QOMARUNA Journal of Multidisciplinary Studies

E-ISSN: 3026-5266
https://jurnal.uqgresik.ac.id/index.php/qjms

Predicting Computer Numerical Control (CNC) Machine Downtime using Ensemble Learning Approaches

Amir Aminuddin [1]*, Adam Shariff [1], Kamarulzaman Mahmad Khairai [2]

[1] Centre of Mathematical Science, Universiti Malaysia Pahang Al-Sultan Abdullah, Gambang, Malaysia, 26300
[2] Faculty of Industrial Management, Universiti Malaysia Pahang Al-Sultan Abdullah, Gambang, Malaysia, 26300
* Correspondence: amiramin5587@gmail.com

Received: 31 August 2025
Revised: 2 December 2025
Accepted: 13 December 2025

Citation:
Aminuddin, A., Shariff, A.., & Mahmad Khairai, K. (2025). Predicting Computer Numerical Control (CNC) Machine Downtime using Ensemble Learning Approaches. *Qomaruna: Journal of Multidisciplinary Studies*, 3(1), 59–74. https://doi.org/10.62048/qjms.v3i1.137.

ABSTRACT

Computer Numerical Control (CNC) machining is a subtractive manufacturing technique that removes layers of material from a blank or workpiece to create a specific product. With increasing global competition, minimizing downtime during production is essential to maximize machine availability and productivity. This study investigates the application of machine learning models, specifically Extreme Gradient Boosting (XGBoost) and Random Forest (RF), to forecast CNC machine downtime from multiple failure sources. The study uses data collected from 16 CNC machines at Company A in Malaysia over an extended period. The data contain key variables for each downtime event, such as machine ID, failure type, start date/time, end date/time, and downtime duration in minutes. Failure types are categorized into several groups, including mechanical, electrical, and tool malfunctions. After hyperparameter tuning, the XGBoost model outperformed the RF model, achieving a Mean Squared Error (MSE) of 0.4017, Root MSE (RMSE) of 0.634, and Mean Absolute Error (MAE) of 0.470 on the test set, while the RF model yielded higher errors, with an MSE of 1.2654, RMSE of 1.125, and MAE of 0.943. These results demonstrate the superiority of the XGBoost model over RF in predicting future CNC downtime, as indicated by its lower prediction errors. Future work should focus on refining the model with larger, more diverse datasets and exploring its integration into AI-based decision support systems to enhance machine availability and operational efficiency.

Keywords: *Computer Numerical Control (CNC), machine learning, predictive maintenance maintenance, manufacturing*

ABSTRAK

Pemesinan Computer Numerical Control (CNC) merupakan teknik manufaktur subtraktif yang menghilangkan lapisan material dari bahan awal atau benda kerja untuk menghasilkan produk tertentu. Seiring dengan meningkatnya persaingan global, meminimalkan waktu henti (downtime) selama proses produksi menjadi sangat penting untuk memaksimalkan ketersediaan mesin dan produktivitas. Penelitian ini mengkaji penerapan model machine learning, khususnya Extreme Gradient Boosting (XGBoost) dan Random Forest (RF), untuk memprediksi waktu henti mesin CNC yang berasal dari berbagai sumber kegagalan. Penelitian ini menggunakan data yang dikumpulkan dari 16 mesin CNC di Perusahaan A di Malaysia selama periode waktu yang cukup panjang. Data tersebut mencakup variabel-variabel utama untuk setiap kejadian downtime, seperti identitas mesin, jenis kegagalan, waktu mulai, waktu selesai, dan durasi downtime dalam satuan menit. Jenis kegagalan diklasifikasikan ke dalam beberapa kategori, termasuk kegagalan mekanik, listrik, dan perkakas. Setelah dilakukan penyetelan hiperparameter, model XGBoost menunjukkan kinerja yang lebih unggul dibandingkan model RF, dengan nilai Mean Squared Error (MSE) sebesar 0,4017, Root Mean Squared Error (RMSE) sebesar 0,634, dan Mean Absolute Error (MAE) sebesar 0,470 pada data uji. Sebaliknya, model RF menghasilkan tingkat kesalahan yang lebih tinggi, dengan MSE sebesar 1,2654, RMSE sebesar 1,125, dan MAE sebesar 0,943. Hasil ini menunjukkan keunggulan model XGBoost dibandingkan RF dalam memprediksi downtime mesin CNC di masa mendatang, sebagaimana tercermin dari nilai kesalahan prediksi yang lebih rendah. Penelitian selanjutnya disarankan untuk menyempurnakan model dengan menggunakan dataset yang lebih besar dan lebih beragam, serta mengeksplorasi integrasinya ke dalam sistem pendukung keputusan berbasis kecerdasan buatan untuk meningkatkan ketersediaan mesin dan efisiensi operasional.

Kata kunci: Computer Numerical Control (CNC), machine learning, pemeliharaan prediktif, manufaktur

Introduction

CNC machining is a fundamental component of modern manufacturing and supports key sectors such as automotive, aerospace, medical, and oil and gas (Keller et al., 1982; Newman et al., 2012) By automating the production of mechanical components through pre-programmed CAD models, CNC machines provide the precision, consistency, and flexibility needed for efficient mass production (Soori et al., 2023, 2024). Their role continues to grow as manufacturers adopt advanced technologies to remain competitive in an evolving industrial landscape (Ye et al., 2018)

CNC machine maintenance is critical to ensuring the reliability and productivity of these systems. Three primary maintenance strategies are commonly used: reactive, preventive, and predictive maintenance (Muchiri et al., 2014; Roosefert Mohan et al., 2021). Reactive or breakdown maintenance involves repairing machines only after they fail, often leading to costly downtime and potential damage. Preventive maintenance takes a more proactive approach by scheduling routine servicing based on a fixed timeline, regardless of the machine's actual condition, reducing breakdowns but sometimes resulting in unnecessary maintenance. The most advanced strategy, predictive maintenance, uses real-time data and analytics to monitor machine conditions, predicting failures before they occur. This approach minimizes downtime and costs by addressing issues, if necessary, based on actual wear and usage.

Despite advancements in maintenance strategies, unplanned downtime, unexpected machine stoppages caused by malfunctions or failures, remains a critical challenge in manufacturing (Cochran et al., 2001). Such interruptions reduce overall equipment effectiveness (OEE), slow production, delay deliveries, and generate substantial financial losses, with some manufacturers reporting costs of up to $260,000 per hour (output.industries, 2024; Xiao et al., 2021). The issue is especially severe in high-mix, low-volume environments where operational flexibility is essential (Jauregui Becker et al., 2015).

