EasyChair Preprint

№ 5806

Predictive Maintenance System Integrated with Periodic Maintenance: Machine Learning and Classical Approaches

Kamer Pı̈nar Bahar, Gazi Bilal Yı̈ldı̈z and Banu Soylu

June 15, 2021

Predictive Maintenance System Integrated with Periodic Maintenance: Machine Learning and Classical Approaches

Kamer Pınar Bahar[1], Gazi Bilal Yıldız[2] and Banu Soylu[1]

[1] Erciyes University, Industrial Engineering, Kayseri, Türkiye
kpinarbahar@gmail.com, bsoylu@erciyes.edu.tr

[2] Hitit University, Industrial Engineering, Çorum, Türkiye
bilalyildiz@hitit.edu.tr

ABSTRACT

With the fourth industrial revolution, the preventive and predictive maintenance, replaced the corrective maintenance, have been popular in recent years. Preventive maintenance is a scheduled maintenance strategy applied in order to reduce failures. On the other hand, predictive maintenance strategy requires to monitor equipment continuously and to analyze the data. The main objective of the predictive maintenance is to predict problems on the equipment that may lead to stops and to maximize utilization of the machine/equipment. It is reasonable to eliminate some failures with preventive maintenance while predictive maintenance can be applicable to eliminate others. For this purpose, we present a few criteria to determine maintenance strategy that will be applied to eliminate failures. Smartly integrating predictive and preventive maintenance will help to improve sustainability of the system. In this study, the preventive maintenance period is determined considering classical approaches such as Weibull analysis. We analyzed the failures of a specific machine for a time period. We also collected data about the system, environment and the machine condition during failures. We utilized machine learning algorithms in order to predict the type of possible failure and associations. The proposed decision support system helps to update the maintenance program with respect to results of machine learning methods. We perform a real-life case study and present our results.

KEYWORDS – Industry 4.0, Predictive maintenance, Weibull analysis, Machine learning.

1 INTRODUCTION

With the development of new technologies, a new industrial revolution, called industry 4.0, has been appeared. In recent years, improvements in technology and social media have significantly influenced customers in terms of product innovation, quality, variety and delivery on time. Satisfying these needs requires establishing the facility with capability of self-awareness, self-prediction, self-comparison, self-reconfiguration and self-maintenance [1]. Industry 4.0 brings enormous importance for the business in terms of maximizing productivity and ensuring just-in-time manufacturing and self-maintenance. Unexpected failures decrease productivity, deteriorate schedules and lead to financial loses. However, earlier detection of failures can prevent these problems.

Maintenance strategy is divided into two categories as *unplanned* and *planned*. Unplanned maintenance is often called as reactive maintenance or corrective maintenance. In this maintenance strategy, the repair is performed only when a failure occurs. Unplanned maintenance leads to long breakdown time and high maintenance cost. Planned maintenance strategy refers to proactive strategy. The aim of this maintenance is to minimize breakdown, reduce maintenance cost and maximize equipment performance. This maintenance contains all necessary activities that should be carried out on the equipment or machine. There are two most popular types of planned maintenance strategies as preventive and predictive maintenance. Preventive maintenance is a planned activity that is performed regularly on the equipment to reduce probability of failure and breakdown time. Moreover, a planned maintenance of equipment helps to extend equipment lifetime and reduces spare parts inventory. Predictive maintenance refers to condition-based or condition monitoring approach, i.e., the maintenance based on the analysis of current condition of the machine or the system. Predictive maintenance approach uses tools such as vibration monitoring, process parameter monitoring, thermography, tribology, visual inspection, sensor data etc. [2].

Maintenance approaches for diagnostic and prognostic purposes can be grouped into three main categories: statistical approaches, artificial intelligence approaches and model-based approaches. Compared to other approaches, artificial intelligence approaches have been applied progressively [3].

Machine Learning (ML), with artificial intelligence, is an effective tool for developing intelligent predictive algorithms in many applications. ML approaches deal with big dimensional data and extract hidden relationships within data. Hence, ML provides a powerful predictive tool for maintenance applications [4].

