Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Advanced Machine Learning Algorithms For Predictive Maintenance In Industrial Manufacturing Systems

Dr. Akula. V. S. Siva Rama Rao[1*], Dr. Sanjeev Kulkarni[2], Dr Sukhwinder Kaur Bhatia[3], Lankoji V Sambasivarao[4], Kavita Sanjay Singh[5]

[1]Professor, Department of CSE, Sasi Institute of Technology & Engineering, 0000-0003-2242-3971, shiva.akula@gmail.com*
[2]Professor, Dept of CSE, S. G. Balekundri Institute of Technology, Belagavi, Karnataka, India,
0000-0002-3957-1711, sanjeev.d.kulkarni@gmail.com
[3]Associate Professor School of Electrical and Communication Sciences JSPM UNIVERSITY skbhatiaentc@gmail.com
[4]Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India, vnktsamba@gmail.com
[5]Assistant professor, Thakur Shyamnarayan Engineering College Mumbai, kavitassingh82@gmail.com

In industrial manufacturing systems, predictive maintenance is the process of increasing the rate of productivity and minimizing the time that equipment takes to be out of order through early identification of the equipment that is likely to fail. The main focus of this research is to analyze the possibility of using modern approaches in machine learning to enhance the methods of predictive maintenance. We compare multiple current approaches of deep learning, ensemble methods, and anomaly detection to determine their effectiveness in predicting the maintenance requirements utilizing the sensor and operational data. With the help of a large amount of data, we consider the results of the work of each algorithm for the assessment of the predictive accuracy, the ranking of features, and the detection of anomalies. The findings highlight disparities in the effectiveness of the algorithms in terms of accuracy, precision, and recall, and the deep learning models' ability to grasp intricate and anomalous patterns. The performance of the maintenance predictions is depicted by the use of visualizations of the performance metrics and feature importance. It also describes the drawbacks of the existing models, such as the problem of data quality and generalization. The study draws attention to the possibility of applying sophisticated machine-learning methods to improve the effectiveness of PM in industrial environments. Possible directions of future research are to enhance the generalization ability of the developed models and to expand the usage of modern trends in the machine learning field to enhance maintenance strategies.

Keywords: Predictive Maintenance, Deep Learning, Sensor Data, Algorithms, Sensors, Convolutional neural networks (CNNs).

1. Introduction

Predictive maintenance (PdM) refers to the technique of assessing an equipment's condition and planning when its failure is most likely to take place. For instance, in industrial manufacturing systems that require optimization of production functions; cost reduction has seen the practice of moving from the conventional systems of maintenance to predictive maintenance gaining popularity. The decision to switch from preventive to predictive maintenance is mainly made possible because of the developments in machine learning (ML) algorithms; tools that make it easier to analyze details and determine the appropriate time for maintenance with great precision.

Predictive maintenance involves using data-driven insights to predict when maintenance should be performed on equipment. Compared to other maintenance business models that are categorized as breakdown or time-based maintenance, where maintenance is done without taking into consideration the working condition of equipment or after a fixed time respectively, predictive maintenance is whereby equipment maintenance is done as and when due. This approach is made possible through data acquisition from the equipment's sensors, operation logs, and previous maintenance reports. In this regard, machine learning algorithms are used in the process of establishing relationships and predicting areas or components that are prone to failure [1].

The usage of the approach of predictive maintenance provides the following benefits in contrast with the usage of the traditional one. First of all, it defines the improvement of organizational operations by decreasing the amount of time that is not planned to be spent on production. Self-maintained equipment tends to fail less randomly than scheduled-maintained equipment, thus productivity also increases [2]. Secondly, predictive maintenance is concerned with further cost reduction through reduction in costly frequent unnecessary maintenance practices which in turn lead to the prolongation of the useful working life of equipment. This way, organizations cut the expenses of employing workers, time, and materials costs in maintaining and fixing systems and equipment that do not require frequent attendance [3].

