World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:11, No:4, 2017

An Alternative Approach for Assessing the Impact of Cutting Conditions on Surface Roughness Using Single Decision Tree

S. Ghorbani, N. I. Polushin

Abstract**—**In this study, an approach to identify factors affecting on surface roughness in a machining process is presented. This study is based on 81 data about surface roughness over a wide range of cutting tools (conventional, cutting tool with holes, cutting tool with composite material), workpiece materials (AISI 1045 Steel, AA2024 aluminum alloy, A48-class30 gray cast iron), spindle speed (630-1000 rpm), feed rate (0.05-0.075 mm/rev), depth of cut (0.05-0.15 mm) and tool overhang (41-65 mm). A single decision tree (SDT) analysis was done to identify factors for predicting a model of surface roughness, and the CART algorithm was employed for building and evaluating regression tree. Results show that a single decision tree is better than traditional regression models with higher rate and forecast accuracy and strong value.

Keywords**—**Cutting condition, surface roughness, decision tree, CART algorithm.

I. INTRODUCTION

THE main goal in modern manufacturing industry is to manufacture low cost, high-quality products in short time. Turning operation is one of the most common machining processes for cutting and specially for finishing machined parts. In a turning operation, relative motion take places between cutting tool and workpiece, which affects the surface finish and tool life [1]-[3]. Surface roughness is one of the important aspects in mechanical design, which is used to evaluate the quality of the product since it affects the friction, corrosion and wear resistance, fatigue strength, lubricant, heat generation and product life [4]. In a turning operation it is an important task to avoid chatter vibrations as they cause tool breakage, tool wear, dimensional errors and unacceptable surface finish [5], [6]. In a turning operation, severe chatter vibrations occur due to a dynamic motion between the work piece and the cutting tool [7]-[10]. Therefore, the majority of works attempt to reduce vibration during a machining process by increasing dynamic stiffness of a machining system, changing its main natural frequency or feedback-controlled actuators, which can be achieved by using special coating on a cutting insert, vibration damper or toolholder made of material with high damping capability [11]-[14].

In a turning operation, selecting the proper cutting parameters is of great importance in order to achieve high cutting performance. Several mathematical models [15]-[21] based on statistical regression techniques have been proposed to establish cause and effect relationship between cutting parameters and surface roughness. Then, an objective function is formulated to solve the optimal cutting parameters using optimization techniques. The equations, for any combinations of factor levels in a range specified, all have the form: $Y=b_0+b_1x_1+b_2x_2+b_3x_3+...+b_px_p$ where, Y represents the estimated surface roughness value, b_0, b_1, b_2, b_3, …, b_p are estimates of the regression parameters and x_1, x_2, x_3,…, x_p are the logarithmic transformation of independent parameters (such as material hardness (HRB), feed rate (mm/rev), tool nose radius (mm), depth of cut (mm) and cutting speed (m/min)). However, the studies provided primary available estimates of optimal cutting parameters and effect relationship between cutting parameters and surface roughness; they applied regression techniques as a method to estimate surface roughness and have difficulties in showing the important factors affecting on surface roughness. In addition, it is likely that the assumptions that are made in a regression technique may be violated, because linear regression techniques need assumptions to be made, including assumptions about the normality, linearity and homoscedasticity of the data among others [22].

The prediction of significant cutting parameters for surface roughness is not easy to accomplish by using deterministic equations. Hence, it is suited to decision trees as they are primarily aimed at the recognition of a random pattern in a given set of input values. Decision tree method enables to model relationship between variables without strong model assumptions. This method also identifies the ‘‘important’’ variables in classifying objects/observations through the built tree and basic functions when many variables are considered. In decision trees method the resulting classification model can be easily interpreted in comparison with other classification techniques [23]. A single decision tree (SDT) analysis can be used both for classification and regression problems, which is able to deal with collinear data, to exclude insignificant variables, and to allow asymmetrical distribution of samples [24].

