

Application of Data Mining in Failure Estimation of Cold Forging Machines: An Industrial Research

Buse TURKOGLU¹, Murat KOMESLI^{2*}, Mehmet Suleyman UNLUTURK³

¹ Norm Group, Department of Research and Development, 10007 Sk. 1/1, AOSB, 35620, Cigli, Izmir, Turkey
buse.turkoglu@norm-fasteners.com.tr

² Yasar University, School of Applied Sciences, Department of Management Information Systems, Universite cad. 37, 35100, Bornova, Izmir, Turkey
murat.komesli@yasar.edu.tr (*Corresponding author)

³ Yasar University, Faculty of Engineering, Department of Software Engineering, Universite cad. 37, Bornova, Izmir, 35100, Turkey
mehmet.unluturk@yasar.edu.tr

Abstract: The industrial companies are now reaching out for solutions that would enable them to reduce the number of manufacturing defects in production so that they may be able to compete and maintain their sustainability in the market. All production processes need to be uninterrupted. This study utilizes data mining algorithms to turn the data created by machines into information. These data mining algorithms are effective tools for reducing the cold forging machine downtime. Furthermore, the selected data mining methodology, the J48 model, generates meaningful results for a large real-life data set and predicts the error according to a behavioral model. The J48 model successfully detected 28 failures from this data set which suggests that it can be a promising method for reducing the periods of downtime of the cold machine.

Keywords: Data mining, Industrial systems, Predictive maintenance, Cold forging machines, Failure estimation.

1. Introduction

The machine productivity has become one of the most important factors in manufacturing products of higher quality among the industrial firms. It is necessary to plan regular maintenance and to solve possible future problems without interrupting the production process in order to increase the machine productivity. For this reason, the companies have started to monitor the machines by utilizing sensors and automation systems. However, this monitoring process requires highly qualified workforce. In order to understand and infer the meaningful results from big data, the data mining techniques should be applied. The operations that need to be performed to extract meaningful information from raw data may simply be called data mining. Data mining is a powerful current technology that helps companies to focus on the data. It allows users to sort through large data sets to identify patterns and establish relationships to solve problems such as manufacturing defects in production through data analysis. The aim of this study is to estimate the failure pauses of the cold forging machines used in an industrial company by applying the data mining methods like clustering, classification and decision tree on real-life data sets.

Real-life data was obtained from 2014 until 2017. The data were cleaned, the relevant features were selected and finally, they were processed by using the several different data mining algorithms. The goal of this study is to ensure that the machines

work efficiently and the production process is not interrupted because of the unplanned pauses. This will increase the quality of the final product and will reduce the consumer complaints.

This paper is structured as follows: In the following section, a background information on data mining is provided. In section 3, a comprehensive literature review has been done. In section 4, materials and methods have been depicted. Suggestions are made in the conclusion section.

2. Background Information

Data mining is used in many applications such as image processing, automation, control systems engineering, optimization, etc. Since the algorithms for each research area are not the same, there are various data mining applications (Koh & Tan, 2005). This paper analyzes several solutions to prevent pauses of industrial machines by using the data mining methodologies. Data mining methods are examined in two ways: predictive and descriptive. Common algorithms in data mining are; classification, estimation, time series analysis, associating, clustering and summarizing. In the classification algorithm, the data related to the class is called labeling. The classification algorithm builds the classification model by learning these labels. Then, the test data is processed to make an estimation according to the training set learned. In this study, the data was

predicted as “yes / no” using maintenance error records in the classification.

The clustering algorithms process data similar to the classification algorithms. The main difference between clustering and classification is that the data to be obtained as the output is defined in the classification algorithm, but this is not the same in clustering algorithm. In the classification algorithm, classes are trained by adding them as inputs. New results are obtained according to the classes in the data model. In the clustering algorithm, the data is added to a common group in such a way that data in the same group is more similar to each other than to those in other common groups.

