International Journal on Emerging Technologies 16(1): 45-53(2025)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Analysis of Predictive Mechanical Maintenance using Artificial Intelligence, Machine Learning and Data Science

Suraj Shashikant Patil[1*] and Gajanan N. Thokal[2]
[1]M.Tech. Scholar, Department of Mechanical Engineering,
Pillai College of Engineering, New Panvel (Maharashtra), India.
[2]Associate Professor, Department of Mechanical Engineering,
Pillai College of Engineering, New Panvel (Maharashtra), India.

(Corresponding author: Suraj Shashikant Patil)*
(Received 22 October 2024, Revised 20 December 2024, Accepted 15 January 2025)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: It is quite clear that machine downtime due to a sudden machinery breakdown will cost the organization a lot of money. Organizations need to avoid this by using an innovative maintenance methodology. There are also available machine learning algorithms that can be utilized in the statistical checks to ascertain the primary cause of the problems that were not envisioned due to larger datasets that are available to the companies. The integration of Artificial Intelligence, Machine Learning, and Data Science has emerged as a transformative approach to preventive mechanical maintenance, offering profound enhancements in the reliability and operational efficiency of industrial machinery. The author's aim is to predict the failure of mechanical components using Artificial Intelligence, Machine Learning, and Data Science in Mechanical Maintenance, with a particular focus on milling machine components, including the boom roller, copping roller, guide roller, and support material device. We used different supervised machine learning algorithms like Linear Regression, Gradient Boosting, Random Forest, Decision Tree, K-nearest neighbors, and Support Vector Machine. The findings reveal that the Support Vector Machine model delivers the highest accuracy than other algorithms, predicting failures with precision rates of 75% for the boom roller, 63.64% for the copping roller, 53.85% for the guide roller, and an impressive 69.23% for the Support Material Device. Additionally, the Mean Absolute Error analysis for the Support Vector Machine model indicates minimal prediction errors of 1 to 4 days. This research highlights the tangible benefits of implementing Artificial Intelligence-driven predictive maintenance in industrial settings, including cost savings, improved machinery performance, and enhanced safety standards.

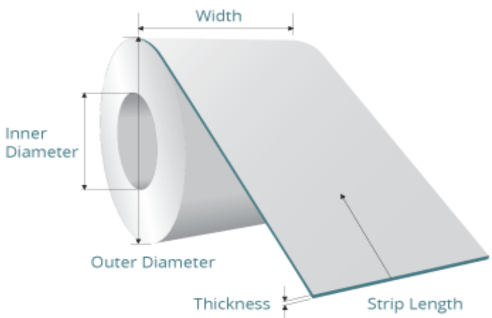
Keywords: Artificial intelligence, Data science, Machine learning, Predictive maintenance, Python, Data analysis, Data visualization, Mean absolute error.

INTRODUCTION

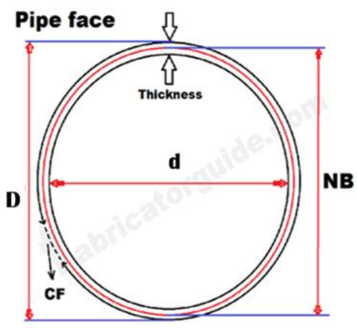
Mechanical components inevitably degrade over time due to wear, tear, and operational stresses, leading to costly downtimes and potential safety hazards. At the same time, we can base our choices on our prior knowledge and experience. To upgrade and make our time more successful we have to look deeply as to what, when, and how to do it and then identify the data to use. And most importantly, we must act wisely (Nacchia *et al.*, 2021). Preventive maintenance, aimed at proactively identifying and addressing issues before they escalate into failures, has emerged as a crucial strategy for ensuring the smooth operation of machinery. Traditionally, preventive maintenance schedules were based on fixed intervals or accumulated operating hours, often resulting in unnecessary servicing or missed maintenance opportunities. Industry is becoming 'smarter' by introducing local intelligence in equipment in the form of machine learning (Nacchia *et al.,* 2021). With advancements in technology, particularly in the fields of Artificial Intelligence (AI), Machine Learning (ML), and Data Science, A fundamental change has taken place in the approach to mechanical maintenance. By utilizing data mining techniques and machine learning algorithms like J48, predictive models can be developed to analyze historical data, assess risk, and enhance decision-making, aligning with the goals of predictive mechanical maintenance to prevent failures and optimize operations (Sameh *et al.*, 2024).

With the use of machine learning algorithms, AI-driven predictive maintenance examines equipment data to predict when repair is necessary before a breakdown happens. AI can detect any problems early by tracking the functioning of the equipment in real time, enabling operators to do preventative maintenance. This method lowers maintenance expenses, minimizes downtime, and enhances overall machine reliability (Jambol *et al.*, 2024). Instead of relying solely on predetermined schedules or reactive approaches, predictive maintenance leverages real-time data and advanced analytical techniques to forecast equipment failures before they occur. By analyzing large volumes of real-time and historical data, AI models can detect patterns and anomalies in machine behavior, allowing for the early identification of potential issues (Joy *et al.*, 2024). ML models trained on such data can then predict potential failure events, enabling maintenance teams to take pre-emptive action, such as scheduling maintenance activities during planned downtime or replacing worn components before catastrophic failures occur.

We will explore how these technologies are utilized to predict failures of critical components, focusing on spiral pipe mill machine components such as Boom Roller, Guide Roller, SMD (Support Material Device), and Copping Roller. The report will examine the various parameters shown in Fig. 1, considered in predictive maintenance models, including Speed of machine (S), Diameter of pipe (D), Thickness of coil(T) and ratios such as Diameter of pipe/ Thickness of coil (D/T), Thickness of coil/ Diameter of pipe (T/D), Diameter of pipe/Speed of machine (D/S) and Thickness of coil/Speed of machine (T/S). Through a comprehensive analysis of real-time readings and historical failure data, we will highlight the benefits, challenges, and implications of adopting predictive maintenance strategies powered by AI, ML, and Data Science in industrial settings.



