A Study of Random Forest Machine Learning based Predictive Maintenance in Industrial Maintenance

M. Sheik Syed Sulaiman

Technology Lead, Software Industry, Bengaluru, Karnataka, India.

Abstract - Industrial maintenance is the upkeep and repair of equipment and machines.It includes the task of equipment maintenance, Inspection and repairs.It aims to keep the machinery running efficiently and safely and prevents downtime.One of the most prevalent challenges in the industries like logistics, aircraft etc is the need to reduce the costs and delays, while maintaining and improving the operational reliability. Currently, these industries are trying to leverage data and technological progress to better predict and manage the maintenance efforts through predictive maintenance.Poor maintenance planning and inaccurate overview of maintenance can lead to overstocking of surplus machine parts and devastating financial results for carriers and led to logistic cancellations.Prediction requires machine learning models based on large amounts of data for each system component. The proposed system acquires the data about the machine parts through a standard input mechanism. The system will collect, store and send audio data for processing in the edge devices as well as the cloud. It will use sensor data visualization tool to analyse system health. Machine health reports will be periodically generated, and tools to forecast machine failures will be used. The system generates alarms and sends notifications to the concerned officials of the industry. The proposed methodology will achieve real-time computing prediction of failures of industrial machines. Industries focuses on process optimization, reducing costs and increasing efficiency.

Keywords - Predictive Maintenance, Preventive Maintenance, Reactive Maintenance, Logistics Industry, Operational Reliability, Industrial Maintenance.

I. INTRODUCTION

In Industries, machinery requires continual maintenance for proper working. To increase operational reliability and cost saving measures, these industries follow the maintenance programs. There are three well-known types of maintenance: reactive, preventive and predictive maintenance.

• Reactive Maintenance: This approach involves waiting for a piece of equipment to fail and then repairing or replacing it. Reactive maintenance is typically the least expensive option in the short term, but it can lead to increased downtime and higher costs in the long run.

• Preventive Maintenance: This approach involves performing maintenance tasks on a fixed schedule, such as replacing parts or conducting inspections, regardless of whether or not the equipment is showing signs of wear or malfunction. Preventive maintenance can help reduce the likelihood of unexpected failures but can also lead to unnecessary maintenance and downtime.

• Predictive Maintenance: This approach involves using data and analytic to predict, when a piece of equipment is likely to fail and then taking action to prevent the failure from occurring. Predictive maintenance can help minimize downtime and reduce maintenance costs, but it requires significant data collection and analysis investment.

Nowadays, many companies still follow a periodic maintenance or condition-based maintenance approach. While in periodic maintenance, industrial machines are maintained at regular intervals, condition-based maintenance involves defining threshold values for particular sensors, which trigger maintenance operations when exceeded. Periodic maintenance often leads to a waste of personnel and material since, in many cases, maintenance is not necessary and could be postponed.

Among all the maintenance approaches, predictive maintenance using machine learning is a powerful tool that can help organizations improve equipment reliability and reduce costs associated with unplanned downtime. Predictive maintenance is a technique used to predict equipment failure and prevent unplanned downtime by analyzing data from various sources, including sensors, historical maintenance records, and other operational data. Machine learning is an important tool used in predictive maintenance to automatically identify patterns and anomalies in this data, which can help identify potential issues before they become critical. Machine learning models can be trained using historical data to identify patterns that indicate equipment failure is likely to occur. These models can then predict when a piece of equipment is likely to fail and alert maintenance teams to take action before the failure occurs. Using machine learning in predictive maintenance can lead to significant cost savings by reducing downtime, improving equipment reliability, and extending equipment lifespan. It also enables maintenance teams to focus their efforts on the most critical issues rather than spending time on routine maintenance tasks.

In this research work, we propose Random Forest machine learning algorithm for predicting the type of failure that can arise in the machine.

II. LITERATURE

Several survey studies were published previously on the predictive maintenance in various industries by various researchers using various techniques, as summarized below briefly.

Gian Antonio Susto has proposed a methodology that has been demonstrated for a semiconductor manufacturing implant-related maintenance task and shown to guarantee better performance than classical PvM approaches and a single SVM classifier distance-based PdM alternative [1]. This case study has also shown that SVMs offer superior performance to kNN classifiers when implementing MCPdM and that, in general, MC-PdM-known also consistently outperforms PvM approaches.

The Industrial Internet of Things is used to gain valuable insights from machine information along with the data kind components for predictive modelling [2]. The researchers have explored using auto regressive integrated moving average forecasting trauma plating machines to predict quality defects, downtime and maintenance. Machine learning has been proven to be an important component in the industrial Internet of Things for quality management and quality control. It enhances performance and improves the manufacturing process.