In the estimation model, the regression algorithms which are linear and logistic regression type of analyses are used. The linear regression formula (Khan, 2019) is given below:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon \quad (1)$$

The relationship between the two data is examined. In the test data set, a relationship is established on the learned linear graphical model. The error of the model is represented by the difference between the predicted and the real data, denoted by ε . It is desirable that the value of ε for Y_i (i -th sample) should be equal to zero. The logistic regression formula (Anon, Multiple Linear Regression Formula, 2019) is used instead of the linear regression when there is no direct effect on the data. The error is measured by the actual distance of the coordinates as estimated from the linear regression.

$$\text{observed} \rightarrow y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p + \varepsilon \quad (2)$$

$$\text{predicted} \rightarrow y' = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p \quad (3)$$

$$\text{error} \rightarrow \varepsilon = y - y' \quad (4)$$

In the time series, the trends of the value calculated by taking the data of the past time are taken into consideration and the forecasting process about the future is carried out. Possible scenarios in the future data are simulated by taking into account the characteristics of the data set. Decisions on accepting credit applications in the banking sector or market research for a new product can be made in time series. There are various time series pattern in usage. These are: periodic trend, increasing / decreasing linear trend, increasing / decreasing curriculum trend. Time series algorithm learns the obtained data to create a pattern and try to create the future values in the same structure. For the time series to be consistent, the obtained data

used in the estimation process must be acceptable when compared to the common input data sets. If a reasonable time series is not defined, the algorithm will give an average value which is known as purely random error.

In the association algorithm, it is tried to establish a meaningful relationship by looking at the correlation between two variables called X and Y . Thus, the meaningful data can be generated according to the relation among the attributes. This is frequently used in sales and marketing strategies. The association algorithm operates according to two formulas given in (5) and (6). The set support value of the data is calculated by utilizing the values in the data set. Trust formula is shown as:

$$X \Rightarrow Y: (X \cap Y)/N \quad (5)$$

(N : Total number of items in the data set)

The confidence between the binary values is found by calculating the relation formula shown below:

$$X \Rightarrow Y: \text{support}(X, Y) / \text{support}(X) \quad (6)$$

Thus, when a new data set is transmitted, the estimation is made based on the connection model (Lai & Cerpa, 2001). The summarization function is often used in text mining. In the analysis of the texts in social networks, grouped together with the summarization function and the sentiment analysis, the idea mining studies are mainly performed. An important feature of the summarization function is that it works with the big data set. The most widely accepted representation in the literature is the use of graph theory in order to show that individuals or entities are depicted as edges as well as nodes and relations (Seker, 2015).

The failure of the machine is referred to as an interruption or a stopping the system where one or more variables are affected. The most common way to detect or predict errors is to analyze combination of thermal and physical characteristics or produced parts by using information gathered and collected from software and hardware. Accurate analysis of the information and early detection of the error have critical importance for the surveillance of the system (Amooee, Minaei-Bidgoli, & Bagheri-Dehnavi, 2011).

Fault detection is defined as detecting abnormal process behaviors. Error detection techniques are divided into two categories: Model-Based Approaches and Data-Based Approaches (Cugueró-Escofet, et al., 2016). Model-based

approaches require the modeling of the process flow which validates normal behavior by comparing each input for error detection. The fault is identified by finding the defect according to the comparison model.

Unlike model-based approaches, database-based detection approaches require a preliminary process knowledge such as large amounts of process history. The data can be identified in two different ways in the system: quantitatively and qualitatively. Qualitative background information can be obtained by expert systems and QTA (Qualitative Trend Analysis) methods. The neural networks and statistical methods can be utilized to obtain the quantitative information (Gertler, 1988).

Because of the difficulty in establishing comparative models for real-time industrial processes, model-based fault detection techniques are not as common as database-based fault detection techniques. For this reason, the error detection was accomplished by utilizing database systems in this study. The results of the tested algorithms are compared with the ones obtained by using the “hypothesis testing” method.

3. Literature Review

Bastos, Lopes & Pires (2012) collected data from similar machines, stored them in a database, and tried to increase productivity by estimating the errors while using a data mining method. The system they built consists of three modules: operational, model, virtual. The operational module consists of the physical machines and the sensors connected to them. This module also performs the normalization task of the data. Data mining is done in the model structure. The core structure focuses on the behavior and events and tries to reveal information. The virtual platform module is the warning and information portal that shows the results of the data mining. In this paper, it is mentioned that different algorithms can be applied depending on the type of data to be analyzed:

- Decision Trees: Decisions are made in the form of tree structure utilizing the rules. This method uses numerical values.
- Artificial Neural Network (ANN): This structure imitates the human brain like learning and self-correcting and it is the most commonly used method in data mining.
- Genetic Algorithms: This technique can also work with all data types. However, it is usually used for optimization problems.