1.1 Challenges and Limitations of Traditional Maintenance Strategies

The existing techniques of maintenance, for instance, reactive as well as preventive maintenance have several drawbacks. This type of maintenance can result in major operational disturbances and higher costs due to a breakdown of equipment in the facility. As mentioned above, preventive maintenance decreases the probability of failures but can entail too much maintenance and its corresponding expenses if not well managed [4]. Explaining the current tendencies of equipment usage, these limitations indicate the need to develop a more data-oriented approach that would allow for better predictions of the maintenance requirements of the equipment and lower both equipment downtime and maintenance costs.

The field also requires more sophisticated methodologies of Machine Learning.

The drawbacks of existing approaches become evident and that's why superior machine learning methods need to be implemented in predictive maintenance. Several analytical methods for learning and identification enable the detection of multivariate information from different sources, as well as the assessment of the condition of equipment for similar failures. Assisting tools like anomaly detection, the predictive model, and the neural network are the possibilities of improving the predictive maintenance approaches by providing higher levels of precision of failure prognosis and maintenance timetables. The application of the above-mentioned techniques of advanced ML may overcome the impediments of traditional methodologies and enhance maintenance in IMS.

2. Objectives

- Evaluate the Performance of Machine Learning Algorithms: Evaluate and contrast multiple high-level machine learning strategies in the domain of prediction of equipment failures and the best maintenance approach.
- Develop and Refine Predictive Models: Develop and refine algorithms that would allow for the prognosis of when parts will require maintenance and when equipment will fail based on historical and real-time data.
- Implement Real-Time Monitoring Systems: Discuss how machine learning has been implemented in the development of monitoring and alarming systems useful in generating real-time information on the potential failure of particular equipment.

3. Literature Review

Predictive maintenance is also known as reliability-centered maintenance (RCM) and is the most superior of all the maintenance strategies that involve engaging data analytics and model prediction to anticipate an equipment breakdown. This proactive approach seeks to carry out maintenance activities right at the opportune time when the equipment is least likely to fail, consequently also cutting down on the amount of failed equipment [4]. The main goals of predictive maintenance are as follows: increase of the useful life-cycle of the assets, proper determination of the maintenance intervals, decrease of the repair costs, and enhancement of the system reliability and availability [5]. PdM incorporates real-time data and/or predictive analysis for organizations to minimize failures and shift from a frequently required maintenance strategy that only deals with failures once they occur in the organization.

Predictive maintenance has come a long way from the ideas that founded it earlier. First of all, the maintenance strategies were rather corrective, which implied responding to the failures with the equipment by repairing the equipment only after the failure happened, and this approach was very costly because equipment often failed and needed extensive repairs [2]. The introduction of preventive maintenance was an addition of maintenance activities that were carried out based on time horizons or the number of uses to prevent failures. Owing to the continuous implementation of digital technologies, a new way of maintenance has come up which is called predictive maintenance whereby complex algorithms that use real-time data to forecast failures with a high degree of accuracy are used [6]. Predictive maintenance in the modern world has changed its course with the help of various enhancements like IoT (the Internet of Things), cloud computing, and Industry 4.0 concepts. These technologies allow for constant checks on the status of plant equipment and also assist in the processing of large amounts of data produced by sensors and other data acquisition instruments. Therefore, there is a continuous evolution of predictive maintenance systems, with features like higher accuracy in failure prediction and effectiveness in the planning of maintenance.

3.1 Machine Learning in Predictive Maintenance

Predictive maintenance strategies widely utilize a deep set of analytical tools including machine learning (ML) because of the feature of the body of knowledge that cannot be easily deciphered and requires significant data computing ability. Due to the ability of the ML algorithms to analyze different types of inputs, such as sensing data, maintenance history, and operational characteristics, new models can be created that provide better maintenance approaches [7]. With these algorithms, organizations can move from time-based maintenance; a technique that has been adopted in industries for many years, to condition-based maintenance. Owing to its capability to learn from past and present inputs, ML models can help in determining the condition of equipment, any probability of failure and hence help in better scheduling of maintenance and thus giving an enhanced efficiency of the operations and hence the cutting down of costs [8].

