Predictive Maintenance of Banbury Mixer Using Machine Learning Methods: A Case Study of a Tire Manufacturing Factory

Pattaraporn Nueasri[1], Chanokporn Smuthkalin[2,*], Anan Butrat[1], Siriwan Polset[3] and Nuntarat Sookpunya[4]

[1] Department of Industrial Management Engineering, Faculty of Industrial Technology, Valaya Alongkorn Rajabhat University under the Royal Patronage, Khlong Nueng, Khlong Luang, Pathum Thani, 13180, Thailand

[2] Department of Technology Management, Faculty of Industrial Technology, Valaya Alongkorn Rajabhat University under the Royal Patronage, Khlong Nueng, Khlong Luang, Pathum Thani, 13180, Thailand

[3] Department of Mechatronics and Robotics Engineering, Faculty of Industrial Technology, Valaya Alongkorn Rajabhat University under the Royal Patronage, Khlong Nueng, Khlong Luang, Pathum Thani, 13180, Thailand

[4] Department of Mechanical Engineering Technology, Faculty of Industrial Technology, Valaya Alongkorn Rajabhat University under the Royal Patronage, Khlong Nueng, Khlong Luang, Pathum Thani, 13180, Thailand

***Corresponding Author E-mail:** chanokporn.s@vru.ac.th

Received: Sep 26, 2024; **Revised:** Dec 17, 2024; **Accepted:** Dec 24, 2024

Abstract

This study investigates parameters of the Banbury mixer, focusing on Material Rubber Sheet (MRS) production duration, to develop a machine learning-based failure detection system. Proposed data preprocessing methods aim to uncover patterns of the relationship between current and previous batch parameters. While data differencing suffices for SVM and RF models, it minimally impacts ANN. Solely employing data cleaning renders RF suitable for model creation. SVM outperforms RF in failure detection but may produce occasional false alarms. Performance evaluation indicates RF accurately detects 13 out of 29 failures, surpassing SVM detecting 16 failures with 2 false alarms. The SVM model from data cleaning combined with differencing reduces waste by 55.17%, surpassing RF by 6.89%. Future research could explore advanced preprocessing for failure cause categorization and leverage sophisticated techniques.

Keywords: Banbury mixer, Failure detection, Machine learning, Data preprocessing, Waste reduction

1. Introduction

In the manufacturing industry, Banbury mixing machines stand as indispensable tools for achieving precise blends of rubber, plastics, and assorted chemical compounds, priming them for subsequent processing. These machines boast a distinctive design, featuring dual rotors that pivot in opposing directions within a figure-eight shaped chamber. This configuration engenders a potent shearing action, facilitating the thorough integration of diverse ingredients and additives. In this study, the tire manufacturing factory under scrutiny relies heavily on Banbury mixers, deploying a total of six units to fulfil its operational demands. These mixers play a pivotal role in producing three primary compound types: Material Rubber Sheet (MRS), non-productive compounds, and productive compounds. With their versatility and efficiency, Banbury mixers emerge as linchpins in the factory's production process, enabling the precise formulation of compounds essential for crafting high-quality tires.

The challenge at hand revolves around the unexpected periods of downtime that disrupt the flow of production. Equipment downtime, characterized by the interval during which machinery remains inactive, whether due to unanticipated malfunctions or scheduled maintenance pauses, poses a significant threat to productivity, revenue streams, and customer satisfaction within industrial settings. Among the myriad factors contributing to equipment downtime, deviations in temperature and torque settings emerge as primary culprits impacting the functionality of Banbury mixers. Identifying these parameters as critical indicators, efforts are underway to scrutinize their patterns and discern early warning signals of impending failures. However, the scope of this endeavour is constrained by the complexity of Banbury mixer operations. Each mixer is tasked with producing a diverse array of compounds, with each compound necessitating distinct parameter configurations. Hence, the study focuses on compounds with the highest production utilization to optimize resource allocation effectively. Material Rubber Sheet (MRS) emerges as a focal point in this investigation, accounting for a substantial portion of production activity. With 1,246 batches produced out of 7,975, constituting 15.62% of total available time, MRS commands significant attention. Notably, MRS production predominantly occurs during night shifts, when maintenance teams are typically unavailable, further accentuating the need for pre-emptive failure detection mechanisms in this critical operational window.

Furthermore, the parameters pertinent to the Banbury mixing machine, serving as pivotal indicators of potential failure states, are delineated as follows:

- Mixing time: denoting the duration over which materials undergo blending for a single batch of compound.
- Time between batches: signifying the interval during which the Banbury mixer remains inactive, typically comprising a 900-second setup time and a 30-second material feeding period per batch.
- Temperature: reflecting the thermal conditions generated within the machine throughout the mixing process.
- Torque power: representing the rotational force exerted by the machine on the materials during mixing operations.
- Electrical energy: indicative of the power consumption of the motor driving the mixer.
- Material weight: denoting the mass of the substances undergoing mixing within the machine.
- Failure status: defined by instances where the time lapse between batches exceeds the predetermined threshold of 900 seconds.

Real-time recording of all parameters is seamlessly integrated into the operational framework, enabling comprehensive data collection on a batch-by-batch basis. Furthermore, through vigilant observation, the case study company can discern potential failure states, particularly when fluctuations in torque values occur abruptly, as depicted in **Figure 1**. However, the analytical scope extends beyond torque alone, with an additional four parameters recorded and available for study to unveil patterns indicative of impending failures, as illustrated in **Figure 2**.

The proposed predictive maintenance (PdM) model aims to preemptively identify equipment failures immediately after the conclusion of each mixing batch, enabling swift remedial action through the reconfiguration of all parameters before commencing the subsequent batch. This approach stands in stark contrast to the current process, wherein failures often go undetected until after materials have been fed into the mixer, as depicted in **Figure 3**. The imperative for this proactive strategy lies in its potential to mitigate downtime by averting unforeseen disruptions. If a failure is detected before it occurs, the subsequent repair time remains within the realm of normal setup durations (900 seconds).

Conversely, failure detection post-material feeding can result in substantially prolonged repair times, ranging from 901 to 95,216 seconds, with a mean time to repair (MTTR) of 3,268.51 seconds, as illustrated in **Figure 4**. Factors contributing to this extended MTTR include the necessity to extract materials already fed into the mixer, thereby compounding downtime. Moreover, material wastage is a consequential concern if failures are only detected post-mixing initiation, as incomplete batches must be discarded. Consequently, the predictive model under examination promises threefold benefits: reduction in downtime, heightened productivity, and mitigation of material wastage stemming from prematurely fed batches.

