



Decision tree-based parametric analysis of a CNC turning process

S.S. Dandge^a and S. Chakraborty^{b,*}

a. *Department of Mechanical Engineering, Government Polytechnic, Murtizapur, Maharashtra, India.*

b. *Department of Production Engineering, Jadavpur University, Kolkata, West Bengal, India.*

Received 7 May 2019; received in revised form 7 January 2020; accepted 18 February 2020

KEYWORDS

CNC turning process;
Decision tree;
CART;
CHAID;
Parameter;
Response.

Abstract. Computer Numerical Control (CNC) is a manufacturing concept where machine tools are automated to perform some predefined functions based on the instructions fed to them. CNC turning processes have found wide-ranging applications in modern-day manufacturing industries due to their capabilities to produce low-cost high-quality parts/components with very close dimensional tolerances. In order to exploit the fullest potential of a CNC turning process, its different input parameters should always be set to the optimal level for operation. In this paper, two classification tree algorithms, i.e., Classification And Regression Tree (CART) and Chi-squared Automatic Interaction Detection (CHAID) are applied to study the effects of various turning parameters on the responses and identify the best machining conditions for a CNC process. It is perceived that the obtained settings almost match with the observations of the earlier researchers. The CART algorithm outperforms CHAID with respect to higher overall classification accuracy and lower prediction risk.

© 2021 Sharif University of Technology. All rights reserved.

1. Introduction

In manufacturing industries, machining is one of the major operations performed to remove unwanted material from the workpiece surface to attain the desired shape of the final product/component while fulfilling the customers' needs. Therefore, machining involves shaping the component by removing material. This can be achieved by using a tool whose material is harder than the component to be molded, which is removed by shear deformation in the form of chips [1]. Amongst various machining operations, turning is the most

popularly adopted process. To meet the increasing requirements of low-cost and high-quality products, higher productivity, higher dimensional accuracy, and lower surface finish, Computer Numerical Control (CNC) technology is constantly replacing traditional turning operations. The precision and accuracy that can be achieved through the use of CNC turning operations cannot be achieved by traditional material removal processes [2]. The performance of any of the machining operations can usually be characterized by the combination of its various input parameters and outputs (responses). The input parameters of CNC turning operation mainly include feed rate, cutting speed, depth of cut, spindle speed, tool nose radius, machining time, type of the work material, cutting tool type, cutting fluid used etc. On the other hand, Material Removal Rate (MRR), Surface Roughness (SR), the amplitude of vibration, Tool Life (TL), Power

*. *Corresponding author.*

E-mail address: s_chakraborty00@yahoo.co.in (S. Chakraborty)

Consumption (PC), Cutting Force (CF), and acoustic emission are the major process outputs. In CNC machines, these input parameters need to be more effectively controlled so as to achieve the target response values, minimize machining time and cost, minimize tool wear, impart more flexibility, generate complicated shapes, provide higher repeatability, attain high dimensional tolerance, reduce power consumption etc. It has been observed that there exist complex relationships between the input parameters and responses in CNC turning operations. Understanding these relationships and proposing parametric optimization techniques to improve the performance of CNC turning processes play a pivotal role in a manufacturing environment.

Optimization of a CNC turning operation has already been identified as a complex problem due to the involvement of several different and often contradictory objectives, like maximization of MRR and minimization of SR, maximization of turning efficiency, and minimization of power consumption etc. It is the primary objective of any optimization tool to identify the optimal values of various CNC turning parameters so as to achieve better process performance [3]. Usually, in manufacturing industries, the concerned machinists select the most suitable settings of different input parameters based on their knowledge and expertise. Sometimes, machining data handbooks have also been consulted to meet these requirements. However, the proposed machining parameters are far from their optimal values which could hinder the goal of achieving better performance. Due to the rapid development of CNC technology and the availability of large amounts of related data, it is now impossible to achieve the best machining performance only by deploying conventional optimization technologies. Therefore, the need for more advanced tools is ardently felt to fulfill the above-mentioned objective. In this paper, the application of a data mining tool in the form of the development of decision trees is explored to determine the optimal parametric settings of different input parameters in a CNC turning operation. Two decision tree generation algorithms, i.e. Classification And Regression Tree (CART) and CHi-squared Automatic Interaction Detection (CHAID) are employed here to investigate the effects of the considered CNC turning parameters on the responses. The relative performance of both the algorithms is also compared with respect to solution accuracy and prediction risk.