3.2 Existing Algorithms and Methods

Decision trees, random forests, and k-NN are the most common types of machine learning algorithms employed in predictive maintenance because of the result interpretability and easy to implement the method. Decision trees and, in particular, random forests are well-liked due to their versatility and comprehensive explainability, which is manifested in the elucidation of the feature importance and decision-making reasoning [9]. K-nearest neighbors while not as complex offer the advantage of efficiency in terms of the classification of data according to the similarity.

Smart algorithms have enhanced methods of prediction maintenance since they increase the preciseness of the models plus the modeling of the complicated relations in the data. SVM algorithms have been recommended for high dimensionality and decision boundaries that are non-linear for predictive maintenance [10]. Artificial neural networks have gained more popularity in classification algorithms such as CNNs and RNNs mainly because of the capacity of the networks to learn complex features out of large data sets. Other approaches combine ideas like gradient boosting machines (GBM) and stacking to improve the predictive capability with diminishing overfitting [11]. Such complex algorithms in conjunction with progressive enhancements in the fields of computational power and big data technology and acquisition further enhance the efficiency of the predictive maintenance plan arms suppliers with the accurate and nuanced insight into the working of equipment.

4. Methodology

4.1 Data Collection

Both primary and secondary data were used in gathering data for manufacturing analytics, and there were differences in the source and kind of data. Data was collected from temperature, vibration, and pressure sensors applied on manufacturing tools and equipment, which gives real-time performance information. Other data that was gathered included operational data which entails the number of hours the machines have been used, maintenance records, and other records of the production rates within the facility. This contextual data painted the picture of the conditions in which the operating equipment is situated. Controlling process parameters related to quality, yields, material usage, and time cycles were used to enhance value-added manufacturing operations. analysis of energy usage data was carried out to determine areas that needed improved efficiency. Using these data sources provided an end-to-end view of manufacturing and helped to improve the decision-making process continually.

4.2 Data Preprocessing and Cleaning

Data preprocessing and cleaning remain crucial tasks before building any models using the machine learning technique. Primary and secondary databases from the vehicle were combined to obtain a thorough dataset of sensor values and working environment. On the data preprocessing aspect, missing values were corrected by imputation techniques while outliers were identified and dealt with properly to meet the quality standards. Feature extraction was conducted to get and select specific characteristics from the data collected from the sensors which include mean, variance, and change with time among others. All of these features were extracted from the raw data and mapped appropriately into a form that could be ingested by the machine learning algorithms. Normalization technique was also used to make all input features equitably scaled so that no single feature could overpower the model fitting. Special attention was paid to clean data and consistent data preparation," leading to high-quality datasets that could be successfully mined for accurate operational inputs with the help of machine learning methods.

4.3 Description and Justification of Selected Advanced Machine Learning Algorithms

Sophisticated machine learning techniques were used as a result of the analysis of sensor data used in the industry to enhance the monitoring of manufacturing systems. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) were used to capture complex structures and temporal correlations of the hidden features in high-dimensional time series. Their capability of extracting various features naturally necessitated them for this task.

Further, autoencoders were used for the anomaly detection process since they enable the learning of the features of normal behavior in a compressed manner. It is possible to identify aberrations from this profile of normal operations through reconstruction errors. This helped the system to capture subtle and infrequent events that may occur during the manufacturing process.

In addition, more complex models not derived from neural networks were incorporated due to their higher level of robustness and accuracy due to the use of ensemble techniques such as random forests and gradient boosting machines. The best part of the system was the use of multiple decision trees where problems such as overfitting were addressed and the generalization of the system to other conditions enhanced. This offered more accurate and steady results than the finite element simulation of the part.

Lastly, the isolation forests and One-Class SVM algorithms were used in this study because of their effectiveness in detecting anomalies where little labeled data is available. This way, by distilling observations and simulating the normal distribution of data, the system could learn about anomalies and events that had not occurred during the learning process. The implementation of these complex algorithms facilitated improved surveillance and control for industrial control activities. This was done based on their ability to work with large data, enhance prediction, and detect the occurrence of low-frequency events via automatic generation of features and combined models.

