İŞLETME ARAŞTIRMALARI DERGİSİ
JOURNAL OF BUSINESS RESEARCH-TURK
2021, 13(4), 3728-3741
https://doi.org/10.20491/isarder.2021.1352

Fault Detection in Manufacturing Companies with Ensemble Machine Learning Method

Deniz DEMIRCIOGLU DIREN [a] **Semra BORAN** [b]

[a] Sakarya University, Distance Education Research and Application Center, Sakarya, Turkey.ddemircioglu@sakarya.edu.tr
[b] Sakarya University, Department of Industrial Engineering, Sakarya, Turkey. boran@sakarya.edu.tr

ARTICLE INFO

Keywords:
Multivariate control chart
Ensemble machine learning
Bagging
Boosting
Stacking

Received 20 October 2021
Revised 10 December 2021
Accepted 20 December 2021

Article Classification:
Research Article

ABSTRACT

Purpose – Multivariate control charts cannot be indicative of which variable is the cause of the out-of-control signal. To keep the process under control, the cause of the out-of-control signal must be determined correctly. The study, it is aimed to predict the variable that causes the out-of-control with the highest accuracy when there is 2 sigma and 3 sigma shift from the mean.

Design/methodology/approach – The method used in the study is machine learning-based detection analysis. The data set was taken from a company that produces furniture connecting part. Sample values were collected from the enterprise. Then the under-control samples were detected from these. According to these samples' mean and standard deviation values, data was produced in such a way that 2 sigma and 3 sigma shifts occur from the mean for training the machine learning algorithms. To predict the out-of-control samples three individual machine learning algorithms and three ensemble methods (Bagging, Boosting and Stacking) were used. In addition, 3 stacking models were developed using combinations of the individual algorithms.

Findings – When the results are examined, higher accuracy has been reached by using a model developed with the stacking method than individual algorithms. The highest accuracy rates have been achieved as 69.00% for 2 sigma and 85.75% for 3 sigma shift with the stacking 3 models developed based on the stacking method.

Discussion – The issue of detecting out-of-control signals in the quality processes in manufacturing companies with the least error was examined and the results were discussed.

1. Introduction

Examining the variables separately with traditional univariate control charts causes loss of time and the relationship between variables (Montgomery,2009). To remove these disadvantages, multivariate control charts should be used in a complex process where one than more variable affect. The most widely used multivariate control charts are Hotelling T^2 (Montgomery, 2009; Mason and Young, 2002; Arıoglu and Fırat, 2005). However, the values of all variables in the control chart are shown with a single value by statistical calculation. The shortcoming of such charts is that it is not possible to make comments about the variables separately by examining all variables together (Aparisi and Sanz,2010). Therefore, the variable that causes the out-of-control state is not easy to detect. However, in order to keep the process under control, it is necessary to quickly find the cause (s) that negatively affect the process. In addition to multivariate control charts, statistical or machine learning methods can be used to determine the cause of the out-of-control.

Statistical or machine learning methods are used to predict out of control causes in the literature. Some statistical methods are Mason Young Tracy (MYT) (Mason et al.,1995), principal component analysis (Edward-Jackson,1985; Bakdi et al.,2017) and discriminant analysis (Murphy,1987). The most widely used method among statistical methods is MYT decomposition method, which was developed specifically for Hotelling T^2 chart. MYT method tries to determine the cause of the out-of-control signal by investigating the correlation between dependent and independent variables (Mason et al.,1995). When the studies are examined, it is concluded that statistically it can predict the causes of out-of-control signals encountered in previous periods, but the error rate of the results is not known and there is no prediction for the future periods. These two issues are important shortcomings for the MYT method (Niaki and Abbasi, 2005).

Suggested Citation
Demircioglu Diren, D., Boran, S. (2021). Fault Detection in Manufacturing Companies with Ensemble Machine Learning Method, *Journal of Business Research-Turk*, 13 (4), 3728-3741.

Machine learning algorithms that can be used as an alternative to statistical methods for classifying out-of-control situations are very popular and overcome the deficiencies. These algorithms are advantageous because they learn from previous examples and decide on the result of a new sample and the accuracy of the prediction can be evaluated with performance criteria. It is very important to determine the causes of out-of-control situations in the process in the most accurate and fastest way. Machine learning algorithms that use many different classification and prediction models are important guides in this regard. There are studies in the literature with different perspectives and methods on this subject. For the analysis of out-of-control signals, ensemble methods by constructing the same algorithm in parallel or sequentially (Guh and Shiue, 2008; Yu et al., 2009; Alfaro et al., 2009; Du et al.,2012; Cheng and Lee, 2012; Yang, 2015; Jiang and Song, 2017; Asadi and Farjami, 2019) as well as individual machine learning algorithms (Aparisi and Sanz, 2010; Niaki and Abbasi, 2005; Chen and Wang, 2004; Cheng and Cheng, 2008; He et al, 2013; Li et al., 2013; Huda et al., 2014; Song et al.,2017; Shao and Lin, 2019) are popular in recent years. As is known, there are many different ensemble machine learning algorithms. Many studies have shown that ensemble algorithms are more successful than using individual machine learning algorithms (Guh and Shiue, 2008; Yu et al., 2009; Yang, 2015; Asadi and Farjami, 2019; Alpaydın, 2012). In addition to using the same algorithms in parallel or sequentially with boosting and bagging methods, using different algorithms together to overcome the deficiencies of the algorithms has recently become popular in different fields other than quality control (Oza and Tumer,2008). There are studies on quality control charts using Boosting (Yu et al.,2009; Alfaro et al.,2009; Du et al.,2012; Yang,2015; Asadi and Farjami,2019) and bagging methods (Cheng and Lee, 2012; Jiang and Song,2017). With the stacking method, there are studies about the students at academic risk (Lauría et al.,2018), such as to detect network intrusion (Demir and Dalkılık,2018), to estimate the warfarin dose (Ma et al.,2018).

