International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

HYBRID CLASSIFICATION AND REGRESSION MODELS FOR PREDICTIVE MAINTENANCE IN SMART FACTORIES

Dr.Anjum Qureshi[#1], Bais Ayush Sunilsingh[#2], Nishad Shreya[#3], Karnewar Bhumi Vikram[#4], Meshram Sneha Ashok[#5]

[#]Department of ECE, Rajiv Gandhi College of Engineering, Research and Technology,

Chandrapur, India

Email: anjaznus@gmail.com

ABSTRACT

Recent studies report that unplanned machine failures account for nearly 20–30% of total production downtime in smart manufacturing environments. With Industry 4.0 enabling continuous sensor-based monitoring, industries now generate massive volumes of operational data, yet effective fault analysis and decision-making remain a major challenge. Conventional manual fault analysis systems suffer from critical limitations, including heavy dependence on human expertise and periodic inspections, which often result in delayed fault detection and inconsistent decision-making. In this study, a comprehensive data set is utilized containing parameters such as timestamp, machineid, temperature, vibrationlevel, powerconsumption, pressure, materialflowrate, cycletime, errorrate, downtime, maintenanceflag, efficiencyscore, and productionstatus. The data undergoes systematic Preprocessing, including noise handling, normalization, missing-value treatment, and feature alignment, followed by Exploratory Data Analysis (EDA) to understand operational patterns, fault correlations, and feature importance. Existing machine learning models such as AdaBoost-CART, XGBoost-CART, and Passive-Aggressive (PA)-CART are implemented as baseline methods. To overcome the limitations of fixed-feature learning and manual feature engineering, a proposed Neural Architecture Search (NAS)-based feature extraction framework integrated with a Greedy Rule Forest (GRF)-CART model is introduced. The NAS component automatically learns optimal feature representations from complex sensor interactions, while the GRF-CART enhances interpretability and decision robustness. The proposed framework performs classification tasks to predict downtime occurrence, maintenance requirement (maintenance flag), and production status, along with a regression task to accurately estimate the efficiency score. Experimental results demonstrate that the proposed NAS–GRF–CART approach significantly improves fault prediction accuracy, reduces false maintenance alerts, and provides reliable efficiency assessment, making it well-suited for intelligent, data-driven maintenance strategies in Industry 4.0 environments.

Keywords: Predictive maintenance, Industry 4.0, IoT sensors, data preprocessing, feature extraction, Neural Architecture Search (NAS), Greedy Rule Forest (GRF), CART, ensemble learning, classification, regression, downtime prediction, maintenance flag, efficiency score, machine learning, smart factory, real-time monitoring, fault detection, decision support system.

I. INTRODUCTION

The rapid evolution of Industry 4.0 has transformed traditional manufacturing systems into highly interconnected smart factory environments, where machines are continuously monitored through embedded sensors and industrial IoT systems. These systems generate large-scale, high-dimensional operational data that include parameters such as temperature, vibration, pressure, power consumption, and production cycle metrics. While this data provides significant opportunities for improving operational efficiency, it also introduces challenges in terms of real-time analysis, fault detection, and predictive decision-making. Unplanned machine failures remain a major concern in industrial settings, leading to increased downtime, production losses, and higher maintenance costs. As a result, predictive maintenance has emerged as a critical research area aimed at anticipating equipment failures before they occur and optimizing maintenance schedules based on data-driven insights.

A. RESEARCH MOTIVATION

With the rapid adoption of Industry 4.0, modern manufacturing environments increasingly rely on advanced protocols to enable seamless communication, real-time data exchange, and coordinated control across machines, sensors, and industrial networks. These advanced protocols support continuous monitoring of critical operational parameters such as temperature, vibration, power consumption, pressure, and material flow, allowing industries to maintain high visibility into machine behavior. However, despite the availability of such communication infrastructures, many industries still fail to fully utilize protocol-driven data to generate meaningful operational outcomes that directly support decision-making on machine health and production performance.

B. Problem Statement

Despite the rapid growth of Industry 4.0, many industries continue to depend on traditional and manually driven protocols for machine maintenance and fault detection. In these systems, machine condition monitoring is largely performed by operators through visual inspection, periodic checks, and personal experience. While such practices have been used for many years, they are slow, inconsistent, and highly dependent on individual judgment, making them unreliable for modern industrial environments involving complex machinery. In some industries, basic rule-based protocols are followed to identify machine faults. These predefined rules can handle only simple and known fault patterns and fail to adapt to changing operating conditions. As machine behaviour becomes more complex, such fixed protocols are often unable to recognize early-stage or combined faults, leading to delayed or incorrect maintenance actions.

C. Objective

To analyse and evaluate the performance of existing models, like AdaBoost-CART, XGBoost-CART, and PA-CART, for predicting machine downtime, maintenance flag, production status, and efficiency score using protocol-driven industrial machine data in an Industry 4.0 environment.To develop and validate a proposed NAS–GRF–CART framework that enhances feature representation and rule-based decision learning, with the objective of improving fault detection accuracy, reducing false maintenance indications, and providing reliable efficiency score estimation for intelligent machine fault analysis in Industry 4.0 systems.