- Clustering: The similar data in the data set is grouped by common field or class.
- Visualization: The process utilized in data mining to provide the user a deeper understanding of the results.

The data mining functions used in this paper are C 5.0, Neural Network, Apriori and K-Means. The C5.0 algorithm is used on a large number of data sets. The decision points are determined according to the entropy on the decision tree. In the Apriori algorithm, one attribute is examined at a time and it tries to reveal the relationship with the other attributes. In the decision tree structure, k-1th element is selected from a list of k elements. Individual elements are pruned if any of them has insufficient lower frequency structure. Then, the remaining elements are processed. The purpose of the algorithm is to predict the future probability of the machine failure based on the existing records and the support maintenance interventions which are in planning.

In another implementation of the same study, Bastos, Lopes & Pires (2014) aimed to predict errors by building a model of maintenance activities generated in industrial units. While data was being analyzed, the results were compared by using the RapidMiner (Anon, Rapidminer, 2018) program. The research consists of four phases:

- Understanding the problem: The decision support system is divided into three components: alarms, actions and predictions. The paper published by Bastos, Lopes & Pires (2014) focuses on increasing the efficiency and continuity of production.
- Data Preparation: Inconsistent, missing, and dirty data was removed in their study. So, data mining function can operate more efficiently. A random machine with data objects and rules has been chosen to obtain the results without affecting performance.
- Data Mining: Ten attributes in total within 529 records were processed in their dataset. The training set has had 330 records collected for 8 months and the test set has had 199 records collected for 3 months. There is a 71.3 % likelihood that the confidence interval [0.0, 0.25] covers the true classification error of the model on unseen data.
- Evaluation and Interpretation: The model provides the desired trust and accuracy after examining their test results.

The tests were carried out on one machine for sampling. The proposed approach has been verified with respect to the empirical results.

Jean Nakamura's article published in 2007 (Nakamura, 2007) estimates the failure of machines by looking at the results of the temporary vibration data. The difficulty in estimating the fault in industrial machines is that the faulty is not linear. For this reason, there is not any available tool or algorithm for estimating machine failures.

Eight numerical columns were used for testing and training data. CaseID was chosen as the unique value. Common data is grouped for the data preparation. Two copies of the database were created for testing and training, and 17% of the data is reserved for testing. The Time-to-Failure (TTF) system uses a CBR (Case-Based Reasoning) system to map the library data stored in a newly created problem. If there is no match, the indirect match is searched using the generalization step. Analysis was carried out by the Windows SPSS and Waikato University's WEKA programs. The CBR, SPSS Linear Regression, Weighting Parameters, WEKA Algorithms and DataSet, WEKA (Multilayer Perceptron, Logistic, J48 tree, K *) are the compared algorithms. The best results were obtained with Logistic Regression algorithm which was selected as the best algorithm in the paper. The second best algorithm chosen was the Multilayer Perceptron.

The equipment manufacturers have difficulties in increasing efficiency and reducing the cost of the service. The predictive maintenance strategy solves these difficulties in order to plan for the optimum time. However, there are two difficulties involved in the predictive maintenance strategy. They are few known features about the machines and the absence of a machine power used to handle them when large data is gathered. The data mining and prognostic systems should be used as a solution. In this method, the first step is to examine the data generated by the data mining technique so as to determine the relevant properties at the optimal level. The second step is to use this model for the predicted features and for observing the action in live data (Zong, 2017).

Data mining algorithms are generally treated in three different categories and common data mining techniques are shown in Figure 1:

- Interpretability; Level of understanding the data.
- Predictive Accuracy; The ability of predicting new entries.
- Computational Efficiency; Performance of working with the big data sets.

