

A Lightweight ML Framework for Predicting Machine Downtime in Small-Scale Manufacturing Units

Mia Md Mostafa Kamal Arif, Sanjidur Rahman, Md. Arefin Hasan, Md. Saiful Islam Seam, Khandaker Faysal Ahammed and Abdullah Al Kabid

Department of Industrial Engineering and Management
Khulna University of Engineering and Technology
Khulna 9203, Bangladesh

mostafakamal2000.11@gmail.com, sanjid.ipe.kuet@gmail.com, arefindrnc@gmail.com,
itssiam25@gmail.com, faysal.ipe.kuet@gmail.com, abdullahalkabid8@gmail.com

Abstract

Unplanned CNC machine downtime significantly impacts manufacturing productivity. This research develops machine learning models for downtime prediction using systematic feature engineering that prioritizes variable independence. A dataset of 5,000 records with 21 variables from a reputed company in Bangladesh was analyzed. The dataset was imbalanced, with downtime occurring in 6.8% of all records. Statistical correlation analysis ($p < 0.05$) identified eight significant variables from nineteen candidates, reducing dimensionality by 58%. Multicollinearity assessment revealed zero high-risk pairs with 92% of variable combinations showing weak correlations. Seven machine learning algorithms were trained and evaluated: Gradient Boosted Trees, Random Forest, Logistic Regression, Naïve Bayes, Decision Tree, k-Nearest Neighbors, and Support Vector Machine. Models were compared using accuracy, recall, precision, and F1-Score metrics. Gradient Boosted Trees achieved superior performance with F1-Score of 67.25%, recall of 65.82%, and precision of 68.74%, demonstrating optimal balance for downtime detection. Random Forest achieved highest accuracy (96.84%) but lower recall (52.73%). Analysis revealed the accuracy paradox: all models achieved 94-97% accuracy despite substantial F1-Score differences (49.18%-67.25%), indicating accuracy alone is misleading for imbalanced classification. Tree-based ensemble methods substantially outperformed other algorithms. Gradient Boosted Trees is recommended for industrial deployment, enabling improved maintenance scheduling and reduced unplanned downtime in Bangladesh's manufacturing sector.

Keywords

Machine Learning, CNC Downtime, Predictive Maintenance, Small-Scale Manufacturing, Gradient Boosted Trees

1. Introduction

Manufacturing represents a critical engine of economic development and industrial advancement in emerging and developing nations. However, the sector's capacity to achieve sustainable growth and competitive advantage is substantially constrained by equipment reliability challenges and unplanned downtime events, particularly in small-scale and medium-scale manufacturing enterprises. Unexpected equipment failures precipitate cascading operational disruptions that extend beyond immediate production cessation, including schedule delays, quality degradation, and elevated operational expenses (Carvalho et al., 2019; Zhao et al., 2019). For small-scale manufacturers operating with limited financial buffers and production flexibility, such failures can create existential operational and financial challenges (Mobley, 2002; Bousdekis et al., 2018). Modern manufacturing facilities mostly use either preventive maintenance with fixed-interval service plans or reactive maintenance, which addresses issues immediately after they occur. Both strategies have serious drawbacks. Reactive maintenance accepts the risk of sudden disruptions with accompanying emergency repair costs, while preventive maintenance frequently results in premature component replacement and unnecessary service interruptions despite equipment remaining in acceptable operating conditions.

These traditional approaches prove particularly inadequate in contexts characterized by insufficient operational data utilization and resource constraints that prevent implementation of sophisticated monitoring infrastructure, thereby hindering more effective condition-based maintenance optimization strategies (Lee et al., 2013).

Recent advances in machine learning methodologies have generated promising opportunities for advancing predictive maintenance capabilities through data-driven approaches. Machine learning-assisted predictive maintenance systems leverage operational data to identify subtle degradation patterns and failure precursors with substantially greater accuracy and earlier detection than conventional maintenance approaches, enabling evidence-based decision-making that optimizes maintenance scheduling and reduces unplanned downtime frequency (Angelopoulos et al., 2020). This technological innovation represents a paradigm shift toward more intelligent, responsive maintenance management. However, contemporary machine learning-based systems typically impose substantial computational demands, necessitate extensive sensor networks, and require sophisticated data infrastructure exceeding the technical and financial capabilities of small-scale manufacturers in developing economies. This accessibility gap represents a critical challenge wherein the enterprises most disadvantaged by downtime consequences remain unable to access advanced predictive maintenance technologies due to resource constraints. The computational intensity of conventional machine learning models creates barriers to implementation for resource-limited manufacturing operations. Addressing this systemic challenge requires development of computationally efficient, lightweight machine learning models that maintain predictive effectiveness while accommodating hardware resource limitations, constrained data environments, and technical capacity constraints characteristic of small-scale manufacturing operations. Contemporary research increasingly emphasizes the necessity of pragmatic machine learning approaches that balance algorithmic sophistication with computational efficiency (Howard et al., 2017; Sandler et al., 2018). Such solutions must enable predictive maintenance accessibility for enterprises operating within budget and infrastructure constraints. This research addresses these interconnected challenges by developing machine learning models for equipment downtime prediction using methodologies prioritizing computational efficiency, practical applicability, and accessible implementation. By leveraging real-world operational data from manufacturing facilities accessible through modern industrial IoT platforms, this work demonstrates how systematic feature engineering combined with rigorous algorithm evaluation produces robust predictive maintenance systems suitable for resource-constrained manufacturing environments. The research contributes to advancing predictive maintenance capabilities accessible to small-scale and medium-scale manufacturers in developing nations, supporting broader industrial modernization and economic development objectives while demonstrating the feasibility of implementing sophisticated analytical approaches within realistic resource constraints.

