SME 4.0: Machine learning framework for real-time machine health monitoring system

Velmurugan K1, Venkumar P1* and Sudhakara Pandian R2

1Department of Mechanical Engineering, Kalasalingam Academy of Research and Education, Virudhunagar, Tamilnadu, 626126, India.
2School of Mechanical Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, 632014, India.

*Corresponding Authors Email: p.venkumar@klu.ac.in

Abstract. Over the past ten years, manufacturers and consumers have become increasingly interested in the applications of smart, sustainable, and autonomous systems in the industry and in everyday life. Due to the recent industrial revolution (Industry 4.0), most of the existing Small and Medium-sized Enterprises (SMEs) also want to adapt their work environment into the smart system by the applications of these technologies such as the Industrial Internet of Things (IIoT), Artificial Intelligence (AI) techniques, Machine Learning Algorithm (MLA), Internet Communication Technology (ICT), and Cyber-Physical System (CPS). Because they are very much interested in maximizing productivity, machine availability, reliability, and customer satisfaction in this competitive industrial world. This research study particularly focuses on the Predictive Maintenance (PdM) activity of critical machines and their components in the SME based on the maintenance history dataset through the application of the supervised machine learning algorithm such as Logistics Regression (LR) and K-Means (K-Nearest Neighbor) approaches. In accordance, the real-time case study is presented in SMEs in the southern region of Tamil Nadu, India with two-phase activities. Initially, the optimal failure rate of the machines is predicted by the utilization of LR trained models. Then trigger the man-machine communication and suitable decision-making process of service and maintenance activity through the application of the K-Mean approaches. The main objective of this research study is to organize the smart PdM activity of the smart factory systems in SMEs with the application of MLA based on the real-time maintenance history dataset.

Keywords: Supervised Machine Learning Algorithm, Predictive Maintenance, Industrial Internet of Things, Cyber-Physical System, Industry 4.0, Small and Medium-sized Enterprises.

1. Introduction

Many industry experts and researchers have summarized the revolution of the present conditions in the industry from its inception into four parts. Accordingly, in the early eighteenth century steam-powered machines and workers used more energy to produce goods and as a result could not meet the needs of a sufficient number of goods and customers in a timely manner, declaring this to be the first revolution of the industry. Subsequently, in the nineteenth century, they introduced electric powered machines to resolve some of the problems of the previous industrial revolution, thus reducing production time, the manpower of workers and meeting the needs of production materials and customers but not completely solving the problems. They pointed to these actions as the second revolution in the industry. In the twentieth century, computer-integrated machines further reduced production time and produced more products to meet the challenges of increasing productivity in the industry and thus meeting the needs of customers, although the share of workers was somewhat reduced and the demand for skilled workers who knew about the computer increased as the third revolution in the industry. Finally, in the 21st century, they decided to produce products with autonomous Machines in the industry with or without the help of workers who introduced various sophisticated technologies which they named the Industrial Revolution 4.0, which they call the Industry 4.0. Among the latest trends, the implementation of the smart factory system in the SMEs remains a major challenge as many SMEs in the South of India continue to manufacture with a more primitive structure due to the lack of knowledge, application, and awareness of the latest Industry 4.0 tools and technologies. But such companies are ready to transform their industry into a smart structure at a low cost. Applications of sophisticated technologies are the most essential function for implementing a smart factory system. Some important key technologies of Industry 4.0 are shown in figure 1. All of these technologies are currently being used in the highlands to transform a normal industry into a smart. The most important part of all this is the technology of
the IoT, IIoT because its functionality is the main source for all other technologies. This technology includes the following functions: 1. Monitoring, 2. Controlling, 3. Creating and 4. Computerizing. This allows for continuous monitoring of machinery, machine tools, and manufacturing materials, control of changes in them, optimized decision making and computerization of all these, and the transmission of information about it to the person concerned in a timely manner through wireless internet and Bluetooth assistance.