In another study by Marina Paolanti, the methodology was implemented in the experimental environment on the example of a real industrial group, producing accurate estimations [3]. Data has been collected by various sensors, machine PLCs and communication protocols and made available to the Data Analysis Tool. The proposed PdM methodology allows dynamical decision rules to be adopted for maintenance management, which is achieved by training a Random Forest approach on Azure Machine Learning Studio. Preliminary results show a proper behaviour of the approach on predicting different machine states with high accuracy (95%) on a data set of 530731 data readings on 15 different machine features collected in real time from the tested cutting machine.

Similarly, Weiting Zhang and other authors proposed a methodology demonstrated as a universal end-to-end predictive maintenance methodology for the time-to-failure prediction of industrial machines [9]. This includes a universal sensor handling and feature extraction approach based on integral values, feature transformation, feature selection, adjustable target class labeling, training of different machine learning models, and their hyper parameter tuning. The results show that a feed-forward neural network with multi-class labeling managed to achieve the best prediction quality in terms of accuracy, F1-score, and kappa of 97.79% and 94.03%, respectively.

K. Liulys, in his study, shows that the process of regularly checking the equipment is performed so that the machine does not break down. However, preventive maintenance costs many irrelevant and repetitive checking [10]. It leads to greater costs that cannot be afforded frequently. Therefore, a new concept called predictive maintenance can be implemented in such machines to reduce the downtime cost and decrease the number of checking. Through predictive maintenance, the total time required for all the checking and the total cost spent on all the checking is reduced tremendously. Machine learning is the process of predicting output values by implementing programs on huge gathered data of core values.

Farzam Farbiz has introduced a cognitive analytic based framework for machine condition monitoring and anomaly detection. We applied the proposed framework to an industrial robot use case and validated the proposed approach [11]. The machine model generated by the proposed framework can learn the machine performance adaptive with the new data and detect anomalies of the robot movement in real-time. Although the proposed solution has demonstrated promising results and its unsupervised machine learning model can classify the data despite noise and outliers, it should be noted that it has been applied to one use-case study so far.

In a more recent study, the paper's authors conducted a survey of PdM of industrial equipment [12]. Initially, a brief introduction of the industrial PdM scheme was proposed, demonstrating the challenges faced. The paper's main aim is to provide a systematic overview of the PdM, propose an industrial PdM scheme for automatic washing equipment, and demonstrate the challenges faced when conducting a PdM research study.

Condition monitoring has been prevailing in industrial machinery failures for years. The idea of protecting against failures has been in the industries to reduce the cost of machine maintenance. [13] the paper describes a system that collects data from 30 industrial pumps at a thermo-chemical plant. This data is gathered and processed through the Random Forest Algorithm to establish relevant information. The paper has articulated the challenges that arise while implementing machine learning algorithms on the data and Systems performance.

A study by Amruthnath, N. and Gupta shows fault detection has been an important subject prevailing in industries in recent years [15]. Detecting faults is important to reduce the break time cost early and ensure proper running of the manufacturing processes. The research paper has proposed unsupervised learning for predictive maintenance by the system on the exhaust fan. Algorithms such as PCAT2 statistic, Fuzzy C-Means clustering, Hierarchical clustering, K-Means, and model-based clustering are used by the researchers for predictive maintenance. These algorithms are used, and the best one suitable for predictive maintenance has been proposed to ensure robustness.

II. PROPOSED APPROACH

A. Functional Requirements
a. Data Storage

The data for predictive maintenance is stored and collected from the cloud in a time-series format. The data is stored in two types of formats:

• Main Storage: A data lake is used to store the data through an online cloud. It is stored here for preprocessing and analyzing the data.

• Backup File: A readily available dataset is stored in .csv format and is used for developing a machine learning model and analyzing the data.

b. Data Transfer

It involves data movement from the cloud to the proposed system for analysis. The data transfer occurs through a secure communication channel provided by the cloud.

c. Data Pre-Processing and Processing

• Data Cleaning: Data cleaning is performed as the raw data may contain irrelevant or undesired elements. It manages the feeding of missing data and noisy values through procedures such as outlier detection and others.

• Data Transformation: Data Normalization is performed to scale the data values into the desired range. In this phase, attributes are selected as per suitability, and new attributes from existing or known attributes are generated.