Examples of decision tree applications include multi component fault diagnosis of a rotational mechanical system [25], assessing the workpiece surface roughness to support decision making in a machining process [26]. Reference [27] applied decision tree method for process planning in the machining process to determine cutting operations and

S. Ghorbani and N. I. Polushin are with the Department of Functional Nanosystems and High-temperature Materials, National University of Science and Technology “MISiS”, Russia, Moscow, No. 4, Leninsky Prospect, Moscow, Russia (phone: +79269801269, email: siamak.ghorbani@yahoo.com).

sequences, to select machine tools and cutting tools, to calculate machining parameters and generate CNC part programs. Reference [28] used decision tree in their investigation and they stated that the proposed approach has a higher recognition rate than other methods on the same dataset.

The present study develops and presents a new expert system to assess surface roughness using a single decision tree (SDT) model and results compared with linear regression models.

II. Materials

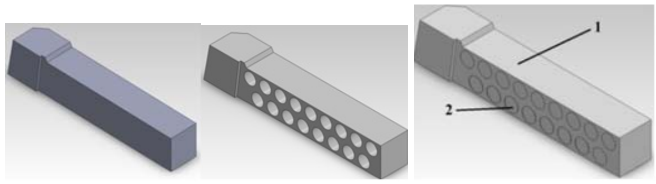
Machining experiments were performed at lathe machine model 16K20VF1 (Russia) with a maximum power of 5.5 kW and maximum spindle speed of 1600 rpm. The conventional cutting tool, cutting tool with horizontal holes in the toolholder arranged in a chess-board pattern, and cutting tool with horizontal holes filled up with epoxy-granite, with general specification of PCLNR 2525M12 made of AISI 5140, were used (Fig. 1). As it can be seen from Fig. 1 (c) the holes (Ø 10 mm) of the cutting tool are filled up with epoxy granite, the physical and mechanical characteristics of which are provided in Table I. As a cutting tool insert, we used the Carbide rhombic cutting insert (CT35M coated with TiC), which is manufactured by Sandvik Coromant. In this study AISI 1045 steel, AA2024 aluminum alloy and A48-class30 gray cast iron were used as workpieces with 65 mm diameter and 200 mm length. During turning operation, in each trial, the rust layers were removed by using a new cutting insert in order to reduce the effect of homogeneity of the workpiece material on the experimental result. The effect of wear during machining process was minimized by using a new cutting insert in each trial. This research applies the Taguchi approach to design experiments. Taguchi method is one of the important tools used in the industry to shortage product design, develop time and produce lower product cost. This method also takes into consideration the effect of uncontrollable factors on the response. Besides, Taguchi method is highly flexible and can allocate different levels of factors, even when the numbers of the levels of factors are not the same [29]. The cutting parameters were spindle speed (s), feed rate (f), depth of cut (d) and tool overhang (l). Three levels were specified for each of the factors as shown in Table II.

In this study, the average surface roughness (Ra) was measured, which was performed by means of a profile meter model 130 (Russia). To calculate the Ra, four different points of machined surface in 90° increments around the circumference were obtain and then the average value was calculated for roughness values [3].

III. Method

A. Background

Machine learning, statistical analysis, and other data mining are processes that use a variety of data analysis tools to discover models, patterns and relationships in data used to make predictions. A major goal of the analysis is to determine the best decisions. Data mining takes advantage of advances in the fields of artificial intelligence and statistics.



(a) (b) (c)

Fig. 1 (a) conventional cutting tool (b) cutting tool with holes in toolholder; (c) modified cutting tool filled up with epoxy granite: 1—toolholder and 2 — epoxy granite

TABLE I
PHYSICAL AND MECHANICAL CHARACTERISTICS OF EPOXY GRANITE

Parameter	Epoxy-granite
Density (kg/m^3)	2400–2600
Strength stress (MPa)	
Compression	150-160
Tensile	15-20
Elasticity module (MPa*10^{-4})	3.5–4.0
Poisson's ratio	0.25–0.40
Thermal conductivity (W/(m*K))	1.7–1.75
Linear expansion coefficient (1/°C)	(12–16)*10^{-6}
Damping ratio	0.6

TABLE II
CUTTING PARAMETERS AND THEIR LEVELS

Variables	Level 1	Level 2	Level 3
Spindle speed (rpn)	630	800	1000
Feed rate (mm/rev)	0.05	0.06	0.075
Depth of cut (mm)	0.05	0.1	0.15
Overhang (mm)	41	50	65

Both disciplines have been applied in pattern recognition and classification. Decision trees are simple and powerful data mining that can be used as a model for classification and prediction problems under uncertainty [30]. They provide unique capabilities to supplement, complement, and substitute for traditional statistical techniques of analysis such as multiple linear regressions. Decision tree describes graphically the decisions to help people to obtain a target value through the classification and analysis [31].

