International Journal of Industrial Engineering & Production Research December 2023 Vol. 34, No. 4: 1-17
DOI: 10.22068/ijiepr.34.4.1

RESEARCH PAPER

Single Machine Preemptive Scheduling Considering Energy Consumption and Predicting Machine Failures with Data Mining Approach

Ali Qorbani[1], Yousef Rabbani[2*] & Reza Kamranrad [3]

Received 13 January 2023; Revised 2 July 2023; Accepted 10 September 2023;

ABSTRACT
Prediction of unexpected incidents and energy consumption are some industry issues and problems. Single machine scheduling with preemption and considering failures has been pointed out in this study. Its aim is to minimize earliness and tardiness penalties by using job expansion or compression methods. The present study solves this problem in two parts. The first part predicts failures and obtains some rules to correct the process, and the second includes the sequence of single-machine scheduling operations. The failure time is predicted using some machine learning algorithms includes: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, and k-nearest neighbors. Results of comparing the algorithms, indicate that the decision tree algorithm outperformed other algorithms with a probability of 70% in predicting failure. In the second part, the problem is scheduled considering these failures and machine idleness in a single-machine scheduling manner to achieve an optimal sequence, minimize energy consumption, and reduce failures. The mathematical model for this problem has been presented by considering processing time, machine idleness, release time, rotational speed and torque, failure time, and machine availability after repair and maintenance. The results of the model solving, concluded that the relevant mathematical model could schedule up to 8 jobs within a reasonable time and achieve an optimal sequence, which could reduce costs, energy consumption, and failures. Moreover, it is suggested that further studies use this approach for other types of scheduling, including parallel machine scheduling and flow job shop scheduling. Metaheuristic algorithms can be used for larger dimensions.

KEYWORDS: *Data mining; Energy; Just-in-time; Machine failure; Single machine preemptive scheduling.*

1. Introduction

Problems of operations sequence are widely now used in many manufacturing systems. The most substantial decision made in scheduling problems is the assignment of considered resources and the sequence of jobs. Jobs and machines are the most significant parts of scheduling models. Most scheduling problems assume that a machine works persistently during the scheduled plan, while this assumption may be incorrect in some modes due to factors such as repair and maintenance operations, failures, or other constraints related to machine availability. Accordingly, a machine may stop working after a while owing to repair, maintenance, or failures. The study first uses machine learning algorithms to predict failure or failure time then the proposed model measures the time of repairs or between failure and machine idle until the repair time and machine availability. The model measures this time randomly using exponential distribution because this distribution occurs when equipment failure is caused by a failure in one of its pieces. Moreover, the exponential distribution function has no memory, and failure time is considered in the mathematical model at the next step to achieve more precise scheduling. Scheduling is an important and considerable issue that should

*Corresponding author: Yousef Rabbani
Rabbani@semnan.ac.ir

1. *Department of Industrial Engineering, Faculty of Engineering, Semnan University, Semnan, Iran.*
2. *Department of Industrial Engineering, Faculty of Engineering, Semnan University, Semnan, Iran.*
3. *Department of Industrial Engineering, Faculty of Mechanical Engineering, Semnan University, Semnan, Iran.*

be addressed before starting production. Some bottlenecks in the scheduling part must be solved. Single-machine operations' sequence has its complexities making it an NP-hard problem. Accidental machinery failures are one of the gaps in scheduling cases, leading to improper scheduling. Such failures increase some costs, including the cost of job acceleration and the cost of maintenance. The failures also may lead to lost sales and mistrust. Previous studies have used data mining to discover the rule and connections in different scheduling problems rather than predicting failure factors.

In contrast, the extant study used simulated data for this purpose because it could be a suitable model to implement in the manufacturing industry. [1] forecasted the machinery situation in the oil and gas industry using recurrent neural networks (RNNs). [2] studied the optimization of a two-level assembly system. They assumed that the machine is exposed to accidental failure and proposed a mathematical model to design a comprehensive plan for two-level assembly systems under stochastic times and machine failure. They also designed a preemptive maintenance plan to reduce repair and maintenance. They used Genetic Algorithm (GA) to indicate the model's effectiveness. Some studies have examined machine failure forecasts in manufacturing, optimizing inventory, repair, and maintenance. Refer to studies [3-8]. For further information. [9] studied single-machine scheduling and examined the proposed mathematical model on 144 stochastic numerical problems. Their results indicated that the proposed model could generate up to six optimal solutions. The extant study aimed at continuing this process using their mathematical model. To do this, this study designed a model to balance energy consumption, rotation speed, and torque of the machine and just-in-time delivery.

This paper has addressed the single-machine scheduling problem in a real-world case. First, past data are received, analyzed, and assessed to identify factors causing failures. Next, this study forecasts the failures using machine learning algorithms to improve the sequence of operations in real conditions and achieve the study's main goals (reduction of costs and energy consumption). Moreover, the extant study uses a mathematical model to provide machine idle time based on pre-forecasts. Moreover, we obtain some rules by using and analyzing data mining results to optimize the machining process in addition to single-machine scheduling, which is useful for the manufacturing process to minimize costs and achieve just-in-time delivery.

2. Literature Review

Numerous studies have been conducted on single-machine scheduling, and researchers attempt to test and improve all aspects of this problem to provide an optimal scheduling system and sequence of operations for these problems. Previous studies have been mentioned herein. [10] studied single-machine scheduling problems with controllable processing time with an electricity pricing approach to minimize total energy considering financial and environmental constraints. Moreover, they used a fuzzy control approach integrated with a genetic algorithm. [11] studied the problem of minimizing completion time in single-machine scheduling by considering flexible repair and maintenance and job release. They also proposed a mathematical mixed-integer programing model, heuristic algorithm, and branch and bound algorithm for this problem. [12] studied single-based production scheduling problems and maintenance planning. [13] studied single-machine scheduling with nonnegative inventory constraints and controllable processing times. They also used an exact and heuristic algorithm to solve this problem. To solve this problem, [14] examined the problem of minimizing total cost by considering just-in-time delivery. They used the Hodgson method and scheduling algorithm for this purpose. [15] considered a data mining-based approach to discover former unknown dispatch priority rules for the problem of single-machine scheduling and, after integrating a Naïve Bayesian classifier with their proposed method, suggested an optimal scheduling sequence. [16] used a classic differential evolution algorithm as the underlying framework for optimization. They analyzed job shop scheduling by using the fuzzy job-shop scheduling problem. [17] solved a single-machine scheduling problem using a multi-objective metaheuristic algorithm. In this scheduling problem, job interference exists; hence, they used a Greedy-based non-dominated sorting genetic algorithm (GNSGA-III) to solve this problem. [18] examined the single-machine scheduling problem with flexible maintenance under human resource constraints using exact and metaheuristic approaches. They used the Guided Local Search (GLS) algorithm to solve the large instance of the problem. [19] used heuristic methods to solve single-machine scheduling by consideration of periodic maintenance time. They also used a mixed integer linear programing model to solve the problem optimally. [20] used an exact solution to solve single-machine scheduling problems with periodic maintenance and sequence-dependent setup times. [21] addressed the problem of operations sequence with a single machine in which they used controllable processing times and unavailability periods to minimize job completion time or makespan. [22] investigated minimizing the total tardiness, the total energy cost, and the disruption to the original schedule in the job shop with new arrival jobs using a dual heterogeneous island parallel genetic algorithm with an event-driven strategy. [23] considered a workshop planning problem with reverse flow under uncertainty and suggested simulated annealing algorithms for small instances while proposing discrete harmony search for medium and large-sized instances. [24] proposed a new hybrid approach composed of data mining, simulation, and dispatching rules to use extracted rules for job shop scheduling in real-time and reduce makespan. In another study, [25] used a hybrid approach composed of dispatching rules, data mining, GA, and simulation to solve job-shop scheduling problems. [26] addressed the flexible job shop scheduling problem by considering release time to minimize total weighted tardiness using dispatching rules. They also used a random forest algorithm to select the best-extracted rule among the rules proposed for scheduling.