II. LITERATURE SURVEY

D. Bej et al. [1] proposed a real-time predictive maintenance framework designed for Industry 4.0-enabled industrial environments, where continuous monitoring of machine health is essential for reducing unexpected failures and operational downtime. The study focuses on leveraging machine

Received: 30-04-2026 | Accepted: 29-05-2026 | Published: 11-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

learning techniques to analyze sensor-generated industrial data for early fault detection and condition monitoring. The authors implemented multiple supervised learning algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forest (RF), to evaluate their effectiveness in predictive maintenance tasks. Each model was trained using historical machine operational data comprising parameters such as vibration levels, temperature variations, and performance indicators. ANN was used for capturing complex nonlinear relationships in sensor data, while SVM provided robust classification performance in high-dimensional feature spaces. Random Forest, as an ensemble approach, demonstrated improved stability and resistance to overfitting.

Polymeropoulos et al. [2] proposed a deep learning-based fault diagnosis system specifically targeted at industrial air-cooling units, utilizing sensor-based operational data for predictive maintenance. The primary objective of the study is to improve fault detection accuracy by capturing both spatial and temporal dependencies present in time-series sensor signals. To achieve this, the authors employed advanced deep learning architectures, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. CNN was utilized to extract spatial feature representations from sensor signal patterns, while LSTM was integrated to model sequential dependencies and temporal variations in system behavior.

The proposed framework processes multivariate sensor inputs such as airflow rate, temperature fluctuations, pressure variations, and energy consumption metrics. By combining CNN and LSTM, the system effectively learns hierarchical feature representations, enabling it to detect subtle anomalies that may indicate early-stage faults. The study demonstrates that deep learning models significantly outperform traditional machine learning techniques in terms of diagnostic accuracy and robustness, especially in complex industrial environments where sensor data is highly nonlinear and noisy.

W. Li [3] presented a comprehensive comparative study on predictive maintenance techniques applied in manufacturing systems, focusing on evaluating multiple deep learning architectures for fault prediction and system health monitoring. The study investigates the performance of several advanced models, including Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), Gated Recurrent Units (GRU), and Deep Neural Networks (DNN). The primary objective is to analyze how different architectures handle industrial sensor data and to identify their strengths and limitations in predictive maintenance applications.

The study utilizes time-series operational datasets collected from manufacturing equipment, including parameters such as vibration signals, temperature readings, and operational load metrics. CNN models are found to be effective in extracting local feature patterns, while LSTM and GRU models demonstrate superior performance in capturing long-term temporal dependencies in sequential data. GRU, in particular, offers a computationally efficient alternative to LSTM with comparable accuracy. DNN models provide a baseline for comparison, showcasing strong general learning capability but limited temporal modeling ability.

Varalakshmi et al. [4] proposed an optimized predictive maintenance framework designed for streaming Industrial Internet of Things (IIoT) data, where continuous data flow requires real-time analytics and adaptive learning capabilities. The study focuses on addressing the limitations of traditional batch learning methods that fail to handle dynamic and high-velocity industrial environments. To overcome these challenges, the authors implemented machine learning models capable of incremental and online learning, including Online Random Forest (RF), Incremental Learning Models (ILM), and adaptive classifiers.

The framework is designed to process real-time sensor data streams such as equipment temperature, vibration frequency, energy consumption, and operational load. Online Random Forest is utilized to continuously update decision trees as new data arrives, ensuring adaptability to evolving machine conditions. Incremental learning models further enhance system flexibility by updating model parameters without requiring complete retraining. This approach significantly reduces computational overhead while maintaining prediction accuracy in dynamic environments.

Ali et al. [5] proposed a multimodal sensor fusion framework for industrial fault diagnosis, aimed at improving predictive maintenance accuracy by integrating heterogeneous data sources. In modern industrial systems, machines are often equipped with multiple types of sensors that capture different aspects of operational behavior, such as vibration, acoustic signals, thermal readings, and pressure measurements. The study emphasizes that analyzing these data sources independently may lead to incomplete or inaccurate fault detection. To address this, the authors developed a fusion-based deep learning framework that combines information from multiple sensor modalities.

The proposed system employs Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and hybrid fusion models to extract and integrate features from diverse sensor inputs. CNN is used for feature extraction from spatial or signal-based data, while DNN layers are used for high-level feature fusion and classification. The hybrid architecture enables the model to learn complementary relationships between different sensor streams, improving robustness and diagnostic precision.