(A) Parameters of coil.



(B) Parameters of pipe.

Fig. 1. Parameters of coil.

LITERATURE REVIEW

Mechanical components inevitably degrade over time due to wear, tear, and operational stresses, leading to costly downtimes and potential safety hazards. At the same time, past knowledge and experience aid the decision-making process. We need to study, analyze, and use data systematically, but above all, wisely, to improve ourselves while minimizing efforts (Nacchia *et al.*, 2021). Predictive Maintenance (PdM) is a key strategy that uses real-time data to estimate a machine's remaining usable life (RUL) and diagnose a malfunction. This is especially effective for industrial machinery, where safety is of utmost importance because of the high cost and potential harm to people. Machine learning is a technology that uses data to make accurate predictions. The use of machine learning in PdM has significantly reduced costs and guaranteed the protection of human life (Adryan and Sastra 2021). Therefore, to facilitate complicated decision-making in the manufacturing industry, we can further exploit the great potential of the AI and ML techniques, if the problems are properly formulated based on the understanding of system properties (Huang *et al.*, 2020).

According to reports, supervised learning is the most used prediction strategy. The most often used techniques, comprising 40% of papers, were found to be Random Forest, Support Vector Machine, and Neural Networks. Of these, 67% were associated with Deep Neural Networks, with a prevalence of long short-term memory. However, no robust approach that is, no one ideal performance across a range of case tests—works well for every situation (Nacchia *et al.*, 2021). Pure physical models or hybrid models may still be the best option under various conditions, not machine learning, is the basic implication inferred over here but modern machine learning adds to the repertoire of tools required for PdM. However, this does not imply that machine learning will completely supplant previous strategies. Pure physical models or hybrid models may still be the best option in some circumstances (Theissler *et al.*, 2021).

A. Aim of the current study

The aim of this study is to predict the failure of mechanical components using AI, ML, and Data Science in Mechanical Maintenance. This study investigates the application of these cutting-edge technologies in predicting mechanical failures, with a particular focus on milling machine components, including the boom roller, copping roller, guide roller, and support material device.

B. Objectives of the current study

1. To develop an AI-based system that can automatically predict the failure of components based on various input data.
2. Utilize data analytics to develop optimized maintenance schedules, to identify potential production issues well before lead to machine shut-down, reducing unnecessary maintenance and minimizing downtime.
3. Improve the efficiency of resource utilization by minimizing the frequency of inspections and maintenance activities, while ensuring machinery operates reliably and safely.

METHODOLOGY

The methodology of the project starts with the collection of data from a milling machine, with an emphasis on important operating characteristics including speed, thickness, and diameter. Certain parts of the milling machine, such as the support material device (SMD), copping roller, guiding roller, and boom roller are chosen for in-depth examination. Subsequently, the next step involves the analysis of component data to identify key patterns and trends. After that, data is visualized using programs like Matplotlib to help better understand the correlations between different variables. Next, a comprehensive analysis of machine learning methods is carried out to identify which models like SVM, Random Forest, or Linear Regression best suit the predictive maintenance requirement. Once appropriate machine learning algorithms have been chosen, they are trained and verified using historical data. Finally, a comparison of actual vs. AI-predicted data is performed to evaluate the accuracy of the predictions, ensuring the reliability of the maintenance forecasts.

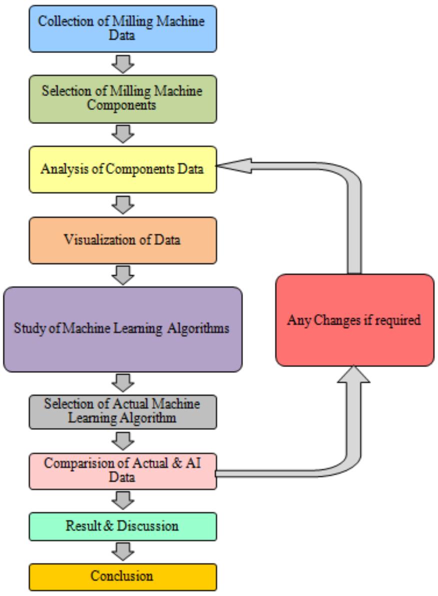


Fig. 2. AI and ML Integration Methodology for Predictive Maintenance.

SELECTION OF COMPONENTS

In the area of predictive maintenance for industrial machinery, such as spiral pipe milling machines, selecting the right components for analysis is critical to effectively forecast and prevent failures (Tessoni and Amoretti 2022). For this project, the components chosen boom roller, copping roller, guide roller, and SMD were identified based on their significance in operational integrity and historical failure patterns.

A. Important of component selection

1. Criticality and Functionality: Each selected component has a crucial function in the overall functioning and performance of the spiral pipe milling machine. Understanding their failure modes and implications is essential for maintaining operational continuity and minimizing downtime.

2. Historical Failure Analysis: Prior analysis of real-time data revealed that the boom roller, copping roller, guide roller, and SMD frequently experienced failures. These failures were documented based on factors such as operational hours, load conditions, and environmental influences.

3. Impact on Production: Failures in these components can significantly disrupt production schedules and lead to costly repairs or replacements. Predicting failure dates allows proactive maintenance interventions, reducing the risk of unplanned downtime and optimizing machine uptime (Jambol *et al.,* 2024).

B. Methodology and data collection

1. Real-Time Data Acquisition: Data collection involved gathering operational metrics such as rotational speed, temperature, vibration levels, and specific performance indicators unique to each component. This real-time data provided a comprehensive view of component health and operational status.