In general, a decision tree consists of tests or attributes nodes linked to two or more sub-trees and leafs or decision nodes labeled with a class which means the decision [32]. The outcome, which is related to the one of the sub-trees, is computed by a test node by attributing values of an instance. The classification of an instance is started at root node of the tree. The determination of the outcome for the instance is performed and continued applying the suitable sub-trees only if this node is a test node. Then, the class of the instance is predicted by its level when a leaf is encountered [32]. Reference [33] has popularized the application of decision trees to classification in machine learning. In 1986 a well-known tree-growing algorithm for inducing decision trees ID3 was introduced by [33]. In 1993 Quinlan's ID3 was upgraded with an algorithm called C4.5 [28]. They use the statistical calculation of information gain from a single attribute to build a decision tree. The algorithm basically chooses the attribute that provides the maximum degree of discrimination between classes locally. Theoretical concepts related to decision trees can be found in many text books [29], [34].

B. CART Algorithm

Reference [35] developed the Classification and Regression Trees (CART) methodology in their paper "Classification and Regression Trees", which generates binary decision trees. CART is a classification method which uses data for constructing decision trees and it is a nonparametric technique that can select from among a large number of variables those and their interactions that are most important in determining the outcome variable to be explained. The CART can be applied either as a classification tree or as a regressive tree depending on weather the response variable is categorical of continuous. In order to construct a CART tree the data were split repeatedly and then they were defined using a single explanatory variable. After which the partition of data into two homogeneous mutually exclusive groups is performed at each split in order to keep the tree small, which is equal to the number of final groups. Until the tree grows, the splitting process continues [36]. The classification trees are for dependent variables with a finite number of unordered values and prediction error, which are measured with regard to misclassification cost. While the regression trees are designed for dependent variables with ordered discrete or continuous values and prediction error which measured in terms of squared difference between the predicted and observed values [37].

A regression tree building is similar to a classification tree, which is centered on three major components: (1) a set of questions of the form: is $X \leq d$? where X is a variable and d is a constant. The response to such questions is yes or no; (2) goodness of split criteria for choosing the best split on a variable; and, (3) the generation of summary statistics for terminal nodes (unique to a regression tree). In regression tree, the least squared deviation (LSD) impurity measure is used for splitting rules and goodness of fit criteria. The LSD measure $R(t)$ is simply the weighted within node variance for node t, and it is equal to the resubstitution estimate of risk for the node. It is defined as:

$$R(t) = \frac{1}{N_w(t)} \sum_{i \in t} w_i f_i (y_i - \bar{y}(t))^2 \tag{1}$$

$$\bar{y}(t) = \frac{1}{N_w(t)} w_i f_i y_i \tag{2}$$

$$N_w(t) = \sum_{i \in t} w_i f_i \tag{3}$$

where, $N_W(t)$ represents the weighted number of records of the node t, w_i is the weighting field value for record i (if any), f_i is the frequency field value (if any), y_i is the target field value, and $\bar{y}(t)$ is the dependent variable mean (target field) of the node t.

The following equation can be used for determining the LSD criterion function for split s at node t [38]:

$$Q(S,T) = R(t) - R(t_L) - R(t_R) \tag{4}$$

where, $R(t_R)$ and $R(t_L)$ represent the sum of squares of the right and left child nodes, respectively.

The split s is selected in order to maximize the Q(s,t). Stopping rules determines how the splitting nodes in the specific tree are stopped by the algorithm. Tree growth is continued until at least one stopping rule is triggered by leaf node. The next conditions prevent a node to bg split: a. The model uses all predictors which have records in the node with the same value. b. The number of elements in the node is no more than the minimum parent node size (user defined). c. If the number of records in any of the child nodes resulting from the node's best split is less than the minimum child node size (user defined). d. The best split can be selected by decreasing the node impurity (user defined). In regression trees, $\bar{y}(t)$ is the predicted category of each terminal node.

