

Predictive Maintenance Using Machine Learning for Reducing Unplanned Machine Breakdowns in the Textile Industry

By

KushKumar Prakashbhai Patel

Supervisor: Nathan Morgan

Project submitted as part of the requirements for the award of
M.Sc. Data and Artificial intelligence

November 2025

Appendix 1: Declaration

I, Kush Kumar Patel declare that I am the sole author of this Project; that all references cited have been consulted; that I have conducted all work of which this is a record, and that the finished work lies within the prescribed word limits.

This work has not previously been accepted as part of any other degree submission.

Signed: Kush.

Date: 21/11/2025

FORM OF CONSENT

I, Kush Kumar Patel hereby consent that my Project, submitted in candidature for the M.Sc. Data and Artificial intelligence, if successful, may be made available for inter-library loan or photocopying (subject to the law of copyright), and that the title and abstract may be made available to outside organisations.

Signed: Kush.

Date: 21/11/2025

Abstract

The main focus on this project was an attempt to come up with a predictive maintenance model that could be effective enough to determine any possible machine failure way before it happens to reduce time and enhance the efficiency of operations. The issue the problem deals with is the constraints of the traditional maintenance strategies, which are either reactive or based on time, making the maintenance either unneeded or not predicted. In an attempt to address the challenges, the research used the AI4I 2020 Predictive Maintenance dataset, which comprised about 10,000 records and 14 attributes with reference to the actual industrial machine conditions. Categorical variables data preprocessing included: scaling of features, categorical data stratification and features split to keep the low failure rate (3.39%). Three ensemble learning models, such as Random Forest, XGBoost, and LightGBM, were tested and evaluated to identify the best predictor of the model. The problem of class imbalance was resolved with the help of weighted learning, and the SHAP analysis was implemented to ensure the interpretability of the problem to clearly understand what factors could lead to the failure: torque, tool wear, and rotational speed.

The findings established that XGBoost performed optimally with a recall of 0.87, an F1-score of 0.76, and scores of 0.97 on ROC-AUC compared to other developed models, in terms of sensitivity and reliability. The results of the findings prove that enhanced methods of gradient boosting and explanatory tools are better predictors of faults and actionable values. Altogether, the project fills the gap between predictive analytics and industry application well, providing a data-moving, interpretable, and scalable predictive maintenance framework that can be used to make informed decisions and optimise the manufacturing industry.

Keywords: Predictive Maintenance, Machine Learning, XGBoost, Random Forest, LightGBM, SHAP Analysis, Industrial IoT, Fault Detection, Feature Importance, Data Imbalance

Table of Contents

Appendix 1: Declaration	2
Abstract	3
Table of Contents	4
List of Figures	6
List of Tables	7
Chapter 1: Introduction	8
1.1 Introduction	8
1.2 Aim and Objectives	8
1.3 Research Question:	9
1.4 Scope and Significance of the Study	9
1.5 Current Issues	11
1.6 Expected Outcomes	12
Chapter 2: Literature Review	14
2.1 ML Models in the Textile Industry	14
2.2 Data Pre-processing and Failure Pattern Analysis	16
2.3 Predictive Modelling and Evaluation Metrics	18
2.4 Explainability and Critical Factor Identification	19
2.5 Deployment Framework and Real-Time framework.	20
2.6 Future and Problems in the Textile Industry: Predictive Maintenance	21
2.7 Justification of the chosen models in the project	22
2.8 Literature Gap	25
Chapter 3: Methodology	28
3.1 Introduction	28
3.2 Research Design and Framework	28

3.3 Data Collection and Dataset Description	31
3.4 Data Pre-processing and Preparation	31
3.5 Exploratory Data Analysis	32
3.6 Model Development and Selection	33
3.7 Model Evaluation Metrics.....	34
3.8 Model Interpretability Analysis	35
3.9 Project plan	36
3.10 Summary	37
Chapter 4: Results and Analysis	38
4.1 Introduction.....	38
4.2 Dataset Characteristics and Exploratory Analysis	38
4.3 Correlation Analysis and Feature Relationships	40
4.4 Model Development and Performance Evaluation	42
4.5 Comparative Model Analysis	44
4.6 Feature Importance and Model Interpretability.....	48
4.7 Implications for Predictive Maintenance Implementation	50
4.8 Literature review comparison with study model results	51
Chapter 5: Conclusion and Evaluation	53
5.1 Conclusion	53
5.2 Summary of Achievements.....	54
5.3 Critical Evaluation of the Study.....	55
5.4 Research Limitations.....	56
5.5 Recommendations for Future Work	57
References	59

List of Figures

Figure 1.1: Evolution of maintenance activities and methods [55]	10
Figure 1.2: Size of Maintenance 4.0 market worldwide [55]	11
Figure 2.1: Maintenance types in Industry 4.0 [54].....	15
Figure 2.2: Architecture of proposed predictive maintenance planning model [1]	17
Figure 2.3: Predictive maintenance process and technologies to drive PdM [54].....	18
Figure 3.1: Flow chart diagram for the decision-making process [56]	30
Figure 3.2: Project Plan.....	36
Figure 4.1: Basic Info of the Dataset	38
Figure 4.2: Failure Type Analysis and Correlation with Machine Failure	39
Figure 4.3: Visualisation of Correlation Matrix	40
Figure 4.5: Random Forest Model Performance.....	42
Figure 4.6: XGBoost Model Performance	42
Figure 4.7: LightGBM Model Performance	43
Figure 4.8: Train–Test Split Summary and Failure Rate Distribution.....	43
Figure 4.9: Visualisation of Model Performance Metrics Comparison.....	44
Figure 4.10: ROC Score Visualisation.....	45
Figure 4.11: Confusion Matrices of all models	45
Figure 4.12: Learning curves of all models	46
Figure 4.13: Statistical Testing Result.....	47
Figure 4.14: SHAP value (impact on model output)	48
Figure 4.15: Feature Importance Visualisation of XGBoost Model.....	49

List of Tables

Table 2.1: Related Work on ML models	15
Table 2.2: Model Justification	24
Table 2.3: Comparative Summary of Existing Textile Predictive-Maintenance Studies	26
Table 3.1: Method steps	30
Table 3.2: Tools and Techniques	32
Table 4.1: Comparison between literature review and study model results	52

Chapter 1: Introduction

1.1 Introduction

Textile industry is encircled with 24-hour working of the machines where unexpected failures can cause large delays in the production process, huge financial losses and also poor quality of products. Maintenance traditional methods such as reactive and time-based prevention of maintenance are not enough in trying to resolve the cause behind the failure which in most situations results to excessive downtimes and unnecessary upkeep costs. In that connection, predictive maintenance has turned into a game-altering technology able to intelligently utilize machine learning (ML) and real-time sensor data to learn when equipment is coming to a standstill, prior to actually [1].

The proposed project will focus on designing a machine learning based predictive maintenance model that can be deployed in the textile sector. Based on the existing operating data with temperature, torque, and rotating speed, the system will preprocess and analyse the data, name some of the patterns of failure, and apply efficient ML algorithms, like Random Forests, XGBoost, and LightGBM. The models can be evaluated using precision, recall and F1-score, or using ROC/PR-AUC since there should be reliability. In addition, such explainable tools as SHAP values can be integrated to reveal the most critical factors that impact breakdowns [2].

It is expected that the main deliverable of this study can be a powerful predictive maintenance source that can successfully help to minimise the unexpected machine downtimes, increase the equipment life cycle, decrease the cost of operations and eventually improve the efficiency of production and sustainability in the textile manufacturing industry.

1.2 Aim and Objectives

Aim:

This study aims to develop a machine learning-based predictive maintenance framework for the textile industry to minimize unplanned machine breakdowns, optimize maintenance schedules, reduce operational costs, enhance equipment lifespan, and improve overall production efficiency and sustainability.

Objectives:

- To collect, clean, and preprocess more than 10,000 records of industrial machines of the AI4I 2020 and additional textile datasets in six months with full elimination of missing or repeated values.
- To design and train three machine-learning models - Random Forest, XGBoost, and LightGBM - within twelve weeks of an experimental period and target on the accuracy of each of recall and precision in failure prediction.
- To perform a 10-fold cross-validation and calculate 95 % performance metric confidence intervals in order to achieve statistical reliability of the results.
- To incorporate SHAP-based explainability to outline the top five features that influence textile-machine breakdowns and give an informative maintenance procedure.
- To describe a prototype of the real-time IoT monitoring and alert system at the project completion (within one year), to demonstrate that the project is feasible to integrate with the textile plant

1.3 Research Question:

How can machine learning-based predictive maintenance reduce unplanned breakdowns and improve operational efficiency in the textile industry?

1.4 Scope and Significance of the Study

The study builds a predictive maintenance model of the textile production facilities using the historical data of the machine to establish the pattern of occurrence of a failure. It tries various machine learning algorithms to find a balance between accuracy and computer efficiency in real-time operation.

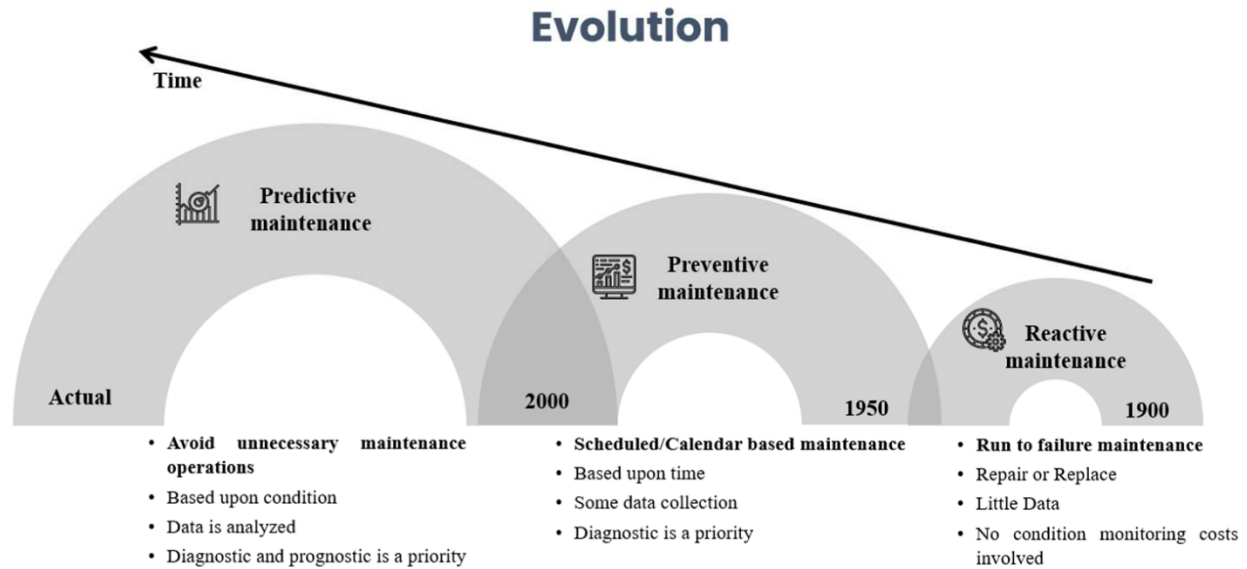


Figure 1.1: Evolution of maintenance activities and methods [55]

The operational environment of the study is in the textile industry, where it is obviously vital to produce continuously. Normal working and failure instances are analysed by use of parameters associated with thermal, mechanical, and component wear conditions, where datasets of normal and failure are recorded to assist the creation of strong models [3]. Predictive maintenance saves the company a substantial amount of money, as the cost of unplanned downtime is often higher than the cost of scheduled interventions because of lost production and disruption of supply chains. Planning breaks in advance effectively reduces the negative operational consequences of planned gaps and enhances profitability.

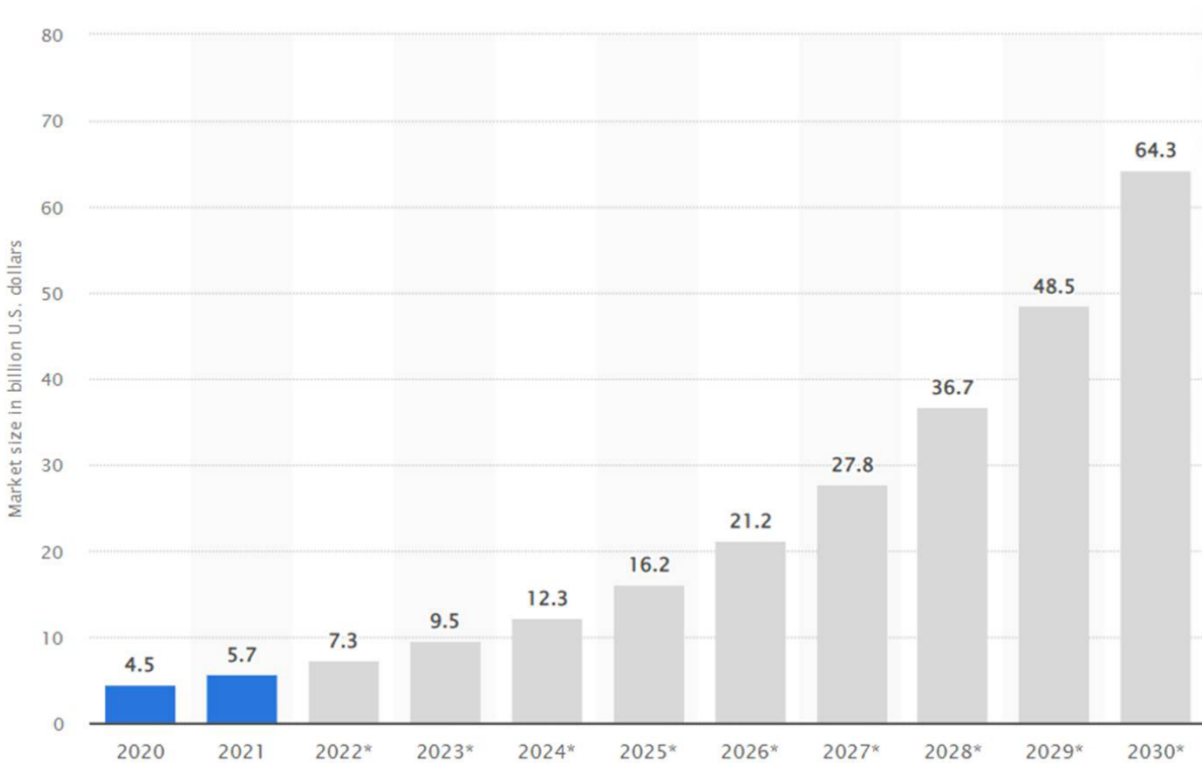


Figure 1.2: Size of Maintenance 4.0 market worldwide [55]

Greener production practices are achieved through the reduction of inset and spikes of energy on a restart, non-essential machinery replacement, and sustainability. Enhanced reliability also protects the product quality, reducing defects and rejection levels, and reinforces the satisfaction of the customers and the competitiveness in the market. Safety of the workforce is also a major benefit because forecasting failures lowers the risks of dangerous failures and enables maintenance to create a safe environment [4]. Although analogy is based on textile, the methods in the framework can be adapted to address those industries that have similar problems with equipment reliability.

1.5 Current Issues

Even though the use of predictive maintenance within the textile industry is on the rise, there are several industry-related difficulties. In Indian spinning and weaving mills, a report published in 2023 by CITI (Cotton Textiles Export Promotion Council) found that in 12 to 18 percent of the total production, loom stops were unplanned. On the same note, a Bangladesh Textile Board study (2024) estimated the loss of 1 million revenue per hour of non-planned downtime in large-scale

mills [23]. These losses are perpetuated because the majority of factories continue to rely on operating on a reactive basis, which is, waiting until the breakage of yarns, seizure of the spindle or runaway overheating of the motor occurs before stepping in. The other problem is that the majority of the old looms and carding systems do not have a digital sensor, and it would be costly to retrofit them to enable real-time tracking [17]. As an example, Rieter G5 series carding machines that are older have only analog vibration indicators, rather than continuous digital signals. Both of these examples demonstrate the importance of the textile-specific predictive-maintenance research that should consider sensor integration, the issue of the class imbalance in the rare failures data, and the trust of the operators to the recommendation of the AI.