In this study, we analyze the historical data collected from a machine, and present a decision support system using the results of Weibull analysis and machine learning methods. As the main contribution, we proposed an integrated strategy as different from the literature. First, we categorized all failures into three groups as electrical failures, mechanical failures and pollution. We collected data about the temperature, failure time, raw-material processed and number of products until failure. Then, we explored the associations between the levels of these factors and each failure. We also introduced these attributes into the machine learning algorithms such as SVM, decision tree etc. For final decision of modifying maintenance program, we evaluated the results of Weibull analysis and machine learning methods. This leads to make dynamic maintenance actions to eliminate failures.

2 LITERATURE REVIEW

In the literature, there are many studies about preventive maintenance such as reliability-based model, maintenance cost optimization model etc. Some of them are based on the Weibull distribution to calculate failure rate and system reliability. Koçer [5] developed the preventive maintenance strategy based on the reliability and minimizing maintenance cost. Weibull distribution parameters were used to calculate system failure rate coefficient and reduction rate of equipment age after each maintenance. Doostparast et al. [6] carried out preventive maintenance based on the deteriorating components in the system. Minimizing cost-based model and reliability-based model were integrated. Weibull parameters were used to calculate reliability of component.

The studies about predictive maintenance are relatively new. Li and He [7] stated that there are two application techniques of predictive maintenance. The first one is the classification approach, which predicts the condition of the machine. Second one is the regression approach, also called Remaining Useful Life (RUL), predicts the time left until the next failure.

Machine Learning (ML) algorithms such as support-vector machine (SVM), random forest (RF), artificial neural networks (ANN) and k-nearest neighbours (K-NN) are successfully applied to design predictive maintenance applications. Carvalho et al. [8] review the studies applying ML methods in predictive maintenance.

Praveenkumar et al. [9] employed SVM algorithm to identify failures in automotive transmission boxes. To classify gear as good and faulty, vibration data was used in SVM. Li et al. [10] predicted alarm faults in a bearing of a rail network by using SVM algorithm.

Durbhaka and Selvaraj [11] used k-means algorithm to classify types of faults in the wind turbines. They compared results of K-NN and SVM algorithms with k-means algorithm. Eke et al. [12] applied k-means algorithm to extract clusters in a dissolved gases data in the oil of a transformer. Uhlmann et al. [13] identified clusters with k-means algorithm using data collected from a laser melting machine. On the other hand, Amruthnath and Gupta [14] compared a few clustering algorithms such as hierarchical clustering, k-means, fuzzy c-means clustering and model-based clustering to detect failures in an exhaust fan.

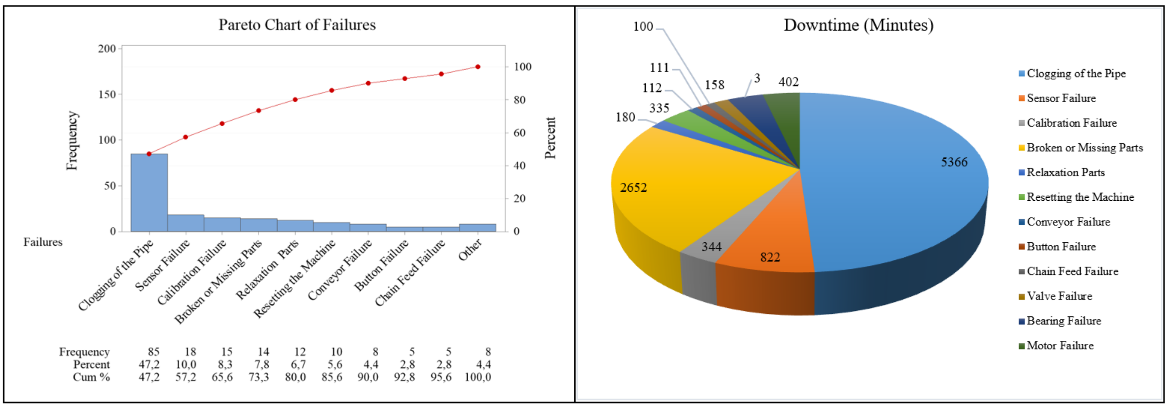
Biswal and Sabareesh [15] carried out predictive maintenance in order to classify the condition of a wind turbine as healthy and faulty. Certain characteristic features of healthy and faulty states were extracted by using ANN algorithm. Kolokas et al. [16] predicted condition of an engine by using ANN algorithm. They collected data about some features such as temperature, engine pressure, fuel and coolant bleed.