2. Failure Analysis Criteria: Components were selected based on the frequency of failure incidents observed over a specified period. The analysis focused on identifying patterns or anomalies in data that correlated with impending failures, using techniques such as time-series analysis and anomaly detection (Quatrini *et al.*, 2020).

C. Predictive model development

1. Feature Engineering: Relevant features extracted from real-time data included performance metrics specific to each component. This involved calculating metrics such as speed, thickness, and deviations from optimal operating conditions.

2. Machine Learning Algorithms: Utilizing machine learning algorithms, predictive models were trained to forecast failure dates for the selected components (Tessoni and Amoretti 2022). Algorithms such as regression, decision trees, and neural networks were employed to analyze historical data patterns and predict future failure events.

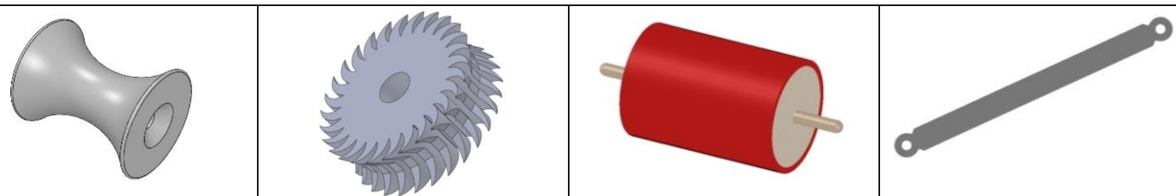
D. Components

1. Boom Roller: Positioned above the forming section, the boom roller plays a crucial role in controlling the diameter and pitch of the spiral, ensuring uniformity and precision in the final product, as shown in Fig. 3 (A).

2. Copping Roller: Positioned at the entry point of the machine, the copping roller ensures that the strip's edges are properly aligned and trimmed before entering the forming section, as shown in Fig. 3 (B).

3. Guide Roller: Positioned strategically along the length of the machine, these rollers play a vital part in maintaining the proper alignment and tension of the strip, ensuring smooth and consistent formation of the spiral pipe, as shown in Fig. 3 ©.

4. SMD: Positioned underneath the strip, the SMD applies pressure and support to prevent distortion or deformation as the strip is shaped into a spiral, ensuring that the pipe maintains its desired shape and structural integrity, as shown in Fig. 3 (D).



(A) Boom roller (B) Copping roller (C) Guide roller (D) Support Material Device

Fig. 3. Components of the Milling Machine.

MACHINE LEARNING

Machine learning is the field that studies computer algorithms to make accurate predictions and responses in specific situations or to act smart. Speaking, machine learning uses past knowledge to figure out how to create better conditions down the road. Machines gain knowledge and skills from existing data. So, machine learning is the development of programs that allow us to pick relevant data from various sources, assess that data, and use it to forecast how the system will act in similar or different scenarios (Subasi, 2020). Machine learning is a branch of artificial intelligence whose goal is to create algorithms that allow computers to understand data and use it to make predictions or choices. In contrast to traditional programming, which requires explicit instructions, machine learning systems learn how to do a task via practice (Surden, 2021). The goal of machine learning, a branch of artificial intelligence, is to create algorithms that let computers understand data and utilize it to forecast or make choices. In contrast to traditional programming, which requires explicit instructions, machine learning systems learn how to do a task via practice. This learning process involves identifying patterns in data and using these patterns to make informed predictions or decisions. ML techniques support the resolution of numerous business issues, including clustering, associations, forecasting, classification, regression, and others (Savitha *et al.*, 2023).

Machine learning has the potential to revolutionize preventive mechanical maintenance by enabling predictive capabilities that were previously unattainable. By leveraging data from machinery and advanced ML algorithms, organizations can achieve significant improvements in operational efficiency, cost savings, and equipment reliability. As ML technology continues to advance, its role in preventive maintenance will likely become even more integral, driving further innovation and benefits across various industries.

The different supervised machine learning algorithms like Linear Regression, Gradient Boosting, Random Forest, Decision Tree, k-nearest neighbors, and Support Vector machines are used to predict the failure date of components and to find which algorithm is best suited according to the problem.

DATA ANALYSIS

Data analysis serves as the foundational step in the ML pipeline, guiding decisions from data pre-processing to model selection, validation, and ongoing refinement. It ensures that ML models are not only accurate but also robust, interpretable, and aligned with real-world data dynamics and requirements. Data analysis guides the process of uncovering hidden insights within the dataset. Exploratory Data Analysis (EDA) makes visible any patterns or issues in the data, which results in a hypothesis (Rahmany *et al.*, 2020). We mainly focus more on real-time data analysis of the milling machine dataset and finding insights. We used Python programming language for data analysis with libraries like NumPy, Pandas, and scikit-learn. Python has become one of the most popular interpreted programming languages, along with Perl, Ruby, and others (Hope, 2020).

Data analysis plays a pivotal role in understanding the performance and failure patterns of milling machine components. This report delves into the application of correlation, R^2 coefficient, and linear regression methods to establish relationships between key parameters and the failure days of components such as the boom roller, copping roller, guide roller, and support material device (SMD). The following parameters were analyzed about the failure days of the components: Speed of machine (S), diameter of pipe (D), thickness of coil (T), diameter/thickness(D/T) ratio, thickness/diameter(T/D) ratio, diameter/speed (D/S) ratio, thickness/speed(T/S) Ratio.

A. Correlation analysis

An examination of correlation was used to measure the strength and direction of the linear relationships between each parameter and the failure days of the components. Correlation coefficients close to +1 indicate a strong positive relationship, while coefficients close to -1 signify a strong negative relationship. A coefficient near 0 suggests no linear relationship.

To find the correlation of two variables, In Python, we used the most commonly used function to find the correlation between two variables corr() from the Pandas library and the excel correl function to validate the result.