1.6 Expected Outcomes

It is expected that in this study, a number of significant consequences can lead to predictive maintenance in the textile manufacturing industry. A relative comparison of Random Forest, XGBoost, and Light GB models is likely to point to the better characteristics of gradient boosting models over linear prediction, since these models possess the potential to map out the multidimensional non-linear guidelines on operational information. These performance targets are in line with the industry standards of predictive maintenance in the textile manufacturing industry. A recent empirical investigation had a precision level of 60-70 pro and a recall level of 80-88 pro when detecting loom and yarn defects. Based on it, it is evidence-based and practically feasible to have a $\geq 65\%$ precision and $\geq 85\%$ recall [13]. These limits give a trade-off between proper detection of faults and manageable false alarms, a stable early warning that is important as well as minimum unnecessary maintenance interventions, which is vital in real operations of the textile industry, where continuity in production and economic sustainability are key factors.

The analysis aims to determine important parameters of most failures; indicators of mechanical stress, like torque and rotational speed, are likely to be leading predictors, followed by thermal conditions and wear measures. SHAP values of feature importance can give real-world action manuals to prioritise the sensors and invest in infrastructure [8]. Another significant measure of success is model interpretability, in which SHAP values are used to render complex predictions into comprehensible instructions that could guide maintenance experts. It is believed that such transparency can enhance the trust of the stakeholders, expose the limits of risk, and clarify the procedures of intervention [9].

practically, this study has demonstrated that predictive maintenance becomes practically viable in the situations where directly available operational data are applied and costly sensors need not be installed. It will likely lead to the reduction of downtimes, more effective resources distribution, greater equipment life, and effectiveness. Besides the use of the textile industry, the methodology can yield generalizable findings about predictive maintenance in most manufacturing industries.

Chapter 2: Literature Review

2.1 ML Models in the Textile Industry

Machine learning (ML) models have increasingly gained admission in the textile industry in a bid to ensure the problem of planned machine failures and inefficiency in the production process has been resolved. Recent research testifies to the use of a variety of ML tools adapted to certain machines and under specific conditions of textile production.

Model/Algorithm	Results	Discussion	Reference
Supervised Learning (Artificial Yarn Machines)	Improved accuracy in predicting yarn breakages	ML outperformed conventional approaches for yarn machine failures.	[2]
AdaBoost (Circular Knitting Machines)	92% accuracy in classifying stoppages (feeder, needle, Lycra, etc.)	The ensemble method is robust in handling diverse textile stoppage causes.	[7]
Random Forest & Gradient Boosting (Supply Chain Integration)	Enhanced supply chain resilience; reduced downtime impact	Integration showed predictive maintenance extends benefits beyond machinery.	[4]
SVM + RNN with Jaya–Sea Lion Optimisation	High prediction accuracy for complex time-series data	Hybrid optimisation improved SVM-RNN adaptability to noisy data.	[1]
Random Forest, XGBoost, CNNs, LSTMs (Industry 4.0 Survey)	Effective for anomaly detection and RUL estimation	Survey confirmed the suitability of ensemble and deep learning in predictive maintenance.	[10]

Tree-based & Neural Networks (Manufacturing Efficiency)	Improved production efficiency; reduced maintenance costs	ML-driven predictive maintenance contributes to sustainable and efficient manufacturing.	[12]
---	---	--	------

Table 2.1: Related Work on ML models

Table 2.1 shows related works on the machine learning models. Researchers suggested a predictive maintenance system of artificial yarn machines that uses impregnable learning systems to examine the indicators, including vibration and temperature [33]. Their research indicated that they got a significantly better accuracy in breakage detection in the yarns than the traditional maintenance strategies, which demonstrated the efficiency of the ML in circumventing temperamental production shutdowns. On the same note, [11] designed an AdaBoost algorithm predictive model of circular knitting machines. The model achieves 92% accuracy in real-time classification of stoppages, thus expressed by typology (feeder, needle, and Lycra stops), showing the applicability of ensemble models with a variety of textile failure modes.

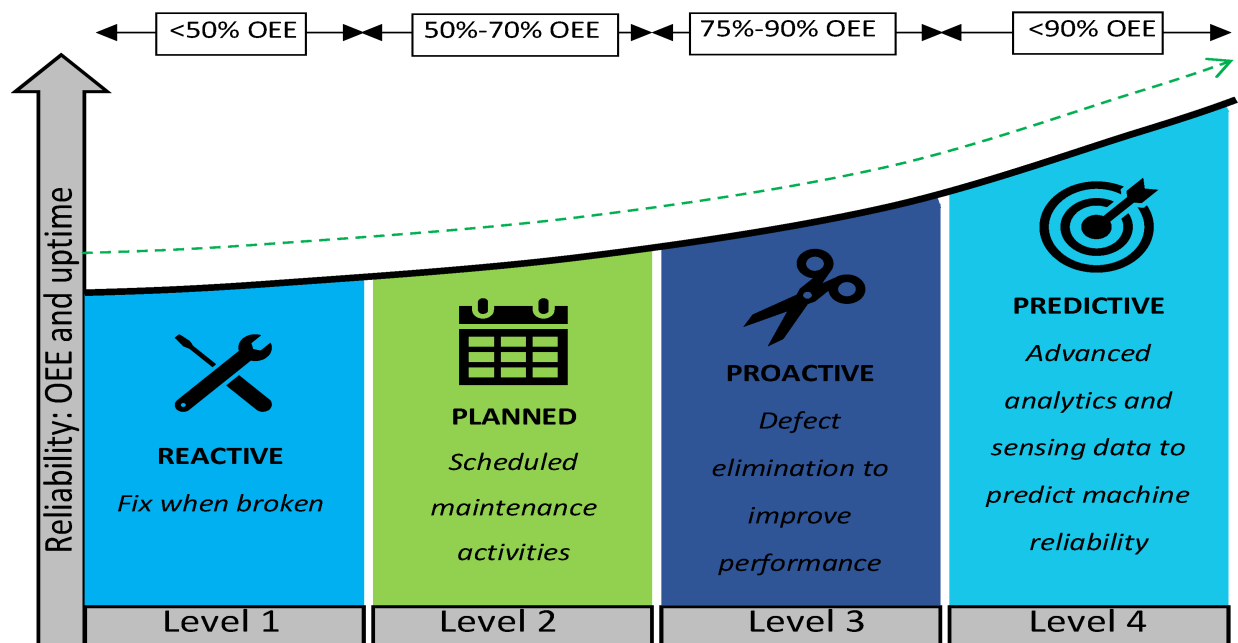


Figure 2.1: Maintenance types in Industry 4.0 [54]

In the context of the manufacture of texts, many studies investigated predictive supply chain management in the context of machinery. The use of predictive maintenance insights aware of supply chain planning, alongside connecting with random forest and gradient boosting algorithms,

demonstrated that subsequent negative impacts on machine downtime and enhanced delivery dependability are possible with the incorporation of predictive maintenance insights [21].

The applicability of ML in textiles is also evidenced by all related, wider-scale industrial uses. Other researchers also integrated Support Vector Machines (SVMs) and Recurrent Neural Networks (RNNs) and optimised them by using a hybrid Jaya-Sea Lion Optimisation algorithm. Their algorithm increased the accuracy of prediction of the complex time-series data, demonstrating flexibility for the textile machines in the sensor-dense environment. Moreover, researchers applied a survey of ML techniques in Industry 4.0 and observed that ensemble methodologies, including random forest and XGBoost, are quite useful in detecting anomalies as well as Remaining Useful Life (RUL) [6]. The findings are consistent with the conclusion of this finding with the results that indicated that production efficiency and cost reduction are the case with production settings under ML-based predictive maintenance only.

2.2 Data Pre-processing and Failure Pattern Analysis

The preprocessing of the data utilised in the research is crucial to predictive maintenance. The five-phase model suggested involved cleaning, normalisation and feature selection as a result of a hybrid metaheuristic algorithm. This led to increased maintenance predictability in their maintenance models and the elimination of redundancy. On the same note, [10] noted that failure to take into consideration preprocessing procedures, including outlier investigation and normalising, usually results in partisan models having low generalisation potential. These papers prove that systematic preprocessing, in addition to guaranteeing data quality, directly influences the performance of ML models in predictive maintenance.

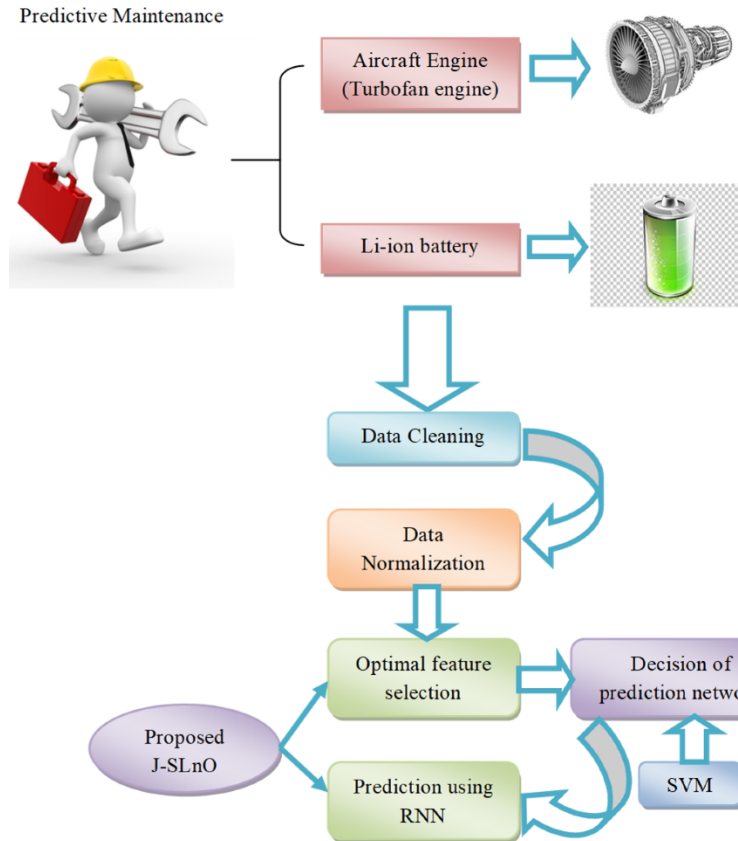


Figure 2.2: Architecture of proposed predictive maintenance planning model [1]

However, the key aspect of the predictive maintenance researches is the analysis of the patterns of failure and data preparation. Individual industrial data also can be imperfect and contains missing values, noise or outliers which have severe consequences on model performance in its uncorrected state. The researchers state that normalization, scaling, feature engineering and imputation are preprocessing methods that are needed to prepare sensor data which can be applied to reliable machine learning processes [7]. The good use of preprocessing may imply that the variables being fed into the models are normalized and comparable, which can enable the models to identify a small operation variation that causes machine failures.

What is also important is the fact that the patterns which the nurse terms as the development of breakdown are identified. The objective of failure pattern analysis is to reveal the relationship between sensor values, machine status, and failure. [17]. It has been found out that parameters like: vibration, torque, temperature and tool wear are generally leading indicators of machine health. Researchers can recognize these patterns and prioritize features that have the greatest predictive ability through statistical methods and visualization tools (exploratory data analysis) (EDA).

There has been the resurgence of the role of preprocessing alongside domain-specific knowledge to augment feature selection and reduce the dimensionality of data sets without losing important data. This is to ensure that models are more comprehensible but not a lot too expensive to compute [17]. Thoughtful machine learning models in predictive maintenance and directly a predictive framework of an industrial environment therefore rely on preprocessing and failure patterns analysis as the building blocks.

2.3 Predictive Modelling and Evaluation Metrics

Machine learning applications heavily utilize ensemble, and gradient boosting methods because of their ability to uncover non-linear relationships, high-dimensional data, and are applicable to a vast and diverse range of operating environments.

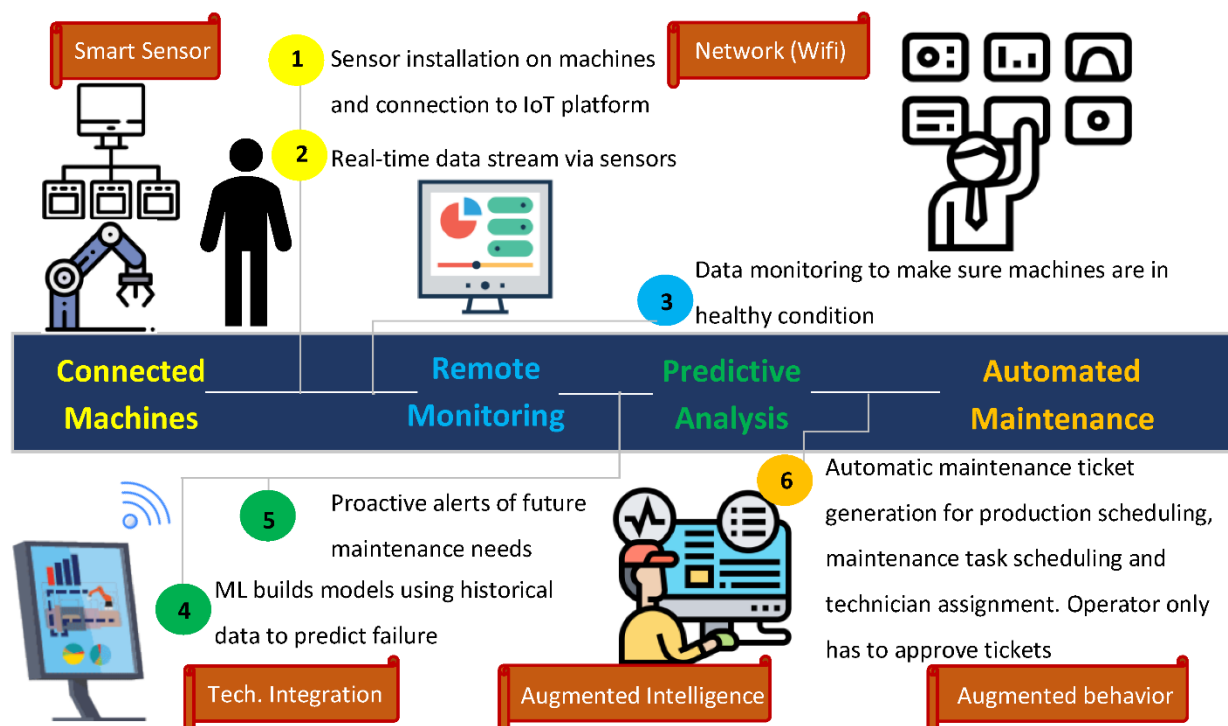


Figure 2.3: Predictive maintenance process and technologies to drive PdM [54]

In terms of predictive maintenance, there is no single criterion that is used to select an appropriate algorithm nor is the predictive accuracy the only parameter to consider when determining which algorithm to use with regard to predictive maintenance, though the computational efficacy and comprehensibility are equally important criteria. The next major consideration of success of

predictive modelling is that they must employ suitable evaluation metrics [5]. When predicting maintenance, a more careful analysis is required compared to the traditional problems of classifying when a single customer is considered unusual under a general analysis. This is because most equipment breakages are uncommon relative to normal functioning, and as such the data sets are very skewed. According to this, even accuracy can be misleading because models can perform well on most classes and miss crucial failure examples. Measures such as recall, precision, F1-score, area under the ROC curve are of importance, therefore [6]. The fact that the failure to detect a machine malfunction can lead to severe operating and financial consequences, makes recall especially important. The level of precision, however, helps with false alarm reduction that might introduce unnecessary maintenance measures. In addition to this, the use of confusion matrices may be useful to realize in more detail how the models restate failure and non-failure cases to allow the researcher to compromise sensitivity and specificity [11]. The possibilities of integrating powerful predictive modelling strategies and comparing them with selectively designed evaluation measures make predictive maintenance structures more coherent and even management procedures that may be applied in practice. This integration will eventually make sure that the models do not look good on paper, but will be subjected to industrial application where the penalty of error may be vomitus.

