Proceedings of the 8th IEOM Bangladesh International Conference on Industrial Engineering and Operations Management, Dhaka, Bangladesh, December 20-21, 2025

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/BA08.20250158
Published: December 20, 2025

A Comparative Analysis of Machine Learning Models for Predictive Maintenance in the Ready-Made Garments (RMG) Industry

Samiur Rahman Wasi
Department of Computer Science and Engineering
International Islamic University Chittagong, Bangladesh
Chattogram, Bangladesh
samwaseee@gmail.com

Abstract

In Bangladesh's vital ready-made garment (RMG) sector, unplanned machine downtime is a major cause of operational inefficiencies and financial loss. A solution is provided by Predictive Maintenance (PdM), a crucial element of Industry 4.0, which uses sensor data to predict machine faults before they happen. Using the benchmark AI4I 2020 Predictive Maintenance dataset, this paper conducts a rigorous comparative analysis of three baseline models—Logistic Regression, Random Forest, and XGBoost—for machine failure prediction. To ensure a robust and reliable evaluation for this high-imbalance problem, models were assessed using 5-fold stratified cross-validation, focusing on their average F1-Score, Precision, and Recall. The findings unequivocally demonstrate that the XGBoost classifier (Avg. F1-Score: 0.731) provides the most balanced and reliable performance, significantly outperforming Random Forest (Avg. F1-Score: 0.571) and Logistic Regression (Avg. F1-Score: 0.235). Furthermore, a Business Impact Analysis (BIA) translates this performance into a practical metric, estimating the XGBoost model could reduce downtime-related costs by 70.7%. The model's primary predictors are logical physical elements, with torque and tool wear being the most important, according to an explainability (SHAP) study. Rather than proposing a new algorithm, this paper builds a methodological bridge from benchmark datasets and industrial application. We provide a practical, end-to-end framework that: (1) identifies the most reliable baseline model (XGBoost) using robust 5-fold cross-validation, (2) translates its performance into a clear financial case using a Business Impact Analysis (BIA), and (3) validates its logic using SHAP explainability. This provides a direct, data-driven recommendation for RMG factories.

Keywords

Predictive Maintenance, Machine Learning, RMG, XGBoost, Industry 4.0, Business Impact Analysis, Explainable AI (XAI).

1. Introduction

With millions of workers and a substantial GDP contribution, the Ready-Made Garments (RMG) sector is the foundation of Bangladesh's economy. Operational efficiency is not just a goal but also a vital survival element in this high-volume, low-margin, and time-sensitive industry. Unplanned machine downtime is a major threat to this efficiency. The expense is instant and exponential when a crucial piece of equipment, such as an automated sewing machine, cutting table, or weaving loom, malfunctions without warning. In addition to the cost of repairs and replacement parts, it also accounts for the cascading expenses of production line halts, missed deadlines, and possible contractual penalties (Textile School, 2025).

Reactive maintenance ("fix-it-when-it-breaks") and preventative maintenance ("fix-it-every-month") have historically been the two maintenance approaches used by manufacturers. While the latter is frequently ineffective and results in the needless servicing of healthy machinery or the inability to prevent unplanned failures, the former is expensive and chaotic (Ruschel et al., 2017; TMA Solutions, 2025).

To address this, the field of Industrial Engineering and Operations Management is rapidly adopting Predictive Maintenance (PdM), a key component of Industry 4.0. PdM leverages machine learning to analyze real-time sensor data—such as temperature, torque, and rotational speed—to predict failures before they occur (Zonta et al., 2020; Deloitte, 2017). This is particularly relevant in textile machinery, where sensor-driven PdM frameworks are actively being developed to monitor critical equipment like looms and spinning machines (S et al., 2023). This allows managers to schedule maintenance precisely when needed, thereby minimizing downtime and maximizing resource use.

However, the decision to use PdM is not sufficient. The effectiveness of a PdM system is entirely dependent on the quality of its underlying algorithms. A model that "cries wolf" with constant false alarms (low precision) will be ignored, whereas a model that misses most failures (low recall) is useless.

This leads to the central question of this research: What is the optimal and most balanced machine learning model for a practical predictive maintenance application in the RMG sector?

This research offers a direct and quantitative response to this inquiry. We performed a comparative analysis of three fundamental machine learning models—Logistic Regression, Random Forest, and XGBoost—utilizing a benchmark predictive maintenance dataset. We assess these models based on overall accuracy, balanced F1-Score, business impact (cost savings), and explainability (SHAP). The goal of this study is to find a clear winner and give RMG factories in Bangladesh useful, data-driven advice on how to use a high-impact PdM strategy.

