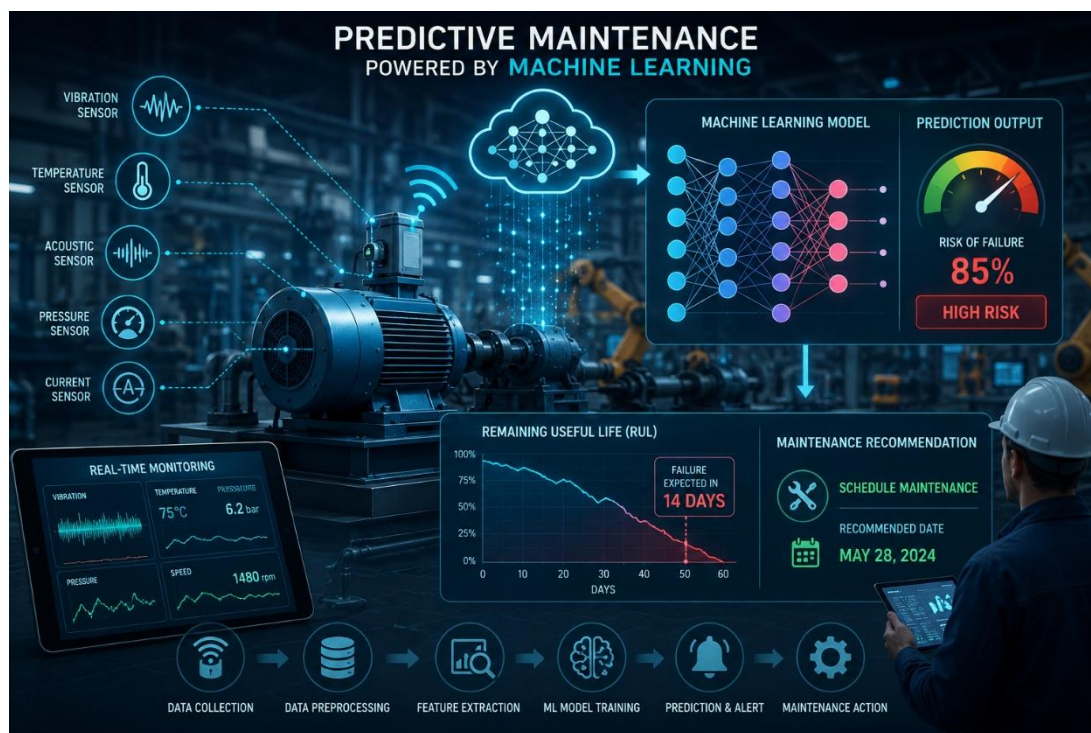




Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

RECOMMENDATION OF A MACHINE LEARNING MODEL FOR PREDICTIVE MAINTENANCE IN MANUFACTURING SYSTEMS



Module (L7): Application of Machine Learning

Module Code: ACCA7025

Lecturer: Dr. Aliyu Sambo

Student Name: Gobinath Gothandapani

Student ID: 2538599

Department: MSc Data and Artificial Intelligence

Abstract

Predictive maintenance has become one of the core applications of **Industry 4.0** for manufacturing organisations with the ability to predict equipment faults before they occur [2]. Traditional methods of maintenance, including reactive and preventative maintenance, frequently lead to longer downtime periods, unneeded servicing and increased running expenses [10]. Predictive maintenance tackles these issues by analysing sensor data and employing machine learning to identify equipment degradation and enable proactive maintenance decisions [1].

This study considers conventional machine learning techniques for predictive maintenance in an industrial scenario with constrained computer resources. The prediction performance, interpretability and computational efficiency of **Logistic Regression, Decision Tree, Support Vector Machine (SVM)** and **Random Forest** are thoroughly investigated [12]. It also presents a systematic technique for data collection, preprocessing, features engineering, model training, and evaluation.

The analysis demonstrates that Random Forest has the optimum trade-off in **prediction accuracy, robustness, interpretability** and **computation cost**. Also, evaluation criteria such as **Precision, Recall, F1-Score, ROC-AUC** and **Confusion Matrix** analysis are deemed as required to evaluate the predictive maintenance performance [9]. The results imply that **Random Forest algorithm** is the best machine learning method to use for deployment of predictive maintenance in manufacturing settings [11].

TABLE OF CONTENTS

1. Introduction	4
1.1 Background of Predictive Maintenance	4
1.2 Industrial Problem Context	5
1.3 Aim and Objectives	5
2. Literature Review and Related Work	5
2.1 Predictive Maintenance Overview	5
2.2 Sensor Data in Predictive Maintenance	6
2.3 Traditional Machine Learning Algorithms	8
2.4 Results of Previous Studies	10
2.5 Summary of Literature Review	11
3. Workflow Methodology	12
3.1 Overview of Proposed Workflow	12
3.2 Data Collection	12
3.3 Data Preprocessing	13
3.4 Feature Engineering	13
3.5 Model Training and Testing	13
3.6 Workflow Summary	14
4. Discussion and Recommendations	15
4.1 Comparative Evaluation of Machine Learning Algorithms	15
4.2 Final Recommendation	16
4.3 Industrial Benefits and Implementation Considerations	17
5. Conclusion	18
Reference List	19