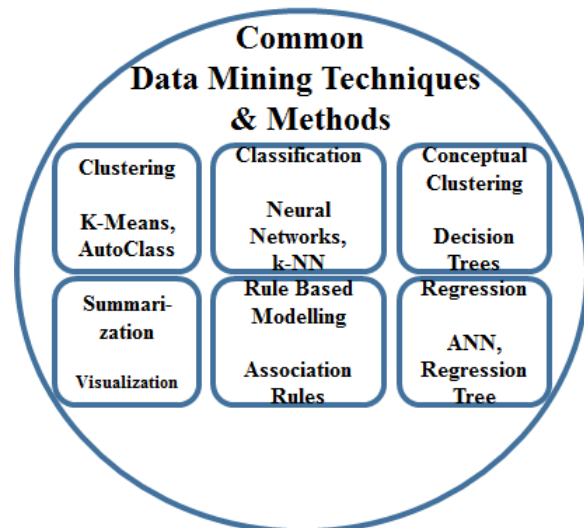


Figure 1. Common data mining techniques

One of the techniques used in data preparation is ISF (Independence Significance Feature). By using this technique, unusable attributes are quickly eliminated. While data model is being created, the data mining technique should be selected. In this model, Classification and Regression Trees or CART decision tree model and k nearest neighbor (KNN) model (Amaricai, 2017) were selected. A five-fold cross-validation method was used to test the accuracy of the compared algorithms. According to the results obtained by cross validation, the CART algorithm is more consistent and the maintenance time is planned with the Weighted Mean Slope (WMS) model.

The estimation of failures in spare parts is one of the topics that inventory management is interested in. Nachirat Rachburee, Samatachai Jantararat, Wattana Punlumjeak analyzed the prediction spare part failure (Rachburee, Jantararat, & Punlumjeak, 2012). For this work, the maintenance time data set between 2013 and 2016 was considered for the two models used in their analysis: the ANN (Artificial Neural Network) and the SVR (Support Vector Machine for Regression) models.

11 attributes and 10 faulty parts were selected in the data set. Between the compared models, the accuracy rate of 87.07% was achieved in the ANN model and 88.24% was achieved in the SVR model. When the parts were examined separately, SVR model was chosen as the most suitable model with 92.7% accuracy. The SVR model was applied when results were satisfactory. However, the relationship techniques can be used in rule-based technique for further studies.

According to Wang et al. (2017) preventive maintenance techniques ensure that equipment is

Table 1. Sample data used for this study

DSINIFI	AOLCU	BOYM	OZELURUN	NUMUNE	QTY	HMDCINS	HMD_TAV	HMD_TIP	BAKIM
10.9	18	55	1	0	19000	2	0	1	0
10.9	18	55	1	0	19000	2	0	1	1
8.8	24	70	0	0	50000	8	0	1	0
8.8	24	70	0	0	50000	8	0	1	0
8.8	24	70	0	0	50000	8	0	1	0
8.8	24	70	0	0	50000	8	0	1	0
8.8	24	80	0	0	40000	8	0	1	0
8.8	24	50	0	0	40000	8	0	1	0
8.8	24	50	0	0	40000	4	0	0	0
8.8	24	50	0	0	40000	8	0	1	0
8.8	24	50	0	0	40000	4	0	0	0
8.8	24	60	0	0	60000	8	0	1	1
8.8	24	60	0	0	60000	8	0	1	1
8.8	24	60	0	0	60000	8	0	1	0
8.8	24	60	0	0	60000	8	0	1	0
8.8	24	60	0	0	60000	8	0	1	0
8.8	24	55	0	0	12000	8	0	1	0
8.8	24	55	0	0	12000	8	0	1	0
8.8	24	65	0	0	15000	8	0	1	0
8.8	24	65	0	0	15000	8	0	1	0
8.8	24	90	0	0	18000	8	0	1	0
8.8	24	90	0	0	18000	8	0	1	0
8.8	24	100	0	0	20000	8	0	1	0
10.9	20	60	1	0	10000	7	1	1	0
10.9	18	55	1	0	19000	2	0	1	0

properly serviced, thereby increasing the quality of service provided to customers and reducing the cost of unplanned downtime. Various devices have features such as logging, fault actions or system messages for fault conditions. For this reason, the failure records in ATMs (Automated Telling Machine) are taken as an example in this article. For data selection, seven months of data were received from 1000 ATMs in North America. There are 128 different fault types and 5 different devices in this data. In this experiment, InfoGain, GainRatio, ReliefF and Symmetrical Uncert sorting algorithms are used in Weka program. Thus, the properties that are not relevant to the data and create deflection should be removed. Based on the results, the InfoGain function and toolkit have been selected based on the performance/result selection. After the data set was prepared, tests were carried out with four models in the Weka program. Based on the correlation rate, the XGBoost clustering function is the most efficient result for this data set.