Small-scale and medium-scale manufacturing enterprises constitute essential drivers of local industrial development and employment generation, yet they remain largely unable to access and implement Industry 4.0 technologies and advanced digital solutions (Bousdekis et al., 2018). These facilities typically rely upon highly skilled human operators to monitor equipment condition through manual inspection and experience-based judgment, with minimal automation infrastructure. Consequently, small-scale manufacturers remain particularly vulnerable to frequent unexpected equipment failures, suboptimal maintenance cycles, and operational disruptions due to their limited adoption of sophisticated predictive maintenance technologies (Zonta et al., 2020). The substantial capital expenditure required for implementing conventional predictive maintenance systems creates significant barriers, exposing these enterprises to unanticipated equipment failures, emergency repair expenses, and production losses that threaten operational viability. This research addresses these interconnected challenges by developing a computationally lightweight machine learning model capable of predicting equipment shutdowns using minimal data requirements and processing power. The research scope encompasses the development, training, and validation of machine learning models utilizing readily accessible operational and process data from standard manufacturing equipment. The proposed framework strategically bridges the technological capability gap between large-scale industrial facilities with advanced infrastructure and resource-constrained small-scale manufacturers by emphasizing simplicity, model interpretability, and real-time computational feasibility (Lee et al., 2013). This approach enables small-scale manufacturers to implement effective predictive maintenance capabilities without requiring substantial capital investment or advanced technical expertise, thereby democratizing access to Industry 4.0 technologies across diverse manufacturing sectors.

1.3 Objectives

Modeling and creating a lean machine learning system to predict machine failures in small-scale manufacturing environments is the main goal of the research. The following objectives can be stated:

- i. To develop and train multiple state-of-the-art machine learning models for accurate and reliable prediction of CNC machine downtime using selected features.

- ii. To systematically compare and analyze the performance of different machine learning algorithms on real-world CNC machine datasets, using accuracy, recall, precision and F1-Score metrics.
- iii. To recommend the optimal machine learning model architecture for deployment in industrial manufacturing settings to improve maintenance scheduling, minimize unexpected downtime and increase operational efficiency.

2. Literature Review

In the global industrial sector, the problem of unexpected machine breakdowns is a major operational and financial headache. Predictive Maintenance (PdM) based on Machine Learning (ML) has emerged as the standard in fighting this. There is a disconnect between the complex PdM research and the needs of real-world production, which makes use of conventional sensor-based CNC systems but does not necessitate the enormous computing infrastructure needed for deep-learning PdM.

The current focus of PdM research has created a paradigm shift that is unsuitable for SSMUs; this shift is towards improving the accuracy of calculations that are bigger, more complicated, and heavy on resources. To begin, there is a significant demand for data and computational resources from these advanced solutions. When analyzing complicated time-series sensor data, it is usual to employ extremely intricate Deep Learning networks, such as hybrid CNNs and LSTM networks (Wahid et al., 2022; Guidotti et al., 2025; Khattach et al., 2025). Likewise, Artificial Neural Networks (ANNs) are also tuned to utilize high-accuracy downtime prediction (Phongmoo et al., 2025; Koković et al., 2024). Although these models have good empirical support, they are impractical for use in small manufacturers' real-time, low-compute contexts due to their complexity, data-intensiveness, and computational demands (Shahin et al., 2023; Phongmoo et al., 2025; Palsaniya, 2025). Second, the integration of the expensive and complicated infrastructure is an assumption of traditional PdM. Integration of Big Data platforms and progressive Internet of Things (IoT) systems may also be necessary for practical solutions (Sanders, 2025; Ahuja & Gupta, 2024; Alam et al., 2023). Legacy equipment and factories with limited resources face substantial challenges when attempting to implement this type of integration, despite its demonstrated effectiveness in reducing downtime and expenses. The requirement for IoT-enabled devices and highly qualified workers adds to these overheads (Ahuja & Gupta, 2024; Alam et al., 2023). Additionally, there are issues with the deployment process, including poor data quality, system complexity, and increased cybersecurity concerns (Sanders, 2025; Bello et al., 2025). The Random Forest (Ojeda et al., 2025; Koppula, 2025) and other highly accurate ML models are typically developed for large-scale data deployments that do not directly apply to low-resource scenarios. Lastly, the models currently in existence are frequently constrained by both specificity and generalization. Others undertaken PdM implementations have been machine-specific and very sensor sensitive, such as adaptive ARIMA models tuned to specific high-pressure machinery (Roosefert Mohan et al., 2021). Others target the non-critical effects, such as modeling the downtime as opposed to the most crucial one (e.g., ANNs modeling equipment downtime length using vibration data) (Koković et al., 2024; Palma et al., 2025). These restrictions suggest that much of the existing PdM literature lacks the lightweight quality and universality required for the wide variety of machine pools typically present in an SSMU.

