Comparative Analysis of Machine Learning Models for Predictive Maintenance in HVAC Constructions

K. Alice[1*], K. Surya[2]

Associate Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai[1]
Assistant professor, Department of Electronics and Communication Engineering, SRM Easwari Engineering College, Chennai[2]
Correspondence: alicesureshme@gmail.com[1]

Abstract: Predictive maintenance is vital for ensuring the longevity and optimal performance of building systems, such as Heating, Ventilation, and Air Conditioning (HVAC) units. By utilizing historical sensor data, machine learning techniques can forecast potential failures, enabling timely maintenance interventions. This study evaluates and compares the performance of three prominent machine learning Models-Logistic Regression (LR), Random Forest (RF), and Gradient Boosting (GB) in predicting HVAC unit failures. A simulated dataset encompassing features like temperature, humidity, pressure, usage hours, and maintenance logs was used for this comparison. Finally, Gradient Boosting demonstrated the highest performance, with an accuracy of 87 %, precision of 0.86, recall of 0.88, F1-score of 0.87, and ROC-AUC of 0.92. These results underscore the superior predictive capabilities of Gradient Boosting and its potential in enhancing maintenance strategies for critical building systems. The results suggest that Gradient Boosting is the most effective model for enhancing predictive maintenance strategies in building systems, offering valuable insights for timely and efficient maintenance interventions.

Keywords: Predictive Maintenance, Gradient Boosting, HVAC Systems, Machine Learning, Random Forest, Logistic Regression, Classification Metrics, ROC-AUC

1. INTRODUCTION

Building maintenance plays a vital role in infrastructure management, influencing public safety, operational continuity, and cost-effectiveness. Poor maintenance practices can lead to early system failures, escalating expenses, and structural degradation that endangers the safety and functionality of buildings. The lack of an effective building maintenance management system, coupled with the absence of advanced predictive maintenance tools, presents significant obstacles to the sustainable development of both current and future infrastructures.

Traditionally, building maintenance has been reactive, addressing issues only after failures occur. This reactive approach is often the reason for unplanned downtime, elevated repair costs, and inefficient use of resources. In response to the above mentioned challenges, there has been a global shift towards more proactive maintenance strategies. Predictive maintenance, which aims to forecast potential failures by analyzing historical and real-time data, has become a key component of modern building management systems. This shift enhances operational efficiency, extends equipment lifespans, and reduces maintenance costs [1,2].

Machine learning (ML) has evolved as an unignorably powerful enabler of predictive maintenance. Unlike conventional methods that rely on predetermined schedules or reactive measures, ML models can analyze large datasets from building systems to identify patterns and predict when failures are likely to occur. This allows facility managers to optimize maintenance activities, minimizing downtime and avoiding unnecessary repairs.

This research explores the adaptation of ML models to predict system failures in buildings using sensor data. Specifically, we compare the evaluated outcomes of three widely-used algorithms—Logistic Regression, Random Forest, and Gradient Boosting [3]. Each model presents distinct advantages in terms of accuracy, interpretability, and efficiency.

The objective is to identify which model performs best in predicting failures, providing facility managers with valuable insights for optimizing building maintenance strategies. Moreover, this study seeks to bridge the gap in building management practices and promote sustainable infrastructure development [4, 5].

2. LITERATURE SURVEY

Numerous studies have explored predictive maintenance frameworks and their effectiveness across various building systems. Almobarek et al. (2023) [6] developed a predictive maintenance framework for cool water systems in commercial buildings, utilizing a decision tree model that achieved over 98% accuracy in fault prediction and reduced maintenance costs by more than 20%. This illustrates the considerable potential of machine learning in enhancing system reliability. Bouabdallaoui et al. (2021) [7] introduced a comprehensive five-step predictive maintenance framework that includes data collection, processing, model development, fault notification, and model refinement. When applied to HVAC systems in a sports facility using IoT devices and a building automation system, the framework demonstrated promise in forecasting system failures. However, challenges related to data availability and feedback mechanisms indicated areas that require further improvement.