Table 1: Correlation of the parameters and day of component failure with the help of python's corr() from the Pandas.

Parameters	Boom Roller	Copping Roller	Guide Roller	SMD
Speed	-0.47	0	-0.4	-0.6
Diameter	0.64	0.43	0.39	0.71
Thickness	0.18	0.15	0.32	0.45
D/T	**0.84**	0.57	0.33	0.34
T/D	-0.84	**-0.64**	**-0.44**	-0.44
D/S	0.67	0.43	0.4	**0.73**
T/S	0.28	0.15	0.33	0.51

Table 2: Correlation analysis of the parameters and day of component failure with the help of excel's correl function.

Parameters	Boom Roller	Copping Roller	Guide Roller	SMD
Speed	-0.47	0	-0.4	-0.6
Diameter	0.64	0.43	0.39	0.71
Thickness	0.18	0.15	0.32	0.45
D/T	**0.84**	0.57	0.33	0.34
T/D	-0.84	**-0.64**	**-0.44**	-0.44
D/S	0.67	0.43	0.4	**0.73**
T/S	0.28	0.15	0.33	0.51

B. Linear regression

Linear regression models were fitted to quantify the linear relationship between each parameter and the failure days. These models provide insights into the rate of change in failure days concerning changes in the parameters, offering a predictive framework for understanding component reliability.

To find the linear regression of two variables, In Python, we used the Linear Regression class from sklearn.linear_model module in the Scikit-Learn library for more advanced linear regression analysis.

Table 3: Linear regression analysis of the parameters and day of component failure with the help of python's sklearn.linear_model.

Parameters	Boom Roller	Copping Roller	Guide Roller	SMD
Speed	-75.07	0	-318.56	-142.1
Diameter	0.12	0.07	0.07	0.14
Thickness	5.04	4.15	7.76	9.24
D/T	2.18	1.64	1.31	0.73
T/D	-20942.97	-17721.56	-16434.18	-11395.08
D/S	0.57	0.35	0.36	0.68
T/S	38.23	20.75	40.51	51.82

C. Coefficient of determination

The R^2 coefficient or coefficient of determination, was calculated to assess how well the variation in each parameter explains the variability in the failure days of the components. A higher R^2 value (closer to 1) indicates that the parameter is more predictive of the component's failure days. The coefficient of determination (R-squared) is more informative and truthful, therefore suggests the usage of R-squared as a standard metric to evaluate regression analyses in any scientific domain (Chicco *et al.*, 2021).

Table 4: Coefficient of determination analysis of the parameters and day of component failure with the help of python.

Parameters	Boom Roller	Copping Roller	Guide Roller	SMD
Speed	0.22	0	0.16	0.36
Diameter	0.41	0.18	0.15	0.5
Thickness	0.03	0.02	0.1	0.2
D/T	**0.7**	0.33	0.11	0.11
T/D	0.7	**0.41**	**0.19**	0.2
D/S	0.45	0.18	0.16	**0.54**
T/S	0.08	0.02	0.11	0.26

DATA VISUALIZATION

Data visualization tools have become an essential part of the data analysis process (Lavanya *et al.*, 2023). Data visualization plays a pivotal role in understanding complex datasets by visually representing relationships, trends, and patterns. An important component of the scientific method is data visualization. A scientist will be able to explain their findings to others and understand their own data with the help of effective visuals (Waskom, 2021). Understanding the relationships between key parameters and the failure times of machine components is crucial for predictive maintenance and optimizing operational efficiency. This study utilizes Matplotlib, a Python library for data visualization, to explore and analyze the relationships between various parameters of milling machine components and their failure times. The components studied include the boom roller, copping roller, guide roller, and support material device (SMD). The parameters explored are Diameter/Thickness (D/T), Thickness/Diameter (T/D), and Diameter/Speed (D/S) ratios with respect to the number of days until component failure. By plotting these relationships using scatter plots, we aim to gain insights into the performance characteristics and failure patterns of these components.

Data visualization graphs

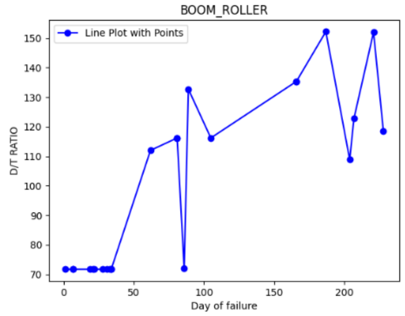
The components and parameters analyzed in this study shown in Figure include:

1. Boom Roller: Diameter/Thickness (D/T) ratio vs. Day of failure, shown in Fig. 4 (A).

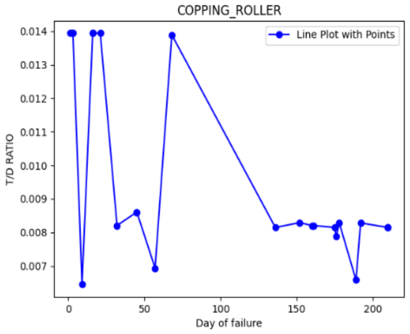
2. Copping Roller: Thickness/Diameter (T/D) ratio vs. Day of failure, shown in Fig. 4 (B).