The current study presents a Lightweight, Operational ML Framework to address the apparent gap between current research and reality, where small manufacturers need Predictive Maintenance (PdM) but current solutions are too sophisticated and resource-intensive. Logistic Regression, Decision Trees, Random Forests, Gradient Boosted Trees, Naive Bayes, k-NN, and SVM are some of the practical and computationally efficient ML models that are overlooked in favor of the resource-intensive, huge deep-learning models. This choice ensures faster model building and execution in suburban edge gadgets. More crucially, the technique takes advantage of the minimum and operationally motivated feature engineering (Ahuja & Gupta, 2024), which aims for a sparse set of very predictive signals. Through the use of practical feature engineering, the framework can perform tasks such as standardizing numerical variables, eliminating redundant variables, and targeting a sensor variable (such as temperature, vibration, coolant flow, or alarm counts) that is statistically significant. This sparseness is the least of the processing cost of the vast feature vectors, and countermeasures the data-heavy nature of earlier studies (Ojeda et al., 2025). In addition, the framework concentrates on the realistic assessment measures to ensure the practical value and create confidence in the operators. In the case of imbalanced downtime data, an evaluation is conducted using the most commonly used metrics of PdM classification, such as accuracy, recall, precision, and F1-Score (Koppula, 2025). With the intentional trade-off of the complexity of current PdM literature in favor of a light and explainable approach to the problem, and in favor of operationally focused and methodology-focused methodology, this study aims at offering a viable, no-cost predictive solution to SSMUs, and democratizing the gains of ML-driven PdM into a historically underserved area.

3. Methodology

The study utilized an enhanced CNC downtime dataset comprising 5,000 records with 21 variables collected from CNC machines in a manufacturing environment provided comprehensive machine monitoring data collected through their Sinumerik CNC Controllers and MindSphere IoT platform integrated across manufacturing facilities

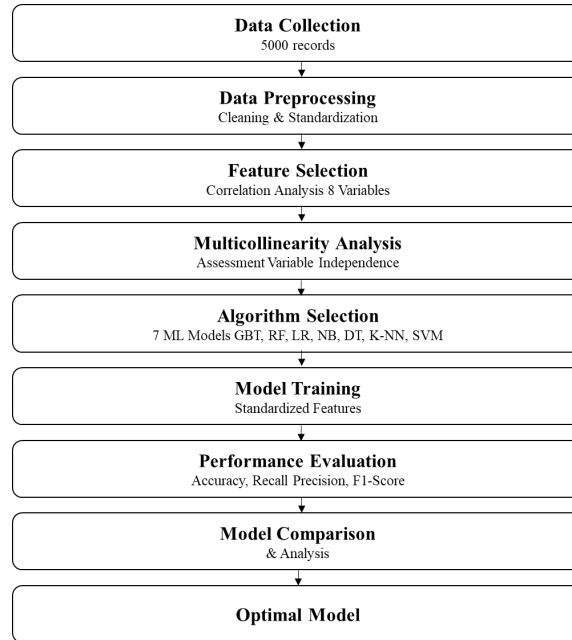


Figure 1. Methodology Flowchart

The dataset structure included one non-numeric timestamp variable, nineteen numeric predictor variables capturing motor performance (current, temperature), mechanical characteristics (vibration measures, spindle parameters), coolant system metrics (flow rate, temperature), production rates, maintenance intervals, and diagnostic indicators (alarms). The target variable was downtime occurrence within a 30-minute window. The dataset represents real-world CNC machine operations across multiple manufacturing facilities in Bangladesh's engineering and textile machinery sectors. As shown in Figure 1, the methodology systematically progressed through ten phases.