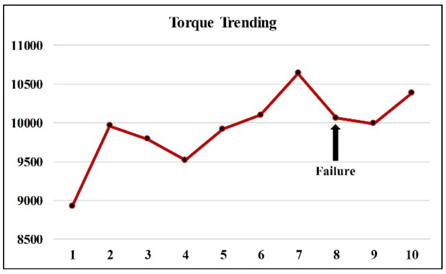

Figure 1 Trending of Torque

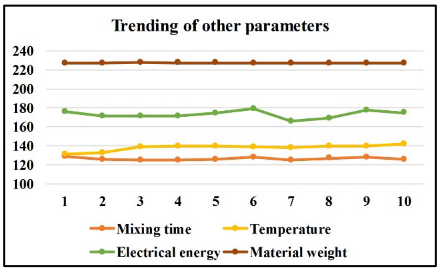

Figure 2 Trending of Banbury mixer's parameters

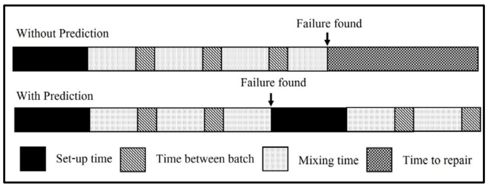

Figure 3 The concept of this predictive maintenance

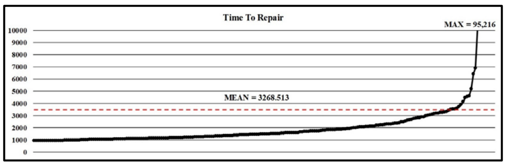

Figure 4 A Time to repair for a failure occurring

Existing predictive maintenance (PdM) methods have made significant progress in industries but often fall short when addressing the specific challenges posed by Banbury mixers. These mixers operate under complex and dynamic parameters, such as torque power and temperature, which complicate failure detection and prediction. Traditional PdM approaches using sensor-based monitoring or machine learning models often focus on generalized industrial equipment and may not account for the unique patterns and anomalies associated with Banbury mixing processes.

For instance, one study implemented a machine learning-based PdM model in manufacturing, successfully reducing downtime and enhancing equipment reliability through real-time monitoring of operational conditions and failure prediction using IoT sensors [1]. Similarly, PdM for automotive systems integrated structured data with machine learning algorithms to predict remaining useful life and minimize unexpected breakdowns, showcasing its applicability in equipment with repetitive, measurable wear patterns. While these methods are effective, their generality limits their application to specialized equipment like Banbury mixers, where operational parameters interact in non-linear and machine-specific ways.

Moreover, existing predictive maintenance approaches may struggle with capturing the unique operational dynamics of Banbury mixers or handling imbalances in failure occurrences versus normal operations. These limitations can lead to inefficiencies and missed opportunities for timely interventions.

As **Figure 5**, using the application of data cleaning methods, the performance of four predictive maintenance models—Support Vector Machine (SVM), Random Forest (RF), 1-layer Artificial Neural Network (1-layer ANN), and 2-layer Artificial Neural Network (2-layer ANN)—was assessed. Among these, the Random Forest model was the only one capable of detecting failures, successfully identifying 14 failure instances. The remaining models (SVM, 1-layer ANN, and 2-layer ANN) failed to detect any failures under the same conditions. Moreover, no additional performance measurements, such as downtime reduction, productivity improvements, or waste reduction, were observed during this analysis. This suggests that while the Random Forest model excels in failure detection, its broader impact on operational metrics may require further investigation and optimization.

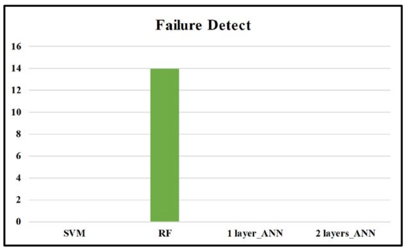

Figure 5 Applying the Existing Methods

This gap highlights the necessity for a tailored PdM approach, leveraging advanced data preprocessing techniques and machine learning to address the nuanced challenges of Banbury mixers, such as identifying correlations between operational metrics and failures. By focusing on these complexities, this study aims to bridge the limitations of existing methods and provide a robust solution for predictive maintenance in specialized industrial contexts as comparison to the previous studies as **Table 1**.

However, the adoption of predictive maintenance (PdM) introduces not only potential benefits but also risks, notably the possibility of increased downtime due to erroneous failure predictions. Traditionally, when a failure is identified (FF), the time required for repair of the affected batch is reduced from the mean time to repair (MTTR) to the setup time (ST). Conversely, if the failure remains undetected (FNF), the batch's repair time remains at the MTTR level. However, there exists a third scenario, termed the fake failure found (FFF) case, which arises when the predictive model erroneously identifies a failure that does not actually exist. In this situation, the time needed for repair of the batch increases from the usual time between batches to the setup time (ST).

This research aims to address significant downtime and material waste in the Banbury mixing process by developing a comprehensive predictive maintenance (PdM) system. The study introduces an advanced data preprocessing technique to uncover critical patterns within mixer parameters, forming the foundation for a robust failure detection model powered by machine learning. By analyzing data from ongoing batches, the research applies three distinct machine learning methodologies to evaluate model accuracy and effectiveness. Additionally, it proposes an engineering framework to assess performance improvements from PdM implementation and estimate associated risks, ultimately contributing to enhanced operational efficiency and reduced production waste.

2. Literature Review

2.1 Previous Data Preprocessing

Data preprocessing stands as a foundational step in the development of predictive maintenance (PdM) models, pivotal in readying the data for subsequent analysis and modeling endeavors. This preparatory phase encompasses a range of techniques, including but not limited to:

- Data cleaning: This involves finding, removing, and replacing bad or missing data, such as outliers, noise, or errors [2–4].
- Data transformation: This involves changing the data from one domain to another, such as time to frequency, or linear to nonlinear. This can help reveal hidden patterns or features in the data [5],[6].
- Data scaling or normalization: This involves changing the range or distribution of the data, such as rescaling, standardizing, or normalizing. This can help reduce the effect of different units or scales on the data analysis [7].
- Data feature extraction: This involves reducing the dimensionality of the data by selecting or creating relevant features that capture the most information. This can help improve the performance and interpretability of the data analysis [8–11].

Nevertheless, these data preprocessing techniques do not encompass data transformation, which involves converting the recorded values at specific time points into different values between two time points. This transformation yields patterns that can be classified and used to predict failures.

