Scientia Iranica E (2020) 27(6), 3233–3241

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
http://scientiairanica.sharif.edu

Integration of machine learning techniques and control charts in multivariate processes

D. Demircioğlu Diren*, **S. Boran, and I. Cil**

Department of Industrial Engineering, Sakarya University, Sakarya, 54050, Turkey.

Received 31 January 2018; received in revised form 25 March 2019; accepted 16 July 2019

KEYWORDS
Multivariate control chart;
Naive Bayes-kernel;
K-nearest neighbor;
Decision tree;
Artificial neural networks;
Multi-layer perceptron;
Deep learning.

Abstract. Using multivariate control chart instead of univariate control chart for all variables in processes provides more time and labor advantages that are of significance in the relations among variables. However, the statistical calculation of the measured values for all variables is regarded as a single value in the control chart. Therefore, it is necessary to determine which variable(s) are the cause of the out-of-control signal. Effective corrective measures can only be developed when the causes of the fault(s) are correctly determined. The present study was aimed at determining the machine learning techniques that could accurately estimate the fault types. Through the Hotelling T^2 chart, out-of-control signals were identified and the types of faults affected by the variables were specified. Various machine learning techniques were used to compare classification performances. The developed model was employed in the evaluation of paint quality in a painting process. Artificial Neural Networks (ANNs) was determined as the most successful technique in terms of the performance criteria. The novelty of this study lies in its classification of the faults according to their types instead of those of the variables. Defining the faults based on their types facilitates taking effective and corrective measures when needed.

1. Introduction

In order to survive in the competitive environment of globalization, a company should take into consideration both quality and price factors. In this respect, the importance of statistical process control is felt more than ever with increase in the complexity of processes. The selection of the appropriate quality control charts by the companies will be of importance in controlling the processes and reducing variability. Generally, a process is comprised of more than one quality variable in real life. It is impossible to control several variables simultaneously while using univariate control charts. Therefore, control charts with multiple variables such as Hotelling T^2 [1], multivariate cumulative sum control chart [2], and multivariate exponentially weighted moving average [3] have been developed to evaluate multiple variables.

Multivariate control chart has an advantage. It takes into account the relationship among variables, thus saving time and labor. However, it cannot detect the cause of the out-of-control signal, since all quality variables are calculated as a single value on the control chart and this is considered a major disadvantage. This is why it is necessary to utilize some methods to find the variables that cause out-of-control signals. In the literature, several studies have been conducted that employ different methods and approaches to detecting the cause of out-of-control signals in multivariate processes. While some studies have focused on the unnatural pattern recognition in control charts, others

*. *Corresponding author. Tel.: 90-264-2955540*
E-mail addresses: ddemircioglu@sakarya.edu.tr (D. Demircioğlu Diren); boran@sakarya.edu.tr (S. Boran); icil@sakarya.edu.tr (I. Cil)

doi: 10.24200/sci.2019.50377.1667

emphasize the mean and variance. Methods used for detecting the variables can be divided into two categories of statistical and machine learning techniques. The statistical methods are used to identify several variables that cause out-of-control signals, including discriminant analysis [4]; Mason, Young, and Tracy (MYT) decomposition approach [5]; principal component analysis [6]; and causation-based T^2 decomposition [7]. These methods are not able to predict unexpected new situations. To this end, it is useful to employ machine learning techniques that can predict the unexpected new situations by learning from the historical data. Artificial Neural Networks (ANNs) [8,9], Support Vector Machine (SVM) [10–12], and hybrid methods [13–15] consider pattern recognition for detecting the variables that cause out-of-control signals in multivariable control charts. Moreover, ANNs [13–20], SVM [21], Decision Tree (DT) [22–24], K-Nearest Neighbor (K-NN) [25], and hybrid methods [26,27] are mostly used in studies considering mean, variance, or both for detecting the variable(s) that cause out-of-control signals in multivariate control charts.

The main objective of this study is to classify the out-of-control signals based on fault types that occur in a multi-variable process by the most appropriate machine learning techniques. To this end, ANNs, DT, K-NN, Naive Bayes-kernel (NB-k), Multi-Layer Perceptron (MLP), and Deep Learning (DL) are compared, which are commonly used in real life processes [24].