![Diagram of Industry 4.0 Key Terms]

**Figure 1. Advanced Technologies of Industry 4.0**

1.1 Predictive Maintenance

Generally the maintenance is categorized into three different types such as Design out Maintenance (DM), Preventive Maintenance (PM) or Scheduled Maintenance (SM), and Breakdown Maintenance (BM). Industry 4.0 PM function is further classified into another section like Predictive Maintenance (PdM). The PdM is closely similar to the PM because both activities are carried out before the failure occurs and prevent the sudden breakdown of the manufacturing machine in the industries [3, 4]. In PM action the service and maintenance carried out based on the defined time schedule (weekly, monthly, quarterly, and yearly) whether the machine is failed or not fully depends on the defined interval time in the manufacturing plant. In PdM action slightly differs from the PM action because it’s carried out based on the real-time behavior of the machine there is no time constraints for doing that action in the production plant. The main objective of this PdM is to minimize the uncertainty in diagnosing the machine failure and maximize the availability and Remaining Useful Life of the machine and its components in the industry.

2. Literature review

In this section organize the critical overview of various research articles with the application of smart maintenance and factory systems. An already published research article summarizes the smart functions, benefits, challenges, and applications of essential items used in everyday life [1]. A technical reference article also describes how to build smart and sustainable manufacturing plants using existing state-of-the-art technologies, as well as the challenges, uses, and opportunities of these technologies [2]. It is possible to diagnose the lifespan of machines used in industry with mathematical modeling before it is repaired and to implement an optimal and optimal maintenance process and to provide a better maintenance checklist on how to maximize the performance of those machines [3, 4]. An important study explains how the smart system (Industrial Revolution 4.0) is implemented in the industry with process descriptions of machines in the small and medium manufacturing industry [5, 6]. A research paper briefly describes how to implement the forecast maintenance process in the industry using various architectural levels of advanced technologies such as the Internet of Things and the cyber physical system [7]. The ML based smart predictive model of the Heating Ventilation and Air Condition Systems (HVAC) in the industry 4.0 is demonstrated. The accuracy and precision of the train, test data error evaluation measured through the application of the LR and Random forest supervised MLA techniques [8, 9, 10]. Proposed the smart injection molding operation in the
foundry is presented through the utilization of the real time machine data with big data analysis techniques. This smart framework supports the construction of the computerized manufacturing activity with minimum down time in the working environment [11]. The real time predictive maintenance model of the wind turbine is illustrated. The behavior of the critical components in the wind turbine is monitored through the implementation of the random forest algorithm and big data environment. This proposed model applied into various applications such as Apache Spark, Apache Kafka, Apache Mesos and HDFS [12]. The recent applications, major challenges and opportunities of the MLA and big data platform in the industry 4.0 is illustrated. The major challenge in implementing the ML enabled prediction model and its various phases of activities are presented in this research article [13]. Illustrated the significance and highlights feed forward Artificial Neural Network (ANN) approach to develop the quality and shelf-life prediction of the roasted coffee flavored sterilized drink through the application of the ML techniques. The different back propagation ANN model is investigated to measure the sensory quality of the coffee makers [14]. The real time case study of the spring factory is presented, the computerized machine status predicted through the utilization of the inexpensive add on triaxial sensors with the ANN approach in the spring production factory [15]. Described the evaluation of supplier selection in the industry through the application of the hybrid fuzzy decision model [16]. From this overview to understand the present challenges and opportunities and research gaps in the future industrial revolution for the competitive world.