Lu et al. (2007) [8] proposed a condition-based maintenance strategy that models system deterioration as a stochastic dynamic process. By employing structural time series, state-space modeling, and Kalman filtering, they were able to predict future system deterioration, allowing for decisions to be made based on failure probabilities and cost analyses. This method provided a dynamic balance between preventive maintenance and economic factors. In a more recent study, Arora et al. (2023) [9] adapted an existing predictive maintenance methodology for residential heating systems, benefiting the reliability department at BOSCH Thermotechnology. Their research showcased the versatility of predictive maintenance across various domains by optimizing its implementation for heating systems.

Omar et al. (2019) [10] designed a mathematical model for building maintenance based on global standards and expert input. Using the Weighted Sum Model (WSM) and SPSS analysis, they identified critical maintenance items and proposed their model for quick evaluation and decision-making. This framework provides architects and engineers with a structured tool for predictive maintenance.

Sanzana et al. (2022) [18], investigates the application of machine learning (ML) techniques for predictive maintenance specifically in cooler environments. The authors evaluate various ML algorithms, including supervised and unsupervised methods, and discuss their effectiveness in analyzing sensor data and operational metrics to anticipate equipment failures. By focusing on the unique challenges posed by cooler conditions, the study highlights the need for tailored approaches in data collection and analysis. The authors support their findings with practical case studies that demonstrate the benefits of implementing predictive maintenance strategies, such as improved equipment reliability and reduced downtime, while also suggesting future research directions to enhance model accuracy through hybrid approaches and integration of domain knowledge.

Recent advancements in machine learning have revolutionized predictive maintenance, especially in industrial systems. Supervised learning algorithms like Support Vector Machines (SVM), Decision Trees, and Neural Networks have proven effective in predictive tasks. Studies focusing on building systems, particularly HVAC units, have found Random Forests and Gradient Boosting to be highly effective in fault detection. Ensemble learning techniques, which combine multiple models, often outperform single-model approaches due to their ability to minimize variance and bias (Kasiviswanathan et al. 2024) [12,13].

Despite the progress in the field, there is a lack of comparative analysis between simpler models such as Logistic Regression and more complex ensemble methods like Random Forest and Gradient Boosting. This study fills this gap by comparing the performance of these models in predictive maintenance for building systems.

3. METHODS AND MATERIALS

The proposed research work, we applied three machine learning models to a dataset derived from sensor data of HVAC units. The dataset included features such as temperature, humidity, pressure, usage hours, and maintenance logs. The proposed framework for predictive maintenance of HVAC systems utilizes simulated historical sensor data, incorporating features like temperature, humidity, pressure, and usage logs, with the target variable representing maintenance needs.

Data pre-processing involves addressing missing values, normalizing continuous variables, and dividing the dataset into training and testing sets. Attribute selection was performed to identify the most relevant predictors, thereby enhancing model efficiency. We trained and evaluated three machine learning models and evaluated performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.

A comparative analysis of these models was carried out to identify the most effective approach for optimizing predictive maintenance strategies. The workflow of the proposed methodology is illustrated in Figure 1.

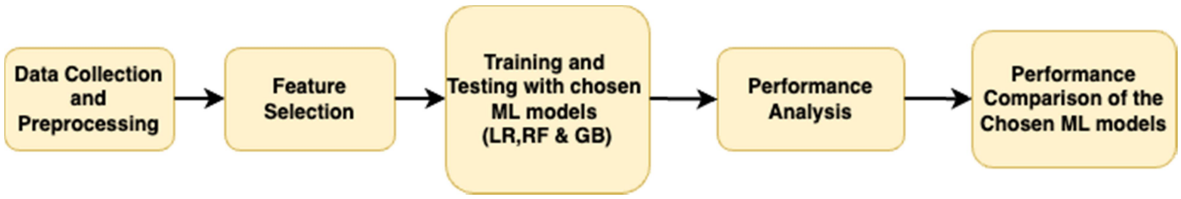

Figure 1: Work flow of the proposed framework.