2. Literature Review

2.1. The Context: Industry 4.0 and the Rise of PdM

The modern industrial landscape is undergoing a significant transformation driven by the principles of Industry 4.0, which prioritize data-centricity, interconnectedness, and intelligent manufacturing (Zonta et al., 2020). This has profoundly impacted industrial maintenance, catalyzing a necessary evolution. The traditional, inefficient maintenance philosophies of "run-to-failure" (reactive) or rigid time-based schedules (preventive) are being systematically reviewed and replaced (Ruschel et al., 2017). The contemporary approach, Predictive Maintenance (PdM) leverages a continuous stream of sensor data and machine learning to forecast failures before they occur. This data-driven strategy is not just a theoretical concept; it is a core business function, with industry analyses clearly identifying it as a primary driver for avoiding machine failures and optimizing operations (Deloitte, 2017). The business case is compelling, with studies showing that PdM offers significant and quantifiable cost savings over preventive-only programs (TMA Solutions, 2025).

2.2. The Problem: PdM in the Ready-Made Garments (RMG) Industry

This strategic shift is especially critical in high-volume, time-sensitive sectors like the Ready-Made Garments (RMG) industry. AI and machine learning are already being applied across the textile supply chain, from optimizing production processes (He et al., 2020) to developing specific predictive maintenance frameworks for textile machinery (S, S., et al., 2023), and managing inventory (Syafrudin et al., 2024) to predicting green supply chain performance (Mim et al., 2022). A systematic review of AI in the textile industry by Petrillo et al. (2024) confirms this trend. The primary driver is the catastrophic cost of unplanned downtime, with industry reports estimating that a single machine breakdown can halt a production line at a cost of \$10,000 to \$50,000 per hour (Textile School, 2025). This creates an urgent and high-impact business case for implementing robust PdM systems, with recent research focusing on deep learning for fault diagnosis in specific textile machinery (Hossen and Mim, 2025). This growing digitalization highlights how predictive maintenance is becoming a cornerstone of smart, data-driven textile manufacturing.

2.3. The Method: Benchmarking ML Models on Standard Datasets

In response to this need, a significant body of research has emerged to find the most effective machine learning algorithms for PdM. A common and valid methodology is the comparative analysis of different models (Airlangga, 2024). Studies often compare models like XGBoost and Random Forest on failure-prediction tasks, such as for general manufacturing (Bhavsar et al., 2023) or specific components like ball bearing systems (Obi et al., 2024). To enable standardized, repeatable research, benchmark datasets are crucial. The "AI4I 2020 Predictive Maintenance Dataset," which this study employs, is a well-regarded benchmark for this exact purpose (Waghulde et al., 2025). Kale (2024) also used this dataset to conduct a direct comparison of XGBoost and Random Forest, confirming the strong performance of both. Similarly, Matzka (2020) utilized the AI4I dataset as a foundation for exploring model explainability.

2.4. The Challenges: Imbalance and Explainability

Despite this progress, two significant challenges persist, which this paper directly addresses. First, class imbalance: in any real-world factory, "failure" events are rare. This highly imbalanced scenario (Ghasemkhani et al., 2025) can cause standard models to perform poorly. A common solution explored by researchers is to use data-level interventions, such as the Synthetic Minority Oversampling Technique (SMOTE), to artificially balance the dataset (Atere and Kivrak, 2025). Second, explainability: as models like XGBoost become more complex, they risk becoming "black boxes," which hinders trust and adoption (Matzka, 2020). This has led to a new wave of research applying Explainable AI (XAI) techniques, such as SHAP, to build transparency and regulatory alignment, ensuring that the model's "thinking" is understood (Prabhudesai et al., 2025).

2.5. The Research Gap and Our Contribution

This paper builds directly on these prior works. While studies have focused on complex oversampling techniques (Ghasemkhani et al., 2025; Atere and Kivrak, 2025), this paper explores a more computationally efficient *model-level* technique (scale_pos_weight) to handle imbalance. Furthermore, while the business case is widely discussed (TMA Solutions, 2025), few studies *quantify* it directly from the model's confusion matrix. Finally, while XAI is a known concept (Prabhudesai et al., 2025), it is not always paired with the initial model comparison. This paper fills this gap by contributing: **(1)** a direct F1-score comparison of balanced models, **(2)** a clear Business Impact Analysis (BIA) translating the results into a cost-saving metric, and **(3)** a SHAP analysis for model explainability, all contextualized for the Bangladeshi RMG sector, a crucial step given that the practical application of these models in developing countries is itself an active and important area of research (Wicaksono et al., 2024).