Studies on the effects of maintenance and repair types on machine scheduling have received great attention. The implementation of support decision systems allows organizations to minimize service costs, maximize job time, and improve productivity [27]. Common machine learning methods have been proposed for preemptive maintenance and repair problems [28-30]. Machine learning methods consist of using classifiers such as support vector machines, decision trees, random forests, and Naive Bayes. Moreover, [31] widely used the periodic repair and maintenance approach. [32] studied the problem of integrated production planning and preemptive maintenance for parallel machine scheduling to minimize machine unavailability and production time simultaneously. They also used a multiobjective non-dominated sorting genetic algorithm to solve this problem.

According to the analysis of results obtained from previous studies, the extant study supplemented previous studies. Hence, this study addressed failures and failures in single-machine scheduling. Moreover, this study examined the effect of these failures on the sequence of operations. This paper also aimed at reducing energy consumption by considering the rotation speed and torque of the machine. The rules can be achieved through data mining. In the proposed model, the compression and expansion time of jobs were linked to rotation speed and torque, so this process affects energy consumption because reduced rotation speed may lead to a considerable decline in energy consumption [33]. For instance, job acceleration requires increased rotation speed, which leads to higher energy consumption and failure. This study attempts to optimes this process by applying available constraints in the model. Considering failures has been one of the critical challenges that managers face. Hence, the extant study tends to overcome this challenge to provide a roadmap for others and develop in solving optimization problems.

3. Theoretical Foundations

3.1. Single-machine scheduling

Single-machine scheduling or single-resource scheduling is an optimization problem that aims at an optimal operation sequence and minimizing some objectives, such as cost and job lateness. The solution methods of single-machine scheduling are used as a pattern to solve other models. Various scheduling models have been introduced for single-machine problems, including the weighted sum of completion time, maximum lateness, number of late jobs, complete lateness, and total weighted lateness. Some sudden failures cause variations in single-machine scheduling. Product delivery in due time is one of the main objectives that this study pursues. There are some random failures in the industry and production process; therefore, these failures and disorders lead to variations in machine scheduling. The failures usually include machine failure, order cancelation, tool failure, unexpected accidents, and power outages These failures may stem from new arrival jobs and rework. The extant study mostly examines machine and tool failures during single-machine scheduling. The mentioned failures highly affect the sequence of jobs. Deciding how to plan jobs and time is a basic issue in single-machine scheduling.

3.2. Data mining

Data mining collects important and practical tools to reduce anomalies and detect or predict failures in production systems. Data mining is the discovery of interesting, unexpected, or valuable structures in large data sets [34]. Various algorithms are used in data mining based on the problem type. The methods used in data mining are closely linked to machine learning algorithms. Machine learning algorithms are placed in different categories, including regression, classification, and clustering, and are used in different contexts [35]. In terms of predictor maintenance and repair, machine learning has become such an important case that an increasing number of papers have used machine learning for this problem because of the interpretability of machine learning models [36]. Maintenance and repair decision support systems have been boosted by the Internet of Things, big data, and machine learning. These systems play a vital role in ensuring the maintenance and reliability of equipment in industries by converting large data sets to useful knowledge and consequences [37-39]. One can use classification and modeling approaches to design an instant and computational model for estimating failure probability, predicting future failure time, predicting failure rate, and so forth. Classification, clustering, and rule-based methods can make the proper decision. Nonlinear regression-based approaches can be used to estimate continuous quantities in the system. The serial approach is used to predict machine failure based on robust algorithms. Classification algorithms are supervised algorithms that tend to achieve an accurate forecast. The present paper used a decision tree classifier algorithm, which generates simple and comprehensible results and good rules.

3.3. Decision tree algorithm

The decision tree algorithm is one of the widely used data mining algorithms. The decision tree is a predictor model used for regression and classification models in data mining. When the tree is used for classification tasks, it is known as a classification tree. In the decision tree structure, the prediction obtained from the tree is explained in the frame of some rules. Each route from the root to the leaf of the decision tree expresses a rule, and the leaf is finally labeled based on the class with the highest record. One plots a tree graph when a problem is solved based on this algorithm. In doing so, the middle nodes are question nodes of its leaf representing the classification. The problem considered in the present study aims at predicting machine failure. An initial classification should be done to create the model. The dataset should be labeled because this algorithm is a supervised method. The decision tree is created by selecting a feature or column with the highest entropy. First, a feature is found to do classification precisely by using this feature. Second, the first step is repeated by finding a feature with the highest entropy for the dataset. The process continues in the dataset until there is no feature with the highest good entropy, no new feature, or all features are reviewed. Moreover, Gini and entropy values are measured by the equations below, where P_i indicates the probability of the relevant class.

$$\text{Gain(X,A)= entropy(X) - entropy(X,A)} \quad (1)$$

$$\text{entropy(X)}= -\sum_{i=1}^{n} P_i log_2(P_i) \quad (2)$$

$$\text{entropy(X,A)} = \sum_{i=0}^{m} \frac{|X_i|}{|X|} entropy(X_i) \quad (3)$$

$$\text{Gain-Rate(X,A)} = \frac{gain(X,A)}{entropy(X,A)} \quad (4)$$

3.4. Dataset

One important step in data mining is inserting and preprocessing data. This dataset included information about numerical, serial, and historical features (1000 observations). This study used 80% of training and 20% of test data. In these observations, failures are shown as binary values (0,1), in which 1 represents breakdown while 0 indicates a lack of failures. Moreover, five features (air temperature, process temperature, rotation speed, torque, and tool wear) exist, and their scatter plots have been depicted in Figures 1-5.

