Integration of Machine Learning Techniques and Control Charts for Multivariate Processes

Deniz DEMIRÇIOĞLU DİREN¹, Semra BORAN² and İbrahim ČİL³

¹ Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey, ddemircioglu@sakarya.edu.tr
² Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey, boran@sakarya.edu.tr
³ Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey, icil@sakarya.edu.tr

*corresponding author ddemircioglu@sakarya.edu.tr Tel:90-264-2955540

ABSTRACT

Using multivariate control chart instead of establishing univariate control chart for all variables in processes provides time and labor advantage. In addition, it is considered in the relations between variables. However, the statistical calculation of the measured values of all variables is seen as a single value in the control chart. Therefore, it is necessary to determine which variable(s) is the cause of the out of control signal. Effective corrective measures can only be developed when the causes of the fault(s) are determined correctly. The aim of the study is to determine the machine learning techniques that will accurately estimate the type of fault. With the Hotelling $T^2$ chart, out of control signals are identified and the types of faults affected by the variables are defined. Various machine learning techniques are used to compare classification performances. The developed model was applied in the evaluation of the paint quality in a painting process. ANN was determined as the most successful techniques according to performance criteria. The novelty of the study is to classify the fault according to the types of faults, not the variables. Defining the faults according to its types will enable to take effective corrective actions quickly.

Keywords – Multivariate Control Chart, Naive Bayes-Kernel, K-Nearest Neighbor, Decision Tree, Artificial Neural Network, Multi-Layer Perceptron, Deep Learning
1 INTRODUCTION

In order to survive in the competitive environment of globalization, company has to take into consideration quality as well as the price factor. So that the importance of statistical process control is increasing day by day because of processes become more complex. The selection of the appropriate quality control chart by the companies will be importance to control the process and reduce the variability. Generally, a process is consist of more than one quality variable in real life. It is impossible to control several variables simultaneously with univariate control charts. For this reason, control charts with multi variable such as Hotelling $T^2$ [1], Multivariate Cumulative Sum Control Chart [2] and Multivariate Exponentially Weighted Moving Average [3] have been developed to evaluate the multiple variables.

The advantage of the multivariate control chart is that it takes into account the relationship between the variables and saves time and labor. Unlike these advantages, the main disadvantage is that they cannot show the cause of the out of control signal because of all quality variables are calculated as a single value on the control chart. For this reason, it is necessary to use some methods to find out the variables that cause of the out of control signals. In literature, there are several studies that used different methods and approaches about detecting the cause of out of control signal in multivariate process. Some studies focus on unnatural pattern recognition on control chart, while others are focus on mean or variance or both shifts. Methods for detecting the variables can be divided two class such as statistical methods and machine learning techniques. The statistical methods that used to identify the variable that the cause of out of control signal are; discriminant analysis method [4], Mason Young Tracy decomposition method [5], principal component analysis [6] and causation-based $T^2$ decomposition [7]. However, these methods do not have the ability to predict new conditions. For this reason, it is useful to use the machine learning techniques, which predict the new situation by learning from the historical data. Artificial Neural Network (ANN) [8,9],
Support Vector Machine (SVM) [10-12], and hybrid methods [13-15] were used in studies based on pattern recognition for detecting the variable(s) that cause of out of control signal in multi variable control charts. Similarly, ANN [13-20], SVM [21], Decision Tree (DT) [22-24], K-Nearest Neighbor (K-NN) [25] and hybrid methods [26,27] were mostly used in studies based on mean and or variance or both shifts for detecting the variable(s) that cause of out of control signal in multivariate control charts.

In this study, it is aimed to classify the out of control signals according the fault types that occur in a multi variable process by the most appropriate machine learning techniques. For this, ANN, DT, K-NN, Naive Bayes-kernel (NB-k), Multilayer Perceptron (MLP) and Deep Learning (DL) were commonly used in real life process [24] and the performance of learning were compared.

In addition, another novelty of the study is to make classification according to fault types instead of variable. Thus, experts determine the value range, high, medium and low, of the variable causing the fault. As the fault is known that occurred from which one of low, medium or high values of the variable, the correct corrective actions related to the variable can be determined quickly.

The intended model was used to determine the process variables affecting the quality of the paint in the painting process. Six different machine learning techniques were compared according to four performance criteria likes as classification accuracy, squared error (SE), squared correlation ($R^2$) and root mean squared error (RMSE). ANN was found to be the most successful technique to detect the fault type of a new sample from the process.