3. Guide Roller: Thickness/Diameter (T/D) ratio vs. Day of failure, shown in Fig. 4 (C).

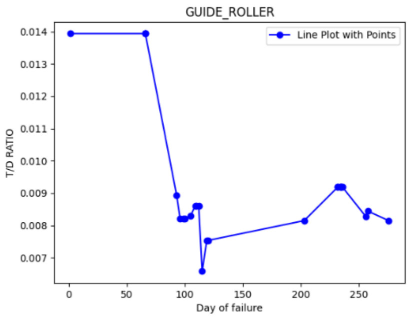
4. SMD: Diameter/Speed (D/S) ratio vs. Day of failure, shown in Fig. 4 (D)



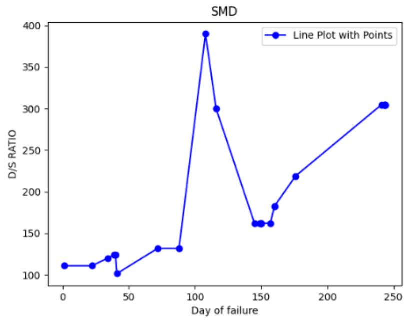
A. Boom roller- D/T ratio vs. Day of failure.



B. Copping roller- T/D ratio vs. Day of failure.



C. Guide roller- T/D ratio vs. Day of failure.



D. SMD- D/S ratio vs. Day of failure.

Fig. 4. Data visualization graphs.

RESULTS

In the project, focused on the various machine learning algorithms Gradient Boosting, Linear Regression, Decision Tree, Random Forest, KNN, SVM were applied to predict the maintenance needs of milling machine components: boom roller, copping roller, guide roller, and SMD.

SVM achieved the highest accuracy rates of 75% for the boom roller, 63.64% for the copping roller, 53.85% for the guide roller, and an impressive 69.23% for SMD.

Error Analysis: MAE calculations revealed the predictive capabilities in terms of days shown in Table 10:

Boom roller: ±0.67 days
Copping roller: ±1.36 days
Guide roller: ±1.38 days
SMD: ±3.65 days

These results highlight SVM's ability to provide reliable predictions. Among the algorithms tested, SVM emerged as the most effective, demonstrating superior performance across the components.

In Tables 5-8, the actual failure date of the components is shown in the 1st column & other columns show the predicted failure date of the component by the respective algorithm. Di shows the difference in the predicted failure date by algorithms & actual failure date of the component.

Where the negative (-) sign of difference (Di) indicates that the algorithm predicts the failure before the actual failure. A positive (+) sign of difference (Di) indicates that the algorithm predicts the failure after the actual failure. Zero (0) in difference (Di) indicates the exact prediction of failure date.

Table 5: Actual date and predicted date by algorithms for boom roller.

Actual Date	Linear Regression		Gradient Boosting		Random Forest		Decision Tree		K - Nearest Neighbors		Support Vector Machine	
	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di
07-08-2023	19-08-2023	12	23-08-2023	16	21-08-2023	14	24-08-2023	17	19-08-2023	12	16-08-2023	9
21-08-2023	22-08-2023	1	27-08-2023	6	26-08-2023	5	28-08-2023	7	23-08-2023	2	19-08-2023	-2
28-08-2023	10-09-2023	13	11-09-2023	14	10-09-2023	13	12-09-2023	15	06-09-2023	9	03-09-2023	6
11-09-2023	13-09-2023	2	19-09-2023	8	18-09-2023	7	20-09-2023	9	14-09-2023	3	11-09-2023	0
26-09-2023	28-09-2023	2	03-10-2023	7	03-10-2023	7	03-10-2023	7	01-10-2023	5	26-09-2023	0
12-10-2023	14-10-2023	2	19-10-2023	7	19-10-2023	7	19-10-2023	7	18-10-2023	6	11-10-2023	-1
28-10-2023	06-11-2023	9	06-11-2023	9	04-11-2023	7	07-11-2023	10	29-10-2023	1	28-10-2023	0
08-11-2023	12-11-2023	4	23-11-2023	15	19-11-2023	11	24-11-2023	16	05-11-2023	-3	07-11-2023	-1
20-11-2023	19-11-2023	-1	05-12-2023	15	01-12-2023	11	02-12-2023	12	27-11-2023	7	20-11-2023	0
09-12-2023	09-12-2023	0	14-12-2023	5	14-12-2023	5	14-12-2023	5	14-12-2023	5	06-12-2023	-3
26-12-2023	29-12-2023	3	07-01-2024	12	03-01-2024	8	07-01-2024	12	03-01-2024	8	26-12-2023	0
14-01-2024	15-01-2024	1	21-01-2024	7	20-01-2024	6	19-01-2024	5	19-01-2024	5	14-01-2024	0

Table 6: Actual date and predicted date by algorithms for copping roller.

Actual Date	Linear Regression		Gradient Boosting		Random Forest		Decision Tree		K - Nearest Neighbors		Support Vector Machine	
	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di
26-08-2023	03-09-2023	8	10-09-2023	15	06-09-2023	11	11-09-2023	16	04-09-2023	9	02-09-2023	7
13-09-2023	28-08-2023	-16	31-08-2023	-13	30-08-2023	-14	29-08-2023	-15	28-08-2023	-16	05-09-2023	-8
13-09-2023	26-09-2023	13	29-09-2023	16	29-09-2023	16	30-09-2023	17	25-09-2023	12	25-09-2023	12
26-09-2023	26-09-2023	0	30-09-2023	4	30-09-2023	4	30-09-2023	4	27-09-2023	1	25-09-2023	-1
07-10-2023	10-10-2023	3	14-10-2023	7	14-10-2023	7	14-10-2023	7	14-10-2023	7	09-10-2023	2
20-10-2023	24-10-2023	4	30-10-2023	10	25-10-2023	5	31-10-2023	11	19-10-2023	-1	20-10-2023	0
02-11-2023	04-11-2023	2	13-11-2023	11	11-11-2023	9	13-11-2023	11	04-11-2023	2	02-11-2023	0
15-11-2023	12-11-2023	-3	27-11-2023	12	21-11-2023	6	28-11-2023	13	20-11-2023	5	14-11-2023	-1
27-11-2023	02-12-2023	5	29-11-2023	2	30-11-2023	3	29-11-2023	2	03-12-2023	6	29-11-2023	2
10-12-2023	29-12-2023	19	02-12-2023	-8	02-12-2023	-8	30-11-2023	-10	29-11-2023	-11	09-12-2023	-1
22-12-2023	29-12-2023	7	07-01-2024	16	03-01-2024	12	08-01-2024	17	30-12-2023	8	25-12-2023	3

Table 7: Actual date and predicted date by algorithms for guide roller.