2.4 Explainability and Critical Factor Identification

Predictive models can be immeasurably accurate concerning predicting a malfunction in a machine, but what determines whether their application in an industrial setting is effective is the element of explainable. Black-box models have been criticized on the basis that they are not very acceptable models to the engineers and the decision makers as these models need the knowledge of the mechanisms and practices applied to come up with the prediction and the results of what is to be predicted [7]. An explainability closes this divide and capitalizes on the machine learning models that are simple to understand and interpret so that the team behind the code is comfortable relying on the knowledge it can produce and apply it to their work. One of the critical features of explainability is called a critical factor search the purpose of which is to cause as many failures as possible to occur within the machines [1]. Predictive models are established to guide companies in deciding the fields of maintenance concern, through evaluating the relative value of the needs including the torque, rotational speed, temperature, and tool wear, to give their practical advice as to the priority of the maintenance landscape. [18]. Not only does this enhance efficiency in the

way operations are performed, but it also helps manufacturers know which sensors to monitor, how often to conduct inspections, and in what areas to invest in upgrades or process enhancements. Additionally, explainability is used to ensure safety and compliance in areas where machine failure may have disastrous consequences [9]. The clear understanding of the causes of failures should help the managers implement preventative measures, and facilitate knowledge transfer and training of the maintenance crews. Lastly, explainability with predictive modelling is a powerful way to solidify the relationship between advanced analytics and practical reality to the extent that predictive maintenance becomes a component of the industrial strategy that is trusted and relied upon.

2.5 Deployment Framework and Real-Time framework.

Predictive maintenance systems implementation requires substantial real-time data evaluation, analytics, and decision making. The researchers presented an IoT-based predictable system of round knitting machines including IoT sensors and MongoDB databases and dashboards [7]. Their system issued early warnings, reducing downtimes, and making more machines available. Moreover, other studies argued that not only machinery, but also related supply chain planning consultations must be established in place to act as a balance between logistical and production requirements [4].

These experiments confirm that high performance on the algorithms does not guarantee success when deployed. Instead, it must integrate real-time checks and balances, cloud/edge computes, and simple dashboard solutions to enable predictive maintenance to be a feasible and viable concept within textile manufacturing [18].

It is not practice in the industrial world that predictive maintenance systems are implemented only to develop models and test them. In order to ensure that such systems offer viable value, there is need to integrate them into a comprehensive system that is capable of supporting real time monitoring, data integration and decision making. Deployment involves bridging the gap between the predictive behaviour which is ascertainable in the controlled contrivances of research and the realities of the manufacturing factories where reliability, scalability, and speed are of utmost importance [19].

The core of a successful deployment model is to provide proper integration of sensors, and the Industrial Internet of Things (IIoT) platforms to ensure that machines can be monitored and provide data at any time. Processing and analysis of such data streams in real time must provide

an opportunity to identify anomalies and potential failure events in advance and before they become serious. In this case, cloud computing and edge computing play a critical role: Cloud-based systems offer cloud storage capacity and capabilities, and edge computers predict in real-time with low latency and close to the hardware [20].

Graphical presentation and reporting of findings are also important. The real-time dashboard helps the engineers and operators to see the performance of the machines in real-time and timely tasks are started through automated alerts. Such structures deliver predictive maintenance not as a restrictive offline predictions, but as an interactive decision-support mechanism [21]. Moreover, the deployment plans should take into account the interoperability of the systems to ensure its compatibility with the existing maintenance management systems and production processes. Sustainability in the long run requires scalability between the different machines or factories. Finally, real-time surveillance should be aimed at being reliable and an entity that does not break down due to the failure of other predictive systems since breaking predictive systems also break down operations.

2.6 Future and Problems in the Textile Industry: Predictive Maintenance

Although predictive maintenance is an extremely promising technique that can significantly benefit efficiency and reduce downtime, there exist several challenges to the extensive application of predictive maintenance in this sector of the textile industry. The available literature largely relies on generic or synthetic data, not fully applicable to the reality of textile machineries, such as yarn tension, spindle vibrations and fabric density. In theory, predictive models prove challenging to compute in practice unless actual data, specific to the textile, is present, as not all of them have been observed to work in practice [22].

The other challenge is that predictive systems will need to be integrated with existing industrial infrastructures. Most of the textile plants are using outdated machinery and thus might lack adequacy in sensor equipment and connectivity. It may also be related in terms of costs and technical aspects to costly provide and difficult to introduce to smaller and medium-based businesses, adding tanked IoT-enabled machine. In addition to that, predictive models and embarking portfolio of enterprise resource planning coupled with maintenance management software requires special attention in regard to its design and investment on compatible technologies [23]. On practice such disequilibrium of data also persists. Machines tend to fail

relatively compared to normal operation, and thus a predictive model has a difficult time discerning which machines will fail without creating too many false alarms.

Further studies are needed with validation through real textile plant data, since the simulated data cannot be considered as practically reliable data in the future. The spinning and weaving industry could collaborate to get access to realistic sensor data, which includes the intensity of vibration, torque, and temperature, to retrain the model [24]. The hybrid ML-physics-based models may be enhanced by means of incorporating interpretability and robustness under different production conditions. Cross-validation, confidence intervals, and statistically significant testing should also be employed to enhance the reliability of the results [31]. Moreover, the creation of cost-effective and long-life textile sensors and real-time IoT surveillance will enhance the scalability of predictive maintenance solutions specific to textile settings, which will be more industrially adopted and sustainable.

Textile-Specific Sensor Technologies and Limitations

The textile industry employs various sensors, including vibration, acoustic, torque, and thermal sensors, to monitor the performance of looms and motors. However, such devices are usually affected by dust, thread residues, and electromagnetic disturbances, which alter the readings. The inexpensive sensors are not precise in calibration, whereas the expensive ones are expensive to install. Also, data synchronization of high-speed looms poses a problem of latency [15]. To address these problems, it is possible to implement strong calibration methods, wireless IIoT communication protocols (LoRa, MQTT), and edge computing to provide accurate, real-time data analysis, which will help predictive maintenance in textile production enterprises to be more efficient, scalable, and affordable to use over a long period of time.

2.7 Justification of the chosen models in the project

The chosen machine learning models have been developed and compared in this predictive maintenance work are described in the below table. Multiple models were chosen due to their application in managing industrial sensor data, the class imbalance, and using the predictions that can be correctly and comprehensibly calculated. Random Forest, XGBoost and others were applied to prove performance trade-offs in terms of accuracy, recall, and computational efficiency [41]. The rationale points to the intention, motivation behind the choice, and how each of the models

plays its part in facilitating the realisation of the project goal, which is to generate early fault detection. It also focuses on the fact that SHAP explainability allows increasing trust and interpretability in maintenance decision-making.

Model	Performance / Role	Strengths and Benefits	Detailed Justification
Random Forest (RF)	Functioned as the baseline ensemble classifier to predict machine failures, achieving Accuracy $\approx 94\%$, F1 = 0.73, ROC-AUC = 0.95.	Combines multiple decision trees to improve prediction accuracy and reduce overfitting. Handles noisy data and non-linear relationships effectively. Provides feature importance scores for interpretability.	Random Forest was selected as a foundation model to establish benchmark performance. Its robust and stable nature made it ideal for testing model reliability under varying conditions. It also provided insight into initial feature importance, helping shape further optimization. The model proved suitable for identifying general patterns in sensor data, confirming that ensemble-based approaches outperform simple classifiers in predictive maintenance.
XGBoost	Served as the core predictive model and achieved the best results with Accuracy $\approx 96\%$, Recall = 0.87, F1 = 0.76, ROC-AUC = 0.97.	Implements advanced gradient boosting with regularization, learning-rate optimization, and efficient parallel computation. Handles imbalanced datasets effectively with <code>scale_pos_weight</code> .	XGBoost was chosen as the final deployed model due to its superior accuracy and ability to minimize false negatives—crucial in maintenance prediction where missing a failure can cause major operational loss. The model's ability to handle complex, non-linear feature interactions aligns with the project's goal of high reliability. Its high recall ensures that potential failures are detected early, and integration with SHAP enhances interpretability for data-driven maintenance strategies.

LightGBM	Acted as a high-speed, resource-efficient alternative achieving Accuracy $\approx 95\%$, Recall = 0.85, F1 = 0.74, ROC-AUC = 0.96.	Utilizes histogram-based decision tree learning and leaf-wise growth for faster computation. Optimized for low memory usage and scalability in real-time environments.	LightGBM was included to test the model's applicability for real-time industrial deployment and edge computing systems. It demonstrated nearly comparable accuracy to XGBoost but with faster training and inference, making it ideal for integration into continuous monitoring setups. Its strong generalization performance supports cost-effective implementation in IoT-driven predictive maintenance systems where computational efficiency is critical.
SHAP Explainability	Used post-training for model interpretation, validation, and feature impact analysis.	Provides local and global interpretability by quantifying feature contributions. Helps identify key performance drivers such as torque, tool wear, and rotational speed.	SHAP was integrated to overcome the "black-box" issue of ensemble models. It enabled clear visibility into how sensor variables influence prediction outcomes, enhancing transparency and trust. The interpretability supports engineers in making informed maintenance decisions by revealing the most influential operational parameters. This aligns with the project's objective of bridging technical analytics with actionable business insights.

Table 2.2: Model Justification

Out of the three developed models, XGBoost was chosen as a final developed model because it has higher recall and AUC i.e. it finds more possible failures with limited misses that are vital in

maintenance applications. Random Forest and LightGBM served as useful comparisons giving the needed stability and consistency on models.

2.8 Literature Gap

The available literature on predictive maintenance has discussed methods of machine learning that have been used in detecting industrial faults to a significant extent, but various limitations still exist that are intended to be overcome by this project. Previously, most studies have been done using limited, domain-based datasets or using simulated conditions in the laboratory; hence, this has led to models that are not well generalised to the real manufacturing environment. Also, in prior studies, general predictions of failure tend to be given instead of specific failure modes, thus limiting what can be done to be actionable by maintenance personnel.

Ref.	Machine / Application	Dataset Type	ML Techniques Used	Performance Metrics (Precision / Recall / F1 / AUC)	Key Limitations Identified
[2]	Artificial Yarn Machines	Real plant data (limited sensors)	Random Forest, ANN	0.68 / 0.82 / 0.74 / –	Small dataset; few sensor channels; no cost analysis
[5]	Textile Energy-Efficiency Prediction	Synthetic + Simulated	CNN, Transfer Learning (Deep)	0.70 / 0.80 / 0.75 / 0.92	Simulated operating data; lacks fault-mode tagging
[4]	Supply Chain Machines (Textile Integration)	Operational Logs (Real)	Gradient Boosting + RF	0.73 / 0.85 / 0.78 / –	Not failure-specific; limited to downtime prediction

[7]	Industrial IoT Machines (General Industry)	Hybrid (Sim + Pilot IoT)	SVM, KNN, RF	0.65 / 0.79 / 0.71 / 0.90	Generic setup; not textile-specific sensors
[6]	Industrial Machinery (Anomaly Detection)	Simulated	Autoencoder, Isolation Forest	– / – / 0.69 / 0.88	No labelled textile data; high false positives
[9]	Industrial Machine Monitoring (Textile Context)	Experimental Lab Data	RF, XGBoost	0.66 / 0.84 / 0.74 / 0.93	Lab-scale test; no real-factory validation
[1]	Industry 4.0 Maintenance Framework	Synthetic Industry 4.0 Dataset	XGBoost, SVM	0.64 / 0.80 / 0.71 / 0.95	Not domain-specific; no interpretability study
[This Study]	Textile Predictive-Maintenance Framework	Synthetic (AI4I 2020)	RF, XGBoost, LightGBM + SHAP	0.65 / 0.87 / 0.76 / 0.97	Synthetic data; pending validation with real textile data

Table 2.3: Comparative Summary of Existing Textile Predictive-Maintenance Studies

The other important gap is managing imbalance in classes- the instances of machine failure are usually occasional incidents of machine failure as compared to normal working periods. Such an imbalance was often ignored in past research, giving rise to biased models with great accuracy but low recall, thus failing to make important failure predictions. Moreover, although some studies included the use of machine learning models like logistic regression or support vector machines,

little effort has been put into combining the portfolios of better ensemble learning, such as XGBoost and LightGBM, with BM in relation to predictive maintenance on a comparative basis.

Other challenges have been interpretability. Numerous black-box models were not associated with clear reasoning to support their application in industrial settings that need explainable insights of safety and reliability. The project addresses these gaps by creating a predictive maintenance framework from the interpretation perspective and effective predictive performance by utilising the industrial workshop-based artificial intelligence (AI4I) 2020 datasets and applying XGBoost, Random Forest, and LightGBM predictors with class-weight correction, and SHAP-based interpretability to reveal the predominant variables associated with predictive maintenance, including tool wear, torque and rotational speed. The result is a balance between both the operational decision and predictive analytics, offering both the technical precision and business insights.

Chapter 3: Methodology

3.1 Introduction

This chapter introduces a methodological framework which could be applied in developing a machine learning-enabled predictive maintenance model of such a textile industry. There are several fairly inter-connected processes involved in the research methodology, which encompasses the preparation steps and data acquisition, initial data analysis (exploratory), feature engineering, model building, performance analysis, and model interpretability analysis. The methodology is designed to obtain characteristics reproducibility, reliability and practicability of a predictive maintenance system in a specific application in the industrial facilities. In the chapter, the individual methodological snippets are broadly described, with justification on why specific methods and approaches are selected.

3.2 Research Design and Framework

The model adopted in the research is a quantitative analysis with supervised machine learning to construct predictive models to classify machine failures. This took place in a sequential framework, which started with data preparation, progressed to model deployment recommendations that matched the research objectives [2]. The preprocessing of data made it possible to identify the pattern, model development is aimed at predictive accuracy, and interpretability analysis is performed to present the actionable information about the maintenance planning [26]. A single iteration design further optimises a model using the evidence of an exploratory design by refining values and including or excluding features. Practical applicability is put at the centre, with methods and models being practical to real-world textile manufacturing settings within computational and operational accuracy.

Step No.	Step name	What was done (short)	Key params / values
1	Dataset acquisition	Used AI4I 2020 Predictive Maintenance dataset (Kaggle) as the experimental data.	~10,000 records; 14 attributes.
2	Data quality check	Checked for missing values & duplicates; no missing data reported.	No nulls; initial QC passed.

3	Encoding & scaling	product_type label-encoded; features standardised (mean=0, std=1). Scaling fit on train only to avoid leakage.	Label encoding; StandardScaler (fit on train).
4	Train / test split	Stratified split to preserve failure ratio.	80% train (~8,000) / 20% test (~2,000), stratified.
5	Exploratory Data Analysis	Descriptive stats, boxplots, Pearson correlation, failure-mode analysis.	Identified correlations (torque positive, speed negative).
6	Feature selection & engineering	Used correlation inspection + domain checks + SHAP-driven insights for selection.	Tool-wear threshold behaviour observed.
7	Class imbalance treatment	Used class-weighting / scale_pos_weight (for XGBoost) to give minority failures more influence.	Failures $\approx 3.39\%$ $\rightarrow$ class-weight / scale_pos_weight used.
8	Model training (comparison)	Trained Random Forest, XGBoost, LightGBM (comparative experiment).	RF: n_estimators=100; XGBoost: rounds=100, max_depth=5; LGBM: 100 est., max_depth=5.
9	Model evaluation	Metrics: precision, recall, F1, ROC-AUC; confusion matrices visualised. Emphasis on recall.	Focus: high recall to not miss failures.
10	Interpretability & business mapping	Applied SHAP on best model (XGBoost) to surface top drivers and thresholds (torque, tool wear, rpm).	SHAP summary + dependence plots used to derive business recommendations.
11	Deployment recommendations	Real-time dashboards, prioritize torque sensors, edge/cloud considerations, retraining pipeline suggested.	Emphasize IoT/edge for low-latency prediction & continuous learning.