The paper is organized as follows; the proposed model and the methods that were used in study is presented in Section 2, the results of the implementation were shown in Section 3,
the comparison of method results were presented in Section 4 and finally study was summarized in Section 5.

2 MATERIALS AND METHODS

2.1 Hotelling $T^2$ Control Chart

Hotelling $T^2$ control charts, provided by Hotelling (1947), developed for simultaneous monitoring of associated p dimensional quality variables of a multivariate process [1]. This control chart is derived by adapting $T^2$ statistic, a distance measure based on normal distribution, to the graph. If $X_1, X_2, \ldots, X_p$ are p correlated quality characteristics (variable), the parameters are unknown, the control chart is formed by historical sample data and sample size is 1 then $T^2$ is given Eq. 1.

$$T^2 = (X_i - \bar{X})'(S)^{-1}(X_i - \bar{X})$$  \hspace{2cm} (1)

where $S$ is variance-covariance matrix of sample and $\bar{X}$ is mean vector of sample.

The upper control limit (UCL) and lower control limits (LCL) are given by Eq.2 and Eq.3 respectively

$$UCL = \frac{(m-1)^2}{m} \beta_{a,p/2,(m-p-1)/2}$$  \hspace{2cm} (2)

$$LCL=0$$  \hspace{2cm} (3)

where $m$ is the observation size in the historical data, $\beta_{a,p/2,(m-p-1)/2}$ is beta distribution, $p/2$ and $(m - p - 1)/2$ are parameters of distribution.

2.2 Machine Learning Techniques

Machine learning is a technique based on logical or binary operations that models a problem according to the present. It performs data analysis according to automatic calculation procedures [28]. The techniques are divided into supervised and unsupervised. In supervised
techniques, the class numbers and the relations between input and output are predefined. Whereas unsupervised machine learning techniques do not involve these values.

In this study, NB-k, K-NN, DT, ANN, MLP and DL classification and prediction techniques were used in as machine learning techniques.

2.2.1. Naive Bayes-Kernel Technique

Naive Bayes kernel (NB-k) is a simple classifier and based on Bayesian theorem. It assumes independency between each of the class. NB-k techniques determines conditional probability for relationship between each variable and class for each item. It is used when the number of values is high. Bayesian theorem is presented in Eq. 4.

\[
P(C|X) = \frac{P(X|C)P(C)}{P(X)}
\]

(4)

where, \( P(C|X) \) is the posterior probability, \( p(X|C) \) is the probability of \( X \) when is given \( C \) and \( P(C) \) is the probability of obtaining class.

It has been observed that the accuracy rates of Naive Bayes have been increase when it is used with kernel density function [29, 30].

In the kernel density function method, the bandwidth is determined after the kernel number is selected. There are many different methods for determining the bandwidth. Otherwise, bandwidth is determined by expert opinion [31].

2.2.2. k-Nearest Neighbor Technique

k-Nearest Neighbor (KNN) which was proposed in 1951 is one of the simplest pattern recognition method that classify according to the given k value by the nearest neighbors class and it is a non-parametric supervised classification method [32]. Unlike other supervised learning techniques, it does not have a training phase. The classes in a data set is determined
by historical data. Each sample in the data set to be classified in the test phase is processed individually. In order to determine the class of this sample, number of k must be chosen, k is neighbors number of the unknown sample. The k-nearest neighbors are determined based on some distance functions such as Euclidean, Manhattan, Mahalanobis and Minkowsk. Euclidean is generally preferred over this distance functions types. In a comparative study of distance measurements methods, best performance in these methods was found as Mahalanobis [33], but in study, it was preferred Euclidean distance (EUD) because of the suitability to the Gaussian distribution and ease of use [34]. This function that given in Eq.5, is the straight distance from \( p = (p_1, p_2, \ldots, p_n) \) to \( q = (q_1, q_2, \ldots, q_n) \)

\[
\text{EUD}(p, q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}
\]

(5)

2.2.3. Decision Tree Learning Technique

The decision tree (DT), which predicts the target variable according to various input variables, is a widely used classifier among machine learning techniques [35]. There are three types of nodes in a tree such as, a root, non-terminal and leaf.

