall effectiveness of data analysis insights (Fan et al., 2021).

Time-series data plays a crucial role in various application scenarios, leading to the development of diverse approaches for its preprocessing. These approaches include strategies for cleaning, segmenting, representing, normalising, comparing and aggregating time-series data. Combining these operations creates a preprocessing workflow with significant flexibility. Designing an effective preprocessing pipeline requires incorporating

[1] https://archive.ics.uci.edu/dataset/791/metropt+3+dataset

domain knowledge (Bernard et al., 2012).

3.5 Handling Class Imbalance

The methodology employed in this research involved the use of "class_weights" to address the unique characteristics of time-series data. This approach was chosen as there is no clear consensus on the applicability or effectiveness of techniques such as over/undersampling for certain types of data (Yang et al., 2024).

According to Moniz et al. (2016), time-series forecasting presents significant challenges due to the non-stationary nature of the data, which creates a complex environment for predictive tasks. One common issue is the imbalanced distribution of the target variable, where certain intervals, though critical to the user, are underrepresented. Traditional regression methods often focus on the average behaviour of the data, which contrasts with the goal in many time-series forecasting tasks that aim to predict rare events. To address this, resampling strategies are frequently employed, modifying the learning data distribution to favour a specific bias.

3.6 Model Development

The model selected for this analysis is Random Forest, chosen for its versatility and effectiveness in handling structured datasets. While the primary objective is not solely to achieve high predictive accuracy, the focus is on leveraging Random Forest as a baseline to explore the interpretative capabilities of SHAP in the context of maintenance prediction. The study aims to examine how SHAP can uncover the key factors influencing model predictions, providing insights for improving the model's performance and also maintenance decision-making. Furthermore, the study seeks to evaluate the feasibility of using a method like SHAP in environments with constrained computational resources and time-sensitive requirements.

3.7 Limitations in Transferability and Generalisability

The findings are specific to the MetroPT3 dataset and may not generalise directly to other industries settings. Future research should explore the applicability of this methodology to diverse equipment and operational contexts.

4 Design Specification

This section outlines the framework, techniques and methodologies employed in this research to explore the interpretability of Random Forest models in Predictive Maintenance (PdM) using SHAP. The focus is on identifying the key features influencing failure predictions and evaluating how effectively actionable insights can be delivered in resource-constrained environments. This study aims to propose a straightforward approach that balances predictive accuracy and transparency while addressing the challenges posed by highly imbalanced time-series datasets.

The project is structured into several key stages in order to meet the research objectives. First, the dataset is selected, the MetroPT3 dataset, a time-series collection of machine sensor readings. Understanding the dataset and its context is crucial, as it reflects real-world Predictive Maintenance scenarios. The dataset includes sensor readings related to air compressors, which are critical for ensuring the functionality of metro systems.

Data preprocessing constitutes the next stage, addressing binary target creation (status), irrelevant features, rolling statistical feature generation (mean and standard deviation), forward and backward filling and class imbalance are applied to maintain data integrity while preserving the sequential nature of the dataset. Feature selection leverages correlation analysis to identify and retain relevant predictors for modelling.

The modelling phase utilises Random Forest and to address the dataset's severe imbalance, class weights are applied during training, ensuring fair representation of the minority class (status=1). Model evaluation is conducted using metrics such as precision, recall, and F1-score to assess the effectiveness of failure predictions. Hyperparameter optimisation and threshold tuning are also performed to enhance the model's performance under.

Finally, SHAP is integrated to interpret model predictions, providing insights into feature contributions, highlighting the contributions of specific predictors in individual failure cases. By integrating Random Forest and SHAP, this research demonstrates a transparent approach to Predictive Maintenance that aligns with computational and operational constraints.

5 Implementation

The implementation focuses on developing and executing a pipeline to analyse and interpret machine failure predictions using Random Forest and SHAP to elevate the model's interpretability. The pipeline is designed to process time-series data and produce outputs that help to understand which features play an important role in the model. The final implementation integrates data preprocessing, model training, evaluation and explainability.