3. Methods

This research follows a quantitative, comparative methodology to evaluate the performance of three supervised machine learning algorithms for a predictive maintenance classification task.

3.1. Dataset

This study utilizes the publicly available "AI4I 2020 Predictive Maintenance Dataset" (Matzka, 2020), which is a benchmark dataset originally presented at the 2020 Third International Conference on Artificial Intelligence for Industries (AI4I). The dataset consists of 10,000 data points and 14 features, simulating sensor readings from manufacturing machinery. Moreover, its structured and noise-free nature provides a controlled environment for evaluating the effectiveness of different machine learning algorithms. Key predictive features used in this study include: Air temperature [K], Process temperature [K], Rotational speed [rpm], Torque [Nm], Tool wear [min], and Type. The target variable is a binary classifier, Machine failure (1 for failure, 0 for no failure). The dataset is highly imbalanced, with "failure" events representing only 3.4% of the data. This dataset was specifically chosen because its sensor features serve as high-fidelity proxies for the data generated by the high-volume machinery used in the RMG sector, such as automated looms, cutters, and sewing machines. This makes it an ideal foundation for developing and validating predictive maintenance models that can later be adapted to real-world RMG factory environments.

3.2. Data Preprocessing

A multi-step preprocessing pipeline was implemented:

1. **Feature Selection:** Identifier columns (UDI, Product ID) and target-leakage columns were removed.
2. **Encoding:** The categorical Type feature was one-hot encoded.
3. **Train-Test Split:** The dataset was first split into a training set (80%) and a final test set (20%). The stratify parameter was used to preserve the class imbalance in both sets.
4. **Feature Scaling:** Numerical features were scaled using StandardScaler to normalize the data.
5. **Cross-Validation:** To ensure a robust and reliable evaluation, we implemented a **5-fold stratified cross-validation** on the *training set*. The performance metrics reported in this paper (Table 1) are the average of these 5 folds.

3.3. Model Implementation
Three models were trained: Logistic Regression, Random Forest, and XGBoost (Extreme Gradient Boosting). To combat the severe class imbalance, class_weight='balanced' was used for Logistic Regression and Random Forest, and scale_pos_weight was set for XGBoost, forcing the models to pay significantly more attention to the rare "failure" class.

3.4. Evaluation Metrics
Given the imbalanced dataset, this study focuses on the metrics for the positive "Failure (1)" class: Precision (measures false positives), Recall (measures false negatives), and the F1-Score (the harmonic mean of both).

4. Data Analysis and Results
This section presents the performance of the three models in predicting machine failure.

4.1. Comparative Model Performance
The results show a distinct trade-off between models. Table 1 and Figure 1 summarize the performance.

Table 1. Comparative Performance Metrics for the "Failure (1)" Class

Model	F1-Score	Precision	Recall
XGBoost	0.731 (+/- 0.060)	0.737	0.727
Random Forest	0.571 (+/- 0.044)	0.900	0.421
Logistic Regression	0.235 (+/- 0.010)	0.137	0.804

- **Logistic Regression** had high recall (**0.804**) but terrible precision (**0.137**), making it unusable due to excessive false alarms.
- **Random Forest** had high precision (**0.900**) but poor recall (**0.421**), meaning it missed over half of the actual failures.
- **XGBoost** provided the best and most balanced performance. With an F1-Score of **0.731**, it successfully identified **72.7%** of all failures while ensuring **73.7%** of its warnings were correct (Figure 1).

