

Prediction of Milling Machine Failure Using a Random Forest Machine Learning Model

Olaniyi Emmanuel Ogundipe¹ and Sunday Emmanuel Kayode^{2*}

¹Department of Mechanical Engineering, Federal University Oye-Ekiti, Nigeria

^{2*}Department of Agricultural and Environmental Engineering, Federal University of Technology, Akure, Nigeria

Abstract

Machine failures in manufacturing industries result in significant production losses, increased maintenance costs, and reduced equipment reliability. Predictive maintenance has emerged as an effective strategy for minimizing unplanned downtime by forecasting equipment failures before they occur. This study developed a machine learning-based predictive maintenance model for milling machine failure prediction using a Random Forest classifier. Operational data comprising air temperature, process temperature, rotational speed, torque, and tool wear were utilized as input variables. A dataset containing 10,000 observations was obtained and divided into training (70%) and testing (30%) subsets. The Random Forest model was trained using 100 decision tree estimators and evaluated using accuracy, precision, recall, F1-score, and confusion matrix metrics. The model achieved training accuracy, precision, recall, and F1-score values of 99.34%, 84.03%, 96.15%, and 89.69%, respectively. Testing results yielded accuracy, precision, recall, and F1-score values of 98.53%, 71.29%, 82.76%, and 76.60%, respectively. Confusion matrix analysis showed that the model correctly classified 2,884 non-failure cases and 72 failure cases. The results indicate that Random Forest is highly effective for milling machine failure prediction and can support proactive maintenance decisions, thereby reduce downtime and improve operational efficiency in manufacturing environments.

Keywords: Predictive Maintenance, Milling Machine Failure, Machine Learning, Random Forest, Condition Monitoring, Manufacturing Systems.

1. Introduction

1.1 Background and Context

Milling machines are among the most widely used machine tools in manufacturing industries due to their ability to produce components with high dimensional accuracy and surface quality. Continuous operation of these machines is essential for maintaining production efficiency and product quality. However, milling machines are vulnerable to failures caused by factors such as excessive tool wear, spindle defects, bearing degradation, overheating, and electrical faults [1]

Traditional maintenance strategies, including corrective and preventive maintenance, are often inadequate because they either respond after failure occurs or rely on fixed maintenance schedules that may not reflect the actual condition of the machine. Such approaches often lead to unnecessary maintenance actions or costly unplanned downtime [2], [3]. Predictive maintenance has gained considerable attention as an advanced maintenance strategy that utilizes operational data to predict equipment failures before they occur. The integration of machine learning techniques into predictive maintenance systems has significantly enhanced fault diagnosis and failure prediction capabilities. Machine learning algorithms can analyse large datasets, identify hidden patterns, and establish relationships between machine operating conditions and failure events [3].

Recent advancements in sensor technologies and industrial data acquisition systems have facilitated the collection of real-time machine condition data. These developments have enabled the deployment of intelligent predictive maintenance systems capable of supporting proactive maintenance decisions [4].

This study focuses on the development of a Random Forest-based predictive model for milling machine failure prediction using operational parameters including air temperature, process temperature, rotational speed, torque, and tool wear. The aim is to improve equipment reliability, reduce downtime, and optimize maintenance scheduling in manufacturing systems.

2. Materials and Methods

2.1 Data Collection

The dataset used in this study was obtained from milling machine operations at the Project Development Institute (PRODA), Enugu State, Nigeria. Five operational parameters were selected as predictor variables:

- i. Air temperature (K)
- ii. Process temperature (K)
- iii. Rotational speed (rpm)
- iv. Torque (Nm)
- v. Tool wear (minutes)

The dataset consisted of 10,000 observations representing different operating conditions and corresponding machine failure states.

2.2 Data Preparation

Data preprocessing was performed to improve model performance and reliability. The preprocessing stage involved:

- i. Data cleaning

- ii. Removal of inconsistent records
- iii. Handling missing values
- iv. Feature normalization
- v. Data partitioning

The dataset was divided into training and testing subsets, with 70% used for model training and 30% used for testing.