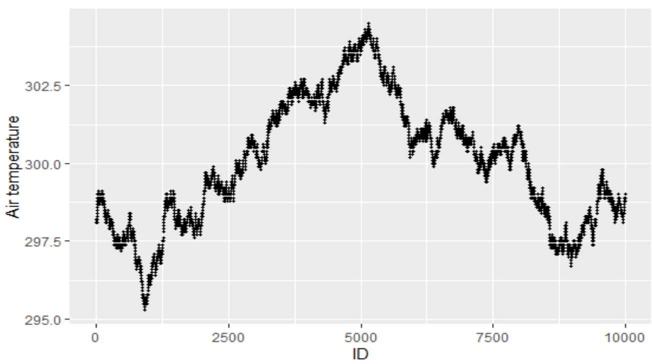

Fig. 1. Scatter plot of air temperature

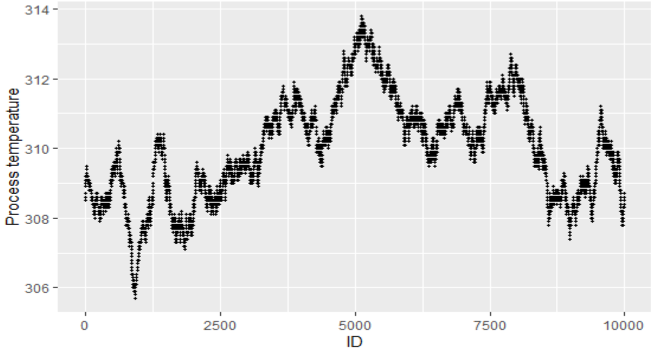

Fig. 2. Scatter plot of process temperature

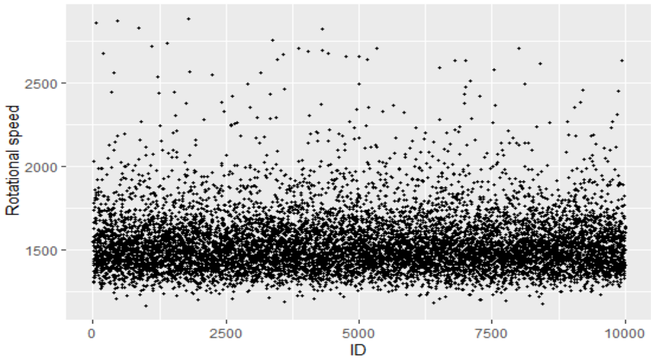

Fig. 3. Scatter plot of rotation speed

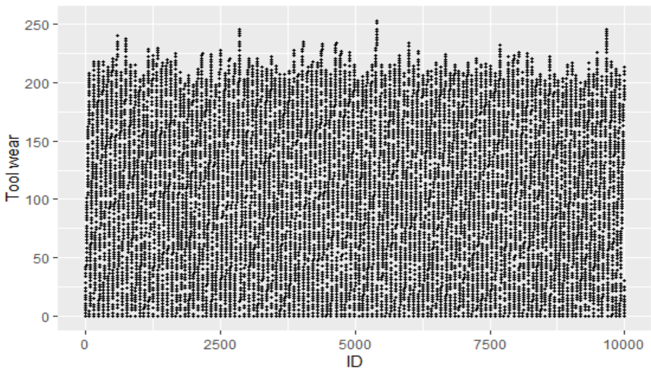

Fig. 4. Scatter plot of tools wear

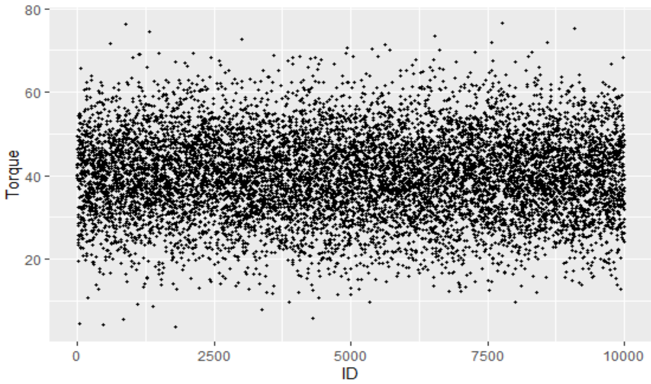

Fig. 5. Scatter plot of torque

In this dataset, machine failure consists of four independent modes. The first mode is a failure caused by tools and pieces. If the parameter of tool wear varies between 200 and 240min, the devices will most likely be damaged, and the tool must be replaced. If the difference between air temperature and process temperature is lower than 6.8 Kelvin and the rotation speed of tools is less than 1380 rpm, the machine will fail, called heat loss-cased failure. Power outage is another failure factor in which the process will fail if the result of torque multiplied by rotation speed does not equal the power required for the process. Accordingly, the machine will fail and runs outs of power if the power is less than 3500W or higher than 9000W. Overpressure is another parameter causing failure. Accordingly, the process will fail if the sum of tool wear and torque parameters exceeds 11000 N m/min. All the abovementioned factors cause machine failure by adjusting the machine failure label on 1. Table 1 reports three rows of the problem's dataset.

Tab. 1. Problem’s dataset

Air temperature	Process temperature	Rotational speed	torque	Tools Wear	failure
298/8	309/2	١٤٢٥	53/9	135	٠
298/8	309/2	١٤١٢	44/1	١٤٠	٠
298/8	309/1	٢٨٦١	4/6	١٤٣	١

3.5. Classification of failures

This analysis used six classifier algorithms to assess the accuracy of results and compare them: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, and k-nearest neighbors. Table 2 reports the classification results. The algorithms were run through Python software, Sklearn Library, and a system with 8Gb Ram and Core i7 CPU.

Tab. 2. Results of classification

Model name	Accuracy (Percentage)	Time (seconds)	time/accuracy
KNN	٩٧	٠,١١٨٩	٢٨١,٢٣٩
logistic regression	٩٧,٣	٠,٦٥	١٢٠,٢٠٥
decision tree	٩٧,٩٥	٠,٢٦	١٩٠,٢٩١
Random forest	٩٦,٩	١,٣٥	٨٣,٢٣٩
Support vector machine	٩٦,٩٥	٧,٠٨	٣٦,٤١٣
Naïve Bayes	٩٦,١	٠,٠٣	٥٢٣,١٨٥

On the other hand, the decision tree algorithm outperformed other algorithms in forecasting failures. As shown in Table 2, implementing this algorithm requires a longer time than other algorithms. The decision tree algorithm could accurately forecast 69.8% of machine failures and accurately predicted the lack of machine failures in 98.86% of cases. The mentioned rates are greater than the results obtained by other algorithms. Table 3 reports the accurate forecast percentage of failures.

Indicators reported in Tables 2, and 3 assess the efficiency of classification algorithms, while an integrated index (accuracy and precision) is used in the operating feature plot shown as a curve (Figure 6).