3. Materials and Method of smart maintenance system

In this research the recent industry 4.0 tools and technologies are utilized to implement the smart working environment in the production plant such as AI technique, MLA, IIoT and ICT. Here the most widely used supervised MLA is utilized to predict the optimal maintenance model and proposed the smart maintenance management system in the small and medium scale manufacturing industries. In the present working environment a set of data has been obtained before the MLA applied here the maintenance datasets are gathered through the monitoring the real time behavior of the manufacturing machines in the SMEs. These supervised ML analysis techniques consist of different phases of activities subsequently such as Pre Processing dataset, Trained the dataset, utilization of the suitable machine learning model and finally evaluate the optimal solution of the given problems [8, 9, 10]. The detailed illustration of sequential stages designed in the MLA techniques is demonstrated in the upcoming sections. Initially the raw maintenance data was prepared by collecting the real time monitoring IIoT sensors in this stage the original maintenance data are unstructured, incomplete, inconsistent and much noisier because of that data preprocessing is performed. Then structured maintenance datasets are transformed into the input values of the MLA selected for the testing and training of the maintenance datasets. That test maintenance datasets will be utilized to implement the prediction model and also extract the new set of maintenance data to be obtained. Finally to evaluate the trained maintenance dataset for measuring the errors in the trained dataset and analyze the optimal solution of the test dataset of the maintenance in the given SMEs.

3.1 Case study

In this research, we choose the sensors and switch production system of the SMEs (XYZ private Limited) cited in the southern region of Tamilnadu, India. That small and medium scale industry was producing numerous electronic devices for the leading automobile industries. They work with some important manufacturing processes (Stamping, Molding, Soldering, Cutting, Punching and Embossing etc) with the utilization of the hydraulic and pneumatic pressing machine in the manufacturing plant. They are facing a lot of challenges in organizing the optimal maintenance management system and maximum availability of the machines in the production plant. For this research analysis the IIoT sensors placed on the manufacturing machine to continuous monitoring activity and generate maintenance datasets based on the real time behavior of the machines. The main objective of this proposed research framework is to implement the predictive maintenance management system and extend the Remaining Useful Life of machines and its critical equipment
in the SMEs. The detailed demonstration of this research analysis is illustrated in below.

3.2 Description of the maintenance dataset

In this real time maintenance case study a dataset constructed by columns was utilized. It consists of the optimal failure rates of the machines and the real time values gathered by the IIoT sensors in the manufacturing plant. That dataset was used to analyze the behavior of the machines and its critical components of the production systems and predicted if the machines are repaired to maintain the reliability and availability of the machines in an optimal [8, 9, 10]. This maintenance dataset totally contains 13800 (Thirteen thousand eight hundred) failure rate records (Target rate of failure) monitored by the IIoT sensor networks, installed in five machine works on twenty four hours in last one year (5*24*365) of the production plant their corresponded repair rate of the machine to predict the independent variable machine availability. The structure of the maintenance dataset and its different variables are illustrated in Table 1.

Table 1. Maintenance Dataset model

| S.No | Column     | Data description             |
|------|------------|------------------------------|
| 1    | Machine ID | Machine Identification      |
| 2    | Date       | Date of Monitor              |
| 3    | Timestamp  | Monitoring time              |
| 4    | Target failure rate | Failure rate measured by the sensor |
| 5    | Actual failure rate | Optimal failure rate for the machine |
| 6    | Availability | Availability of Machine      |
| 7    | System     | System model                 |

3.3 Maintenance dataset Pre Processing

The numerous researchers have already used various classification and search algorithms to select features and preprocessing the collected raw data in the industries [8, 9, 10]. The graphical representation of the machine learning train model is shown in figure 2. This proposed model organized a range for the normal failure rates and three type of information sharing (Machine indication lamp blaring, Alarm and Short message service) to the three different responsible persons (Machine operator, Plant supervisor and Maintenance engineer) that the extreme variation of failure rates and therefore a possible machine repaired in the production plant. These are illustrated in Table 2. As follows

1. Normal: within the failure rate envelope region (Below the optimal values)
2. Up normal: exceed the failure rates envelope region (Smaller variations in the optimal values). It classified as extreme failure rates and cause of possible machine repairs.
3. Extreme up normal: exceed the failure rates envelope region (Higher variations in the optimal values). It classified as extreme failure rates and cause of possible machine repairs.