2.3 Random Forest Model Development

Random Forest is an ensemble learning algorithm that combines multiple decision trees to improve prediction accuracy and reduce overfitting. Each tree is trained using bootstrap sampling, and final predictions are obtained through majority voting (figure 1 and 2)

The Random Forest classifier was implemented using 100 estimators. The algorithm was trained using the selected operational parameters to predict machine failure conditions.

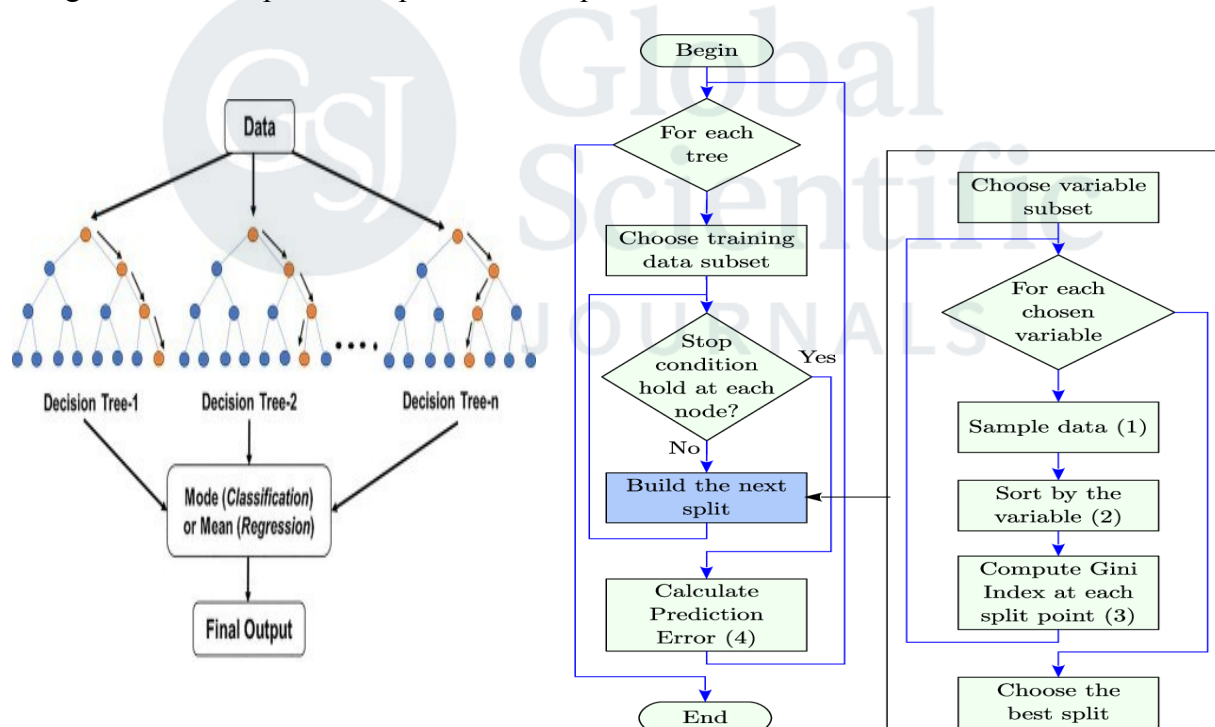


Figure 1: Random Decision. Forest
Figure 2:Flow chart of random forest algorithm.

2.4 Performance Evaluation

Model performance was evaluated using:

- i. **Confusion Matrix:** A confusion matrix is a classification performance evaluation tool that summarizes a model's predictions using four outcomes: true positives

(correctly predicted positive cases), true negatives (correctly predicted negative cases), false positives (negative cases incorrectly predicted as positive), and false negatives (positive cases incorrectly predicted as negative). It provides a comprehensive assessment of how well a classification model performs.

ii. Accuracy

Accuracy is a performance metric used to measure the percentage of correctly classified observations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

iii. Precision

Precision measures the proportion of true positives among the positive predictions. It is gotten by dividing the number of true positives by the sum of true positives and false positives.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

iv. Recall

Recall, also called "True Positive Rate" or "Sensitivity" measures the proportion of true positives among the actual positives

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

v. F1-Score

The F1-score is a metric that combines precision and recall in to a single score.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where:

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

2.5 Data Visualization

The dataset consists of 10,000 observations with binary target outputs, where 0 represents no machine failure and 1 represents machine failure. The distribution of these target values is presented in Figure 3. Figure 4 further categorizes the failure instances into specific failure types, including tool wear failure, overstrain failure, heat dissipation failure, and no failure, providing insight into the major causes of machine failure within the dataset.