In addition, this study is novel in that it classifies out-of-control signals based on fault types instead of variable types. Thus, experts determine the range of values for variables, i.e., high, medium, and low, which cause the fault. As the range in which the fault occurs is determined, the corrective measures with respect to the variables are immediately taken.

The intended model is used to determine the process variables affecting the quality of paint in the painting processes. Six different machine learning techniques are compared based on four performance criteria including classification accuracy, Squared Error (SE), squared correlation (R^2), and Root Mean Squared Error (RMSE). ANNs was found the most successful technique in detecting the fault type for a new sample in the process.

The paper is organized as follows: the proposed model and the methods used in the study are presented in Section 2. The results of the application are given in Section 3. The methods are compared in Section 4. Finally, the study is concluded in Section 5.

2. Materials and methods

2.1. Hotelling T^2 control chart

Hotelling T^2 control charts, presented by Hotelling (1947), are developed for simultaneous monitoring of the 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 $X_1, X_2, \cdots, X_p$ are p-correlated quality characteristics (variables), the parameters are unknown, the control chart is formed by historical sample data, and the sample size is one. Then, T^2 is given in Eq. (1):

$$T^2 = \left(X_i - \bar{X}\right)' (S)^{-1} \left(X_i - \bar{X}\right), \tag{1}$$

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

The Upper Control Limit (UCL) and Lower Control Limit (LCL) are given in Eqs. (2) and (3), respectively:

$$UCL = \frac{(m-1)^2}{m} \beta_{\alpha,p/2,(m-p-1)/2}, \tag{2}$$

$$LCL = 0, \tag{3}$$

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

2.2. Machine learning

Machine learning is a technique functioning on the basis of logical or binary operations for modeling a problem. 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 the input and output are predefined. However, unsupervised machine learning techniques do not include these values.

In this study, NB-k, K-NN, DT, ANNs, MLP, and DL classification and prediction techniques are used as subcategories of machine learning techniques.

2.2.1. Naive Bayes-kernel (NB-k) technique

NB-k is a simple classifier based on Bayesian theorem. It considers each class independent from others. NB-k techniques determine the conditional probability for the relationship between each variable and class. It is used when the values are high. In this regard, Bayesian theorem is presented in Eq. (4):

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}, \tag{4}$$

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

It was observed that the accuracy rates of the Naive Bayes increased when implemented in kernel density function [29,30].

In kernel density function, the bandwidth is determined after determining the kernel number. There are many different methods for determining the bandwidth. Otherwise, bandwidth is determined according to an expert's opinion [31].

2.2.2. K-Nearest Neighbor (KNN) technique

KNN, which was proposed in 1951, is one of the simplest pattern recognition methods functioning on the given value for k by the nearest neighbors' class. 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 dataset are determined by the historical data. Each sample in the dataset to be classified in the test phase is processed individually. In order to determine the class of each sample, k, which is the number of neighbors in an unknown sample, must be determined first. The K-NNs are determined based on some distance functions such as Euclidean, Manhattan, Mahalanobis, and Minkowsk. Euclidean is generally preferred over the others. In a comparative study of distance measurement methods, Mahalanobis was found to have the best performance [33]. However, in this study, Euclidean Distance (EUD) is preferred, because it is appropriate for the Gaussian distribution and easy to use [34]. This function is given in Eq. (5), showing the straight distance from $p = (p_1, p_2, \cdots, p_n)$ to $q = (q_1, q_2, \cdots, q_n)$:

$$EUD(p, q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}. \tag{5}$$

2.2.3. Decision Tree (DT) learning technique

The DT that predicts the target variable through different input variables is a widely used classifier among machine learning techniques [35]. There are three types of nodes in a tree, namely the root, non-terminal, and leaf.

DT starts with the root node determined by entropy criterion. The non-terminal and leaf nodes are separated to start with the highest entropy value. The entropy formula is shown in Eq. (6):

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i), \tag{6}$$

where, X is a discrete random variable that can take n possible values $(x_1 \cdots x_n)$.

The root, non-terminal, and leaf nodes represent all training cases and their subsets. The root and non-terminal nodes include a variable test value. The training cases are divided into two or more subsets according to the results of the test. The tree is pruned by removing the branches with low statistical validity. The obtained results of this classifier are feasible to understand and interpret [36]. The classification rules consist of routes, each leading from the root node to a leaf node. In addition, the tree is more comprehensible via the if-then rules. Therefore, it is preferred over difficult interpretation techniques even if it is less successful.