IV. Application of Decision Tree and the Results

In this section of study, a single decision tree was used to develop model and to compare the results with linear regression models. In the first step, the related factors in the decision tree model were predicted using the classification and regression tree (CART) algorithm. The CART is an algorithm that performs a binary split, where only two child nodes are formed from the parent node, where the alpha value for the criteria of splitting and merging was set at 0.05. Besides, the weights for misclassification, costs were set asymmetrically in order to make up for the imbalance in data distribution. At the end of a training process, the model with the lowest error was selected as the final model. For qualitative evaluation of the models, the statistical measures such as the normalized mean square error, the correlation between actual and predicted, root mean squared error and mean absolute percentage error were used. The single decision tree diagrams and the error statistics of calculated significant cutting parameters by CART are illustrated in Fig. 2 and Table III, respectively.

TABLE III
Results of the Error Statics Calculated Surface Roughness

Error	SDT
Correlation between actual and predicted values	0.9165
Maximum error	1.0459
RMSE (Root mean squared error)	0.2748
MSE (Mean squared error)	0.0755
MAPE (Mean absolute percentage error)	16.141
NMSE (Normalized mean square error)	0.3377

The information displayed in each node in Fig. 2, depends on whether it is part of a classification tree (categorical target variable). Five lines of information are presented in this node: 1. Node number – The top line displays the number of the node. 2. Predictor variable used for split – The second line displays the name of the predictor variable that was used to generate the split from the parent node 3.