The 'If-Then' decision rules generated based on the applications of CART and CHAID classification algorithms constitute a more powerful knowledge representation to understand the effects of different input parameters on the responses for the considered CNC turning process. When these rules are organized as a non-overlapping decision set, they become quite easy to interpret, even by a non-technical end-user. They

follow a general structure, i.e. if the given CNC turning conditions are met, then certain response values can be attained or predicted. They are probably the most interpretable prediction models, semantically resembling the natural language and human thinking process. They also provide valuable information on how to make the final decision and why certain conditions are met. The rule generation process using the CART and CHAID algorithms has high speed and scalability and is almost robust against the presence of outliers in the input dataset. The decision rules usually generate sparse models, which do not contain many features and draw final conclusions based on only a few binary statements. They can be generated from large-scale datasets containing numerical and categorical information.

2. Literature survey

Gupta et al. [4] employed Taguchi method along with fuzzy modeling for multi-objective optimization of a CNC turning process while considering cutting speed, feed rate, depth of cut, tool nose radius, and cutting environment as the input parameters, and SR, tool life, CF and power consumption as the responses. Mukherjee et al. [5] introduced the application of the Taguchi method to identify the optimal operating levels of speed, feed, and depth of cut to maximize MRR during CNC turning of SAE 1020 material. Marko et al. [6] applied the Particle Swarm Optimization (PSO) technique to determine the optimal settings of three CNC cutting parameters, i.e. cutting speed, feed rate, and cutting depth for achieving the desired values of cutting force, SR and tool life. Saini and Pradhan [7] conducted CNC turning operation on aluminum alloy 8011 to investigate the effects of cutting speed, feed, and depth of cut on MRR and SR using an integrated Taguchi-fuzzy approach. While taking into account tool nose radius, cutting speed, feed rate, and depth of cut as the input parameters, Vasudevan et al. [8] combined grey theory and fuzzy technique with the Taguchi method to optimize SR, tangential CF, and MRR in CNC turning of GFRP/epoxy composite materials. Saini and Pradhan [9] studied the effects of three CNC turning parameters, i.e. spindle speed, feed rate, and depth of cut on SR properties of 316L stainless steel, EN24 alloy steel, and Ti6Al4V alloy materials. While considering cutting speed, spindle speed, feed rate, and depth of cut in a CNC machining operation, Aghdeab et al. [10] determined the minimum SR values using the Simulated Annealing (SA) technique. Camposeco-Negrete [11] identified feed rate and depth of cut as the two most important input parameters in rough turning operation of AISI 6061 T6 aluminum material to minimize energy consumption and SR and maximize MRR. Sarıkaya and Güllü [12] presented the

application of Taguchi-based Grey Relational Analysis (GRA) to identify the optimal combination of the type of the cutting fluid, fluid flow rate, and cutting speed for having better values of flank wear, notch wear, and SR during machining Haynes 25 material under minimum quantity lubrication condition.