Tab. 3. prediction results

Model name	The percentage of correct predictions of no failures	Incorrect percentage of predicted failures	Correct percentage of failure prediction
K nearest neighbor	۹۹/۱	۸۴.۳	۲۵/۷
logistic regression	۹۹/۵۹	۷۳/۰۲	۲۶/۹۸
decision tree	۹۸/۸۶	۳۰/۱۶	۶۹/۸۴
Random forest	۱۰۰	۹۸/۴۱	۱/۵۹
Support vector machine	۱۰۰	۹۶/۸۳	۳/۱۷
Naïve Bayes	۹۸/۶۱	۸۰/۹۵	۱۹/۰۵

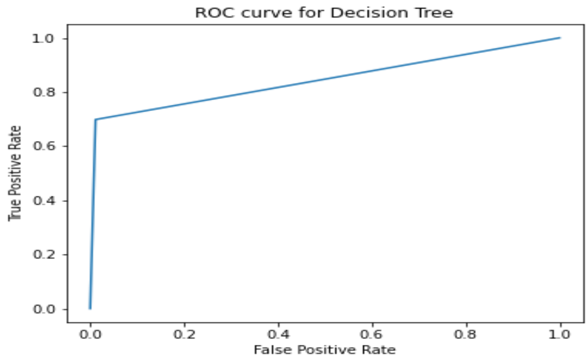

Fig. 6. Operating feature of decision tree classification algorithm

The numerical value of AUROC varies between 0 and 1, indicating the detection power or accuracy of the results of a test. Test results' accuracy depends on the test's ability to indicate the difference between integer positive and negative results. If this value is near 1, the data are usually placed above the bisector line, the positive integer rate is high, and the test method has an appropriate detection power or accuracy. The relevant model indicated a high assignment rate of this model. Moreover, five variables were used to set the decision tree. The end nodes or leaves of the tree equaled 14, and the mean deviation of residual and error rate of classification equaled 1.1012 and 0.01638, respectively. Moreover, Figure 7 depicts a graphical schematic of the tree.

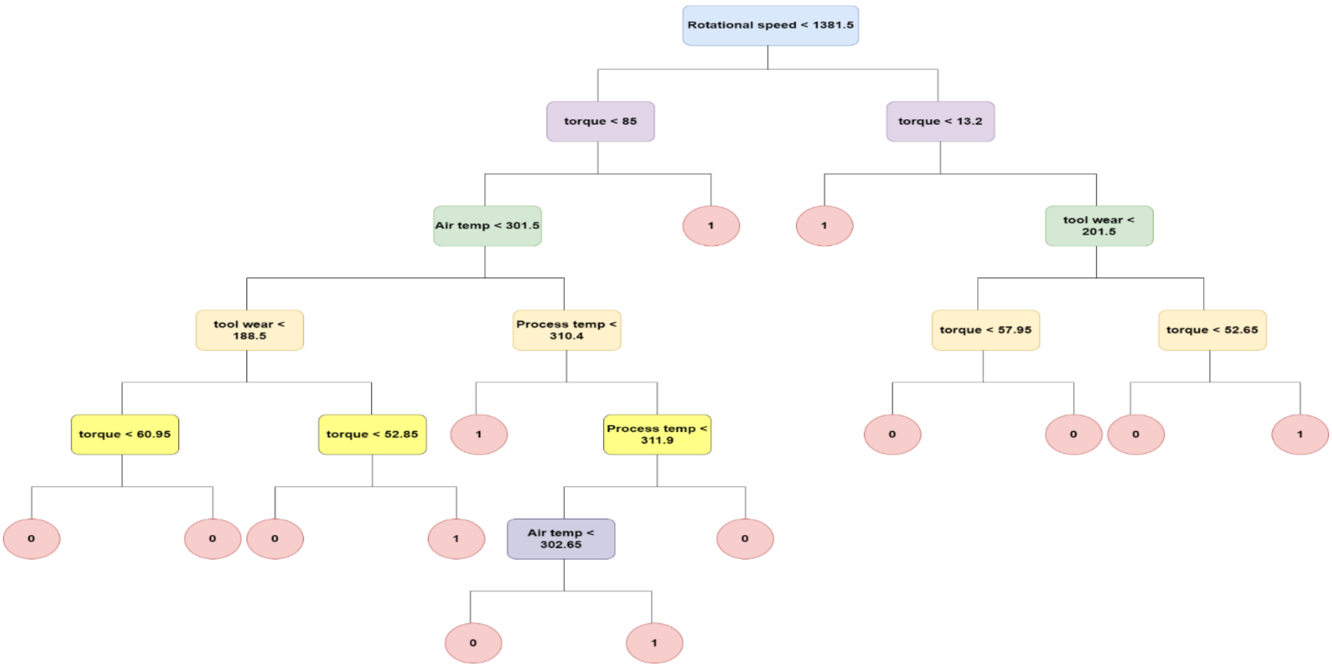

Fig.7. Graphical schematic of classifier decision tree

For example, the rules from the decisions tree algorithm indicate that the machine will not likely fail if rotation speed, tool wear, and torque values are less than 1381.5, 201.5, and 57.95, respectively.

4. Problem Definition

This part has formulated a nonlinear mathematical model for a single-machine scheduling problem in which job interruption, release time, and machine idle time are allowed. First, decision parameters and variables are defined then a mathematical model is designed for single-machine scheduling, considering failures to minimize the sum of penalties. In this model, P_i Indicates the normal processing time of the job with no extra processing cost, $\mathrm{d_i}$ represents the due date of the product, B_i Indicates lateness penalty through the time when the job is completed after the due date, and α_{i} represents early delivery when the job I completed sooner than due time. If a one-unit increase or decrease occurs in processing time (this increase or decrease is matched with the rotation speed and torque of the machine), the compression (c_i) or expansion ($c_i^{'}$) costs are increased or decreased. Each job can be compressed or expanded to the maximum extent of allowed decline or enhancement. Moreover, initial rotation speed and torque are parameters and have a constant value. The aim is to determine the sequence of jobs and the optimal amount of job processing time compression or expansion in each machine simultaneously to minimize earliness or tardiness penalties and the cost of compressing to expanding jobs. It is also aimed at minimizing energy consumption and machine failure. This part used former analyses about machine failure forecasts using data mining to consider model disruptions. Data mining is used to achieve rules about the rotation speed or torque causing failure and machine unavailability or availability times using forecast results in the mathematical model. This model measures optimal rotation speed and torque variables by considering job compression and expansion times separately. Therefore, energy consumption can be calculated and optimized considering the abovementioned variables. Using an exponential distribution, this model randomly measures the maintenance time or the duration between machine failure and setup. This distribution is used because it occurs when a failure in its pieces causes equipment failure.