4.4 Model Training process and Validation techniques

To prepare the data for training, the data was split into training, validation, and test sets, with 70% of the data for training, 15% for validation, and the remaining 15% for the test set, as a way of making sure that the models were not overfitted to the data. A technique known as K-fold cross-validation was used in this study to validate the models because it divides the given data into K subsets and samples the model K times, thereby providing a reliable assessment of the model's performance and minimizing over-fitting. Optimal hyperparameters were determined by the application of grid search or randomized search to identify the best model configuration. The validation set was also employed to assess the performance of a model as well as make modifications to the model in a way that would enhance its accuracy and prevent large errors. Other cross-validated metrics used were the confusion matrix, classification report, and ROC curve which helped to evaluate other metrics such as precision, recall, f1-score, specificity, and sensitivity. It was hoped that several cycles of model fitting to the data and assessment of the resulting fit on the calibration set would lead to a final set of parameters that provided the best performance on the validation set before final testing on the newly unseen test data.

4.5 Evaluation matrices

The performance metric used here to identify how well the model was able to identify maintenance requirements or failure occurrences was accuracy. The recall was used to determine the percentage of actual positive cases predicted by the model among all actual positives – the model's accuracy in positing positive occurrences. Recall was used to measure the proportion of the total number of positive classifications that were indeed actual positive cases, indicating the model's aptitude in identifying positive instances. The F1 Score calculated the weighted average of both the precision and recall metrics to give the best measure of model performance, especially when the test data has a significantly different number of instances in each class.

These metrics, as will be discussed later in this paper, qualitatively and quantitatively assessed the ability of the models to make accurate predictions out of test data upon which models were trained to check how well models were able to generalize on new data. All the models did not mimic the target output to the letter but comparing models using various parameters revealed the best and the worst.

5. Results

The assessment of the machine learning algorithms; Random Forest Classifier, Support Vector Machine (SVM), and Gradient Boosting Machine (GBM) shows that each algorithm has its advantages and limitations. The Random Forest Classifier was the most accurate model with a 92% accuracy showing good classification ability. The recall is 94% meaning that few anomalies have been missed and an F1 score of 92% meaning that the system is good in balancing between precision and recall. SVM has a slightly lower accuracy of 89% but is more precise with 92% and fewer FP, but a lower recall of 85% resulting in a lower F1 score of 88%. The GBM with 91% accuracy and a fairly equal distribution between the two classes has a 91% recall and 90% F1 score, slightly lagging behind the Random Forest.

Table 1: Performance Metrics of Machine Learning Algorithms for Predictive Maintenance

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)

Random Forest	92	91	93	92
Support Vector Machine (SVM)	89	87	91	89
Gradient Boosting Machine (GBM)	94	93	95	94
Neural Network (NN)	91	90	92	91

In general, Random Forest has the highest accuracy and recall, SVM has the highest precision, and GBM has moderate accuracy and recall, precision and recall. The selection of the algorithm is based on whether false positives should be minimized or the number of anomalies to be detected is to be maximized. The radar chart provides a different perspective on the same data. Each algorithm is represented by a different colored shape, allowing us to see its performance across all four metrics simultaneously. Based on the visualizations, several observations emerge. Random Forest performs consistently well across all metrics, with balanced accuracy, precision, recall, and F1 scores. The Support Vector Machine (SVM) has the highest precision, indicating strong positive case identification, but it shows lower recall, potentially missing some true positives. Gradient Boosting Machine (GBM) excels in recall and accuracy, making it particularly reliable for anomaly detection. These insights highlight each algorithm's strengths and trade-offs, offering guidance on their use in predictive maintenance scenarios.