These significant operational and financial risks underscore the need to accurately understand and predict downtime occurrences. Analyzing patterns in downtime frequency, duration, causes, and trends provides insights into dominant failure modes, supports resource prioritization, and improves asset management decisions (Muchiri et al., 2014). However, reactive approaches based solely on historical data struggle to anticipate random failures, such as sudden component breakdowns or power disturbances, making them inadequate for preventing costly disruptions (Alaswad & Xiang, 2017).

To address this issue, incorporating predictive analytics, real-time monitoring, and sensor-based systems can provide timely alerts and early warnings of potential issues, preventing downtime before it occurs (Lee et al., 2014). Maintenance strategies shift from reactive firefighting to proactive risk management by estimating the likelihood of failures in advance. Predictive modeling techniques, including machine learning and deep learning, have recently gained significant traction for their ability to enhance predictive maintenance across various industries (Hussain & Jan, 2019; Wan et al., 2019).

Predictive maintenance methods commonly include both classical approaches and advanced machine learning techniques. Time series forecasting techniques, such as autoregressive integrated moving averages (ARIMA), have traditionally been used to predict downtime and failure events based on historical data (Alaswad & Xiang, 2017) and (Traini et al., 2019). Pavlyshenko, (2019) mentioned that ARIMA for prediction model is closer to results of neural networks, which has the added advantage of less memory space and computational time. However, ARIMA's linear assumptions limit its effectiveness in handling the nonlinear and noisy data typical of CNC operations (Özel & Karpat, 2005; Sun et al., 2006). Complex interrelationships, such as the non-linear effects of tool wear and cutting conditions, often require more sophisticated models.

This is where machine learning (ML) comes into play, which offers significant advantages in predictive maintenance by addressing the complexities that classical approaches struggle to capture (Pimenov et al., 2023; Soori et al., 2023). ML, as a branch of artificial intelligence (AI), allows machines to learn from data, identify patterns, and make decisions without human intervention. ML algorithms can capture complex relationships between machine conditions, operational parameters, and failure modes, which lead to more precise predictions of equipment failures and maintenance needs (Soori et al., 2024). This capability is important for enhancing CNC machine operations, where real-time data analysis can help optimize processes, reduce downtime, and improve overall efficiency.

For example, Support Vector Machines (SVM) are used for condition monitoring and predicting the remaining useful life of machines. At the same time, Long Short-Term Memory (LSTM) networks handle time-series data effectively, especially in capturing long-term dependencies in nonlinear datasets. Other algorithms, including neural networks (Widodo & Yang, 2011), recurrent neural networks (RNN) (Malhotra et al., 2016), convolutional neural networks (CNN), and tree-based ensemble methods like Random Forest (Zhou et al., 2019), excel at processing noisy and non-linear time series data. Overall, data-driven approaches like deep learning have demonstrated superior performance to traditional model-based methods in predictive maintenance tasks, particularly in minimizing human intervention and optimizing machine performance.

However, there is still a gap in the existing literature on the lack of comprehensive, data-driven models that can forecast overall CNC machine downtime arising from diverse failure sources. Prior studies to date have focused on modeling specific failure modes, such as tool wear (Abellan-Nebot et al., 2012) or software crashes (Soori et al., 2024) using domain-specific approaches. While these studies provide valuable insights, unplanned CNC downtime often stems from various complex and interrelated causes, including mechanical breakdowns, electrical faults, and operational errors (Djurdjanovic et al., 2003). Each failure type may exhibit distinct patterns and trends not captured by models designed to predict individual failure modes (Jantunen, 2002).

There is a need for ML techniques capable of handling heterogeneous industrial failure data to predict overall CNC downtime. Such models would provide a more comprehensive view of multiple failure types, allowing for more effective maintenance planning and overall uptime improvement. Current models optimized for single failure modes fail to address the broader, interconnected nature of CNC machine failures.

Furthermore, while many studies have explored predictive maintenance for CNC machines, the rapidly evolving manufacturing environment and the complexity of machine operations require continuous exploration and refinement of ML algorithms (references). Downtime patterns vary based on machine age, usage intensity, and operational conditions. Advances in ML algorithms offer opportunities for improved accuracy, faster computation, and better generalization across diverse datasets. Research into more sophisticated and flexible models is essential for enhancing machine reliability and reducing production disruptions in real-world applications.

This study addresses this research gap by developing ML models that use heterogeneous downtime data from multiple failure sources, like mechanical failures, electrical faults, software errors, etc, to predict aggregated CNC downtime trends. Therefore, this study develops and compares two ensemble learning models: Random Forest and XGBoost, to forecast total CNC downtime using heterogeneous failure data from 16 machines. The contribution is twofold: (1) providing a plant-level downtime forecasting model that accounts for multiple failure sources, and (2) demonstrating the superiority of an optimized XGBoost model for industrial downtime prediction.

Method

This study adopts a quantitative approach, utilizing time series analysis and machine learning techniques to forecast downtime trends in CNC machining operations. The research focuses on predicting unplanned downtime, a critical issue in manufacturing industries where machine availability is key to maintaining productivity. The study utilizes secondary data collected from multiple CNC machines at Company A in Malaysia, consisting of operational and maintenance records over an extended period., which contains operational and maintenance records over an extended period. This dataset provides a comprehensive basis for understanding historical downtime patterns and informing predictive models. Figure 1 summarizes the systematic methodology, aligning with the standard data science process. The process begins with data collection and preparation, which involves cleaning, transforming, and organizing the raw data into a format suitable for analysis. In this case, the secondary data includes various machine-related variables, such as operational hours, instances of downtime, environmental factors, and maintenance logs. The data is prepared by removing outliers, handling missing values, and ensuring consistency across different periods to improve the accuracy of the subsequent analysis

