Predictive Maintanance: Machine Downtime Prediction Using Data Analytics & Machine Learning

Prof. Shailendra Shende[1], Assistant Professor, Department of Computer Science and Engineering, Government College of Engineering Chandrapur, Maharashtra, India

Janhvi Mohurle[2], Aliya Kazi[3], Viniya Jagtap[4], Prachi Mesharam[5], Arpita Gajghate[6] Final year, Department of computer science and engineering, Government college of engineering, Chandrapur, Maharashtra, India.

Abstract: Machine downtime significantly impacts production efficiency and operational costs in various industries. Predictive maintenance leverages data analytics and machine learning to forecast machine failures before they occur, reducing unexpected downtime and maintenance costs. This paper explores the implementation of machine downtime prediction models using historical machine data, sensor readings, and AI-based predictive techniques. The study compares various machine learning algorithms, including Regression models, Decision trees, Support vector machine and Random Forest Classifier to determine their efficacy in predicting machine failures.

Key Words: Predictive Maintenance, Preventive Maintenance, Neural Networks, downtime costs, Maintenance cost, Machine learning, Internet of things, Random Forest Algorithm.

1. INTRODUCTION

In modern manufacturing and industrial operations, machine downtime presents a significant challenge, leading to productivity losses, increased maintenance costs, and potential disruptions in supply chains. Unplanned downtime, in particular, can result in severe financial consequences, making predictive maintenance a crucial area of study. Machine downtime prediction aims to minimize unexpected failures by leveraging data-driven approaches, including machine learning, statistical modeling, and real-time monitoring.

This research explores various methodologies for predicting machine downtime of CNC milling machine used for cutting. The main CNC machine parameters include cutting speed, spindle speed, feed rate, depth of cut, and others. By analyzing historical failure patterns and operational data, predictive models can forecast potential breakdowns, allowing for timely intervention and optimized maintenance schedules. This proactive approach not only reduces downtime but also extends machine lifespan and improves overall operational efficiency.

2. THEOROTICAL FRAMEWORK

Machine downtime prediction is grounded in several key theoretical concepts, including predictive maintenance, machine learning, reliability engineering, and data-driven decision-making. This framework provides a foundation for understanding how predictive analytics can be applied to minimize unplanned machine failures and improve operational efficiency.

Predictive maintenance is a data-driven maintenance strategy that utilizes condition-monitoring techniques to anticipate failures before they occur. PdM relies on theories from reliability-centered maintenance (RCM) and failure mode and effects analysis (FMEA) to identify critical failure points and develop intervention strategies. The core principle is that machine degradation follows identifiable patterns that can be analyzed and predicted.

Modern downtime prediction leverages machine learning algorithms and statistical models to analyze historical failure data. Theoretical foundations include:

- **Supervised Learning Models:** Classification and regression algorithms (e.g., Decision Trees, Support Vector Machines) that predict failure probabilities based on historical inputs.
- **Unsupervised Learning Models:** Clustering techniques (e.g., K-Means) for anomaly detection in sensor readings and operational patterns.

This theoretical framework integrates predictive maintenance theories, machine learning techniques, reliability engineering, and IoT-driven monitoring to establish a comprehensive foundation for machine downtime prediction. By leveraging these theories, industries can develop more accurate predictive models, reduce operational disruptions, and enhance equipment longevity.

Previous research on predictive maintenance has focused on time-series forecasting, anomaly detection, and classification models. Techniques such as logistic regression, support vector machines (SVM), random forests, and deep learning have been employed to predict downtime events. The integration of the Internet of Things (IoT) and cloud computing has further enhanced the capabilities of PdM systems.

3. METHODOLOGY

This study employs a structured approach to downtime prediction.

SOFTWARE REQUIREMENTS

BACKEND-

1. Install virtual environment in VScode.
2. Database server : MySQL.
3. Language : Python
4. Import libraries like Num py , Pandas , Scikit learn , Matplotlib,...etc.