Studies can be grouped according to the used techniques as well as the data used. In some studies, the out-of-control state is based on pattern recognition and the type of mean or variance shift (Zhao et al., 2017; Zhang and Cheng, 2015). Some studies are based on shift magnitude (Niaki and Abbasi, 2005; Chen and Wang, 2004; Li et al., 2013). Some studies aimed to determine the variable (s) of the cause of the out-of-control signal (Niaki and Abbasi, 2005; Cheng and Cheng, 2008; Shao and Lin, 2019; Gonzalez-de la Parra and Rodriguez-Loaiza, 2003).

In the study, a new machine learning methods based on stacking algorithm is proposed to determine the cause out of control signal in multivariate control signal in manufacturing companies. The success of the proposed model has been tested with a furniture fastener production real-life application. Estimates were obtained with models created by ensemble algorithms with the stacking algorithm in a heterogeneous manner in 3 different ways and the results were compared with all cases. The most successful algorithm will be selected by comparing the prediction accuracy rates of all individual and ensemble algorithms. However, according to our limited literature research, it has been observed that there is no work with stacking algorithm on quality control charts.

The study can be summarized as follows; In chapter 2, the methods used in the study are explained. In the third chapter, the model developed for the purpose of the study is included. In the chapter 4 , real life application is made on the model. Finally, in the chapter 5, the results of the study and recommendations are presented.

2. Methods

2.1. Hotelling T^2 Control Chart

In multivariate processes, it is more complex and difficult to detect shifts in mean and / or variance than in univariate processes. The measured values of all variables are converted to a single statistical value, T^2 presented in Equation 1 (Hotelling, 1947).

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

$\bar{X}$ and S, represent the sample mean vector and covariance matrix of these observations, respectively

$\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix}$ is quality vector.

The structure of the control chart consists of two phases. In the first phase stage, it is checked whether the process is under control. In the second phase, future production process is monitored. Equation 2 represents the first phase control limits.

$$ÜKS = \frac{m-1^2}{m}\beta_{\alpha,\frac{p}{2},\frac{(m-p-1)}{2}},$$
$$AKS = 0 \qquad (2)$$

2.2. Machine Learning

The method used in the study is machine learning-based detection analysis. Machine learning algorithms enable computers to learn from data and make inferences. It is divided into two as supervised and unsupervised learning. The label value is not specified in the unsupervised learning algorithms and makes its own inference. In supervised algorithms, the result is taught to the system by entering the label value (Alpaydın, 2012). Since input and target values are determined in the study, supervised algorithms will be used. For this reason, firstly individual machine learning algorithm then ensemble machine learning algorithm were used for comparing the prediction accuracy rates.

2.2.1. *Individual Machine Learning*

Decision Tree Algorithm: A decision tree is a flow chart-like tree structure in which each node represents a test on an attribute value, each branch represents a test result, and tree leaves represent classes or class distributions (Ali et al.,2015). Tree learning algorithms start from the root of the whole learning set and decide where to best divide each step. It explores how to best divide each section within itself, stops when no more divisions are needed, and forms a leaf node (Alpaydın,2012). It is a classifier and predictor commonly used among machine learning techniques that predict the target variable according to various input variables. Decision tree learning algorithms are used for supervised learning and allow you to see the logic for the interpretation of data that is not like black box algorithms. A mistake that can be easily made when working with decision tree learning is that when changing the tree for specific purposes, a tree that does not model the properties of all data can be encountered. This is an overfitting of machine learning (Mitchell,1997).

Naïve Bayes Algorithm: Naïve Bayes is simple statistical Bayesian classifier and algorithm for calculating the probability of hypotheses to estimate the values of unobserved variables based on Bayes theorem (Bhatia, 2010). Bayes' theorem is shown as in Equation 3 (Salvador-Meneses et al., 2019).

$$P(H|X) = \frac{P(X\,|\,H)P(H)}{P(X)} \qquad (3)$$

where, X is considered evidence, H is some hypothesis, $P(X)$ is the probability of the X (independent), $P(X)$ is the probability of H occurring when X is known to be, $P(H)$ is the prior probability, $P(X\,|\,H)$ is the posterior probability of X conditioned on H.

Naive Bayes classifiers analyze the relationship between each property and class to obtain a conditional probability for the relationship between attribute values and class (Mitchell, 1997), and at the end of comparisons, the high probability class becomes the predicted class label (Han et al., 2012).

K- Nearest Neighbor Algorithm: K- Nearest Neighbor (K-NN) is a non-parametric method used for classification and regression, and basically the points closest to the new point are searched. Also it is easy to run the K-NN algorithm. K represents the amount of the nearest neighbors of the unknown point. The K amount value of the algorithm must be selected to predict the results. A data is classified by majority vote of its neighbors. The distance between all training points to the sample point is calculated and the point with the least distance is called the nearest neighbor (Bhatia, 2010). There are some distance calculations to calculate the K-Nearest Neighbor, but for convenience and efficiency, the Euclidean distance is often used. The Euclidean distance is calculated as in Equation 4 (Salvador-Meneses et al.,2019).

$$d(X,Y) = \sqrt{\sum_{i=1}^{n} d_i(x_i, y_i)} \quad (4)$$

where $d_i(x_i, y_i)$ is the distance between two observations in the i-th attribute.