3.1 DATA DESCRIPTION

The dataset used for this study was simulated to mimic real-world HVAC sensor data. Table 1 summarizes the key features and target variable of the simulated HVAC sensor data.

Table 1: Key features and target variable of sensor data

Variables	Description	Range/Values
Temperature (°C)	Temperature readings from the HVAC system	15 - 35 °C
Humidity (%)	Humidity levels in the environment	30% - 70%
Pressure (kPa)	Pressure readings from the HVAC system	90 - 120 kPa
Usage Hours	Number of hours the HVAC unit had been in operation	Varies (Continuous)
Maintenance Logs	Binary indicator of whether maintenance had been performed (0: No, 1: Yes)	0 or 1
Target (Failure)	Binary indicator of system failure (0: No Failure, 1: Failure)	0 or 1

3.2 SELECTION OF ML MODELS FOR COMPARATIVE ANALYSIS

In the comparative investigation of ML models for predictive maintenance, selecting the right models is essential to harness their complementary strengths. We chose Logistic Regression, Random Forest, and Gradient Boosting because of their unique characteristics and benefits:

3.2.1 Logistic Regression (LR)

LR is valued for its straightforward implementation and ease of interpretation. It provides binary classification by modeling the relationship between features and the probability of an event (Cramer 2002) [11].

i) Simplicity and Interpretability: Its simplicity makes it a suitable baseline model, offering fast training and prediction times. However, its major limitation is its inability to capture non-linear relationships, which can restrict its performance in more complex scenarios typical in predictive maintenance. Despite this limitation, Logistic Regression's interpretability makes it valuable in situations where understanding the relationship between individual features and outcomes is important. For example, in predictive maintenance, the model can offer insights into the contribution of each feature to the likelihood of equipment failure.

ii) Probabilistic Output: Logistic Regression generates probabilistic outputs, which provide more than just binary classifications. This can be particularly useful for threshold-based decision-making in predictive maintenance, where a probability score might be used to prioritize maintenance actions.

3.2.2 Random Forest

Random Forest is an ensemble learning technique that constructs multiple decision trees and combines their results to improve both accuracy and stability. Each tree is trained on a randomly selected subset of the data, and the final prediction is made through majority voting for classification tasks or averaging for regression tasks. This approach reduces the risk of overfitting that individual trees often face, making the model more robust. Random Forest also ranks features by their importance, providing insights into which variables contribute most to the prediction. Overall, it is highly effective for complex datasets and large-scale predictive tasks. [14,15].

i) Robustness and Flexibility: Its robustness makes it well-suited for complex datasets with diverse feature types, as it mitigates the risk of over fitting, which can be a concern with simpler models like Logistic Regression. In predictive maintenance, where data may come from various sources like sensors, logs, or environmental factors, Random Forest can handle the heterogeneity effectively. Additionally, Random Forest provides insights into feature importance, which can help identify the most critical factors driving system failures, thus aiding maintenance prioritization.

ii) Non-linear Relationships: Unlike Logistic Regression, Random Forest is capable of capturing non-linear relationships between features. This makes it particularly valuable for predictive maintenance tasks, where complex interactions between equipment age, usage patterns, and environmental conditions might affect failure rates. Random Forest's ability to model such interactions enhances its predictive power in maintenance scenarios.