2.2.4. Artificial Neural Networks (ANNs) technique

ANNs technique, one of the most effective learning methods known today, is a robust approach to estimating a problem by learning how to interpret real-world data [37]. ANNs is a calculation model used to simulate the human nerve cells [38].

A network is formed by input, hidden, and output layers. The output of each neuron is computed by the weights of the nodes in the previous layer. Transfer functions such as sigmoid, tangent-sigmoid (tansig), and logarithmic-sigmoid (logsig) are applied to the input of the hidden node to determine the results.

To train ANNs, the dataset is divided into two parts, namely training and test data. ANNs must be trained through some learning techniques such as Levenberg-Marquardt backpropagation (trainlm) and quasi-Newton backpropagation (trainbfg) to achieve the best result. Back-Propagation Neural Networks (BPNNs) is the most popular type among all neural networks. BPNNs employs a supervised learning method. The training data are randomly selected from combinations of inputs and outputs and the data are used for testing. A well-trained ANNs model is able to define a relationship among inputs and outputs, even without a mathematical relationship. In case errors reach their minimum, the process stops. Otherwise, it is modified to connection weights to obtain the desired results.

2.2.5. Multi-Layer Perceptron (MLP)

MLP is a technique for attaining nonlinear decision nodes. It is an ANNs structure that can be used for classification and regression. If used for classification, the MLP can apply non-linear discriminators and then, approximate the nonlinear functions of the input for regression. It comprises the input, hidden, and output layers. The input layer transmits the inputs from the external world to the hidden layer. Then, this information is transmitted to the next layer when processed. There can be more than one hidden layer. Finally, the information is sent to the output layer [39].

2.2.6. Deep Learning (DL)

DL, beginning to function as a part in machine learning in 2006, includes multiple hidden layers of ANNs [40]. It takes into account non-linear processing in multiple layers and controlled or uncontrolled learning factors. The technique operates through taking the output of the previous layer as input [41]. It was proved to

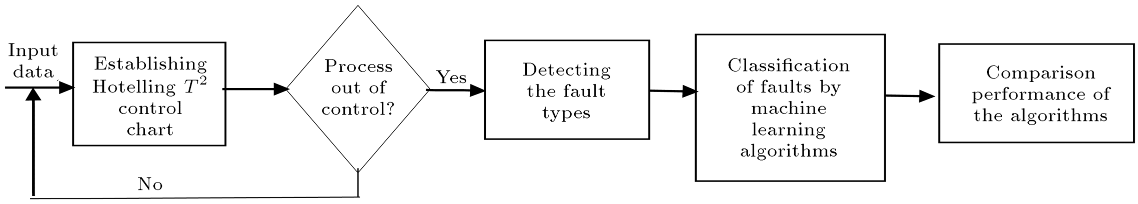

Figure 1. The proposed model.

be successful in solving complex structures and thus, applicable 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

Different techniques are compared, using several performance criteria, to determine which machine learning technique is the most appropriate for the process. In addition to accuracy, which is the most frequently used criterion in the literature, other criteria considered in studies are precision/recall, Receiver Operating Characteristic (ROC) curve area, SE, correlation, etc. [43,44].

To evaluate and compare the predictions of the performances of techniques, four performance criteria were used in this study including accuracy, SE, squared correlation coefficient (R^2), and RMSE. The equations of the performance criteria are presented in Table 1.

2.2.8. Proposed model

This study aims to determine the fault types from the historical data and determine which type of fault the new samples coming from the process belong to. The proposed model consists of four steps:

- Establishing a Hotelling T^2 control chart by historical measurement values to identify samples that make the process out-of-control;
- Detecting the fault types that occur in the process and then, considering the classes according to these fault types;
- Classifying the fault types of the sample in data set with machine learning techniques;
- Comparing the techniques according to performance criteria such as accuracy, SE, etc.

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

3. Results

In this study, machine learning techniques were applied to the painting process of an automotive supplier company, which produced chairs, door panels, and bumper products. The paint quality of the door panel was also analyzed. The dyed parts were dried and fixed in a drying cabinet. The most significant variables affecting this process were the cabinet temperature, pressure, and humidity.