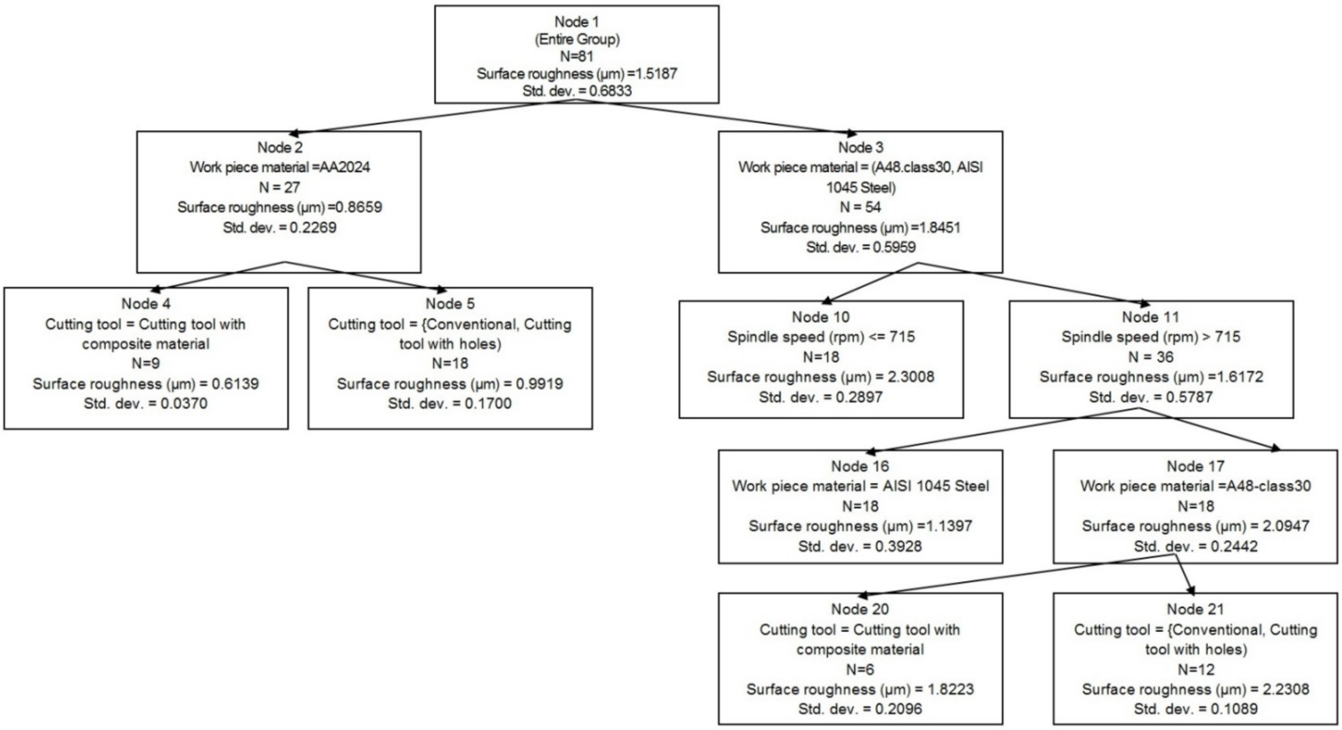


Fig. 2 Single decision tree generated by CART algorithm

Record counts – The "*N=nn*" shows how many rows (*N*) were placed in this node. 4. The name of the target variable and the mean value of the target variable – the next-to-bottom line displays the name of the target variable and the mean value of the target variable for all rows in this node 5. The standard deviation – the bottom line displays the standard deviation for the mean target value.

In decision tree the relative importance of input parameters can be found by algorithm itself determining the important parameters through branching of inputs, and knowledge of decision tree can help us choose parameters and assess the dependencies between related attributes. As can be seen, the greatest number of branching was performed using workpiece material. Therefore, workpiece material is the most important parameters for surface roughness. Table IV shows the relative importance of variables on surface roughness.

In the next step, after analyzing the prediction model of surface roughness, the obtained results were compared with traditional regression models.

TABLE IV
RELATIVE IMPORTANCE OF VARIABLES ON ESTIMATED SURFACE ROUGHNESS

Variable Importance	
Workpiece material	100.0
Spindle speed (rpm)	21.50
Cutting tool	20.08
Feed rate (mm/rev)	20.04
Depth of cut	15.58
Overhang (mm)	11.82

The regression equations and correlation between actual and predicted values are [15]-[21]:

$$R_a = 12.942 - 014.02f - 0.038V - 0.00445H, \mathrm{R}^2=0.672 \quad (5)$$

$$R_a = 8.6 - 0.00017V + 28.2f + 3.74d + 0.688r + 1.244f * a, \mathrm{R}^2=0.867 \quad (6)$$

$$R_a = 2.74 - 0.011V + 0.00117 * frequency + 261 * \ duty\ cycle, \mathrm{R}^2= 0.776 \quad (7)$$

$$R_a = 2.1066 - 0.0011V + 0.004f - 0.00976a\ \mathrm{R}^2=0.867 \quad (8)$$

$$R_a = 1.481 - 4.727 * 10^{-3}V + 9.817f + 0.1276a\ \mathrm{R}^2=0.504 \quad (9)$$

$$R_a = 1.9596 - 5.582 * 10^{-3}V - 2.706f + 0.071a + +0.025V * f + 1.244f * a, \mathrm{R}^2=0.47 \quad (10)$$

$$R_a = 3.179 + 9.826f - 0.009V - 0.922a, \mathrm{R}^2=0.608 \quad (11)$$

where, *V* is cutting speed, *f* is feed rate, *a* is depth of cut, *r* is nose radius and *H* is material hardness. It is clearly seen that the results generated by decision tree are more accurate in comparison with regression models with higher recognition rate, forecast accuracy and strong practical value. In predicting the surface roughness (R^2=0.9165 for decision tree and R^2=0.6806 for traditional regression models in average).

V. Conclusion

A decision tree uses the values of a set of predictor variables to predict the value of a variable. In this study, the prediction of surface roughness was done using single decision tree based on workpiece material, cutting tool, spindle speed, feed rate, depth of cut and tool overhang on the surface roughness. It was concluded that decision tree, in contrast with traditional regression models, represents rules, and the significant parameter is determined by the algorithm itself through the branching of inputs. Besides, decision tree does not require parametric assumptions about the training data to be met. Moreover, they can easily accommodate nonlinear relationships to outcome and are readily understood. Decision tree can be used to make inferences to understand the "big picture" of the model.