• Data Processing: In this stage, Random Forest machine learning algorithm is applied to process the data. This algorithm is used to predict the data using values such as pressure, volume, temperature and others.

B. System Architecture

The system architecture for the predictive maintenance system consists of several components such as industrial machines, cloud for storage and a user interface screen.

The data collected from the cloud will be given to the model development part, which is processed and trained using Random Forest machine learning algorithm and the predicted results will be displayed on the user interface screen.

The proposed methodology will discuss the system's behaviour and maintenance schedule through predictive analysis. The system will be continuously assessed through the gathered data on pressures, vibrations, temperatures, power consumption. Random Forest Algorithm will be implemented on the gathered data to predict their behaviour. Through these algorithm, data will be processed to obtain the changes in the values of parameters like vibration, volume, etc., so that the breakdown period of any equipment can be predicte
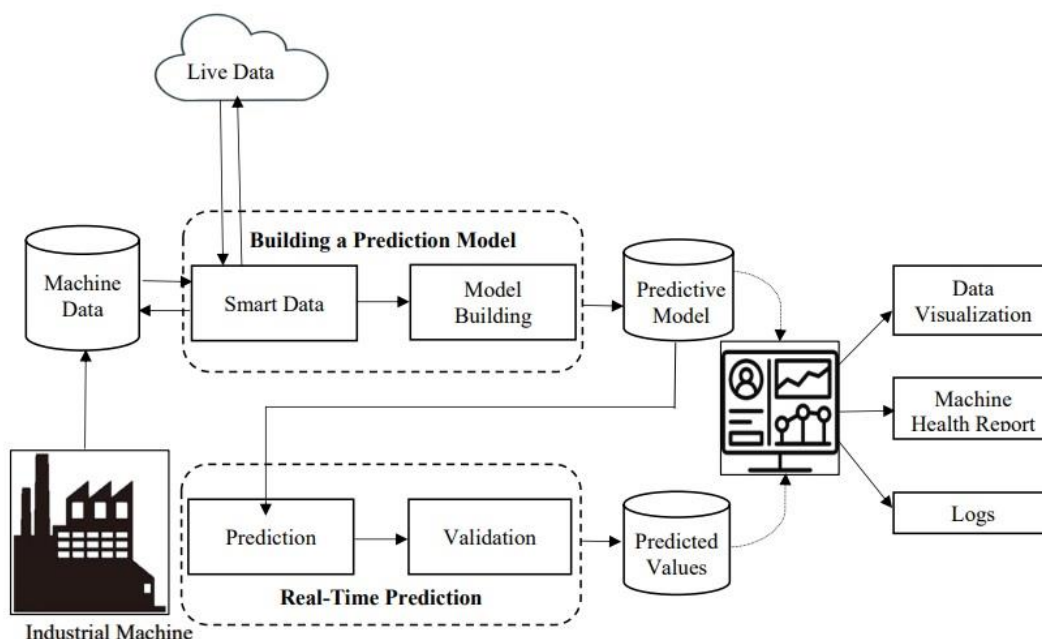

Fig. 1 System Architecture

III.   RANDOM FOREST ALGORITHM

Random Forest is the most popular machine learning algorithm used for predictive maintenance. It can be used for both classification and regression problems in machine learning. It is based on concept of ensemble learning method which combines multiple classifiers to solve a complex problem and to improve the performance of the model.

Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of over fitting.

Based on past data, Random Forest can be used in predictive maintenance to forecast machine failures or breakdowns. The algorithm analyses large amounts of data from sensors, devices, and other sources to find patterns and trends that point to possible issues. The algorithm then makes predictions about potential failures or breakdowns using this knowledge.

One of the key advantages of using Random Forest for predictive maintenance is its ability to handle large, complex datasets with many variables. The algorithm can handle numerical and categorical data and automatically select the most important features for making predictions.
Tool Wear Failure (TWF) - The tool will be replaced or fail at a randomly selected tool wear time between 200 – 240 mins.

1. Heat Dissipation Failure (HDF) - Heat dissipation causes a process failure if the difference between air- and process temperature is below 8.6 K, and the tool's rotational speed is below 1380 rpm.

2. Power Failure (PWF) - The product of torque and rotational speed (in rad/s) equals the power required for the process. If this power is below 3500 W or above 9000 W, the process fails.