On the other hand, the exponential distribution function has no memory. Moreover, the time required for repairing or servicing devices or pieces is called system maintainability time. As the machine failure time is a stochastic variable, maintainability time can be considered a random variable because it does not take any definite value. Assume that maintenance time follows a negative exponential density function; in this case, the maintenance rate remains constant, and its density function is obtained through equation (5).

$$f(x) = \lambda e^{-\lambda x} \qquad x \geq 0 \tag{5}$$

where $\boldsymbol{\lambda}$ indicates the average number of maintenance in the time unit and $\frac{1}{\lambda}$ indicates the average maintenance time. Maximum Likelihood Estimation (MLE) was used to calculate λ and achieve an estimated value of λ. The following table reports the number of machine failures in the system. The exponential distribution is fitted to this data based on the assumption of independent failures to estimate their λ values. Table 4 reports the number of failures per period.

Tab. 4. Number of failures per period

period number	Number of failures	period number	Number of failures
١	١٣	١٤	٢٤
٢	٥	١٥	٨
٣	٩	١٦	٧
٤	١٣	١٧	٩
٥	١٠	١٨	٩
٦	٦	١٩	٧
٧	٩	٢٠	٦
٨	٨	٢١	٩
٩	٦	٢٢	٩
١٠	٨	٢٣	٨
١١	١٥	٢٤	٦
١٢	٥٤	٢٥	٧
١٣	٥٧	٢٦	١٧

Table 5 indicates the results of the MLE method.

Tab. 5. MLE results

Estimated amount	12.22713
Gradient	٧,٤٣٨٣٣٢ $\times 10^{-8}$
Number of iterations to solve	١١

Therefore, Mean Time to Repair (MTTR) is measured as follows:

$$MTTR=\frac{1}{\lambda} \tag{6}$$

Therefore, the MTTR contributes to a more precise scheduling calculation in a mathematical model to estimate machine unavailability times. Sets and indexes, parameters, decision variables, and relations of the mathematical model have been defined in the following tables.

Tab. 6. Sets and indexes

N	Number of jobs
M	Number of situations
i	Jobs' index
k	Situation's index

Tab. 7. Parameters

p_i	Normal processing time
r_i	Release time
α_i	Earliness penalty
β_i	Tardiness penalty
F_i	Cost of job compression
$F_i^{'}$	Cost of job expansion
d_i	Job due time
L_k	Maximum compression time related to each situation
L_k'	Maximum expansion time related to each situation
RS_i	Rotation speed of the machine in the job
TO_i	Machin torque in job i
V	Fixed cost per kWh
M	Positive large number
SPM_k	Start time of machine unavailability in situation k
FPM_k	End of machine unavailability in situation k

Tab. 8. Decision variables

C_i	Completion time of piece i
y_{ik}	Equals 1 if the piece "I" is scheduled in the kth situation, 0, otherwise
E_i	Earliness time
T_i	Tardiness time
X_{ik}	Interruption duration of the kth situation of the piece i
S_i	Start time
H_k	Compression time in each platform
H_k'	Expansion time in each platform
Q_i	Compression time of each job
Q_i'	The expansion time of each job
$s_i^{"}, s_i'$	Auxiliary variable for calculating processing time
c_k	Completion time of peace in situation k
H_k	Compression time in each platform

4.1. Mathematical model

Z=min $\left(\sum_{i=1}^{N}(\alpha_i E_i + \beta_i T_i + F_i Q_i + F_i^{'} Q_i^{'})\right) + (V * PM_i)$ ۱

$$\sum_{i=1}^{n} Y_{ik} \leq 1 \qquad K = 1,2,\dots,m$$ ۲

$$Y_{ik} \leq X_{ik} \leq p_i . Y_{ik} \qquad i = 1,2,\dots,n \;\; k = 1,2,\dots,m$$ ۳

$$\sum_{k=1}^{m} X_{ik} = p_i \qquad i = 1,2, \dots, n$$ ٤

$$c_k \geq c_{k-1} + \sum_{i=1}^{n} X_{ik} - H_k + H'_k + FPM_k - SPM_k \qquad k = 2,3, \dots, K$$ 5

$$c_k \geq \sum_{i=1}^{n} X_{ik} + Y_{ik} . r_i \qquad k = 1, \dots, K$$ 6

$$C_i = max_{k=1}^{m} (Y_{ik} . c_k) \qquad i = 1,2, \dots., n$$ 7

$$S'_i = min_{k=2}^{m} (c_{k-1} . M(1 - Y_{ik}) \qquad i = 1,2, \dots., n$$ 8

$$S''_i = r_i + M(1\_Y_{ik}) \qquad i = 1,2, \dots., n \ , k = 1$$ 9

$$S_i = min(S'_i, S''_i) \qquad i = 1,2, \dots., n$$ 10

$$\sum Y_{ik} . H_k = Q_i \qquad i = 1,2, \dots, n \quad k = 1,2, \dots, K$$ 13

$$\sum Y_{ik} . H'_k = Q'_i \qquad i = 1,2, \dots, n \quad k = 1,2, \dots, K$$ 14

$$T_i \geq C_i - d_i \qquad i = 1,2, \dots, n$$ 15

$$E_i \geq d_i - C_i \qquad i = 1,2, \dots, n$$ 16

$$L_k \geq H_k$$ 17

$$L'_k \geq H'_k$$ 18

$$RS'_i = \frac{RS_i \times P_i}{P_i - Q_i + Q'_i}$$ 19

$$TO'_i = \frac{TO_i \times P_i}{P_i + Q_i - Q'_i}$$ 20

$$PM_i = \frac{2 * \pi * RS'_i * TO'_i}{60} * P_i + Q_i - Q'_i \qquad i = 1,2, \dots, N$$ 21

$$LRS'_i \leq RS'_i \leq URS'_i$$ 22

$$LTO'_i \leq TO'_i \leq UTO'_i$$ 23

$$T_i \geq 0$$ 24

$$E_i \geq 0$$ 25

$$y_{ik} \in 0,1$$ 26

The objective function includes five parts: earliness penalty, tardiness penalty, compression cost of jobs, expansion cost of jobs, and energy consumption cost. Constraint (2) requires that only one interrupted part of a job is operatable in each situation. Constraint (3) ensures that the length of each interrupted part of a job assigned in the sequence is greater than the time unit and equals or is less than its processing time. This constraint guarantees that the length of each interrupted part is a continuous variable between its upper and lower bounds. Constraint (4) ensures that the length of all interrupted parts of a job equals its processing time. Constraints (5) and (6) measure the completion time of situation K considering the release time of jobs. If job interruption is allowed, the job can be interrupted and continued at another time. Therefore, a job can be interrupted several times and consist of repair and maintenance time. Constraint (7) calculates the completion time of each job by determining the completion time of the last interrupted part of each job. Moreover, constraint (8) determined the start time of the first interrupted part of each job from situation 2 to k. Constraint (9) calculates the possible start time of the interrupted part of each job in the first situation. Finally, constraint (10) indicates the real start time of each job based on the computation results of constraints (8) and (9). Constraints (13) and (14) indicate the expansion and compression extent of each job. Constraints (15) and (16) calculate earliness and tardiness costs, respectively. Constraints (17) and (18) indicate the compression and expansion extent in each platform of jobs. Constraints (19) and (20) indicate optimal rotation speed and torque. Constraint (21) measures the energy consumed for each job. As mentioned, earliness cost is linked to start time while tardiness costs depend on job completion. Constraints (22) and (23) report upper and lower limits related to the rotation speed and torque of the machine, which were obtained from the data mining results. Constraints (24), (25), and (26) provide logical binary and nonnegative requirements for decision variables.