Acknowledgment

The work was carried out with financial support from the Ministry of Education and Science of the Russian Federation in the framework of Increase Competitiveness Program of NUST «MISiS» (№ К4-2015-053).

References

[1] C. Thomas, M. Katsuhiro, O. Toshiyuki and Y. Yasuo, "Metal Cutting", Great Britain, 2000.
[2] V. A. Rogov and S. Ghorbani, "Research on selecting the optimal design of antivibrational lathe tool using computer simulation," Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, vol. 229, no. 3, 2015, pp. 162–167.
[3] V. A. Rogov and S. Ghorbani, "The Effect of Tool Construction and Cutting Parameters on Surface Roughness and Vibration in Turning of AISI 1045 Steel Using Taguchi Method," Modern Mechanical Engineering, vol. 4, 2014, pp. 8–18.
[4] J. Serge, "Metal Cutting Mechanics and Material Behavior," Technische universitiet Eindhoven, 1999.
[5] K. K. Rama and J. Srinivas, "Study of Tool Dynamics with a Discrete Model of Workpiece in Orthogonal Turning," International Journal of Machining and Machinability of Materials, vol. 10, no. 1-2, 2011, pp. 71–85.
[6] S. Ghorbani and N. I. Polushin, "Effect of Composite Material on Damping Capacity Improvement of Cutting Tool in Machining Operation Using Taguchi Approach," World Academy of Science, Engineering and Technology, International Journal of Chemical, Molecular, Nuclear, Materials and Metallurgical Engineering, vol. 9, no. 12, 2015, pp. 1222–1232.
[7] K. Ramesh and T. Alwarsamy, "Investigation of Modal Analysis in the Stability of Boring Tool using Double Impact Dampers Model Development," European Journal of Scientific Research, vol. 80, no. 2, 2012, pp. 182–190.
[8] M. Dogra, V. S. Sharma and J. Dureja, "Effect of Tool Geometry Variation on Finish Turning – A Review," Journal of Engineering Science and Technology Review, vol. 4, no. 1, 2011, pp. 1–13.
[9] K. Yusuke, M. S. Doruk, A. Yusuf, S. Norikzau and S. Eiji, "Chatter Stability in Turning and Milling with In Process Identified Process Damping," Journal of Advanced Mechanical Design, Systems and Manufacturing, vol. 4, no. 6, 2010, pp. 1107–1118.
[10] L. V. Martinez, J. C. Jauregui-Correa and E. Rubio-Cerda, "Analysis of Compliance between the Cutting Tool and The Workpiece on the Stability of a Turning Process," International Journal of Machine Tools and Manufacture, vol. 48, 2008, pp. 1054–1062.
[11] S. Kanase and V. Jadhav, "Enhancement of Surface Finish of Boring Operation using Passive Damper," Indian Journal of Applied Research, vol. 2, no. 3, 2012, pp. 68–70.
[12] L. N. Devin and A. A. Osaghchii, "Improving Performance of cBN Cutting Tools by Increasing their Damping Properties," Journal of Superhard Materials, vol. 34, no. 5, 2012, pp. 326–335.
[13] V. A. Rogov, S. Ghorbani, A. N. Popikov and N. I. Polushin, "Improvement of cutting tool performance during machining process by using different shim," Archives of Civil and Mechanical Engineering, 10.1016/j.acme.2017.01.008.
[14] S. S. Abuthakeer, P. V. Mohanram and G. Mohan Kumar, "Prediction and Control of Cutting Tool Vibration Cnc Lathe with Anova and Ann," International Journal of Lean Thinking, vol. 2, no. 1, 2011, pp. 1–23.
[15] K. Arun Vikram and Ch. Ratnam, "Empirical model for surface roughness in hard turning based on analysis of machining parameters and hardness values of various engineering materials," International Journal of Engineering Research and Application, vol. 2, no. 3, 2012, pp.3091–3097.
[16] L. B. Abhang and M. Hameedullah, "Optimal machining parameters for achieving the desired surface roughness in turning of steel," The Journal of Engineering Research, vol. 9, no. 1, 2012, 37–45.