3.1. Establishment of Hotelling T^2 control chart

The quality of a part was determined based on the variables and the data collected about 581 samples in the process. The sample size was assumed on the scale of one.

To examine the quality status of the sample values, the Hotelling T^2 control chart was used, as depicted in Figure 2. Each point on the control chart embodied the values of all the sample variables, as shown in Eq. (1). Control limits were calculated using Eqs. (2) and (3).

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

3.2. Determination of fault types

The causes of out-of-control signals must be accurately determined to eliminate the faults. In order to make it easier to identify these causes, it is necessary to determine the types of faults. To this end, the range of values for variables was determined by quality experts through taking into consideration the evaluated sample data. The value ranges are presented in Table 2. The

Table 1. Performance criteria.

Criterion	Equation	Notation
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}100\%$	TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative sample number in the test set.
SE	$\sum_{i=0}^{n}\left(\theta_i-\hat{\theta}_l\right)^2$	θ_i: Actual values; $\hat{\theta}_l$: predicted values
R^2	EV/TV	EV: Explained Variation; TV: Total Variation
RMSE	$\sqrt{\frac{\sum_{i=0}^{n}\left(\theta_i-\hat{\theta}_l\right)^2}{n}}$	θ_i: Actual values; $\hat{\theta}_l$: Predicted values

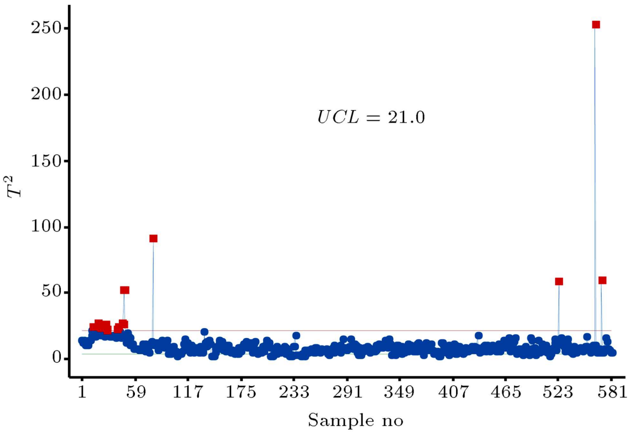

Figure 2. Hotelling T^2 control chart.

Table 2. Value ranges of variables.

Variable	Low (L)	Normal (N)	High (H)
Temperature (°C)	19–20	20.1–21	21.1–22
Pressure (N/m^2)	10–14	14.1–20	20.1–26
Humidity (%)	70–72	72.1–74	72.1–75

encountered fault types and the number of faultless samples in the dataset were identified according to the value ranges shown in this table.

Different fault types could appear in a painting process as a result of the drying cabinet atmosphere. The value ranges of variables and the consequent fault types are shown in Table 3. If the pressure was kept high and the humidity low, then the air condition of the cabinet was not suitable for the mixture of paint. Consequently, the painted part, which was defined as Fault Type 1, would be exposed to paint sag. In the presence of dust in the low-pressure cabinet, Fault Type 2 occurred, which was defined as observing dust on the surface of the product. If the temperature of the cabinet was not high enough, scratches were formed because the part was not completely dry, thus the formation of Fault Type 3.

Samples of Fault Type 1 were the 3rd, 4th, 5th, 6th, 7th, 8th, 9th,10th, and 45th; those of Fault Type 2 were the 15th, 17th, 18th, 19th, 20th, 21st, 24th, 27th, 28th, 29th, and 33rd; and samples of Fault Type 3 were the 47th, 48th, 63rd, 79th, 524th, 564th, and 571st. Other samples of data were faultless.

In this study, unlike other studies, the root cause was not determined only by variables and the value ranges for the variables were also determined. For example, not only temperature was found as a variable causing error type 3, but also its range was determined, which was low. By specifying the value ranges, decisions about corrective actions could be easily made and the process of improvement was accelerated.

3.3. Implementation of machine learning techniques

Supervised machine learning techniques are used to classify the process faults. NB-k, KNN, DT, NN, MLP, and DL are classification and prediction techniques for detecting the fault class in the process dataset. Rapid miner Studio 7.6 is also implemented to apply the techniques.