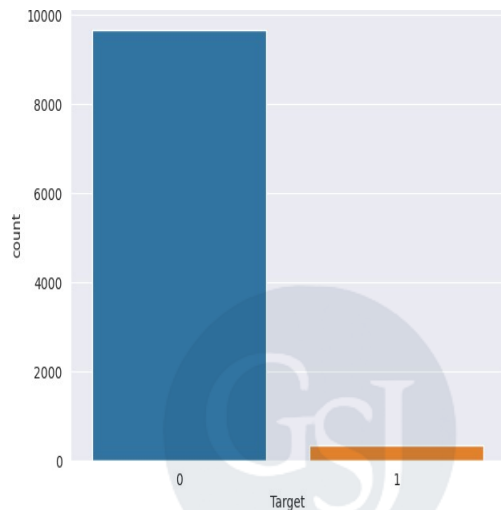


Figure 3: Target values

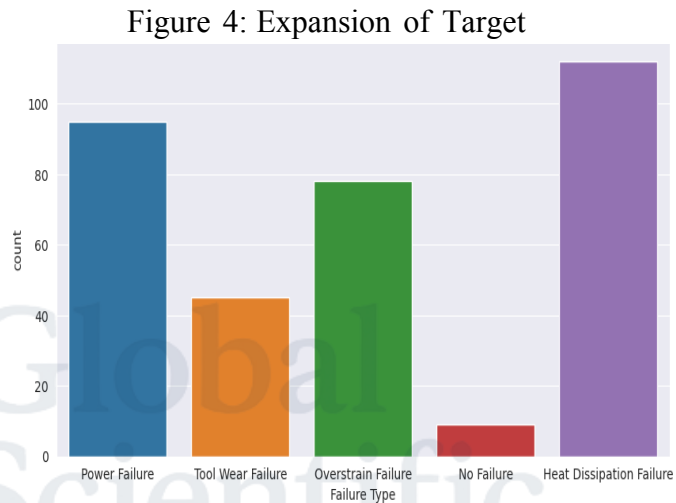


Figure 4: Expansion of Target

value one on causes of power failure.

3. Results and Discussion

3.1 Model Performance

Table 1 presents the performance metrics of the developed Random Forest model.

Table 1. Performance Metrics of Random Forest Model

Model	Accuracy	Precision	Recall	F1
Training	0.99342	0.84033		0.96153
	0.89686			
Testing	0.98533	0.71287		0.82758
	0.76595			

The model achieved excellent performance during both training and testing phases, indicating strong generalization capability and low overfitting.

3.2 Confusion Matrix Analysis

A confusion matrix was built using our validation data as a proxy of unseen data to evaluate how well our model has learned in classifying machine failure and no machine failure from our data-set. Figure 5 shows the confusion matrix with true label and predicted label. True label signifies the actual target of our data-set while predicted label depicts the model target output. By summing the roll, we have 2899 no machine failure target and 101 machine failure target. However, of the 2899 no machine failure target, our model correctly predicted 2884 no machine failure as true positive (TP) and 15 of no machine failure as false positive (FP). On the other hand, the model correctly predicted 72 of machine failure target as true negative (TN) out of 101 machine failure target and 29 as false negative (FN).

The obtained results are comparable to those reported by [5], who applied explainable artificial intelligence techniques to the AI4I 2020 Predictive Maintenance Dataset and observed that machine learning models achieved very high accuracy in distinguishing healthy machines from faulty machines. Similar to the present study, the majority of non-failure cases were correctly classified, demonstrating the ability of ensemble learning algorithms to capture the complex relationships among operational variables and machine health conditions.

The confusion matrix also reveals that the model produced only 15 false positives out of 2,899 healthy machine instances, indicating excellent specificity. This finding agrees with the work of [6], who reported that Random Forest classifiers generally exhibit high specificity and robustness in predictive maintenance applications because they aggregate predictions from multiple decision trees, thereby reducing classification errors and overfitting.