Table 1 Comparison of data preporcessing and performance measurement between previous studies and the proposed method

Previous Studies	Data Preprocessing						Performance Measurement							
	Data cleaning	**Data**	**Data scaling**	**Data feature extraction**	**Data differencing**	**Data absolute difference**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**ROC AUC**	**Downtime**	**Productivity**	**Waste reduction**
O. Masmoudi et al. [2]	✓													
A. Canito et al. [3]	✓													
Z. Znaidi [4]	✓													
I. Mahmud et al. [5]		✓												
S. Putchala et al. [6]		✓												
S. Lu et al. [7]			✓											
F. Gatta et al.[8]				✓										
S. Panagou et al.[9]				✓										
F. Calabrese et al. [10]				✓										
A. Shylendra et al. [11]				✓			✓							
H. Xu and J. A. Prozzi [12]							✓							
B. P. Mota et al. [13]							✓				✓			
A. F. Azyus [14]								✓						
P. Ngwa and I. Ngaruye [15]								✓	✓					
A. J. Alfaro-Nango et al. [16]									✓					
S. Gautam et al. [17]									✓	✓	✓			
K. Patel and A. Shanbhag [18]														
Proposed					✓	✓						✓	✓	✓

2.2 Previous Performance Measurement

PdM serves as a proactive technique aimed at evaluating asset conditions and executing maintenance actions when necessary to uphold them in their prime operational state. Assessing the performance of a PdM model necessitates selecting suitable metrics that align with business objectives and problem characteristics. Depending on whether the problem manifests as a classification or regression scenario, distinct metrics come into play. In classification problems, where the model predicts whether an asset will fail within a specific time frame, several common metrics are outlined as follows.

- Accuracy: the proportion of correct predictions among all predictions. This metric is suitable for balanced datasets, where the classes are equally distributed. However, it can be misleading for imbalanced datasets, where one class is much more frequent than the other. For example, if only 1% of the assets fail, a model that always predicts no failure will have 99% accuracy, but it will not be useful for PdM [12–14].
- Precision: the proportion of true positives among all positive predictions. This metric measures how reliable the model is when it predicts a failure. A high precision means that the model rarely makes false alarms, which can reduce unnecessary maintenance costs. However, precision does not account for the missed failures, which can also be costly [15],[16].
- Recall: the proportion of true positives among all actual positives. This metric measures how good the model is at detecting failures. A high recall means that the model can capture most of the failures, which can prevent breakdowns and production stoppages. However, recall does not account for the false alarms, which can also be wasteful [16–18].
- F1-score: the harmonic mean of precision and recall. This metric balances both aspects of the model's performance and gives a higher score when both precision and recall are high. This metric is suitable for imbalanced datasets, where the positive class (failure) is rare and important to detect [16].
- ROC AUC: the area under the receiver operating characteristic curve. This metric measures how well the model can distinguish between the two classes, regardless of the chosen threshold. A high ROC AUC means that the model can assign higher probabilities to the positive class (failure) than to the negative class (no failure). This metric is also suitable for imbalanced datasets, as it is not affected by the class distribution [13],[16].

In the context of predictive maintenance (PdM), the risk pertains to the potential loss or harm that may arise from utilizing the model to guide maintenance decisions. Estimating this risk involves contemplating both the uncertainty surrounding the model's predictions and the impact of these predictions. However, current measurements often fail to account for the risk associated with model implementation, neglecting to consider consequences when erroneous failure signals prompt preemptive actions.

2.3 ML Classification methods

PdM involves employing machine learning techniques to monitor system conditions and anticipate maintenance requirements or failures. This approach is instrumental in cost reduction, enhancing reliability, and optimizing system performance. Various machine learning methods can be applied in PdM, selected based on factors such as data type and availability, system complexity, and desired outcomes. The choice of algorithms in this study is critical for addressing the complexities of failure detection in Banbury mixers. The Radial Basis Function (RBF) kernel in Support Vector Machines (SVM) is selected due to its ability to handle non-linear relationships within the dataset, effectively transforming features like torque, power, and temperature into a higher-dimensional space for better separability. Similarly, the Random Forest (RF) algorithm is employed for its strength in managing feature interactions and robustness against noise in imbalanced datasets. Additionally, Artificial Neural Networks (ANN) are utilized for their adaptability in capturing complex, non-linear patterns through layers of interconnected nodes. The combination of these methods ensures a comprehensive approach, leveraging the unique strengths of each to enhance the accuracy and reliability of failure detection models. Several commonly used methods include:

- Random forest (RF) is an ensemble method that combines multiple decision trees to produce a more accurate and robust prediction. Each tree is trained on a random subset of the data and features, and the final prediction is obtained by averaging or voting the predictions of all the trees. RF can handle large and complex datasets, deal with missing values and outliers, and provide feature importance measures. However, RF can also be slow to train and test, prone to overfitting, and difficult to interpret [19],[20].
- Support vector machine (SVM) is a kernel-based method that finds the optimal hyperplane or boundary that separates the data into different classes. SVM can handle nonlinear and high-dimensional data, achieve high accuracy and generalization, and provide a clear margin of separation. However, SVM can also be sensitive to noise and outliers, require careful selection of the kernel and parameters, and suffer from scalability issues [13],[15],[16],[21–23].
- Artificial neural network (ANN) is a biologically inspired method that consists of interconnected layers of artificial neurons that process and learn from the data. ANN can approximate any complex function, capture nonlinear and interactive relationships, and adapt to changing data and environments. However, ANN can also be computationally expensive, prone to overfitting and local minima, and hard to explain and debug [24].

In practice, Support Vector Machine (SVM) classification is typically employed to distinguish between failure and non-failure states, particularly when dealing with smaller datasets comprising two or three groups. However, given the large and intricate datasets involved in this study, Random Forest (RF) emerges as a preferred method due to its capability to yield highly accurate predictions. Additionally, Artificial Neural Networks (ANN) are incorporated in this investigation for their adaptability to fluctuating data and environments, facilitating comparative analyses with other methods.

3. Methodology of research

3.1 Notation

Indices

i = Sequence of batch i

Parameters

n = Total number of compound
x_i = Training data of feature of batch i
y_i = Target data of feature of batch i
MT_i = Mixing time of batch i
TbB_i = Time between Batch of batch i
$Temp_i$ = Temperature of batch i
ToP_i = Torgue Power of batch i
EE_i = Electrical Energy of batch i
MW_i = Material weight of batch i
F_i = Failure Status of batch i
DT = Downtime
FF = Failure Found
FNF = Failure Not Found
$MTTR$ = Mean Time To Repair
FFF = Fake Failure Found
ST = Setup time

3.2 Collecting data

The data pertaining to the compound under study was automatically recorded over the course of a month, spanning 24-hour periods, and subsequently subjected to statistical analysis. The dataset comprises a total of 1,241 batches of MRS. A systematic presentation of statistical measures, organized methodically as Mean, Standard Deviation, Maximum, Minimum, and Range, is delineated in **Table 2**, illustrating the structured nature of the analysis.