Riccardo Accorsi et al. (2017) published an article about the case-based maintenance systems in the

case of complex production systems with different methods and related models. All models were run on real-time data and according to the results, the decision tree model was determined as being the most effective technique (Accorsi et al., 2017).

4. Materials and Methods

Different approaches implemented are presented for fault detection. Among the compared algorithms, the algorithm that maximizes the efficiency and gives the most reliable result is determined. The steps of the comparison work are shown in the present section. This work is carried out in cooperation with an important supplier in the automotive industry and a manufacturer for companies such as Ford, and Tesla.

Data preparation is the most important part of data mining. At this stage of the study, data needs to be converted to acceptable form for mining algorithms. For this reason, it starts with the characteristics of the product produced on the machine. Then, deviated values are determined, the data is cleaned by deleting NULL values

Table 2. Comparison of the Algorithms

Algorithms	Correctly Classified Instances	Root Mean Squared Error
IB1 instance-based classifier	48.8722 %	0.6494
Naive Bayes Classifier	73.4962 %	0.4472
J48 pruned tree	77.4436 %	0.418
KStar Beta Version	60.1504 %	0.5199
Locally weighted learning	76.8797 %	0.4173

and correcting the inconsistent values manually, and constant variables are ignored. According to the processed product, it is determined if the records represent a faulty or a not faulty case. Some algorithms process files with the extension “.csv”, while some of them process “.xls” or “.xlsx” format files. For this reason, the final stage of data preparation is the formatting of the file containing the data. The data set obtained is presented in Table 1.

There are nine attributes and one result column in the dataset. DSINIFI is a numerical value which determines the strength class of the manufactured material. The values in this attribute can be of the following values: 8.8 or 10.9. The AOLCU column is also a numerical value between 18 and 24 that shows the surface area of the material. The size of the material produced in the BOYM column is another numerical values and ranges from 50 to 100. The OZELURUN column is a Boolean value that can be either 0 or 1. This column depicts if the material is a standard or a special product. According to the author’s experience, the failure rate in standard products is lower than that of special products. For this reason, the OZELURUN column was added. NUMUNE column accepts Boolean values that show the total amount of production produced as samples for the customer. According to the author’s experience, sample production affects the fault records. The QUANTITY value is a numerical value that shows the total production amount. HMDCINS, HMD_TAVLAMA and HMD_TIP columns contain information about the raw material of the product produced in the machine. The BAKIM column lists the Boolean value of the faulty record.

The results of five different algorithms compared by using the WEKA program are shown in Table 2. These algorithms were compared against correctly classified instances and root mean squared error. Finally, J48 was chosen as the most efficient algorithm, with 77% accuracy. The company which provides data analyzes the material-based

pauses and takes precautions at the appropriate time by using the J48 algorithm. The relation between material and raw material is established by taking the fault data into consideration. This increases the productivity by avoiding the stops.

Table 3 shows the confusion matrix with respect to the results of J48 algorithm. It shows that 28 failures would have been caught if the J48 data mining algorithm had been utilized in this company.

Table 3. Confusion matrix

n = 2028	Predicted: NO	Predicted: YES	
Actual: NO	TN = 1534	FP = 119	1653
Actual: YES	FN = 347	TP = 28	375
	1881	147	

The legend information of the confusion matrix is the following:

- True Positives (TP): These are cases in which yes was predicted (they have the failure), and they do have the failure.
- True Negatives (TN): “No” was predicted, and they don’t have the failure.
- False Positives (FP): “Yes” was predicted, but they don’t actually have the failure. (Also known as a “Type I error.”)
- False Negatives (FN): “No” was predicted, but they actually do have the failure. (Also known as a “Type II error.”)
- n: Total number of items in the data set

RapidMiner is a software platform for software engineers that combines data preparation, machine learning, and predictive model development (Prytz, 2014). The rule model generated by using RapidMiner is shown in Figure 2. First, the data was sent as an input to the decision tree model. Then, by using the Naïve Bayes classification function, the class of the data was determined. Generated rules were estimated through the induction method.