Paolanti et al. [17] aimed to apply predictive maintenance approach on a cutting machine. They employed Random Forest approach to predict different states of machine. Çakır et al. [18] used a few ML algorithms such as SVM, linear discrimination analysis (LDA), RF, decision tree (DT), and K-NN. The data has some features such as vibration, sound, rotational speed and temperature. Ayvaz and Alpay [19] carried out predictive maintenance for production lines in manufacturing. They compared ML algorithms such as RF, XGBoost, Gradient Boosting, MLP Regressor, support vector regression (SVR) and AdaBoost to find the most suitable prediction model.

Association rules have also been applied to predictive maintenance. Antomarioni et al. [20] aimed both predicting components breakages through association rule mining and determining the optimal set of components to repair to improve the overall plant's reliability, under time and budget constraints using integer linear programming. Foguem et al. [21] used association rule mining to extract information on fault causes in a drilling process. Djatna and Alitu [22] applied association rule mining in a total productive maintenance strategy.

3 PROBLEM DEFINITION

In this study, we first analyzed the failures of a specific machine for a time period. Figure 1(a) illustrates the Pareto chart of failures occurred in last two years. As can be seen from the chart, 80% of failures are due to five causes. That is if these causes are eliminated, then 9364 minutes downtime will not appear as shown in Figure 1(b).



(a) (b)

Figure 1. Pareto chart of failures and Pie chart of downtimes

We categorized these failures into three categories as electrical, mechanical and pollution. We integrated preventive and predictive maintenance for failures and determined the criteria given below in order to categorize maintenance requirements. It is better to apply preventive maintenance in order to eliminate failures if

- the diagnosis requires to stop the machine and screw off,
- collecting data is expensive,
- the maintenance period is deterministic and given in a manual.

For example, these criteria are satisfied for electrical and mechanical failures of the considered machine while the pollution can be eliminated with periodic cleaning.

First, we analyze the failure data collected from the pillow filling machine with respect to shifts as shown in Figure 2. When the machine works in a single shift, the frequency of failures is the highest between 11 and 12 o'clock. When double shifts are applied, the number of failures in the night shift (00:00-08:00) is lower than the day shift (08:00-20:00).

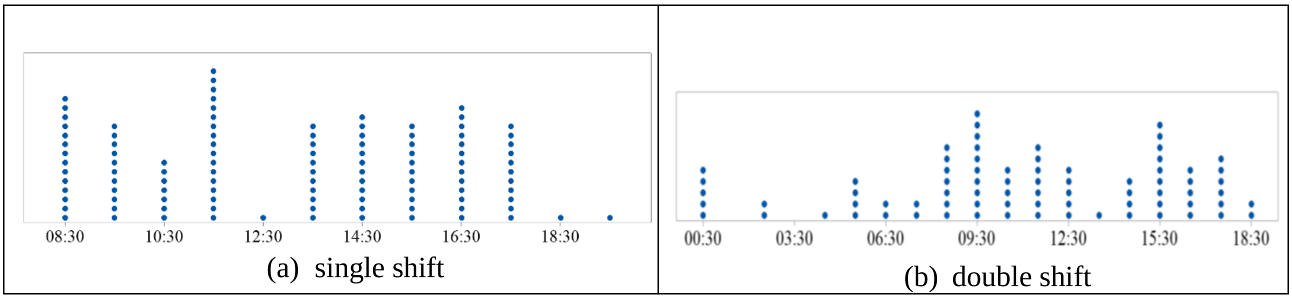


Figure 2. Dot-plot of failures with respect to shifts

We also analyzed the effect of weather condition, i.e., outside temperature, on the machine failures. According to dot-plot in Figure 3, as the outside temperature increases, the number of failures shows an increasing trend.