However, the model recorded 29 false negatives among the failure cases, indicating that some actual failures were classified as healthy conditions. This observation is consistent with findings reported by [7], who noted that predictive maintenance datasets are often highly imbalanced, causing machine learning models to favour the majority class and making failure detection more challenging. Similar studies have shown that although Random Forest models achieve excellent overall accuracy, their sensitivity toward rare failure events is usually lower than their ability to identify normal operating conditions.

Compared with other machine learning approaches, the performance achieved in this study is highly competitive. [8] reported that Random Forest models often outperform traditional classification techniques such as Decision Trees and Logistic Regression due to their superior generalisation capability and resistance to noise. Likewise, [13] found that ensemble-based algorithms consistently achieve high predictive accuracy in machine condition monitoring and fault diagnosis tasks.

The confusion matrix results further demonstrate that the developed model possesses strong predictive capability for practical predictive maintenance applications. The high number of correctly classified non-failure cases and the satisfactory detection of failure cases suggest that the model can effectively support maintenance decision-making by identifying machines that are likely to fail while minimising false alarms. These findings are consistent with previous predictive

maintenance studies that identified Random Forest as one of the most reliable machine learning algorithms for machine failure prediction.

The confusion matrix analysis confirms that the developed model performs comparably with state-of-the-art predictive maintenance models reported in the literature, achieving high classification accuracy while maintaining a relatively low number of misclassifications.

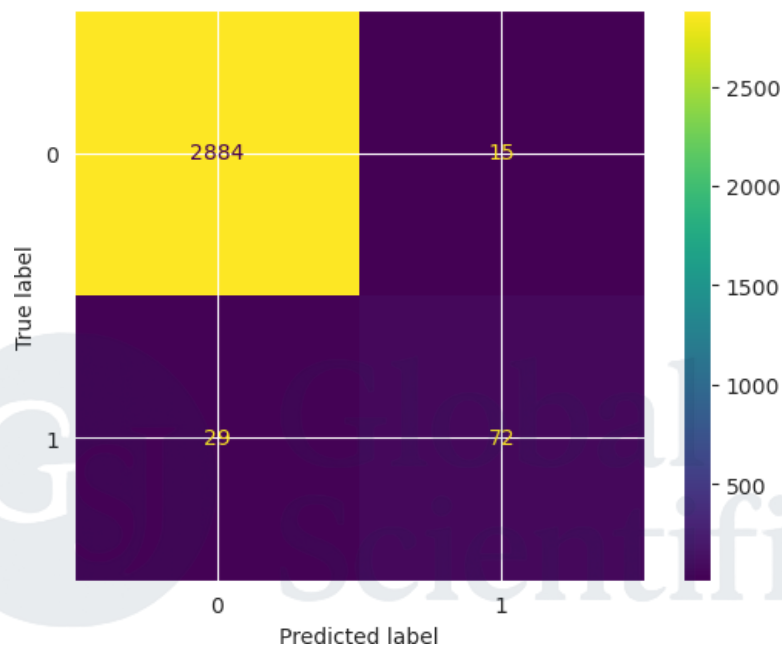


Figure 5: Confusion Matrix

3.4 Possibility of machine failure with respect to input parameter

Figure 6 below illustrates the influence of the five operational parameters on machine failure probability. The results indicate that machine failure probability increases as air temperature rises from 296 K to 302 K, decreases slightly around 304 K, and then increases substantially beyond 304 K. This observation agrees with the findings of [8]B, who reported that elevated operating temperatures are among the major indicators of machine degradation and impending failure because excessive heat accelerates wear and component deterioration.

The effect of process temperature on machine failure was relatively insignificant below 313 K; however, the probability of failure increased once this threshold was exceeded. Similar observations were reported by [5], who noted that process temperature becomes a critical predictor only when it deviates substantially from the normal operating range. This suggests that machines can tolerate moderate temperature variations, whereas extreme temperatures contribute significantly to fault occurrence.