3.2.3 Gradient Boosting

Gradient Boosting is known for its superior predictive performance, achieved through the iterative correction of errors from previous models. It works well with different loss functions and adapts effectively to complex patterns within the data [16,17].

i) High Predictive Accuracy: For predictive maintenance, where subtle trends in sensor data or operational logs could signal upcoming failures, Gradient Boosting's ability to refine its predictions iteratively makes it particularly useful. However, Gradient Boosting comes with the trade-off of longer training times and potential over fitting if not carefully tuned. Nonetheless, its accuracy makes it a powerful model for more intricate maintenance datasets.

ii) Handling of Outliers: One of Gradient Boosting's strengths is its reduced sensitivity to outliers, thanks to its iterative learning process. In practical applications of predictive maintenance, sensor anomalies or unusual data points are common, and Gradient Boosting's ability to handle these outliers without sacrificing performance enhances its robustness in real-world settings.

3.3 TRAINING AND TESTING OF MACHINE LEARNING MODELS

The dataset was split into 70% for training and 30% for testing to provide an impartial assessment of the model's performance.This train-test split strategy allowed for sufficient data to train the models while reserving a portion for evaluating how well the models generalize to unseen data.

- **Logistic Regression**: The model was trained using a standard linear relationship between the input features and the binary target variable. This approach allows the model to estimate probabilities and classify outcomes based on the learned weights, making it a baseline model for comparison.
- **Random Forest**: Random Forest was trained using 100 decision trees, with each tree constructed using a random subset of features. This method improves generalization by reducing the risk of overfitting and increasing model robustness, particularly in handling non-linear relationships and diverse feature types.
- **Gradient Boosting**: The Gradient Boosting model was trained with a learning rate of 0.1 and 100 boosting iterations. Through this iterative process, the model sequentially corrects errors from previous iterations, enhancing its ability to capture complex patterns in the data and improving overall predictive performance.

Each model was assessed on the testing set to evaluate its accuracy, precision, recall, F1-Score, and ROC-AUC, providing a complete understanding of their effectiveness in predictive maintenance tasks.

4. RESULTS AND DISCUSSION

The performance comparison of three ML classifiers reveals that Gradient Boosting consistently outperforms the other models across all key metrics as presented in Table 2 and Figure 2.

Table 2: Comparison of performance metrics for the chosen ML models

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC |
Logistic Regression	0.82	0.80	0.85	0.82	0.88
Random Forest	0.84	0.83	0.86	0.84	0.90
Gradient Boosting	0.87	0.86	0.88	0.87	0.92

Gradient Boosting emerges as the best performer with an accuracy of 0.87, indicating it correctly predicts outcomes 87% of the time, followed by Random Forest at 0.84 and Logistic Regression at 0.82. In terms of precision, Gradient Boosting again leads with 0.86, showing that it minimizes false positives better than Random Forest 0.83 and Logistic Regression 0.80. This suggests that Gradient Boosting and Random Forest are more suitable when avoiding false positives is critical.

When it comes to recall, the ability the measures to identify true positives, Gradient Boosting scores highest at 0.88, slightly ahead of Random Forest at 0.86 and Logistic Regression at 0.85. This means Gradient Boosting is more reliable in scenarios where missing true positives, such as in fraud detection or medical diagnoses, is a greater concern.

Figure 2: Performance comparison of ML models in Predictive Maintenance in Building Systems

The F1-score, which balances precision and recall, follows a similar pattern. Gradient Boosting has the highest score of 0.87, indicating the best trade-off between precision and recall, while Random Forest and Logistic Regression have F1-scores of 0.84 and 0.82, respectively.

Lastly, the ROC-AUC values reinforce Gradient Boosting's superior performance, with the model achieving a score of 0.92, the highest among the three. This means it has the best ability to distinguish between positive and negative cases. Random Forest follows with a ROC-AUC of 0.90, while Logistic Regression scores 0.88. Overall, Gradient Boosting is the top-performing model across all metrics, making it the best choice for tasks requiring both high prediction accuracy and class separation. Random Forest is a strong alternative, while Logistic Regression, though effective, falls behind in precision and overall classification quality. Overall, Gradient Boosting is recommended for complex scenarios, while Random Forest and Logistic Regression are viable alternatives where interpretability and speed are prioritized.