Decision tree starts with the root node determined by entropy criterion. The non-terminal node and the leaves are separated to start with the highest entropy value. The entropy formula is shown in Eq.6.

\[
H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)
\]

(6)

where \( X \) is a discrete random variable and can take \( n \) possible values \( (x_1, \ldots, x_n) \).

The root, non-terminal root and leaf represent all training cases, a subset of the training cases. First two nodes root and non-terminal nodes include a variable value test. It is divided training cases into two or more subsets according to results of test. The tree is pruned by
removing the branches with little statistical validity. Produced results by this classifier are simple to understand and interpret [36]. The classification rules consist of each route leading from root node to leaf node. In addition, tree is understandable with If-Then rules, so it can be preferred over difficult interpretation techniques even if less successful.

2.2.4. Artificial Neural Network Technique

Artificial neural network (ANN) technique, one of the most effective learning methods known today, are a robust approach to estimate the problem by learning how to interpret real world data [37]. ANN is a calculation model used to simulate the human nerve cell [38].

A network is occur with an input, hidden and an output layer. The output of each neuron is computed by the weights of the nodes in the previous layer. Transfer function such as sigmoid, tangent-sigmoid (tansig), and logarithmic-sigmoid (logsig) is implemented to the input of the hidden node to determine hidden node results.

To train ANN, the data set is divided into two parts as training and test data. ANN must be train with some learning techniques such as Levenberg–Marquardt backpropagation (trainlm) and quasi-Newton backpropagation (trainbfg) to achieve the best result. The back-propagation neural network (BPNN) is the most popular type of neural network. BPNN uses a supervised learning method. The training data are randomly selected from combinations of inputs and outputs. Other data are used for testing. A well-trained ANN model has capability to define a relationship between inputs and outputs without having a mathematical relationship. If error is reach to minimum value, process is stopped. Otherwise, it is modified connection weights for receiving better results.

2.2.5. Multi-Layer Perceptron

Multi-layer perceptron (MLP) is a technique of having nonlinear decision nodes. MLP is an artificial neural network structure and it can be used for classification and regression. If used for classification, these MLPs can apply non-linear discriminators and, if so, can
approximate the nonlinear functions of the input for regression. It has an input, hidden and output layers. The input layer transmits the inputs from the external world to the hidden layer. Then this information transmitted to the next layer after processed. There can be more than one hidden layer. Finally, the information is sent to the output layer [39].

2.2.6. Deep Learning

Deep learning (DL), which began to part in machine learning in 2006, includes multiple hidden layers of artificial neural networks [40]. It takes into account non-linear processing in multiple layers and controlled or uncontrolled learning factors. The technique works by taking the output of the previous layer as input [41]. It has been shown to be successful at solving the complex structures and so be applied to different fields [42]. DL can be understood as a method for improving results and optimizing processing times in various computations [41].

2.2.7. Performance Criteria of Machine Learning Techniques

Several performance criteria allow techniques to be compared to determine which of the machine learning techniques are more suitable for the process. In addition to the accuracy which is the most frequently used criteria in the literature, there are different criteria used in many different studies like precision/recall, receiver operating characteristic (ROC) curve area, squared error, correlation etc. [43,44].

For evaluate and compare the predict performance of the techniques, four performance criteria were used in this study. These are; accuracy, squared error (SE), squared correlation coefficient ($R^2$), root mean squared error (RMSE). The equations of the performance criteria is described in Table1.

2.2.8. Proposed Model
The purpose of the study is to determine the fault types from the historical data and to estimate which type of fault the new samples coming from the process belongs to. The proposed model consists of four steps.

- Establishing a Hotelling $T^2$ control chart by historical measurement values, to identify samples that caused the process to be out of control
- Detecting the fault types that occur in process then consisting the classes according to these fault types
- Classification the fault types of the sample in data set with machine learning techniques
- Comparison of the techniques according to performance criteria like accuracy, squared error etc.

The architecture of the proposed model is shown in Figure 1.

3. RESULTS

In this study, machine learning techniques were implemented to a painting process of an automotive supplier company. In the company, seat, door panel and bumper products are produced. A door panel’s paint quality was analyzed. Parts are dried in the drying cabinet for fixing the dye. The most important variables affecting this process were selected as, the temperature, pressure and humidity of the cabinet.

3.1. Establishment of Hotelling $T^2$ Control Chart

The quality of the part was determined according to these variables and data of 581 samples from the process were taken. Sample size is one.