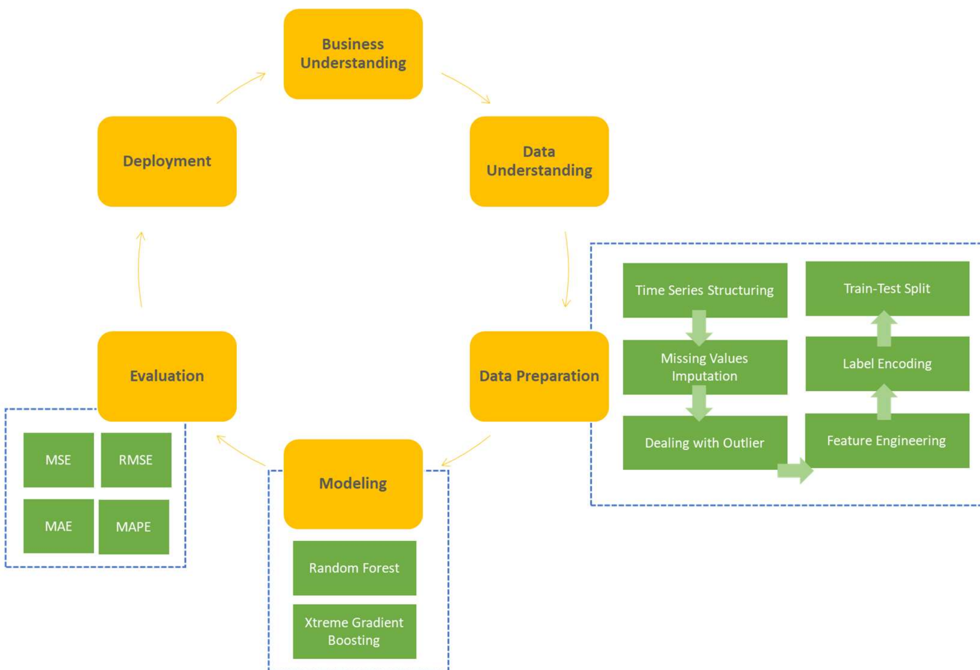

Figure 1. Research Methodology flow

Data Collection

The secondary data for this study were collected from Company A's historical CNC maintenance records, covering approximately 14 months, from January 2022 to early March 2023. The process of downtime data collection is illustrated in Figure 2. The dataset, obtained with permission from the company's technical support team, includes detailed reports of downtime events across 16 CNC machining assets in the facility.

The dataset captures key variables for each downtime event, such as machine ID, failure type, start date/time, end date/time, and downtime duration in minutes. Failure types are categorized into several groups, including mechanical, electrical, and tool malfunctions. Data was collected daily, with downtime incidents consolidated into a single record per day. This level of granularity enables a comprehensive analysis and model development aimed at accurately forecasting future downtime events.

The dataset contains 458 daily records, with 65 instances of CNC machine downtime. Each record specifies the downtime duration for that day and includes details of any logged failure events. Days without downtime incidents were also recorded, indicating zero downtime hours. This complete dataset allows for a thorough analysis of CNC downtime patterns over 14 months, providing valuable insights for predictive modeling.

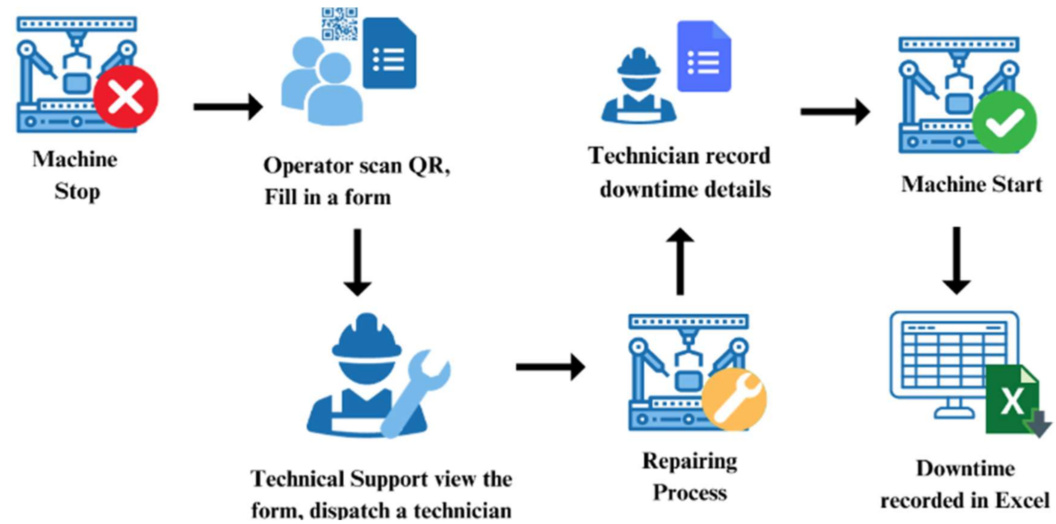

Figure 1. Downtime Data Collection Process Flow

DATE	MACHINE NO.	PROBLEM TYPE	TIME			DOWN TIME		
			START DOWN	START REPAIR	JOB COMPLETE	DOWN TIME	REPAIRING TIME	IDLE TIME
1/7/2022	CNC 1	SLANT LEADS	8:00:00 AM	8:00:00 AM	8:10:00 AM	0.17 hrs	0.17 hrs	0.00 hrs
1/7/2022	CNC 11	SLANT LEADS	8:00:00 AM	8:10:00 AM	8:20:00 AM	0.33 hrs	0.17 hrs	0.17 hrs
1/24/2022	CNC 5	UNEVEN WINDING	8:00:00 AM	8:06:00 AM	8:35:00 AM	0.58 hrs	0.48 hrs	0.10 hrs
1/25/2022	CNC 5	SLANT LEADS	12:45:00 AM	1:00:00 AM	2:00:00 AM	1.25 hrs	1.00 hrs	0.25 hrs
1/25/2022	CNC 10	SLANT LEADS	1:00:00 AM	3:00:00 AM	3:30:00 AM	2.50 hrs	0.50 hrs	2.00 hrs
1/25/2022	CNC 5	MECHANICAL	3:00:00 AM	3:00:00 AM	3:45:00 AM	0.75 hrs	0.75 hrs	0.00 hrs
1/26/2022	CNC 4	TOOLING	11:00:00 AM	12:00:00 PM	12:30:00 PM	1.50 hrs	0.50 hrs	1.00 hrs
2/6/2022	CNC 11	SLANT LEADS	9:15:00 PM	9:50:00 PM	10:05:00 PM	0.83 hrs	0.25 hrs	0.58 hrs
2/8/2022	CNC 1	STRIPPING	1:30:00 AM	1:35:00 AM	1:57:00 AM	0.45 hrs	0.37 hrs	0.08 hrs
2/8/2022	CNC 11	SLANT LEADS	1:00:00 AM	1:58:00 AM	2:10:00 AM	1.17 hrs	0.20 hrs	0.97 hrs
2/10/2022	CNC 16	STRIPPING	1:00:00 AM	1:05:00 AM	1:20:00 AM	0.33 hrs	0.25 hrs	0.08 hrs
2/10/2022	CNC 13	MECHANICAL	1:00:00 AM	1:20:00 AM	1:30:00 AM	0.50 hrs	0.17 hrs	0.33 hrs
2/16/2022	CNC 2	STRIPPING	8:30:00 AM	8:45:00 AM	9:10:00 AM	0.67 hrs	0.42 hrs	0.25 hrs
3/11/2022	CNC 1	STRIPPING	8:20:00 AM	8:25:00 AM	9:05:00 AM	0.75 hrs	0.67 hrs	0.08 hrs
3/18/2022	CNC 16	UNEVEN WINDING	12:00:00 PM	2:30:00 PM	3:20:00 PM	3.33 hrs	0.83 hrs	2.50 hrs
3/22/2022	CNC 14	STRIPPING	9:00:00 AM	9:15:00 AM	11:00:00 AM	2.00 hrs	1.75 hrs	0.25 hrs
3/24/2022	CNC 7	SLANT LEADS	8:00:00 PM	8:15:00 PM	9:00:00 PM	1.00 hrs	0.75 hrs	0.25 hrs
3/24/2022	CNC 2	STRIPPING	11:45:00 PM	12:30:00 AM	2:50:00 AM	3.08 hrs	2.33 hrs	0.75 hrs