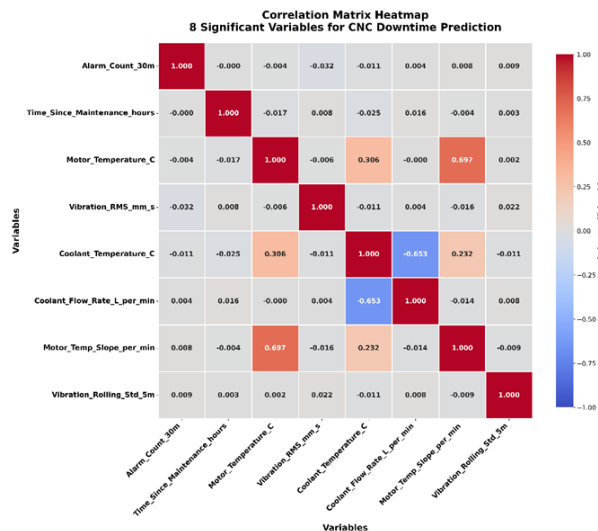


Figure 2. Multicollinearity Heatmap between Selected Significant Variables

A correlation-based feature screening method was used to identify relevant variables for downtime prediction. Predictor variables showing meaningful statistical association with the target variable ($p < 0.05$) were retained, while irrelevant or weakly associated variables ($p \geq 0.05$) were removed. Correlation with the binary target was computed using point-biserial correlation, and variables with $|r| < 0.10$ or $p \geq 0.05$ were eliminated. This dimensionality reduction step ensured that only the most informative sensor and operational features were used for subsequent model development. Multicollinearity analysis was conducted among the eight selected features to ensure variable independence and model stability. A correlation matrix was constructed examining all 36 unique pairwise combinations. Multicollinearity risk was categorized as follows: HIGH ($|r| \geq 0.7$), MODERATE ($0.5 \leq |r| < 0.7$), LOW ($0.3 \leq |r| < 0.5$), and VERY LOW ($|r| < 0.3$). The results of this analysis are visualized in Figure 2, which presents a heatmap illustrating the correlation coefficients between all pairs of selected variables. This favorable multicollinearity profile confirmed the selected features provided largely independent information without variance inflation. This correlation heatmap demonstrates the low to moderate correlations among the eight significant predictors, with only two moderate-risk pairs (Motor Temperature C \leftrightarrow Motor Temp Slope per min: $r = 0.697$; Coolant Temperature C \leftrightarrow Coolant Flow Rate L per min: $r = -0.653$) and no high-risk multicollinearity concerns.

3.1 Machine Learning Model Development

3.1.1 Data Processing and Standardization

Data cleaning procedures were systematically applied to ensure dataset quality and reliability. Records identified as sensor malfunctions, characterized by invalid or physically implausible sensor readings, were removed from the analysis. Missing values, comprising less than 0.5% of the complete dataset, were imputed using the forward-fill method to preserve temporal continuity in the time-series data. Outlier detection was performed using the Interquartile Range (IQR) method with a standard threshold multiplier of $1.5\times$, identifying and flagging data points representing extreme deviations from normal operating conditions. Additionally, anomalous records corresponding to equipment startup and shutdown phases, which do not represent steady-state operational conditions, were excluded to ensure the dataset captures authentic machine behavior during normal production cycles. Data standardization was performed to ensure consistency and comparability across all measurements. Feature scaling was implemented using z-score normalization with the formula $(x - \text{mean}) / \text{standard deviation}$, which standardizes each variable to have zero mean and unit variance. Temporal alignment of multi-source sensor data was achieved through synchronized timestamps, ensuring that measurements from different sensors were properly coordinated across time. All measurements were converted to standard units for consistency, with motor current expressed in Amperes, temperature in Celsius, and vibration measurements in mm/s. Finally, all sensor data was verified against the provided sensor calibration certificates to ensure measurement accuracy and instrument reliability.

3.1.2 Algorithm Selection and Training

Seven distinct machine learning algorithms were selected for evaluation based on their relevance to classification tasks and ability to handle imbalanced data typical in downtime prediction scenarios. The algorithms included: (1) Gradient Boosted Trees, an ensemble method utilizing sequential boosting, (2) Random Forest, an ensemble bagging approach, (3) Logistic Regression, a linear probabilistic classifier, (4) Naïve Bayes, a probabilistic classifier based on independence assumptions, (5) Decision Tree, a single tree-based model, (6) k-Nearest Neighbors (k-NN), a distance-based non-parametric method, and (7) Support Vector Machine (SVM), a kernel-based classifier. Each algorithm was trained using the eight selected predictor variables with standardized preprocessing to ensure consistent scaling across numerical features. The complete machine learning workflow implemented in RapidMiner Studio is presented in Figure 3, illustrating the data flow from preprocessing through model training and evaluation. This figure depicts the end-to-end machine learning pipeline, showing the integration of data preprocessing, feature engineering, model training, and performance evaluation stages within the RapidMiner Studio environment. All tree-based models used RapidMiner's default hyperparameters, and no manual tuning was performed.