5.1 Data Preprocessing and Transformation

The raw dataset was transformed to prepare it for modelling. Key preprocessing steps included: creating a binary target variable ('status') to mark failure events. As The purpose of Predictive Maintenance is to identify and prevent failures proactively. Creating a binary target variable (status) to mark failure events (failure = 1, no failure = 0) was essential for framing the problem as a classification task. This simplified the model's objective to predict failure events based on sensor readings.

In addition to that, a correlation matrix was computed, showing the pairwise Pearson correlation coefficients between all features and the target variable. This matrix provided a visual and numerical representation of the relationships in the data. Features with weak or negligible correlations to the target variable were identified and excluded.

Rolling statistical features, such as mean and standard deviation, were engineered for key analog and digital sensors to capture temporal patterns and trends over time. These features enhanced the model's ability to detect patterns leading to failures by leveraging both short-term and long-term sensor behaviours. engineering rolling statistical features (mean and standard deviation) for selected analog and digital sensor readings.

These steps were taken in order to maintain the temporal integrity of the dataset. Additionally, the dataset was thoroughly examined for class imbalance (Fig.2), this imbalance can lead to biased models that overly focus on the majority class while neglecting the minority class as discussed in the Related Work section.

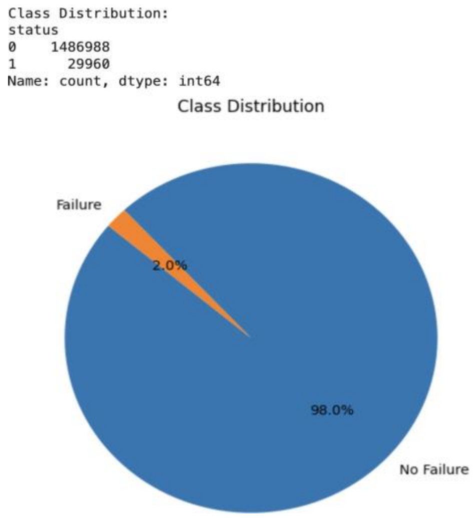

Figure 2: Class Imbalance.

Traditional methods for handling class imbalance, such as oversampling or undersampling, were not suitable for this dataset due to its time-series nature. These methods disrupt the temporal order, which is critical for maintaining the integrity of time-series data. Instead of modifying the data distribution, class weights were used to adjust the model's loss function. The way it works is: Higher weights are assigned to the minority class (failures), encouraging the model to pay more attention to these underrepresented points during training. Using class weights did not alter the sequential structure of the data, ensuring the time-series integrity was intact. In Predictive Maintenance, missing a failure (false negatives) can have severe consequences, such as unplanned downtime or equipment damage. Class weights help mitigate this risk by ensuring that the minority class is not overlooked, improving the model's ability to predict failures accurately.

5.2 Model Development

A Random Forest model was trained (Fig. 3) on the preprocessed dataset to predict failure events. The implementation included a k-fold cross-validation approach. Model performance was evaluated using metrics such as precision, recall and F1-score. Outputs included confusion matrices and feature importance rankings for each fold.

```
Training Random Forest Model...
Random Forest Model Training Completed.

Testing Classification Report:
              precision    recall  f1-score   support

  No Failure       1.00      1.00      1.00    607201
     Failure       0.95      0.73      0.83      6510

    accuracy                           1.00    613711
   macro avg       0.98      0.86      0.91    613711
weighted avg       1.00      1.00      1.00    613711
```

Figure 3: First Random Forest Model.

5.3 Explainability with SHAP

The initial SHAP implementation (Fig. 4) focused on interpreting the predictions of the Random Forest model for a specific test sample, particularly one with borderline classification probabilities. By isolating a borderline test sample (one with classification probabilities close to the decision threshold), the goal was to explore the underlying decision-making process of the model in a scenario where its predictions were less certain. This approach allowed for a focused examination of how the model evaluated key features to determine a prediction. The sample to be explained was isolated and formatted as a DataFrame (to ensure compatibility with SHAP). A TreeExplainer was initialised using the best-performing Random Forest model from the training process (After Hyperpara- meter tuning and threshold tuning). SHAP values were then calculated for the selected sample, specifically targeting class 1 ("Failure"). These values quantified the contribution of each feature to the prediction outcome. Then, a SHAP waterfall plot was generated to visually illustrate how individual feature values influenced the model's prediction for the failure class, incorporating the base prediction value to contextualise the result.