To examine the quality control status of the sample values the Hotelling $T^2$ control chart was used established (Figure 2). Each point on the control chart contains the values of all the
variables of that sample according to the calculation in Eq. 1. Control limits were calculated using Eq.2 and Eq.3.

In control chart, 27 samples were out of the control. Minitab 17 was used to establish control chart.

3.2. Determination of Class of Fault Types

The causes of out of control must be accurately determined to eliminate the faults. In order to make it easier to identify the cause of the fault, it is necessary to determine the types of the faults. For this reason, the range of values for variables was determined with quality experts by taking into consideration the evaluated sample data. The value ranges are as in Table 2. The encountered fault types and the faultless samples number in data set are identified according to value ranges in this table.

Different kinds of fault types due to drying cabinet can appear in painting process. Value ranges of variable and the types of faults that occur because of these ranges are shown in Table 3. If the pressure is high and the humidity is low, then the air conditions of the cabinet are not suitable for the mixture of paint thus paint sag occur on painted part and it is defined as fault type 1. If there is dust in the cabinet and the pressure of the cabinet is low, dusting on the surface of the product is seen, this is defined as fault type 2. If the temperature of the cabinet is not sufficient high, scratches will form because the part is not completely dry and this is also defined as fault type 3.

Samples of fault type1 are, 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th, 45th, samples of type 2 are 15th, 17th, 18th, 19th, 20th, 21st, 24th, 27th, 28th, 29th, 33rd, samples of type 3 are 47th, 48th, 63rd, 79th, 524th, 564th, 571st and other samples of data are faultless.

In this study, unlike other studies, root cause is not only determined with variable. The value range of the variable was also determined. For example, in addition to detecting that the
variable causing the error type 3 is temperature, this refers that the temperature is low. Knowing the value range allows to make the right decision about corrective actions and accelerate the process of improvement.

3.3. Implementation Machine Learning Techniques

Supervised machine learning techniques were used to classify process faults. NB-k, KNN, DT, NN, MLP and DL are classification and prediction techniques have been used to detect the fault class in the process data set. Rapid miner Studio 7.6 was used for implementing the techniques.

Performance criteria which described in the previous section were compared to find out which technique is better acceptable for the data set. In this study, the models were tested by cross validation. Cross validation divides the data set into selected numbers for treats part of the set as training data. Then technique repeat the selected number of times, each time with different test data. The average of the accuracy rates obtained at the end of each classification also generate the overall accuracy of the technique. Considering past studies number of fold of cross validation was determined as 10 [45]. By using stratified-sampling, the same rate was obtained for training and testing each time.

3.3.1 Naïve Bayes-Kernel Technique

A grid application was used to determine the optimum relationship between the number of kernel (k) and bandwidth. Gaussian kernel function was used for kernel number selection. Bandwidth was chosen to minimize the means square error (MSE). The best accuracy performance was obtained with a bandwidth of 0.1 and a kernel value of 2.
3.3.2. k-Nearest Neighbor

The most effective factor for the k-nearest neighbor technique's accuracy performance is the value of k. The value that maximizes the performance ratio of the technique is chosen for k. The k value has been tried from 1 to 20, but only some few intermediate values are shown in Table 4. The performance ratio was not change for values after k = 13. The best accuracy value was reached for k = 3. For larger k values, general and class accuracy rates decreased. Because dataset was not containing nominal data numerical measure and Euclidean distance was used for finding nearest neighbor.

3.3.3. Decision Tree

The highest entropy value calculated for 3 variables was found as the temperature variable. Therefore, according to the information gain criterion, the root node was determined as the temperature. This rule has been followed up from the top node to the bottom node, and the humidity variable has been pruned (Figure 3).

3.3.4. Neural Network

Learning data set consists of variables and fault types. The network structure has 3 inputs are; temperature, pressure and humidity and has 4 outputs are; Fault Type1, Fault Type2, Fault Type3 and Faultless. In order to minimize the mean squared error of training and testing [46], the number of hidden layer neuron was determined to be 6. The parameters and functions are used in network is shown Table 5. The best learning rate is achieved with 200 training cycles.
3.3.5. **Multi-Layer Perceptron**

Both the number of training cycles used for the neural network and for MLP training were chosen 10. The number of MLPs per ensemble and number of folds of cross validation was chosen 10 and 4, respectively and sampling type was determined as stratified.