FRONTEND-

1. For website : basic knowledge of HTML , JS.
2. Framework: streamlit python library .

3.1 Data Collection: Sensor data, machine logs, and maintenance records are gathered.

Data sources- The dataset for downtime prediction is obtained from:

- **Machine logs and event records** (error codes, failure timestamps)
- **Maintenance records** (scheduled and unscheduled maintenance history)
- **Production data** (machine workload, operational hours, speed)

Data acquisition methods-

- **Historical failure logs:** Extracting downtime records from manufacturing databases.

Operator and maintenance reports: Collecting qualitative insights on machine conditions.

Model is trained on CNC milling machine online dataset from Kaggle.

3.2 Data Preprocessing: Data preprocessing is a critical step in the data science and machine learning pipeline. It involves transforming raw data into a clean and usable format to improve the performance and accuracy of models. Handling missing values, feature selection, and normalization.

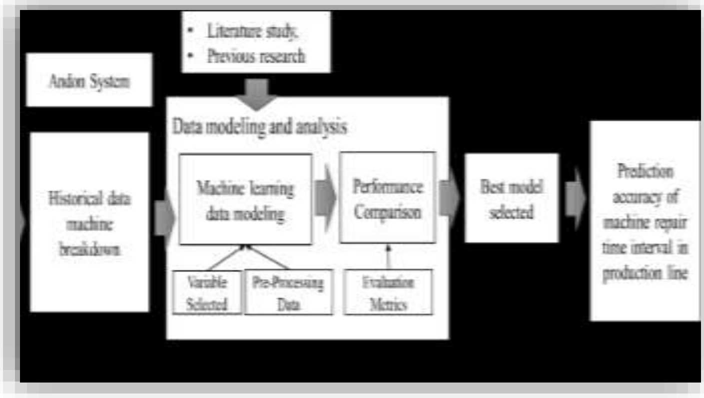

Fig. Research Framework

Data Cleaning

- Handling missing values through winsorization (e.g., mean/median replacement or interpolation).
- Removing redundant or inconsistent records to improve data quality.

Feature Engineering

- Extracting relevant features (e.g., temperature fluctuations, vibration anomalies).
- Creating lag variables for time-series modeling.
- Encoding categorical variables (e.g., machine type, failure category) into numerical values.

Data Normalization and Transformation

- Scaling sensor readings using Min-Max scaling or Standardization (Z-score).
- Converting timestamps into cyclical features (e.g., time-of-day, day-of-week).
-

3.3 Model Selection: Comparing machine learning algorithms such as Random Forest, Regression, Decision tree, and Support vector machine
Several predictive models are evaluated for machine downtime prediction:

- DECISION TREE
- SUPPORT VECTOR MACHINE
- RANDOM FOREST CLASSIFIER
- LINEAR REGRESSION
- KNN

Therefore, **RANDOM FOREST CLASSIFIER** algorithm is accurate for this model.

3.4 Model training and testing: Training model on gathered data. Testing if it meets all requirements.

TRAINING PROCESS-

- **Splitting the dataset**: 70% training, 20% validation, 10% testing.
- **Hyper parameter tuning**: Using **Grid Search** or **Bayesian Optimization** to optimize model performance

EVALUATION METRICS-

To measure the effectiveness of the predictive models, the following metrics are used:

- **Accuracy**: Measures overall correctness of predictions.
- **Precision & Recall**: Evaluates the model's ability to correctly predict failures.
- **F1 Score**: Balances precision and recall for an optimal assessment.

3.5 Deployment: Deploying it as a website. Creating a prediction website for users to make it easy for machine downtime prediction.
Deployed the model in Streamlit app.

```
print(f"\n❌ Error during prediction: {e}")
raise
```

4. EXPERIMENTAL RESULTS

The models are tested on real-world industrial datasets. The performance of different models is analyzed, highlighting the best-performing techniques. Feature importance analysis reveals critical parameters influencing downtime prediction.

Voltage(volts)	Torque(Nm)	Cutting(kN)	Prediction	Failure_Probability
335	24.0553	3.58	Machine_Failure	0
368	14.2029	2.68	Machine_Failure	0
325	24.0493	3.55	Machine_Failure	0.01
360	25.86	3.55	Machine_Failure	0
354	25.5159	3.55	Machine_Failure	0
319	25.5213	3.55	Machine_Failure	0.01
438	25.4547	3.58	Machine_Failure	0
334	34.973	2.02	No_Machine_Failure	1
278	32.5193	2.88	No_Machine_Failure	1
379	25.6186	3.93	Machine_Failure	0.04
396	25.4614	3.67	Machine_Failure	0.02
366	25.6131	3.59	Machine_Failure	0.01

Above image indicates the result given by prediction model. It calculates the probability of machine failure.