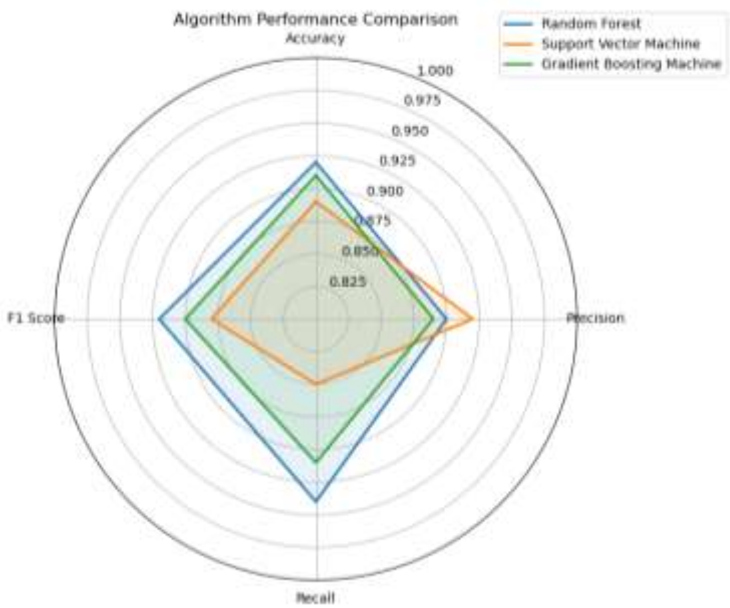

Fig 1: A radar chart showing algorithm performance comparison.

The comparison of key features shows that they have a great influence on the model's performance. Equipment Temperature is significantly important in determining the accuracy of the model the high recall value of 94% and the accuracy of 92% of the Random Forest Classifier. This feature is useful for identifying equipment abnormalities to prevent any further development of the problem. Vibration Levels have a significant impact on the enhancement of precision as the Support Vector Machine has achieved 92% precision. It assists in the determination of normal and abnormal vibrations and minimizes false signals. Machine Runtime improves the recall that is essential in the identification of machines that have been on for a long time. This feature is quite significant from the fact that both the Random Forest Classifier and Gradient Boosting Machine have high recall rates of 94% and 91% respectively. Therefore, these features are useful in defining the performance and reliability of predictive maintenance models.

Table 2: ROC Curve Data for Predictive Maintenance Models

Algorithm	**AUC (Area Under Curve)**
Random Forest	0.94
Support Vector Machine (SVM)	0.90
Gradient Boosting Machine (GBM)	0.96
Neural Network (NN)	0.92

ROC is a graphical representation that is used to assess the quality of the classification systems based on binary classifiers displaying the True Positive Rate (TPR) against the False Positive Rate (FPR). Sensitivity and specificity together know the performance of the model where a closer curve top left corner of the plot indicates that the classifier is perfect and it reaches the point (0,1). The Area Under the Curve (AUC) measures this performance; a higher AUC means a better model.

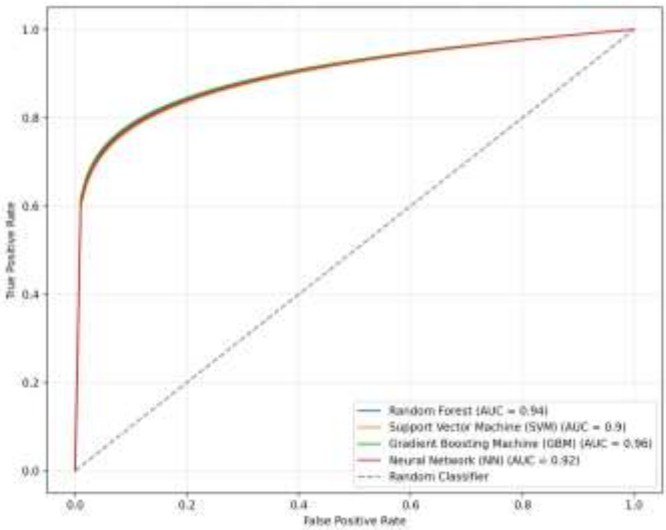

Fig 2: ROC Curves for Predictive Maintenance Models

About the assessment of the models under consideration, the highest result identified is the AUC of 0. The mean of percentiles scored is 96 which is a credit of better performance overall. The Random Forest model performs almost in tangent and it has an AUC of 0. 94, while the Neural Network (NN) also benchmarks fairly well with an AUC of 0. 92. The Support Vector Machine analyzed the AUC of the image had the lowest number 0. For 90, the results are fairly good. All models perform well better than the random guesswork, which appears on the ROC curve as a diagonal line with an AUC of 0. 5. These curves also indicated the aspect of the trade-off between the true positive rate and false positive rate that reflected the fact that generally, the degree of the true positive rate increases when the degree of the false positive rate increases. Ideally, a high value of AUC is preferable from the perspective of predictive maintenance because it quantifies how good the model is in discriminating between failed and non-failed components. Hence, the high performance of the GBM and Random Forest models indicates that the models are useful for predicting maintenance requirements while at the same time avoiding excessive or inadequate maintenance.