5. CONCLUSION

The analysis reveals that Gradient Boosting consistently outperforms Logistic Regression and Random Forest across all key metrics. Its superior predictive accuracy, higher recall, and robustness in distinguishing between maintenance and non-maintenance scenarios make it the most suitable model for predictive maintenance in building systems. However, Random Forest also performs well, especially in feature-rich environments where interpretability (via feature importance) is valued. Logistic Regression, while not as strong in handling complex relationships or non-linearity's, remains a reliable and interpretable baseline model, especially useful in cases where transparency is paramount. In practical applications, Gradient Boosting should be the model of choice when accuracy and the ability to handle complex data are priorities. However, for faster, more interpretable models, Random Forest or Logistic Regression may still be viable options depending on the specific use case.

CONFLICTS OF INTEREST: The authors declare no conflict of interest.

REFERENCES

1. L. Breiman, "Bagging predictors," Machine Learning, vol. 24, no. 2, pp. 123-140, 1996.

2. Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139, 1997.
3. X. Li, et al., "Ensemble learning for predictive maintenance," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 393-404, 2019.
4. J. R. Quinlan, C4.5: Programs for Machine Learning. San Francisco, CA: Morgan Kaufmann, 1993.
5. H. Zhang, et al., "Boosting-based fault detection in HVAC systems," *Energy Procedia*, vol. 143, pp. 725-731, 2017.
6. M. Almobarek, K. Mendibil, and A. Alrashdan, "Predictive maintenance 4.0 for chilled water system at commercial buildings: A methodological framework," Buildings, 2023. Available: https://doi.org/10.3390/buildings13020497.
7. Y. Bouabdallaoui, Z. Lafhaj, P. Yim, L. Ducoulombier, and B. Bennadji, "Predictive maintenance in building facilities: A machine learning-based approach," S*ensors*, vol. 21, 2021. Available: https://doi.org/10.3390/s21041044.
8. S. Lu, Y. Tu, and H. Lu, "Predictive condition-based maintenance for continuously deteriorating systems," *Quality and Reliability Engineering International*, vol. 23, 2007. Available: https://doi.org/10.1002/qre.823.
9. S. Arora and M. Rabe, "Predictive maintenance: Assessment of potentials for residential heating systems," *International Journal of Computer Integrated Manufacturing*, vol. 36, pp. 1876-1900, 2023. Available: https://doi.org/10.1080/0951192X.2023.2204471.
10. N. Omar, W. Hatem, and H. Najy, "Predictive modeling for developing maintenance management in construction projects," *Civil Engineering Journal*, 2019. Available: https://doi.org/10.28991/CEJ-2019-03091297. [Accessed: Sept. 16, 2024].
11. J. S. Cramer, "The origins of logistic regression," *Tinbergen Institute Technical Report*, vol. 119, pp. 167–178, 2002. Available: https://doi.org/10.2139/ssrn.360300.
12. S. Kasiviswanathan, T. B. Vijayan, and S. John, "Performance analysis of the ensemble model in anaemia detection from unmodified smartphone-captured conjunctiva images," *IETE Journal of Research*, pp. 1–12, 2024. Available: https://doi.org/10.1080/03772063.2024.2367046
13. L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001. Available: https://doi.org/10.1023/A:1010933404324.
14. T. K. Ho, "Random decision forests," in Proc. of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 1995, pp. 278–282.
15. T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998. Available: https://doi.org/10.1109/34.709601.
16. J. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
17. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. New York, NY: Springer, 2009.
18. M. R. Sanzana, M. O. M. Abdulrazic, J. Y. Wong, T. Maul and C. -C. Yip, "Analysis of Machine Learning Techniques for Predictive Maintenance in Cooler Condition," *2022 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Penang, Malaysia, 2022, pp. 1-4, doi: 10.1109/ISPACS57703.2022.10082814.