[17] Sivarao, T. J. S. Anand, Ammar, Shukor, "RSM based modeling for surface roughness prediction in laser machining," International Journal of Engineering & Technology, vol. 10, no. 4, 2010, pp. 26–32.
[18] M. F. F. Ab. Rashid and M. R. Abdul Lani, "Surface roughness prediction for CNC milling process using artificial neural network," Proceedings of the World Congress on Engineering, vol. 3, June 30 - July 2, 2010, London, U.K., pp. 1–6.
[19] Amit Kumar Gupta, "Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression," International Journal of Production Research, vol. 48, no. 3, 2010, pp. 763–778.
[20] Ersan Aslan, Necip Camuscu, Burak Birgoren, "Design optimization of cutting parameters when turning hardened AISI 4140 steel (63 HRC) with Al2O3 + TiCN mixed ceramic tool," Materials and Design, vol. 28, 2007, pp. 1618–1622.
[21] J. P. Davim, "A note on the determination of optimal cutting conditions for surface finish obtained in turning using design of experiments," Journal of Materials Processing Technology, vol. 116, 2001, pp. 305–308.
[22] H. Byeon, "The Risk Factors of Laryngeal Pathology in Korean Adults Using a Decision Tree Model," Journal of Voice. vol. 29, no. 1, 2015, pp. 59–64.
[23] Y. Zhao and Y. Zhang, "Comparison of decision tree methods for finding active objects," Advances in Space Research, vol. 41, 2008, pp. 1955–1959.
[24] N. C. Coops, R. H.Waring, C. Beier, R. Roy-Jauvin, T. Wang, "Modeling the occurrence of 15coniferous tree species throughout the Pacific Northwest of North America using a hybrid approach of a generic process-based growth model and decision tree analysis," Applied Vegetation Science, vol. 14, 2011, pp. 402–414.
[25] M. Saimurugan, K. I. Ramachandran, V. Sugumaran, N. R. Sakthivel, "Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine," Expert Systems with Applications, Vol. 38, 2011, pp. 3819–3826.
[26] B. Filipic, M. Junkar, "Using inductive machine learning to support decision making in machining processes," Computers in Industry, vol. 43, 2000, pp. 31–41.
[27] K. Wang, "An integrated intelligent process planning system (IIPPS) for machining," Journal of Intelligent Manufacturing, vol. 9, 1998, pp. 503–514.
[28] I. Aydin, M. Karakose, E. Akin, "An approach for automated fault diagnosis based on a fuzzy decision tree and boundary analysis of a reconstructed phase space," ISA Transactions, vol. 53, 2014, pp. 220–229.
[29] P. J. Ross, "Taguchi Techniques for Quality Engineering," McGraw-Hill International Book Company, OH, 1996.
[30] M. Kantardzic, Data Mining: "Concepts, models, methods, and algorithms," John Wiley and Sons, 2003.
[31] J. W. Han, M. Kamber, "Data mining: concepts and techniques," 2nd ed. CA: Morgan Kaufmann Publishers, 2001.
[32] J. R. Quinlan, "C4.5: Programs for machine learning, "San Mateo, (CA): Morgan Kaufman; 1993.
[33] D. Hand, M. Heikki, S. Padhraic, "Principles of data mining," MIT press, 2001.
[34] J. R. Quinlan, "Induction of decision trees, "Machine Learning, vol. 1, 1986, pp. 81-106.
[35] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, "Classification and regression trees," Belmont, Wadsworth Statistical Press, 1984.

[36] G. de'Ath and K. E. Farbricus, "Classification and regression trees: A powerful yet simple technique for ecological data, analysis," Ecology, vol. 81, no. 11, 2000, pp. 3178–3192.
[37] W. Y. Loh, "Classification and regression trees, "WIREs Data Mining and Knowledge Discovery, vol. 1, 2011, pp. 14–23.
[38] M. K. Ayoubloo, A. Etemad-Shahidi, J. Mahjoobi, "Evaluation of regular wave scour around a circular pile using data mining approaches," Applied Ocean Research, vol. 32, 2010, pp. 34–39.