Figure 2. Sample of Manufacturing plant CNC downtime record from January 2022 to early March 2023

Data Preparation

The raw CNC downtime data from Company A's maintenance records required substantial preprocessing steps before time series modeling and forecasting (see Figure 3). While containing the historical timeline of downtime events, the raw data had several limitations, making it unsuitable for downtime occurrence predictive modeling without preprocessing:

- Irregularity: The records were unevenly spaced in time, with gaps on failure-free days. This posed challenges for forecasting algorithms relying on consistent time intervals between observations.
- Lack of explanatory variable: Only fundamental downtime duration values were captured, without additional temporal or explanatory variables needed to model failure patterns.
- Structure: Machine IDs and failure types were stored as unintelligible string labels rather than encoded categorical variables.
- Data leakage: Temporal interrelationships could be exploited if time series splitting was not done sequentially.
- Non-normality: Downtime duration data was right-skewed, needing transforms for effective modeling.

To address these challenges and prepare the data for time series modeling, the following critical preprocessing steps were implemented:

- Time Series Restructuring: The raw data contained temporal gaps, which can degrade forecasting performance. Restructuring into an evenly spaced series ensured consistent intervals between observations critical for time series algorithms.
- Feature Engineering: Additional variables representing temporal cycles and failure intervals were derived to provide more explanatory power for modeling downtime patterns over time.
- Encoding: Converting machine IDs and failure types into numeric categories enabled the models to recognize these variables' distinct values and patterns (see Tables 1 and 2).
- Train-Test Split: Temporal splitting prevented data leakage and overfitting by ensuring the models were evaluated on unseen future data.
- Transformation: Data distributions that do not match model assumptions can reduce accuracy. Log transforms normalized the raw downtime data to fit the required data normality.

After preprocessing, the engineered dataset had the necessary structure, explanatory variables, and time-ordered train-test split to enable robust machine learning-based forecasting of CNC downtime.

Table 1. Encoded Problem Type

Problem type	Encoded label
Slant leads	1
Uneven Winding	2
Mechanical	3
Tooling	4
Stripping	5
Electrical	6
Red Lead	7
Damage Air Coil	8
Pneumatic	9
Uneven lead length	10
Wire Scratch	11

Table 2. Encoded Machine Number

Machine Number	Encoded Label
CNC 1	1
CNC 2	2
CNC 3	3
CNC 4	4

Machine Number	Encoded Label
CNC 5	5
CNC 6	6
CNC 7	7
CNC 8	8
CNC 9	9
CNC 10	10
CNC 11	11
CNC 12	12
CNC 13	13
CNC 14	14
CNC 15	15
CNC 16	16

Generating Predictor Using Probabilistic Approach

Two supplementary predictors—machine number and problem type—were introduced using a probabilistic approach to enhance the forecasting models with additional helpful information for predicting downtime duration. This step was necessary because the machine number and specific problem type were only determined once a downtime event occurred and thus could not be known in advance. To account for this uncertainty, the probability p of downtime occurrence for each machine i was calculated using *Eq* .1-3 with the results shown in Table 3:

$$p(\text{Downtime Occurence}) = \frac{\text{Number of Downtime}}{\text{Total Number of Days}} \tag{1}$$

$$p(\text{Downtime Occurence}|\text{Machine } i) = \frac{\text{Count of machine } i \text{ downtime occurence}}{\text{Count of downtime occurrence}} \tag{2}$$

$$p(\text{Problem } j|\text{Machine } i) = \frac{\text{Count of machine } i \text{ downtime occurence with problem } j}{\text{Count of downtime occurrence}} \tag{3}$$

Where : i =1, 2, …, 16 and j = 1, 2, …,11

Table 2. Generated Probability of Machine Downtime Given Machine Number and Problem Type

	Machine Number												
Prb	1	2	3	4	5	6	7	8	10	11	13	14	16
1	0.182	0.000	0.000	0.111	0.250	1.000	1.000	0.000	0.600	0.333	0.000	0.000	0.500
2	0.091	0.000	0.500	0.111	0.250	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.167
3	0.091	0.500	0.000	0.111	0.250	0.000	0.000	0.000	0.000	0.111	0.500	0.000	0.000
4	0.000	0.000	0.000	0.111	0.000	0.000	0.000	0.667	0.200	0.000	0.250	0.000	0.000
5	0.546	0.333	0.000	0.444	0.250	0.000	0.000	0.000	0.000	0.222	0.000	0.000	0.167
6	0.091	0.000	0.000	0.111	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.111	0.000	0.000	0.000
8	0.000	0.167	0.500	0.000	0.000	0.000	0.000	0.333	0.000	0.111	0.000	0.500	0.167
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.200	0.000	0.000	0.000	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.111	0.000	0.000	0.000

Model Development

Given the study's focus on comparing machine learning techniques for downtime forecasting, two prominent algorithms—Random Forest (RF) and eXtreme Gradient Boosting (XGBoost)—were selected for model development. These algorithms were chosen for their ability to handle complex, nonlinear relationships in data and their proven effectiveness in predictive maintenance applications.

Ensemble Learning Algorithms

Random Forest

Random Forest (RF) is an ensemble learning method that builds multiple decision trees on bootstrapped subsets of the training data and aggregates their predictions to produce a final output (see Figure 4). Each tree is trained on a randomly sampled dataset and uses a random subset of features at each split, promoting diversity and reducing overfitting. During prediction, the outputs of all trees are averaged, resulting in a more stable and accurate forecast.

RF is well suited for downtime forecasting because it can capture nonlinear relationships and high-order interactions inherent in complex industrial data (Genuer et al., 2010). It is robust to noise and outliers, limits overfitting through controlled tree depth, and enables efficient parallel training (Breiman, 2001). These characteristics make RF a strong candidate for forecasting tasks involving irregular and noisy temporal patterns.