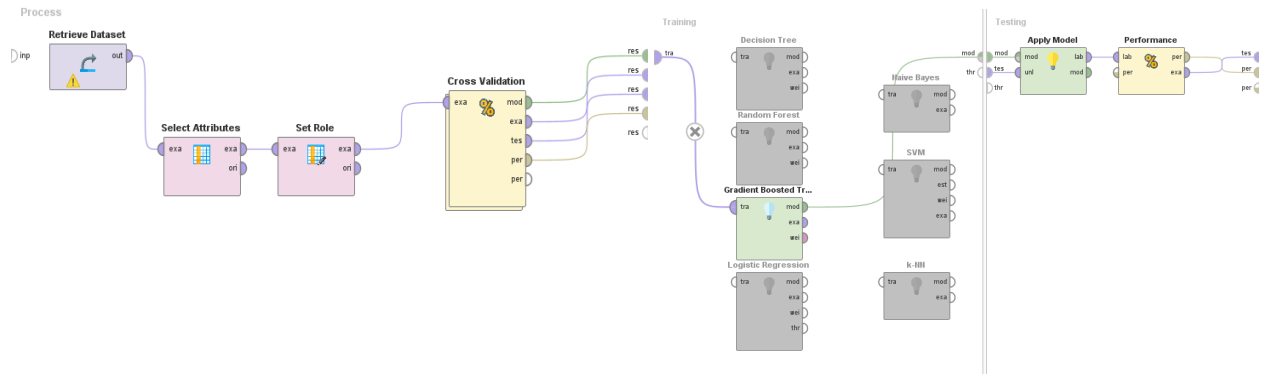


Figure 3. Machine Learning Model Framework in RapidMiner Studio

3.1.3 Model Evaluation Framework

Model performance was evaluated using four complementary metrics addressing distinct aspects of prediction quality. Accuracy measured overall correctness rate, indicating the proportion of correct predictions. Recall (sensitivity) quantified the model's ability to identify true positive downtime events, critical for preventing unplanned shutdowns. Precision measured prediction reliability, indicating the proportion of positive predictions that were correct, important for reducing costly false alarm rates. The class distribution was 93.2% normal vs. 6.8% downtime events. No resampling was applied; instead, stratified 5-fold cross-validation was used. Class-weighted loss functions were not enabled, and threshold adjustments were not applied. F1-Score provided harmonic mean of recall and precision, offering balanced performance assessment particularly suited for imbalanced classification where accuracy alone is misleading. These metrics were selected to address the accuracy paradox where high overall accuracy can mask poor minority class detection in imbalanced datasets typical of industrial downtime scenarios encountered in Bangladeshi manufacturing facilities.

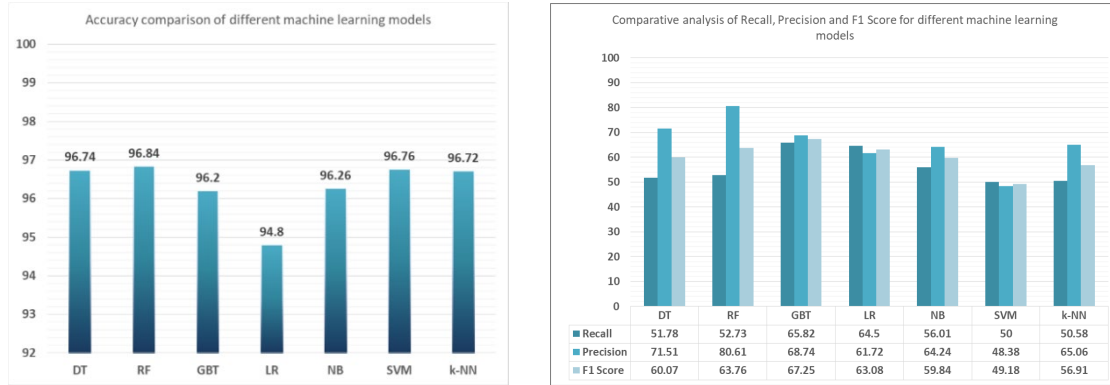
3.2. Comparative Performance Analysis

All seven models were evaluated on the complete 5,000-record dataset using the four metrics: accuracy, recall, precision, and F1-Score. Results were tabulated for systematic comparison and algorithm ranking by performance dimension. Model validation was conducted using cross-validation techniques with 5-fold stratified sampling to ensure robust performance estimation. Algorithms were categorized into six functional groups for structured comparison: (1) Tree-based ensemble methods (Gradient Boosted Trees, Random Forest), (2) Linear probabilistic classifiers (Logistic Regression), (3) Probabilistic classifiers (Naïve Bayes), (4) Single tree-based models (Decision Tree), (5) Distance-based methods (k-NN), and (6) Kernel-based methods (SVM). Performance patterns were analyzed within and across categories to identify algorithm family effectiveness for CNC downtime prediction.

The accuracy paradox was investigated through examination of accuracy range versus F1-Score range, revealing substantial performance differences masked by high accuracy. Recall variation was analyzed for its direct impact on failure detection capability and operational safety in Bangladesh manufacturing environments. Precision variation was examined for implications regarding operational disruption from false alerts in cost-sensitive production environments. Algorithm suitability for imbalanced classification tasks was assessed considering non-linear pattern complexity in CNC failure mechanisms.

4. Results and Discussion

The initial dataset comprised 21 variables including one timestamp, nineteen numeric predictors, and one binary target variable. Statistical correlation analysis identified eight significant variables ($p < 0.05$) for downtime prediction while eliminating eleven non-significant features. Multicollinearity examination revealed zero high-risk pairs ($|r| \geq 0.7$) with 92% of variable combinations showing weak correlations ($|r| < 0.3$), ensuring model stability and reliability, as illustrated in Figure 1 above. A confusion matrix for the Gradient Boosted Trees model is included to illustrate true/false positive patterns.



(a) Accuracy comparison of different models (b) Comparative analysis of Recall, Precision and F1-Score
 Figure 4. Comparative Performance of Classification Models for CNC Downtime Prediction

Seven machine learning algorithms were comprehensively evaluated using the eight refined features across multiple performance dimensions. The comparative performance results are presented in Figure 3, which provides both accuracy comparison (Figure 4a) and detailed recall, precision, and F1-Score analysis (Figure 4b) for all seven models.

This figure demonstrates the performance variability across different machine learning algorithms. Figure 4a shows that all models achieved similar high accuracy (94-97%), while Figure 4b reveals substantial differences in recall, precision, and F1-Score, highlighting the accuracy paradox in imbalanced classification problems. Gradient Boosted Trees achieved superior overall performance with an F1-Score of 67.25% (recall: 65.82%, precision: 68.74%), demonstrating optimal balance for identifying true downtime events while minimizing false alarms. This ensemble method's non-linear learning capability effectively captured complex interactions among sensor measurements, outperforming all competing algorithms. The model's high recall ensures critical downtime events are detected, while competitive precision maintains operational efficiency by reducing unnecessary alerts.