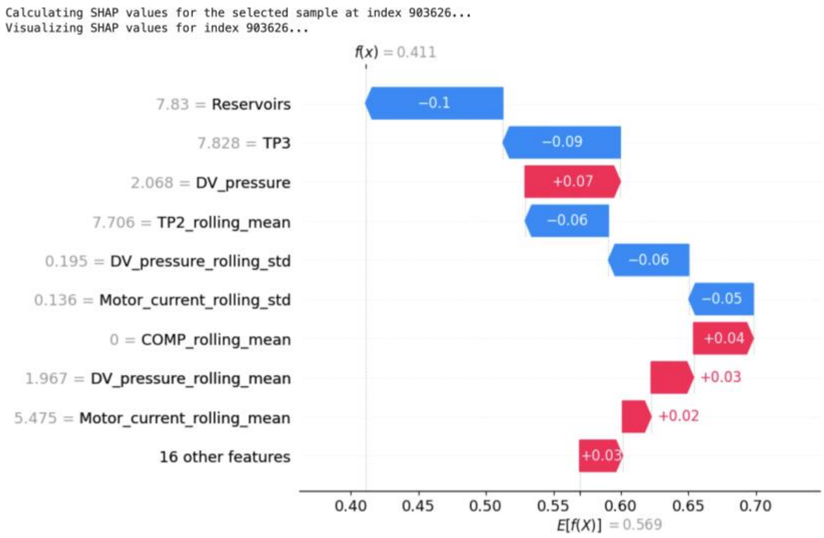

Figure 4: SHAP waterfall plot.

The SHAP waterfall plot (Fig. 4) illustrates the feature contributions for a failure prediction at test sample index 903626. The base model prediction (E[f(X)]) is 0.569, while the final prediction (f(x)) is 0.411. Positive contributions, such as "DV_pressure" and "Motor_current_rolling_mean", push the prediction toward failure, whereas negative contributions, including "Reservoirs" and "TP3", mitigate the failure likelihood. This initial breakdown supports the interpretability of the Random Forest model, enabling maintenance teams to identify and monitor critical features driving predictions.

5.4 Tools

The implementation was carried out in Python, and Google Colab was utilised. The libraries used were: Pandas (for data manipulation and preprocessing), Matplotlib (for visualising data distributions and model evaluation metrics), Scikit-learn (for Random Forest modelling and performance evaluation) and SHAP (for explainability analysis and feature contribution visualisations).

This implementation produced transformed datasets, trained Random Forest models, performance metrics and SHAP visualisations. These outputs provide interpretive insights, supporting the research objectives of enhancing model transparency.

6 Evaluation

The evaluation phase of this study focused on integrating SHAP results to the regression model pipeline in order to have some insights and refine the Predictive Maintenance framework. After interpreting SHAP values, the analysis was extended to investigate the distribution and statistical properties of high-impact features identified by SHAP, namely: Reservoirs (measurement of the pressure downstream from the reservoirs), TP3 (pressure produced at the pneumatic panel), DV_pressure (measurement of pressure drop during the discharge of air dryers in the towers), TP2(pressure on the compressor) rolling_mean, DV_pressure_rolling_std, Motor_current_rolling_std, COMP_rolling_mean, DV_pressure rolling_mean, and Motor current rolling mean. These features were deemed most influential in the prediction of machine failures, and their patterns were thoroughly explored (Fig. 5).

Statistical Summary of High-Impact Features:

		Non-Failure	Failure
Reservoirs	mean	8.987906	8.377688
	std	0.627032	0.422010
TP3	mean	8.987300	8.376050
	std	0.627787	0.422141
DV_pressure	mean	0.040628	1.989085
	std	0.330024	0.278665
TP2_rolling_mean	mean	1.108296	8.125115
	std	1.536312	0.772425
DV_pressure_rolling_std	mean	0.072287	0.224687
	std	0.130117	0.093538
Motor_current_rolling_std	mean	1.911744	0.192143
	std	0.535293	0.270006
COMP_rolling_mean	mean	0.866345	0.008331
	std	0.182782	0.068129
DV_pressure_rolling_mean	mean	0.040679	1.979349
	std	0.294824	0.199811
Motor_current_rolling_mean	mean	1.759777	5.513191
	std	1.036038	0.466228

Figure 5: High importance features after SHAP.