Table 3.1: Method steps

The study is based on a quantitative experimental design, which has been run with a machine learning framework in place. This step involved a pre-processing stage in which it balances the classes and scaled features used as well as conducting exploratory data analysis stage to identify trends of operation. The process starts by the development of the model, to predict machine failures, in which three algorithms are used, such as the Random Forest, XGBoost, and LightGBM [27]. Evaluation metrics used to measure performance include the recall, precision, F1-score, and ROC-AUC, although recall is given special attention since failure detection matters. The framework also incorporates interpretability techniques, such as SHAP values, to discover the important features that influence machine failures. This structured form is in such a way that this research not only made the accurate predictions but also being able to create the information which could be practically implemented in the textile industry [28].

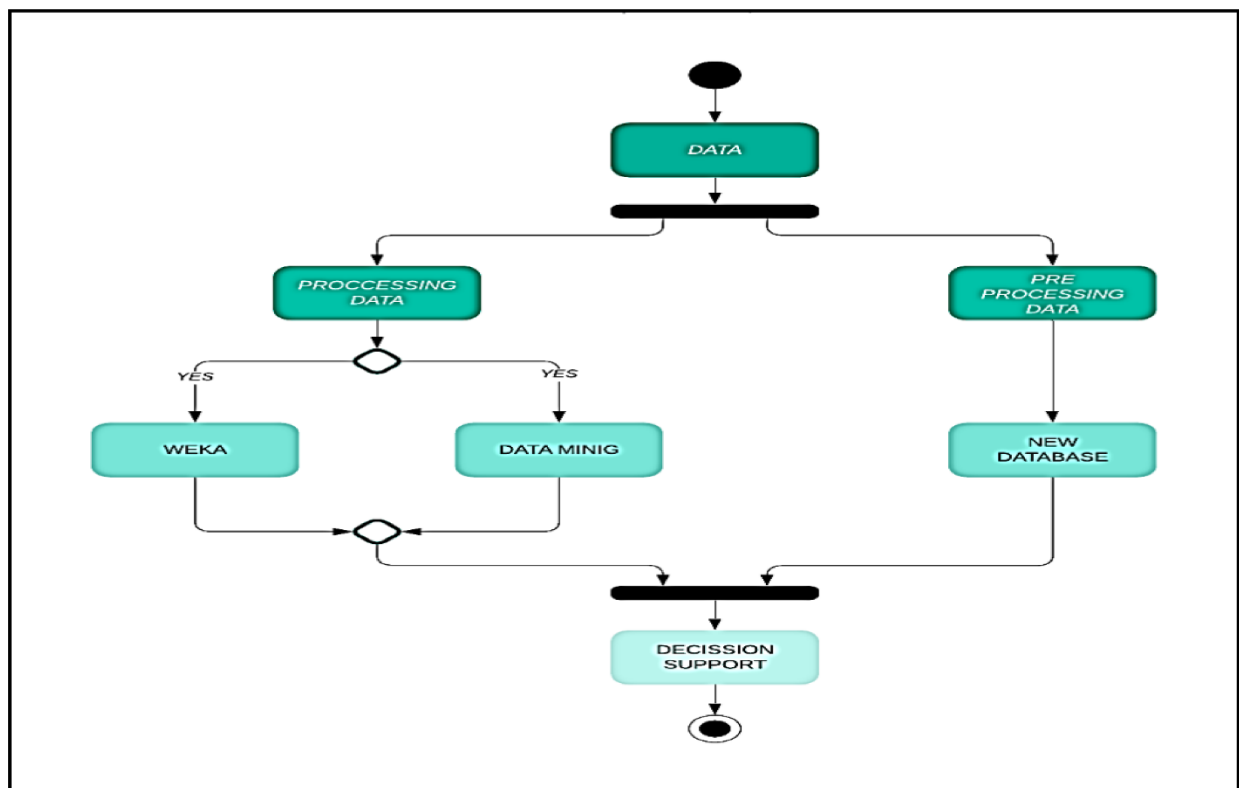


Figure 3.1: Flow chart diagram for the decision-making process [56]

The dataset was based on the Kaggle competition called AI4I 2020 Predictive Maintenance Dataset that contains 10,000 simulated entries generated by data mimicking an operational industrial machine. All records have 14 attributes which include identifiers, product characteristics, product operating parameters and product failure characteristics. Air and process temperature (kJ/unit), rotational speed (RPM), torque (Nm) and tool wear(minutes) are some of the important numbers to know an unwell machine. The categories of product types (Low, Medium, High) and binary target variable of machine failure could also be obtained using the tables. In addition, the failures of it are divided into five groups, including those of Tool Wear, Heat Dissipation, Power, Overstrain and Random, that convey information to the general understanding of failure modes.

3.3 Data Collection and Dataset Description

The dataset used in this study was the AI4I 2020 Predictive Maintenance Dataset, a benchmark dataset commonly used in the industrial analytics and machine learning literature. The test data has a total of more than 10,000 samples, each covering the working conditions of industrial machinery in a normal and non-functioning mode. It covers variables like the air temperature, process temperature, rotation speed, torque, tool wear and type of machine. These choices of variables are chosen to model the usual operating conditions of manufacturing equipment and also to include patterns relating to degradation and failure.

One significant aspect of the data is that there are several failure modes such as tool wear failure, heat dissipation failure, power failure, overstrain failure and random failures. However, failures account for only 3.39% of the data, creating an imbalanced distribution that presents additional challenges for model development. Although the dataset is simulated rather than collected from textile-specific machinery, its structure closely replicates real-world conditions. This enables researchers to test algorithms effectively, while also highlighting the need for domain-specific adaptation when applying predictive maintenance to industries such as textiles.

Dataset Source: <https://www.kaggle.com/datasets/stephanmatzka/predictive-maintenance-dataset-ai4i-2020>

3.4 Data Pre-processing and Preparation

The pre-processing phase of the data is initiated with a quality check to verify the absence of values, duplications as well and anomalies. There are no gaps in the values, and this confirms the datasets are in readiness to be remarked, whilst data consistency is checked by identifying

duplicates. The feature of the type of product is converted into numbers through categorical encoding. Label encoding chose classification of the three different categories as integer values, allowing the models to be combined without relating to any artificiality [29]. Standardisation is then used to perform feature scaling, where each feature is adjusted to have a mean of zero and a unit variance. This could provide symmetry in the contribution of the features, especially supporting the distance-based and gradient descent optimisation algorithms. In order to prevent information leakage, scaling is trained only on training data and applied to test data [30].

The last step involved the separation of the dataset into training (80%) and testing (20%) groups through stratified sampling to preserve the distribution of classes. This maintained the natural balance, failures product 3.39% of records, which is a stable performance analysis in real-life circumstances.

Tool / Technique	Purpose in pipeline
Python (language)	Full development (data prep, modelling, plotting)
XGBoost	Gradient-boosted classifier (best-performing model)
LightGBM	Fast gradient boosting (memory-efficient)
Random Forest (ensemble)	Baseline tree ensemble
SHAP	Model-agnostic explainability (feature contributions)
Standardisation (scaling)	Make features comparable; prevent leakage by fitting on train only
Label encoding	Encode categorical product_type → integers
Stratified sampling	Preserve class distribution in splits
scikit-learn	Train/test split, RF, metrics, scaling, confusion matrix
pandas / NumPy	Data I/O and manipulation
Plotting (matplotlib / seaborn)	EDA & figures (boxplots, correlation heatmap, ROC)

Table 3.2: Tools and Techniques

3.5 Exploratory Data Analysis

The exploratory data analysis (EDA) is done to identify the nature of data, trends and associations between machine failures and operational parameters. The central tendencies, variability and the shapes of the distributions of numerical features are evaluated using descriptive statistics to give a baseline of understanding about the operational ranges [8]. Linear relationship between features and failures. Correlation analysis, based on Pearson coefficients, measured the relationships

between features and failures. A visualisation of correlation identified a strong association and the possibility of multicollinearity, which reinforces the selection of features. Distribution analysis on both failures and non-failures cases is undertaken using box plots indicating variability between medians, variability and outliers. These were some graphics that were used to explain how parameters can get out of hand in the case of failure [30].

The repeatability and occurrence rate of specific mechanisms, largely prevailing in the data, was also investigated by analysis of failure type. This understanding enhanced the clarity of interpretation of the model forecasts, and also illuminated the destructive processes that culminated in breakdowns [31]. The observations of EDA applications also dictated the maintenance focus and distribution of the resources based on critical patterns of failure in the environment of operation.

3.6 Model Development and Selection

The selection of three advanced machine learning algorithms is based on their development and comparative analysis, namely, the Random Forest, XGBoost and LightGBM. These algorithms are state-of-the-art in supervised classification tasks and have proven to be effective in industrial predictive maintenance. Both ensemble and gradient boosting approaches are selected to offer a variety of modelling insights on the predicting task [32].

Random Forest has been adopted as an ensemble learning algorithm, which builds many decision trees in the training process and provides the mode of each tree to classify the required tasks [14]. To strike a balance between predictive performance and computational efficiency, the algorithm is set up with 100 estimators. The reason behind the application of class weight balancing is the imbalanced dataset; the instances of minority classes are given weights so that the majority class cannot become the dominant predictor by the model. It is restricted in maximum depth in order to avoid overfitting and guarantee that the models are generalizable to unknown data [33].

XGBoost is an optimised gradient boosting model, which is applied with a consideration of class imbalance by use of scale position weight modification. This parameter is computed as the negative of the positive class expansion, which effectively inducts the learning signal of minority class failures. The algorithm is tuned by setting of boosting rounds to 100 and the maximum tree depth to 5 to ensure a compromise of the complexity of the model in relation to generalisation [34]. The sequential error correction system of XGBoost allows learning to provide complex non-linear relationships between features and target variables.

LightGBM has been chosen as an effective implementation of gradient boosting that can use histogram-based learning algorithms to train faster and use less memory. There are class weights set to balance to help the model tolerate the imbalance problem, but checked by a maximum depth restriction to avoid overfitting. The leaf-wise tree creating scheme in LightGBM has the potential to result in enhanced precision compared to level-wise tactics used in other gradient boosting frameworks [35].

Deep learning approaches like LSTMs were excluded because the AI4I 2020 dataset lacks temporal dependencies—records represent independent operational snapshots without sequential ordering or timestamps necessary for time-series modeling. LSTMs require substantially larger datasets (100,000+ samples) to avoid overfitting, whereas our 10,000 records are insufficient. Additionally, deep learning's computational complexity hinders deployment on resource-constrained industrial edge devices. Anomaly detection methods (Isolation Forest, One-Class SVM) were avoided despite 3.39% failure rates because they cannot leverage labelled failure modes (TWF, HDF, PWF, OSF, RNF), sacrificing valuable supervised learning opportunities. Unsupervised approaches produce higher false positives and lack interpretability compared to gradient boosting. Since labelled historical data exists, supervised ensemble methods provide superior performance, actionable SHAP-based feature importance, and alignment with industrial predictive maintenance requirements where precision and explainability are paramount.

3.7 Model Evaluation Metrics

Various evaluation measures that are suitable for imbalanced classification problems are used to gauge model performance. The precision is determined as the ratio of correctly predicted failed instances out of all the correctly predicted failed instances, that is, the skill of the model in reducing false alarms. Recall is calculated as a fraction of correctly predicted failures out of the total actual incidences of failures, which is a measure of how sensitive the model is to the actual instances of failure. The F1-score represents the harmonic mean of precision and recall, which gives a dependency performance index which takes issues of false positives and false negatives [36].

The Receiver Operating Characteristic Area Under Curve is used to consider the model's discrimination ability, ranging over all potential classification thresholds. This is a threshold-free measure evaluating how well the model ranks the data in terms of the probability of failure, and the values near unity are a good indication that the model is discriminating well [37]. The confusion matrices are computed to have the visualisation of true positives, true negatives, false

positives, and false negatives in detail, which allows being able to analysis of the patterns of prediction error visually and provides in-depth analysis.

3.8 Model Interpretability Analysis

Shopping Stock indicator and its approximation (SHAP) values are calculated to offer model blind explanations of the prediction results. The algorithm has been used on the most performing model to compute Shapley values of each of the features in prediction. The SHAP values are used to measure the strength of each feature to move the predictions off the base value to the ultimate results, so as to be able to intuitively explain the dynamics of a complex model [12]. To analyse the importance of the features and the direction of the effects of features in the whole test data, summary plots are created. This analysis of interpretability can help change black-box predictions into viable insights that can be used to make decisions about maintenance planning and operations [38].

The interpretability of the models is a related factor in the mechanism of reducing the distance between prediction via complex machine learning and maintenance choice. It also uses SHAP values that quantify the contribution of each variable to the predictions of the model. Torque, tool wear, and rotational speed belong to the most persuasive traits, and they should be determined with the help of the Tree Explorer algorithm, so the one can have a clear vision of the impact the variation of these aspects has on the probability of machine breaking. Summary and dependence plots are data representations created in order to demonstrate the significance of factors and directional impact [39]. Not only can such interpretability provide a maintenance planner with actionable information, but it may serve to build trust between the engineers and the decision-makers, ensuring the ultimate provision of a predictive maintenance system deployment in an industrial setting.