3.3.6. **Deep Learning**

Tangent function was chosen as the activation function to be used by the neurons. The hidden layer size in the model was determined 50 and the epochs which means the iteration number of dataset was obtained as 10.

4. **DISCUSSION**

In the study, six machine-learning techniques (NB-k, KNN, DT, ANN, MLP and DL) have been compared. The comparison criterion was chosen as classification accuracy, SE, R² and RMSE. Performance evaluation of the classification techniques is given in Table 6.

Techniques with values greater than 60% based on R² were considered. These are ANN, MLP and NB techniques. Among them, ANN with the highest classification accuracy (97.43 +/- 1.14) and with the lowest SE (0.023 +/- 0.007) and with the lowest RMSE (0.150 +/- 0.024) was selected as the most successful technique. It is also necessary to note that the performances of the other techniques are also close to each other.

5. **CONCLUSION**

The aimed study contains monitoring the process by multivariate control chart and then classification the fault types by machine learning techniques. The proposed model was implemented in a painting process in automotive Supplier Company. Paint quality was evaluated according to temperature, pressure and humidity variables of process. It was aimed
to determine which classes as fault type 1, fault type 2, fault type 3 or faultless of each sample that taken from the process is belongs to. For this, the class of sample was predicted by six different machine learning techniques as NB-k, KNN, DT, ANN, MLP and DL. The accuracy and error performance of the techniques were compared. According the accuracy %, squared error and root-mean squared error the best technique is ANN. However, the performance of techniques are close to each other. It can be said that the use of machine learning techniques are appropriate for prediction and classification problems such as human error, machine error. Quality improving will be provided by correct classifications of fault type with machine learning techniques. Furthermore, for the new products, time is saved by finding out which class the product belongs to without the need for control chart.

The contribution of study to literature;

- In the multivariate control chart, a large number of machine learning techniques were applied and the performance were compared.
- Unlike other studies, it does not only determine the variable that caused the fault, but also the value (large, normal, and high) that the variable takes. In this way, the correct determination of corrective actions will be easier to determine and products will be prevented from being faulty.

As a future direction of our study, we plan to use the ensemble methods like, boosting, bagging and vote to increase the performance and we also think to extend the model by considering the uncertainty in the data. Therefore, modelling the current problem by utilizing Neutrosophic sets [47] or Pythagorean fuzzy sets [48] are the potential research areas in which the uncertainty problem can be handled.

REFERENCES
[1] Hotelling, H. “Multivariate Quality Control Illustrated by Air Testing of Sample Bombsights”, *Techniques of Statistical Analysis*, 2(5), pp.111-152 (1947).

[2] Woodall, W.H., and Ncube, M.M. “Multivariate Cusum Quality Control Procedures,” *Technometrics*, 27(3), pp. 285-292 (1985).

[3] Lowry, A.A., Woodall, W.H., Champ C.W. et al.”A Multivariate Exponentially Weighted Moving Average Control Chart”, *Technometrics*, 34(1), pp.46-53 (1992).

[4] Murphy, B.J. “Selecting out of Control Variables with The T$^2$ Multivariate Quality Control Procedure”, *The Statistician*, 36, pp.571-583 (1987).

[5] Mason, R.L., Tracy, N.D. and Young, C.H. “Decomposition of T$^2$ for Multivariate Control Chart Interpretation”, *Journal of Quality Technology*, 27(2), pp. 99–108 (1995).

[6] Kourtis, T. and MacGregor J.F. “Multivariate SPC Methods for Process and Product Monitoring”. *Journal of Quality Technology*, 28(4), pp.409-428 (1996).

[7] Li, J., Jin, J. and Shi, J. “Causation-Based T$^2$ Decomposition for Multivariate Process Monitoring and Diagnosis”, *Journal of Quality Technology*, 40(1), pp. 46-58 (2008).

[8] Midany, T.E., Baz, A.A.E. and Elwahed, M.S. A. “A Proposed Framework for Control Chart Pattern Recognition in Multivariate Process Using Artificial Neural Networks”, *Expert Systems with Applications*, 37(2), pp.1035-1042, (2010).

[9] Addeh, A., Khormali, A. and Golilarz, N. A. “Control Chart Pattern Recognition Using RBF Neural Network with New Training Algorithm and Practical Features”, *ISA Transactions*, 79, pp.202–216 (2018).