A statistical summary of the high-impact features was conducted, providing insights into their central tendencies and variability. Additionally, the correlation of these features with the target variable (status) was analysed to quantify their predictive strength. This analysis revealed significant trends and highlighted key predictors that were strongly associated with failure events.

Based on the SHAP, a list of High-Impact feature was created and its distribution was analysed as can be seen in Figures 6, 7 and 8.

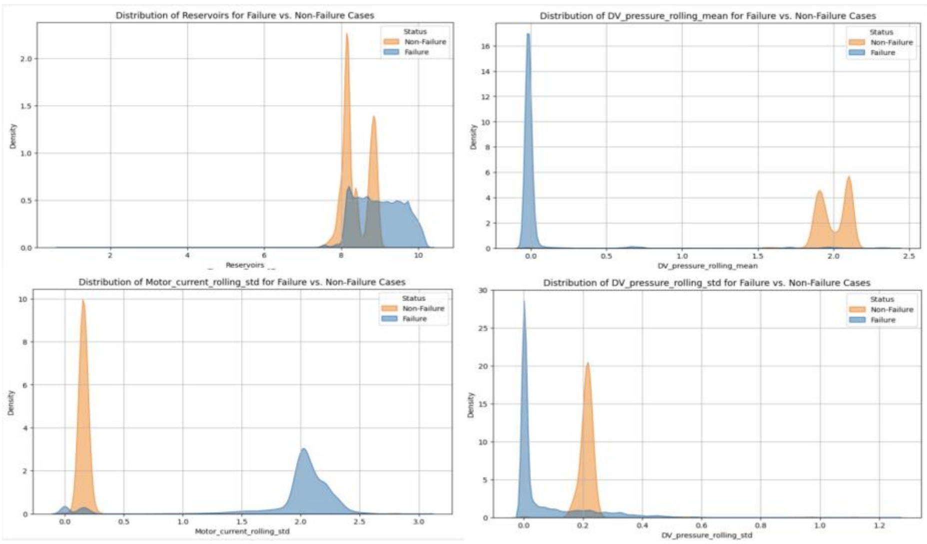

Figure 6: Class distribution of High-Impact features (Part 1).

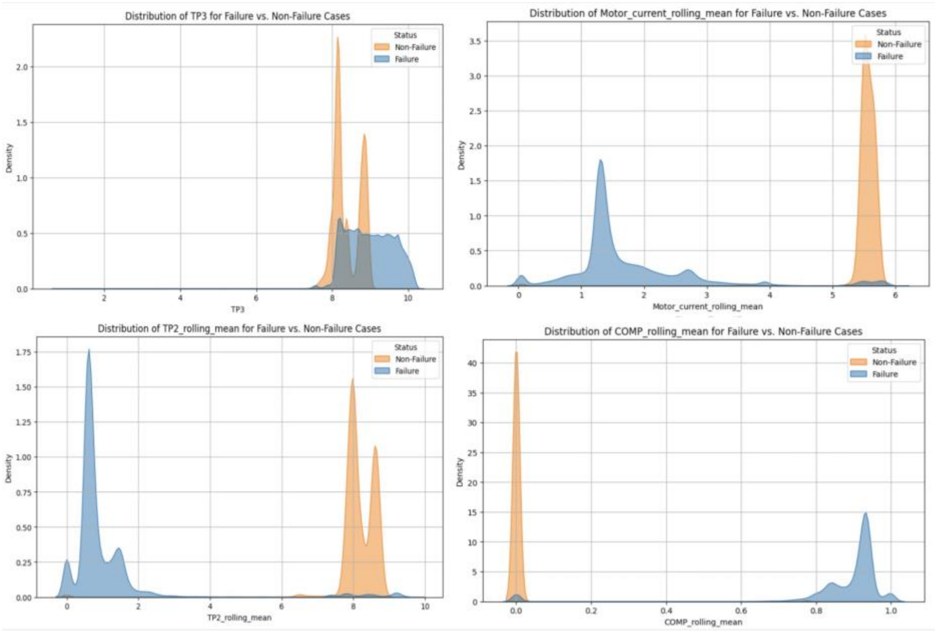

Figure 7: Class distribution of High-Impact features (Part 2).