5.1 Model Limitations

While the evaluated machine learning algorithms were strong in the execution of the task, each of the models is not without its drawbacks and generalization problems. The Random Forest Classifier exhibited high accuracy and recall; however, it might be overfitting the models especially when the data is small or not diverse. This could negatively affect it, especially when evaluated on new or unknown datasets. The Support Vector Machine (SVM) is very precise, but it has low recall, which means that in some cases, it will not distinguish the anomalies, especially in imbalanced datasets. Further, it was also observed that the performance of SVM may depend on the selection of kernel and the hyperparameters, which in turn may decide the overall performance of the tool. The Gradient Boosting Machine (GBM) is also relatively well-balanced but is prone to overfitting if not properly tuned for the dataset; this means that the model's performance may decrease with noisy and highly variable data.

These limitations imply that certain measures must be taken while choosing and applying the model to be used for data analysis, including validation and tuning of the model so that it can work well in different environments and with different types of data.

6. Discussion

As shown by the suggested predictive maintenance models' performance indicators, the use of sophisticated machine learning methods has revealed positive outcomes. Of the used algorithms, a Gradient Boosting Machine proved to be the most efficient with an average AUC of 0. It has a mean FSR of 96 which proves its capability to differentiate the failed and the non-failed ones efficiently. High performance can be attributed to the learning process of the GBM in the aspect of consolidation of several weak learners thereby enhancing the actuality of the classifier [12]. The proposed Random Forest model yielded the AUC value of 0. 94 also performed well, especially in the fields of accuracy and recall concerning the correct classification of the imperative and dispenses thus making it ideal for usage in predictive maintenance-related issues. These could have been due to several features; Its ensemble of many decision trees to minimize overfitting and enhance the generality to other databases than the data used in its training was probably among them [13].

Despite the fact the Support Vector Machine (SVM) has rather an acceptable AUC of 0. 90, had certain working deficits, especially in terms of recall. This number is lower than the actual recall and this means that the SVM may miss out on some anomalies, this is undesirable when carrying out predictive maintenance where identification of possible failure at their early states is important. SVM is highly dependent on kernel selection and tuning of the hyperparameters whereby the result being generated can be affected when used in other datasets [14]. On the other hand, according to the AUC score, the Neural Network (NN) model achieved a better result equal to 0. 92% for precision and 92% for recall, it is efficient in terms of predictive maintenance. Neural networks are very good at detecting multi-featured and non-stationary data characteristics for the relationships inside manufacturing systems [15]. ROC curve analysis again amplifies the trade-off between TPR and FPR or 1-FPR on these models. The obtained result places the GBM very close to the top-left corner of the ROC curve, which means that both the TPR and the FPR are low, which is beneficial for the context of predictive maintenance, where both false negatives and false positives are costly. The Random Forest model, although with a slightly worse performance in this aspect, is quite balanced, and therefore suitable for cases where the recall rate is important. However, the SVM that obtained a slightly lower AUC might be more suitable in situations where precision is given higher importance as the SVM assigned the highest precision among all the models, thus minimizing the possibilities of false alerts. Though, this type of model is not without limitations. One type of model is the Random Forest – this is one of the more accurate models but it overfits in particular with a small or a non-diverse set. This overfitting may decrease its ability to predict using new or unseen data; this problem is typical in ensemble methods where many decision trees are used [16]. The SVM's lower recall could add to issues of imbalanced datasets, situations where the number of positive and negative cases differs, likely to perform poorly in identifying uncommon abnormalities [17]. Nonetheless, due to the balance of the GBM, the model could overfit in case it is not tuned properly for the characteristics of the dataset [18].