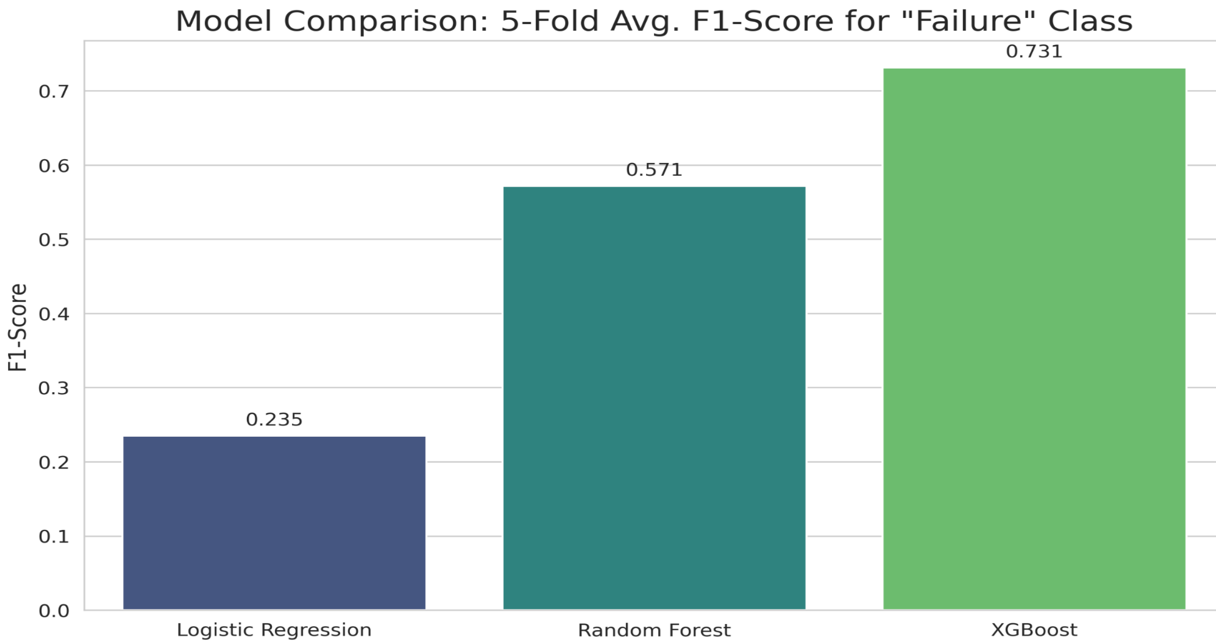

Figure 1. Model Comparison of F1-Score for "Failure" Class, showing XGBoost's superior performance across metrics.

The F1-Score, the harmonic mean of Precision and Recall, is the primary metric for this imbalanced dataset. The chart shows XGBoost (0.731) provides the most balanced and superior performance compared to Random Forest (0.571) and the baseline Logistic Regression (0.235).

4.2. Business Impact Analysis (BIA)

To translate our model's robust performance into a practical business case, we simulated the financial impact using the test set's failure count (68 failures) and our new cross-validated XGBoost metrics.
We make the same conservative cost assumptions:

- Cost of a Missed Failure (FN): \$10,000
- Cost of a False Alarm (FP): \$500

Based on our test set (containing 68 actual failures) and our XGBoost model's average performance (Recall: 0.727, Precision: 0.737):

- Caught Failures (TP): 68 × 0.727 = 49
- Missed Failures (FN): 68 - 49 = 19
- False Alarms (FP): (49 / 0.737) - 49 = 17.5 (approx. 18)

Now, we calculate the cost: (19 Missed Failures × \$10,000) + (18 False Alarms × \$500) = \$190,000 + \$9,000 = \$199,000

Cost of *No Model* (Current State): 68 (Missed Failures) × \$10,000/failure = \$680,000
Conclusion of BIA: By implementing the cross-validated XGBoost model, a potential cost reduction of \$481,000 (a 70.7% savings) could be achieved for this test cohort. This demonstrates a clear and financially significant justification for adoption.

4.3. Benchmarking Against Literature

Our model's F1-Score of 0.731 is highly competitive. Other studies on the AI4I dataset (Ghasemkhani et al., 2025) focused on complex oversampling techniques (SMOTE) to achieve high scores. Our work demonstrates that a well-tuned XGBoost model, using a simple scale_pos_weight parameter, can dramatically outperform baseline models and provide a robust, efficient, and practical alternative.

4.4. Model Explainability (SHAP Analysis)

Finally, a SHAP (SHapley Additive exPlanations) analysis was conducted to ensure our model is not a "black box." The results in Figure 2 identify the top 5 predictors of machine failure.