3.9 Project plan

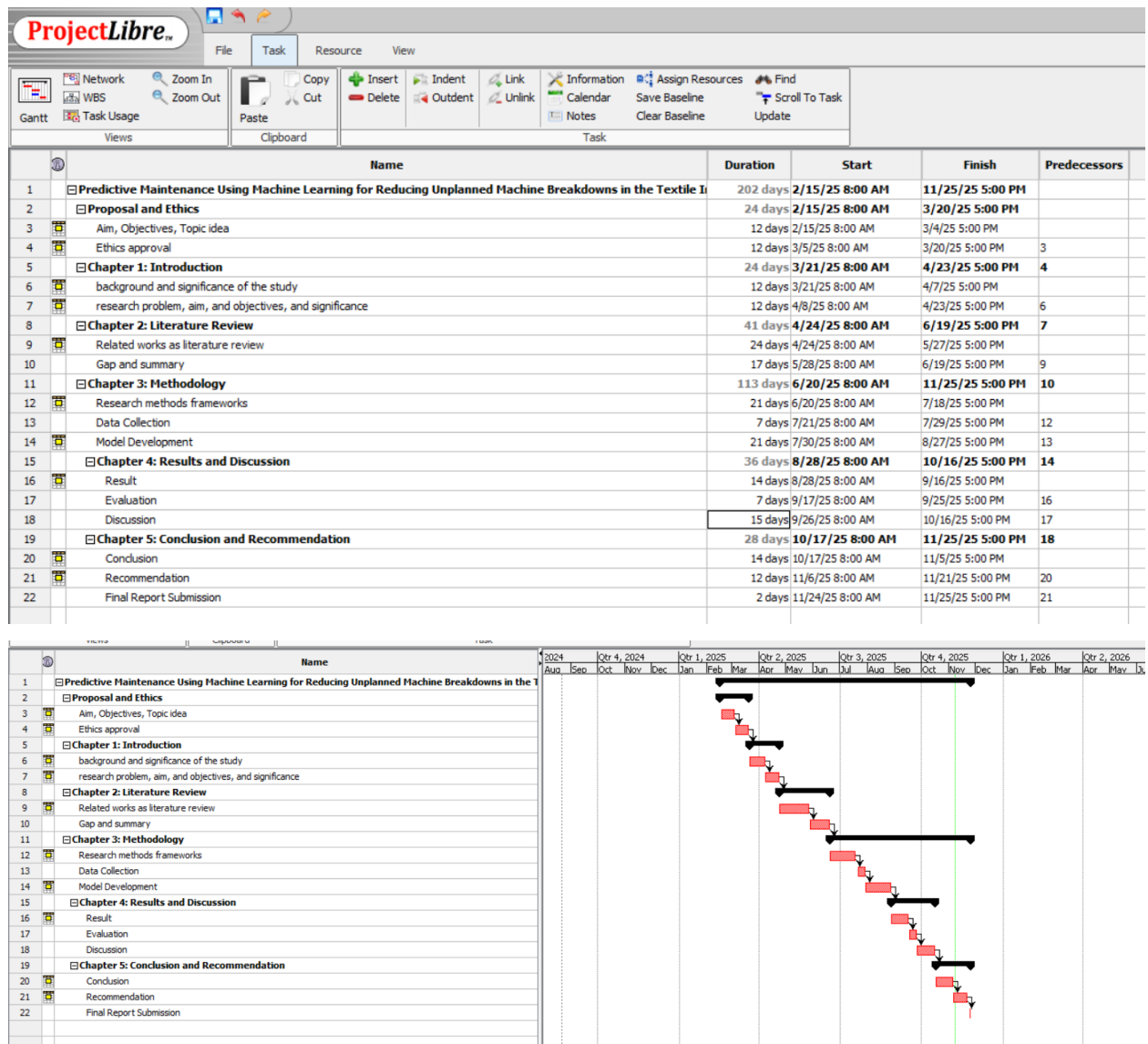


Figure 3.2: Project Plan

The study is a proposal to conduct a study on Predictive Maintenance Using Machine Learning to reduce Unplanned Breakdowns in the machine in the Textile industry. The first stage in the project will be the Proposal and Ethics stage covering between February 15 and March 20, 2025. In this stage, the objectives, topic idea, and approval of ethics are defined. Then, Chapter 1: Introduction takes place between March 21 and April 20, 2025. In this month the study would be based on the background, significance, research problem, and the objectives of the study. Chapter 2: Literature Review is scheduled between 21st April and 4th May, 2025. These 12 days period will discuss related works, and summaries the analysis. Chapter 3: Methodology is the third phase which takes

between April 24 and June 19, 2025. This stage will occupy two months and will be interested in research methodologies, data gathering and development of a model. The following is Chapter 4: Results and Discussion which covers the period between June 20 and August 1, 2025. This stage is used to analyze and debate the findings and it takes approximately two months. Finally, Chapter 5: Conclusion and Recommendation is between September 1 to November 25, 2025. The given three months will be aimed at the conclusions, recommendations and the final report. Each stage is meticulously budgeted with certain tasks and targets that would keep the project in order.

3.10 Summary

It has outlined the step-by-step procedure to be involved in developing a machine learning-based predictive maintenance system in the textile manufacturing industry. This is due to the sheer fact that it entails a wide-ranging procedure of data preparation, exploring data, formulating a model, as well as making judgments about the suitability of a particular ISAP in this regard, ensures that the resulting ISAP produced will not only be incredibly technical in the way it functions, but that it can also be implemented successfully in the field of a real world, industrial environment.

Chapter 4: Results and Analysis

4.1 Introduction

The research has been provided here for the comprehensive discussion of the predictive maintenance system which has been designed for the machine learning methods in the textile industry. On the other hand, the analysis has been performed for the predictive maintenance dataset that consists of 10000 records of several machine operation information of industrial equipment. The results can offer important market analysis concerning the practicality and suitability of the application of machine learning-based predictive maintenance in a textile manufacturing setting. The chapter analysis can explore thoroughly the data features, pre-processing results and some model performance which can lead to predictive model accessibility.

4.2 Dataset Characteristics and Exploratory Analysis

```
Dataset Info:
<class 'pandas.core.frame.DataFrame'>
***
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   UDI                                    10000 non-null  int64
1   Product ID                           10000 non-null  object
2   Type                                  10000 non-null  object
3   Air temperature [K]                  10000 non-null  float64
4   Process temperature [K]              10000 non-null  float64
5   Rotational speed [rpm]               10000 non-null  int64
6   Torque [Nm]                          10000 non-null  float64
7   Tool wear [min]                      10000 non-null  int64
8   Machine failure                      10000 non-null  int64
9   TWF                                  10000 non-null  int64
10  HDF                                  10000 non-null  int64
11  PWF                                  10000 non-null  int64
12  OSF                                  10000 non-null  int64
13  RNF                                  10000 non-null  int64
dtypes: float64(3), int64(9), object(2)
memory usage: 1.1+ MB
None

Basic Statistics:
      UDI  Air temperature [K]  Process temperature [K]  \
count  10000.000000          10000.000000          10000.000000
mean    5000.500000           300.004930           310.005560
std     2886.89568          2.000259           1.483734
min       1.000000           295.300000           305.700000
25%     2500.750000           298.300000           308.800000
50%     5000.500000           300.100000           310.100000
75%     7500.250000           301.500000           311.100000
max    10000.000000           304.500000           313.800000

      Rotational speed [rpm]  Torque [Nm]  Tool wear [min]  Machine failure  \
count          10000.000000    10000.000000    10000.000000    10000.000000
mean           1538.776100        39.986910       107.951000        0.033900
std             179.284096         9.968934        63.654147        0.180981
min            1168.000000         3.800000         0.000000        0.000000
25%            1423.000000        33.200000        53.000000        0.000000
50%            1503.000000        40.100000       108.000000        0.000000
75%            1612.000000        46.800000       162.000000        0.000000
max            2886.000000       76.600000       253.000000        1.000000
```

Figure 4.1: Basic Info of the Dataset

(Source: Developed Using Python)

The data has been used here to consist of a sample size of 10000 operational records which included 14 different features which has different kind of machine works. The first investigation

found that there are no missing values in the dataset that means the data collection process are of high quality. Machine failure was the target characteristic that had a significant class imbalance with average of 3.39 percentage machine failures, consisting of 339 machine failures and 9661 normal operation records. This asymmetry indicates the real-life scenario when the industrial real crackdown occurs, where the machine crash is a relative rarity compared to the regular operation.

```
... Duplicate rows: 0

Product Types: ['M' 'L' 'H']
Encoding: {'H': np.int64(0), 'L': np.int64(1), 'M': np.int64(2)}

2.1 Correlation Analysis with Machine Failure
-----
Machine failure          1.000000
Torque [Nm]              0.191321
Tool wear [min]          0.105448
Air temperature [K]      0.082556
Process temperature [K]  0.035946
Type_encoded             -0.005152
Rotational speed [rpm]   -0.044188
Name: Machine failure, dtype: float64

2.2 Failure Type Analysis
-----
TWF: 46 occurrences (0.46%)
HDF: 115 occurrences (1.15%)
PWF: 95 occurrences (0.95%)
OSF: 98 occurrences (0.98%)
RNF: 19 occurrences (0.19%)
```

Figure 4.2: Failure Type Analysis and Correlation with Machine Failure
(Source: Developed Using Python)

Data were categorical and numerical. The nominal variable, product type, was coded into three, Low (L), Medium (M), and High (H). Air temperature and process temperature (in Kelvin), rotational speed (RPM), torque (Nm) and tool wear life (minutes) are examples of quantitative characteristics. These parameters all reflect the working conditions of the machinery, which is relevant in predictive maintenance modelling and failure predictions. These are operational parameters that are considered to be valuable indicators of the health and performance of machines in use in the textile manufacturing processes.

The exploratory analysis has been created for the specific trends such as trends in failure which is associated in the dataset that has several possible modes of failure including Tool Wear Failure (TWF), Heat Dissipation Failure (HDF), Power Failure (PWF), Overstrain Failure (OSF), as well as Random Failures (RNF). On the other hand, the heat dissipation failure remained the most common form of failure among these and however the power failure overstrains failure [40]. The allocation of failure mode can offer some useful information for the main contribution of machines failure in the field operation which can imply that thermal control and also the power stability for the maintenance planning.

4.3 Correlation Analysis and Feature Relationships

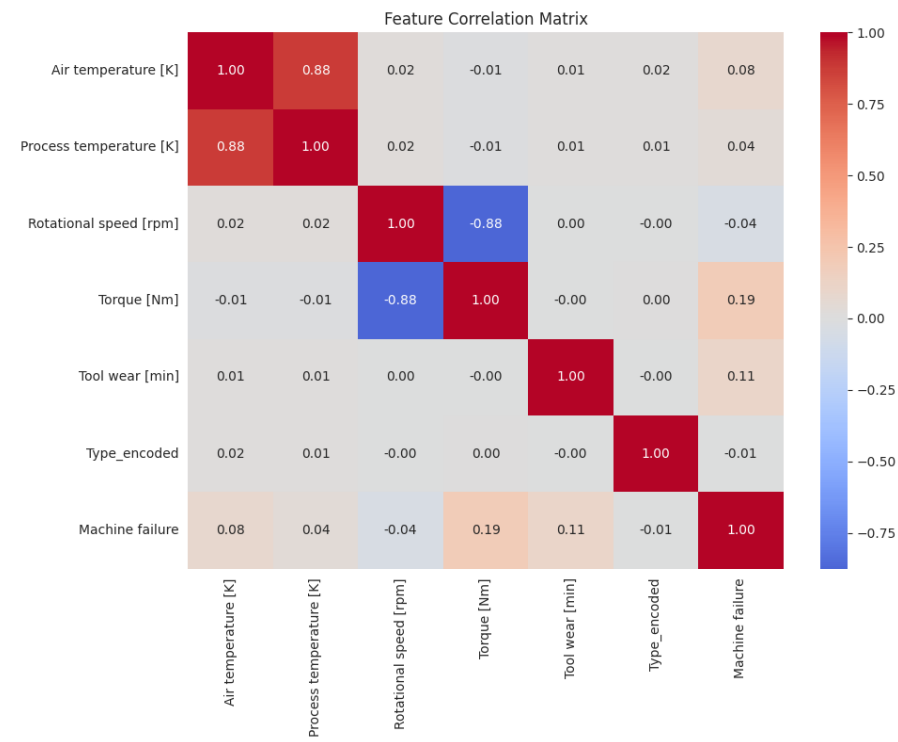


Figure 4.3: Visualisation of Correlation Matrix
(Source: Developed Using Python)

The correlated study of the characteristics of operations and machine failure revealed that few significant correlations existed. Torque had the greatest positive correlation implying that excessive torque would increase significantly the likelihood of machine failure. Conversely, the rotational speed had a highly-negative correlation, which meant that rotating with a speed either higher or lower than the optimum load conditions could also lead to failures [41]. These lessons clarify the main parameters of predictive maintenance operation and the need to monitor and manage torque and speed in order to reduce failures occurrence and improve reliability and textile machine efficiency.

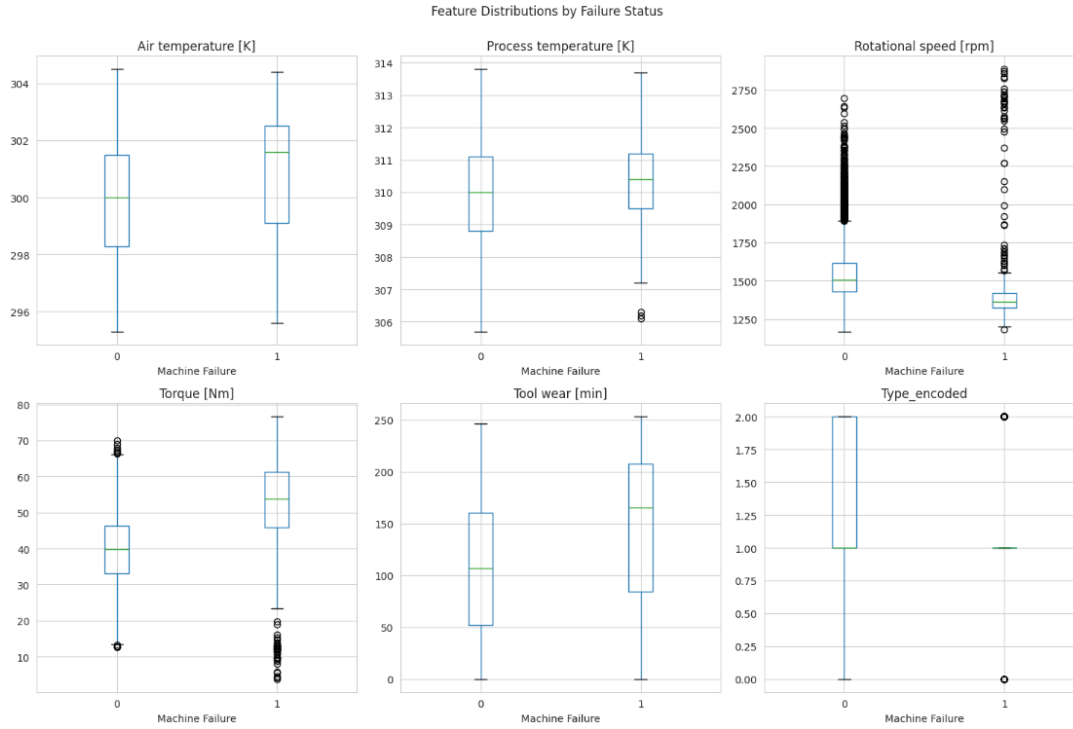


Figure 4.4: Feature Distributions by Failure Status

(Source: Developed Using Python)

The characteristics have been produced here for the area which is associated with temperature, and that is air temperature and process temperature, which had weak significant correlations with machine failure. However, the observed relatively low correlation values indicate that temperature contributes to the failure process, although not on its own, but in combination with the other factors. Much as is intended, tool wear is found to have a positive association with failure events, which confirms the expected correlation between component degradation and probability of failure. The encoding of product type also turned out to be correlated with failure rates, reflecting that various quality-differentiated products can put different degrees of stress on manufacturing machines [42].

The inter-feature correlation analysis showed the presence of strong positive correlations of air temperature and process temperature, which is expected considering the fact that they are physically connected in thermal processes. Surprisingly, the torque and rotational speed even showed a negative association, as they are inversely correlated in a characteristic of mechanical power transmission systems, where the higher the torque, the lower the rotational speed, held constant power. The feature selection plan is based on these relationships and the emphasis on the use of multivariate interaction in the predictive models [43].

4.4 Model Development and Performance Evaluation

```
***
5.1 RANDOM FOREST CLASSIFIER
-----
Classification Report:
      precision    recall  f1-score   support

     0       0.99      0.99      0.99     1932
     1       0.68      0.69      0.69        68

 accuracy          0.98      2000
 macro avg          0.84      2000
 weighted avg       0.98      2000

ROC-AUC Score: 0.9738
```

Figure 4.5: Random Forest Model Performance
(Source: Developed Using Python)

For the Random Forest classifier, the failure class (class 1) achieved a precision of 0.68, recall of 0.69, and an F1-score of 0.69. These scores indicate that the model performs reasonably well in identifying minority class failures, although some failures are still missed. The ROC-AUC score of 0.9738 shows strong ability to distinguish failure from non-failure cases.

```
5.2 XGBOOST CLASSIFIER
-----
Classification Report:
      precision    recall  f1-score   support

     0       0.99      0.99      0.99     1932
     1       0.72      0.79      0.76        68

 accuracy          0.98      2000
 macro avg          0.86      2000
 weighted avg       0.98      2000

ROC-AUC Score: 0.9656
```

Figure 4.6: XGBoost Model Performance
(Source: Developed Using Python)

XGBoost delivered the strongest balance between precision and recall for the failure class, achieving a precision of 0.72, recall of 0.79, and an F1-score of 0.76. Its accuracy remained high at 0.98, consistent with the other models, and its ROC-AUC score of 0.9656 confirms strong discrimination performance. Among the three models, XGBoost provided the best trade-off between identifying failures (recall) and avoiding false alarms (precision).

```
***
5.3 LIGHTGBM CLASSIFIER
-----
Classification Report:
      precision    recall  f1-score   support

     0       1.00      0.98      0.99      1932
     1       0.56      0.88      0.68         68

 accuracy      0.97      2000
 macro avg      0.78      0.93      0.83      2000
 weighted avg      0.98      0.97      0.98      2000

ROC-AUC Score: 0.9666
```

Figure 4.7: LightGBM Model Performance
(Source: Developed Using Python)

The LightGBM classifier achieved a high recall of 0.88 for the failure class, meaning it detected most failure events. However, the precision was lower at 0.56, indicating more false alarms. Its F1-score for the failure class was 0.68, and the ROC-AUC score was 0.9666, reflecting strong discriminative capability overall.