3.3 Proposed Data Preprocessing

Data preprocessing constitutes a fundamental step in the domain of data mining and analysis, facilitating the transformation of raw data into a structured format suitable for interpretation and analysis by computational systems. As depicted in **Figure 6**, data preprocessing occurs subsequent to the collection of raw data and precedes its analysis by ML models.

Table 2 The statistical data of parameters

Stat	MT_i	$Temp_i$	ToP_i	EE_i	EE_i
Mean	135.63	150.51	8,487.60	224.20	221.29
SD	15.75	7.47	1,131.7	29.12	14.62
Max	304.00	181.00	10,970.00	302.20	263.19
Min	94.00	118.00	3,020.00	119.80	208.72

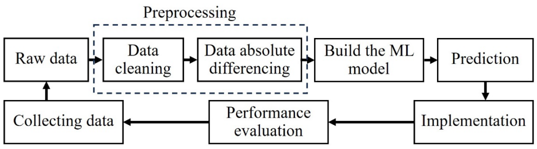

Figure 6 Steps of predictive maintenance using ML

The data absolute difference, proposed in this study, entails calculating the discrepancy between two consecutive data points, specifically the difference between the current value and the preceding value. This technique aims to predict failures based on changes observed in features such as mixing time, temperature, torque power, electrical energy, and material weight. Each training data point (x_i) is computed by subtracting the feature values of the current batch (f_i) from those of the previous batch (f_{i-1}), yielding an absolute difference, as depicted in Eq. (1).

$$x_i = |f_i - f_{i-1}| \quad (1)$$

Furthermore, each target data point (y_i) corresponds to the failure status of the subsequent batch (F_{i+1}), wherein a value of 1 indicates an impending failure in the next batch, while a value of 0 signifies the absence of an impending failure in the subsequent batch, as represented by Eq. (2).

$$y_i = F_{i+1} \quad (2)$$

For the computation of all Banbury mixer features, **Figure 7** illustrates an example of data preprocessing utilizing the data absolute difference technique.

3.4 ML models

The hyperparameter settings for this study were determined through a sensitivity analysis conducted within predefined ranges to ensure optimal performance and computational efficiency. For the ANN model, the number of nodes per layer was varied between 3 and 10, and one-layer and two-layer configurations were tested. These models, implemented using the Keras library in Python, were optimized with the Adam algorithm and employed the Rectified Linear Unit (ReLU) activation function to address the vanishing gradient problem. The models were trained for 100 epochs with a batch size of 10, employing binary cross-entropy as the loss function for binary classification tasks. For the SVM model, implemented using Scikit-learn, a radial basis function (RBF) kernel was applied, and the dataset was split into a training set (70%) and a test set (30%) for model evaluation. The RF model underwent a sensitivity analysis of tree counts ranging from 50 to 200, with bootstrap sampling and feature subsets utilized during node splitting to improve model robustness. These ranges and configurations were chosen to identify the best balance between prediction accuracy and computational efficiency, ensuring that the selected hyperparameters were empirically validated and suitable for the study's objectives.

3.5 Performance evaluation

An important benefit of implementing Predictive Maintenance (PdM) lies in its ability to effectively mitigate instances of unplanned downtime. By leveraging PdM strategies, organizations can proactively identify potential equipment failures or deterioration trends. This proactive approach equips maintenance teams with the insight needed to intervene in a timely manner, thereby preventing unplanned downtime events that disrupt production schedules and result in significant financial losses.

For performance evaluation, this study proposes a method to assess the total downtime following the implementation of the prediction model. The evaluation considers three scenarios: Failure Found (FF), Failure Not Found (FNF), and Fake Failure Found (FFF). In instances of FF and FFF, maintenance staff take proactive measures to set up the Banbury mixer before materials are fed in, resulting in a repair time equal to the setup time (ST). Conversely, in the case of FNF, maintenance actions are taken after material feeding, leading to repair times dependent on the identified issues. However, Mean Time To Repair (MTTR) was utilized in this study, derived from actual recorded data during implementation duration. Consequently, the total downtime is calculated by multiplying the number of FF by ST, adding the number of FNF multiplied by MTTR, and finally incorporating the number of FFF cases multiplied by ST, as depicted in Eq. (3).

$$Total\ DT = (No.of\ FF \times ST) + (No.of\ FNF \times MTTR) + (No.of\ FFF \times ST) \quad (3)$$
$$Waste\ reduction = Average\ MW\ \times (No.of\ current\ FNF\ -\ No.of\ new\ FNF) \quad (4)$$

i	$\lvert MT_i - MT_{i-1}\rvert$	$\lvert TbB_i - TbB_{i-1}\rvert$	$\lvert Temp_i - Temp_{i-1}\rvert$	$\lvert ToP_i - ToP_{i-1}\rvert$	$\lvert EE_i - EE_{i-1}\rvert$	$\lvert MW_i - MW_{i-1}\rvert$	F_{i+1}
2	$\lvert MT_2 - MT_1\rvert$	$\lvert TbB_2 - TbB_1\rvert$	$\lvert Temp_2 - Temp_1\rvert$	$\lvert ToP_2 - ToP_1\rvert$	$\lvert EE_2 - EE_1\rvert$	$\lvert MW_2 - MW_1\rvert$	F_3
3	$\lvert MT_3 - MT_2\rvert$	$\lvert TbB_3 - TbB_2\rvert$	$\lvert Temp_3 - Temp_2\rvert$	$\lvert ToP_3 - ToP_2\rvert$	$\lvert EE_3 - EE_2\rvert$	$\lvert MW_3 - MW_2\rvert$	F_4
4	$\lvert MT_4 - MT_3\rvert$	$\lvert TbB_4 - TbB_3\rvert$	$\lvert Temp_4 - Temp_3\rvert$	$\lvert ToP_4 - ToP_3\rvert$	$\lvert EE_4 - EE_3\rvert$	$\lvert MW_4 - MW_3\rvert$	F_5
n	$\lvert MT_n - MT_{n-1}\rvert$	$\lvert TbB_n - TbB_{n-1}\rvert$	$\lvert Temp_n - Temp_{n-1}\rvert$	$\lvert ToP_n - ToP_{n-1}\rvert$	$\lvert EE_n - EE_{n-1}\rvert$	$\lvert MW_n - MW_{n-1}\rvert$	0

Figure 7 Data absolute difference of Banbury mixer's features

Furthermore, the reduction in downtime observed in this study directly correlates with an increase in the number of batches produced without delay. Within the domain of industrial engineering, augmenting production volumes is synonymous with enhancing productivity. Productivity, defined as the ratio of output to input in a production process, quantifies the efficiency with which a system generates goods or services using available resources. Enhancing productivity has multifaceted benefits, including cost reduction, profit augmentation, and bolstered competitiveness. In this study, the output is represented by the total number of batches produced, while the input is delineated as the total production time, as illustrated in Eq. (5).

$$Productivity = \frac{Total\ No.of\ batches\ produced}{Total\ production\ time} \tag{5}$$

Moreover, waste reduction entails the strategic use of fewer materials and energy to curtail waste generation and conserve natural resources. In this study, waste reduction is achieved by pre-emptively detecting failures before materials are fed into the system, thus preventing material wastage. The waste reduction metric is calculated by multiplying the average material weight by the difference between the numbers of current Failure Not Found (FNF) instances and the numbers of new FNF instances, as demonstrated in Eq. (4).