[10] Wang, X. “Hybrid Abnormal Patterns Recognition of Control Chart Using Support Vector Machining”, *International Conference on Computational Intelligence and Security*, pp.238-241 (2008).
[11] Li, T.L., Hu, S., Wei, Z. et al. “A Framework for Diagnosis the Out of Control Signals in Multivariate Process Using Optimized Support Vector Machine”, Mathematical Problems in Engineering, pp. 1-9 (2013).

[12] Xanthopoulos, P. and Razzaghi, T. “A Weighted Support Vector Machine Method for Control Chart Pattern Recognition”, Computers & Industrial Engineering, 70, pp.134–149 (2014)

[13] Wang, C.H., Dong, T.P. and Kuo, W. “A Hybrid Approach for Identification of Concurrent Control Chart Patterns”, Journal of Intelligent Manufacturing, 20(4), pp.409–419 (2009).

[14] Lu, C.J., Shao, Y.E and Li, P.H. “Mixture Control Chart Patterns Recognition Using Independent Component Analysis and Support Vector Machine”, Neurocomputing, 74 (11), pp.1908-1914 (2011).

[15] Zhang, M. and Cheng, W. “Recognition of Mixture Control Chart Pattern Using Multiclass Support Vector Machine and Genetic Algorithm Based on Statistical and Shape Features”, Mathematical Problems in Engineering, pp.1-10 (2015).

[16] Chen, L.H. and Wang, T.Y. “Artificial neural networks to classify mean shifts from multivariate $\chi^2$ chart signals”, Computers and Industrial Engineering., 47(2-3), pp.195-205 (2004).

[17] Niaki, S.T.A. and Abbasi, B. “Fault Diagnosis in Multivariate Control Charts Using Artificial Neural Networks”, Quality and Reliability Engineering International, 21, pp.825-840 (2005).

[18] Aparisi, F. and Sanz, J. “Interpreting the Out-of-Control Signals of Multivariate Control Charts Employing Neural Networks”, International Journal of Computer and Information Engineering, 4(1), pp.24-28 (2010).
[19] Atashger, K. and Noorossana, R. “An Integrating Approach to Root Cause Analysis of A Bivariate Mean Vector With A Linear Trend Disturbance”, *The International Journal of Advanced Manufacturing Technology*, 52(1-4), pp.407-420 (2011).

[20] Masood, I. and Hassan, A. “Pattern Recognition for Bivariate Process Mean Shifts Using Feature-Based Artificial Neural Network”, *The International Journal of Advanced Manufacturing Technology*, 66(9-12), pp.1201–1218 (2013).

[21] Du, S., Lv, J. and Xi, L. “On-Line Classifying Process Mean Shifts in Multivariate Control Charts Based On Multiclass Support Vector Machines”, *International Journal of Production Research*, 50(22), pp.6288–6310 (2012).

[22] Guh, R-S. and Shiue, Y-R. “An Effective Application of Decision Tree Learning for On-line Detection of Mean Shifts in Multivariate Control Charts”, *Computers and Industrial Engineering*, 55(2), pp.475-493 (2008).

[23] He, S., Wang, G.A., Zhang, M., et al. “Multivariate Process Monitoring and Fault Identification Using Multiple Decision Tree Classifiers”, *International Journal of Production Research*, 51(11), pp.3355–3371 (2013).

[24] Jiang, J. and Song, H., M. “Diagnosis of Out-of-control Signals in Multivariate Statistical Process Control Based on Bagging and Decision Tree” *Asian Business & Management*, 2(2), pp.1-6 (2017).

[25] Jadhav, S.D. and Channe, H.P. “Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques”, *International Journal of Science and Research*, 5(1), pp.1842-1845 (2016).

[26] Cheng, C.S. and Cheng, H.P. “Identifying the Source of Variance Shifts in The Multivariate Process Using Neural Networks and Support Vector Machines”, *Expert Systems with Applications*, 35(1-2), pp.198-206 (2008).
[27] Salehi, M., Bahreininejad, A. and Nakhai, I. “On-Line Analysis of Out-Of-Control Signals in Multivariate Manufacturing Processes Using a Hybrid Learning-Based Model”, Neurocomputing, 74(12-13), pp.2083-2095 (2011).

[28] Ceylan, Y., Usta, K., Yumurtacı, H. et al. “An ESR Study on 22, 4 Diaminotoluene Exposed to Gamma Rays and Application of Machine Learning”, Acta Physica Polonica A, 130(1), pp.184-187 (2016).