Table 2. Failure rate variation

| Rate of failure | Description                        | Light indication |
|-----------------|------------------------------------|------------------|
| Normal          | If (Target - Actual failure rate) < 0.005 | Green           |
| Up Normal       | If (Target - Actual failure rate) > 0.005 | Orange          |
| Extreme Up Normal | If (Target - Actual failure rate) >+0.01 | Red             |

Two different labels of columns are added to the maintenance datasets ‘Variation range’ and ‘Filter range of variation’ in the first column the values obtained from the variation between the ‘Target and Actual failure rates of the particular machines in the production plant. In the next column ‘Filter range of variation’ consists of binary conversion 0 for normal failure rate and 1 for initiating the alarms of the extreme failure rate variations in the manufacturing plant of the industry.
4. Result and Discussion

Once the maintenance datasets were preprocessed, this extended maintenance datasets was utilized to classified into the Train dataset and Test dataset then applied the MLA for organize the optimal and real time machine health prediction framework (Prognostic Health Management System) in the production plant of the SMEs. This framework analyzed the maintenance datasets and validated the effect of failure rate changes through the application of the most widely used classification algorithms of supervised learning MLA such as Logistic Regression and K-Mean approaches. The LR technique was applied to evaluate the accuracy of failure rate variations in the prediction framework and K-Mean approach utilized to identify the optimal, nearest responsible maintenance and service center based on the stored information of the service center details. The LR approach is an extended version of the linear regression approach; the only difference is that it has classified the variables in binomial (0 and 1). The implementation of this research study maintenance datasets was preprocessed successfully and categories the given variables (V) are stored in binary values. Applying the LR approach we assume that V=1, when the IIoT sensor sends on extreme failure rate and V=0, when the monitored failure rate (Target failure rate) is within the optimal envelope region of failure rate range. Table 3 illustrates the proposed model maintenance datasets of the given manufacturing machine (machine ID: HPMC010) in the last year. In this perspective machine health monitoring and control action developed through the application of the IIoT in the manufacturing unit of the SMEs. Based on the failure rate variation the decision making of preventive maintenance activity initiated through different real time information sharing approaches such as Machine top lamp indication, Audio or voice indication, and short message service indication.

Table 3. Perspective smart preventive maintenance dataset

| S. No | Date & Time stamp | Machine ID | Actual Failure rate | Target Failure rate | Variation range | Lamp indication |
|-------|-------------------|------------|---------------------|---------------------|-----------------|----------------|
| 1     | 01.01.20&9.00 am   | HPMC010    | 0.0336              | 0.00738             | 0.00378         | Green          |
| 2     | ……..              | HPMC010    | 0.0336              | ……..               | ……..           | ……..          |
| 3     | ……..              | HPMC010    | 0.0336              | ……..               | ……..           | ……..          |
| 4     | 07.07.20&9.30 am   | HPMC010    | 0.0336              | 0.01039             | 0.00678         | Orange         |
| …….. | ……..              | HPMC010    | 0.0336              | ……..               | ……..           | ……..          |
| …….. | ……..              | HPMC010    | 0.0336              | ……..               | ……..           | ……..          |
| …….. | 10.12.20&10.00am   | HPMC010    | 0.0336              | 0.04344             | 0.03982         | Red            |

Considering the above conditions that the manufacturing machines in the production system is occurring a repair by recording the extreme variation of the failure rate is given in the following equations

\[ P(V = 0) = 1 - P(V = 1) \]  

(1)
\[ V = f(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n) \]  
\[ Z = B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n \]  
\[ e^Z = e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)} \]  
\[ \text{Prediction} = \frac{1}{1 + e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)}} \]  

If \( Y > \text{Optimal failure rate} \) Classify as 1
Else \( Y < \text{Optimal failure rate} \) Classify as 0

\[ B_i = B + \alpha(Y - \text{Prediction}) \times \text{Prediction} \times (1 - \text{Prediction} \times X_j) \]