Hence, this study will undertake a comparative analysis of these three performance indicators between the current operational scenario and a hypothetical scenario wherein the Predictive Maintenance (PdM) model is implemented.

4. Results

The results utilized Google Colab for model training, which provides access to powerful computational resources, including GPUs.

Moreover, the hyperparameter configurations for each machine learning model were derived from a comprehensive sensitivity analysis to ensure the best possible outcomes in terms of prediction accuracy and computational efficiency as **Table 3**.

Table 3 Hyperparameter configurations derived from Sensitivity Analysis

Ex.	Data pre-processing	ML model	Selected configuration
1	Cleaning	SVM	Kernel: RBF, Training-Test Split: 70%-30%
2		RF	Trees: 100, Bootstrap Sampling, Feature Subset
3		1-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10
4		2-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10
5	Cleaning + Difference	SVM	Kernel: RBF, Training-Test Split: 70%-30%
6		RF	Trees: 100, Bootstrap Sampling, Feature Subset
7		1-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10
8		2-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10
9	Cleaning + Absolute Difference	SVM	Kernel: RBF, Training-Test Split: 70%-30%
10		RF	Trees: 100, Bootstrap Sampling, Feature Subset
11		1-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10
12		2-layer ANN	Nodes: 5, Epochs: 100, Batch Size: 10

4.1 Data absolute differencing

Following the application of three data pre-processing methods including data cleaning, data differencing and data absolute difference, comparisons were made concerning the values of torque power and temperature, which constitute the primary parameters under scrutiny. These specific techniques like data differencing and absolute difference were chosen because they effectively highlight changes between consecutive batches, which are critical for identifying failure signals. The data absolute difference method particularly excels at segregating and grouping failure and non-failure instances, facilitating clearer differentiation between the two. This targeted approach enhances the detection of patterns associated with equipment failure, making it more robust than traditional preprocessing methods that may struggle to capture fine-grained changes. The orange data points signify instances where the failure status is true, representing a binary category with potential significance within the dataset's context. The three images exhibit variations in the following aspects.

Figure 8 illustrates a scatter plot matrix accompanied by histograms on the diagonal after data cleaning, providing a visual representation of the relationship between two variables and the distribution of each variable. This depiction suggests that the failure status and non-failure status are not easily distinguishable from each other. Consequently, in such instances, it may prove challenging or even impossible for machine learning models to discern patterns indicative of failure signals.

Figure 9 depicts a series of scatter plots and histograms visualizing data pertaining to two variables, torque power and temperature, following the application of data differencing. Notably, the change values between two batches of non-failure instances are clustered around the center, whereas the change values between two batches of failures are dispersed. This observation suggests that in this scenario, machine learning models may be capable of discerning patterns indicative of failure signals.

Figure 10 presents a collection of scatter plots and histograms illustrating data concerning two variables, torque power and temperature, following the application of data absolute difference. Notably, the data values between failure and non-failure instances occupy distinct regions. By utilizing the absolute difference, negative values associated with failures are shifted and clustered alongside positive values. This approach effectively segregates and groups failure and non-failure instances, facilitating clearer differentiation between the two. Consequently, this method may render the detection of failure signals easier compared to data cleaning or data difference techniques.

As a result, the data absolute difference method reliably captures patterns in changes between two points, notably between parameters of the current batch and the preceding batch.

4.2 Prediction comparison

Currently, the Predictive Maintenance (PdM) model has not been implemented. Data collected over the course of a month serves as the basis for comparison between the actual operational scenario and a hypothetical scenario wherein the PdM model is applied. Various variables are measured in the actual situation, including mean time to repair (MTTR), total number of batches, production time, average time per batch, and average material weight. The total number of batches produced amounted to 839, with a total production time of 59.32472 hours. The average time per batch was 140.1323 minutes, and the average material weight was 219.0172 kg. The MTTR was calculated at 1,490.821, indicating a significant downtime for the system. Within the actual situation, 29 failure signals were identified. This number of failures will be compared to the results obtained from models predicting failure based on data from the 839 batches.

As **Figure 11**, in Experiments 1–4, employing data cleaning alone, only the Random Forest (RF) model successfully predicted and identified 14 failure signals. Additionally, a single signal was identified by a two-layer Artificial Neural Network (ANN), albeit inaccurately classified as a failure. In Experiments 5-8, combining data cleaning with data differencing, all models successfully detected failure signals. Specifically, in Experiment 5, the model correctly identified 16 Failure Found (FF) signals, while incorrectly identifying 13 Failure Not Found (FNF) signals and 2 Fake Failure Found (FFF) signals by using SVM. Comparatively, RF emerged as the most effective classifier. Experiment 6 produced similar results, with the exception of an increase in FFF signals to 10. However, employing a one-layer ANN in Experiment 6 led to a substantial increase in FFF signals, reaching 193. Conversely, the worst performance was observed with the two-layer ANN model, which yielded 13 FF signals, 16 FNF signals, and 202 FFF signals. In Experiments 9–12, combining data cleaning with data absolute difference, SVM and RF models (Experiments 9 and 10, respectively) mirrored the results obtained in Experiments 5 and 6. This suggests that applying data absolute difference did not significantly impact SVM and RF models. However, when focusing on results from ANN models, the one-layer ANN model exhibited decreased performance, with a reduction in FF signals to 10 and an increase in FNF signals to 19, alongside a decrease in FFF signals to 127. Conversely, the two-layer ANN model showed improved performance, with an increase in FF signals to 17, a decrease in FNF signals to 12, and a reduction in FFF signals to 37.