[29] Moore, A. and Zuev, D. “Internet Traffic Classification Using Bayesian Analysis Techniques”, Sigmetrics, 33(1), pp.50-60 (2005).

[30] Murakami, Y. and Mizuguchi, K. “Applying the Naive Bayes Classifier with Kernel Density Estimation to the Prediction of Protein-Protein Interaction Sites”, Bioinformatics, 26(15), pp.1841-1848 (2010).

[31] Kuter, S., Usul, N. and Kuter, N. “Bandwidth Determination for Kernel Density Analysis of Wildfire Events at Forest Sub-District Scale”, Ecological Modelling, 222(17), pp.3033-3040 (2011).

[32] Fix, E. and Hodges, J. “Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties”, Technical Report 4, USAF School of Aviation Medicine, Randolph Field, Texas, USA (1951).

[33] Williams, J. W. and Li, Y. “Comparative Study of Distance Functions for Nearest Neigbor”, Advanced Techniques in Computing Sciences and Software Engineering, Elleithy, Khaled (Ed.), pp.79-84 (2010).

[34] Kataria, A. and Singh, M. “A Review of Data Classification Using K-Nearest Neigbor Algorithm”, International Journal of Emerging Technology and Advanced Engineering, 3(6), pp.354-360 (2013).
[35] Yüksel, A.S, Çankaya, S.F. and Üncü İ.S. “Design of a Machine Learning Based Predictive Analytics System for Spam Problem”, Acta Physica Polonica A, 132(1), pp.500-504 (2017).

[36] Quinlan, J. R. “Improved Use of Continuous Attributes in C4.5”, Journal of Artificial Intelligence Research, 4, pp.77-90 (1996).

[37] Mitchell, T.M., “Machine Learning”, McGraw-Hill Science/Engineering/Math, 81, (1997).

[38] Tsai, K-M. and Luo, H-J. “An Inverse Model for Injection Molding of Optical Lens Using Artificial Neural Network Coupled with Genetic Algorithm”, Journal of Intelligent Manufacturing, 28(2), 473–487 (2017).

[39] Alpaydın, E. “Introduction to Machine Learning”, Second Edition, The MIT Press, England, pp.203-247 (2010)

[40] Hinton, G. E., Osindero, S. and Teh, Y-W. “A Fast Learning Algorithm for Deep Belief Nets”, Neural Computation, 18, pp.1527–1554 (2006).

[41] Vargas, R., Mosavi, A. and Ruiz, R. “Deep Learning: A Review”, Advances in Intelligent Systems and Computing, 5(2) (2017).

[42] LeCun, Y, Bengio, Y. and Hinton, G. “Deep learning,” Nature, 521, pp.436–444 (2015).

[43] Seker, S.E. and Ocak, I. “Performance Prediction of Roadheaders Using Ensemble Machine Learning Techniques”, Neural Computing and Applications, 31(4), pp.1103-116 (2019).

[44] Rodrigues, E.O., Pinheiro, V.H.A., Liatsis. P. et al. “Machine Learning In The Prediction of Cardiac Epicardial and Mediastinal Fat Volumes”, Computers Biology and Medicine, 89, pp.520–529 (2017).
Figure List

Figure 1: The proposed model
Figure 2: Hotelling $T^2$ control chart
Figure 3: Decision tree graph of process
Figure 2: Hotelling $T^2$ control chart

Figure 3: Decision tree graph of process

Table List

Table 1: Performance criteria
Table 2: Value ranges of variables
Table 3: Fault classes
Table 4: Accuracy ratios for k value selection
Table 5: The parameters and functions of network
Table 6: Comparing the performance of techniques

| Criteria      | Equation                                      | Notation                              |
|---------------|-----------------------------------------------|---------------------------------------|
| Accuracy      | $TP + TN$                                      | $TP: \text{true positive} \quad TN: \text{true negative}$ |
|               | $TP + TN + FP + FN$                            | $FP: \text{false positive} \quad FN: \text{false negative}$ |
|               | $\frac{TP + TN}{TP + TN + FP + FN} \times 100\%$ | $\text{sample number in the test set.}$  |
\[
SE = \sum_{i=0}^{n} (\theta_i - \hat{\theta}_i)^2
\]
\(\theta_i: \) actual values  
\(\hat{\theta}_i: \) predicted values

\[R^2 = \frac{EV}{TV}\]
EV: Explained variation  
TV: Total variation

\[
RMSE = \sqrt{\frac{\sum_{i=0}^{n} (\theta_i - \hat{\theta}_i)^2}{n}}
\]
\(\theta_i: \) actual values  
\(\hat{\theta}_i: \) predicted values