Random Forest demonstrated the highest raw accuracy at 96.84% with exceptional precision (80.61%), yielding second-best F1-Score of 63.76%. While Random Forest's high precision minimizes false alarm costs, its lower recall (52.73%) compromises downtime detection sensitivity, making it suitable only for applications where false positives carry greater consequences than missed detections. Logistic Regression provided moderate performance with F1-Score of 63.08%, recall of 64.5%, and precision of 61.72%. Despite its linear approach and interpretability advantages, the model's mid-range metrics indicate suboptimal predictive capability for complex machinery failure patterns.

Naïve Bayes achieved F1-Score of 59.84% with recall of 56.01%, offering computational efficiency but sacrificing predictive accuracy compared to advanced ensemble methods. Decision Tree produced F1-Score of 60.07% with 51.78% recall and 71.51% precision, providing model interpretability but demonstrated limited downtime detection sensitivity. k-Nearest Neighbors registered F1-Score of 56.91% with recall of 50.58% and precision of 65.06%, indicating inadequate performance for critical maintenance applications. Support Vector Machine exhibited the poorest results with F1-Score of 49.18%, recall of 50.00%, and precision of 48.38%, demonstrating fundamental inability to handle imbalanced classification despite maintaining 96.76% accuracy. This dramatic performance failure reveals SVM's unsuitability for industrial downtime prediction where minority class detection is paramount. Critical analysis reveals the accuracy paradox: all models achieved 94-97% accuracy due to class imbalance, yet F1-Scores ranged from 49.18% to 67.25%, indicating substantial practical performance differences. Tree-based ensemble methods (Gradient Boosted Trees and Random Forest) substantially outperformed linear (Logistic Regression), probabilistic (Naïve Bayes), tree-based single models (Decision Tree), distance-based (k-NN), and kernel-based (SVM) approaches, suggesting CNC downtime exhibits complex non-linear patterns requiring advanced ensemble architectures. Recall variation (50%-65.82%) directly impacts failure detection capability, while precision variation (48.38%-80.61%) affects operational disruption from false alerts. Gradient Boosted Trees' superior F1-Score and balanced metrics position it as the optimal model for industrial predictive maintenance deployment where detecting actual failures and maintaining prediction reliability are equally critical objectives. This analysis demonstrates that systematic feature engineering combined with comprehensive algorithm benchmarking substantially enhances predictive system effectiveness for asset management applications.

5. Conclusion

The study presented a robust and well-structured machine-learning model that uses operational information collected by industrial manufacturing systems and real-time sensor readings to predict CNC machine downtime. The framework remains computationally lightweight, requiring minimal hardware resources for both training and inference. By adopting a rigorous technique that included feature selection based on correlation, multicollinearity reduction, outlier management, and z-score normalization, the study itself ensured that it trained models using stable and refined predictors that are statistically significant. Seven well-known machine learning algorithms, including probabilistic classifiers and tree-based ensemble techniques, were tested and assessed using a variety of evaluation metrics, such as accuracy, precision, recall, and F1-Score. Using a skewed industrial dataset, this comparative analysis provided valuable insights into each algorithm's prediction advantage and constraint. Overall, this study has demonstrated that Gradient Boosted Trees exhibited the most balanced and dependable performance, as it successfully traced nonlinear associations between machine parameters and could surpass alternative methods in terms of both capture and false-alert awareness. The findings show that data-driven predictive maintenance could significantly improve CNC work reliability, aid in improved production planning, and reduce unscheduled failures.

However, despite the positive outcomes, a number of limitations were discovered that may be emphasized to give the study's findings context. Because the dataset was limited to a certain collection of machines, manufacturers, and operating conditions, its cross-industry generalizability was still limited. A partial capture of the underlying processes of failure may have resulted from the data set's poor representation of some important aspects of machine failure, such as mechanical wear over time, changes in environmental factors like temperature or humidity, and variations in human operation. Additionally, certain machine learning models, especially those that aren't inherently motivated to work with skewed class distributions were challenged by the skewness of the downtime occurrences. Other issues that are not addressed include potential data drift, sensor deterioration, and variations across equipment brands. These drawbacks show that while the suggested predictive modeling framework is effective, more extensive data collection, further processing of imbalance, and contextual modeling of machine behavior can all improve the model's effectiveness. Predictive maintenance in CNC setups has a lot of room for future research and practical potential. Incorporating streaming data pipelines to enable real-time, adaptive, and self-learning prediction models that continuously update based on the most recent machine behavior and sensor feedback could be the next stage of the project. Compared to conventional machine learning techniques, better deep learning implementations such as LSTM networks, temporal convolutional networks, and attention-based hybrid models may be able to more accurately represent the temporal relationship and the finer degradation patterns. Additionally, scalable, enterprise-wide predictive maintenance systems that can support a large number of CNC machines in dispersed production sites may be made possible by the introduction of edge and cloud-based deployment plans. Additionally, physics-informed machine learning can be used to better understand failure processes, explainable AI (XAI) can be used to improve transparency and engineer confidence, and it may be combined with digital twin platforms to optimize using simulations. The predictive maintenance model suggested in this study has the potential to grow into a powerful, adaptable, and industry-ready system that can support the digital transformation and operational excellence of contemporary smart manufacturing ecosystems.