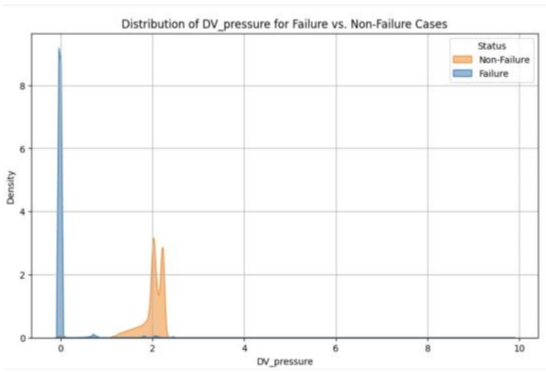

Figure 8: Class distribution of High-Impact features (Part 3).

Taking into account the insights provided by both SHAP and statistical findings, transformations and aggregations were applied to enhance the dataset further. New Feature Creation: Categorical bins were defined for DV pressure and Motor current to capture operational thresholds that could indicate impending failures. This was particularly useful for interpreting and acting upon sensor data effectively. Thresholds for Early Warning: Specific thresholds were established for features in order to try to implement an early warning system. Rolling statistical measures were monitored to flag abnormal trends indicative of failure risks.

Existing features were refined to align with the SHAP-detected importance, focusing on improving their predictive utility. A new Random Forest model was trained using the updated dataset that included the newly engineered features and refined data. Class weights were recalculated to address any remaining imbalance in the dataset, ensuring fair representation of failure cases.

After training, threshold tuning (Fig. 9) was performed once again in order to optimise model performance, targeting the best trade-off between precision, recall and F1-score. This step aimed to enhance the reliability of the failure predictions while maintaining a practical focus on actionable insights for early interventions (Fig. 10).

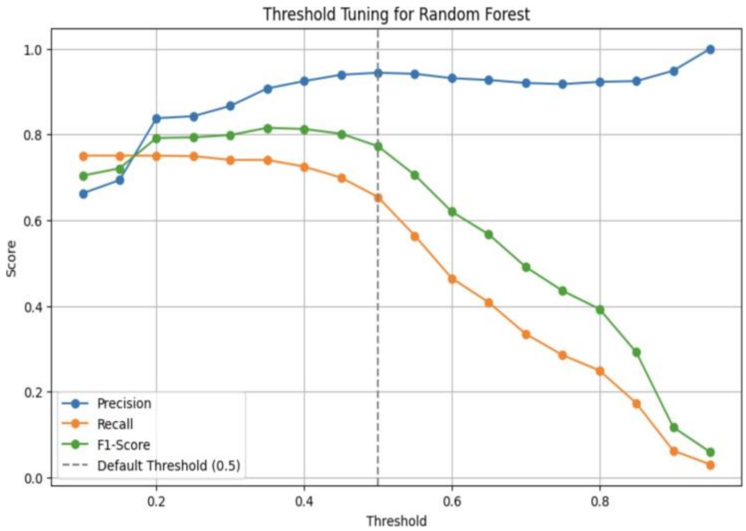

Figure 9: Optimal threshold tuning after SHAP analysis.

```
Optimal Threshold: 0.40
Precision: 0.94, Recall: 0.74, F1-Score: 0.83
Confusion Matrix at Optimal Threshold:
[[606908    293]
 [  1688   4822]]

Classification Report at Optimal Threshold:
              precision    recall  f1-score   support

  No Failure       1.00      1.00      1.00    607201
     Failure       0.94      0.74      0.83      6510

    accuracy                           1.00    613711
   macro avg       0.97      0.87      0.91    613711
weighted avg       1.00      1.00      1.00    613711
```

Figure 10: Random Forest after feature selection with SHAP analysis.