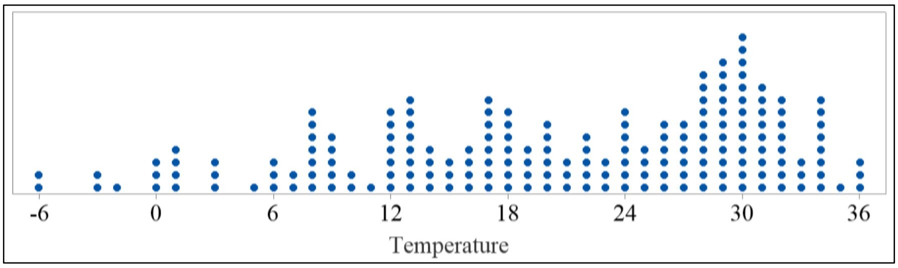


Figure 3. Dot-plot of failures with respect to temperature

4 METHOD

The preventive maintenance should not be disregarded when the predictive maintenance strategy is in progress. The proposed decision support system is summarized in Figure 4. First, historical data was collected from the considered machine. Secondly, the data was categorized into three failure groups: "Mechanical Failures", "Electrical Failures" and "Pollution". Predictive maintenance strategy is suitable according to above criteria in order to eliminate electrical and mechanical failures, while preventive maintenance strategy is applied to all failures in general. When the mean time between failures is analyzed for pollution, results showed that pollution failures have Weibull distribution. Because of this, preventive maintenance period has been calculated by using

Weibull distribution. The data has features such as failure time, number of products produced until failure, outside temperature, and raw materials of the last three lot at just in failure time. Then, several ML methods were applied to predict failure type. Finally, the result of the ML method is feed into Stage 4 to redetermine the maintenance period. Thus, we expect a steadier system in future.

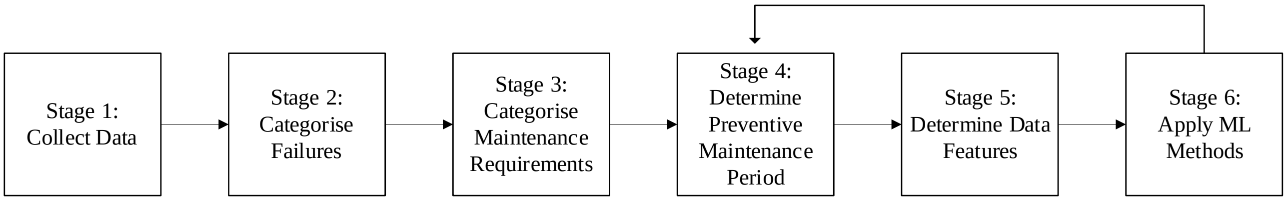


Figure 4. Processing flow chart of applying ML method

Association rule is one of the popular machine learning methods. Association rule is based on machine learning method for exploring and interpreting relations between transactions in large datasets. This learning method aims to identify strong relations that discovered in datasets. In order to select interesting rules from the set of all possible rules, constraints on various measures of significance and interest are used. The best-known constraints are minimum thresholds on support and confidence. In case of maintenance, an item can be assumed as a level of attribute and a transaction can be assumed as the set of levels that have failure in a given time interval. Assume that set T includes N number of transactions.

Let A be the set of failures {mechanical, pollution, electrical} and F is an element from A, i.e. $F \subseteq A$. We assumed that failures are independent from each other. Thus a transaction includes exactly one failure. Let I be the set of attribute levels and L be the subset of I, i.e. $L \subseteq I$, such that $F \cap L = \emptyset$. Support is an indication of how frequently the itemset appears in the dataset. Formulation of support is shown in Eq. (1).

$$Support(F \Rightarrow L) = \frac{count(\{F \cup L\})}{N} \tag{1}$$

Accordingly, where $count(\{F \cup L\})$ counts the number of transactions including this union set. Confidence is an indication of how often the rule has been found to be true. It is the conditional probability of transactions covering set L given the occurrence of failure F. Accordingly, equation of confidence is given in Eq. (2).

$$Confidence(F \Rightarrow L) = \frac{count(\{F \cup L\})}{count(\{F\})} \tag{2}$$

4.1 COMPUTATIONAL RESULTS

In this section, we explained results of Weibull analysis and the machine learning methods. Our failure data includes 186 samples belonging to last 2 years and 7 attributes given in Table 1. According to this table, failure type is the class attribute. The temperature ranges between -6 and 36 and we discretized it by selecting 10 °C as the threshold. In Figure 3, we can easily observe this threshold point until where failures are relatively fewer while failures increase after that point. Dot-plots for other attributes are also analyzed and the thresholds are determined in the same way.