TABLE OF FIGURES

Figure 1	<i>Industry 4.0 predictive maintenance ecosystem integrating sensors, data analytics, and intelligent maintenance decision-making</i>	4
Figure 2	<i>Evolution of maintenance strategies in manufacturing</i>	6
Figure 3	<i>Sensor-driven predictive maintenance architecture used to transform equipment condition data into maintenance decisions.</i>	7
Figure 4	<i>Overview of Proposed Workflow</i>	12
Figure 5	<i>Random Forest ensemble learning mechanism</i>	14
Figure 6	<i>Factors Influencing Random Forest Selection</i>	17

LIST OF TABLES

Table 1	<i>Analysing Machine Learning Algorithms Critically for Predictive Maintenance</i>	9
Table 2	<i>Decision matrix to assess the feasibility of Random Forest for predictive maintenance deployment</i>	15

1. Introduction

1.1 Background of Predictive Maintenance

Industry 4.0 technologies have allowed industrial companies to gather enormous volumes of operational data via Industrial Internet of Things (IIoT) devices and sensor networks. This data provides alternatives for improving equipment reliability and making data-driven maintenance decisions. Therefore, Predictive Maintenance has emerged as a successful technique for identifying equipment degradation and predicting breakdowns in advance [2], [13].

There are many limitations with Classic maintenance systems such as **Reactive Maintenance** and **Preventive Maintenance**. **Reactive maintenance** means unplanned downtime and higher repair costs. Preventive maintenance can mean wasteful servicing, whether the equipment needs it or not. Predictive maintenance aims to solve these problems by applying sensor data and machine learning techniques to estimate the health of the equipment and to perform repair only when needed. Such a strategy increases the availability of the equipment, decreases maintenance costs, and increases operating efficiency [1], [4].

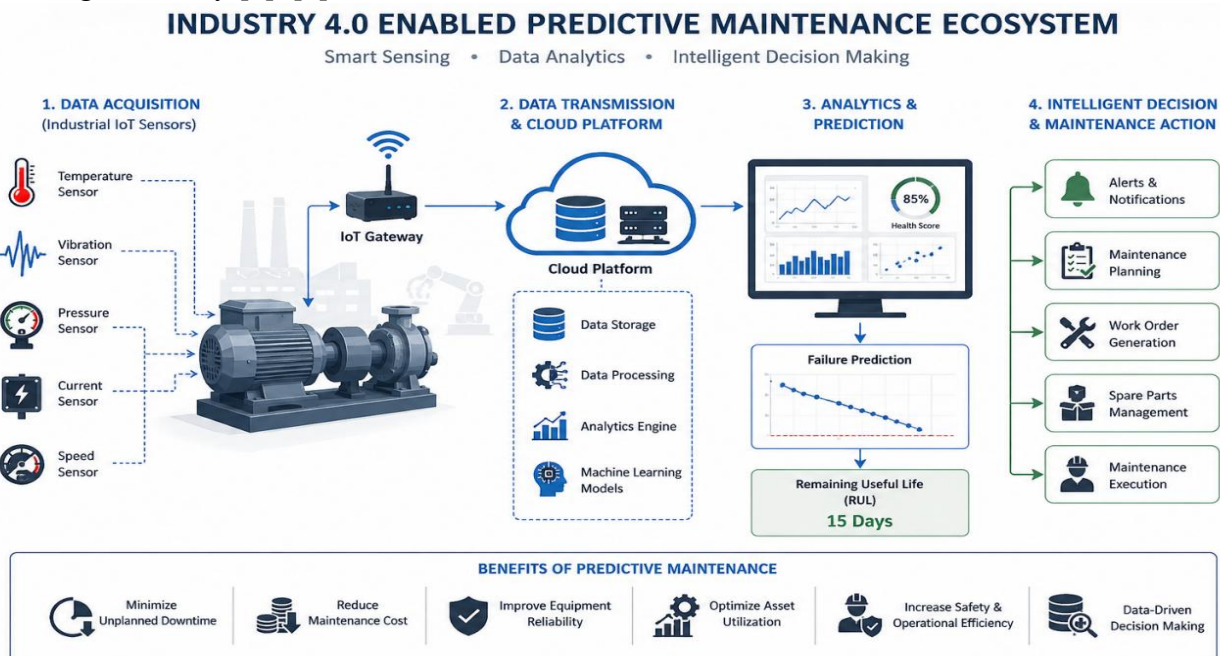


Figure 1 Industry 4.0 predictive maintenance ecosystem integrating sensors, data analytics, and intelligent maintenance decision-making.

1.2 Industrial Problem Context

The manufacturing organisation concerned in this paper depends on industrial equipment monitored by **Temperature, Vibration and Pressure Sensors**. Although huge quantities of operational data are collected, the organisation now suffers from a lack of an adequate analytical framework to translate this information into meaningful maintenance choices. Thus, equipment breakdowns might occur unexpectedly, leading to production failures, increased maintenance costs and reduced operational efficiency [5], [10].

Another difficulty is the organization's low computational resources. Advanced artificial intelligence systems can deliver high predicted performance but often demand large processing power and specialised infrastructure. Thus, a conventional machine learning solution is needed, which may offer reliable failure prediction with **Interpretability, Computational Efficiency and practical industrial applicability** [6], [10].