2.2.2. ***Ensemble Machine Learning***

Ensemble machine learning algorithms are available as an alternative to the use of individual machine learning algorithms. These are often seen to increase the success of individual methods (Yu et al., 2009; Yang, 2015; Breiman, 1996). The generalization ability of an ensemble, that is, multiple algorithms, is generally stronger than using basic algorithms individually. Because they can make weak algorithms that provide a little better classification than random predictions, with strong algorithms that can make very precise predictions. Therefore, individual algorithms are also called weak algorithms (Zhou, 2012). Bagging, Boosting and Stacking methods which are the methods of Ensemble machine learning used to create community algorithms. It is thought that instead of using the machine learning algorithms individually, using the ensemble methods will increase the predictive accuracy. However, when creating ensemble algorithms, it is important to consider the basic algorithm parameters used in the best way and to combine them correctly. Otherwise, besides not achieving successful results, time and cost losses will occur (Alpaydın,2012).

Bagging Method: The method generates multiple models by taking different samples from the data set and voting through the prediction results of the models to obtain a general prediction result (Breiman,1996). The bagging method developed by Breiman (1996) is the oldest, simple and effective ensemble algorithm (Zhang and Ma, 2012). Bagging aims to increase classification accuracy by producing a model that combines multiple classifiers of the same type (Rokach, 2010). In this algorithm, decisions are made by majority vote from different classifiers (Gowda et al., 2018). The bagging method, in which basic learners are combined in parallel, is a multi-class method (Zhou,2012).

Boosting Method: In boosting ensemble method, weights are given by estimating the samples, then the weights of the faulty samples are updated according to the prediction results of the previous model, a new model is established and classification estimates are obtained (Freund and Schapire,1999). In this method, which allows each algorithm to use the output of the previous algorithm as input by connecting the algorithms in series, each classifier is affected by the performance of the previous algorithm. It places more emphasis on classification errors made by classifier algorithms first (Rokach,2010). In other words, instead of voting equally for each classifier as in bagging, it is weighted based on the performances of each classifier (Lantz,2013).

Stacking Method: The stacking ensemble developed by Wolpert (1992) is a machine learning algorithm that combines multiple different individual algorithms and learns the outputs of the basic classifiers to achieve higher predictive accuracy (Wolpert,1992). Therefore, stacking algorithm was used in the study. The Stacking model consists of Base and Meta learner as seen in Figure 1. The meta model tries to learn where and how the base models make prediction errors and correct them. Therefore, the two models should be trained with different data. It is useful for base models to use different learning algorithms to complement each other, not same type. The meta model is more flexible and less biased because it is trained (Alpaydın,2012).

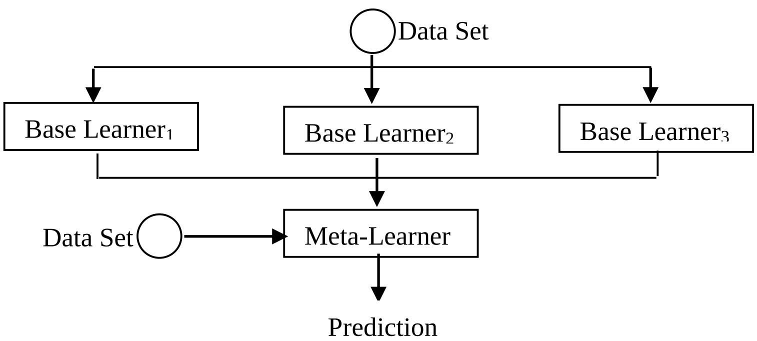

Figure1. Stacking Ensemble Model

Here, the basic individual algorithms are called the base learner, while the combiner is called the meta learner. The basic idea is to train the base learners using the original training data set and then create a new data set to train the meta learner, where the outputs of the base learners are considered as input features when in the original state. Base learners are often produced by applying different learning algorithms, and therefore, stacked communities are often heterogeneous (Zhou, 2012). The metadata set consists of predictions of all algorithms (Onan, 2018).

2.3. *Performance Criteria of Machine Learning Algorithms*

The performances of models built with machine learning algorithms should be evaluated according to certain criteria to predict the classification and predictive performance on future data (Raschka, 2018). There are many criteria to measure the performance of algorithms. However, the criteria to be used differ according to the intended use. When regression is used, measurements of errors are taken, and when classification is used, measurements of accuracy are taken (Zheng, 2015). Since the classification was generated throughout the study, accuracy and sensitivity were employed.

Accuracy is defined as the number of samples correctly classified from all samples. Accuracy is the most popular method and uses a very simple equation as shown in Equation 5 to calculate.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}\ 100\ \% \qquad (5)$$

Given that TP, TN, FP, FN are the number of True Positive, True Negative, False Positive and False Negative respectively.

3. Proposed Model

The model of the study consists of four steps.

- Collecting the data of the process and determining the out-of-control signals: Hotelling T^2 multivariate control chart is used to monitor the process.
- Generating of data: At this stage, to examine 2 sigma and 3 sigma shifts, data is generated by simulation from the mean according to the distribution, average and standard deviation of the data under control.
- Training the machine learning algorithms developed with individual (Decision Tree (DT), K-Nearest Neighbor (K-NN) and Naive Bayes (NB)) and Stacking method with these generated data for detecting the variable which cause of out-of-control signal: In order to determine the causes of out-of-control situations, the success of individual machine learning algorithms was first examined. Then, the same algorithms were combined homogeneously, and the prediction accuracies were investigated by using bagging and boosting ensemble algorithms.
- Selecting the algorithm with the highest prediction accuracy by comparing according to the classification accuracy performance criteria: Finally, estimates were obtained with models created by ensemble algorithms with the stacking algorithm in a heterogeneous manner in three different ways and the results were compared with all cases. The most successful algorithm will be selected by comparing the prediction accuracy rates of all individual and ensemble algorithms.