### Table 2: Value ranges of variables

| Variables     | Low(L) | Normal(N) | High(H) |
|---------------|--------|-----------|---------|
| Temperature (°C) | 19-20  | 20.1-21   | 21.1-22 |
| Pressure (N/m²)    | 10-14  | 14.1-20   | 20.1-26 |
| Humidity (%)       | 70-72  | 72.1-74   | 72.1-75 |

### Table 3: Fault classes

| Sample Number | Temperature | Pressure | Humidity | Fault Classes   |
|---------------|-------------|----------|----------|-----------------|
| 9             | N           | H        | L        | Fault Type 1    |
| 11            | N           | L        | N        | Fault Type 2    |
| 7             | L           | N        | N        | Fault Type 3    |
| 554           | N           | N        | N        | Faultless       |

### Table 4: Accuracy ratios for k value selection

| Overall Accuracy |
|------------------|
| k=3 | 96.05% +/- 2.54% |
| k=5 | 95.88% +/- 2.87% |
| k=7 | 95.89% +/- 2.42% |
| k=9 | 95.20% +/- 1.66% |
| k=11| 95.02% +/- 1.39% |
| k>13| 95.19% +/- 1.01% |

### Table 5: The parameters and functions of network

| Parameters and Functions |
|--------------------------|
| Training function        | Levenberg- Marquardt    |
| Network type             | Feed forward back propagation |
| Transfer function        | Sigmoid                 |
| Training cycle           | 200                     |
| Total function           | Weighted sum            |

### Table 6: Comparing the performance of techniques

| Techniques | \(R^2\) | Accuracy % | SE     | RMSE    |
|------------|---------|------------|--------|---------|
| NB(k)      | 0.694 +/- 0.332 | 96.91 +/- 2.96 | 0.027 +/- 0.022 | 0.150 +/- 0.067 |
| k-NN       | 0.414 +/- 0.362 | 96.05 +/- 2.54 | 0.032 +/- 0.019 | 0.172 +/- 0.054 |
| DT         | 0.478 +/- 0.127 | 96.74 +/- 1.42 | 0.033 +/- 0.006 | 0.180 +/- 0.017 |
| ANN        | 0.653 +/- 0.116 | 97.43 +/- 1.14 | 0.023 +/- 0.007 | 0.150 +/- 0.024 |
| MLP        | 0.601 +/- 0.160 | 97.08 +/- 0.68 | 0.029 +/- 0.010 | 0.168 +/- 0.028 |
| DL         | 0.332 +/- 0.290 | 96.36 +/- 1.20 | 0.030 +/- 0.011 | 0.172 +/- 0.028 |
Deniz Demircioğlu Diren received the B.E. and M.E. degrees in Industrial Engineering from Sakarya University, Sakarya, Turkey, 2007 and 2011, respectively. She is research assistant and doctoral candidate in Industrial Engineering at Sakarya University. Her researches focus on Quality Management, Statistical Process Control, Artificial Intelligence, Machine Learning, Data Mining and Multi-Criteria Decision Making.

Semra Boran received the B.E and M.E. degrees in Industrial Engineering from Istanbul Technical University, Istanbul, Turkey and Ph.D. degree in Industrial Engineering from Yildiz Technical University, Istanbul, Turkey. She is Associate Professor in Department of Industrial Engineering, Sakarya University. Her current research interest includes Quality Management, Statistical Process Control, Machine Learning and Multi-Criteria Decision Making.

Ibrahim Cil is a Professor of Department of Industrial Reengineering at the Sakarya University. He holds a B.S. from Istanbul Technical University, and a M.S. from the Yildiz Technical University, and a Ph.D. in Industrial Engineering from Istanbul Technical University. His teaching and research specialties are in the fields of Decision Sciences, Economics, Multi criteria Decision Making, Computer Science, Decision support systems, Data Mining, Expert Systems and Artificial Intelligence, Engineering, Manufacturing strategies, Lean Manufacturing, Facility Planning and Business Reengineering.