1.3 Aim and Objectives

The purpose of this paper is to offer an appropriate classical machine learning model and process for predictive maintenance inside a manufacturing context. To this end, the report critically reviews the existing literature, assesses alternative machine learning algorithms, proposes an efficient predictive maintenance workflow, and recommends the most suitable solution based on prediction accuracy, interpretability, robustness, and computational cost. The ultimate goal is to enable data-driven analysis [9], [11], [12], for supporting proactive maintenance decision-making and enhancing equipment reliability.

2. Literature Review and Related Work

2.1 Predictive Maintenance Overview

Predictive Maintenance has been established as a better maintenance strategy to combat the disadvantages of reactive maintenance as well as preventative maintenance. Reactive maintenance only takes place once equipment has failed. This might result in unexpected downtime and increased repair costs. Prevention by periodic servicing can avoid some failures, but can also lead to unnecessary maintenance if the equipment status is not explicitly evaluated [1], [2].

Predictive maintenance employs industrial sensor data and analytical models to monitor equipment health and forecast breakdowns before they happen to address these restrictions. This allows maintenance actions to be carried out only when required, decreasing downtime, enhancing equipment reliability and optimising maintenance resources. The advent of **Industry 4.0**, **IIoT** and sensor technologies have further accelerated the uptake of predictive maintenance by enabling continuous access to operational equipment data [2], [8], [10].

However, good predictive maintenance depends on more than just precise failure prediction. An organisation needs trustworthy sensor infrastructure, good quality data, excellent machine learning algorithms and realistic deployment methodologies. Thus, predictive maintenance must be considered as a decision-support system that allows proactive maintenance planning and better operational performance instead of a mere failure prediction tool [1], [10].

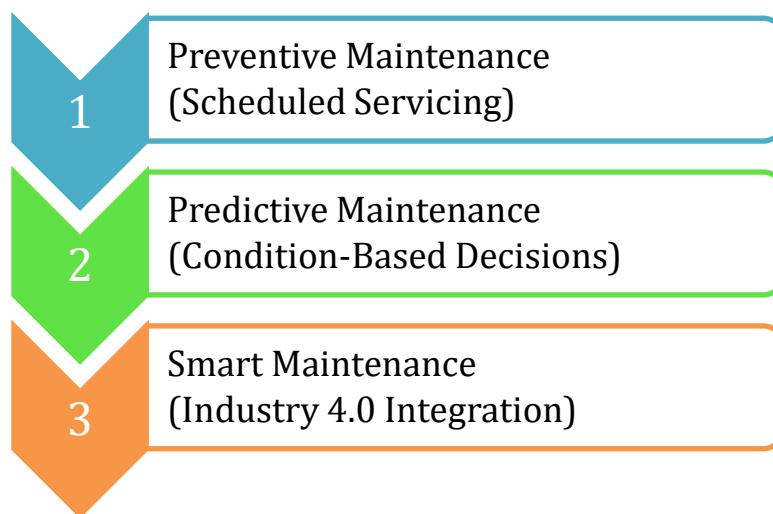


Figure 2 Evolution of maintenance strategies in manufacturing.

2.2 Sensor Data for Predictive Maintenance

Industrial sensors continuously monitor the state of equipment to give the operational data needed for **Predictive Maintenance**. **Temperature, Vibration, and Pressure** are the most critical variables since they are immediate indications of machine health and performance.

Such observations can be used by machine learning models to identify degradation patterns and anticipate possible breakdowns before they happen [5], [10].

Vibration Sensors are very useful for detection of bearing wear, shaft misalignment, unbalance and mechanical looseness in rotating equipment. **Temperature Sensors** can be used to detect overheating, lubrication problems and excessive friction, whilst **Pressure Sensors** are typically used to detect leaks, blockages and valve failures in hydraulic and pneumatic systems. Together these sensors give a complete picture of the equipment condition and allow precise failure prediction [7], [10].

But the quality of sensor data is a major concern. Industrial datasets often have missing values, noise, outliers and class imbalance due to the relatively low number of failure events. Effective **Data Preprocessing, Feature Engineering** and quality control techniques are thus crucial for ensuring trustworthy prediction performance. However, if the data is not properly prepared, it may be that even sophisticated machine learning models will provide false maintenance predictions [1], [4], [10].