Asiltürk et al. [13] presented the combined application of Taguchi methodology and Response Surface Methodology (RSM) to study the effects of spindle rotational speed, feed rate, depth of cut, and tool tip radius on SR properties of Co28Cr6Mo alloy. Kumar et al. [14] identified the optimal parametric combination of cutting speed, feed rate, and depth of cut during CNC machining of AISI 1045 steel material to have the most preferred values of tool wear rate and MRR. During the CNC turning of AA7075 material, Maheswara Rao and Venkatasubbaiah [15] concluded that feed and cutting speed were the most significant parameters affecting SR of the machined components. Klancnik et al. [16] employed gravitational search algorithm and non-sorting genetic algorithm-II as the multi-objective optimization tools for a CNC turning process. Cutting speed, depth of cut, and feed rate were considered as the input parameters, and SR, CF, and tool life were the responses for the said process. Based on Taguchi experimental design plan, Bilga et al. [17] studied the effects of cutting speed, feed rate, depth of cut, and nose radius on some energy consumption responses, like energy efficiency, power factor, and active energy consumed during CNC rough turning operation of EN 353 alloy steel material. Based on Taguchi's L_9 experimental design plan, Kushwaha and Singh [18] studied the effects of cutting speed, feed rate, and depth of cut on SR and MRR during CNC turning of Inconel 625 material using coated carbide tool. Based on a developed model, Nataraj and Balasubramanian [19] determined the optimal settings of three CNC turning parameters, i.e. cutting speed, depth of cut, and feed rate in order to minimize SR, the intensity of vibration, and work-tool interface temperature. Nayak and Sodhi [20] applied the RSM technique to evaluate the relationship between three CNC turning parameters, i.e. depth of cut, feed rate, and cutting speed, and two responses, i.e. MRR and SR. The desirability function approach was later adopted to determine the optimal parametric settings for the considered process. During CNC turning of aluminum alloy, Sahoo et al. [21] integrated weighted principal component analysis with RSM technique to minimize SR and tool vibration while taking into account spindle speed, feed rate, and depth of cut as the input parameters. Based on the developed RSM-based equations, Mandal et al. [22] determined the optimal values of spindle speed, feed rate, and depth of cut for achieving favorable values of MRR, SR, and power in a CNC turning process. The corresponding Pareto

fronts for the conflicting objectives were also proposed. Suresh et al. [23] determined the optimal settings of cutting speed, feed rate, tool nose radius, and depth of cut in a CNC turning operation for having minimum SR and maximum MRR values.

Akkuş [24] considered cutting speed, feed rate, and depth of cut as the process parameters in CNC machining of AISI 1040 steel material and identified their optimal settings for having minimum SR values. Bhanu Prakash et al. [25] treated spindle speed, feed rate, and depth of cut as the input parameters, and MRR and SR as the responses during CNC turning operation of AlSi₇ Mg material. Taguchi method and GRA technique were later adopted to optimize the said process. Gadekula et al. [26] determined the best parametric settings for spindle speed, feed rate, and depth of cut when using the Taguchi method to perform CNC turning operations on high-carbon and high-chromium steel workpiece materials. The MRR and SR were the responses for the considered process. Palanisamy and Senthil [27] proposed a combined application of the grey system and fuzzy logic approach to optimize cutting speed, feed rate, and depth of cut in a CNC turning process for achieving minimum values of SR and power consumption. Sahoo et al. [28] developed second-order RSM-based regression models to investigate the effects of three CNC turning parameters, i.e. spindle speed, feed rate, and depth of cut on two responses, i.e. SR and tool vibration. Weighted Aggregate Sum Product Assessment (WASPAS) method was later adopted for parametric optimization of the considered process. While performing CNC turning operation on aluminum alloy, Saravanakumar et al. [29] employed Taguchi method to determine the optimal settings of feed, speed and depth of cut to attain minimum values of SR and roundness error. Using RSM technique, Nataraj et al. [30] examined the impacts of feed rate, cutting speed and depth of cut on the work-tool interface zone temperature and SR during CNC turning operation of LM6 reinforced metal matrix composites. Vasudevan et al. [31] combined principal component analysis with GRA technique to explore the effects of feed rate, depth of cut, cutting speed and tool nose radius on MRR, CF and different SR parameters during machining of glass fibre reinforced polymer composites. Rao et al. [32] performed multi-objective optimization of MRR and SR during CNC turning of stainless steel 304 work materials. Cutting speed, feed rate and depth of cut were taken into account as the input parameters for the said process. Vijay Kumar et al. [33] endeavoured to investigate the effects of feed rate, depth of cut and spindle speed on MRR and SR while machining EN 19 stainless steel materials using a CNC turning centre. Taguchi's L_{18} orthogonal array was incorporated for conducting the experimental trials. Arun Vikram et al. [34] applied Taguchi method to explore the

interactions between feed rate, depth of cut and spindle speed, and SR, MRR and tool temperature. Later, GRA technique was employed to determine the best parametric condition for attaining better response values during wet machining of materials with low machinability. Chau et al. [35] integrated Taguchi method, adaptive neuro-fuzzy inference system and teaching-learning-based optimization algorithm for parametric optimization of a CNC turning process.