The architecture of the proposed model as shown in Figure2.

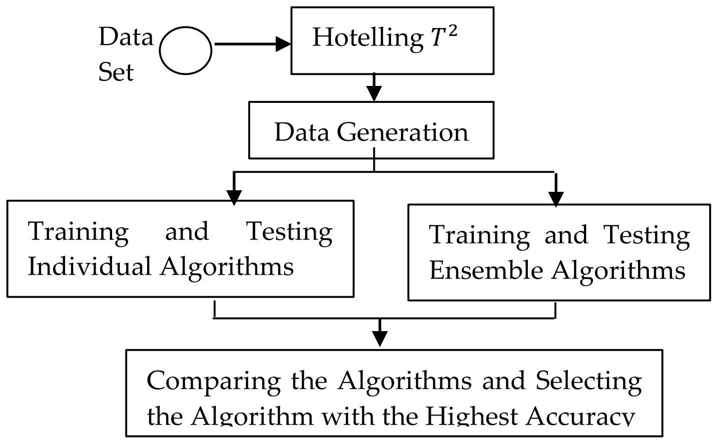

Figure2. The Architecture of the Proposed Model

In this study, the quality of furniture fastener production was examined. The production process of the aluminum part consists of drilling, cutting, and bending processes.

3.1. *Collecting Data and Hotelling T^2 Control Chart*

In order to prove the model in the study, the aluminum furniture connecting part, which is the most problematic in the business, has been examined. Furniture connecting part is presented in Figure 3.

Figure 3. Furniture Connecting Part

The out-of-control signals analyzed in the production of part were monitored with the Hotelling T^2 control chart and the causes of the out-of-control signals were detected. The variable namely quality characteristics examined in terms of quality during the production of the part were determined as x_1: diameter of the hole, x_2: bending angle and x_3: length of the part. Firstly, 1230 measured values of these quality characteristics collected in the process were evaluated in the Hotelling T^2 control chart by using Equation 1 and Equation 2. The out of control signals conditions were detected as seen in Figure 4 that obtained with MINITAB software.

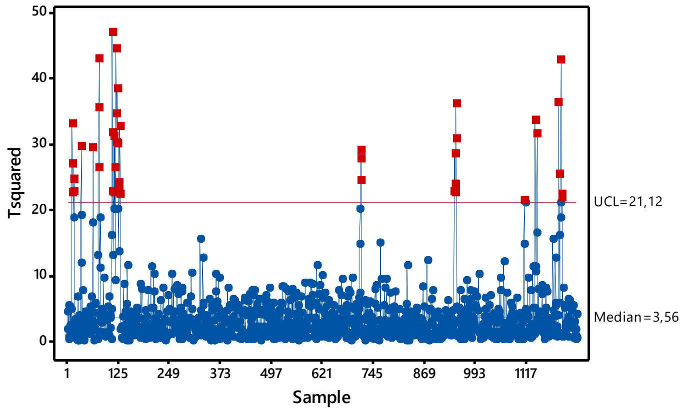

Figure 4. Hotelling T^2 Control Chart

Out of control signals can be easily seen through the chart, but it is not possible to see which variables group is the cause of the signal by looking at the chart. In this study, the causes of out-of-control signal were estimated by classification method. All possible out-of-control signal causes should be taught to the classification and prediction model. Therefore, based on the under-control samples, data were generated for all quality characteristics with 2 sigma and 3 sigma shifts from the process mean. For this reason, first 128 out of control sample that shown in Figure 3 were eliminated from data sets. When the remaining samples were examined, the samples under control were obtained as shown in Figure 5.

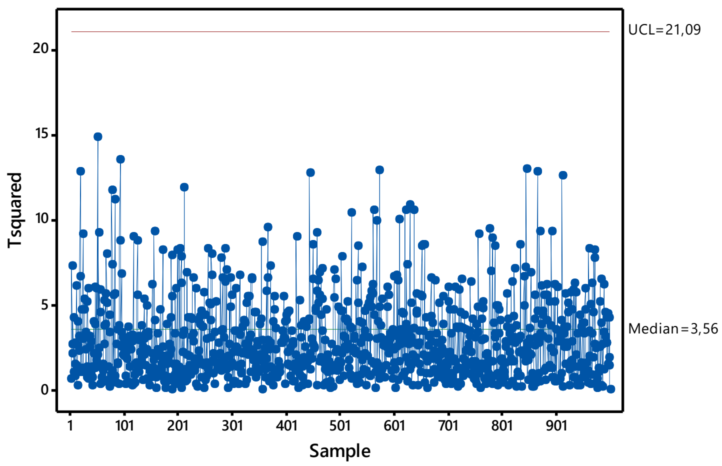

Figure 5. The Hotelling T^2 Control Chart of Under Control Samples of Process

3.2. *Generating Data*

For training the machine learning algorithms, data were generated according to the measurement values of the under-control samples' distribution (normal distribution), mean of x_1: 7.93, x_2: 85.680, x_3: 135.11 and standard deviation of x_1: 0.2636, x_2: 1.347, x_3: 4.70. Data generation was made by considering variable or variable groups that may cause out of control signal.