Table 1. Attributes used in machine learning

Attribute	Range	Discretization threshold and levels
C: Temperature (°C)	[-6, 36]	1: (≤ 10), 2: (>10)
FTime: Failure time	[00:00, 20:00]	1: (08:00-20:30), 2: (00:00-08:00)
#P: number of products produced until failure	[0, 2800]	1: [0-1200], 2: [1201-2800]
RType: Raw material type of the lot at failure time	[1, 16]	1: cause high number of failures, 2: cause avg. number of failures, 3: cause less number of failure
FType: Failure type		1: Electrical, 2: Mechanical, 3: Pollution

Table 2 presents the distribution of attribute levels with respect to failure types over 186 samples. According to this table level 2 of attribute **C**, level 1 of attribute **#P** and level 1 of attribute **Ftime** are frequently observed in failures.

Table 2. Frequency of attribute levels with respect to failure types

	Failure Type		
Attribute levels	Mechanical	Pollution	Electrical
Level 1 of C	8	6	20
Level 2 of C	36	**79**	37
Level 1 of FTime	41	**74**	**55**
Level 2 of FTime	3	11	2
Level 1 of #P	37	**70**	**48**
Level 2 of #P	7	15	9
Level 1 of Rtype	16	39	24
Level 2 of Rtype	6	23	14
Level 3 of Rtype	22	23	19

Table 3. Associations with respect to different support and confidence levels

Support	Confidence	Ftype	Levels			
			C	#P	Rtype	FTime
0.4	0.6	Pollution	2			
0.3	0.5	Pollution		1		
0.3	0.5	Pollution	2			
0.3	0.5	Pollution	2	1		
0.3	0.5	Pollution				1
0.3	0.5	Pollution		1		1
0.3	0.5	Pollution	2			1
0.25	0.4	Pollution	2	1		1
0.25	0.4	Electrical		1		
0.25	0.4	Electrical				1
0.2	0.35	Electrical		1		1
0.2	0.35	Mechanical				1

Table 3 shows the associations of attribute levels with failure types according to different support and confidence levels. As can be seen from table, level 2 of attribute C, level 1 of attribute #P and level 1 of Ftime are highly associated with failure types. Interestingly, attribute Rtype is not associated with failure types. Most of the associations are due to pollution which mostly appears when temperature > 10°C, number of products produced until failure ≤ 1200 and failure time 08:00-

20:30. Actually, this is reasonable since hot weather conditions affect the machine, the machine highly operates at 08:00-20:30 hours. According to these results, if a cleaning operation is performed within a predefined period, 39% ($=\frac{74}{186} * 100$) of downtimes can be eliminated. As the weather is hotter, the cleaning period could be shorter.