The performance criteria described in the previous section were compared to find the most appropriate technique for the dataset. In this study, the models were tested by cross validation. Cross validation divides the dataset into selected numbers and treats the portions as the training data. Then, the technique is repeated for a specific number of times, each time with different test data. The average of the accuracy rates obtained at the end of each classification ensures the overall accuracy of the technique. According to the previous studies, the number of cross validation folds was estimated at 10 [45]. By using stratified-sampling, the same rate was obtained for training and testing each time.

3.3.1. NB-k technique

A grid application was used to determine the optimum relationship between the number of kernels (k) and bandwidth. Gaussian kernel function was used for kernel number selection. Bandwidth was chosen to minimize the Mean Squared Error (MSE). The best accuracy was achieved with a bandwidth of 0.1 and a kernel value of 2.

3.3.2. K-NN

The most effective factor in the accuracy of the K-NN technique is the value of k. The value that maximized the performance ratio of the technique was selected for k. The k value has been tried for 1 to 20 times, but only some few intermediate values are shown in Table 4. The performance ratio did not change for values after $k = 13$. The best accuracy was reached for $k = 3$. With

Table 3. Fault classes.

Sample number	Temperature	Pressure	Humidity	Fault class
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.

k value	Overall accuracy
$\mathbf{k=3}$	**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%

larger k values, the general and class accuracy rates decreased. Since the dataset did not contain nominal data, numerical measure and EUD were used to find the nearest neighbor.

3.3.3. DT

The highest entropy value calculated for three variables belonged to temperature. Therefore, according to the information gain criterion, the root node was determined as temperature. This rule was followed from the top node to the bottom and the humidity variable was pruned as shown in Figure 3.

3.3.4. Neural network

The learning dataset consisted of variables and fault types. The network structure was comprised of three inputs including temperature, pressure, and humidity and four outputs including Fault Type 1, Fault Type 2, Fault Type 3, and faultless. In order to minimize the mean SE for both training and testing [46], the number of hidden layers for the neurons was set to six. The parameters and functions used in the network are shown in Table 5. The best learning rate was achieved with 200 training cycles.

3.3.5. MLP

The number of training cycles used for both neural network and MLP training was assumed to be 10. The number of MLPs per ensemble and that of cross

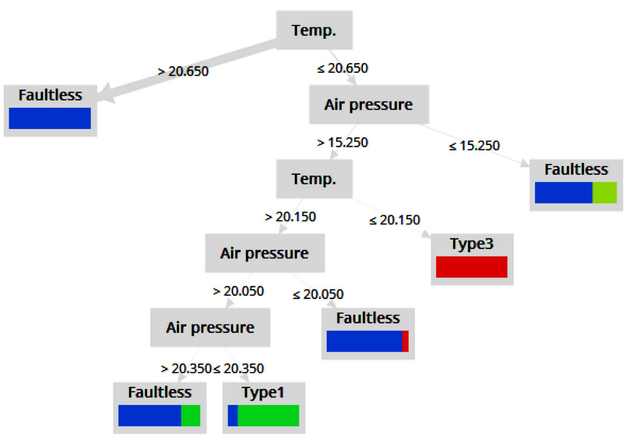

Figure 3. Decision Tree (DT) graph of the process.

Table 5. The parameters and functions of network.

Parameters and functions	
Training function	Levenberg-Narquardt
Network type	Feed forward back propagation
Transfer function	Sigmoid
Training cycle	200
Total function	Weighted sum

validation folds were considered 10 and 4, respectively, and the sampling type was determined as stratified.

3.3.6. DL

Tangent function was selected as an activation function used by neurons. The size of hidden layers in the model was determined 50 and the epochs, i.e., the number of iterations for the dataset, was considered 10.

4. Discussion

In this study, six machine learning techniques, including NB-k, KNN, DT, ANNs, MLP, and DL, were compared. The criteria considered for comparison were classification accuracy, SE, R^2, and RMSE. Performance evaluation of the classification techniques is illustrated in Table 6.

Techniques with values greater than 60% based on R^2 included ANNs, MLP, and NB. Among them, ANNs with the highest classification accuracy (97.43 + / − 1.14), the lowest SE (0.023 + / − 0.007), and the lowest RMSE (0.150 + / − 0.024) was selected as the most successful technique. Of note, the performances of other techniques were quite the same.