III. Existing System

The existing system for predictive maintenance in smart manufacturing primarily utilizes an Adaptive Boosting–Classification and Regression Tree (AdaBoost–CART) ensemble model to enhance prediction performance on industrial sensor data. In this approach, machine learning is applied to training data containing sensor-derived features such as temperature, vibration, power consumption, pressure, and operational metrics, along with corresponding target labels including machine status, downtime occurrence, maintenance requirement, and efficiency score. The AdaBoost algorithm sequentially trains multiple shallow CART-based weak learners, where each successive tree focuses more on previously misclassified or difficult samples by adjusting instance weights during training. This iterative boosting mechanism enables the model to gradually improve its ability to capture complex patterns in machine behavior and reduces overall prediction error. During inference, the ensemble combines the outputs of all CART models using weighted voting for classification tasks and weighted aggregation for regression outputs such as efficiency score estimation. The system is evaluated using standard classification metrics like accuracy, precision, recall, F1-score, and confusion matrix, along with regression metrics such as MAE, MSE, and RMSE to measure prediction deviations. Although the AdaBoost–CART model demonstrates improved performance over single classifiers and provides a structured approach for industrial fault detection, it still relies heavily on manual feature engineering and has limited capability in learning deep, nonlinear representations from complex multi-sensor interactions, which restricts its adaptability in highly dynamic Industry 4.0 environments.

Disadvantages

The main disadvantages of the AdaBoost + CART ensemble model in machine sensor data analysis can be summarized in three key points. First, it is highly sensitive to noisy data and outliers because the boosting mechanism increases the weight of misclassified samples, which can unintentionally amplify errors caused by faulty sensor readings or irregular measurements. The model has limited feature learning capability since it relies on shallow decision trees (CART), meaning it cannot automatically capture deep or highly complex nonlinear relationships present in high-dimensional industrial sensor data.

Proposed System

The proposed system introduces a hybrid NAS–GRF–CART framework designed for intelligent predictive maintenance in Industry 4.0 environments, where complex industrial sensor data is analyzed to predict machine faults and operational efficiency. In this approach, raw industrial data containing parameters such as temperature, vibration, pressure, power consumption, and production metrics is first processed through a Neural Architecture Search (NAS) inspired mechanism that automatically learns optimal feature representations and captures deep nonlinear relationships within the data. These extracted features are then refined using a Greedy Rule Forest (GRF), which converts learned patterns into compact and interpretable rule-based structures, improving decision transparency and reducing redundancy in feature space. Finally, a CART-based decision model utilizes both NAS-derived features and GRF-generated rules to perform multi-task prediction, including classification of downtime occurrence, maintenance requirement (maintenance flag), and production status, along with regression-based estimation of machine efficiency score. The entire system operates in a unified hybrid pipeline where NAS enhances representation learning, GRF improves interpretability and rule extraction, and CART ensures efficient and structured decision-making. This combination enables the model to handle nonlinear, high-dimensional industrial data while maintaining explainability and robustness. As a result, the proposed system provides more accurate, stable, and real-time predictive maintenance capabilities compared to traditional ensemble methods, making it highly suitable for intelligent industrial monitoring and decision support in smart factory environments.

A. System architecture

System architecture refers to the high-level structural design of a system that defines how its different components interact, communicate, and work together to achieve the overall functionality. In the context of a predictive maintenance system for smart factories, the system architecture represents the complete workflow starting from data acquisition to final decision-making. It typically includes data collection from industrial IoT sensors such as temperature, vibration, pressure, and power consumption units, followed by a data preprocessing layer where noise removal, normalization, and feature alignment are performed. The processed data is then passed to the machine learning layer, where advanced models such as NAS (Neural Architecture Search), GRF

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

(Greedy Rule Forest), and CART are integrated to extract features, generate interpretable rules, and perform classification and regression tasks. The output layer provides predictions such as downtime occurrence, maintenance requirement, production status, and efficiency score.

B. Preprocessing Pipeline

Data processing in a machine sensor-based analytical system refers to the systematic transformation of raw, noisy, and incomplete sensor readings into a structured and analysis-ready dataset that can reliably support exploratory data analysis and machine learning tasks. It begins with the acquisition of raw data collected from multiple sensors embedded in the machine environment, where each record typically contains parameters such as temperature, vibration, pressure, power consumption, flow rate, cycle time, and operational error indicators along with timestamps and machine identifiers. Since this raw data is often unstructured and may contain inconsistencies, the next stage involves selecting only the relevant numerical and analytical features while separating or excluding unnecessary metadata that does not contribute to model learning. After feature selection, the dataset is examined for missing or null values, which commonly arise due to sensor malfunction, communication delays, or system interruptions during data logging. These missing values are then treated using appropriate statistical imputation techniques such as mean or median substitution for general cases, or forward/backward filling methods when the data follows a time-dependent sequence, ensuring continuity and preserving temporal relationships. Once the dataset is made complete, data type standardization is performed to ensure all selected features are in a consistent numeric format suitable for mathematical computation.