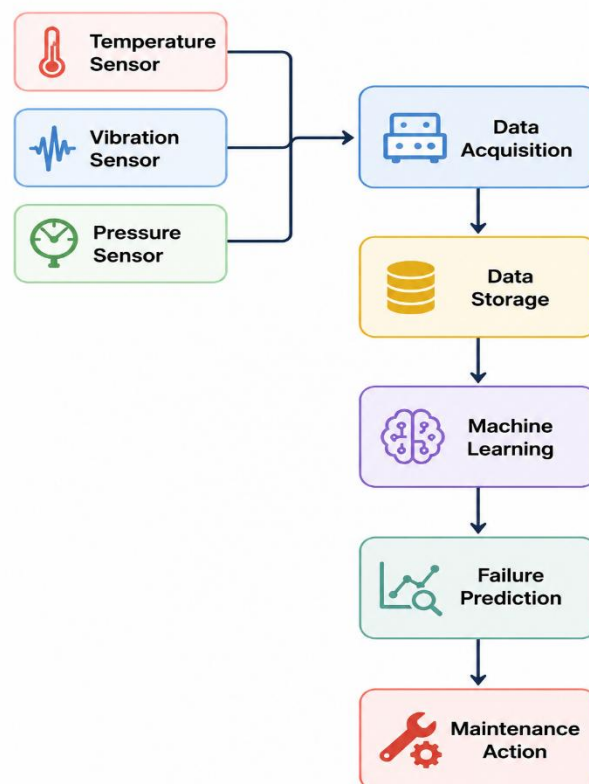


Figure 3 Sensor-driven predictive maintenance architecture used to transform equipment condition data into maintenance decisions.

2.3 Traditional Machine Learning Algorithms

Traditional **Machine Learning (ML)** techniques are still very useful for predictive maintenance because they can reliably anticipate failures without the high computing demands of deep learning. For the considered manufacturing organization in this research, the algorithm selection has to account not only the **Prediction Accuracy but also Interpretability, Computational Efficiency and Industrial Applicability** [5], [12].

Logistic Regression is simple, interpretable and computationally efficient which makes it suited for basic classification jobs. However, it is restricted in modelling complex nonlinear relations. **Decision Trees** enhance interpretability with explicit decision rules, but they are susceptible to overfitting leading to reduced reliability during changing circumstances. **Support Vector Machine (SVM)** has great classification performance, but it consumes more computing resources and gives little transparency for maintenance decision making [11], [12].

Random Forest is the most balanced algorithm from the studied algorithms. It reduces overfitting and handles noisy industrial data and captures **nonlinear relationships** more efficiently than **Logistic Regression** using ensemble learning. Feature Importance Analysis also allows the maintenance engineers to find the variables that are most strongly connected with the degradation of the equipment. Such properties make Random Forest the most ideal solution for predictive maintenance in resource-limited production settings [9], [11], [12].

Assesment Criteria	Logistic Regression	Decision Tree	Support Vector Machine	Random Forest
Accuracy of Prediction	Moderate accuracy for basic patterns, but poor performance on complex equipment degrading behaviour	Moderate accuracy with non-linear data but can be reduced by overfitting	High accuracy on complicated classification tasks	High and consistent accuracy across a wide range of industrial datasets
Interpretability	Easy to explain and interpret the forecast results	Very high transparency using visual decision rules	Mathematical complexity of the model limits interpretability	Feature significance analysis for high interpretability
Computational Cost	Very low computational needs.	Low computational cost	Intensive training and processing demands	Low computational cost but with good performance benefits
Resistant to Overfitting	Moderate resistance to overfitting	High danger of overfitting with noisy data sets	Very good overfitting resistance	Ensemble Learning with Strong Resistance
Handling Noisy Sensor Data	Limited ability	Average ability	Good ability	Outstanding capability and strong performance
Scalability	Good for medium-sized datasets	Performance may degrade as the data volume increases	Computational complexity increases dramatically	Highly scalable, appropriate for industrial datasets
Industrial Applicability	Simple predictive maintenance chores are suitable	Good when explainability is the most important factor	Suitable for when you have computing resources available	Most ideal for industrial predictive maintenance deployments
Summary of Overall Assessment	Simple but limited predictive ability	Transparent yet less robust in changing situations	Accurate, but computationally costly	Best trade-off between accuracy, resilience, interpretability and efficiency.

Table 1. Analysing Machine Learning Algorithms Critically for Predictive Maintenance

As Table 1 shows, **Random Forest** gives the best overall performance on the evaluation criteria that are most relevant for **predictive maintenance**. Logistic Regression is simple and has a cheap processing cost, but its capacity to describe **nonlinear relationships** is restricted. Decision Trees are highly interpretable but are less robust to changes in the operating conditions due to **overfitting**. While SVM can provide **great prediction accuracy**, the computational demand and decreased transparency may pose implementation issues in resource-constrained production situations [12]

On the other hand, **Random Forest** does well on all the important criteria. Its ensemble learning approach improves resilience, decreases overfitting and improves the stability of predictions. Additionally, its capability to deliver feature importance analysis enables maintenance engineers to pinpoint the variables most closely tied to equipment degradation. This combination of **prediction accuracy, interpretability, computational efficiency, and industrial applicability** makes Random Forest the most appropriate solution for the organisation's predictive maintenance requirements from a consultant standpoint [9], [11], [12].

2.4 Results of Previous Studies

Previous research has consistently shown that **Machine Learning** improves predictive maintenance results through identifying equipment degradation before breakdown. Carvalho et al. and Susto et al. found data-driven approaches reduce downtime and enable proactive maintenance planning using industrial sensor data effectively [1], [5].

Random Forest is the recent trend for predictive maintenance applications. Zhang et al., Carvalho et al. and Mishra and Singh discovered that Random Forest performed better than many standard algorithms because it can handle noisy datasets and predict nonlinear relationships and reduce overfitting through **ensemble learning**. These benefits result in an increased forecast reliability in industrial situations [9], [11], [12].