The analysis further revealed that increasing rotational speed increased the probability of machine failure, with speeds above 2600 rpm resulting in a high likelihood of failure. This finding is

consistent with the work of [6], who demonstrated that excessive rotational speed increases mechanical stress, vibration, and frictional forces, thereby accelerating equipment deterioration and increasing failure rates. Similar trends were also observed by [9] in predictive maintenance studies involving rotating machinery. Regarding torque, values between 10 Nm and 40 Nm had minimal influence on machine failure probability, whereas torque values exceeding 50 Nm significantly increased failure likelihood, leading to complete failure at approximately 70 Nm. This result corroborates the findings of [10], who established that excessive torque loading is strongly associated with accelerated wear and mechanical breakdown in industrial equipment. The observed threshold behaviour suggests that machines can operate efficiently within a specified torque range but become increasingly vulnerable when subjected to overload conditions.

Tool wear exhibited little influence on machine failure probability between 0 and 150 minutes. However, the probability increased markedly between 200 and 250 minutes, with failure occurring almost certainly at tool wear values above 250 minutes. This finding is in agreement with studies by [11] and [12], which identified tool wear as one of the most significant predictors of machine failure. Their studies showed that prolonged tool usage results in progressive degradation of cutting performance, increased energy consumption, and eventual machine malfunction. Overall, the observed relationships between machine failure probability and the selected operational parameters are consistent with previous predictive maintenance studies, which have identified temperature, rotational speed, torque, and tool wear as critical indicators of machine health. The results further demonstrate the capability of machine learning models to capture nonlinear interactions among these variables and accurately predict machine failure before catastrophic breakdown occurs.