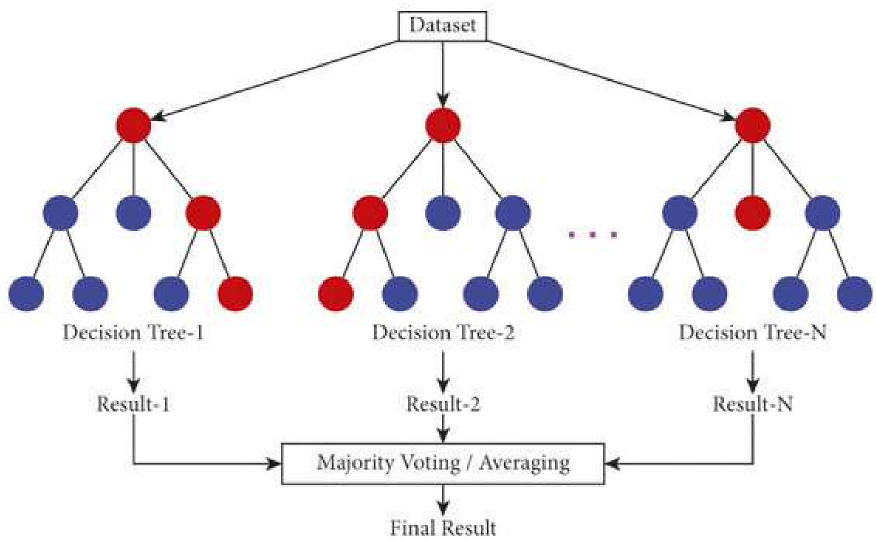

Figure 4. Graphical Representation of Random Forest (Khan et al., 2021).

Extreme Gradient Boosting (XGBoost)

XGBoost is a tree-based ensemble method that applies gradient boosting to iteratively improve model accuracy (Chen & Guestrin, 2016) (see Figure 5). Unlike standard boosting, it incorporates strong regularization—such as shrinkage and tree pruning—to reduce overfitting as new trees are added. The algorithm also uses second-order gradient information to more efficiently minimize the loss function, enabling faster and more precise optimization (Mitchell & Frank, 2017) XGBoost offers several computational advantages, including parallel processing, approximate tree learning for speed, and the ability to handle missing data natively. Subsampling of both data instances and features further enhances model generalization. These characteristics make XGBoost particularly effective for time-series forecasting, where its combination of gradient boosting, regularization, and computational efficiency consistently yields high predictive accuracy.

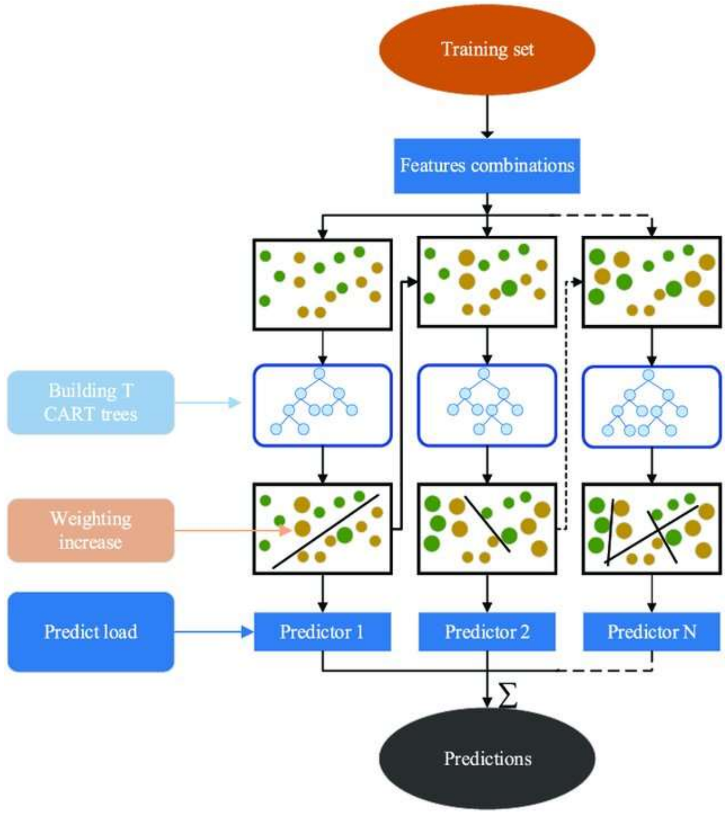

Figure 5. Graphical Representation of XGBoost (Yao et al., 2022).

Model Training

The pre-processed downtime dataset was split into contiguous train and test sets using a rolling origin scheme to simulate real-world iterative forecasting. The first 80% of the time series was the initial training set, with the next 20% as the test set.

The Random Forest and XGBoost models were trained on the training data using 5-fold stratified cross-validation with early stopping to prevent overfitting. In k-fold cross-validation, the training data is split into k groups or folds. Each fold is held out as the validation set, while the remaining *k*-1 folds are used to train the model. Early stopping tracks validation performance, stopping training if metrics worsen for a defined number of iterations to avoid overfitting the training data.

In addition, a rolling origin evaluation was implemented where the test set is progressively rolled forward in time. The models are retrained on newly available data at each iteration to simulate real-world systems where forecasts are generated continuously as new data arrives. Model performance is evaluated by predicting each test partition. This training approach, through temporal cross-validation, early stopping, and rolling origin testing, evaluates the models' ability to generalize to unseen data, which is critical for reliable forecasting performance. It mimics how the models would be retrained and generate forecasts in an operational environment.

Hyperparameter Tuning (Random Search CV)

The Random Search Cross-Validation optimizer improved both RF and XGBoost models. The open-source Python machine learning toolkit *scikit-learn* offers methods for adjusting model hyperparameters. The complete model optimization procedure employing Random Search CV may be summed up as follows:

Step 1: A search space is a bounded domain of hyperparameter values.

Step 2: Sample that domain's points at random.

Step 3: Create a random search instance and fit it to a model using Scikit-Learn. Indicate the number of cross-validation folds and the number of iterations or possible combinations to try.
Step 4: Get the best parameters and model from the optimizer. Compare the base and top random search models to see if the random search produced a superior model.

Results and Discussion

The Results

The study compared Random Forest and XGBoost models for forecasting CNC downtime using 14 months of maintenance records from Company A. As shown in Table 4, XGBoost substantially outperformed RF across all error metrics. Based on these metrics, the Random Forest model performed better in minimizing forecasting errors for the CNC downtime duration data before tuning. The lower values indicate that the Random Forest predictions were numerically closer to the downtime duration.

However, upon further diagnostic analysis using learning curves, overfitting tendencies were revealed in both models. The learning curves plotted model performance on the training set and a held-out validation set as training progressed (see Figures 6 and 7). For the Random Forest model, the training MSE stabilized at around 0.5 after 50 estimators, while the validation MSE remained around 2.5 with no convergence. The XGBoost learning curve showed a similar trend, with the training MSE reaching 0 after 100 iterations but validation MSE stabilizing at a higher value of around 3.5.

These diverging training and validation errors indicate that despite achieving low training MSE, both models were overfitting the data. They failed to generalize well to new unseen data, necessitating corrective measures.