Actual Date	Linear Regression		Gradient Boosting		Random Forest		Decision Tree		K - Nearest Neighbors		Support Vector Machine	
	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di
05-08-2023	21-08-2023	**16**	22-08-2023	**17**	22-08-2023	**17**	22-08-2023	**17**	22-08-2023	**17**	14-08-2023	**9**
25-08-2023	22-08-2023	**-3**	23-08-2023	**-2**	24-08-2023	**-1**	24-08-2023	**-1**	23-08-2023	**-2**	15-08-2023	**-10**
27-08-2023	07-09-2023	**11**	16-09-2023	**20**	13-09-2023	**17**	17-09-2023	**21**	08-09-2023	**12**	05-09-2023	**9**
10-09-2023	08-09-2023	**-2**	19-09-2023	**9**	16-09-2023	**6**	19-09-2023	**9**	14-09-2023	**4**	09-09-2023	**-1**
22-09-2023	22-09-2023	**0**	04-10-2023	**12**	01-10-2023	**9**	05-10-2023	**13**	26-09-2023	**4**	22-09-2023	**0**
30-09-2023	10-10-2023	**10**	11-10-2023	**11**	11-10-2023	**11**	12-10-2023	**12**	10-10-2023	**10**	05-10-2023	**5**
10-10-2023	13-10-2023	**3**	20-10-2023	**10**	21-10-2023	**11**	20-10-2023	**10**	20-10-2023	**10**	14-10-2023	**4**
25-10-2023	26-10-2023	**1**	06-11-2023	**12**	02-11-2023	**8**	07-11-2023	**13**	24-10-2023	**-1**	25-10-2023	**0**
05-11-2023	12-11-2023	**7**	22-11-2023	**17**	14-11-2023	**9**	14-11-2023	**9**	01-11-2023	**-4**	05-11-2023	**0**
20-11-2023	18-11-2023	**-2**	04-12-2023	**14**	28-11-2023	**8**	05-12-2023	**15**	24-11-2023	**4**	18-11-2023	**-2**
05-12-2023	06-12-2023	**1**	14-12-2023	**9**	13-12-2023	**8**	15-12-2023	**10**	13-12-2023	**8**	05-12-2023	**0**
20-12-2023	22-12-2023	**2**	05-01-2024	**16**	01-01-2024	**12**	31-12-2023	**11**	25-12-2023	**5**	20-12-2023	**0**
01-01-2024	07-01-2024	**6**	17-01-2024	**16**	15-01-2024	**14**	15-01-2024	**14**	15-01-2024	**14**	05-01-2024	**4**

Table 8: Actual date and predicted date by algorithms for SMD.

Actual Date	Linear Regression		Gradient Boosting		Random Forest		Decision Tree		K - Nearest Neighbors		Support Vector Machine	
	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di	Predicted Date	Di
04-08-2023	19-08-2023	**15**	23-08-2023	**19**	23-08-2023	**19**	24-08-2023	**20**	19-08-2023	**15**	17-08-2023	**13**
05-08-2023	20-08-2023	**15**	24-08-2023	**19**	24-08-2023	**19**	24-08-2023	**19**	22-08-2023	**17**	18-08-2023	**13**
05-08-2023	21-08-2023	**16**	25-08-2023	**20**	26-08-2023	**21**	26-08-2023	**21**	26-08-2023	**21**	21-08-2023	**16**
21-08-2023	22-08-2023	**1**	26-08-2023	**5**	26-08-2023	**5**	26-08-2023	**5**	26-08-2023	**5**	22-08-2023	**1**
04-09-2023	12-09-2023	**8**	14-09-2023	**10**	11-09-2023	**7**	15-09-2023	**11**	07-09-2023	**3**	03-09-2023	**-1**
22-09-2023	02-10-2023	**10**	30-09-2023	**8**	26-09-2023	**4**	30-09-2023	**8**	22-09-2023	**0**	22-09-2023	**0**
07-10-2023	10-10-2023	**3**	14-10-2023	**7**	14-10-2023	**7**	14-10-2023	**7**	13-10-2023	**6**	11-10-2023	**4**
26-10-2023	30-10-2023	**4**	03-11-2023	**8**	01-11-2023	**6**	04-11-2023	**9**	29-10-2023	**3**	26-10-2023	**0**
06-11-2023	05-11-2023	**-1**	23-11-2023	**17**	15-11-2023	**9**	24-11-2023	**18**	02-11-2023	**-4**	06-11-2023	**0**
20-11-2023	20-11-2023	**0**	05-12-2023	**15**	30-11-2023	**10**	06-12-2023	**16**	25-11-2023	**5**	20-11-2023	**0**
10-12-2023	13-12-2023	**3**	20-12-2023	**10**	16-12-2023	**6**	21-12-2023	**11**	10-12-2023	**0**	09-12-2023	**-1**
30-12-2023	30-12-2023	**0**	10-01-2024	**11**	06-01-2024	**7**	07-01-2024	**8**	30-12-2023	**0**	30-12-2023	**0**
17-01-2024	20-01-2024	**3**	31-01-2024	**14**	28-01-2024	**11**	01-02-2024	**15**	21-01-2024	**4**	19-01-2024	**2**

Table 9: Predicted date accuracy by algorithms for Components (in %).