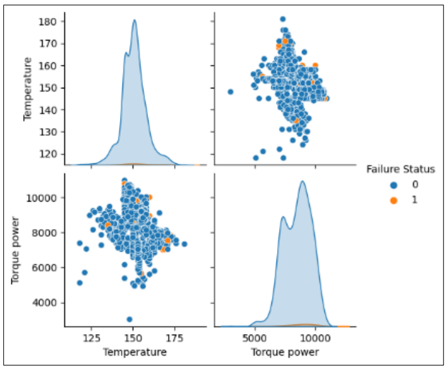

Figure 8 Data visualization of Data cleaning

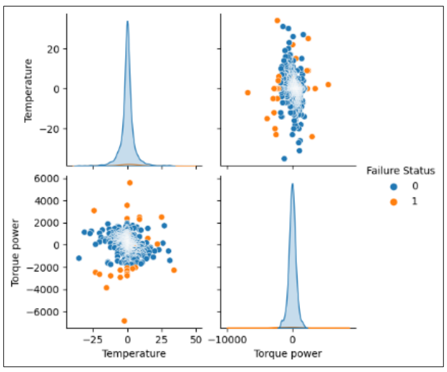

Figure 9 Data visualization of Data differencing

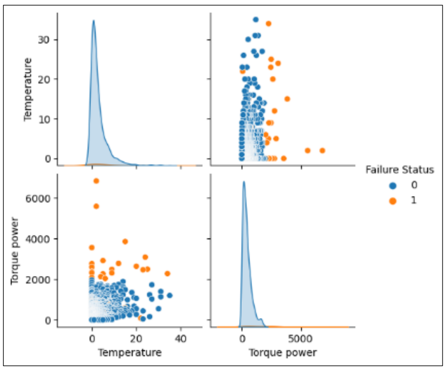

Figure 10 Data visualization of Data absolute difference

Overall, the two-layer ANN model can improve the FF signals using data absolute difference. However, FFF signals cannot reduce to be better than the other cases. The Random Forest (RF) method consistently yields results across all types of data. However, when transitioning to data differencing, the RF model performs less effectively compared to the Support Vector Machine (SVM), which generates a lower number of Fake Failure Found (FFF) signals. Moreover, it is observed that employing absolute values is unnecessary when utilizing RF and SVM models for prediction. Consequently, the optimal scenario for implementation is Experiment 5, which yields the lowest FFF count relative to FF.

4.3 Performance comparison

The analysis of actual data indicates that, in the absence of the Predictive Maintenance (PdM) model, downtime amounted to 43,233.82 seconds as shown in **Table 4**. However, with data cleaning, employing the Random Forest (RF) model proved most effective, resulting in a downtime reduction to 34,962.32 seconds, representing a 19.13% decrease as demonstrated in Experiment 2. Notably, this performance surpassed that of Experiment 5, which achieved a downtime reduction of 17.70%.

Conversely, all models utilizing Artificial Neural Networks (ANN) displayed suboptimal outcomes with a significant occurrence of Fake Failure Found (FFF) signals. This led to downtime increases exceeding 100%, highlighting the necessity of avoiding such scenarios. Furthermore, an examination of batches per hour revealed that the model from Experiment 2 exhibited the highest productivity, with a 7.04% increase compared to other models. With the implementation of Predictive Maintenance (PdM) models, waste reduction can be achieved by preventing materials from feeding into failed mixers. The magnitude of this reduction depends on the number of Failure Found (FF) instances. Based on actual data, waste was calculated at 6,351.50 kg, obtained by multiplying the average material weight (219.0172 kg) by the number of actual failures (29 failure statuses). Experiment 2, which exhibited the most significant reduction in downtime, resulted in a waste reduction of 3,285.26 kg or -48.28% compared to actual data. Furthermore, when compared to Experiment 5, Experiment 2 achieved a waste reduction of 2,847.22 kg or -55.17%, surpassing Experiment 5 by approximately 6.89%. This implies that Experiment 2 and Experiment 5 offer distinct advantages.

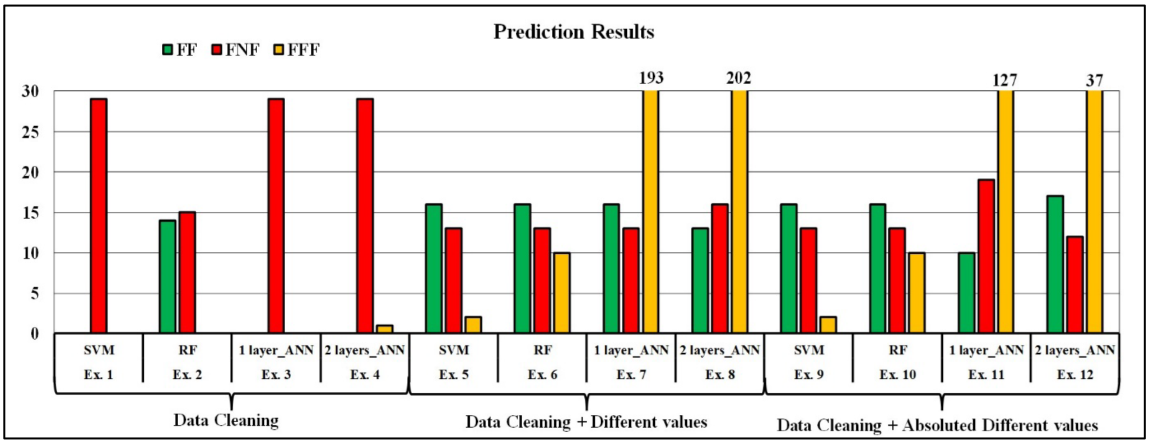

Figure 11 Comparison results between Data preprocessing and ML methods

Table 4 Comparison of Downtime, Productivity and Waste reduction

Ex.	Data pre-processing	ML model	Downtime (Seconds)		Productivity (Batches)		Waste Reduction (Kilogram)	
			Total	Diff (%)	No. of Batch per hour	Diff (%)	Weight	Diff (%)
c	-	-	43,233.82	-	14.14	-	6,351.50	-
1	Cleaning	SVM	43,233.82	0.00%	14.14	0.00%	6,351.50	0.00
2		RF	34,962.32	-19.13%	15.14	7.04%	3,285.26	-48.28
3		1-layer ANN	43,233.82	0.00%	14.14	0.00%	6,351.50	0.00
4		2-layer ANN	44,133.82	2.08%	14.03	-0.77%	6,351.50	0.00

Table 4 Comparison of Downtime, Productivity and Waste reduction (cont.)