5. Experiment

This part of the study addressed the mathematical model to explain this approach more. The extant study decided to use the penalties caused by early or late delivery for the idle time between job setups according to which we will have relations 7 and 8.

$$E_i = \max\,(0, d_i - P_i - S_i) \qquad (7)$$
$$T_i = \max\,(0, C_i - d_i\,) \qquad (8)$$

This model addressed a new problem: job processing times can be divided. In other words, each job may be interrupted alone in several parts or processes, and the length of each interrupted part can be quantified within a certain interval. This method schedules machines and jobs after predicting failures. We use example 1 to clarify the considered approach. Table 9 reports the values of the parameters. We also used Figure 8 to illustrate the graphic of scheduling.

Tab. 9. Parameters' values in example 1

job	p_i	r_i	F_i	F'_i	α_i	β_i	d_i	RS_i	TO_i	SPM_k	FPM_k
١	2	3	2	2	1	2	5	1050	35	1	2
٢	8	1	2	2	3	2	10	1100	38	0	0
٣	6	10	1	2	2	2	١۵	1000	30	0	0
										10	11

Fig. 8. Scheduling three jobs and four situations

As seen, release times of jobs equaled 3, 1, and 10, and no job was available in early time (0). Two times have been assigned to repair and maintenance in this process. Moreover, job 2 has been divided into two parts, and jobs are done in due time, so no earliness or tardiness does not exist because job 2 and job 3 have been compressed into 2 and 3 units, respectively.

5.1. Effect of fixed parameter

In example 1, all parameters remain fixed except for processing time. In addition, compression or expansion time of jobs was not considered. In contrast, job compression and expansion time was considered in example 2. The value of the objective function in example 1 equaled 11 without considering energy cost. In example 2, objective function equaled 11, which 3 units of penalties raised from the cost of compressing and expanding jobs. Therefore, example 2 indicates that the penalty caused by earliness and tardiness declines by three units compared to the former case. The cost of energy consumption in example 2 reached from 2816000 to 2517300 compared to example 1. Figures 9 and 10 illustrate the scheduling of these two examples graphically. Tables 10 and 11 report the values of parameters.

Tab. 10. Parameters' values in example 1

۵job	p_i	r_i	F_i	F'_i	α_i	β_i	d_i	RS_i	TO_i
١	١	١	١	١	١	١	٢۶	١١٠٠	۴٠
٢	٢	١	١	١	١	١	٢۶	١١٠٠	۴٠
٣	٩	١	١	١	١	١	٢۶	١١٠٠	۴٠
4	۵	١	١	١	١	١	٢۶	١١٠٠	۴٠

۵	۹	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰

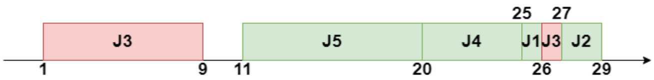

Fig. 9. Scheduling five jobs for example1

Tab. 11. Parameters' values in example 2

۵job	p_i	r_i	F_i	F'_i	α_i	β_i	d_i	RS_i	TO_i
۱	۱	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰
۲	۲	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰
۳	۹	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰
4	۵	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰
۵	۹	۱	۱	۱	۱	۱	۲۶	۱۱۰۰	۴۰

Fig. 10. Scheduling five jobs for example2

5.2. Effect of limited rotation speed and torque

In example 3, some limitations were considered for the model based on the data mining rules. Accordingly, rotation speed and torque cannot vary beyond their upper and lower limits, which was considered in Constraints 22 and 23 of the mathematical model. Figure 11 depicts the scheduling of this problem under the limited rotation speed and torque, and Table 12 reports the values of its parameters.

Tab. 12. Parameters' values in example 3

۳job	p_i	r_i	F_i	F'_i	α_i	β_i	d_i	RS_i	TO_i
۱	۲	۳	۲	۲	۲	۲	۵	۱۲۵۰	۳۸
۲	۸	۱	۲	۲	۲	۲	۱۰	۱۲۵۰	۴۲
۳	۶	۱۰	۲	۲	۲	۲	۱۵	۱۲۵۰	۵۰

Fig. 11. Scheduling under limited rotation speed and torque

The machine's rotational speed must be increased while machine torque must be reduced to accelerate jobs. According to constraints imposed in the model, the rotation speed of the machine cannot exceed 1380. Hence, jobs will be done delay. As seen in Figure 11, jobs 1 and 2 are delivered to customers with one-unit tardiness, while job 3 is done with two-unit tardiness.

5.3. Effect of unlimited rotation speed and torque

Unlike the previous example, no limitation is applied to rotation speed and torque, so optimal rotation and optimal torque can take any quantity. As is seen, Figure 12 depicts the scheduling of this problem under the unlimited rotation speed and torque, and Table 13 reports the values of its parameters. According to the scheduling results, the machine can accelerate jobs by increasing rotation speed, but the excessive rise in this speed leads to the failure of the machine and tools. As shown in example 4, one-unit acceleration has been created in job 2, and its processing time has been reduced from 8 to 7 units. The torque of this job has reached 42 to 38, and its rotation speed has increased from 1300 to 1486, which may cause machine failure. The mentioned failures lead to financial damages to the whole system. In general, it is concluded that excessive increases in the rotation speed of the model and failures caused by an extra rise or decline can be prevented by considering data mining rules. Moreover, energy consumption will be optimized in this way. Moreover, these constraints reduce the solution time of the problem rather than the problem with no constraint.

Fig. 12. Scheduling under unlimited rotation speed and torque

6. Results

In this research, coefficients of earliness and tardiness penalties are stochastic, and the cost of compressing and expanding jobs is considered randomly. The SPM parameter represents the start time of machine unavailability, which is obtained by the conducted forecast. FPM parameter is estimated by exponential distribution after required repairs or replacing tools to find when the machine is available. Processing time and job due time have been considered as presumptions and also the due date is obtained from equation 9:

$$d(1 + (i \times v)) \tag{9}$$

results of a mathematical model with a size of 2-6 jobs have been compared between two modes of applying and not applying job compression and expansion constraints. Table 13 reports the obtained values. Mode A indicates that the model is allowed to apply the constraints related to upper and lower limits of rotation speed and torque. Mode B indicates that the model cannot apply the constraints related to upper and lower limits of rotation speed and torque. This model was run through GAMS software with 8GB Ram and Core i7 CPU.