Components	Linear Regression	Gradient Boosting	Random Forest	Decision Tree	K - Nearest Neighbors	Support Vector Machine
Boom Roller	58.33	0.00	0.00	0.00	16.67	75
Copping Roller	18.18	9.09	0.00	9.09	27.27	63.64
Guide Roller	46.15	7.69	7.69	7.69	15.38	53.85
SMD	30.77	0.00	0.00	0.00	23.08	69.23

MEAN ABSOLUTE ERROR (MAE)

The MAE measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures the accuracy of continuous variables.

$$\text{MAE}=\frac{1}{m}\sum_{i=1}^{m}|X_i - Y_i| \quad (1)$$

Among the algorithms tested, SVM emerged as the most effective, demonstrating superior performance across the components.

Table 10 shows the Mean Absolute Error between the actual date and the date that is predicted by the algorithm.

Table 10: Mean Absolute Error shown by each algorithm for components (in days).

Components	Linear Regression	Gradient Boosting	Random Forest	Decision Tree	K- Nearest Neighbors	Support Vector Machine
Boom Roller	4.00	10.08	8.42	10.17	5.00	0.67
Copping Roller	3.82	6.55	4.64	6.64	2.00	1.36
Guide Roller	3.85	12.38	9.92	11.77	6.23	1.38
SMD	5.92	12.54	10.08	12.92	5.77	3.62

A graphical representation shown in Figure 5 of MAE by algorithm shows the predictive performance for each component. The vertical axis represents MAE, and the horizontal axis lists components. Each algorithm is represented by distinct bars. For example, SVM shows the lowest MAE for SMD (4 days) and consistent accuracy for the boom roller (1 day). This visualization helps identify the best algorithm for each component and areas for improvement.

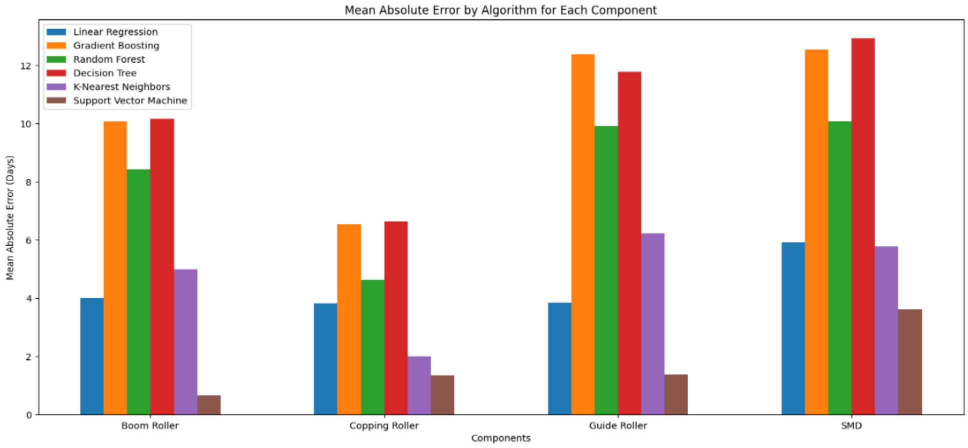


Fig. 5. Graphical Analysis of MAE by Algorithm for Component Performance Evaluation.

DISCUSSION

Predictive mechanical maintenance has been revolutionized by the integration of AI, ML, Data Science, and Python, offering significant improvements in reliability and efficiency. AI and ML algorithms, such as Linear Regression, Gradient Boosting, Random Forest, Decision Trees, K-nearest neighbors, and Support Vector Machines analyze extensive real-time data collected from historical records to identify patterns indicative of potential failures. Python, with its rich ecosystem of libraries like Scikit-learn, NumPy, Pandas, and MatPlotlib, provides powerful tools for implementing these algorithms. These technologies allow for the precise prediction of equipment failures, enabling maintenance teams to address issues proactively rather than reactively. In the context of milling machines, AI and ML models can process data related to the diameter of pipes, the thickness of plates, and machine speed to forecast the failure of components like boom rollers, copping rollers, guide rollers, and support material devices (SMD).

The accuracy shown by SVM, Linear regression, KNN, Gradient Boosting, Random Forest, and Decision tree for boom roller is 75%, 58.33%, 16.67%, 0.00%, 0.00%, and 0.00 respectively shown in Table 9. For copping roller accuracy shown by SVM, KNN, Linear regression, Decision Tree, Gradient boosting, and Random forest is 63.64%, 27.27%, 18.18%, 9.09%, 9.09% and 0.00% respectively shown in Table 9. While for Guide roller ac-curacy shown by SVM, Linear regression, KNN, Gradient boosting, Decision tree and Random forest is 53.85%, 46.15%, 15.38%, 7.69%, 7.69%, and 7.69% respectively shown in Table 9. For SMD, SVM, Linear regression, KNN, Gradient boosting, Random forest, and Decision tree accuracy is 69.23%, 30.77%, 23.08%, 0.00%, 0.00%, and 0.00% respectively shown in Table 9.

SVM achieved the highest accuracy rates of 75% for the boom roller, 63.64% for the copping roller, 53.85% for the guide roller, and an impressive 69.23% for SMD.

CONCLUSION AND FUTURE SCOPE

PdM stands as one of the factor strategies that rely on real-time data to predict machine failures by estimating the RUL. This approach is important for industrial machines where safety takes priority due to the huge costs and potential risk to human life safety (Adryan and Sastra 2021). In conclusion, the project underscores the critical role of AI, ML, and Data Science in modernizing predictive maintenance strategies. By adopting these technologies, industries can move towards proactive maintenance practices that not only enhance reliability and safety but also drive significant cost efficiencies. As technology continues to evolve, integrating advanced analytics and predictive capabilities will remain pivotal in shaping the future of mechanical maintenance.