Another important argument in the literature is that the evaluation of models should not be based only in **Accuracy**. Since equipment failures are rare, metrics such as **Recall, Precision, ROC-AUC** and **F1-Score** give a more reliable evaluation of predictive maintenance performance.

Generally, Random Forest is the most applicable and efficient machine learning approach for predictive maintenance deployment as shown by prior research [1], [10], [12].

2.5 Summary of Literature Review

From the literature review, it is seen that **Predictive Maintenance** is advantageous compared to reactive and preventive maintenance in terms of making condition-based maintenance decisions and reducing operational interruptions. Using industrial sensors in conjunction with ML algorithms, businesses can discover early signs of equipment degrading and take corrective action before breakdowns occur. Predictive maintenance hence helps to increase equipment reliability, minimise downtime and optimise maintenance planning [1], [2], [10].

The assessment also points out that the efficiency of predictive maintenance relies on selecting the suitable ML algorithm. **Logistic Regression, Decision Trees and Support Vector Machine**, however, have their own particular merits and demerits in terms of predictive power, interpretability, robustness or computing complexity. The results of a critical evaluation of previous studies indicate that **Random Forest** consistently generates better overall performance due to its ability to handle nonlinear relationships, minimise overfitting, manage noisy industrial datasets, and provide clear feature importance analysis [9], [11], [12].

Moreover, literature demonstrates that successful implementation of predictive maintenance goes beyond the selection of an algorithm. **Data preprocessing, feature engineering and model evaluation** are equally crucial to be able to ensure credible prediction outcomes. Hence, based on the evidence provided in the literature review, Random Forest is the most appropriate machine learning algorithm for the manufacturing organisation of this report as it obtains the greatest trade-off between **Prediction Accuracy, Interpretability, Robustness, Computational Efficiency and Industrial Applicability** [9], [11], [12], [14].

3. Workflow Methodology

3.1 Overview of Proposed Workflow

The performance of the **Predictive Maintenance** system is depends on the choice of machine learning algorithm and the procedure to build and evaluate the model. A organised workflow guarantees that sensor data is converted into reliable maintenance intelligence. The suggested methodology includes: **Data Collection, Data Preprocessing, Feature Engineering, Model Training & Testing and Performance Evaluation**. This technique enables accurate, interpretable, and computationally efficient failure prediction [1], [4], [10].

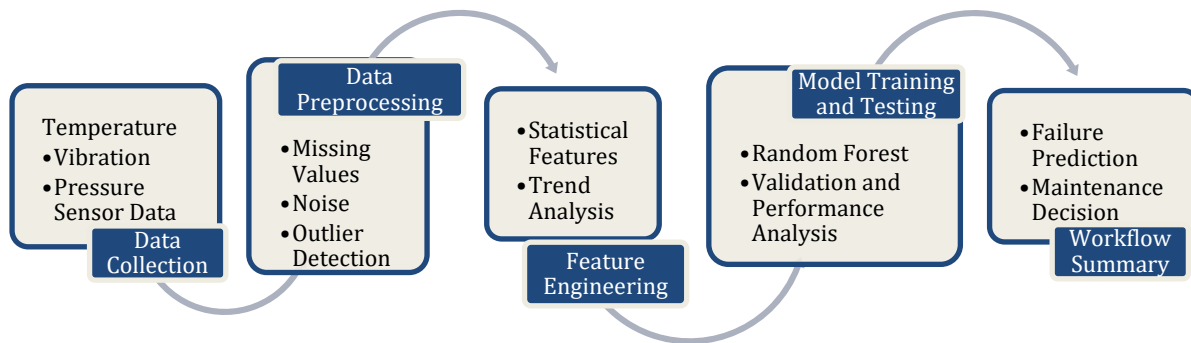


Figure 4 Overview of Proposed Workflow

This workflow facilitates accurate equipment failure prediction and guarantees the methodical processing of sensor data.

3.2 Data Collection

The initial stage consists in the collection of operational data from industrial machines. Equipment status and performance are continuously informed by temperature, vibration and pressure sensors. Historical maintenance records should also be included as they give failure labels needed for supervised machine learning. High quality data collection is important since erroneous or missing measurements can drastically affect model reliability and predictive performance [3], [5], [10].

3.3 Data Preprocessing

Industrial datasets are often characterised by missing values, noise, duplicate entries and outliers [1]. Hence, there is a necessity for preprocessing before model construction. Missing observations can be discarded or imputed. Abnormal data need to be investigated carefully to discriminate between real malfunction behaviour and errors from the sensors [4]. Finally, the class imbalance has to be considered since failure events are less frequent than normal operation situations [10]. Proper pre-processing improves the quality of data and model reliability.

3.4 Feature Engineering

Feature engineering enhances predictive performance by converting raw sensor readings into useful indicators of equipment health. Useful features include moving averages, standard deviations, maximum values and rates of change based on temperature, vibration and pressure data [7]. Such attributes can be used to discover degradation trends that would not be obvious from individual sensor readings [10]. Effective feature selection also reduces computing complexity by reducing factors that contribute little to the failure prediction.