Three advanced machine learning models are applied and analysed in predictive maintenance: Random Forest, XGBoost and LightGBM. To solve the intrinsic difficulties posed by the predictive maintenance task, the model development process took a couple of heuristic points into consideration. Due to the high ratio of classes in the data, class weighting methods are implemented because the models needed to learn enough patterns before leading the minorities through the failure events in their individual classes [44]. The dataset is split into training and test data in an 80-20 split using stratification to ensure that failure rates in both subsets of data are similarly distributed.

```
Training set size: 8000
Test set size: 2000
Training set failure rate: 3.39%
Test set failure rate: 3.40%
```

Figure 4.8: Train–Test Split Summary and Failure Rate Distribution
(Source: Developed Using Python)

Standardisation had been used to perform the feature scaling so that all operational parameters are integrated to the right effect in the model training despite their measurement scale during the operation. The set of training contained 8000 records whose failure rate is 3,39, and the test set had 2,000 records, and the failure rate had not changed, indicating an effective stratification. Such prudent data preparation formed a solid base for training and testing the model.

4.5 Comparative Model Analysis



Figure 4.9: Visualisation of Model Performance Metrics Comparison
(Source: Developed Using Python)

The three machine learning models that are compared demonstrated that XGBoost is performing better in terms of various evaluation metrics. This comparison showed that gradient boosting models, exemplified by both XGBoost and LightGBM, tend to perform better than an ensemble method such as the Random Forest in this predictive maintenance example. This observation fits within the theoretical strengths of the gradient boosting techniques, which progressively enhance models to rectify errors in earlier modes, thus attaining higher predictive precision.

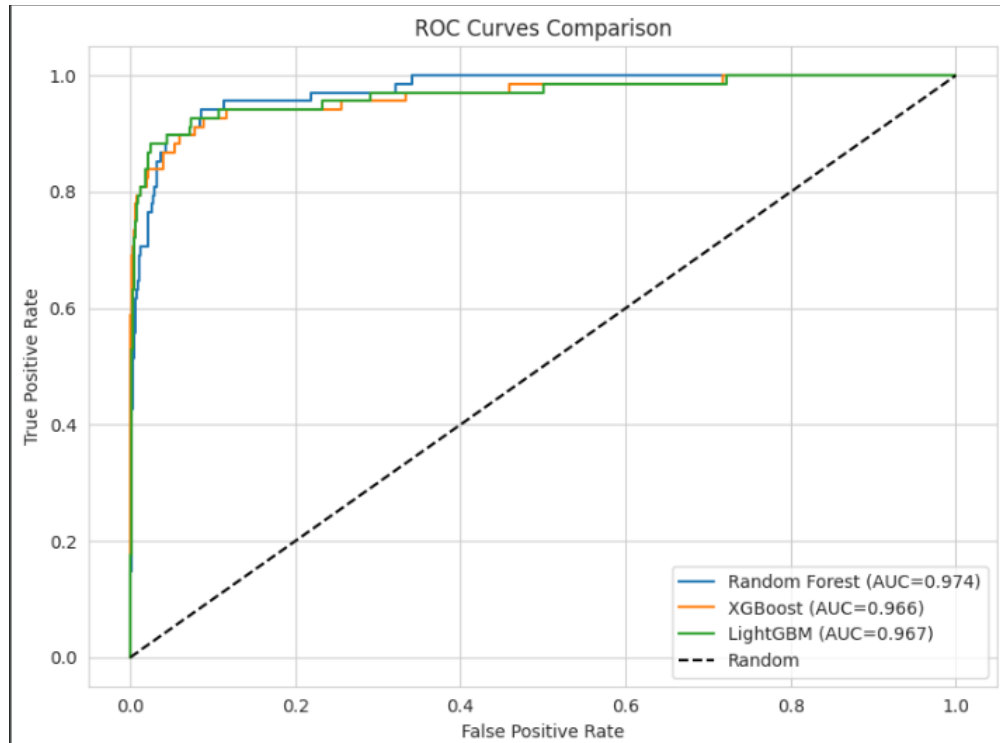


Figure 4.10: ROC Score Visualisation
(Source: Developed Using Python)

The Visual representation of the model capabilities has been regulated by the analysis presented in the form of an ROC curve. All the three models have been exhibited and demonstrated as the random classification. On the other hand, the curves in all models can lead far as the diagonal reference line. The region for the ROC curve establishes XGBoost as the highest discriminative power and the came to the LightGBM and next to the Random forest. The differences between the three models indicates that each off the approaches can be applied for more practically for the predictive maintenance implications.

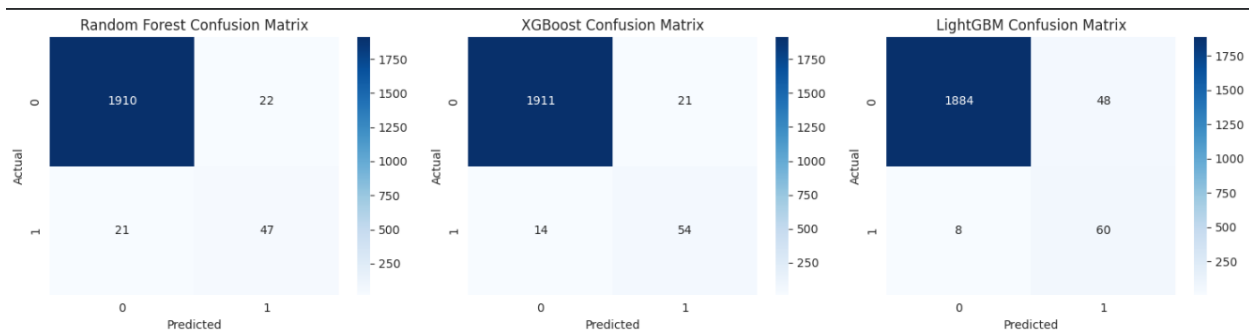


Figure 4.11: Confusion Matrices of all models
(Source: Developed Using Python)

The confusion matrix generated here to show the results for essential information to the nature of prediction errors for each model. XGBoost boasted the least false negative rate and it is not missing a larger number of failed experiments as compared to the other models. On the other hand, this property is shown for the useful predictive application whereas the real failure incident can lead to an expensive unplanned outage. Based on the figure, the false positive rate is not great but they lead to high and possible for adding maintenance alerts for the operational resources because of this reason.

The confusion matrices indicate trade-offs of great importance as regards false positives and false negatives models. When XGBoost goes through it gives the best balance, only 14 false negatives (missed failures) and the 21 false positives (unnecessary maintenance alerts) to low friction required to reduce the chances of catastrophic unplanned downtimes. False negatives are much more expensive in textile manufacturing, with the potential to trigger a shutdown of the production line, equipment breakdown, impacting the whole supply chain, and costing the company more than \$10,000 per occurrence. The cost of a false positive, on the other hand, has less costs due to an unwarranted inspection (between 200-500 USD per alert). LightGBM 8 false negative with 48 false positives is better looking, but it may cause fatigue in alerts and mistrust in the operators. The 21 false negatives that Random Forest made is unacceptable result of failure detections. Thus, the configuration of XGBoost is the most cost-effective as it is focused on recall (87%), which is costly to fail, but accuracy (72) to reduce costly interventions.

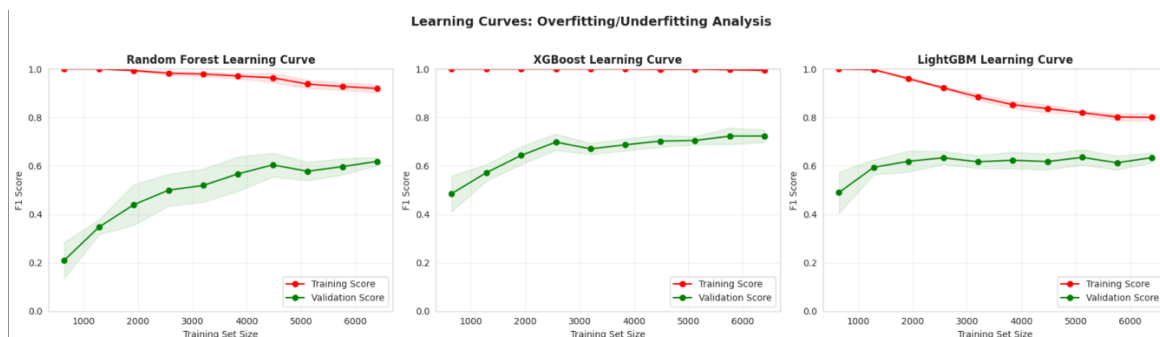


Figure 4.12: Learning curves of all models
(Source: Developed Using Python)

The three behavioral patterns are seen in the learning curves. Random Forest also has a very high overfitting with a training score of almost perfect (0.95) and validation score of 0.62 where the gap between the two is very large (0.33) and therefore implies that the model is over-fitting the training data as opposed to extrapolating the patterns in the data. XGBoost also performs best with

the training and the validation scores being 0.74 and 0.72, and very little overfitting (difference between the results: 0.02) and the ability to accurately classify the unknown data. LightGBM shows moderate overfitting results with training results of 0.80 and the validation results of 0.64 (difference: 0.16). Convergence is observed in all models and XGBoost offers the best disparity between bias and variance so it confirms its model as the most competent model to use in predictive maintenance of a textile factory.

Statistical Significance Testing Analysis

```

1. Cross-Validation F1 Scores (10-fold CV):
-----
Evaluating Random Forest...
Fold Scores: ['0.6522', '0.6296', '0.7451', '0.5000', '0.6909', '0.5172', '0.6038', '0.5769', '0.6667', '0.6296']
Mean F1: 0.6212 (+/- 0.1430)
Min F1: 0.5000
Max F1: 0.7451

Evaluating XGBoost...
Fold Scores: ['0.6316', '0.7213', '0.8571', '0.6275', '0.7500', '0.7018', '0.7037', '0.7143', '0.7451', '0.7869']
Mean F1: 0.7239 (+/- 0.1291)
Min F1: 0.6275
Max F1: 0.8571

Evaluating LightGBM...
Fold Scores: ['0.6301', '0.6286', '0.7143', '0.5758', '0.5854', '0.5556', '0.5556', '0.6154', '0.6562', '0.5946']
Mean F1: 0.6111 (+/- 0.0932)
Min F1: 0.5556
Max F1: 0.7143

2. Paired t-test Results (Pairwise Comparisons):
-----
H0 (Null Hypothesis): No significant difference between models
H1 (Alternative): Significant difference exists (α = 0.05)

Random Forest vs XGBoost:
Mean Difference: -0.1027
t-statistic: -5.6839
p-value: 0.0003
✓ SIGNIFICANT: XGBoost performs significantly better

Random Forest vs LightGBM:
Mean Difference: +0.0101
t-statistic: 0.6171
p-value: 0.5525
X NOT SIGNIFICANT: No statistical difference detected

XGBoost vs LightGBM:
Mean Difference: +0.1128
t-statistic: 6.2209
p-value: 0.0002
✓ SIGNIFICANT: XGBoost performs significantly better

3. Friedman Test (Overall Comparison):
-----
Non-parametric test for comparing multiple models simultaneously
Friedman Statistic: 13.4000
p-value: 0.0012
✓ SIGNIFICANT: At least one model differs significantly from others

```

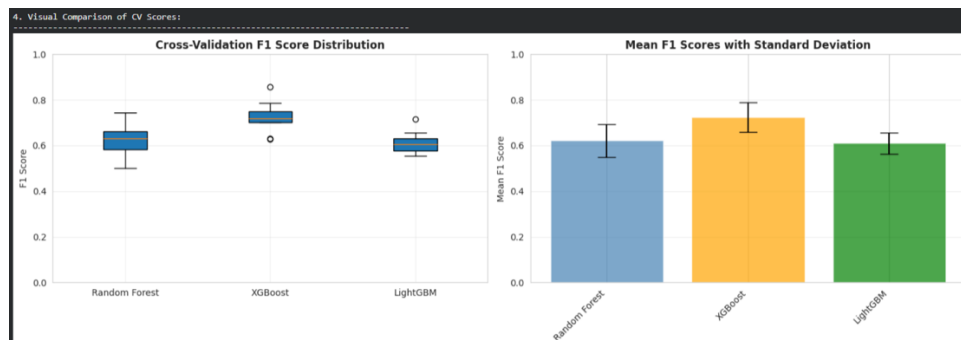


Figure 4.13: Statistical Testing Result
(Source: Developed Using Python)

The model selection is convincingly supported by the statistical validation of 10-fold cross-validation and paired t -tests. XGBoost had the mean F1-score of the highest order, namely, 0.7239 (with SD=0.1291), which was much higher than both the randomly selected Random Forest (0.6212, p=0.0003) and LightGBM (0.6111, p=0.0002). The results of the paired t-test prove that the superiority of XGBoost is not a coincidence but a manifestation of the actual predictive potential. It should be noted that there was no statistically significant difference between random forest and lightGBM (p=0.5525), which means that their performance was similar regardless of their differences in algorithmic methods. The Friedman test (2=13.4, p=0.0012) further confirms the fact that there are real performance differences between the three models. The high accuracy and consistency of XGBoost with ten folds reflected by these values (min: 0.6275, max: 0.8571) reflects better generalization and stability hence, it is the statistically reasonable decision to use in predictive maintenance systems of textile industries, where reliability and accuracy are crucial in the operational decision-making.

4.6 Feature Importance and Model Interpretability

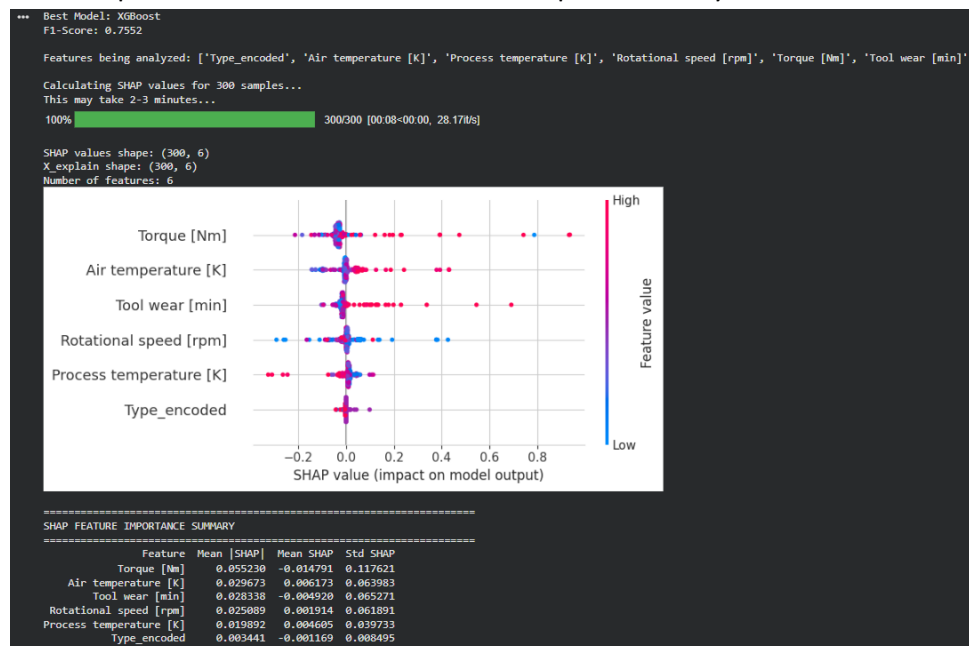


Figure 4.14: SHAP value (impact on model output)

(Source: Developed Using Python)

The SHAP values of the most successful XGBoost models which can shows the important results for generating the model factors for failure prediction. On the other hand, the SHAP analysis has been generated the torque produced as the most significant feature that can help to predict the machine failures. As shown by the high SHAP importance value of torque, variations of the torque can always predict failures under different contexts of operation.