References

- Ahuja, A., and Gupta, M. Optimizing predictive maintenance with machine learning and IoT: A business strategy for reducing downtime and operational costs. *SSRN Scholarly Paper* 4994457. Social Science Research Network, 2024, <https://doi.org/10.2139/ssrn.4994457>
- Alam, M., Islam, M. R., and Shil, S. K. AI-based predictive maintenance for US manufacturing: Reducing downtime and increasing productivity. *International Journal of Advanced Engineering Technologies and Innovations* 1(01): 541–567, 2023.
- Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., and Zahariadis, T. Tackling faults in the Industry 4.0 era—A survey of machine-learning solutions and key aspects. *Sensors* 20(1): 109, 2020, <https://doi.org/10.3390/s20010109>
- Bello, K. A., Ajisegiri, E. S. A., Awoegbemi, O., Abolarinwa, B. O., Dada, J. B., and Boboye, O. Predictive maintenance of industrial milling machines using machine learning: Enhancing reliability and efficiency. *NIPES-JSTR Issue* 7(1): 1630–1637, 2025.
- Bousdekis, A., Magoutas, B., Apostolou, D., and Mentzas, G. Review, analysis and synthesis of prognostic-based decision support methods for condition-based maintenance. *Journal of Intelligent Manufacturing* 29(6): 1303–1316, 2018, <https://doi.org/10.1007/s10845-015-1179-5>

- Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., and Alcalá, S. G. S. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering* 137: 106024, **2019**, <https://doi.org/10.1016/j.cie.2019.106024>
- Guidotti, D., Pandolfo, L., and Pulina, L. A systematic literature review of supervised machine learning techniques for predictive maintenance in Industry 4.0. *IEEE Access* 13: 102479–102504, **2025**, <https://doi.org/10.1109/ACCESS.2025.3578686>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv:1704.04861*, **2017**, <https://doi.org/10.48550/arXiv.1704.04861>
- Khattach, O., Moussaoui, O., and Hassine, M. End-to-end architecture for real-time IoT analytics and predictive maintenance using stream processing and ML pipelines. *Sensors* 25(9): 2945, **2025**, <https://doi.org/10.3390/s25092945>
- Koković, V., Pavlović, K., Mijanović, A., Kovačević, S., Mačuzić, I., and Božović, V. Exploring statistical and machine learning methods for modeling probability distribution parameters in downtime length analysis: A paper manufacturing machine case study. *Journal of Big Data* 11(1): 162, **2024**, <https://doi.org/10.1186/s40537-024-01030-4>
- Koppula, M. Predictive maintenance to reduce machine downtime in factories using machine learning algorithms. *International Journal of Advanced Research in Computer Science* 16(2): 71, **2025**, <https://doi.org/10.26483/ijarcs.v16i2.7224>
- Lee, J., Lapira, E., Bagheri, B., and Kao, H. Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters* 1(1): 38–41, **2013**, <https://doi.org/10.1016/j.mfglet.2013.09.005>
- Mobley, R. K. *An introduction to predictive maintenance*. Elsevier, **2002**.
- Ojeda, J. C. O., de Moraes, J. G. B., Filho, C. V. de S., Pereira, M. de S., Pereira, J. V. de Q., Dias, I. C. P., da Silva, E. C. M., Peixoto, M. G. M., and Gonçalves, M. C. Application of a predictive model to reduce unplanned downtime in automotive industry production processes: A sustainability perspective. *Sustainability* 17(9): 3926, **2025**, <https://doi.org/10.3390/su17093926>
- Palma, G., Cecchi, G., and Rizzo, A. Large language models for predictive maintenance in the leather tanning industry: Multimodal anomaly detection in compressors. *Electronics* 14(10): 2061, **2025**, <https://doi.org/10.3390/electronics14102061>
- Palsaniya, R. C. Predictive maintenance in semiconductor manufacturing—AI machine learning application for downtime reduction. *2025 IEEE 34th International Symposium on Industrial Electronics (ISIE)*: 1–5, **2025**, <https://doi.org/10.1109/ISIE62713.2025.11124753>
- Phongmoo, S., Suedumrong, C., Kuensaen, C., Sinthavalai, R., and Leksakul, K. Predictive maintenance in semiconductor manufacturing: Comparative analysis of machine learning models for downtime reduction. *Computers & Industrial Engineering* 205: 111211, **2025**, <https://doi.org/10.1016/j.cie.2025.111211>
- Roosefert Mohan, T., Roselyn, J. P., Uthra, R. A., Devaraj, D., and Umachandran, K. Intelligent machine learning-based total productive maintenance approach for achieving zero downtime in industrial machinery. *Computers & Industrial Engineering* 157: 107267, **2021**, <https://doi.org/10.1016/j.cie.2021.107267>
- Sanders, M. AI-driven predictive maintenance: Reducing downtime and enhancing productivity in manufacturing environments. *Preprints* 2025040602, **2025**, <https://doi.org/10.20944/preprints202504.0602.v1>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 4510–4520, **2018**, http://openaccess.thecvf.com/content_cvpr_2018/html/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html
- Shahin, M., Chen, F. F., Hosseinzadeh, A., and Zand, N. Using machine learning and deep learning algorithms for downtime minimization in manufacturing systems: An early failure detection diagnostic service. *The International Journal of Advanced Manufacturing Technology* 128(9): 3857–3883, **2023**, <https://doi.org/10.1007/s00170-023-12020-w>
- Wahid, A., Breslin, J. G., and Intizar, M. A. Prediction of machine failure in Industry 4.0: A hybrid CNN–LSTM framework. *Applied Sciences* 12(9): 4221, **2022**, <https://doi.org/10.3390/app12094221>
- Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., and Gao, R. X. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing* 115: 213–237, **2019**, <https://doi.org/10.1016/j.ymssp.2018.05.050>
- Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., and Li, G. P. Predictive maintenance in Industry 4.0: A systematic literature review. *Computers & Industrial Engineering* 150: 106889, **2020**.