There are several classes of causes of out-of-control signal for classifying. The number of classes varies according to the number of variables. More clearly, the mean of a variable is either in-control or out of control. If there are p variables, there is a total of one in- control and 2^p-1 out of control case. For example, in a process with two variables (0,1), (1,0), (1,1), there are 3 cause classes (0: variable under control, 1: variable out of control). Since there are 3 variables in the study, there are 7 different cause classes as: (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1). Here for example (0,0,1) means that x_1 and x_2 is under control and x_3 is out of control.

Because of Hotelling T^2 investigates large shifts; data were generated with 2 and 3 sigma shifts from the mean for all variables according to the cause classes. For example, in the (0,0,1) class, since the third variable has a shift of 2 sigma, calculations have been made according to the mean values $\mu_1, \mu_2, \mu_3 + 2\sigma_3$ (7.93, 85.680, 135.11+2(1.347)). The same calculations were made for all out-of-control situation and the study was carried out over 350 data for 2 sigma and 3 sigma derived by simulation. Since it was desired to be determined in under control products in the study, 50 units of under control product data were added to the data sets and each 2 data sets containing 400 data in were obtained. The first data set was generated for 2 sigma shifts and the second data set for 3 sigma shifts.

3.3. *Machine Learning Algorithms Implementation*

***Individual Machine Learning Algorithms*:** The parameters that make the classification performances of individual machine learning algorithms the best has been estimated heuristic considering the previous information of the data set. The models were re-designed according to each parameter, trained and their results were obtained. Comparison of accuracy rates obtained from algorithms with the determination of appropriate parameter values.

Cross-validation method was used for the training stages of the models established with classification algorithms. Cross-validation parameters are determined based on the values that are decided to increase the performances by experimenting with basic individual algorithms. All models are trained using the same parameters to ensure consistent study.

The parameters used for cross validation are shown in Table 1.

Table 1. The Parameters of Cross Validation

Parameter	Value
Number of folds	10
Sampling type	Automatic

The number of folds is often taken as 10 in studies (Refaeilzadeh et al.,2009; Zhang et al.,2019; Jonathan, 2019; Karimi et al., 2015; Ramezan et al.,2019; Yu and Feng, 2014). Similarly, successful results were obtained with the number of folding 10 in the study. In addition, sampling type was automatically selected and folded sampling was used because the result values are nominal. Multi-class performance criteria are used to evaluate the classification estimation accuracy rates obtained by using that parameter value determined for all algorithms.

In the study, each prediction model was tested with cross validation. Cross validation divides the data into selected numbers to evaluate part of the set as test data. The method is then repeated with the selected number of different test data each time. The mean of the accuracy obtained at the end of each classification constitutes the overall accuracy of the method. The number for cross-validation is 10. Using automatic sampling, the same ratio was obtained for training and testing each time.

Algorithms operate according to specific parameters. The most suitable parameter values were determined for the algorithms used. The parameters and result values that make the best performance of the 3 basic individual machine learning algorithms used are presented below. Rapidminer Studio 7 was used for the results.

Decision Tree: The parameter values used for the decision tree can be seen in the Table 2. Similar to previous studies for the splitting process in the tree (Dreiseitl et al., 2001; Anwar et al., 2014), information gain was determined as criteria. The maximal depth is used to limit the depth of the tree and varies according to the size of the sample data set.

Table 2. The Parameters of Decision Tree

Parameter	Value
Criterion	Information gain
Maximal Depth	6
Confidence	0.1

16 values between 5 and 20 were tried and selected as 6. As an example, accuracy rates corresponding to maximum depth for 3 sigma shifts are shown in Table 3.

The confidence level parameter selected as (0.1) is a parameter used in the pessimistic error calculation of tree pruning. Some parameters other than those specified were also run with the default values of the program.

Table 3. Accuracy Rates Corresponding to Maximum Depth For 3 Sigma Shift

Maximal depth	Accuracy rate
5	78.75
6	**79.25**
7	79.25
8	79.25
9	79.25
10	79.25
11	79.25
12…………..20	73.50

***K-nearest neighbor**:* Since K-NN has both nominal and ordinal values in the data set, it is determined as mixed Euclidean distance from measure type mixed measures. The nearest neighbor number (k) used for classification was determined by testing. The nearest neighbor (k) was tried to be 1, 3, 5, 7, 9, 11 and13. The highest accuracy was found to be k = 1 and k = 3 then decreases as seen in Table 4.

Table 4. Accuracy Rates Corresponding to K for 3 Sigma Shift

K	Accuracy Rate
1	**67.00**
3	67.00
5	66.25
7	65.00
9	63.75
11	63.25
13	61.75

The parameters used for K-NN are shown in Table 5.

Table 5. The Parameters of K-NN

Parameter	Value
K	1
Mixed Type	Mixed Measure
Mixed Measure	Mixed Euclidean distance

The measure type parameter, which is used to determine the nearest neighbors, was chosen as the mixed measures Euclidean distance because the data set contains numerical values and mixed measure was chosen as mixed Euclidean distance that is the most frequently used (Hu et al.,2016).

Naive Bayes Algorithm: The classification in naive bayes algorithm is only based on one parameter, it is made according to Laplace correlations (Anwar et al., 2014). Laplace correlation was used for naive bayes and the number of kernels was found to be 1 most appropriate.

The prediction accuracy results of all individual machine learning algorithms, for 2 and 3 sigma shifts, by considering the selected parameters are shown in Table 6.

Table 6. Accuracy Rate of Individual Machine Learning Algorithms.