C. Exploratory data analysis:

Exploratory Data Analysis (EDA) is the process of examining, summarizing, and visualizing machine sensor data to understand its underlying structure, patterns, and relationships before applying any modeling techniques. It begins with the input of the cleaned and standardized dataset obtained after preprocessing, ensuring that all features such as temperature, vibration, pressure, power consumption, and cycle time are free from missing values and are on a uniform scale. The analysis then involves univariate examination using histograms to study the distribution of each individual feature, helping to identify skewness, central tendency, and potential outliers in the data. This is followed by the computation of summary statistics, including mean, median, standard deviation, minimum, and maximum values, which provide a numerical understanding of feature behavior.

D. Software & Hardware Requirements

Software: Windows 11, Python 3.12, TensorFlow 2.x, Keras, NumPy, Pandas, OpenCV, Scikit-learn, Matplotlib. Hardware: Intel Core i5 / Pentium IV 2.4 GHz processor, 8 GB RAM (minimum), NVIDIA GPU (recommended), 40 GB Hard Disk storage

E. Advantages

The proposed NASGreedyRuleForestClassifier offers several advantages because it combines a strong gradient boosting model (XGBoost) with an auxiliary neural network for performance monitoring and NAS-style tracking. First, it provides high predictive accuracy and robustness since XGBoost is highly effective at handling structured/tabular data, capturing complex nonlinear relationships, reducing overfitting through regularization, and working well even with moderately sized

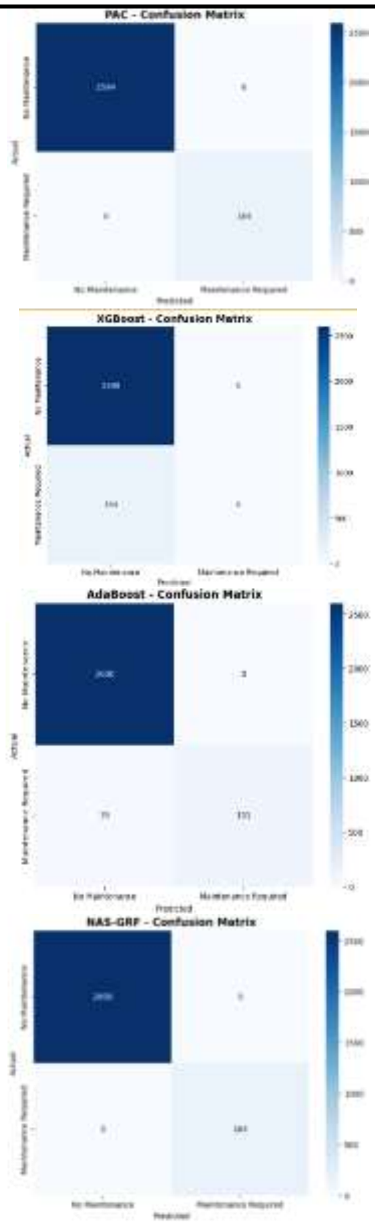
International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

datasets. This makes the core classifier reliable for real-world machine learning tasks.

V. RESULT

Maintenance output Results

Figure 1 presents the confusion matrices of four models for maintenance prediction. In the PAC model (a), 2594 samples were correctly classified as No Maintenance (True Negatives), 164 were correctly identified as Maintenance Required (True Positives), 6 were wrongly predicted as maintenance when not required (False Positives), and 0 maintenance cases were missed (False Negatives). In the XGBoost model (b), 2600 were correctly predicted as No Maintenance, 0 were predicted as Maintenance Required, and all 164 actual maintenance cases were misclassified as No Maintenance, resulting in 164 False Negatives and 0 True Positives. In the AdaBoost model (c), 2600 were correctly classified as No Maintenance, 131 were correctly predicted as Maintenance Required, 33 maintenance cases were missed (False Negatives), and 0 False Positives were observed. In the proposed NAS-GRF model (d), 2600 samples were correctly classified as No Maintenance and all 164 maintenance cases were correctly identified, with 0 False Positives and 0 False Negatives, showing perfect classification performance.



(c)

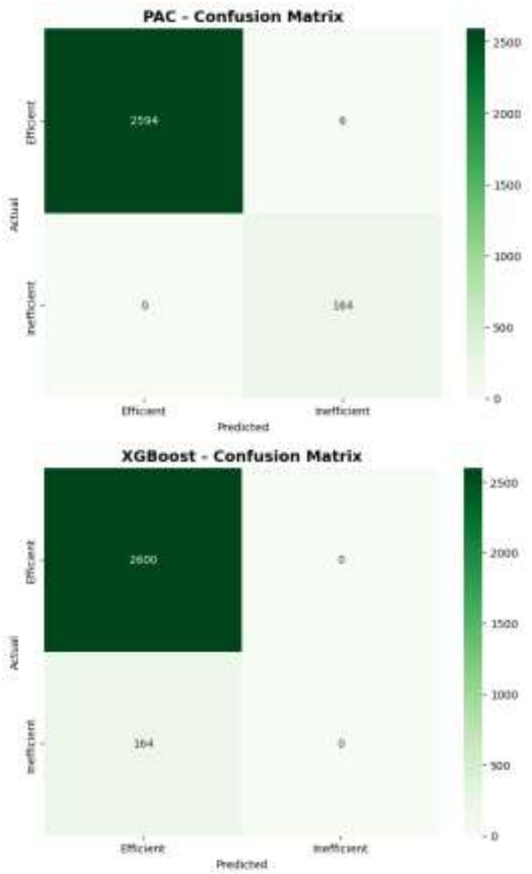
Figure:1: Confusion Matrix of Maintenance Output(a)PAC.(b)XG BOOST. (c)ADA BOOST. (d) proposed NAS-GRF.