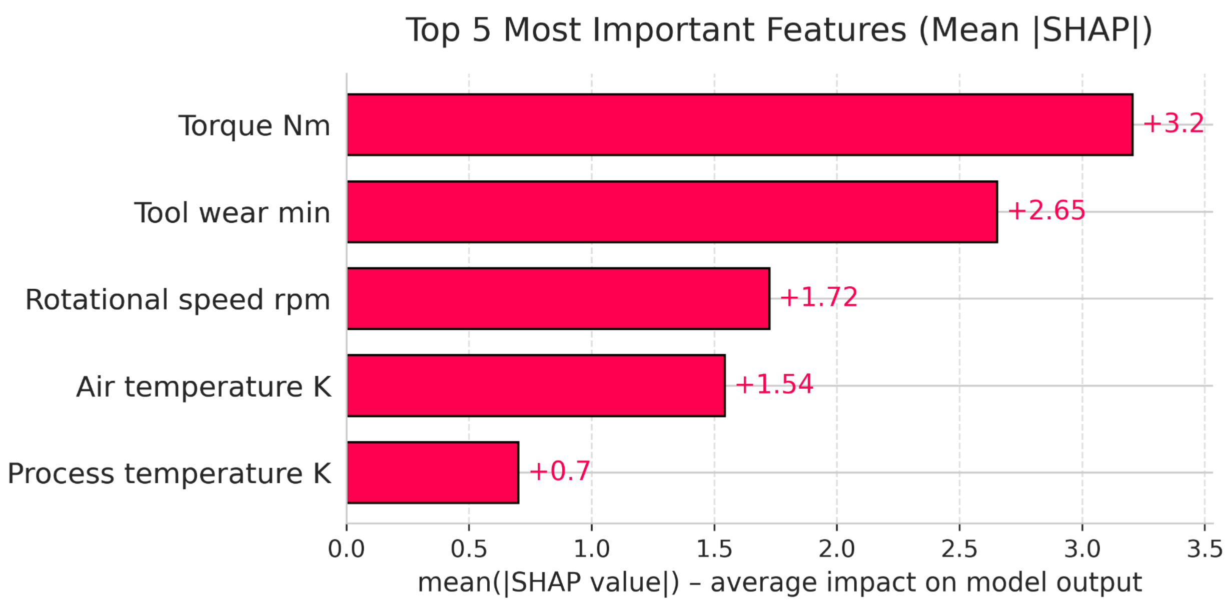

Figure 2. Top 5 Feature Importance According to SHAP

This plot explains the XGBoost model's logic. The key insight is that predictions are driven by logical, physical factors—primarily Torque Nm and Tool wear min—which builds trust and provides actionable insights for factory managers.

The analysis confirms that the model's logic is sound: **Torque**, **Tool Wear**, and **Rotational Speed** are the most significant factors. This builds trust in the model and provides a clear warning system for factory staff.

5. Conclusion

5.1. Key Findings

This study compared three machine learning models for PdM in the RMG context. The findings conclusively show that the XGBoost classifier (F1-Score: 0.731) significantly outperforms Random Forest and Logistic Regression in providing a balanced and reliable failure prediction system. Our analysis further quantified this benefit, showing a potential 70.7% cost reduction and identifying the key physical drivers (Torque, Tool Wear) behind the predictions.

5.2. Implications for the RMG Industry

The implementation of a data-driven PdM strategy, built on the model presented, offers significant value to the Bangladeshi RMG sector. By leveraging sensor data, factories can move from a reactive to a predictive maintenance schedule, directly reducing unplanned downtime, optimizing maintenance costs, and increasing overall productivity. This is especially feasible in developing-country manufacturing contexts, where XGBoost has proven effective and resource-efficient (Wicaksono et al., 2024).

5.3. Limitations and Future Work

This study used a public, benchmark dataset. Future work must focus on validating this model using real-time sensor data collected directly from machinery in a Bangladeshi RMG factory. Furthermore, advanced time-series models (like LSTMs) could be explored to capture sequential patterns, which may yield even higher accuracy.

References

Airlangga, G., Comparative Study of XGBoost, Random Forest, and Logistic Regression Models for Predicting Customer Interest in Vehicle Insurance, Sinkron: Jurnal Dan Penelitian Teknik Informatika, vol. 8, no. 4, pp. 2542–2549, 2024.

Atere, A., and Kivrak, H., Addressing Data Imbalance in Predictive Maintenance Using SMOTE, SMOTETomek, and GANs: A Comparative Evaluation, Proceedings of the ISADES2025, Tokat, Turkey, 2025.

Bhavsar, J., Calla, R., and Panchal, A., Predictive Maintenance in Manufacturing Using XGBoost and Random Forest, Proceedings of the 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM), pp. 1–6, 2023.