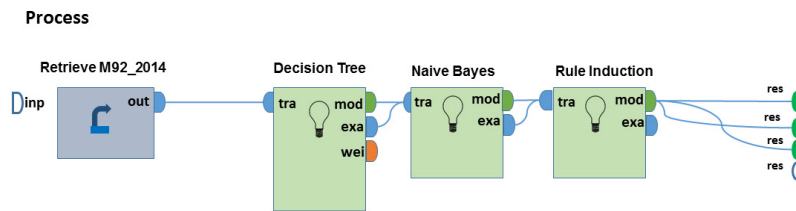


Figure 2. Decision tree model in Rapidminer

RuleModel

if QUANTITY \leq 30500 then No (1255 / 436)

if QUANTITY \leq 68000 and OZELURUN \leq 0.500 and BOYM $>$ 65 and AOLCU = 24 then No (41 / 15)

if QUANTITY \leq 68000 and BOYM $>$ 89.150 and QUANTITY $>$ 43000 and AOLCU = 16 then No (9 / 3)

if QUANTITY \leq 53500 and AOLCU = 24 and HMDCINS $>$ 2.500 then No (22 / 14)

if QUANTITY \leq 68000 and BOYM \leq 48.250 and BOYM $>$ 40.250 and QUANTITY \leq 60500 and AOLCU = 16 then No (17 / 10)

if HMD_TAVLAMA \leq 0.500 and HMDCINS \leq 3.500 and AOLCU = 14 then No (2 / 1)

if QUANTITY \leq 68000 and DSINIFI \leq 9.850 and HMDCINS $>$ 2.500 and BOYM \leq 150 and QUANTITY \leq 47500 and HMDCINS \leq 11 and BOYM \leq 140 and AOLCU = 42871 then No (1 / 0)

if BOYM \leq 62.500 and HMDCINS \leq 7.500 and QUANTITY \leq 82500 and QUANTITY $>$ 61500 and AOLCU = 16 then No (6 / 5)

else No (89 / 82)

correct: 1442 out of 2008 training examples.

All the values were determined according to the above decision tree model called “rule model”. As a result “No” is obtained, when there is no branch found in the decision tree structure. However, the success of the model is approximately 71.8% which correctly classifies 1442 cases out of 2008 records. In this case, the RapidMiner program did not get better results than that of the Weka program. Furthermore, the best algorithm was maintained in J48 model which accomplished 77% accuracy (see Table 2). However, the attained

accuracy is low, but this study is important for the industrial company in question since no such study has been carried out before and 28 failures would have been caught if the selected data mining algorithm had been utilized between 2014 and 2017.

5. Conclusion

The aim of this study is to estimate the failures of cold forging machines by using data mining methods. It has been carried out in one of the main suppliers in the automotive industry which is located in Izmir, Turkey. Firstly, the data was cleaned, normalized and the NULL values were removed. Then, the models were trained and tested by using the Weka and RapidMiner programs. It was observed that the J48 Decision Tree algorithm gave the most efficient result with 77% accuracy and ~ 0.4 error value achieved compared with the other algorithms used in the Weka program. On the other hand, the decision tree that was modeled by RapidMiner gave an accuracy result of 71.8%. For this reason, the use of the J48 algorithm in the Weka program is recommended for the company under discussion. There is no comparable model in this industrial company since no such study has been conducted before. This study correctly caught 28 failures which could be prevented if the selected J48 model was employed in analyzing the machine data between 2014 and 2017.

For future work, data mining functions for all machine groups or all material groups will be further investigated.

Acknowledgement

The research for this paper was financially supported by the company, Norm Group that is located in Cigli, Izmir – Turkey.