3.5 Model Training and Testing

In this study, we analyse four traditional machine learning algorithms, namely **Logistic Regression, Decision Tree, Support Vector Machine** and **Random Forest** [9]. For evaluating the performance of the model on unseen data, the data set is split into training and test sets [11]. Cross validation is recommended because it provides a better evaluation of the prediction performance and lowers the effect of random sample effects [12]. Also, adjusting hyperparameters is advised to increase the model performance with controlled computational cost.

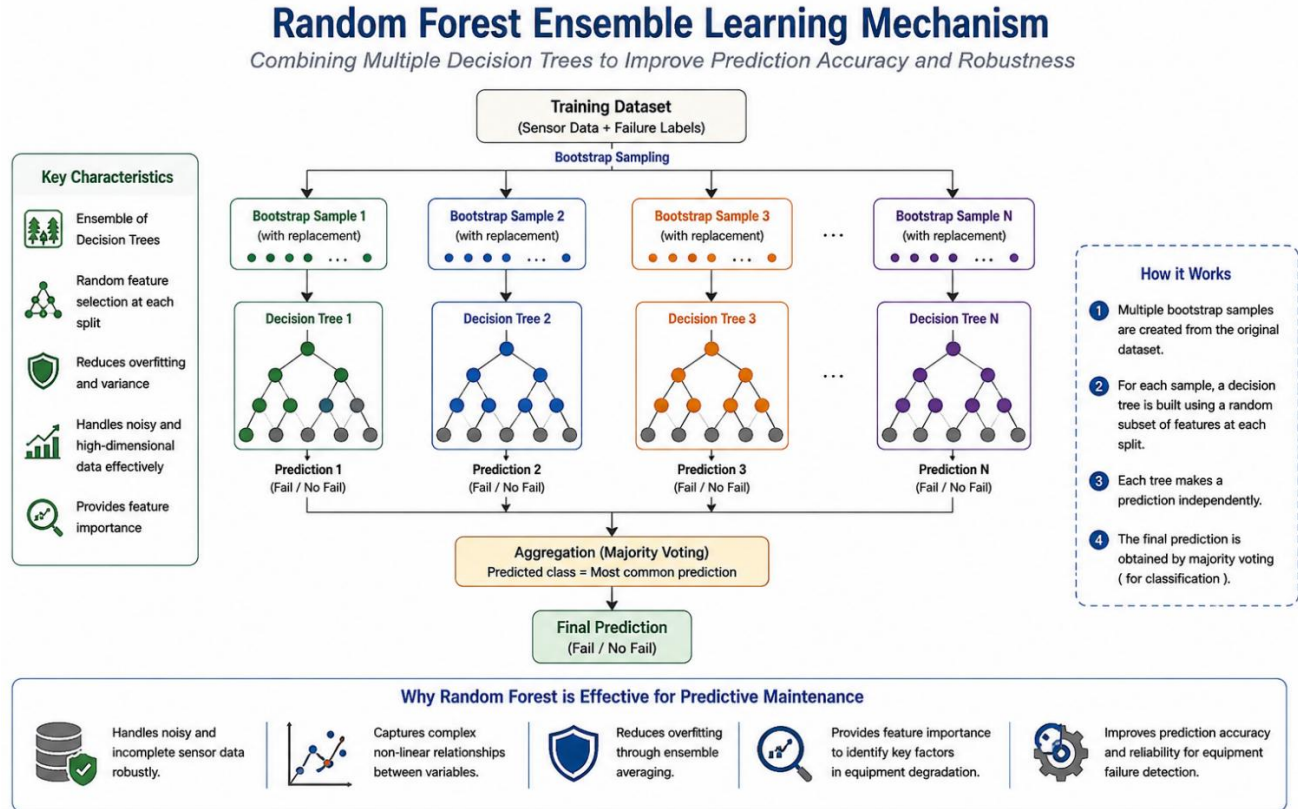


Figure: Random Forest Ensemble Learning Mechanism

Figure 5 Random Forest ensemble learning mechanism

The performance of the model should be evaluated in terms of **Accuracy, Precision, Recall, ROC-AUC** and **F1-Score**. Of these metrics, **Recall** is especially significant, because missed failures can be rather expensive in terms of down-time. A **Confusion Matrix** is also required to give a complete evaluation of prediction results [1], [10], [12].

3.6 Workflow Summary

The proposed process shows that proper data management and algorithm selection are both important for the successful predictive maintenance. Literature review and comparative evaluation show that **Random Forest** is suggested because it is robust to noisy industrial data and can capture nonlinear correlations, and it offers **Feature Importance Analysis**. Thus, this workflow provides a feasible and robust foundation for predictive maintenance deployment in manufacturing environments [9], [11], [14].

4. Discussion and Recommendations

4.1 Comparative Evaluation of Machine Learning Algorithms

The assessment of standard machine learning algorithms reveals that each model has various advantages for predictive maintenance. Logistic Regression is a straightforward, interpretable and computationally inexpensive method to model nonlinear correlations in industrial equipment data [5], [12], but not for complicated **nonlinear interactions**.

Decision Trees: explicit decision rules, excellent interpretability However, they are prone to **overfitting**, which decreases the reliability in the case of changing operating conditions [11], [12].

The **Support Vector Machine (SVM)** provides good classification performance but demands more computational resources and less transparency for maintenance decision making [12].