Tab. 13. Results

			A				B			
N	v	i	obj	Time(sec)	gap	Failure	obj	Time(sec)	gap	failure
3	0.05	1	15004.1	1.3	9.72	0	15003.68	8.52	9.97	1
		2	15003.3	0.73	9.89	0	15003.2	3.53	9.95	1
		3	15048.4	1.64	9.97	0	15000.8	4.66	9.92	1
	0.06	1	15046.6	0.72	9.99	0	15001.3	4.27	9.88	1
		2	15030.1	0.78	9	0	15001	4.64	9.96	1
		3	15031.1	0.81	9.88	0	15000.6	4.15	9.99	1
	0.07	1	15003.7	0.77	9.85	0	15001.3	3.6	9.88	1
		2	15048.2	0.95	0.09	0	15000.8	5.71	9.95	1
		3	15000.6	0.86	9.74	0	15000.4	2.68	9.98	1
Mean			15024	0.95	8.68	0	15001.4	4.64	9.94	1
4	0.05	1	22854.8	0.2	9.32	0	22889	39.9	9.5	1
		2	22825.1	0.37	9.18	0	22882	259.3	9.99	0
		3	22880	0.33	8.97	0	22885	29.78	9.5	1
	0.06	1	22882.2	0.31	8.98	0	22884	89.9	10	0
		2	22881	0.28	8.97	0	22907	38.89	10	0
		3	22880	0.2	8.97	0	22881	90.8	10	0
	0.07	1	22943	0.22	9.56	0	22887	44.8	9.95	1
		2	23025.8	0.45	9.95	0	22885	30.22	9.96	1
		3	22908	0.28	9.39	0	22881	34.51	9.92	0
Mean			22897.7	0.29	9.25	0	22886.7	69.8	9.86	0.44
6	0.05	1	36392.8	1.62	9.15	0	36410.9	1012	19.62	2
		2	36367	2.37	8.72	0	36349.2	1031	17.8	1
		3	36238	0.31	8.61	0	36378.8	1000	19.48	3
	0.06	1	36286.4	1.08	9.48	0	36322.2	1013	20.32	3
		2	36367.4	0.75	9.08	0	36272.7	1014	17.2	1
		3	36374	1	9.1	0	36433	1014	20.4	1
	0.07	1	36386	0.67	8.84	0	36310	1039	18.21	2
		2	36396	2.36	8.28	0	36327	1034	20.37	2
		3	36284	0.36	8.83	0	36266.8	1033	20	1
Mean			36343.5	1.16	8.89	0	36341.1	1021.1	19.26	1.77

The mean objective functions for the results of the solution by GAMS for states A and B equal 24755 and 24743, respectively, and the standard deviation of each equals 8970 and 8978, respectively, which indicates that the mean is not equal. Also, the 95% confidence interval for the mean of the objective function is equal to (-4889,4913), which includes the value of zero, and the p-value is equal to 0.996, which indicates the acceptance of the null hypothesis that the means are equal. to further compare the results of the hypothesis test, we compare the gap averages obtained through the A and B approaches, and statistical results are summarized in Table 14.

Tab. 14. Descriptive statistics

Sample	N	Mean	StDev	SE Mean
gap_A	27	8.94	1.83	0.35
gap_B	27	13.03	4.55	0.88

As it is known, the average gap of B mode is higher than that of A mode and the 95% confidence interval of the test is equal to (-5, 975, -2, 187) which does not include the zero value, and also the p-value equals 0. Therefore, it can be concluded that the null hypothesis, which means equality of means, is rejected and the average value of the gap in the B state is higher. Therefore, it is concluded that the performance of the model is better in terms of the gap. Also, the average solution time and the number of possible failures are higher in B mode and it can be decided that the proposed approach reduces the solution time and that the number of possible failures due to the rotational speed and torque parameters is reduced. Also, in Figure 13, a line diagram is used to compare the gap between the A and B modes.

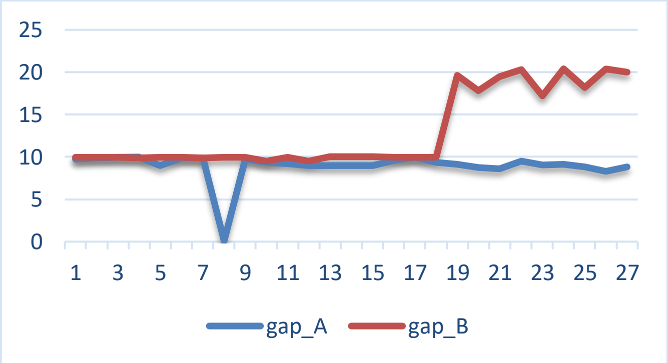

Fig. 13. Comparison of the gap between A and B

7. Conclusion

This study aims at achieving some rules to predict machine failure using a decision tree algorithm. Figure 8 indicates that the decision tree algorithm outperformed other algorithms with a probability of 70% in predicting failure. Then, the failure time was applied to the mathematical model to achieve an optimal sequence of operations. The extant study attempted to address operation sequence, optimization of machine operation, and energy consumption continuously in an integrated problem. It was concluded that the relevant mathematical model could schedule 6 jobs within a reasonable time and achieve an optimal sequence, which could reduce costs, energy consumption, and failures. Moreover, this study suggests further studies use this approach for other types of scheduling, including parallel machine scheduling and flow job shop scheduling. Meta-heuristic algorithms can be used for larger dimensions.

References

[1] Zainuddin, Z., P.A. EA, and M. Hasan, *Predicting machine failure using recurrent neural network-gated recurrent unit (RNN-GRU) through time series data*. Bulletin of Electrical Engineering and Informatics, Vol. 10, No. 2, (2021), pp. 870-878.

[2] Guiras, Z., et al ,.*Optimal maintenance plan for two-level assembly system and risk study of machine failure*. International Journal of Production Research, Vol. 57, No. 8, (2019), pp. 2446-2463.

[3] Mokhtari, H. and M. Dadgar, *Scheduling optimization of a stochastic flexible job-shop system with time-varying machine failure rate*. Computers & Operations Research, Vol. 61, (2015), pp. 31-45.

[4] Riazi, M., et al. *Detecting the onset of machine failure using anomaly detection*

methods. in *International Conference on Big Data Analytics and Knowledge Discovery*. (2019).

[5] Paprocka, I., *Evaluation of the effects of a machine failure on the robustness of a job shop system—Proactive approaches.* Sustainability, Vol. 11, No. 1, (2019), p. 65.

[6] Wang, Z., C.K. Pang, and T.S. Ng, *Robust scheduling optimization for flexible manufacturing systems with replenishment under uncertain machine failure disruptions.* Control Engineering Practice, Vol. 92, (2019), p. 104094.

[7] SobASzek, Ł., A. GolA, and A. Świć, *Time-based machine failure prediction in multi-machine manufacturing systems.* Eksploatacja i Niezawodność, Vol. 22, No. 1, (2020).