Ex.	Data pre-processing	ML model	Downtime (Seconds)		Productivity (Batches)		Waste Reduction (Kilogram)	
			Total	Diff (%)	No. of Batch per hour	Diff (%)	Weight	Diff (%)
5	Cleaning + Difference	SVM	35,580.68	-17.70%	15.06	6.51%	2,847.22	-55.17
6		RF	42,780.68	-1.05%	14.20	0.39%	2,847.22	-55.17
7		1-layer ANN	207,480.68	379.90%	-5.61	-139.70%	2,847.22	-55.17
8		2-layer ANN	217,353.14	402.74%	-6.80	-148.10%	3,504.28	-44.83
9	Cleaning + Absolute Difference	SVM	35,580.68	-17.70%	15.06	6.51%	2,847.22	-55.17
10		RF	42,780.68	-1.05%	14.20	0.39%	2,847.22	-55.17
11		1-layer ANN	151,625.61	250.71%	1.10	-92.19%	4,161.33	-34.48
12		2-layer ANN	66,489.86	53.79%	11.35	-19.78%	2,628.21	-58.62

If the case study company prioritizes downtime reduction, Experiment 2 would be the optimal choice. However, if waste reduction takes precedence, Experiment 5 would be preferable. This performance evaluation emphasizes practical performance evaluations linked directly to operational outcomes like downtime reduction and waste minimization.

5. Discussions

In comparing the proposed Data Absolute Differencing method with previous data preprocessing techniques, several key distinctions emerge. The traditional methods, including data cleaning, transformation, scaling, and feature extraction, aim to enhance data quality, identify hidden patterns, and reduce dimensionality. These techniques are highly effective in ensuring that the data is prepared in a format suitable for analysis by standard machine learning models [1–10]. However, they may struggle to capture the fine-grained changes that signal impending failures, especially when dealing with complex parameters like torque power and temperature. In contrast, the Data Absolute Differencing method focuses on the changes between consecutive batches, which directly highlights discrepancies linked to system failures. By calculating the absolute difference between current and previous data points, this method effectively isolates failure-related signals by clustering negative and positive changes, offering a clearer differentiation between failure and non-failure instances. This focused approach enhances the detection of patterns associated with equipment failure, making it potentially more robust in identifying early failure signs compared to traditional preprocessing techniques.

In comparing the proposed performance evaluation with previous performance measurement methods, several critical differences are evident. Previous performance measurement in predictive maintenance (PdM) largely focuses on classification metrics such as accuracy, precision, recall, F1-score, and ROC AUC [11–17]. These metrics are useful for assessing the predictive capabilities of PdM models, particularly for classification tasks, but they may fall short in capturing the operational and economic impacts of PdM implementation. Metrics like accuracy, for example, can be misleading in imbalanced datasets, while precision and recall highlight specific aspects of model performance without directly reflecting the impact on maintenance decisions. In contrast, the performance evaluation in this study emphasizes a more practical approach by considering the total downtime resulting from three scenarios: Failure Found (FF), Failure Not Found (FNF), and Fake Failure Found (FFF). By calculating downtime based on real-world repair times and correlating it with productivity, this method offers a more comprehensive assessment of the PdM model's effectiveness in reducing unplanned downtime, increasing batch production, and improving productivity. This approach aligns more closely with industrial objectives, directly linking model performance to tangible outcomes such as production efficiency and waste reduction, which are crucial for operational success.

6. Conclusion

This study investigates the parameters of the Banbury mixer, focusing particularly on the production duration of Material Rubber Sheet (MRS), aiming to develop a failure detection system using machine learning methods to address high downtime and material waste resulting from feeding materials into a failed mixer. The proposed data preprocessing techniques aim to identify patterns between parameters of the current and previous batches, compared to traditional data cleaning methods. As a result, the data absolute difference method reliably captures patterns in changes between two points, notably between parameters of the current batch and the preceding batch. However, while data differencing alone suffices for creating models using Support Vector Machine (SVM) and Random Forest (RF), data absolute difference has minimal effect on these methods except for Artificial Neural Networks (ANN). Machine learning methods are adept at identifying data patterns and creating models to detect failures. However, each method requires specific data. With data cleaning alone, only the RF method is suitable for model creation, making it ideal for handling unidentified datasets. By combining data cleaning with data differencing, a preprocessing method is devised to detect failure signals. In this scenario, SVM outperforms RF in failure detection, although it occasionally produces false failure signals. Performance evaluation is crucial in determining the effectiveness of each model. Total downtime, calculated from the number of failure instances, including Failure Found (FF), Failure Not Found (FNF), and Fake Failure Found (FFF), indicates that the RF model from data cleaning detects 13 out of 29 failures with no false failures, outperforming the SVM model from data cleaning combined with data differencing, which detects 16 failures with 2 false failures. Moreover, waste reduction, defined as materials fed into the mixer before failure detection, is crucial. The SVM model from data cleaning combined with data differencing demonstrates superior waste reduction, reducing waste from 6,351.50 kg to 2,847.22 kg (-55.17%), surpassing the RF model by 6.89%. Ultimately, the case study company prioritizes waste reduction and decides to implement the SVM model, considering its superior performance in waste reduction as the primary criterion.

The findings from this research on predictive maintenance (PdM) for Banbury mixers have significant implications for various manufacturing industries. By implementing tailored data preprocessing techniques, such as data absolute differencing, companies can enhance their ability to detect equipment failures proactively, leading to reduced unplanned downtime and improved production efficiency. The study demonstrates that organizations can achieve substantial cost savings by minimizing material waste and optimizing resource utilization, ultimately contributing to greater operational sustainability. These methodologies can be adapted across different sectors that rely on complex machinery, enabling them to implement effective predictive maintenance strategies that enhance overall productivity and competitiveness.

Future investigations could explore advanced data preprocessing techniques capable of categorizing the causes of failure, distinguishing between non-failure (designated as value 0), failures attributed to temperature (value 1), and those attributed to torque (value 2). Moreover, researchers could concern on the temporal aspect of data collection (e.g., night shifts) influence the predictive maintenance strategies for critical manufacturing processes. Furthermore, researchers could study the choice of performance metrics (accuracy, precision, recall, F1-score, ROC AUC)impact the evaluation of predictive maintenance models in imbalanced datasets. Additionally, researchers could explore identifying other parameters that may signal failure instances in cases where failures are not detected. These parameters could originate from external factors or aspects that have not yet been recorded. Furthermore, the application of sophisticated techniques for enhanced accuracy, such as image processing utilizing Convolutional Neural Networks (CNN), could be considered for modelling failure detection using camera systems.

7. References

[1] A. Saxena and K. Goebel, “PHM08 Challenge Data Set.” 2008. Distributed by NASA Ames Prognostics Data Repository. https://data.nasa.gov/download/nk8v-ckry/application%2Fzip.

[2] O. Masmoudi, M. Jaoua, A. Jaoua and S. Yacout, “Data Preparation in Machine Learning for Condition-based Maintenance,” *Journal of Computer Science*, vol. 17, no. 6, pp. 525–538, 2021, doi: 10.3844/jcssp.2021.525.538.