After analyzing components' real-time data with the help of Python and its libraries like NumPy, Pandas, scikit-learn, Matplotlib, sklearn. Linear_model and using several machine learning algorithms like Linear Regression, Gradient Boosting, Random Forest, Decision Tree, K- Nearest Neighbors, Support Vector Machine. We can conclude that the SVM is the best-suited algorithm to predict the component failure date of the milling machine. SVM achieved the highest accuracy rates among all other algorithms that are tested, which is 75% for boom roller, 63.64% for copping roller, 53.85% for guide roller, and an impressive 69.23% for SMD. MAE shows the average difference between the predicted date by algorithms and the actual failure date of the component. The MAE shown by SVM is ±0.67 days for the boom roller, ±1.36 days for the copping roller, ±1.38 days for the guide roller, and ±3.65 days for SMD.

When analyzing parameters for predictive maintenance, common factors such as temperature, speed, and vibration are often prioritized. These are critical indicators of a machine's performance and can provide valuable insights into the likelihood of component failure. However, it is important to recognize that other factors, like lubrication, also play a crucial role in forecasting the failure of components. Proper lubrication reduces friction, minimizes wear, and ensures the smooth operation of mechanical parts. A lack of adequate lubrication can lead to overheating, increased wear, and eventual breakdown, making it a significant factor to consider in predictive maintenance strategies. While gathering real-time data, machines sometimes undergo modifications to increase their capacity and production output. During these periods of modification, the machine may not operate as expected, which complicates the process of analyzing data and predicting component failures. The altered performance metrics during modifications can lead to inaccurate data, making it challenging to assess the true condition of the machine and its components. Consequently, predictive maintenance efforts may be hindered, as the irregular data does not reflect the machine's typical operating conditions.

Future work in the field of predictive mechanical maintenance will likely focus on further enhancing the accuracy and efficiency of AI and ML models. One of the major challenges in creating a predictive maintenance system is the lack of failure data, as the machine/ components are frequently repaired before they fail (Van Dinter *et al.*, 2022). As data collection becomes more sophisticated, incorporating real-time data from IoT devices and advanced sensors will provide richer datasets for analysis. This will enable the development of more precise predictive models.

ABBREVIATIONS

PdM	Predictive Maintenance
AI	Artificial Intelligence
ML	Machine Learning
RUL	Remaining Useful Life
SVM	Support Vector Machines
KNN	K-Nearest Neighbors
SMD	Support Material Device
MAE	Mean Absolute Error
EDA	Exploratory Data Analysis
IOT	Internet of Things
D	Outer diameter of pipe (mm)
T	Thickness of coil (mm)
S	Speed of machine (m/min)
Di	Difference in the predicted failure date by algorithms & actual failure date of the component.

Data Availability Statement. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality agreements with the company.

Conflicts of Interest. None.

REFERENCES

Adryan, F. A. & Sastra, K. W. (2021). Predictive maintenance for aircraft engine using machine learning: Trends and challenges. *Avia, 3*(1).

Chicco, D., Warrens, M. J. & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science, 7,* e623.

Hope, T. M. (2020). Linear regression. In Machine learning (pp. 67-81). Academic Press.

Huang, J., Chang, Q. & Arinez, J. (2020). Deep reinforcement learning based preventive maintenance policy for serial production lines. *Expert Systems with Applications, 160,* 113701.

Jambol, D. D., Sofoluwe, O. O., Ukato, A. & Ochulor, O. J. (2024). Transforming equipment management in oil and gas with AI-Driven predictive maintenance. *Computer Science & IT Research Journal, 5*(5), 1090-1112.

Joy, B., Muhammed, A., John, A. & Balil, M. (2024). AI-Powered Predictive Maintenance: Optimizing Industrial Efficiency.

Lavanya, A., Sindhuja, S., Gaurav, L. & Ali, W. (2023). A comprehensive review of data visualization tools: features, strengths, and weaknesses. *Int. J. Comput. Eng. Res. Trends, 10*(01), 10-20.

Nacchia, M., Fruggiero, F., Lambiase, A. & Bruton, K. (2021). A systematic mapping of the advancing use of machine learning techniques for predictive maintenance in the manufacturing sector. *Applied Sciences, 11*(6), 2546.

Quatrini, E., Costantino, F., Di Gravio, G. & Patriarca, R. (2020). Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. *Journal of Manufacturing Systems, 56*, 117-132.

Rahmany, M., Zin, A. M. & Sundararajan, E. A. (2020). Comparing tools provided by python and r for exploratory data analysis. *International Journal of Information System and Computer Science, 4*(3), 131-142.

Savitha, J., Sneha, D. G., Aiswin, S. & Surendar, R. S. (2023). History of Machine Learning with Its Advantages ad Its Applications. *IJIRT.*

Sameh Ali, Atef Raslan and Lamiaa Fattouh (2024). Model for the Prediction of Default Risk of Funding Requests Using Data Mining. *International Journal on Emerging Technologies, 15*(2), 05–12.

Subasi, A. (2020). Practical machine learning for data analysis using python. Academic Press.

Surden, H. (2021). Machine learning and law: An overview. Research Handbook on Big Data Law, 171-184.

Tessoni, V. & Amoretti, M. (2022). Advanced statistical and machine learning methods for multi-step multivariate time series forecasting in predictive maintenance. *Procedia Computer Science, 200*, 748-757.

Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M. & Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability engineering & system safety, 215,* 107864.

van Dinter, R., Tekinerdogan, B. & Catal, C. (2022). Predictive maintenance using digital twins: A systematic literature review. *Information and Software Technology, 151*, 107008.

Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software, 6*(60), 3021.

How to cite this article: Suraj Shashikant Patil and Gajanan N. Thokal (2025). Analysis of Predictive Mechanical Maintenance Using Artificial Intelligence, Machine Learning and Data Science. *International Journal on Emerging Technologies, 16*(1): 45–53.