The integration of SHAP facilitated understanding of feature importance, enabling targeted transformations and strategic feature engineering. Refining high-impact features and incorporating thresholds led to a model that was more interpretable.
After integrating SHAP for feature-level analysis and using the optimised threshold, the Random Forest model's performance metrics showed a slight improvement. At an optimal threshold of 0.40, the model achieved a precision of 0.94, recall of 0.74 and F1-score of 0.83 for the Failure class. These minor improvements reflect the model's ability to reduce false positives while maintaining reasonable recall levels for failure predictions. The confusion matrix highlights the classification results: for the No Failure class, the model achieved perfect precision and recall (1.00), with 606,908 true positives and only 293 false negatives. For the Failure class, the model identified 4,822 true positives and 1,688 false negatives, achieving a precision of 0.94 and a recall of 0.74. The weighted averages for accuracy, precision, recall, and F1-score were all 1.00, showing a good overall performance. However, macro averages (Precision: 0.97, Recall: 0.87, F1-score: 0.91) provide a more balanced view, highlighting the disparity between the dominant No Failure class and the minority Failure class, underscoring the potential for further refinement in handling imbalanced datasets. In the context of Predictive Maintenance (PdM), these values are considered acceptable, as prioritizing inspections over unexpected machinery downtime due to failures is preferable. This approach allows for the reporting of subtle warning signals to prevent potential issues.

It is worth addressing several important considerations and limitations. Threshold tuning involved a trade-off between precision and recall, which could impact the model's effectiveness in scenarios where identifying all failures is critical. Additionally, the decision to avoid oversampling or undersampling techniques limited the exploration of alternative strategies for addressing class imbalance, potentially constraining the model's performance. Several adjustments were made to enhance the integration of SHAP into the analysis process, although some configurations could not be fully implemented due to technical limitations. For instance, the SHAP configuration was adjusted to use feature perturbation="interventional", providing greater flexibility in handling the dataset. Additionally, additivity checks were disabled to prevent errors arising from discrepancies in the sums of SHAP values, ensuring smoother calculations. To address memory and performance constraints, a manageable subset of data, such as 1,000 rows, was selected for SHAP calculations, to no avail. Although the errors limited the scope of the dependence plots, the SHAP summary plots still provided information about the most impactful features.

7 Conclusion and Future Work

This research set out to address the identification of influential features for predicting failures in time-series data using SHAP and exploring the effectiveness of Random Forest models in generating actionable predictive insights in resource-constrained environments. By leveraging SHAP for interpretability and Random Forest for predictive capabilities, the study sought to balance predictive accuracy with model transparency, a critical requirement for real-world Predictive Maintenance (PdM) applications.

Key findings of this study include the identification of high-impact features such as 'Reservoirs,' 'TP3,' and 'DV_pressure' that significantly influence failure predictions. The integration of SHAP provided important insights that not only enhanced model transparency but also guided the refinement of data preprocessing and feature engineering steps, such as creating rolling statistical features and categorical bins. These insights allowed for more meaningful interpretations of model predictions and a better understanding of failure patterns.

The results revealed that while the initial Random Forest model achieved good precision and recall, SHAP analysis and threshold optimisation improved the model's ability to identify failures. At an optimised threshold, the model achieved a precision of 0.94 and a recall of 0.74 for the Failure class, meaning that the model was able to capture critical failure events with fewer false positives. Despite these advances, challenges such as the imbalanced distribution of failure events and the limitations of time-series stratification presented obstacles that required workarounds, such as using class weights and preserving the sequential integrity of the data during splits.

This study demonstrated the combination of Random Forest and SHAP in predictive maintenance, providing insights while addressing transparency concerns. However, limitations remain. The lack of advanced time-series stratification techniques and the decision not to employ oversampling or undersampling approaches may have influenced the model's ability to fully address class imbalance. Additionally, the focus on a single dataset limits the generalisation of the findings to other Predictive Maintenance scenarios.

The potential of integrating interpretable AI techniques like SHAP into Predictive Maintenance workflows to bridge the gap between complex machine learning models and practical applications should be highlighted. From an industrial perspective, the findings underscore the importance of leveraging data-driven methodologies to enhance maintenance planning and reduce downtime, especially in resource-constrained settings.