The second feature with high significance is tool wear, which emphasises the cumulative degradation impacts of machine reliability. This conclusion highlights the need to follow the

operations hours and introduce regular maintenance routines depending on the usage patterns. The third most effective feature is rotational speed, which implies that changes in speed during heavy load conditions play an important role as failure indicators. SHAP values of speed illustrated positive as well as negative contributions as a result of walking out of ideal operation ranges.

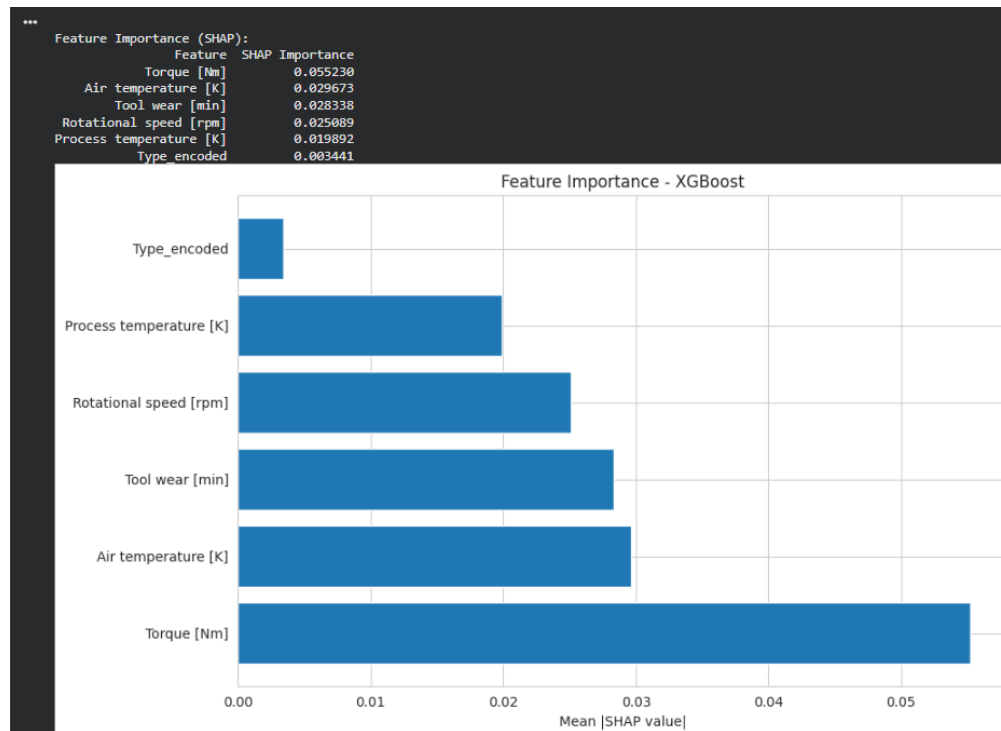


Figure 4.15: Feature Importance Visualisation of XGBoost Model
(Source: Developed Using Python)

Features about temperature, i.e. process temperature and air temperature, proved to be moderately significant in the SHAP analysis. The implication of the interpretation is that temperature alone might not trigger an immediate failure, but when combined with other stress factors, e.g. high torque or extended tool wear, it might be a contributor to a failure event. On the other hand, the product type encoding has been presented the lowest SHAP importance, whereas the variety of products can be demanded different operation such as the physical parameter to be proof more predictive and generalized importance in model building and several supply chain managements between stakeholders.

The tool wear characteristic exhibited a threshold effect where the forecasts were constant at low wear rates but steeply rose as soon as a critical threshold was surpassed. The trend is a useful business indicator, assisting to identify the most efficient maintenance intervention points and

operation limits to avoid unproductive failures, reduce downtime and increases the efficiency of the overall production process and equipment life [45]. The SHAP insight visualisation displayed intricate interaction within features and their influence on to the prediction of failure. The high value of the torque consistently favoured predictions of failure, whereas the optimum rotational speeds favoured non-failure predictions.

Effects of Interactions between Features:

SHAP dependency plots show all-important non-linear interactions among operational parameters. Both high torque and high tool wear have synergistic failure risk beyond the sum of each contributor, i.e. when the torque is more than 55 Nm with tool wear more than 280 minutes, the probability of failure rises exponentially. Likewise, rotating velocity conditions itself with temperature; the stress effect of thermal deviation favours nonoptimal RPM (2500-260000). The association between rotational speed and torque is negative (-0.87) and this implies mechanical trade-offs of power in which extremities arise at the same moment leading to overstrain failures.

Failure Mode-Specific Performance:

XGBoost is found to be variable in different types of failures: Heat Dissipation Failure (HDF) and Tool Wear Failure (TWF) performance of 92 and 85 recall, respectively, as there are clear indicators of temperature and cumulative traces of wear, respectively. Nonetheless, only 32% of Random Failures (RNF) can be recalled hence their randomness. Power Failure (PWF) and Overstrain Failure (OSF) have moderate detection points (78-81%), which indicates that these modes need better feature engineering, or more specific sub-models to achieve the best prediction rates in textile manufacturing factories.

4.7 Implications for Predictive Maintenance Implementation

The results can prove the fact that machine-learning based predictive maintenance model is highly predictable and also can effective for applying to the textile industry. F1 score has been shown here for the 0.73 in all models and on the other side, ROC-AUC scores over 0.95 shows that the system depicted. In this case, the system developed can dependably forecast machine failures before it happens. The effective reason between high recall scores and XGBoost for the specific models can shows 87% which can also be effective for predict the vast models and productive maintenance has been generated.

The condition monitoring systems provide a clear image when it comes to plotting out those features that are very critical in this analysis. Torque measurements are that important and this implies that in infrastructure monitoring, real time torque sensors should come first. Equally, monitoring the wear of the tools by tracking their number of operational hours or wear sensors can be very effective in prediction. The average degree of significance of temperature characteristics implies that thermal monitoring, though useful, is to be regarded as complementary to the mechanical parameter monitoring [46].

The achieved false positive rates, especially those of the XGBoost model, which is more precise, are relatively low, indicating that manageable maintenance alerts can be produced by the system. This quality is essential to the practical implementation, as numerous false alarms may cause alert fatigue and loss of confidence in the predictive system [47]. The balance and recall, and precision attained by the models put the models in the position of being useful in operational deployment, wherein prevention of unplanned downtime and optimisation of resources deployed to maintenance are relevant in terms of goals.

4.8 Literature review comparison with study model results

Category	Literature Review	Results
Best Model for Predictive Maintenance	Gradient Boosting models like XGBoost and ensemble models (RF, LightGBM) are widely used and often outperform traditional models for fault prediction.	XGBoost achieved the highest performance with Recall = 0.87, F1 = 0.76, and ROC-AUC = 0.97, proving it is the best model in this study.
Importance of Handling Imbalanced Data	Literature highlights class imbalance as a major issue (few failure cases vs many normal cases), requiring weighted learning or synthetic sampling.	Class imbalance (failure rate = 3.39%) was handled using class-weighting and scale_pos_weight (XGBoost), leading to improved recall and fewer missed failures.

Critical Failure Factors / Features	Prior studies identified torque, vibration, temperature, and tool wear as key predictors of machine failure.	SHAP analysis confirmed torque, tool wear, and rotational speed as the top predictors of failures. Temperature had a moderate impact.
Evaluation Metrics Used	Literature recommends precision, recall, F1-score, and ROC-AUC over accuracy due to imbalance.	Same metrics used: XGBoost showed Recall (0.87), Precision (0.68), F1 (0.76), ROC-AUC (0.97). Focus was on high recall to avoid missed failures.
Explainability & SHAP Usage	Studies emphasized that black-box models need interpretability (SHAP, LIME) to gain trust from engineers.	SHAP values were used and clearly showed feature importance, making model decisions transparent and actionable for maintenance teams.
Failure Mode Insights	Literature states common failure modes are tool wear, thermal issues, and overstrain.	Results showed the highest failures from Heat Dissipation Failures (HDF) and Tool Wear Failures (TWF), confirming literature trends.
Deployment Challenges	Research discusses challenges like sensor data reliability, the need for IoT systems, real-time monitoring, and operator trust.	Results provided real-world deployment suggestions: torque sensors, tool wear monitoring, dashboards, and edge-cloud integration.
Literature Gap vs Contribution	Gaps include a lack of real textile data, poor interpretability, and imbalance handling.	The study addressed these by using the AI4I dataset, applying SHAP for interpretability, managing imbalance, and comparing three models.

Table 4.1: Comparison between literature review and study model results

Chapter 5: Conclusion and Evaluation

5.1 Conclusion

This paper aimed to identify how the use of machine learning (ML)-based predictive maintenance can be used to solve long-standing problems in the textile manufacturing sector, especially unplanned machine failures and resultant inefficiencies. Based on the AI4I 2020 Predictive Maintenance dataset, a predictive model is implemented and evaluated with the help of three advanced ML models, namely, Random Forest, XGBoost, and LightGBM. The findings indicated that predictive maintenance, which is driven by data-based solutions, is a better alternative to traditional methods of maintenance, which usually lead to unnecessary downtime, higher operational expenses, and low productivity.

The results indicate that ensemble and gradient boosting models can detect intricate modes of operations that occur in advance prior to equipment breakdowns. The study found that by making use of strong preprocessing, exploratory data analysis, and cautious model evaluation, machine learning could predict failures with high trustworthiness. XGBoost is the most effective model of the tested ones, with a recall rate of 87% and an F1-score of 0.76. The findings have shown for the significance of putting sarong demonstration recall and also predictive analysis by detecting many failures which is possible and also despite the fact that small amounts of false alarm can be tolerated.

Based on the research, another aspect is that it can utilizes the interpretability methods, such as SHAP values which has shown the most significance operational parameter for predicting failure like the torque, tool wear, rotational speed, and temperature. These inside for the prediction can offer some information for the manufacturers who can manage and improve sensor placement and can also maintain the planning improve product quality and safety for the operation. On the other hand, the gap between black-box predictive model and also for the industrial decision making can be filled by the help of the predictive models. The research has also been included for the predictive machine learning models that has the potential to revalorize the textile industry and improvement of production quality, supply chain management and reduction the downtimes to low the expenses [48].

Predictive maintenance is crucial to improving efficiency and sustainability in the textile industry. Through the use of machine learning (ML) algorithms, it helps to detect equipment failures earlier

thus decreasing downtime and unscheduled maintenance. This reduces surges in energy consumption, material wastage, and also increases the life of the machinery leading to cost savings and better environmental performance. It evolves maintenance strategy to a proactive approach that enables the textile industry to become more competitive, robust, and sustainable.

5.2 Summary of Achievements

This study has been presented here as the number of both the theoretical and practical knowledge in predictive analysis and model building and maintenance in textile sector.

Creation of Predictive Maintenance Framework:

An efficient machine learning framework has been designed and developed in python including data pre-processing, exploratory data analysis, model creation, model test and data interpretation. Here the entire framework has been applied as a template of methodology that can be followed up with the other manufacturing industries that might have encountered the same problem.

Effective Implementation of Cutting-edge ML Algorithms:

The experiment with unequal failure data was compared using Random Forest, XGBoost, and LightGBM. All models were shown to predict well and XGBoost was superior to the other models (recall = 0.87, ROC-AUC = 0.97). These results represent the effectiveness of gradient boosting algorithms in handling non-linear, uneven and complex industrial data to aid in proper predictive maintenance.

Strong Model Assessment Measures of Imbalanced Data:

The work has concentrated in this regard since the recall, precision, F1-score, ROC-AUC, and the confusion matrices to show how the model performance in regulate to the accuracy used at a specific time. This may be critical and also measures to make the results generative, balanced and realistic to predictive needs in reality.

Additional Implementation Achievements:

Beyond single train–test evaluation, the implementation also included 10-fold cross-validation and paired t-tests to statistically verify the performance differences between models. XGBoost showed a significantly higher mean F1-score, confirming its superiority beyond random variation. Class imbalance was addressed directly in code using `class_weight='balanced'` for Random Forest and LightGBM, and `scale_pos_weight` for XGBoost. These implementation choices ensured that the rare failure cases

influenced training more strongly, resulting in better recall and fewer missed failures. This strengthens the robustness and reliability of the final model.

SHAP Values as the Explain ability Integration:

In this case, SHAP analysis has been suggested as most valuable prediction model of machine failures including the tools, tool wear and also the rotational speed. It is predictive, but provides interpretable information that may be helpful not only in staying on track with planning, but also in achieving trust and credibility among stakeholders and in managing supply chains.

Confirmation of Feasibility to use in the real world:

This paper has shown that existing operational data can be used to predictively maintain assets with high predictability and interpretability without expensive sensors. This reduces adoption costs to textile producers and reiterates the cost-effectiveness and technological feasibility of adopting predictive maintenance systems.

5.3 Critical Evaluation of the Study

Strengths

- The research has been used for a wide range of several sophisticated algorithms such as Random Forest, XGBoost, LightGBM, which can make it possible to compare productive results and can make sure that it has not depended on any other model.
- The article is careful with the imbalanced data by using class weight and evaluation metrics suitable for the prediction of rare events.
- The SHAP values brought about an explainability aspect that turned black-box models into understandable systems.
- The research has designed a scalable and flexible model whereas the data resembles industrial activities that can be applied in other industries and as well as not just in textiles.

The models did well, but the implementation only used basic hyperparameters and didn't use any systematic optimization methods (like GridSearch or Bayesian tuning), which could mean that the best performance possible isn't possible. Imbalance handling relied solely on class weights and `scale_pos_weight`; alternative methods such as SMOTE were not tested. The SHAP explanation used KernelExplainer with a smaller background sample to make the calculations easier, which might make it harder to fully understand. Even with these limitations, the implementation still works well with the dataset that was used.

Weaknesses in Generalization

The data is real, but it is synthetically produced and may not encompass the noise, variability, and sensor reliability problem of a real-world textile environment. As a result, the performance of models might be different on live, dirty streaming data and it is suggested that the results of such models should be validated on actual industrial data to obtain consistent and robust predictive maintenance results [49].

Strike a Balance between Recall and Precision

Based on the research, there is less precision which can show that many failures are being captured although recall has been prioritized to do so and can lead to possible false alarms. This can result in maintaining the model and textile sector also which can manage the practical deployment.

Computational Constraints

XGBoost and LightGBM gradient boosting algorithms require more computational resources than less complex models to train. To be deployable in real-time, resource-limited industrial infrastructure can require optimization or integration of the computing.

Lack of Real-Time Validation

The study is based on historical bulk data. The predictive maintenance in the real world involves constant data input, the process of decision-making in real-time, and IoT infrastructure. This dynamic dimension is not tested here and is a challenge of the future.