[3] A. Canito, M. Fernandes, J. Mourinho, S. Tosun, K. Kaya, A. Turupcu, A. Lagares, H. Karabulut and G. Marreiros, “Flexible Architecture for Data-Driven Predictive Maintenance with Support for Offline and Online Machine Learning Techniques,” in *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, Toronto, Canada, Oct. 13–16, 2021, pp. 1–7, doi: 10.1109/IECON48115.2021.9589230.

[4] Z. Znaidi, M. E. H. Ech-Chhibat, A. Khiat and L. A. E. Maalem, “Predictive maintenance project implementation based on data-driven & data mining,” in *2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Mohammedia, Morocco, May. 18–19, 2023, pp. 1-5, doi: 10.1109/IRASET57153.2023.10152915.

[5] I. Mahmud, I. Ismail and Z. Baharudin, “Predictive Maintenance for a Turbofan Engine Using Data Mining,” in *International Conference on Artificial Intelligence for Smart Community,* Perak, Malaysia, Dec. 17–18, 2022, pp. 677–687, doi: 10.1007/978-981-16-2183-3_65.

[6] S. Putchala, R. Kotha, V. Guda and Y. Ramadevi, “Transformer Data Analysis for Predictive Maintenance,” in *Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems*, Telangana, India, Aug. 13–14, 2021, pp. 217–230, doi: 10.1007/978-981-16-7389-4_21.

[7] S. Lu, Z. Liu and Y. Shen, “Automatic Fault Detection of Multiple Targets in Railway Maintenance Based on Time-Scale Normalization,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 4, pp. 849–865, 2018, doi: 10.1109/TIM.2018.2790498.

[8] F. Gatta, F. Giampaolo, D. Chiaro and F. Piccialli, “Predictive maintenance for offshore oil wells by means of deep learning features extraction,”

Expert Systems, vol. 41, no. 2, pp. 1–13, 2024, doi: 10.1111/exsy.13128.

[9] S. Panagou, F. Fruggiero, M. Lerra, C. D. Vecchio, F. Menchetti, L. Piedimonte, O. R. Natale and S. Passariello, "Feature investigation with Digital Twin for predictive maintenance following a machine learning approach," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 132–137, 2022, doi: 10.1016/j.ifacol.2022.04.182.

[10] F. Calabrese, A. Regattieri, M. Bortolini, M. Gamberi and F. Pilati, "Predictive maintenance: a novel framework for a data-driven, semi-supervised, and partially online prognostic health management application in industries," *applied sciences*, vol. 11, no. 8, 2021, Art. no. 3380, doi: 10.3390/app11083380.

[11] A. Shylendra, P. Shukla, S. Bhunia and A. R. Trivedi, "Analog-Domain Time-Series Moment Extraction for Low Power Predictive Maintenance Analytics," in *2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, Incheon, South Korea, Jun. 13–15, 2022, pp. 9–12, doi: 10.1109/AICAS54282.2022.9869914.

[12] H. Xu and J. A. Prozzi, "Effect of Model Accuracy on Maintenance and Rehabilitation Benefits," *Transportation Research Record*, vol. 2677, no. 6, pp. 773–782, 2023, doi: 10.1177/03611981221150396.

[13] B. P. Mota, P. M. Faria and C. Ramos, "Predictive Maintenance for Maintenance-Effective Manufacturing Using Machine Learning Approaches," in *17th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2022)*, Salamanca, Spain, Sep. 5–7, 2022, pp. 13–22, doi: 10.1007/978-3-031-18050-7_2.

[14] A. F. Azyus and S. K. Wijaya, "Determining the Method of Predictive Maintenance for Aircraft Engine Using Machine Learning," *Journal of computer science and technology studies*, vol. 4, no. 1, pp. 1–6, 2022, doi: 10.32996/jcsts.

[15] P. Ngwa and I. Ngaruye, "Big Data Analytics for Predictive System Maintenance Using Machine Learning Models," *Advances in Data Science and Adaptive Analysis*, vol. 15, no. 01n02, 2022, Art. no. 2350001, doi: 10.1142/S2424922X23500018.

[16] A. J. Alfaro-Nango, E. N. Escobar-Gómez, E. Chandomí-Castellanos, S. Velázquez-Trujillo, H. R. Hernandez-De-León and L. M. Blanco-González, "Predictive Maintenance Algorithm Based on Machine Learning for Industrial Asset," in *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*, Istanbul, Turkey, May. 17–20, 2022, pp. 1489–1494, doi: 10.1109/CoDIT55151.2022.9803983.

[17] S. Gautam, R. Noureddine and W. D. Solvang, "Machine Learning and IIoT Application for Predictive Maintenance," in *Advanced Manufacturing and Automation XII*, Xiamen, China, Nov. 1–2, 2022, pp. 257–265, doi: 10.1007/978-981-19-9338-1_32.

[18] K. Patel and A. Shanbhag, "Exploring ML for Predictive Maintenance Using Imbalance Correction techniques and SHAP," *in 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Prague, Czech Republic, Jul. 20–22, 2022, pp. 1–10, doi: 10.1109/ICECET55527.2022.9873073.

[19] C. Chen, J. Shi, M. Shen, L. Feng and G. Tao, "A Predictive Maintenance Strategy Using Deep Learning Quantile Regression and Kernel Density Estimation for Failure Prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023, doi: 10.1109/TIM.2023.3240208.

[20] E. M. Olariu, R. Portase, R. Tolas and R. Potolea, "Predictive Maintenance - Exploring strategies for Remaining Useful Life (RUL) prediction," in *2022 IEEE 18th International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, Romania, Sep. 22–24, 2022, pp. 3–8, doi: 10.1109/ICCP56966.2022.10053988.

[21] K. A. Manjunatha and V. Agarwal, "Multi-Kernel based Adaptive Support Vector Machine for Scalable Predictive Maintenance," *Annual Conference of the PHM Society*, vol. 14, no. 1, pp. 1-11, 2022, doi: 10.36001/phmconf.2022.v14i1.3198.

[22] S. R. Sabat, "Conditional Predictive Maintenance of Electric Vehicles from Electrical and Mechanical Faults," *International Journal For Multidisciplinary Research*, vol. 5, no. 1, pp. 1–11, 2023, doi: 10.36948/ijfmr.2023.v05i01.1325.

[23] M. H. Abidi, M. K. Mohammed and H. Alkhalefah, "Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing," *Sustainability*, vol. 14, no. 6, 2022, Art. no. 3387, doi: 10.3390/su14063387.

[24] A. Bhide, D. Ghodake, A. Jamle, S Shaikh and S. R. Bhujbal, "Predictive Machine Maintenance Using Tiny ML," *International Journal For Research in Applied Science And Engineering Technology*, vol. 11, no. 4, pp. 4252–4255, 2023, doi: 10.22214/ijraset.2023.51254