Model Optimization

The pre-processed downtime data were divided into sequential training and test sets using a rolling-origin scheme to reflect real-world forecasting. Hyperparameter tuning was then conducted using Randomized Search Cross-Validation to reduce overfitting and identify configurations that generalized well to unseen data.

As shown in Table 4, the tuned XGBoost model substantially outperformed the Random Forest model across all error metrics. Its lower MAE, equivalent to an average prediction error of under 30 minutes, offers meaningful operational advantages. With more reliable downtime forecasts, maintenance teams can plan interventions earlier, prepare spare parts in advance, and allocate personnel more efficiently, reducing the likelihood of prolonged or cascading disruptions. In contrast, the Random Forest model retained higher error levels and exhibited a wider gap between training and validation performance, indicating weaker generalization.

Table 3. Model Error Performance Metrics

	Random Forest		XGBoost	
	Before Tuning	**After Tuning**	**Before Tuning**	**After Tuning**
MSE	5.858	1.265	6.656	0.402
RMSE	2.420	1.125	2.580	0.634
MAE	1.490	0.943	1.563	0.470
MAPE	1.011	1.754	0.986	0.719

The predicted-versus-actual plot (Figure 10) further illustrates XGBoost's ability to follow observed downtime patterns closely, with only a few deviations likely attributable to outlier events. This alignment reinforces the model's practical utility: accurate short-horizon forecasts can support proactive scheduling and improve overall equipment availability in environments where unexpected stoppages are costly.

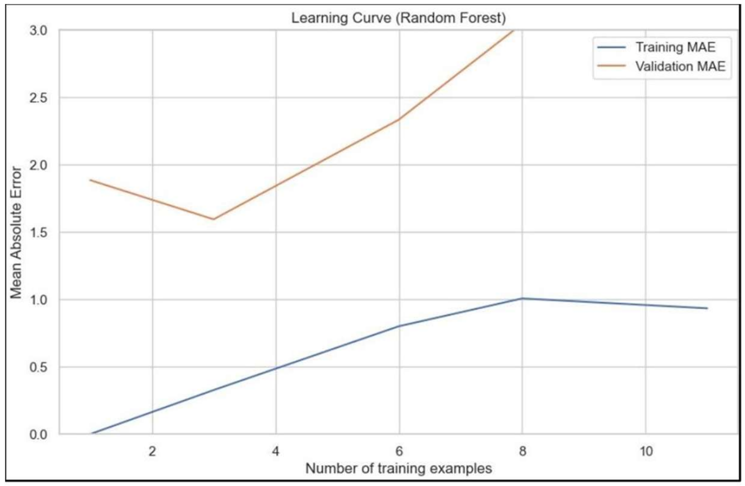

Figure 6. RF learning curve before hyperparameter tuning.

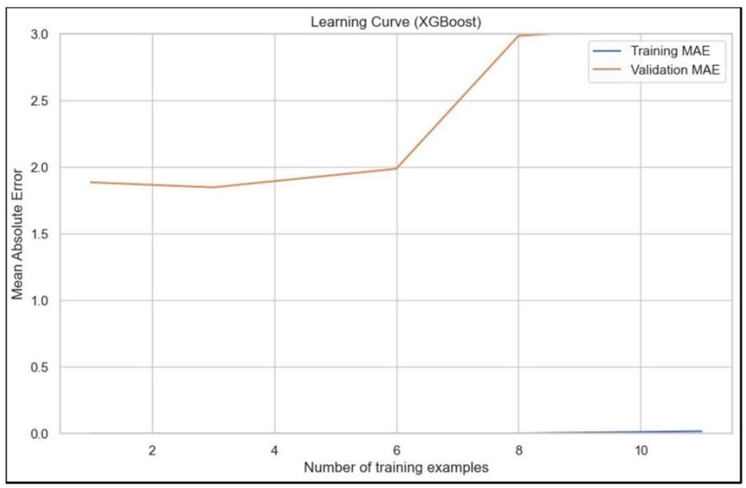

Figure 7. XGBoost learning curve before hyperparameter tuning.

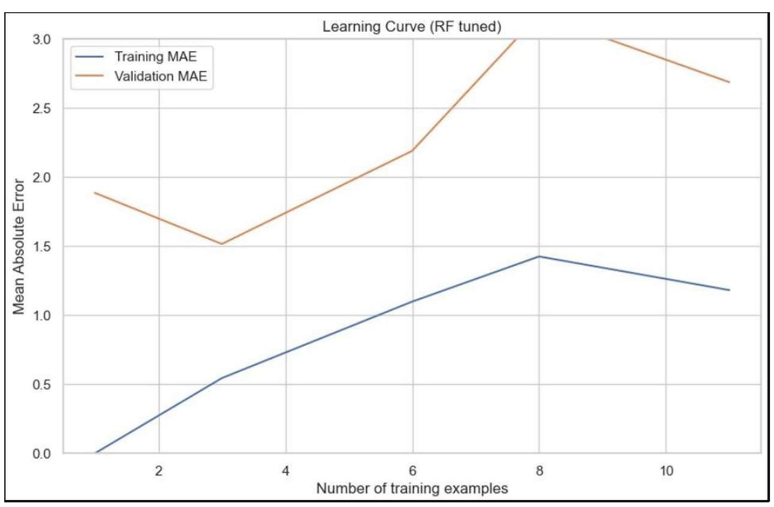

Figure 8. RF learning curve after hyperparameter tuning.

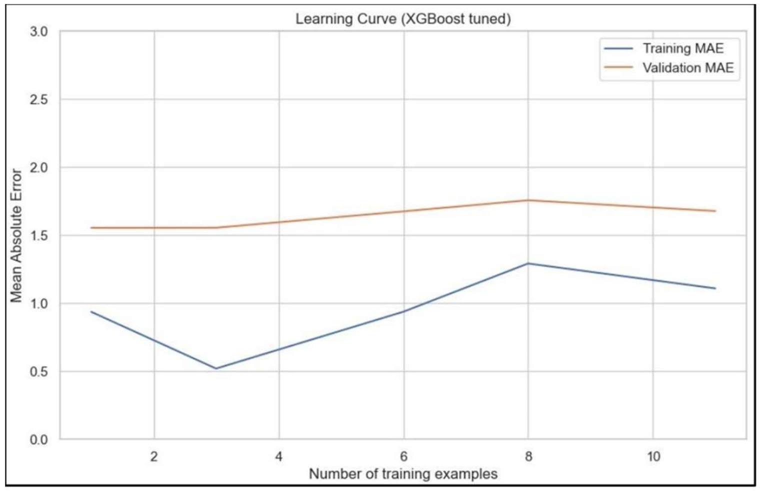

Figure 9. XGBoost learning curve after hyperparameter tuning.