Future research could focus on addressing the limitations identified in this study. For instance, implementing stratified time-series cross-validation techniques could better evaluate the model's performance in imbalanced datasets while preserving temporal dependencies. Additionally, exploring alternative models, such as Gradient Boosting Machines or deep learning approaches like LSTMs, could further enhance prediction accuracy, especially for datasets with complex temporal patterns. Another possibility could be the integration of real-time monitoring systems with predictive models, enabling dynamic updates to feature importance and thresholds based on live data streams.

References

Ahmadzadeh, A., Hostetter, M., Aydin, B., Georgoulis, M.K., Kempton, D.J., Mahajan, S.S., & Angryk, R.A., 2019. Challenges with extreme class-imbalance and temporal coherence: A study on solar flare data. *arXiv preprint*.

Aji, A. S., Sashiomarda, J. A., & Handoko, D. (2020). Predictive maintenance magnetic sensor using random forest method. *Journal of Physics: Conference Series*, *1528*, 012030.
Amer, S., Mohamed, H.K., & Mansour, M.B., 2023. Predictive Maintenance by Machine Learning Methods. *Proceedings of the IEEE Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS)*. DOI: 10.1109/ICICIS58388.2023.10391130.

Batal, I., Sacchi, L., Bellazzi, R. & Hauskrecht, M., 2009. Multivariate Time Series Classification with Temporal Abstractions. *Proceedings of the Twenty-Second International FLAIRS Conference.*

Bernard, J., Ruppert, T., Goroll, O., May, T., & Kohlhammer, J., 2012. Visual-Interactive Preprocessing of Time Series Data. *Proceedings of SIGRAD 2012*, pp. 39–48.

Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C.L.P., 2024. A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review*. DOI: 10.1007/s10462-024-10759-6.

Cummins, L., Sommers, A., Ramezani, S.B., Mittal, S., Jabour, J., & Seale, M., 2024. Explainable Predictive Maintenance: A Survey of Current Methods, Challenges, and Opportunities. *arXiv preprint*.

Davari, N., Veloso, B., Ribeiro, R.P., Pereira, P.M., & Gama, J., 2021. MetroPT-3 Dataset Description for Predictive Maintenance. *INESC TEC*.

Fan, C., Ding, L., Luo, X. & Wang, J., 2021. A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, 9.

Flach, P.A., & Kull, M. (2015). Precision-Recall-Gain curves: PR analysis done right. In Advances in Neural Information Processing Systems, 28.

Gawde, S., Patil, S., Kumar, S., Kamat, P., Kotecha, K., & Alfarhood, S. (2024). Explainable Predictive Maintenance of Rotating Machines Using LIME, SHAP, PDP, ICE. *IEEE Access, 12*, 29345–29359. https://doi.org/10.1109/ACCESS.2024.3367110.

Geurts, P., 2001. Pattern Extraction for Time Series Classification. *Lecture Notes in Computer Science (LNCS), 2168*, pp.115–127. Springer-Verlag Berlin Heidelberg.

He, D., Zhang, X., Ge, C., & Chen, E. (2022). A Novel Reliability-Centered Opportunistic Maintenance Strategy for Metro Train Complex Systems. *IEEE Intelligent Transportation Systems Magazine*, 14(3), 146-153. DOI: 10.1109/MITS.2020.3014080.

Jabbar, M.A., Samreen, S. & Aluvalu, R., 2018. The future of health care: Machine learning. International Journal of Engineering and Technology, 7(4.6), pp.23-25.

Jafarigola, E. & Trafalis, T.B., 2023. A review of machine learning techniques in imbalanced data and future trends. *Preprint submitted to Elsevier, October 11, 2023*.

Johnson, J.M. & Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. *Journal of Big Data, 6*(27). DOI: 10.1186/s40537-019-0192-5.

Karlsson, N. & Bengtsson, M., 2022. Explainable AI For Predictive Maintenance. *Master's Thesis, Halmstad University*.
Katrakazas, C., Antoniou, C. & Yannis, G., 2019. Classification of imbalanced datasets in time-series data for real-time safety assessment. *Proceedings of the TRB Annual Meeting*.

Kedar, M.M. & Mhatre, G.N., 2024. Exploring the Effectiveness of SHAP over other Explainable AI Methods. *International Journal of Scientific Research in Engineering and Management, 8*(6). DOI: 10.55041/IJSREM35556.