Where,
\[ i = 1, 2, 3, \ldots, n \]
\[ \alpha = \text{Constant (0.1 to 0.3)} \]

\[ f(z) = \frac{e^z}{1 + e^z} = \frac{e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)}}{1 + e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)}} \]  
\[ P = P(Y = 1) = \frac{e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)}}{1 + e^{\left(B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n\right)}} \]

\[ \text{Sigmoid} \Rightarrow \ln \left(\frac{P}{1 - P}\right) = B_0 X_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_n X_n \]

The selected set of independent variables of the given research is denoted by \( X_1, X_2, X_3, \ldots, X_n \), where \( n \) is the total numbers of the variable. To measure the probability (\( P \)), we utilized the sigmoid or logit function of the LR mathematical modelling represented in equation 9. It indicates the precision and accuracy of the failure rate variations (0.698032) by means of the LR approach can be identified as the prediction of failure rate values. The availability column of the maintenance datasets was considered a characteristic and as a label the Filter variance column the total maintenance data used was 13800 IIoT sensor monitored records. The optimal accuracy results are obtained based on the 70 percent of X-Trained (9660) and 30 percent of X-Test (4140) dataset values are taken respectively. After completing the LR approach based on the maintenance datasets then used K-Mean search algorithm to categori es the nearest service Centre and available service engineer details with concerned the stored information of the maintenance department records, by measuring the Euclidean distance. In this situation we consider three different features (precision limit) such as Easy to attend service (0-30Km), possible to attend service (31-60Km) and difficult to attend service (61-100Km) for achieving the optimal service and maintenance activity.

\[ d(M, L, N) = \sqrt{(M_1 - M_2)^2 + (L_1 - L_2)^2 + (N_1 - N_2)^2} \]

\[ \text{SSE}_{M} = \sum_{i=0}^{n} d(M_i - C_M); \text{SSE}_{L} = \sum_{i=0}^{n} d(L_i - C_L); \text{SSE}_{N} = \sum_{i=0}^{n} d(N_i - C_N) \]

\[ \text{SSE} = \text{SSE}_M + \text{SSE}_L + \text{SSE}_N \]

Where,
M- Easy to attend service (0-30Km)
L- Possible to access (31-60Km)
N- Difficult to attend service (61-100Km)
Based on this proposed ML enabled smart maintenance framework to optimize the maintenance, manufacturing and quality of the product in the small scale industries. The uncertainty of machine failure and malfunction diagnosis time will be minimized through this proposed framework. The IIoT enabled smart, autonomous machine health monitoring and manufacturing systems illustrated in figure 3. The prospective smart machine health monitoring system will lead to the maximum availability of the critical machine components, maximize the productivity with higher precision and achieved better customer satisfaction. This proposed smart maintenance and machine health monitoring framework consist of some limitations. For example, the single maintenance team should monitor and control the machine activity in the SMEs. Technical skilled employees required to predict the optimal maintenance parameters. Computerized environment needs to analyze the datasets from the machine standard output signals of the particular unit.

5. Conclusion
The proposed smart service and maintenance prediction based on machine learning framework is still on initial stage of implementation in the industry. The outcome of this research is to predict the optimal failure rate of the machines based on the real time data of the IIoT sensors. Trigger the optimal decision making process of the maintenance through the man machine communication technology. The solution of this case study revealed the optimal precision of the failure rates and malfunction of the machines in the manufacturing industries are identified through the utilization of logistic regression approach in supervised machine learning algorithms. In future research this proposed modeling integrated the large volume of real time machine behavior data that are gathered by the IIoT sensors to address the service and maintenance problems in the small scale manufacturing industries. Thus test and train the model with other classification, prediction machine learning algorithms. These trained models will initiate the implementation of algorithms that provide the smart and predictive maintenance framework for the SMEs in the industry 4.0 context.