Table 4. Predicted and actual downtime duration on the test set

Date	Downtime Duration (Hours)		
	XGB Predicted	RF Predicted	Actual
2023-01-08	1.01	1.15	1.08
2023-01-13	1.48	2.86	0.92
2023-01-24	0.96	2.01	0.17
2023-02-10	1.05	2.21	0.50
2023-02-16	1.06	1.64	0.75
2023-02-19	0.84	1.25	2.33
2023-02-23	1.01	1.15	1.00
2023-03-05	1.05	1.84	0.67
2023-03-06	1.47	2.40	1.25
2023-03-08	1.17	3.29	2.33
2023-03-09	1.00	1.59	0.85
2023-03-10	0.97	1.02	1.17
2023-03-11	1.02	1.23	0.83

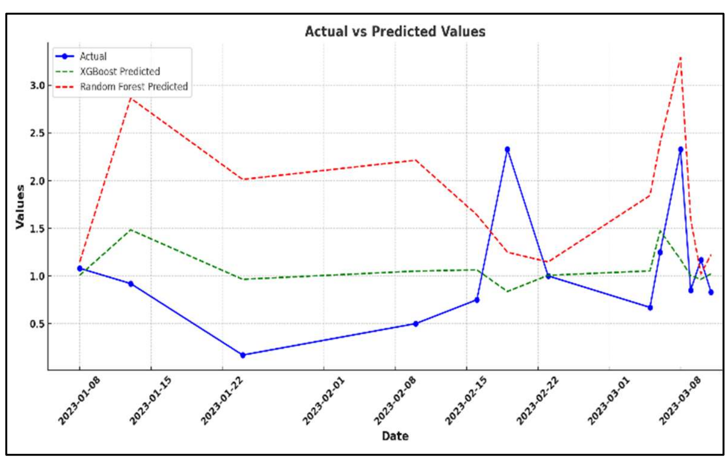

Figure 10. Predicted vs actual plot.

Discussion

This study demonstrated that machine learning models can effectively forecast CNC machine downtime trends when trained on historical maintenance records. The optimized XGBoost model outperformed the Random Forest model in predicting future downtime duration. XGBoost's superior accuracy can be attributed to its gradient-boosting approach, which iteratively minimizes errors through additive expansion. The model's regularized objective function helps control complexity and prevent overfitting, while hyperparameter tuning, performed using randomized search, identified the optimal configuration for the dataset. This combination of boosting, regularization, and hyperparameter optimization allowed XGBoost to learn downtime patterns from the time series data effectively.

In contrast, the initial Random Forest model exhibited signs of overfitting, as reflected by the large gap between training and validation errors. This was likely due to the limited size of the dataset and the absence of early hyperparameter tuning. While the Random Forest model's bagging process creates highly variant decision trees, this variability can result in overfitting when dealing with noisy training data. After tuning, the overfitting was reduced, but XGBoost still achieved lower prediction errors, highlighting the strength of gradient boosting in handling time series forecasting problems involving sparse and irregular data.

Key downtime trends revealed by this study, such as the right-skewed distribution of event durations and the top recurring failure types, are consistent with findings from (Abdallah et al., 2016) and (Lee et al., 2014). However, those studies used traditional statistical forecasting models like ARIMA rather than advanced machine learning techniques. The stable and generalized performance of XGBoost underscores its advantage in capturing complex temporal relationships and improving the accuracy of downtime forecasting.

The practical implication of these findings is that enhanced forecast accuracy can help maintenance teams intervene earlier, reducing unplanned downtime, preventing cascading delays, and potentially lowering associated operational costs. In industrial settings where downtime is extremely expensive, even modest improvements in prediction error can yield significant benefits

One limitation of this study is the relatively small dataset collected from a single manufacturer over one year. Expanding the dataset to include longer durations and data from multiple manufacturers would allow for more robust model performance validation and testing. Besides, incorporating more explanatory variables, such as maintenance history, operating conditions, and detailed failure data, could improve the models' predictive accuracy. Future research should explore these improvements.

Another limitation is that the study primarily focused on predicting downtime duration without delving into the specific causes of failures. Failure type prediction or root cause analysis could provide more actionable insights for targeted maintenance strategies. Additionally, the absence of real-time data in the current model limits its practical application in dynamic environments where immediate decision-making is critical. Future studies should explore integrating real-time sensor data and Internet of Things (IoT) technologies, which could further enhance the model's applicability and responsiveness.

While this study demonstrates the potential of gradient-boosting models for predicting CNC downtime trends, the current approach remains in its early stages. With further refinement and larger datasets, this methodology could contribute to more intelligent maintenance scheduling, improving machine availability and operational efficiency. However, translating this approach into an AI-based decision support system for industrial engineers would require significant additional development, including integration with real-time data and testing in diverse manufacturing environments.

Conclusion

This study demonstrates that XGBoost provides superior performance for forecasting CNC machine downtime using heterogeneous industrial failure data. Beyond statistical improvements, the model's accuracy, particularly its average error of less than 30 minutes, offers practical value for predictive maintenance. More precise forecasts allow maintenance teams to plan interventions earlier, allocate technicians and spare parts more effectively, and mitigate the operational disruptions associated with unexpected stoppages. In manufacturing environments where downtime is costly, even modest gains in prediction accuracy can significantly improve equipment availability and production continuity.

Theoretically, this study contributes a plant-level forecasting approach that integrates multiple failure sources, addressing a gap in prior research that typically focuses on isolated failure modes. The comparison of Random Forest and XGBoost under a rolling-origin evaluation also provides methodological insight into the suitability of ensemble learning for downtime prediction in real industrial settings.

Future work should incorporate real-time sensor data, additional operational variables, and larger multi-site datasets to enhance generalizability. Overall, the findings highlight the potential of advanced machine learning models to support data-driven maintenance decisions and strengthen operational efficiency in modern manufacturing environments.

Declaration of Conflict of Interest

The authors declare that there are no potential conflicts of interest related to this article's research, writing, and/or publication.

References

Abdallah, A. B., Maaroufi, M., & Ouali, M. S. (2016). Dealing with machine deterioration and failures in manufacturing systems. *International Journal of Production Research, 54*(23), 7245–7272.