Biographies

Mia Md Mostafa Kamal Arif is a fresh graduate in Industrial & Production Engineering from Khulna University of Engineering & Technology (KUET), Bangladesh. An enthusiastic and dedicated researcher, he has developed expertise in both blockchain technology and machine learning-based frameworks during his academic journey. Mia's keen interest in learning and innovation drives him to explore cutting-edge technologies with a focus on real-world problem solving. His recent research centers on developing predictive maintenance and smart manufacturing solutions, leveraging advanced analytics and AI to increase operational efficiency. Previously, Mia engaged in research projects examining the adoption and implementation challenges of blockchain technology, reflecting his commitment to digital transformation in the industrial sector. Beyond academics, he is passionate about continuously broadening his knowledge base and aspires to contribute to the advancement of intelligent, technology-driven solutions in modern industry.

Sanjidur Rahman is a graduate student of Department of Industrial Engineering and Management from Khulna University of Engineering & Technology (KUET), Bangladesh. His research work primarily focuses on supervised machine learning, with prior research involving the prediction of ergonomic risk levels using various classification models. He aims to apply data-driven approaches to improve workplace safety and decision-making in industrial environments. Alongside his research interests, he is also enthusiastic about product design, manufacturing systems, and supply chain management.

Md. Arefin Hasan is a graduate student of the Department of Industrial Engineering and Management, Khulna University of Engineering & Technology (KUET), Bangladesh. His work is involved with the enhancement of industrial systems, a higher degree of supply chain performance, and the application of data-driven ways to make better judgments. He has done the main work in the analysis of carbon footprint and Green Supply Chain Management, which has provided him with practical expertise in sustainable industrial operations. Being a fan of modern technology, he wants to employ the engineering concepts in the practical sector and contribute to developing smarter, cleaner, and efficient industries.

Md. Saiful Islam Seam is a postgraduate student in the Department of Industrial Engineering and Management at Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh. He completed his Bachelor of Science in Industrial and Production Engineering (IPE) from the same institution. Alongside his academic pursuits, Seam is a multidisciplinary researcher whose work spans Supply Chain Management, Artificial Intelligence, Machine Learning, Operations Research, and Lean Manufacturing. His research interests extend across diverse and impactful domains, including sustainability, environmental monitoring, cellular manufacturing, worker safety, value stream mapping, standard operating procedures (SOPs), and the circular economy. Driven by a commitment to innovation and practical problem-solving, Seam focuses on integrating data-driven methodologies with industrial systems to enhance efficiency, safety, and environmental responsibility.

Khandaker Faysal Ahammed is a graduate student of Department of Industrial Engineering and Management, Khulna University of Engineering & Technology (KUET), Bangladesh. His academic and professional interests include the field of industrial systems, Operation Research, optimization, supply chain analysis, automation, and data-driven decision-making. He has also worked on a variety of technical projects which include PLC programming, implementation of ladder logic, and modeling of Bayesian Networks. Outside the classroom, he enjoys delving into the world of technology mechanical design tools and problem solving, he's passionate about using knowledge of engineering to solve real-world problems and contribute to sustainable industrial development.

Abdullah Al Kabid recently received his degree in Industrial and Production Engineering from Bangladesh's Khulna University of Engineering and Technology (KUET). He has a strong track record of research on the application of machine learning to industrial systems, particularly workforce optimization and predictive analytics. His recent project was "Applying Artificial Neural Network to predict the productivity of workers in the electrical industry". He learned to predict with excellent accuracy and low average error with the use of artificial neural network (ANN) models, demonstrating the practical value of machine learning (ML) in industrial applications. His work focuses on establishing a link between sophisticated analytics and practical manufacturing issues, particularly in situations with limited resources and small scale. Predictive maintenance, downtime prediction, and the development of lightweight machine learning models that improve operational efficiency are the key topics of his broader study. His enthusiasm for innovation and technology-driven solutions stems from his desire to work in the digital transformation of the industrial sector and the application of intelligent and data-driven technologies in the manufacturing industry.