	2 Sigma Shift	3 Sigma Shift
Algorithm	Accuracy%	Accuracy%
DT	58.25	79.25
K-NN	41.75	67.00
NB	61.25	61.00

When the results are examined, it is seen that the algorithms achieve more successful detections in 3 sigma shift. The reason why the 2 sigma shift is detected with lower accuracies is that the 2 sigma shift corresponds to a slightly lower value. Therefore, the values produced are very close to each other and the classification of values that are so similar to each other is really complex.

3.3.2. ***Ensemble Machine Learning Algorithm***

Bagging and Boosting: Individual algorithms are combined with ensemble methods according to previously determined parameters. The number of iterations of bagging and boosting algorithms is taken as the default value of 10. The accuracy rates of bagging and boosting for 2 sigma and 3 sigma are shown in Table 7.

Table 7. Accuracy Rate of Bagging and Boosting Method

	2 Sigma Shift	3 Sigma Shift		2 Sigma Shift	3 Sigma Shift
Bagging	Accuracy%	Accuracy%	Boosting	Accuracy%	Accuracy%
DT-Bagging	58.25	76.00	DT-Boosting	52.75	76.75
K-NN-Bagging	40.00	67.00	K-NN-Boosting	40.00	67.00
NB-Bagging	60.25	60.00	NB-Boosting	60.75	61.25

When the results of the bagging and boosting methods are examined, it is seen that in the case of only 3 sigma shift on the data set, an improvement of 0.25 is achieved in the NB algorithm with the boosting method. Apart from this, it has been observed that the methods do not provide any improvement and the accuracy rates either decrease or remain constant.

Stacking: There are different options for combining individual algorithms with Stacking. Three individual algorithms separately used to select which algorithm was the main learner and the meta learner. As an example, a model is shown in Figure 6 and it has been developed similarly to other models.

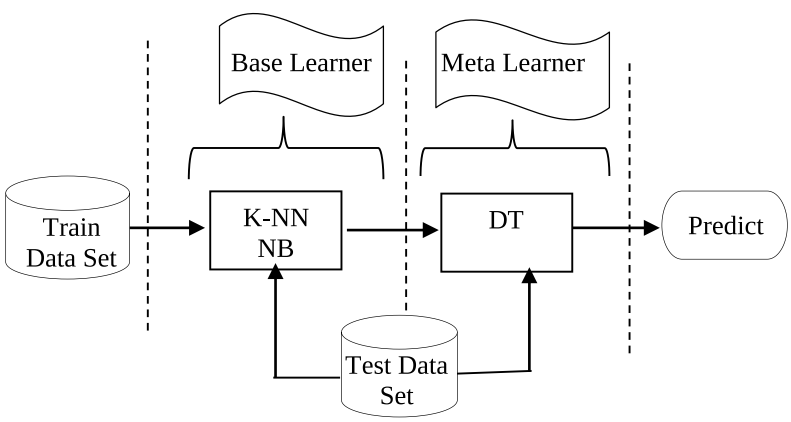

Figure 6. An Example of Stacking Models

The predicted accuracy of 2 sigma and 3 sigma shear conditions are shown in Table 8. To avoid a bias result, the same parameters were used for individual and stacking in all algorithms.

Table 8. Stacking Model Prediction Accuracy

Model	Base Learner	Meta Learner	2 Sigma Shift Accuracy%	3 Sigma Shift Accuracy%
Stacking1	DT+K-NN	NB	56.75	80.50
Stacking2	DT+NB	K-NN	49.25	75.50
Stacking3	K-NN+NB	DT	**69.00**	**85.75**

The highest accuracy of classification was achieved when the data trained with Staking 3 model. The fact that the same model provides high prediction accuracy according to two different shift amounts supports the validity of the model.

4. Result and Discussion

In the study, a production problem was carried out in a company producing furniture fittings. It is aimed to develop a system that will detect the variable that causes an out-of-control signal in a new production process, according to different shifts on mean, as soon as possible. The production process consists of drilling, cutting and bending processes. 3 variables (x_1: diameter of the hole, x_2: bending angle and x_3: length) affect the process simultaneously. The quality has been checked considering the shifts from the mean. For this, 50 data were generated by performing 2 sigma and 3 sigma shifts from the mean separately for each out-of-control situation. By adding under control product data, a data set containing 400 data for both cases was obtained. The two data sets were discussed separately and the prediction accuracies in the two cases were examined. Classification and prediction algorithms in machine learning were used to predict possible quality control failure situations in the future in the process.

The accuracy of individual and stacking algorithms trained according to all these parameters was compared and the most appropriate algorithm was chosen. After the most accurate prediction model was found as stacking 3, the comparison of the ensemble algorithm with the individual algorithms was performed. K-NN and Naive Bayes are base learners and DT is a meta learner in the model of stacking 3. The values of the algorithms are shown in Table 9.

Table 9. Comparing of Performance of Algorithms

Algorithm	2 Sigma Shift Accuracy%	3 Sigma Shift Accuracy%
DT	58.25	78.75
K-NN	41.75	63.75
NB	61.25	75.25
Stacking Model (k-NN+NB/DT)	**69.00**	**85.75**

Three different individual machine learning algorithms and ensemble algorithms were trained to predict new situations by classification and their accuracy and error rates were compared. As a result, the Stacking 3 model was found to be the best answer for this purpose.

5. Conclusion

The results were obtained by using machine learning algorithms firstly individually and then using community methods. When all the results of the prediction accuracy rates were examined, it was seen that the bagging and boosting ensemble methods did not provide a significant improvement compared to the individual use of algorithms. However, with the stacking method, successful results were obtained by providing a significant improvement in prediction accuracy compared to individual algorithms. Since the aim is to classify the out-of-control conditions with the most accurate estimation, stacking 3 with the highest accuracy was chosen as the appropriate classification model to be used in the process. In this model, K-NN and NB are determined as base learner and DT as meta learner. The model result can be said to be the most successful method with 69.00% prediction accuracy for 2 sigma shift and 85.75% for 3 sigma shift.