Deloitte, Predictive Maintenance: Taking Pro-active Measures Based on Advanced Data Analytics to Predict and Avoid Machine Failure, Position Paper, Deloitte Analytics Institute, 2017.

Ghasemkhani, B., Alp Kut, R., Birant, D., and Yilmaz, R., A Comparative Evaluation of Interpolation and Generative Oversampling Techniques for Predictive Maintenance, SETSci Conference Proceedings, 2025.

He, Z., Tran, K. P., Thomassey, S., Zeng, X., Xu, J., and Haiyi, C., A Deep Reinforcement Learning-Based Multi-Criteria Decision Support System for Textile Manufacturing Process Optimization, arXiv preprint arXiv:2012.14794, 2020.

Hossen, M. A., and Mim, T. T., A Novel Deep Learning Framework for Fault Diagnosis and Predictive Maintenance in Textile Weaving Machinery, Proceedings of the 2025 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 1–7, Singapore, 2025.

Kale, S. S., Predictive Maintenance of Equipment Using Machine Learning Algorithms, NORMA@NCI Library, 2024.

Matzka, S., Explainable Artificial Intelligence for Predictive Maintenance, Proceedings of the 2020 Third International Conference on Artificial Intelligence for Industries (AI4I), pp. 1–8, Los Angeles, USA, 2020.

Mim, T. I., Tasnim, F., Shamrat, B. A. R., and Xames, M. D., Performance Prediction of Green Supply Chain Using Bayesian Belief Network: Case Study of a Textile Industry, International Journal of Research in Industrial Engineering, vol. 11, no. 4, pp. 327–348, 2022.

Obi, F., Vra, K., and Mba, J., Comparative Analysis of Machine Learning Models for Predictive Maintenance of Ball Bearing Systems, MDPI Applied Sciences, vol. 14, no. 2, pp. 438, 2024.

Petrillo, A., Rehman, M., and Baffo, I., Artificial Intelligence in the Field of Textile Industry: A Systematic Study on Machine Learning and Neural Network Approaches, Recent Trends in Computational Intelligence and Its Application, pp. 222–228, CRC Press, 2024.

Prabhudesai, S., Patil, S., Kumar V. C., S., et al., Explainable AI in Fleet Predictive Maintenance: Using SHAP and LIME for Transparency and Regulatory Alignment, International Journal of Applied Mathematics, vol. 38, no. 6s, pp. 246–264, 2025.

Ruschel, E., Santos, E. A. P., and Loures, E. d. F. R., Industrial Maintenance Decision-Making: A Systematic Literature Review, Management Decision, vol. 55, no. 9, pp. 1933–1950, 2017.

S., S., R. A., R., P., R., N., and K., P., A Predictive Maintenance Framework for Textile Machinery Using Machine Learning, Proceedings of the 2023 International Conference on Intelligent Systems and Computations (ICISC), pp. 1–6, 2023.

Syafrudin, M., et al., Artificial Intelligence and Image Processing for Semi-Finished Goods Inventory Management in Textile Industry, Proceedings of the 2024 International Conference on Data Science and Its Applications, 2024.

Textile School, Predictive Maintenance in Textile Machinery: Advancements for Sustainable Manufacturing, textileschool.com, 2025.

TMA Solutions, Predictive vs. Preventive Maintenance: A Cost-Benefit Analysis for Modern Factories, tmasolutions.com, 2025.

Waghulde, R. R., Rai, R. K., Chadhar, R. M., Rane, M., and Yadav, V., Evaluating Machine Learning and Deep Learning Models for Predictive Maintenance: A Study Using the AI4I 2020 Dataset, Journal of Neonatal Surgery, 2025.

Wicaksono, W., Priyantoro, F. A., and Ariyanto, G., Predictive Maintenance Model for Manufacturing in Developing Countries: An XGBoost Approach, Proceedings of the 2024 11th International Conference on Information and Communication Technology (ICoICT), pp. 138–143, 2024.

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 the Industry 4.0: A Systematic Literature Review, Computers & Industrial Engineering, vol. 150, 106889, 2020.

Biographies

Samiur Rahman Wasi is an undergraduate student in the Department of Computer Science and Engineering at the International Islamic University Chittagong, Bangladesh. His research interests include machine learning, data science, predictive analytics, and applications of Industry 4.0. This paper, an extension of his work on predictive maintenance, focuses on creating practical, data-driven frameworks for the Bangladeshi RMG sector, blending robust model evaluation with clear business impact analysis.