References
[1] Benjamin K, Sovacoool and Dylan D, Furszyfer Del Rio 2020 Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies Renewable and Sustainable Energy Reviews, 120:109663–109663
[2] Mohamed Abubakr, Adel Abbas T, Italo Tomaz, Mahmoud S, Soliman, Monis Luqman, and Hussien
Hegab 2020 Sustainable and Smart Manufacturing: An Integrated Approach *Sustainability*, 12(6):2280–2280

[3] Velmurugan K, Venkumar P, and Sudhakarapandian R 2019 Design of Optimal Maintenance Policy using Markov Model *Int. J. Eng. Adv. Technol.*, 9:907–917

[4] Velmurugan K, Venkumar P, and Sudhakarapandian R 2019 Reliability Availability Maintainability Analysis in Forming Industry *Int. J. Eng. Adv. Technol.*, 9:822–828

[5] Pai Zheng, Honghui wang, Zhiqian Sang, Zhong Ray Y, Yongkui Liu, Chao Liu, Khamdi Mubarok, Shiqiang Yu, and Xun Xu 2018 Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives *Frontiers of Mechanical Engineering*, 13(2):137–150

[6] Anna Rosaria Boccella, Piera Centobelli, Roberto Cerchione, Teresa Murino, and Ralph Riedel 2020 Evaluating Cen- tralized and Heterarchical Control of Smart Manufacturing Systems in the Era of Industry 4.0 *Applied Sciences*, 10(3):755–755

[7] Li Dong, Ren Mingyue, and Meng Guoying 2017 Application of Internet of Things Technology on Predictive Maintenance System of Coal Equipment *Procedia Engineering*, 174:885–889

[8] Candanedo, Inés Sittón, Elena Hernández Nieves, Sara Rodríguez González, Teresa Santos Martín M, and Alfonso González Briones 2018 Machine learning predictive model for industry 4.0. In *International Conference on Knowledge Management in Organizations*, pp. 501-510

[9] Janitza, Silke, Gerhard Tutz, and Anne-Laure Boulesteix. 2016 Random forest for ordinal responses: prediction and variable selection *Computational Statistics & Data Analysis* 96: 57-73

[10] Candanedo, Inés Sittón, Sara Rodríguez González, Juan Manuel Corchado Rodríguez, Elena Hernández Nieves, and Fernando de la Prieta 2019 Fault predictive model for HVAC Systems in the context of Industry 4.0 In *XVIII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2018)*, pp. 21-26

[11] Lee, Hwaseop, Kwangyeol Ryu, and Youngji Cho 2017 A framework of a smart injection molding system based on real-time data *Procedia Manufacturing*, 11: 1004-1011

[12] Canizo, Mikel, Enrique Onieva, Angel Conde, Santiago Charramendieta, and Salvador Trujillo 2017 Real-time predictive maintenance for wind turbines using Big Data frameworks In *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 70-77

[13] Zhou, Lina, Shimei Pan, Jianwu Wang, and Athanasios, Vasilakos V 2017 Machine learning on big data: Opportunities and challenges *Neurocomputing*, 237: 350-361

[14] Goyal, Sumit, and Gyanendra Kumar Goyal 2013 Machine learning ANN models for predicting sensory quality of roasted coffee flavoured sterilized drink *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal* 2, 3: 09-13

[15] Kuo, Cheng-Ju, Kuo-Cheng Ting, Yi-Chung Chen, Don-Lin Yang, and Hsi-Min Chen 2017 Automatic machine status prediction in the era of Industry 4.0: Case study of machines in a spring factory *Journal of Systems Architecture*, 81: 44-5

[16] Pitchipoo Pandian, Ponmusamy Venkumar, and Sivaprakasam Rajakarunakaran 2013 Fuzzy hybrid decision model for supplier evaluation and selection *International Journal of Production Research*, 13: 3903-3919