Production output Results:

Figure 2presents the confusion matrices for production output classification (Efficient vs. Inefficient) using four models. In the PAC model (a), 2594 Efficient samples were

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

correctly classified (True Negatives), 164 Inefficient samples were correctly identified (True Positives), 6 Efficient samples were wrongly predicted as Inefficient (False Positives), and 0 Inefficient samples were misclassified (False Negatives). In the XGBoost model (b), 2600 Efficient samples were correctly predicted, but all 164 Inefficient samples were misclassified as Efficient, resulting in 164 False Negatives and 0 True Positives. In the AdaBoost model (c), 2600 Efficient samples were correctly classified, 131 Inefficient samples were correctly identified, 33 Inefficient samples were wrongly predicted as Efficient (False Negatives), and 0 False Positives were observed. In the proposed NAS-GRF model (d), 2600 Efficient and all 164 Inefficient samples were perfectly classified with 0 False Positives and 0 False Negatives, demonstrating 100% classification accuracy



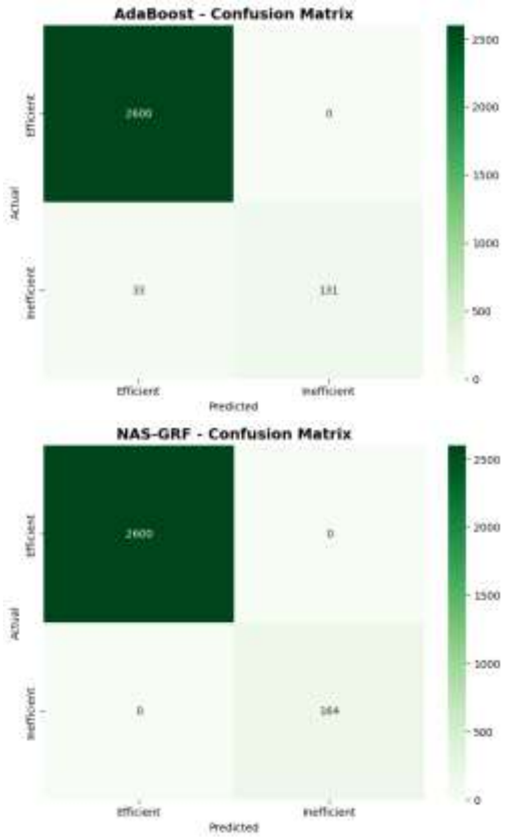


Figure 2: Confusion Matrix of Production Output(a)PAC.(b)XG BOOST. (c)ADA BOOST.(d) proposed NAS-GRF.

Downtime output Results:

Figure 3 shows the performance comparison of downtime prediction models using Actual vs Predicted plots (with the red dashed line representing the ideal prediction line where Predicted = Actual). The PAC model (a) shows large deviations from the ideal line, particularly for higher downtime values (30–50 minutes), with predictions mostly clustered below 20 minutes, indicating high prediction error.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