Lemes, L. C., & Hvam, L. (2019). Maintenance Costs in the Process Industry: A Literature Review. *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 1481-1484. DOI: 10.1109/IEEM.2019.8914712.

Leukel, J., González, J., & Riekert, M., 2021. Adoption of Machine Learning Technology for Failure Prediction in Industrial Maintenance: A Systematic Review. *Journal of Manufacturing Systems, 61*, pp.87–96. DOI: 10.1016/j.jmsy.2021.08.012.

Lipton, Z. C., Elkan, C., & Naryanaswamy, B. (2014). Thresholding Classifiers to Maximize F1 Score. *arXiv preprint arXiv:1402.1892v2*.

Maashi, M. S., Alwhibi, N., Alamr, F., Alzahrani, R., Alhamid, A., & Altawallah, N. (2020). Industrial duct fan maintenance predictive approach based on Random Forest. *Computer Science & Information Technology (CS & IT), 10*(5), 177-184. DOI:10.5121/csit.2020.100516.

Madhu Puttegowda, S., & Ballupete Nagaraju. (2024). Artificial Intelligence and Machine Learning in Mechanical Engineering: Current Trends and Future Prospects. *Engineering Applications of Artificial Intelligence, 142*, 109910.

Moniz, N., Branco, P., & Torgo, L. (2016). Resampling Strategies for Imbalanced Time Series. *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. doi:10.1109/DSAA.2016.37.

Niaz, N.U., Shahariar, K.M.N., & Patwary, M.J.A., 2022. Class imbalance problems in machine learning: a review of methods and future challenges. *Proceedings of ICCA 2022*. DOI: 10.1145/3542954.3543024.

Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., & Frontoni, E., 2018. Machine Learning Approach for Predictive Maintenance in Industry 4.0. *14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*. DOI: 10.1109/MESA.2018.8449214.

Prihatno, A. T., Nurcahyanto, H., & Jang, Y. M. (2021). Predictive Maintenance of Relative Humidity Using Random Forest Method. *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 497-499. DOI:10.1109/ICAIIC51459.2021.9415213.

Puthiya Parambath, S.A., Usunier, N., & Grandvalet, Y. (2014). Optimizing F-measures by cost-sensitive classification. In Advances in Neural Information Processing Systems, 27.

Rai, A., Rout, B. P., Bhushan, & Mehta, A. (2024). Unveiling the power of data: A journey through machine learning techniques. International Journal of Advanced Research in Science, Communication and Technology, 4(7), 42-47.

Rudin, C., 2019. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence, 1*(5), pp. 206–215.

Salman, H.A., Kalakech, A. & Steiti, A., 2024. Random Forest Algorithm Overview. *Babylonian Journal of Machine Learning, 2024*(69-79). [File: *RF.pdf*]. DOI: https://doi.org/10.58496/BJML/2024/007.

Samek, W., Wiegand, T., & Müller, K.-R., 2017. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv preprint*, arXiv:1708.08296v1.

Shafique, U. & Qaiser, H., 2014. A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, 12(1), pp.217–222.

Singh, Y., Bhatia, P. K., & Sangwan, O. (2006). A review of studies on machine learning techniques. *International Journal of Computer Science and Security, 1*(1), 70-84.

Susto, G.A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A., 2015. Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Transactions on Industrial Informatics, 11*(3), pp.812–820. DOI: 10.1109/TII.2014.2349359.

Van den Broeck, G., Lykov, A., Schleich, M., & Suciu, D., 2020. On the Tractability of SHAP Explanations. *arXiv preprint arXiv:2009.08634v1*.

Vinh, T. Q., & Huy, N. T. (2022). Predictive Maintenance IoT System for Industrial Machines using Random Forest Regressor. *2022 International Conference on Advanced Computing and Analytics (ACOMPA)*. DOI:10.1109/ACOMPA57018.2022.00020.

Yang, C., Fridgeirsson, E. A., & Kors, J. A. et al. (2024). Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data. *Journal of Big Data*, *11*(7). https://doi.org/10.1186/s40537-023-00857-7.