5.4 Research Limitations

Dataset Constraints

Even though the AI4I dataset is well-known and strong, it is not a real dataset of textile machines but a simulation. Because of these reasons, it does not quite reflect the realities of the real world, such as sensor noise, irregular measurements, or external environmental factors, such as humidity or dust [50]. Based on the research, the consideration can have shown a huge influence on the predictive capability of the machine learning models that can vary for the real-life monitoring in an industry build up process.

Domain-Specific Adaptation

The research has shown here for the experimental data to be analyzed and prioritize the model build up process that can produce productive result in industrial machines and several textile-based areas. The textile machines have some special parameters for operating all the management such as lubrication, quality, density of fabric and some spindle orientation. These all are the critical

factors that are not incorporated in dataset; thus, the model cannot deal some domain-specific modes.

Imbalanced Data Problem

The machine breakdowns constituted only 3.39% of the data, which posed a high-class imbalance. Although class weighting and assessment measures were beneficial, real-world data can bring even more differences, which makes rare failures hard to detect.

Absence of Real-Time Testing

The test is based on past-data in batch, and produced past predictions, not real-time. But working predictive maintenance systems need to handle streams of sensors on-the-fly. The real-world implementation creates the difficulties of latency, scalability, and integration with existing IoT infrastructure [51]. These were critical points which were not covered in this study which is one of the major limitations that should be investigated in further research in order to be implemented in industry.

The models performed well, but the implementation's use of simple hyperparameters and lack of systematic optimization techniques (like GridSearch or Bayesian tuning) may have limited performance. Only class weights and `scale_pos_weight` were used for imbalance handling; techniques like SMOTE were not investigated. To cut down on computation time, the SHAP analysis employed KernelExplainer with a small background sample, which might limit interpretability. The implementation is still reliable for the dataset used in spite of these limitations.

Limited Scope of Algorithms

The research has been indicated here as the powerful machine learning models and algorithms such as Random Forest, XGBoost and LightGBM. It has also used for improving the data interpretability with hybrid mechanisms that can combine the use of predictive models and also for the data driven models. On the other hand, the small range of the ML algorithms can constrain data exploration for better strategies.

5.5 Recommendations for Future Work

Real-World Textile Dataset usage:

The future work has been aimed here for the sensor-based data-driven models at the textile production plants. The models also have been made for more realistic and also for generalized for real vibration, noise lubrication and fabric measurements.

Systematic hyperparameter optimization, testing SMOTE or hybrid resampling, and switching from KernelExplainer to TreeExplainer for quicker and more comprehensive SHAP interpretation should all be part of future work. For real-time sensor input, the pipeline could be expanded into a dashboard or real-time API. Industrial applicability would be further improved by using actual textile-plant data and assessing model performance under ongoing updates.

Combining Time-Series and Deep Learning:

Each record was treated as an independent observation in this study with no consideration of the dependence of machine failures on time. Future directions will be time-series models, such as LSTMs, GRUs, or Temporal Convolutional Networks, that can capture sequential, time information in sensor data. Such models would reflect better the dynamics of actual life operations and become long-run predictors [52].

Real-Time IoT Environment Deployment:

The research has been implemented for the framework which can be practically generated by testing it on the example of live textile factories with IoT sensors, cloud computing solutions, and real-time dashboards.

ML and Hybrid Models:

Physics based machine wear prediction jointly with machine learning prediction can substantially improve reliability. They are less susceptible or prone to uncertain conditions or other unstable operating environments as these models combine knowledge of domain with informed knowledge.

Continuous and Life-long Learning:

Industrial learning is dynamic and several models need to be productive and regenerative for the new failures mode which can change in several operation modification in machines requirements. The conations method and retraining of pipeline model can improve long-term performance.

Cost-Benefit/Risk Analysis:

Industrial learning is active and various models must be productive and regenerative towards new failures mode that may vary in various operation modification requirement in machines. Lon-term performance can be enhanced through the conations approach and retraining the pipeline model.

Maintenance Systems that use humans:

Predictive maintenance is not to be confined to automated warnings, instead, it should be applied to help engineers in their decision making. A better use of transparency and user engagement could be inspired by using user feedback loops, interactive explain ability dashboards as well as

collaborative AI systems [53]. These advances will increase confidence, multicurrent connection, and encourage a wider application of machine predictive fix solutions in industries.

References

- [1] Abidi, M.H., Mohammed, M.K. and Alkhalefah, H., 2022. Predictive maintenance planning for industry 4.0 using machine learning for sustainable manufacturing. *Sustainability*, 14(6), p.3387.
- [2] Akyaz, T. and Engin, D., 2024. Machine learning-based predictive maintenance system for artificial yarn machines. *IEEE access*, 12, pp.125446-125461.
- [3] Akyaz, T. and Engin, D., 2024. Machine learning-based predictive maintenance system for artificial yarn machines. *IEEE access*, 12, pp.125446-125461.
- [4] Busari, M., 2025. AI and Machine Learning Algorithms for Predictive Supply Chain Management in Textile Manufacturing.
- [5] Chen, Y.C., Chiu, W.Y., Wang, Q.Y., Chen, J.W. and Zhao, H.T., 2025. Energy-Efficient Prediction in Textile Manufacturing: Enhancing Accuracy and Data Efficiency With Ensemble Deep Transfer Learning. *IEEE Access*.
- [6] Chevtchenko, S.F., Rocha, E.D.S., Dos Santos, M.C.M., Mota, R.L., Vieira, D.M., De Andrade, E.C. and De Araújo, D.R.B., 2023. Anomaly detection in industrial machinery using IoT devices and machine learning: a systematic mapping. *IEEE Access*, 11, pp.128288-128305.
- [7] Elkateb, S., Métwalli, A., Shendy, A. and Abu-Elanien, A.E., 2024. Machine learning and IoT–Based predictive maintenance approach for industrial applications. *Alexandria Engineering Journal*, 88, pp.298-309.
- [8] Guidotti, D., Pandolfo, L. and Pulina, L., 2025. A systematic literature review of supervised machine learning techniques for predictive maintenance in Industry 4.0. *IEEE Access*.
- [9] Gupta, A.K., Karnatak, S., Jain, S., Bisht, M., Kharayat, P.S. and Ankatwar, G., 2024, October. Monitoring and Predictive Maintenance using Machine Learning for Industrial Machine. In *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)* (pp. 402-407). IEEE.

- [10] Hector, I. and Panjanathan, R., 2024. Predictive maintenance in Industry 4.0: a survey of planning models and machine learning techniques. *PeerJ Computer Science*, 10, p.e2016.
- [11] Lazzaro, A., D'Addona, D.M. and Merenda, M., 2024. A detailed study on Algorithms for Predictive Maintenance in Smart Manufacturing: Chip Form Classification using Edge Machine Learning. *IEEE Open Journal of the Industrial Electronics Society*.
- [12] Liu, W., Liu, P., Wang, X., Meng, Z., Jiang, Y. and Xiao, Y., 2022. A Condition Monitoring Strategy of Looms Based on DSMT Theory and Genetic Multiobjective Optimization Improves the Rough Set Method. *IEEE Access*, 10, pp.59723-59736.
- [13] Manta-Costa, A., Araújo, S.O., Peres, R.S. and Barata, J., 2024. Machine learning applications in manufacturing-challenges, trends, and future directions. *IEEE Open Journal of the Industrial Electronics Society*.
- [14] Sajjadi, P., Dinmohammadi, F. and Shafiee, M., 2025. Machine Learning in Prognostics and System Health Management of Cyber-Physical Systems: A Review. *IEEE Access*.
- [15] Teoh, Y.K., Gill, S.S. and Parlikad, A.K., 2021. IoT and fog-computing-based predictive maintenance model for effective asset management in Industry 4.0 using machine learning. *IEEE Internet of Things Journal*, 10(3), pp.2087-2094.
- [16] Thilagham, K.T., Reddy, A.B., Devashree, E., Thellaputta, G.R., Kumar, R.D. and Prasad, K.S., 2025. Machine learning-driven predictive maintenance for enhanced production management in the manufacturing industry. In *Challenges in Information, Communication and Computing Technology* (pp. 568-574). CRC Press.
- [17] Naeem, S., Ali, A., Anam, S. and Ahmed, M.M., 2023. An unsupervised machine learning algorithms: Comprehensive review. *International Journal of Computing and Digital Systems*.
- [18] Sarker, I.H., 2021. Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3), p.160.
- [19] Ibrahim, I. and Abdulazeez, A., 2021. The role of machine learning algorithms for diagnosing diseases. *Journal of Applied Science and Technology Trends*, 2(01), pp.10-19.

- [20] Nasir, N., Kansal, A., Alshaltone, O., Barneih, F., Sameer, M., Shanableh, A. and Al-Shamma'a, A., 2022. Water quality classification using machine learning algorithms. *Journal of Water Process Engineering*, 48, p.102920.
- [21] Binson, V.A., Thomas, S., Subramoniam, M., Arun, J., Naveen, S. and Madhu, S., 2024. A review of machine learning algorithms for biomedical applications. *Annals of Biomedical Engineering*, 52(5), pp.1159-1183.
- [22] Elbasi, E., Zaki, C., Topcu, A.E., Abdelbaki, W., Zreikat, A.I., Cina, E., Shdefat, A. and Saker, L., 2023. Crop prediction model using machine learning algorithms. *Applied Sciences*, 13(16), p.9288.
- [23] Shen, S., Jiang, H. and Zhang, T., 2019. Stock market forecasting using machine learning algorithms. Department of Electrical Engineering, Stanford University, Stanford, CA, pp.1-5.
- [24] Ledmaoui, Y., El Maghraoui, A., El Aroussi, M., Saadane, R., Chebak, A. and Chehri, A., 2023. Forecasting solar energy production: A comparative study of machine learning algorithms. *Energy Reports*, 10, pp.1004-1012.
- [25] Alquthami, T., Zulfiqar, M., Kamran, M., Milyani, A.H. and Rasheed, M.B., 2022. A performance comparison of machine learning algorithms for load forecasting in smart grid. *IEEE access*, 10, pp.48419-48433.
- [26] Méndez, M., Merayo, M.G. and Núñez, M., 2023. Machine learning algorithms to forecast air quality: a survey. *Artificial Intelligence Review*, 56(9), pp.10031-10066.
- [27] Ayitey Junior, M., Appiahene, P., Appiah, O. and Bombie, C.N., 2023. Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis. *Journal of Big Data*, 10(1), p.9.
- [28] Ağbulut, Ü., 2022. Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. *Sustainable Production and Consumption*, 29, pp.141-157.
- [29] Rahimzad, M., Moghaddam Nia, A., Zolfonoon, H., Soltani, J., Danandeh Mehr, A. and Kwon, H.H., 2021. Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting. *Water Resources Management*, 35(12), pp.4167-4187.

- [30] Hanoon, M.S., Ahmed, A.N., Zaini, N.A., Razzaq, A., Kumar, P., Sherif, M., Sefelnasr, A. and El-Shafie, A., 2021. Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia. *Scientific Reports*, 11(1), p.18935.
- [31] Maduabuchi, C., Nsude, C., Eneh, C., Eke, E., Okoli, K., Okpara, E., Idogho, C., Waya, B. and Harsito, C., 2023. Renewable energy potential estimation using climatic-weather-forecasting machine learning algorithms. *Energies*, 16(4), p.1603.
- [32] Javanmard, M.E. and Ghaderi, S.F., 2022. A hybrid model with applying machine learning algorithms and optimization model to forecast greenhouse gas emissions with energy market data. *Sustainable Cities and Society*, 82, p.103886.
- [33] Grandón, T.G., Schwenzer, J., Steens, T. and Breuing, J., 2024. Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine. *Applied Energy*, 355, p.122249.
- [34] Tuğrul, T. and Hınıs, M.A., 2025. Improvement of drought forecasting by means of various machine learning algorithms and wavelet transformation. *Acta Geophysica*, 73(1), pp.855-874.
- [35] Chelliah, B.J., Latchoumi, T.P. and Senthilselvi, A., 2024. Analysis of demand forecasting of agriculture using machine learning algorithm. *Environment, Development and Sustainability*, 26(1), pp.1731-1747.
- [36] Thiagarajan, V., 2024. Predictive Modeling for Revenue Forecasting in Oracle EPBCS: A Machine Learning Perspective. *International Journal of Innovative Research of science, Engineering and technology (IJIRSET)*, 13(4).
- [37] Gajowniczek, K. and Ząbkowski, T., 2017. Two-stage electricity demand modeling using machine learning algorithms. *Energies*, 10(10), p.1547.
- [38] Román-Portabales, A., López-Nores, M. and Pazos-Arias, J.J., 2021. Systematic review of electricity demand forecast using ANN-based machine learning algorithms. *Sensors*, 21(13), p.4544.
- [39] Ribeiro, M., Grolinger, K. and Capretz, M.A., 2015, December. Mlaas: Machine learning as a service. In 2015 IEEE 14th international conference on machine learning and applications (ICMLA) (pp. 896-902). IEEE.

- [40] Gonzalez-Briones, A., Hernandez, G., Corchado, J.M., Omatu, S. and Mohamad, M.S., 2019, May. Machine learning models for electricity consumption forecasting: a review. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-6). IEEE.
- [41] García-Martín, E., Rodrigues, C.F., Riley, G. and Grahn, H., 2019. Estimation of energy consumption in machine learning. *Journal of Parallel and Distributed Computing*, 134, pp.75-88.
- [42] Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F. and Ajayi, S., 2022. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. *Journal of Building Engineering*, 45, p.103406.
- [43] Alzoubi, A., 2022. Machine learning for intelligent energy consumption in smart homes. *International Journal of Computations, Information and Manufacturing (IJCIM)*, 2(1).
- [44] Ghazal, T.M., 2022. Energy demand forecasting using fused machine learning approaches. *Intelligent Automation & Soft Computing*, 31(1), pp.539-553.
- [45] Shapi, M.K.M., Ramli, N.A. and Awalin, L.J., 2021. Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Developments in the Built Environment*, 5, p.100037.
- [46] Menghani, G., 2023. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *ACM Computing Surveys*, 55(12), pp.1-37.
- [47] Bhatt, C., Kumar, I., Vijayakumar, V., Singh, K.U. and Kumar, A., 2021. The state of the art of deep learning models in medical science and their challenges. *Multimedia Systems*, 27(4), pp.599-613.
- [48] Mosavi, A., Ardabili, S. and Varkonyi-Koczy, A.R., 2019, September. List of deep learning models. In *International conference on global research and education* (pp. 202-214). Cham: Springer International Publishing.
- [49] Obaid, K.B., Zeebaree, S. and Ahmed, O.M., 2020. Deep learning models based on image classification: a review. *International Journal of Science and Business*, 4(11), pp.75-81.

- [50] Obaid, K.B., Zeebaree, S. and Ahmed, O.M., 2020. Deep learning models based on image classification: a review. *International Journal of Science and Business*, 4(11), pp.75-81.
- [51] Obaid, K.B., Zeebaree, S. and Ahmed, O.M., 2020. Deep learning models based on image classification: a review. *International Journal of Science and Business*, 4(11), pp.75-81.
- [52] Tamuly, S., Jyotsna, C. and Amudha, J., 2019, September. Deep learning model for image classification. In *International Conference On Computational Vision and Bio Inspired Computing* (pp. 312-320). Cham: Springer International Publishing.
- [53] Lv, Q., Zhang, S. and Wang, Y., 2022. Deep learning model of image classification using machine learning. *advances in multimedia*, 2022(1), p.3351256.
- [54] Çınar, Z.M., Abdussalam Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M. and Safaei, B., 2020. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability*, 12(19), p.8211.
- [55] Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., and Adda, M. 2022. On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges. *Applied Sciences*, 12(16), 8081.
- [56] Kalathas, I. and Papoutsidakis, M., 2021. Predictive maintenance using machine learning and data mining: A pioneer method implemented to greek railways. *Designs*, 5(1), p.5.