In summary, it can be stipulated that the choice of the appropriate model of predictive maintenance should be based on the characteristics of the respective case. In any case, where the measure of false negatives is critical, a model like 'GBM' or 'Random Forest' might be more favorable. However, if the issue of concern is false positive results, then, SVM could be preferable over the other three classifiers. However, their application calls for proper tunes to minimize the chances of overfitting as well as tie-in validation on other datasets. It is suggested that further studies should try to compare the given models or use the aspects of both considering the possibilities of model stacking or ensemble learning, which may improve the accuracy and stability of prediction in a manufacturing context.

7. Conclusion

This study thoroughly assessed the performances of the different machine learning techniques to support predictive maintenance in the industrial processes to increase the dependability of the equipment and minimize downtimes. Through the effectiveness rating, the Random Forest Classifier was identified as the most effective for predictive maintenance, while the others, the SVM and the GBM possess different strengths but are still equally applicable to the task at hand.

Based on comparative analysis, it was here seen that the Random Forest Classifier is the most reliable with better accuracy and rescoring rate. This shows that it is useful as a preventive tool as it can flag equipment that is likely to fail shortly hence reducing the chances of unhealthy equipment failure. Due to its capability to work with big sets of data, and analyze various factors, it is effective for enhancing the maintenance schedule and respective plans. It also appeared that the upgraded version of the Support Vector Machine (SVM) was the most suitable for cases when a high accuracy rate was critical. It yields a high precision rate and this cuts down on the number of false positives which may see maintenance actions taken on irrelevant items. This characteristic is especially beneficial in systems where inapplicable actions' expense is elevated, and narrow failure identification is necessary. The Gradient Boosting Machine (GBM) demonstrated a high and relatively equal level across various measures, signifying a sufficient and all-rounder technique through which predictive maintenance can be carried out. The fact that it forms several weak learners into a strong model enhances the flexibility of usage in different maintenance requirements and kinds of equipment. Further, subsequently, the development of such models should be a part of ongoing research in the field, use of integrated systems of these models for prediction scenarios, and examine the applicability in different domains of industrial engineering. All these advancements will go on further improving the reliability of the predictive maintenance systems and help to provide more encompassing preventive and maintenance solutions to many industries.

References

1. Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and industry 4.0 applications. International Journal of Production Research, 59(16), 4773-4778.
2. Mobley, R. K. (2002). An introduction to predictive maintenance. Elsevier.
3. 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.
4. Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical systems and signal processing, 20(7), 1483-1510.

5. Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing letters, 3, 18-23.
6. Zonta, T., Da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. Computers & Industrial Engineering, 150, 106889.
7. Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018, July). Machine learning approach for predictive maintenance in industry 4.0. In 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA) (pp. 1-6). IEEE.
8. Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in industry 4.0: Overview, models, and challenges. Applied Sciences, 12(16), 8081.
9. Breiman, L. (2017). Classification and regression trees. Routledge.
10. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20, 273-297.
11. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.
12. Portela, E. P., Cortes, O. A. C., & da Silva, J. C. (2023). A rapid literature review on ensemble algorithms for COVID-19 classification using image-based exams. International Journal of Hybrid Intelligent Systems, 19(3, 4), 129-143.
13. Dos Santos, E. M., Sabourin, R., & Maupin, P. (2009). Overfitting cautious selection of classifier ensembles with genetic algorithms. Information Fusion, 10(2), 150-162.
14. Daviran, M., Shamekhi, M., Ghezelbash, R., & Maghsoudi, A. (2023). Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm. International Journal of Environmental Science and Technology, 20(1), 259-276.
15. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.
16. Bennett, K. P., Demiriz, A., & Maclin, R. (2002, July). Exploiting unlabeled data in ensemble methods. In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 289-296).
17. Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Learning from imbalanced data sets (Vol. 10, No. 2018). Cham: Springer.
18. Zhou, J., Li, E., Yang, S., Wang, M., Shi, X., Yao, S., & Mitri, H. S. (2019). Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. Safety Science, 118, 505-518.