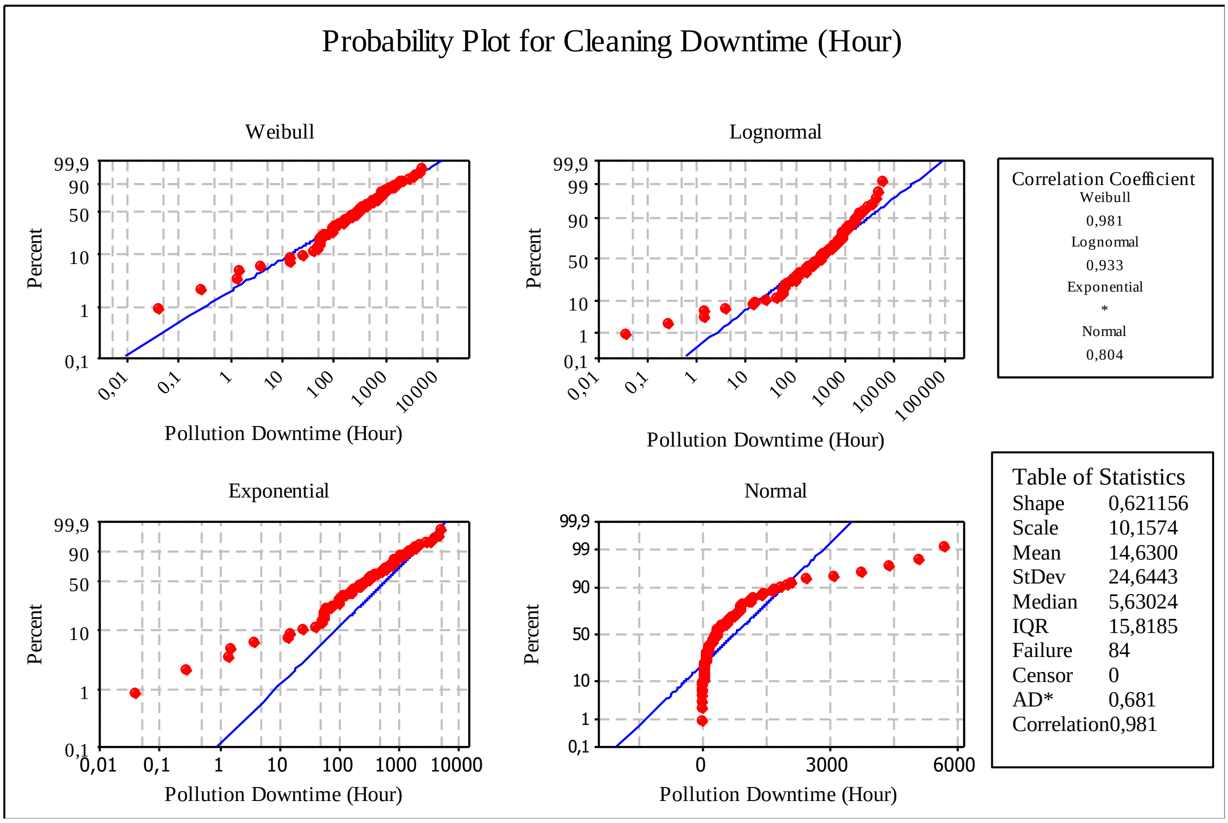


Figure 5. Probability and distribution plots of downtime due to pollution

To determine the maintenance period in Stage 4 of Figure 4, each downtime data has been assumed to be iid and analyzed in terms of Weibull, lognormal, exponential and normal distributions. Minitab result in Figure 5 shows the probability plots of pollution-based failures. As can be seen, the correlation coefficient is highest in Weibull distribution. To eliminate downtimes due to pollution, it is suggested to perform general cleaning of the machine for approximately every 15 days with 95% confidence. The same analysis was performed for electrical and mechanical failures as well. Accordingly, the lognormal distribution with mean 20 days for mechanical failures has the highest correlation coefficient while the Weibull distribution with mean 18 days for electrical failures has the highest correlation coefficient. Therefore, we can perform periodic maintenance for electrical and mechanical equipment together every three weeks.

Table 4: Accuracy of ML methods

ML Algorithms	Best	Average	Std. Dev.
SVM	0.64	0.52	0.054
NB	0.60	0.47	0.077
RF	0.54	0.43	0.049
DT	0.52	0.41	0.055

In this study, we applied ML methods such as SVM, RF, DT, Naïve Bayes (NB). The samples are divided into two sets as training (70% of samples) and test (30% of samples). Each ML algorithm was run 30 times and the results are given in Table 4. According to this table, SVM algorithm gives the best result. After that point we can use the predictions of ML algorithms to double check the maintenance period for failures. For example, if the ML algorithm predicts the next failure due to pollution, we suggest setup a minor cleaning at the end of the current shift. However, if the ML algorithm predicts electrical or mechanical failures even the next scheduled maintenance for that is not close, we suggest updating the maintenance schedule and performing it earlier.

Both the results of association rules and ML methods provided useful information to make better maintenance decisions. Although the Weibull analysis is an efficient tool of determining a maintenance period, the reality requires dynamic actions. The considered methods give the decision maker a flexibility and simultaneous action opportunity. Therefore, we presented a systematic, which is more applicable in real life.

5 CONCLUSION

In the literature, the preventive and predictive maintenance methods are distinctively applied to problems in general. In this study, we construct a decision support system using the outputs of Weibull analysis and ML methods. This integration may help to make dynamic decisions.

We observed that the Weibull distribution fits very well to our data. Therefore, we find reliable the maintenance period estimated by the Weibull analysis. Although these estimations are strong, the real life has dynamism that requires making quick actions. Thus, we modify the maintenance program built with these estimations according to the results of ML methods. We predict the next failure with SVM, NB, RF and DT classification algorithms and analyze associations in order to better understand the relationships between failure types and attribute levels. By using this systematic, the decision maker is free to modify the maintenance program. However, it is clear that as the maintenance frequency increases, the total maintenance cost also increases. In order to decide on modifying the maintenance program, a mathematical model can be used considering associations and results of ML algorithms subject to budget.