5. CONCLUSION

This research article present a review of various methods of machine learning algorithm applies in various parts of manufacturing industry. Finally constructed RFC and tested successfully with various metric measurements. The discussion and literature review also was performed. The determining machine failure is a primary goal of predictive maintenance. All the sensors data sent and processing by learning algorithm for predict the machine failure. The given data is preprocessed because of the real time data.

These data is collected as a raw data from a CNC milling machine industry . So the data are not clean and ordered. The cleaning data carried out for further process which provides higher reliability and accuracy of prediction. Also this data are splitting into two for training and testing. With the help of these dataset, there can create predictive maintenance model. The output from the constructed model will be analyzed and evaluated. The accuracy and precision will be fewer in some period of predictive time. The deep logical analysis of the key demanding situations is challenging task as our future enhancement of our research article.

Also the various segments in manufacturing industry and minimum predictive maintenance data size will provide less accuracy and precise for predictive maintenance.

6. FUTURE SCOPE

Also we recommend do furthering improvements and motivating authors and researchers for the following aspects;

1. The optimum preprocessing technique should be used for raw data analysis to get better accuracy.

2. Combine two or more ML techniques to design the model to achieve better prediction for data acquisition.

3. Machine Learning model approach can be extended for further studied.

4. Classification and anomaly detection algorithms will be combined to keep up exactness of the classification models while not losing anomaly detection benefits. By this manner, PdM can be applied to instrumentality or system that doesn’t have giant data.

7. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who contributed to the successful completion of this research on **Machine Downtime Prediction**.

Firstly, I extend my deepest appreciation to my **supervisor/guide**, **Prof. Shailendra Shende**, for their invaluable guidance, continuous support, and constructive feedback throughout this research journey. Their expertise and encouragement have been instrumental in shaping this study.

I am also grateful to my **institution, Government college of engineering chandrapur**, for providing the necessary resources and a conducive research environment. Special thanks to my **professors, mentors, and colleagues** for their insightful discussions and suggestions that helped refine this work.

This research is a result of collective efforts, and I sincerely acknowledge everyone who played a part in its successful completion.

REFERANCES

- Sathish and Smys "A Survey on Internet of Things (IoT) based Smart Systems" Journal of ISMAC (2020) Vol.02/ No.04 Pages: 181-189, http://irojournals.com/iroismac/ DOI: https://doi.org/10.36548/jismac.2020.4.001
- Carvalho, T.P.; Soares, F.A.; Vita, R.; Francisco, R.D.; Basto, J.P.; Alcalá, S.G. A systematic literature review of machine learning methods applied to predictive maintenance. Comput. Ind. Eng. 2019, 137, 106024.
- Cinar, Z.M.; Nuhu, A.A.; Zeeshan, Q.; Korhan, O. Digital Twins for Industry 4.0: A Review. In Industrial Engineering in the Digital Disruption Era. GJCIE 2019. Lecture Notes in Management and Industrial Engineering; Calisir, F., Korhan, O., Eds.; Springer: Cham, Switzerland, 2020.
- Zhang, W.; Yang, D.; Wang, H. Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. IEEE Syst. J. 2019, 13, 2213–2227.
- Borgi, T.; Hidri, A.; Neef, B.; Naceur, M.S. Data analytics for predictive maintenance of industrial robots. In Proceedings of the 2017 International Conference on Advanced Systems and Electric Technologies (IC_ASET), Hammamet, Tunisia, 14–17 January 2017; pp. 412–417.
- Dai, X.; Gao, Z. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. IEEE Trans. Ind. Inform. 2013, 9, 2226–2238.
- Lee, J.; Lapira, E.; Bagheri, B.; Kao, H.A. Recent advances and trends in predictive manufacturing systems in big data environment. Manuf. Lett. 2013, 1, 38–41.