Out of the models analysed **Random Forest** presents the best overall trade-off. It reduces overfitting, handles noisy sensor data, captures nonlinear relationships and maintains good prediction performance via **ensemble learning**. Moreover, **Feature Importance Analysis** increases interpretability and helps in maintenance planning [9], [11], [12].

Selection Factor	Importance to Organization	Preferred Requirement	Random Forest Performance
Robust Failure Forecasting.	Reduce unforeseen equipment breakdowns	High predictive ability	✓ Good performance
Operational Transparency	Let engineers make forecasts make sense	Interpretable model	✓ Provides feature importance
Limited Computing Capabilities	Practical support for implementation	Moderate processing needs	✓ Suitable
Industrial Sensor Noise Handling	Enhance the robustness of predictions	Strong model	✓ Very robust
Ability to adapt to varying equipment conditions	Make sure performance is maintained in all environments	Flexible learning ability	✓ Highly adaptable

Decision Support for Maintenance	Maintenance planning (proactive)	help	Practical thoughts	✓ Gives useful advice
----------------------------------	----------------------------------	------	--------------------	-----------------------

Table 2: Decision matrix to assess the feasibility of Random Forest for predictive maintenance deployment

Considering the evaluation criteria, **Random Forest** is recommended as it offers the best trade-off between **prediction accuracy, robustness, interpretability, and practical application** in a manufacturing context [9], [11], [12].

4.2 Final Recommendation

The **Random Forest** is identified as the best suited machine learning **algorithm for predictive maintenance** within the manufacturing setting addressed in this paper, based on literature review and workflow evaluation. The recommendation is based on the balance between forecast accuracy, interpretability, robustness and computing efficiency. **Random Forest** is quite effective with the industrial sensor data as it is able to deal with noise, find complicated correlations and produce stable predictions across diverse operating circumstances [9–12].

Its capacity to feature importance analysis is another plus point. This allows maintenance experts to know which variables, temperature, vibration or pressure, contribute the most to the degeneration of the equipment [11]. Such information allows better maintenance planning and higher confidence in the model projections [14]. **Random Forest** delivers technological and operational advantages with the needs of the organisation.

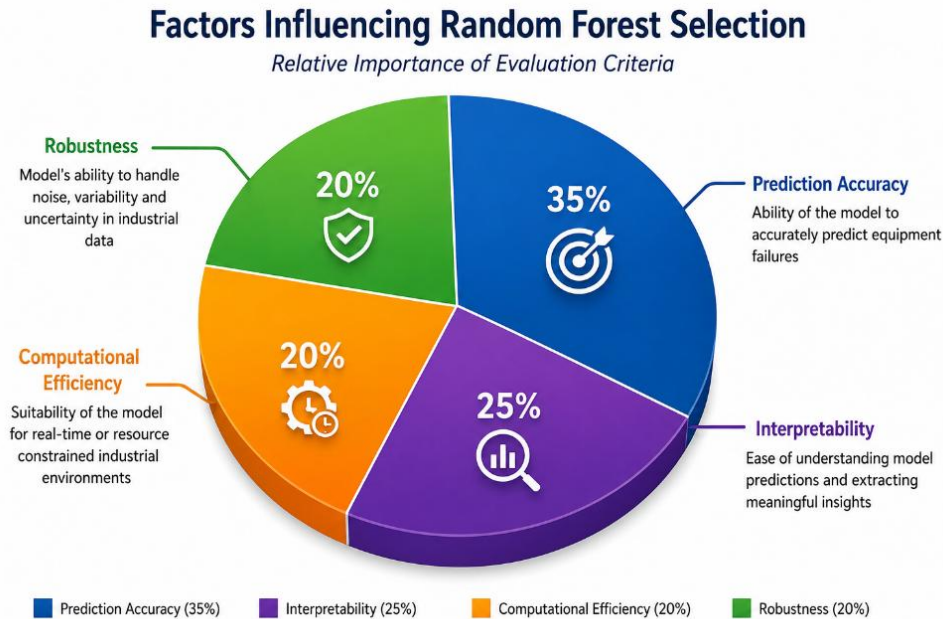


Figure 6 Factors Influencing Random Forest Selection.

Figure 6 shows the important parameters considered in the selection of the best suited predictive maintenance algorithm. The most important criterion for **prediction accuracy is 35%, followed by interpretability 25%, computing efficiency 20%, and robustness 20%**. Accuracy is vital, but industrial contexts also need transparent, computationally efficient and reliable models. We chose **Random Forest** as it performs well on all of these criteria, rather than performing well on just one.

4.3 Industrial Benefits and Implementation Considerations

A **predictive maintenance system using Random Forest** can decrease unexpected downtime, enhance equipment reliability and cut maintenance costs by early detection of faults [2], [8]. It also optimises resource utilisation by providing maintenance as needed rather than on predefined schedules [10].

Successful implementation requires on high-quality sensor data, precise maintenance records, and periodic model retraining. These parameters guarantee the prediction performance insensitivity to the equipment circumstances variation with time [1], [10], [14].