The most fundamental criterion for improving process quality is correctly detecting the variable that causes the product to be faulty. The effectiveness of the multivariate control chart will be increased by this study, and it will contribute to the improvement by correctly determining the variables that produce the out of control state. It has been demonstrated that the ensemble learning algorithms employed for this purpose provide better learning, and that using ensemble algorithms rather than individual algorithms in such investigations will be beneficial. In addition, this work presents a technique that addresses the drawbacks of Hotelling T^2 and other multivariate control diagrams.

The inability to use all individual algorithms can be stated as a limitation of the study. In future investigations, the results will be compared against those of various individual algorithms and community combinations found in the literature. Other than Hotelling T^2, the mentioned machine learning algorithms can be implemented into multivariate control diagrams. In addition, by building an interface with an expert system, it is planned to create software that allows the quality expert in the process to simply enter production data and view the outcomes. Furthermore, by using this established model, better scientific findings will be obtained by deciding on the reason of an out-of-control signal with an automated and learning system rather than a human expert's judgment to heuristic.

REFERENCES

Alfaro, E., Alfaro, J. L., Gamez, M., & Garcia, N. (2009). A boosting approach for understanding out-of-control signals in multivariate control charts. *International Journal of Production Research*, 47(24), 6821-6834.

Ali, M., Eyduran, E., Tariq, M. M., Tirink, C., Abbas, F., Bajwa, M. A., & Jan, S. (2015). Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in Harnai sheep. *Pakistan Journal of Zoology*, 47(6).

Alpaydın, E. (2012), *Yapay Oğrenme*, Boğaziçi University, 3. Editon.

Anwar, H., Qamar, U., & Muzaffar Qureshi, A. W. (2014). Global optimization ensemble model for classification methods. *The Scientific World Journal*, 2014.

Aparisi, F., & Sanz, J. (2010). Interpreting the out-of-control signals of multivariate control charts employing neural networks. *International Journal of Computer and Information Engineering*, 4(1), 24-28.

Arıoglu Salmona, M. Ö., & Fırat, S. (2005). Bir Endüstri İşletmesinde Çok Değişkenli İstatistiksel Süreç Kontrolü. *V. Ulusal Üretim Araştırmaları Sempozyumu, İstanbul Ticaret Üniversitesi*.

Asadi, A., & Farjami, Y. (2019). Online Mean Shift Detection in Multivariate Quality Control Using Boosted Decision Tree learning. *Journal of System Management*, 5(2), 81-106.

Bakdi, A., Kouadri, A., & Bensmail, A. (2017). Fault detection and diagnosis in a cement rotary kiln using PCA with EWMA-based adaptive threshold monitoring scheme. *Control Engineering Practice*, 66, 64-75.

Bhatia, N. (2010). *Survey of nearest neighbor techniques*. arXiv preprint arXiv:1007.0085.

Breiman, L. (1996*). Bagging predictors.* Machine learning, 24(2), 123-140.

Chen, L. H., & Wang, T. Y. (2004). Artificial neural networks to classify mean shifts from multivariate χ2 chart signals. *Computers & Industrial Engineering*, 47(2-3), 195-205.

Cheng, C. S., & Cheng, H. P. (2008). Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines. *Expert Systems with Applications*, 35(1-2), 198-206.

Cheng, C. S., & Lee, H. T. (2012). Identifying the out-of-control variables of multivariate control chart using ensemble SVM classifiers. *Journal of the Chinese Institute of Industrial Engineers*, 29(5), 314-323.

Demir, N., & Dalkiliç, G. (2018). Modi ed stacking ensemble approach to detect network intrusion. *Turkish Journal of Electrical Engineering & Computer Sciences*, 26(1), 418-433.

Dreiseitl, S., Ohno-Machado, L., Kittler, H., Vinterbo, S., Billhardt, H., & Binder, M. (2001). A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *Journal of biomedical informatics*, 34(1), 28-36.

Du, S., Lv, J., & Xi, L. (2012). On-line classifying process mean shifts in multivariate control charts based on multiclass support vector machines. *International Journal of Production Research*, 50(22), 6288-6310.

Edward-Jackson, J. (1985). *Multivariate quality control*. Communications in Statistics-Theory and Methods, 14(11), 2657-2688.

Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780), 1612.

Gonzalez-de la Parra, M., & Rodriguez-Loaiza, P. (2003). Application of the multivariate T2 control chart and the Mason–Tracy–Young decomposition procedure to the study of the consistency of impurity profiles of drug substances. *Quality Engineering*, 16(1), 127-142.

Gowda, S., Kumar, H., & Imran, M. (2018). Ensemble based learning with stacking, boosting and bagging for unimodal biometric identification system, *12th International Conference on Recent Trends in Engineering and Technology*, 30-36.

Guh, R. S., & Shiue, Y. R. (2008). An effective application of decision tree learning for on-line detection of mean shifts in multivariate control charts. *Computers & Industrial Engineering*, 55(2), 475-493.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.

He, S., Wang, G. A., Zhang, M., & Cook, D. F. (2013). Multivariate process monitoring and fault identification using multiple decision tree classifiers. *International Journal of Production Research*, 51(11), 3355-3371.

Hotelling, H. (1947). Multivariate quality control, illustrated by the air testing of sample bombsights. *Techniques of statistical analysis*, 111-184.

Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). *The distance function effect on k-nearest neighbor classification for medical datasets*. SpringerPlus, 5(1), 1-9.

Huda, S., Abdollahian, M., Mammadov, M., Yearwood, J., Ahmed, S., & Sultan, I. (2014). A hybrid wrapper–filter approach to detect the source (s) of out-of-control signals in multivariate manufacturing process. *European Journal of Operational Research*, 237(3), 857-870.

Jiang, J., & Song, H. M. (2017). Diagnosis of out-of-control signals in multivariate statistical process control based on bagging and decision tree. *Asian Business Research*, 2(2), 1.

Jonathan, O., Omoregbe, N., & Misra, S. (2019). *Empirical Comparison of Cross-Validation and Test Data on Internet Traffic Classification Methods. In Journal of Physics: Conference Series*, 1299(1), p. 012044. IOP Publishing.

Karimi, S., Yin, J., & Baum, J. (2015). Evaluation methods for statistically dependent text. *Computational Linguistics*, 41(3), 539-548.

Lantz, B. (2013). *Machine learning with R : learn how to use R to apply powerful machine learning methods and gain an insight into real-world applications*. Birmingham: Packt Publishing Ltd.,66-343.

Lauría, E. J., Presutti, E., Kapogiannis, M., & Kamath, A. (2018). Stacking Classifiers for Early Detection of Students at Risk. *In CSEDU* (1) (pp. 390-397).

Li, T. F., Hu, S., Wei, Z. Y., & Liao, Z. Q. (2013). A framework for diagnosing the out-of-control signals in multivariate process using optimized support vector machines. *Mathematical Problems in Engineering*, 2013.

Ma, Z., Wang, P., Gao, Z., Wang, R., & Khalighi, K. (2018). Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose. PloS one, 13(10), e0205872.

Mason, R. L., & Young, J. C. (2002). Multivariate statistical process control with industrial applications. *Society for Industrial and Applied Mathematics*.

Mason, R. L., Tracy, N. D., & Young, J. C. (1995). Decomposition of T^2 for multivariate control chart interpretation. *Journal of quality technology*, 27(2), 99-108.

Mitchell, T. (1997). *Machine learning*. McGraw Hill.

Montgomery, D. C. (2009). *Introduction to statistical quality control*. John Wiley & Sons.

Murphy, B. J. (1987). Selecting out of control variables with the T^2 multivariate quality control procedure. Journal of the Royal Statistical Society: Series D, The Statistician, 36(5), 571-581.

Niaki, S. T. A., & Abbasi, B. (2005). Fault diagnosis in multivariate control charts using artificial neural networks. *Quality and reliability engineering international*, 21(8), 825-840.

Onan, A. (2018). Particle Swarm Optimization Based Stacking Method with an Application to Text Classification. *Academic Platform Journal of Engineering and Science*, 6(2), 134-141.

Oza, N. C., & Tumer, K. (2008). Classifier ensembles: Select real-world applications. *Information fusion,* 9(1), 4-20.

Ramezan C.A., Warner T.A., & Maxwell A.E. (2019). Evaluation of sampling and cross-validation tuning strategies for regional-scale machine learning classification. *Remote Sensing*, 11(2), 185.

Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:1811.12808.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009), *C Cross-Validation*, Springer, Boston.

Rokach, L. (2010). Ensemble-based classifiers. *Artificial intelligence review*, 33(1), 1-39.

Salvador–Meneses, J., Ruiz–Chavez, Z., & Garcia–Rodriguez, J. (2019). Compressed kNN: K-nearest neighbors with data compression. *Entropy*, 21(3), 234.

Shao, Y. E., & Lin, S. C. (2019). Using a time delay neural network approach to diagnose the out-of-control signals for a multivariate normal process with variance shifts. *Mathematics*, 7(10), 959.

Song, H., Xu, Q., Yang, H., & Fang, J. (2017). Interpreting out-of-control signals using instance-based Bayesian classifier in multivariate statistical process control. *Communications in Statistics-Simulation and Computation*, 46(1), 53-77.

Wolpert, D. H. (1992). Stacked generalization. *Neural networks*, 5(2), 241-259.

Yang, W. A. (2015). Monitoring and diagnosing of mean shifts in multivariate manufacturing processes using two-level selective ensemble of learning vector quantization neural networks. *Journal of Intelligent Manufacturing*, 26(4), 769-783.

Yu, J., Xi, L., & Zhou, X. (2009). Identifying source (s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble. *Engineering Applications of Artificial Intelligence*, 22(1), 141-152.

Yu, Y., & Feng, Y. (2014). Modified cross-validation for penalized high-dimensional linear regression models. *Journal of Computational and Graphical Statistics*, 23(4), 1009-1027.

Zhang, C., & Ma, Y. (2012). *Ensemble machine learning: methods and applications*. Springer Science & Business Media.

Zhang, M., & Cheng, W. (2015). 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.

Zhang, Y., Li, M., Han, S., Ren, Q., & Shi, J. (2019). Intelligent identification for rock-mineral microscopic images using ensemble machine learning algorithms. *Sensors.* 19(18), 3914.

Zhao, C., Wang, C., Hua, L., Liu, X., Zhang, Y., & Hu, H. (2017). Recognition of control chart pattern using improved supervised locally linear embedding and support vector machine. *Procedia engineering*, 174, 281-288.

Zheng, A. (2015). *Evaluating machine learning models: a beginner's guide to key concepts and pitfalls*. O'Reilly Media.

Zhou, Z. H. (2012). *Ensemble methods: foundations and algorithms.* Chapman and Hall/CRC.