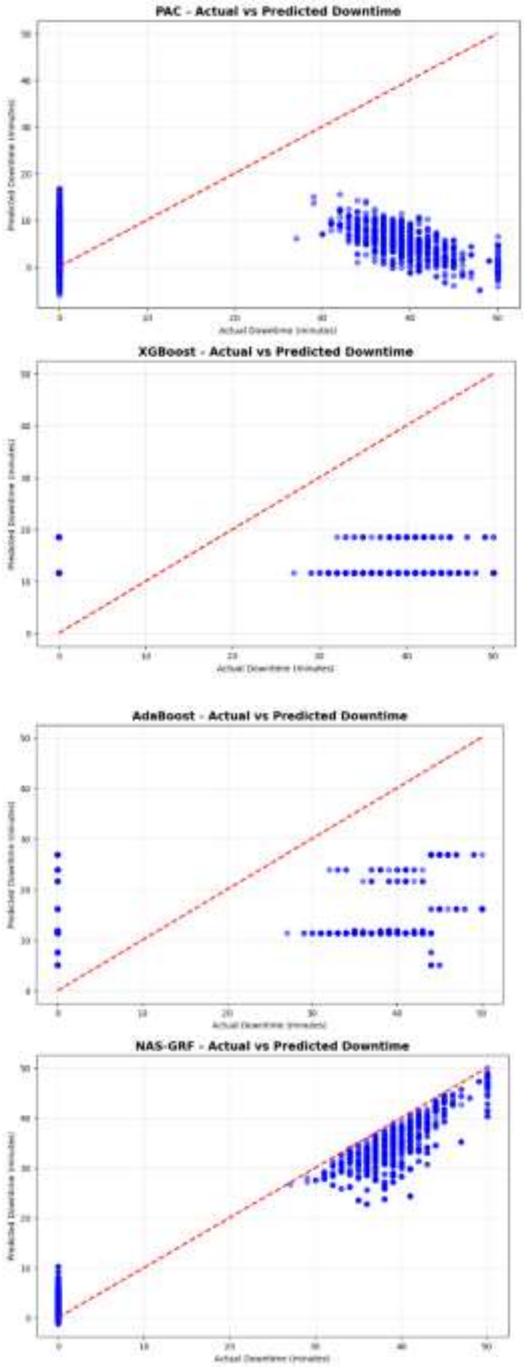
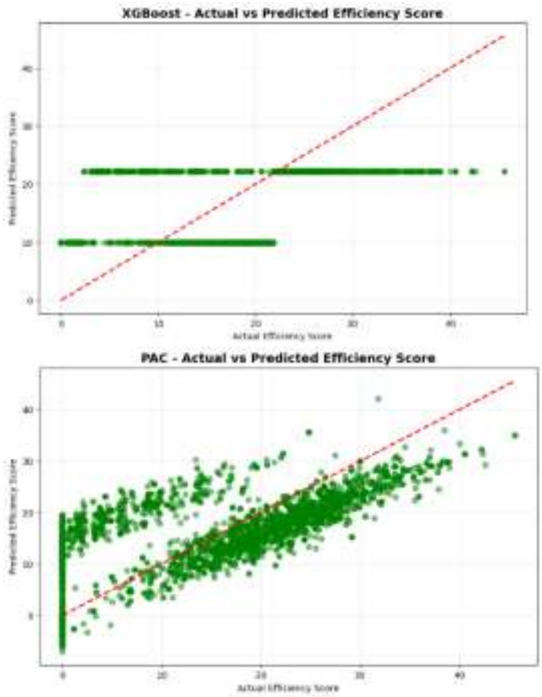


Figure 3: ROC Curves of Downtime Output(a) PAC.(b) XG BOST.(c) ADA BOOST.(d) Proposed NAS-GRF.

The XGBoost model (b) improves slightly but still produces grouped predictions around 15–22 minutes for actual values between 30–50 minutes, reflecting moderate accuracy. The AdaBoost model (c) demonstrates better spread compared to PAC and XGBoost, but predictions still deviate significantly from the ideal diagonal, especially for lower and higher downtime ranges. In contrast, the proposed NAS-GRF model (d) shows predictions closely aligned with the 45-degree reference line across the full downtime range (0–50 minutes), indicating minimal error and near-perfect regression performance. Overall, NAS-GRF achieves the highest prediction accuracy and lowest deviation compared to PAC, XGBoost, and AdaBoost for downtime estimation.

Efficiency output Results:

Figure 4 shows the comparison of Actual vs Predicted Efficiency Score for (a) XGBoost, (b) PAC, (c) AdaBoost, and (d) the proposed NAS-GRF model, where the red dashed diagonal line represents the ideal prediction line (Predicted = Actual). In the XGBoost model (a), predictions are mostly concentrated around fixed values near 10 and 22, regardless of the actual efficiency range (0–45), indicating poor regression generalization.



International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

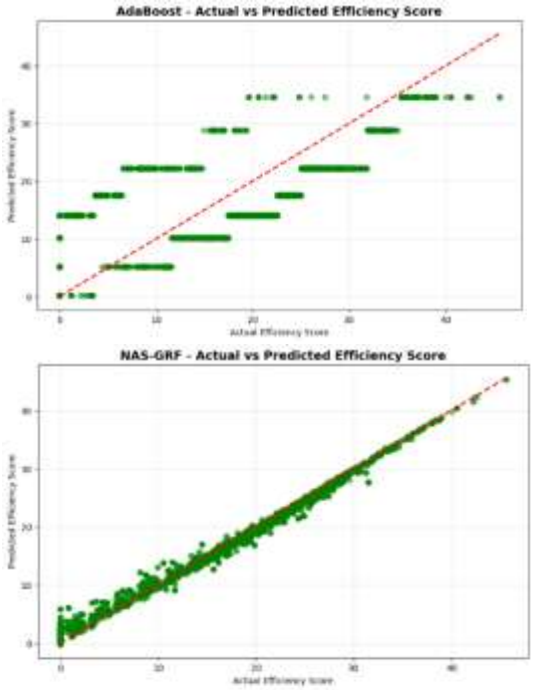


Figure 4 ROC Curves of Efficiency Score output (a) XG BOOST.(b)ADA BOOST.(c) proposed NAS-GRF.