3. Overstrain Failure (OSF) - If the product of tool wear and torque exceeds 11,000 minNm for the L product variant (12,000 for M, 13,000 for H), the process fails due to overstrain. The performance of the system is evaluated using the confusion matrix

Random Forest implementing steps for predictive maintenance
• Collect and clean historical data on equipment failures and maintenance activities.
• Identify relevant variables and features that may be predictive of future failures.
• Train the Random Forest model using the historical data.
• Test the model on new data to evaluate its accuracy and performance.
• Use the model to generate predictions about future failures or breakdowns.

V. RESULTS

Predictive system prototype is successfully built which is capable of sending data from the cloud on the edge device. Using a predictive maintenance dataset from kaggle, the condition of machine such as air temperature, rotational speed, torque, tool wear and process temperature, were monitored.

Random Forest machine learning algorithm was used to train the model on a given data set. Given below are the types of failure that may occur in a Machine:

VI. CONCLUSION

The review of prior research work conducted on predictive maintenance using machine learning is discussed in this paper. The conclusion drawn from the overall discussion is that implementing a predictive maintenance system can lower the costs associated with breakdowns in the various sectors like logistics industry, aircraft industry etc.

The research has been successfully conducted for the maintenance of machine on current data. The proposed system discussed can detect 4 types of faults in the machine. Thus, it can be concluded that the system can reduce the cost of maintenance and maximize an equipment's life. Thus, it can be concluded that the system can reduce the cost of maintenance and maximize an equipment's life.

REFERENCES

[1] Gian Antonio Susto et al., "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812-820, 2015.
[2] [2] Ameeth Kanawaday, and Aditya Sane, "Machine Learning for Predictive Maintenance of Industrial Machines Using IoT Sensor Data," in 2017 8th IEEE International Conference on Software Engineering and Service Science, pp. 87-90, 2017.
[3] [3] Marina Paolanti et al., "Machine Learning Approach for Predictive Maintenance in Industry 4.0," IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications, pp. 1-6, 2018.
[4] [4] R. Yam et al., "Intelligent Predictive Decision Support System for Condition-Based Maintenance," The International Journal of Advanced Manufacturing Technology, vol. 17, no. 5, pp. 383–391, 2001.
[5] [5]Archit P. Kane et al., "Predictive Maintenance Using Machine Learning," ArXiv, 2022.
[6] [6] Ayushi Chahal, and Preeti Gulia, "Deep Learning: a Predictive IoT Data Analytics Method," International Journal of Engineering Trends and Technology, vol. 68, no. 7, pp. 25-33, 2020.
[7] [7] Wo Jae Leea et al., "Predictive Maintenance of Machine Tool Systems Using Artificial Intelligence Techniques Applied to Machine Condition Data," CIRP Life Cycle Engineering (LCE) Conference, vol. 80, 2019.
[8] [8] Joel Anto Williams N et al., "Machine Predictive Maintenance System for Industrial Applications," International Journal of Current Research, vol. 14, no. 05, pp. 21410-21412, 2022.
[9] Weiting Zhang, Dong Yang, and Hongchao Wang, "Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey," IEEE Publication, vol. 13, no. 3, pp. 2213-2227, 2019.
[10] Kaorlis Liulys, "Machine Learning Application in Predictive Maintenance," 2019 Open Conference of Electrical, Electronic and Information Sciences (es- Tream), pp. 1-4, 2019.
[11] Farzam Farbiz, Yuan Miaolong, and Zhou Yu, "A Cognitive Analytics Based Approach for Machine Health Monitoring, Anomaly Detection, and Predictive Maintenance," 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1104-09, 2020.
[12] Marwin Zufle et al., "A Predictive Maintenance Methodology: Predicting the Time-to- Failure of Machines in Industry 4.0," IEEE 19th International Conference on Industrial Informatics (INDIN), pp. 1-8, 2021.
[13] Ido Amihai et al., "An Industrial Case Study Using Vibration Data and Machine Learning to Predict Asset Health," 2018 IEEE 20th Conference on Business Informatics, pp. 178-185, 2018.
[14] P. Strauß et al., "Enabling of Predictive Maintenance in the Brownfield Through Low-Cost Sensors, an IIoT-Architecture and Machine Learning," IEEE International Conference on Big Data (Big Data), pp. 1474-83, 2018.
[15] Nagdev Amruthnath, and Tarun Gupta, "A Research Study on Unsupervised Machine Learning Algorithms for Early Fault Detection in Predictive Maintenance," 5th International Conference on Industrial Engineering and Applications, pp. 355-361, 2018.
[16] B. Balaji et al., "Fault Prediction of Induction Motor Using Machine Learning Algorithm," SSRG International Journal of Electrical and Electronics Engineering, vol. 8, no. 11, pp. 1-6, 2021.