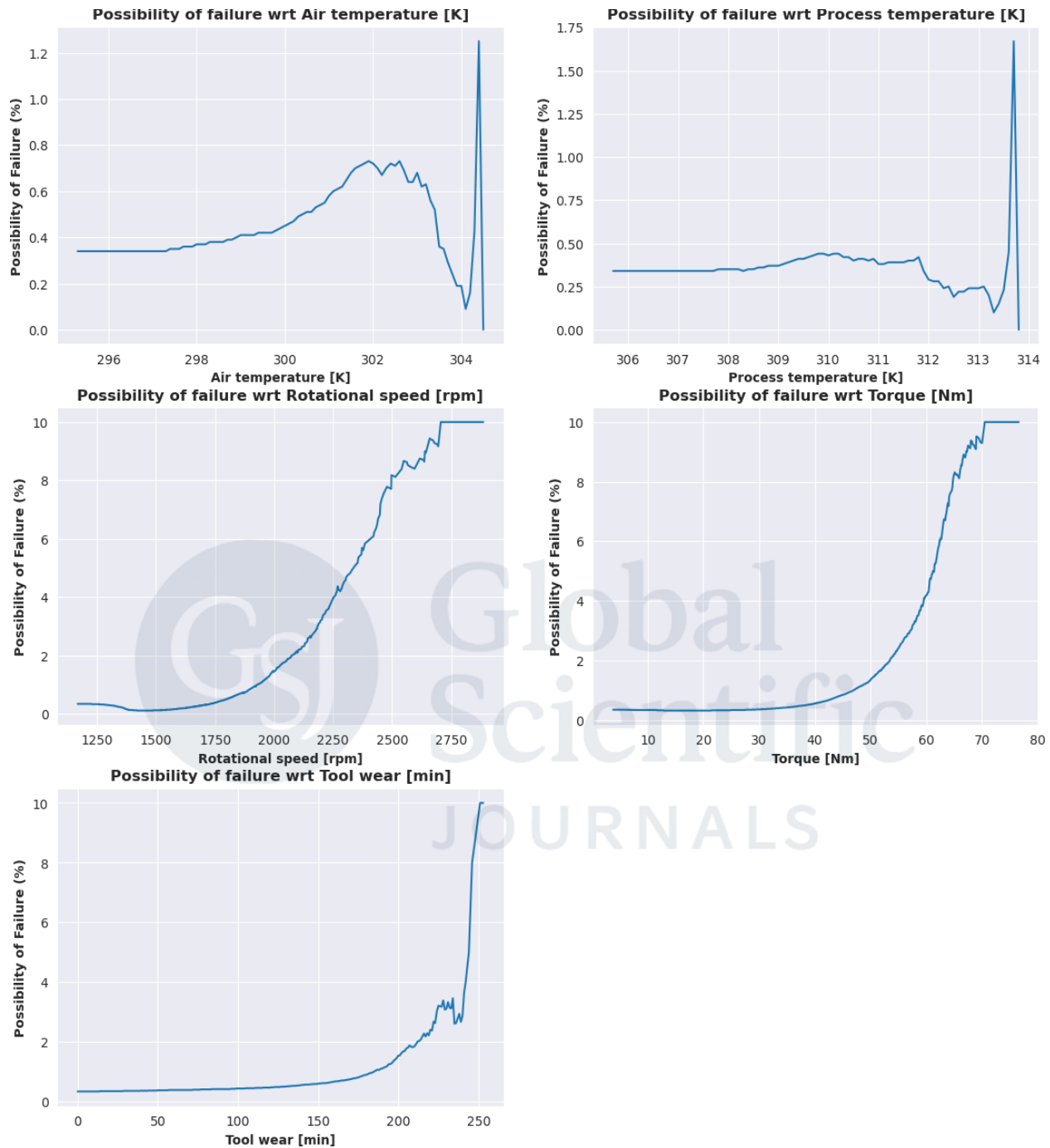


Figure 6: Possibility of machine failure with input parameters

4. Conclusion

A Random Forest machine learning model was successfully developed for predicting milling machine failure using five operational parameters: air temperature, process temperature,

rotational speed, torque, and tool wear. The model achieved a testing accuracy of 98.53%, demonstrating excellent predictive capability. The results indicate that machine learning-based predictive maintenance systems can significantly improve equipment reliability, reduce unplanned downtime, and optimize maintenance resource allocation. The developed model can serve as a decision-support tool for maintenance personnel in manufacturing industries. Future studies should investigate advanced ensemble learning techniques such as XGBoost, LightGBM, and deep learning architectures to further improve prediction accuracy.

References

1. Jansen, F., Holenderski, M., Ozcelebi, T., Dam, P., & Tijsma, B. (2018). Predicting machine from industrial time series data. *Proceedings of the International Conference on Control, Decision and Information Technologies*, 1091–1096
2. Gala, Y., Fernández, A., Díaz, J., & Dorronsoro, J. R. (2016). Hybrid machine learning forecasting of solar radiation values. *Neurocomputing*, 176, 48–59.
3. Achour, L., Bouharkat, M., Assas, O., & Behar, O. (2017). Hybrid model for estimating monthly global solar radiation for southern Algeria. *Energy*, 135, 526–539.
4. Guermoui, M., Rabehi, A., Gairaa, K., & Benkacali, S. (2018). Support vector regression methodology for estimating global solar radiation in Algeria. *European Physical Journal Plus*, 133(1), 1–9.
5. Matzka, S. (2021). Explainable Artificial Intelligence for Predictive Maintenance Applications Using the AI4I 2020 Predictive Maintenance Dataset. *Machines*, 9(12), 354.
6. Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers and Industrial Engineering*, 137, 106024.
7. Zhang, W., Yang, D., & Wang, H. (2019). *Data-driven methods for predictive maintenance of industrial equipment: A survey*. *IEEE Systems Journal*, 13(3), 2213–2227.
8. Bousdekis, A., Magoutas, B., Apostolou, D., & Mentzas, G. (2021). Review, analysis and synthesis of prognostic-based decision support methods for condition-based maintenance. *Journal of Intelligent Manufacturing*, 32(6), 1635–1656.
9. Wang, Y., Ma, E., Yang, T., & Wang, Z. (2022). Machine learning approaches for predictive maintenance in intelligent manufacturing systems: A review. *Applied Sciences*, 12(1), 321.
10. Tien Bui, D., Moayedi, H., Gör, M., Jaafari, A., and Foong, L. K. (2019). Predicting slope stability failure through machine learning paradigms. *ISPRS International Journal of Geo-*

Information, 8(9), 395.

11. Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2021). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115, 213–237.
12. Torabi, M., Mosavi, A., Ozturk, P., Varkonyi-Koczy, A., & Istvan, V. (2018). A hybrid machine learning approach for daily prediction of solar radiation. *International Conference on Global Research and Education*, 266–274.
13. Herath, H., Ariyathunge, S., & Karunasena, G. (2021). *Development of a mathematical model to forecast solar radiation and validating results using machine learning techniques*.