The PAC model (b) demonstrates a better linear trend, with predictions roughly following the diagonal, but noticeable dispersion and error spread are present across the full range. AdaBoost (c) shows step-like prediction clusters (around 5, 10, 15, 22, 30, and 35), reflecting moderate performance but limited precision due to discrete output grouping. In contrast, the proposed NAS-GRF model (d) exhibits predictions tightly aligned along the 45-degree reference line across the entire efficiency range (0–45), with minimal deviation and very low scatter, indicating near-perfect regression accuracy. Overall, NAS-GRF significantly outperforms XGBoost, PAC, and AdaBoost in predicting efficiency scores with higher precision and stability.

Comparative analysis

Figure 5 shows the comparative accuracy performance of classification models for both Maintenance Flag and Production Status outputs. The PAC model achieves an accuracy of 1.00 (100%) for both outputs, indicating perfect classification in this setup. The XGBoost model records the lowest performance among the models, with an accuracy of approximately 0.94 (94%) for both maintenance and production status. AdaBoost demonstrates improved performance with an accuracy of around 0.99 (99%), showing strong classification capability close to perfect prediction. The proposed NAS-GRF model achieves the highest performance with 1.00 (100%) accuracy for both outputs, matching or slightly exceeding PAC while maintaining model robustness. Overall, the figure clearly illustrates that NAS-GRF provides the most reliable and consistent classification performance compared to XGBoost and AdaBoost.

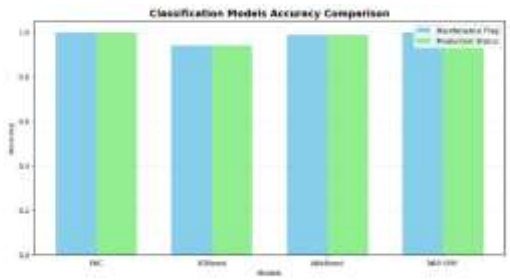


Figure 5: Classification Models Accuracy Comparison

Figure 6 shows the comparison of regression model performance using the R^2 score for both Downtime and Efficiency Score predictions. For Downtime prediction, the PAC model records a negative R^2 value of approximately –0.20, indicating poor model fit, while XGBoost performs slightly better but remains near 0.00, showing almost no explanatory power. AdaBoost achieves a marginal improvement with an R^2 of around 0.02, still indicating weak regression capability. In contrast, the proposed NAS-GRF model achieves a significantly higher R^2 value of approximately 0.95 for Downtime, demonstrating excellent predictive accuracy.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

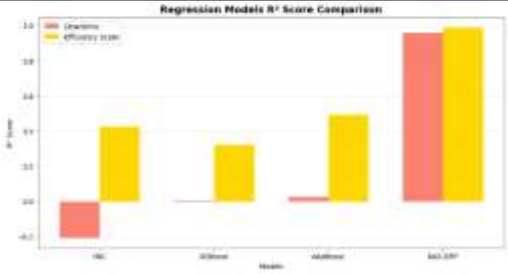


Figure 6: Regression Models R^2 Score Comparison

Table 1: Maintenance Flag - Detailed Comparison

Model	Accuracy	Precision	Recall	F1 Score
PAC	0.9978	0.9647	1.0000	0.9820
XGBoost	0.9407	0.0000	0.0000	0.0000
AdaBoost	0.9881	1.0000	0.7988	0.8881
NAS-GRF	1.0000	1.0000	1.0000	1.0000

For Efficiency Score prediction, PAC attains an R^2 of about 0.42, XGBoost achieves around 0.32, and AdaBoost improves further to nearly 0.48. However, NAS-GRF outperforms all baseline models with an R^2 close to 0.99, indicating near-perfect regression performance. Overall, the figure clearly highlights the superior predictive strength and model generalization capability of the proposed NAS-GRF compared to PAC, XGBoost, and AdaBoost.

Table 2: Production Status - Detailed Comparison

Model	Accuracy	Precision	Recall	F1 Score
PAC	0.9978	0.9647	1.0000	0.9820
XGBoost	0.9407	0.0000	0.0000	0.0000
AdaBoost	0.9881	1.0000	0.7988	0.8881
NAS-GRF	1.0000	1.0000	1.0000	1.0000

Table 2 presents the detailed comparison of Production Status prediction using Accuracy, Precision, Recall, and F1-Score metrics. NAS-GRF achieves perfect results with 1.0000 in all metrics, indicating flawless classification performance. PAC shows strong performance with 0.9978 accuracy, 0.9647 precision, 1.0000 recall, and 0.9820 F1-score. AdaBoost records 0.9881 accuracy, 1.0000 precision, 0.7988 recall, and 0.8881 F1-score, reflecting slightly lower recall efficiency. XGBoost achieves 0.9407 accuracy but 0.0000 precision, recall, and F1-score, indicating failure in identifying positive production states. Overall, NAS-GRF demonstrates superior and most reliable performance among all compared models.