REFERENCES

[1] J. Lee, H.-A. Kao, and S. Yang, "Service innovation and smart analytics for industry 4.0 and big data environment," Procedia CIRP, vol. 16, pp. 3-8, 2014.

[2] Mobley, R. K. (2011). *Maintenance fundamentals*. Elsevier.

[3] Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, *20*(7), 1483-1510.

[4] Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23-45.

[5] Koçer, M. (2017). *CNC kesim makinesi için mükemmel olmayan önleyici bakım politikasının geliştirilmesi ve en iyilenmesi* (Master's thesis, TOBB ETÜ Fen Bilimleri Enstitüsü).

[6] Doostparast, M., Kolahan, F., & Doostparast, M. (2014). A reliability-based approach to optimize preventive maintenance scheduling for coherent systems. *Reliability Engineering & System Safety*, *126*, 98-106.

[7] Li, Z., & He, Q. (2015). Prediction of railcar remaining useful life by multiple data source fusion. IEEE Transactions on Intelligent Transportation Systems, 16(4), 2226-2235.

[8] Carvalho, T. P., Soares, F. A., Vita, R., Francisco, R. D. P., Basto, J. P., & Alcalá, S. G. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, *137*, 106024.

[9] Praveenkumar, T., Saimurugan, M., Krishnakumar, P., & Ramachandran, K. I. (2014). Fault diagnosis of automobile gearbox based on machine learning techniques. Procedia Engineering, 97, 2092-2098.

[10] Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies*, *45*, 17-26.

[11] Durbhaka, G. K., & Selvaraj, B. (2016, September). Predictive maintenance for wind turbine diagnostics using vibration signal analysis based on collaborative recommendation approach. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1839-1842). IEEE.

[12] Eke, S., Aka-Ngnui, T., Clerc, G., & Fofana, I. (2017, August). Characterization of the operating periods of a power transformer by clustering the dissolved gas data. In *2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)* (pp. 298-303). IEEE.

[13] Uhlmann, E., Pontes, R. P., Geisert, C., & Hohwieler, E. (2018). Cluster identification of sensor data for predictive maintenance in a Selective Laser Melting machine tool. *Procedia manufacturing*, *24*, 60-65.

[14] Amruthnath, N., & Gupta, T. (2018, April). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In 2018 5th International Conference on Industrial Engineering and Applications (ICIEA) (pp. 355-361). IEEE.

[15] Biswal, S., & Sabareesh, G. R. (2015, May). Design and development of a wind turbine test rig for condition monitoring studies. In *2015 International Conference on Industrial Instrumentation and Control (ICIC)* (pp. 891-896). IEEE

[16] Kolokas, N., Vafeiadis, T., Ioannidis, D., & Tzovaras, D. (2018, July). Forecasting faults of industrial equipment using machine learning classifiers. In *2018 Innovations in Intelligent Systems and Applications (INISTA)* (pp. 1-6). IEEE.

[17] Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018, July). Machine learning approach for predictive maintenance in industry 4.0. In *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)* (pp. 1-6). IEEE.

[18] Cakir, M., Guvenc, M. A., & Mistikoglu, S. (2021). The experimental application of popular machine learning algorithms on predictive maintenance and the design of IIoT based condition monitoring system. *Computers & Industrial Engineering*, *151*, 106948.

[19] Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications*, *173*, 114598.

[20] Antomarioni, S., Pisacane, O., Potena, D., Bevilacqua, M., Ciarapica, F. E., & Diamantini, C. (2019). A predictive association rule-based maintenance policy to minimize the probability of breakages: application to an oil refinery. *The International Journal of Advanced Manufacturing Technology*, *105*(9), 3661-3675.

[21] Kamsu-Foguem, B., Rigal, F., & Mauget, F. (2013). Mining association rules for the quality improvement of the production process. *Expert systems with applications*, *40*(4), 1034-1045.

[22] Djatna, T., & Alitu, I. M. (2015). An application of association rule mining in total productive maintenance strategy: an analysis and modelling in wooden door manufacturing industry. *Procedia Manufacturing*, *4*, 336-343.