Abellán-Nebot, J. V., Liu, J., & Romero Subirón, F. (2012). Quality prediction and compensation in multi-station machining processes using sensor-based fixtures. *Robotics and Computer-Integrated Manufacturing, 28*(2), 208–219. https://doi.org/10.1016/j.rcim.2011.09.001

Alaswad, S., & Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating systems. *Reliability Engineering & System Safety, 157*, 54–63. https://doi.org/10.1016/j.ress.2016.08.009

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Dalam *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Cochran, D. S., Arinez, J. F., Duda, J. W., & Linck, J. (2001). A decomposition approach for manufacturing system design. *Journal of Manufacturing Systems, 20*(6), 371–389. https://doi.org/10.1016/S0278-6125(01)80058-3

Djurdjanovic, D., Lee, J., & Ni, J. (2003). Watchdog agent—An infotronics-based prognostics approach for product performance degradation assessment and prediction. *Advanced Engineering Informatics, 17*(3–4), 109–125. https://doi.org/10.1016/j.aei.2004.07.005

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters, 31*(14), 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

Hussain, Z., & Jan, H. (2019). Establishing simulation model for optimizing efficiency of CNC machine using reliability-centered maintenance approach. *International Journal of Modeling, Simulation, and Scientific Computing, 10*(6), 1950034. https://doi.org/10.1142/S179396231950034X

Jantunen, E. (2002). A summary of methods applied to tool condition monitoring in drilling. *International Journal of Machine Tools and Manufacture, 42*(9), 997–1010. https://doi.org/10.1016/S0890-6955(02)00040-8

Jauregui Becker, J. M., Borst, J., & van der Veen, A. (2015). Improving the overall equipment effectiveness in high-mix–low-volume manufacturing environments. *CIRP Annals, 64*(1), 419–422. https://doi.org/10.1016/j.cirp.2015.04.126

Keller, A. Z., Kamath, A. R. R., & Perera, U. D. (1982). Reliability analysis of CNC machine tools. *Reliability Engineering, 3*(6), 449–473. https://doi.org/10.1016/0143-8174(82)90036-1

Khan, M. Y., Qayoom, A., Nizami, M. S., Siddiqui, M. S., Wasi, S., & Raazi, S. M. K.-R. (2021). Automated prediction of good dictionary examples (GDEX): A comprehensive experiment with distant supervision, machine learning, and word embedding-based deep learning techniques. *Complexity, 2021*, Article 2553199. https://doi.org/10.1155/2021/2553199

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems: Reviews, methodology, and applications. *Mechanical Systems and Signal Processing, 42*(1–2), 314–334. https://doi.org/10.1016/j.ymssp.2013.06.004

Malhotra, P., TV, V., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). Multi-sensor prognostics using an unsupervised health index based on LSTM encoder–decoder. *arXiv*. https://doi.org/10.48550/arXiv.1608.06154

Mitchell, R., & Frank, E. (2017). Accelerating the XGBoost algorithm using GPU computing. *PeerJ Computer Science, 3*, e127. https://doi.org/10.7717/peerj-cs.127

Muchiri, P. N., Pintelon, L., Martin, H., & Chemweno, P. (2014). Modelling maintenance effects on manufacturing equipment performance: Results from simulation analysis. *International Journal of Production Research, 52*(11), 3287–3302. https://doi.org/10.1080/00207543.2013.870673

Newman, S. T., Nassehi, A., Imani-Asrai, R., & Dhokia, V. (2012). Energy efficient process planning for CNC machining. *CIRP Journal of Manufacturing Science and Technology, 5*(2), 127–136. https://doi.org/10.1016/j.cirpj.2012.03.007

Özel, T., & Karpat, Y. (2005). Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. *International Journal of Machine Tools and Manufacture, 45*(4–5), 467–479. https://doi.org/10.1016/j.ijmachtools.2004.09.007

Output Industries. (2024). *Understanding the real costs of machine downtime in manufacturing*. https://www.output.industries/insights/costs-of-machine-downtime-in-manufacturing

Pavlyshenko, B. (2019). Machine-learning models for sales time series forecasting. *Data, 4*(1), 15. https://doi.org/10.3390/data4010015

Pimenov, D. Y., Bustillo, A., Wojciechowski, S., Sharma, V. S., Gupta, M. K., & Kuntoğlu, M. (2023). Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review. *Journal of Intelligent Manufacturing, 34*(5), 2079–2121. https://doi.org/10.1007/s10845-022-01923-2

Roosefert Mohan, T., Preetha Roselyn, J., Annie Uthra, R., Devaraj, D., & Umachandran, K. (2021). Intelligent machine learning-based total productive maintenance approach for achieving zero downtime in industrial machinery. *Computers & Industrial Engineering, 157*, 107267. https://doi.org/10.1016/j.cie.2021.107267

Soori, M., Arezoo, B., & Dastres, R. (2023). Machine learning and artificial intelligence in CNC machine tools: A review. *Sustainable Manufacturing and Service Economics, 2*, 100009. https://doi.org/10.1016/j.smse.2023.100009

Soori, M., Ghaleh Jough, F. K., Dastres, R., & Arezoo, B. (2024). Sustainable CNC machining operations: A review. *Sustainable Operations and Computers, 5*, 73–87. https://doi.org/10.1016/j.susoc.2024.01.001

Sun, Y., Ma, L., Mathew, J., Wang, W., & Zhang, S. (2006). Mechanical systems hazard estimation using condition monitoring. *Mechanical Systems and Signal Processing, 20*(5), 1189–1201. https://doi.org/10.1016/j.ymssp.2004.10.009

Traini, E., Bruno, G., D'Antonio, G., & Lombardi, F. (2019). Machine learning framework for predictive maintenance in milling. *IFAC-PapersOnLine, 52*(13), 177–182. https://doi.org/10.1016/j.ifacol.2019.11.172

Wan, S., Li, D., Gao, J., & Li, J. (2019). A knowledge-based machine tool maintenance planning system using case-based reasoning techniques. *Robotics and Computer-Integrated Manufacturing, 58*, 80–96. https://doi.org/10.1016/j.rcim.2019.01.012

Widodo, A., & Yang, B.-S. (2011). Machine health prognostics using survival probability and support vector machine. *Expert Systems with Applications, 38*(7), 8430–8437. https://doi.org/10.1016/j.eswa.2011.01.038

Xiao, Y., Jiang, Z., Gu, Q., Yan, W., & Wang, R. (2021). A novel approach to CNC machining center processing parameters optimization considering energy-saving and low-cost. *Journal of Manufacturing Systems, 59*, 535–548. https://doi.org/10.1016/j.jmsy.2021.03.023

Yao, X., Fu, X., & Zong, C. (2022). Short-term load forecasting method based on feature preference strategy and LightGBM–XGBoost. *IEEE Access, 10*, 75257–75268. https://doi.org/10.1109/ACCESS.2022.3192011

Ye, Y., Hu, T., Zhang, C., & Luo, W. (2018). Design and development of a CNC machining process knowledge base using cloud technology. *The International Journal of Advanced Manufacturing Technology, 94*(9–12), 3413–3425. https://doi.org/10.1007/s00170-016-9338-1

Zhou, P., Li, Z., Snowling, S., Baetz, B. W., Na, D., & Boyd, G. (2019). A random forest model for inflow prediction at wastewater treatment plants. *Stochastic Environmental Research and Risk Assessment, 33*(10), 1781–1792. https://doi.org/10.1007/s00477-019-01732-9