REFERENCES

1. Accorsi, R., Manzini, R., Pascarella, P., Patella, M. & Sassi, S. (2017). Data mining and machine learning for condition-based maintenance, *ScienceDirect*, *11*, 1153-1161. Retrieved from <<https://www.sciencedirect.com/science/article/pii/S235197891730447X>>.
2. Amaricai, A. (2017). Design Trade-offs in Configurable FPGA Architectures for K-Means Clustering, *Studies in Informatics and Control*, *26*(1), 43-48, DOI:10.24846/v26i1y201705
3. Amooee, G., Minaei-Bidgoli, B. & Bagheri-Dehnavi, M. (2011). A Comparison Between Data Mining Prediction Algorithms for Fault Detection (Case study: Ahanpishegan co.), *IJCSI International Journal of Computer Science Issues*, *8*(6), 425-431. Retrieved from <<https://arxiv.org/abs/1201.6053>>.
4. Anon. (2018, December 18). *Rapidminer*. Retrieved March 13, 2019, from <<https://rapidminer.com>>.
5. Anon. (2019, March 14). *Multiple Linear Regression Formula*. Retrieved March 14, 2019, from <<http://www.saedsayad.com/mlr.htm>>.
6. Bastos, P., Lopes, I. & Pires, L. (2012). A Maintenance Prediction System using Data Mining Techniques. In *World Congress on Engineering* (pp. 1448-1453). London, U.K., International Association of Engineers. Retrieved from <<http://hdl.handle.net/10198/7298>>.
7. Bastos, P., Lopes, I. & Pires, L. (2014). Application of data mining in a maintenance system for failure prediction, *Safety, Reliability and Risk Analysis: Beyond the Horizon*. London, U.K., Taylor & Francis Group. Retrieved from <<http://hdl.handle.net/10198/8960>>.
8. Cugueró - Escofet, M., Quevedo, J., Alippi, C., Roveri, M., Puig, V., García, D. & Trovó, F. (2016). Model vs. Data based Approaches Applied to Fault Diagnosis in Drinking Water Transport Networks, *Journal of Hydroinformatics*, *18*(5), 831-850, DOI:10.2166/hydro.2016.218
9. Gertler, J. J. (1988). Survey of Model-Based Failure Detection and Isolation in Complex Plants, *IEEE Control Systems Magazine*, *8*(6), 3-11, DOI:10.1109/37.9163
10. Khan, I. (2019, March 14). *How Linear Regression can also be used to do classification?* Retrieved March 2019, from Analytics Buddhu: <<https://analyticsbuddhu.wordpress.com/2016/06/28/how-linear-regression-can-also-be-used-to-do-classification/>>.
11. Koh, H. C. & Tan, G. (2005). Data Mining Applications in Healthcare, *Journal of Healthcare Information Management*, *19*(2), 64-72.
12. Lai, K. & Cerpa, N. (2001). Support vs Confidence in Association Rule Algorithms. In *Proceedings of the OPTIMA Conference*, Chile, Universidad de Talca. Retrieved from <https://www.researchgate.net/profile/Narciso_Cerpa/publication/233754781_Support_vs_Confidence_in_Association_Rule_Algorithms/links/0fcfd512a4907b8aca000000/Support-vs-Confidence-in-Association-Rule-Algorithms.pdf>.
13. Nakamura, J. (2007). *Predicting Time-to-Failure of Industrial Machines with Temporal Data Mining*. University of Washington, Computing and Software Systems, Washington: University of Washington. Retrieved from <<https://s3.amazonaws.com/academia.edu.documents/39357344/jnakamura.pdf>>.
14. Prytz, R. (2014). Machine learning methods for vehicle predictive maintenance using off-board and on-board data, *Halmstad University Dissertations*. Halmstad University Press.
15. Rachburee, N., Jantarat, S. & Punlumjeak, W. (2012). Time Series Analysis for Fail Spare Part Prediction: Case of ATM Maintenance, *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-6), 49-52. Retrieved from <<http://journal.utm.edu.my/index.php/jtec/article/view/2433>>.
16. Seker, S. E. (2015). Çizge Teorisi (Graph Theory), *YBS Ansiklopedi*, *2*(2), 17-29. Retrieved from <http://ybsansiklopedi.com/wp-content/uploads/2015/05/cizge_teorisi.pdf>.
17. Wang, J., Li, C., Han, S., Sarkar, S. & Zhou, X. (2017). Predictive maintenance based on event-log analysis: A case study, *IBM Journal of Research and Development*, *61*(1), 11:121-11:132, DOI:10.1147/JRD.2017.2648298
18. Zong, Y. (2017). *Maintenance Cost and Residual Value Prediction of Heavy Construction Equipment*. University of Alberta, Department of Civil and Environmental Engineering. Edmonton, Canada: University of Alberta. Retrieved March 14, 2019, from <http://yizong-profolio.com/wp-content/uploads/2015/03/Zong_Yi_201701_MSc.pdf>.