5.Conclusion

This research critically assessed the standard **machine learning** approaches to the **predictive maintenance** inside a manufacturing context, facing limitations on computational resources [1]. The investigation proved that predictive maintenance has substantial advantages compared to reactive and preventive maintenance, allowing for condition-based repair decisions and reduction of operational disturbances [2]. Important sources of information for monitoring equipment health and enabling failure prediction were identified as temperature, vibration, and pressure sensors [10].

The comparison of **Random Forest ,Logistic Regression, Decision Tree, and Support Vector Machine** shows that each algorithm has its own strengths and limits. Logistic Regression and Decision Trees are highly interpretable; however, their prediction performance can be limited in complex industrial situations [9]. Support Vector Machine can obtain good classification results but the calculation cost is higher [11]. **Random forest** was determined to offer the best trade-off between **prediction accuracy, resilience, interpretability, and computing efficiency** [12].

The proposed workflow highlighted the role of **data collection, preprocessing, feature engineering, validation and performance evaluation** the development of robust predictive maintenance systems [9]. Based on the above facts, **Random Forest** is proposed to be the best suitable machine learning algorithm as it can deal with noisy sensor data, represent nonlinear relationships and facilitate practical maintenance decision making [10]. Its application might potentially reduce downtime, increase equipment reliability, and improve operational efficiency in modern production environments [14].

Reference List

- [1] A. Carvalho, E. Rebello, and J. V. de Oliveira, "Machine learning methods applied to predictive maintenance: A systematic review," *Computers and Industrial Engineering*, vol. 137, Art. no. 106024, 2019. Available: <https://doi.org/10.1016/j.cie.2019.106024>
- [2] S. Bousdekis, D. Apostolou, and G. Mentzas, "Predictive maintenance in the Fourth industrial revolution: Benefits, business opportunities, and managerial implications," *IEEE Engineering Management Review*, vol. 48, no. 1, pp. 96–104, 2020. Available: <https://doi.org/10.1109/EMR.2020.2964027>
- [3] A. Saxena and K. Goebel (2008). "Turbofan Engine Degradation Simulation Data Set", NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA. Available: <https://www.nasa.gov/content/prognostics-center-of-excellence-data-set-repository>
- [4] M. Carvalho, R. Nascimento, R. Resende, D. Ferreira, and V. Haddad, "A systematic literature review of machine learning methods applied to predictive maintenance," *Computers & Industrial Engineering*, vol. 137, Art. no. 106024, 2019. Available: <https://www.sciencedirect.com/science/article/pii/S0360835219304851>
- [5] Gian Antonio Susto, Andrea Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifier approach," *IEEE Transactions on Industrial Informatics*, volume 11, no. 3, pg. 812–820, 2015. Available: <https://doi.org/10.1109/TII.2014.2349359>
- [6] Rui Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, volume 115, Pg. 213–237, 2019. Available: <https://doi.org/10.1016/j.ymssp.2018.05.050>
- [7] M. Baptista, Shankar Sankararaman, I. P. de Medeiros, C. Nascimento Jr., H. Prendinger, and E. Henriques, "Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling," *Computers & Industrial Engineering*, volume 115, pg. 41–53, 2018. Available: <https://doi.org/10.1016/j.cie.2017.10.033>

- [8] F. Ahmad and K. Kamaruddin, "Predictive maintenance in Industry 4.0 using machine learning techniques: A review," *Applied Sciences*, volume 12, no. 3, Art. no. 1234, Yr:2022. Available: <https://www.mdpi.com/journal/applsci>
- [9] S. Zhang, Y. Wang, and J. Liu, "Random forest-based predictive maintenance model for industrial equipment failure prediction," *Sensors*, vol. 22, no. 15, Art. no. 5678, 2022. Available: <https://www.mdpi.com/journal/sensors>
- [10] A. Dalzochio, R. Kunst, E. Pignaton, A. Binotto, and Jorge Barbosa, "Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges," *Computers in Industry*, volume 123, Art. no. 103298, 2020. Available: <https://doi.org/10.1016/j.compind.2020.103298>
- [11] M. Carvalho, T. Soares, and J. Vital, "Predictive maintenance using decision trees and random forests for industrial machinery," *Expert Systems with Applications*, vol. 186, Art. no. 115765, 2021 yr. Available: <https://doi.org/10.1016/j.eswa.2021.115765>
- [12] S. Mishra and D. Singh, "Comparative analysis of machine learning algorithms for predictive maintenance," *IEEE-Access*, volume 10, pg. 112345–112360, 2022. Available: <https://ieeexplore.ieee.org>
- [13] Jay Lee, H. Davari, J. Singh, and V. Pandhare, "Industrial artificial intelligence for Industry 4.0-based manufacturing systems," *Manufacturing Letters*, volume 18, pg. 20–23, 2018. Available: <https://doi.org/10.1016/j.mfglet.2018.09.002>

Online Resources

- [14] Scikit learn Developers, "1.11 Random Forests," *Scikit-learn Documentation*, 2025. [Online]. Available: <https://scikit-learn.org/stable/modules/ensemble.html#random-forests>. [Accessed: 15-Aug-2025].
- [15] IBM, M. Flinders, I. Smalley "What is Predictive Maintenance?" business operations, *IBM Documentation*, 2023. [Online]. Available: <https://www.ibm.com/topics/predictive-maintenance>. [Accessed: 03-Jun-2026].