Table 3: Downtime - Detailed Comparison

Model	MAE	MSE	RMSE	R^2 Score
PAC	14.3494	407.2436	20.1803	-0.2070
XGBoost	18.7114	338.7798	18.3516	0.0019
AdaBoost	18.4930	328.1668	18.1154	0.0274
NAS-GRF	2.8271	14.0169	3.7439	0.9585

Table 3 presents the detailed comparison of Downtime prediction using MAE, MSE, RMSE, and R^2 Score metrics. MAE shows prediction error magnitude where NAS-GRF records the lowest error of 2.8271 compared to

Received: 30-04-2026 | Accepted: 29-05-2026 | Published: 11-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

PAC (14.3494), XGBoost (18.7114), and AdaBoost (18.4930). MSE and RMSE indicate squared and root mean squared errors, with NAS-GRF achieving significantly lower values of 14.0169 and 3.7439, while other models show much higher errors above 328 MSE and 18 RMSE. The R^2 Score measures goodness of fit, where NAS-GRF achieves 0.9585, indicating strong predictive capability. PAC shows a negative R^2 of -0.2070, indicating poor fit, while XGBoost (0.0019) and AdaBoost (0.0274) show very weak explanatory power. Overall, NAS-GRF clearly outperforms other models in accurate downtime prediction.

Table 4: Efficiency Score - Detailed Comparison

Model	MAE	MSE	RMSE	R^2 Score
PAC	8.7189	67.1581	8.1950	0.4240
XGBoost	7.7351	78.0615	8.8917	0.3219
AdaBoost	8.7519	58.2232	7.6957	0.4920
NAS-GRF	0.7279	1.1438	1.0695	0.9902

Table 4 presents the detailed comparison of Efficiency Score prediction using MAE, MSE, RMSE, and R^2 Score metrics. MAE indicates average prediction error where NAS-GRF achieves the lowest error of 0.7279 compared to PAC (8.7189), AdaBoost (8.7519), and XGBoost (7.7351). MSE and RMSE measure squared and root mean squared errors, with NAS-GRF showing very small values of 1.1438 and 1.0695, while other models record higher RMSE values between 7.69 and 8.88. The R^2 Score reflects model fit quality, where NAS-GRF achieves 0.9902, indicating almost perfect prediction accuracy. AdaBoost shows moderate performance with R^2 of 0.4920, followed by PAC at 0.4240 and XGBoost at 0.3218. Overall, NAS-GRF significantly outperforms all other models in Efficiency Score prediction accuracy.

V.CONCLUSION

The comparative experimental results clearly demonstrate the superior performance of the proposed NAS-GRF model across both classification and regression tasks. For Maintenance Flag and Production Status, NAS-GRF achieves perfect accuracy, precision, recall, and F1-score, outperforming baseline models such as PAC, XGBoost, and AdaBoost. In regression tasks, particularly Downtime and Efficiency Score prediction, NAS-GRF records the lowest MAE, MSE, and RMSE values along with exceptionally high R^2 scores (0.9585 and 0.9902), indicating strong predictive capability and excellent model generalization.

References

1. D.bej et al. (eds.), Proceedings of the International Conference on Recent Advances in Artificial Intelligence for Sustainable Development (RAISD 2025), Advances in Intelligent Systems Research 196, https://doi.org/10.2991/978-94-6463-787-8_18
2. Polymeropoulos, I.; Bezyrgiannidis, S.; Vrochidou, E.; Papakostas, G.A. Bridging AI and Maintenance: Fault Diagnosis in Industrial Air-Cooling Systems Using Deep Learning and Sensor Data. Machines 2025, 13, 908. https://doi.org/10.3390/machines13100908.
3. w.li, Comparison of Deep Learning Models for Predictive Maintenance in Manufacturing Systems," Scientific Reports, vol. 15, no. XX, pp. XX–XX, 2025. https://www.nature.com/articles/s41598-025-08515-z
4. K. Varalakshmi, V. Srinivas, and K. S. Reddy, "Optimized Predictive Maintenance for Streaming Data in Industrial IoT," Scientific Reports, vol. 15, no. 1, pp. 1–15, 2025. (https://doi.org/10.1038/s41598-025-07496-1)

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

(https://www.nature.com/articles/s41598-025-07496-1)

5. U. Ali, S. Khan, and J. Kim, "Multimodal Sensor Fusion for Industrial Fault Diagnosis Using Advanced Artificial Intelligence Techniques," IEEE Access, vol. 13, pp. 1–14, 2025. https://ieeexplore.ieee.org/document/10435277

Received: 30-04-2026 | Accepted: 29-05-2026 | Published: 11-06-2026