5. Conclusion

The main objectives of this study were monitoring processes by multivariate control chart and then, classifying the fault types using machine learning techniques. The proposed model was applied to a painting process in an automotive supplier company. The paint quality was evaluated according to some variables of the process, i.e., temperature, pressure, and humidity. In this study, we tried to specify the classes and categorize them into Fault Type 1, Fault Type 2, Fault Type 3, or faultless for each sample of the process. To this end, the sample classes were predicted by six different machine learning techniques including Naive Bayes-kernel (NB-k), K-Nearest Neighbor (KNN), Decision Tree (DT), Artificial Neural Networks (ANNs), Multi-Layer Perceptron (MLP), and Deep Learning (DL). The accuracy and error of the techniques were compared. In terms of accuracy, Squared Error (SE), and Root Mean Squared Error (RMSE), ANNs was found the best technique. However, the performances of the techniques were almost the same. In addition,

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
ANNs	**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

machine learning techniques were used for prediction and classification of problems such as human and machine errors. Correct classification of fault type ensured quality improvement through machine learning techniques. Furthermore, for the new products, time was saved by finding out to which class it belonged without the need for control chart.

The present study contributes to the related literature in the following manners:

- In the multivariate control chart, a large number of machine learning techniques were applied and their performances were compared;
- Unlike other studies, it determined not only the variable that caused the fault, but also the ranges for the values (large, normal, and high) the variable could take. In this respect, corrective actions were easily taken and hence, the products would be faultless.

As a direction for future studies, the ensemble methods such as boosting, bagging, and vote can be employed to improve the performance and the model can be developed by considering uncertainty in the data. Therefore, modelling the current problem via neutrosophic sets [47] or Pythagorean fuzzy sets [48] are the potential research areas in which the uncertainty problem can be taken into consideration.

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. Kourti, 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., and Liao, Z. "A framework for diagnosis the out of control signals in multivariate process using optimized support vector machine", *Mathematical Problems in Engineering*, **2013**(2), 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*, **2015**(5), pp. 1–10 (2015).

16. Chen, L.H. and Wang, T.Y. "Artificial neural networks to classify mean shifts from multivariate χ^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), pp. 473–487 (2017).

39. Alpaydın, E., *Introduction to Machine Learning*, Second Edn., 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), pp. 53040–53065 (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).

45. Haji, M.M. and Katebi, S.D. "Machine learning approaches to text segmentation", *Scientia Iranica, A*, **13**(4), pp. 395–403 (2006).

46. Rabiei, A., Naja-Jilani, A., and Zakeri-Niri. M. "Application of neural network models to improve prediction accuracy of wave run-up on antifer covered breakwater", *Scientia Iranica, A*, **24**(2), pp. 567–575 (2017).

47. Peng, X. and Dai, J. "A bibliometric analysis of neutrosophic set: two decades review from 1998 to 2017", *Artificial Intelligence Review*, **53**, pp. 199–255 (2020).

48. Peng, X. and Selvachandran, G. "Pythagorean fuzzy set: State of the art and future directions", *Artificial Intelligence Review*, **52**, pp. 1873–1927 (2019).

Biographies

Deniz Demircioğlu Diren received her BE and ME degrees in Industrial Engineering from Sakarya University, Sakarya, Turkey, in 2007 and 2011, respectively. She is currently a research assistant and doctoral candidate in Industrial Engineering at Sakarya University. Her research interests are quality management, statistical process control, artificial intelligence, machine learning, data mining, and multi-criteria decision making.

Semra Boran received her BE and ME degrees in Industrial Engineering from Istanbul Technical University, Istanbul, Turkey, and PhD degree in Industrial Engineering from Yildiz Technical University, Istanbul, Turkey. She is currently an Associate Professor in the Department of Industrial Engineering, Sakarya University. Her current research interests include quality management, statistical process control, machine learning, and multi-criteria decision making.

Ibrahim Cil is a Professor in the Department of Industrial Reengineering at the Sakarya University. He received his BSc from Istanbul Technical University, MSc from Yildiz Technical University, and PhD in Industrial Engineering from Istanbul Technical University. His teaching and research specialties are 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.