[8] Smadi, H.J. and A.K. Kamrani, *PRODUCT QUALITY-BASED METHODOLOGY FOR MACHINE FAILURE ANALYSIS AND PREDICTION.* International Journal of Industrial Engineering, Vol. 18, No. ١1, (2011).

[9] Shokoufi, K. and J. Rezaeian, *An exact solution approach using a novel concept for single machine preemptive scheduling problem in the just-in-time production system.* Journal of Industrial and Production Engineering, Vol. 37, No. 5, (2020), pp. 215-228.

[10] Tsao, Y.-C., V.-V. Thanh, and F-.J. Hwang, *Energy-efficient single-machine scheduling problem with controllable job processing times under differential electricity pricing.* Resources, Conservation and Recycling, Vol. 161, (2020), p. 104902.

[11] Cui, W.-W. and Z. Lu, *Minimizing the makespan on a single machine with flexible maintenances and jobs' release dates.* Computers & Operations Research, Vol. 80, (2017), pp. 11-22.

[12] Liu, Q., M. Dong, and F. Chen, *Single-machine-based joint optimization of predictive maintenance planning and production scheduling.* Robotics and Computer-Integrated Manufacturing, Vol. 51, (2018), pp. 238-247.

[13] Zhou, B. and T. Peng, *New single machine scheduling with nonnegative inventory constraints and discretely controllable processing times.* Optimization Letters, Vol. 13, No. 5, (2019), pp. 1.١١٤٢-١١١

[14] Varela, M.L., et al., *Collaborative paradigm for single-machine scheduling under just-in-time principles: total holding-tardiness cost problem.* Management and Production Engineering Review, Vol. 9, No. 1, (2018).

[15] Premalatha, S. and N. Baskar, *Implementation of supervised statistical data mining algorithm for single machine scheduling.* Journal of Advances in Management Research, (2012).

[16] Gao, D., G.-G. Wang, and W. Pedrycz, *Solving fuzzy job-shop scheduling problem using DE algorithm improved by a selection mechanism.* IEEE Transactions on Fuzzy Systems, Vol. 28, No. 12, (2020), pp. 3265-3275.

[17] Cheng, C.-Y., et al., *Greedy-based non-dominated sorting genetic algorithm III for optimizing single-machine scheduling problem with interfering jobs.* IEEE Access, Vol. 8, (20٢٠), pp. 142543-142556.

[18] Touat, M., F.B.-S. Tayeb, and B. Benhamou, *Exact and metaheuristic approaches for the single-machine scheduling problem with flexible maintenance under human resource constraints.* International Journal of Manufacturing Research, Vol. 17, No. 1, (2022), pp. 22-58.

[19] Perez-Gonzalez, P. and J.M. Framinan, *Single machine scheduling with periodic machine availability.* Computers & Industrial Engineering, Vol. 123, (2018), pp. 180-188.

[20] Nesello, V., et al., *Exact solution of the single-machine scheduling problem with periodic maintenances and sequence-dependent setup times.* European Journal of Operational Research, Vol. 266, No. 2, (2018), pp. 498-507.

[21] Shabtay, D. and M. Zofi, *Single machine scheduling with controllable processing times and an unavailability period to minimize the makespan.* International Journal of Production Economics, Vol. 198, (2018), pp. 191-200.

[22] Luo, J., et al., *Solving the dynamic energy aware job shop scheduling problem with the heterogeneous parallel genetic algorithm.* Future Generation Computer Systems, Vol. 108, (2020), pp. 119-134.

[23] Dehghan-Sanej, K., et al., *Solving a new robust reverse job shop scheduling problem by meta-heuristic algorithms.* Engineering Applications of Artificial Intelligence, Vol. 101, (2021), p. 104207.

[24] Zahmani, M.H. and B. Atmani, *A data mining based dispatching rules selection system for the job shop scheduling problem.* Journal of Advanced Manufacturing Systems, Vol. 18, No. 01, (2019), pp. 35-56.

[25] Habib Zahmani, M. and B. Atmani, *Multiple dispatching rules allocation in real time using data mining, genetic algorithms, and simulation.* Journal of Scheduling, Vol. 24, No. 2, (2021), pp. 175-196.

[26] Jun, S., S. Lee, and H. Chun, *Learning dispatching rules using random forest in flexible job shop scheduling problems.* International Journal of Production Research, Vol. 57, No. 10, (2019), pp. 3290-3310.

[27] Schwendemann, S., Z. Amjad, and A. Sikora, *A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines.* Computers in Industry, Vol. 12, No. 5, (2021), p. 103380.

[28] Bilski, P., *Application of support vector machines to the induction motor parameters identification.* Measurement, Vol. 51, (2014), pp. 377-386.

[29] Calabrese, M., et al., *SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0.* Information, Vol. 11, No. 4, (2020), p. 202.

[30] Schmidt, B. and L. Wang, *Predictive maintenance of machine tool linear axes: A case from manufacturing industry.* Procedia manufacturing, Vol. 17, (2018), pp. 118-125.

[31] Chen, W.-J., *Minimizing number of tardy jobs on a single machine subject to periodic maintenance.* Omega, Vol. 37, No. 3, (2009), pp. 591-599.

[32] Wang, S. and M. Liu, *Multi-objective optimization of parallel machine scheduling integrated with multi-resources preventive maintenance planning* .Journal of Manufacturing Systems, Vol. 37, (2015), pp. 182-192.

[33] Dolipski, M., P. Cheluszka, and P. Sobota, *The relevance of the rotational speed of roadheader cutting heads according to the energy consumption of the cutting process.* Archives of Mining Sciences, Vol. 58, No. 1, (2013), pp. 3-19.

[34] Hand, D.J., *Principles of Data Mining.* Drug Safety, Vol. 30, No. 7, (2007), pp. 621-622.

[35] Mahesh, B., *Machine learning algorithms-a review.* International Journal of Science and Research (IJSR).[Internet], Vol. 9, (2020), pp. 381-386.

[36] Vollert, S., M. Atzmueller, and A. Theissler. *Interpretable Machine Learning: A brief survey from the predictive maintenance perspective*. in *2021 26th IEEE international conference*

on emerging technologies and factory automation (ETFA). (2021).

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

[38] Chen, C., et al., *Predictive maintenance using cox proportional hazard deep learning.* Advanced Engineering Informatics, Vol. 44, (2020), p. 101054.

[39] Schmitt, J., et al., *Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing.* Advanced engineering informatics, Vol. 45, (2020), p.۱۰۱۱۰۱ .

Follow this article at the following site:

Ali Qorbani, Yousef Rabban & Reza Kamranrad. Single Machine Preemptive Scheduling Considering Energy Consumption and Predicting Machine Failures with Data Mining Approach. IJIEPR 2023; 34 (4) :1-17
URL: http://ijiepr.iust.ac.ir/article-1-1694-en.html