The existing literature is full of applications of various mathematical tools and techniques for parametric optimization of CNC turning processes. A great deal of the past researches is dedicated to the applications of the Taguchi method to identify the best settings of different turning parameters for attaining the desired response values. Various multi-criteria decision-making methods, like GRA, and WASPAS, have also been proposed for the same purpose. Based on the experimental data, some researchers have also attempted to develop the corresponding higher-order RSM-based regression equations to describe the relationships between various CNC turning parameters and responses. Metaheuristic algorithms, mainly in the form of SA, PSO, etc., have been later applied to solve those equations in order to identify the optimal values of various input parameters for enhancing the process performance. But, it is quite interesting to notice that there is an immense scarcity of the application of any kind of data mining tool for parametric optimization of CNC machining processes. In this paper, the application of a data mining tool in the form of decision (classification) trees is proposed for the first time for parametric analysis and optimization of a CNC turning process. Using CART and CHAID-based decision trees, the corresponding decision rules in the form of simple and understandable ‘If-Then’ statements are generated to study the effects of each of the considered CNC turning parameters on the responses. The relative performance of both these algorithms is also compared with respect to classification accuracy and prediction risk.

3. Decision tree

Over the years, decision trees have become one of the most popular tools in knowledge discovery and data mining. It mainly deals with exploring large amounts of data to find meaningful patterns [36]. Through applying the decision tree algorithm the effort takes place to solve a given problem in the form of tree representation. It basically employs development of a training model which can predict class or value of target variables based on the decision rules generated from the initial dataset. It belongs to the class of supervised machine learning algorithm of data mining which has the capability to solve both regression and classification

problems. It is also a non-parametric approach having no idea about the pertaining distribution of the data. The developed decision trees usually follow the human thinking process and the logic of interpreting the data is very strong. In a decision tree, the original dataset is broken down into smaller subsets while incrementally developing the associated decision tree at the same time. The final decision tree consists of many decision nodes and leaf nodes. A decision node may comprise two or more branches and a leaf node denotes a classification or a final decision. The topmost decision node containing the complete dataset is known as the root node. This dataset is sequentially split resulting in child nodes during classification. When no further splitting is possible, the final nodes are termed as terminal nodes. Similarly, each decision node of a decision tree relates to an attribute and each leaf node corresponds to a class label. In a decision tree, a series of ‘If-Then’ statements can be finally developed when tracing the path through the different decision and leaf nodes, starting from the root node. The decision trees have several advantages, likeability to deal with both categorical and numerical data, self-explanatory and being easy to interpret, scalability with big data, capability to process datasets having errors or missing values, the requirement of less computational effort, high predictive accuracy etc.

In this paper, two decision tree algorithms, i.e. CART and CHAID are applied for parametric analysis of a CNC turning process. In both the algorithms, each (non-terminal) node identifies a split condition which yields the optimal classification of dependent variables [37]. The details of these algorithms are presented here-in-under.

3.1. CART algorithm

It is a recursive partitioning method used for both regression and classification purposes. In this algorithm, the decision tree is constructed by splitting subsets of the data using all the predictor variables to create two child nodes repeatedly, beginning with the initial dataset. The best predictor variable is chosen based on a variety of impurity or diversity measures. The goal is to generate subsets of the data which are as homogeneous as possible with respect to the target variable [38]. The procedural steps for this algorithm are presented as below [39]:

- Step 1:** The algorithm runs through the entire dataset D to initiate the classification;
- Step 2:** If all datasets in D belong to class P , generate a node P and stop, otherwise, decide a predictor F and produce a decision node;
- Step 3:** Split the samples in D into possible subsets using a predictor selection measure called ‘Gini index’ which is used in splitting for

classification to reduce the impurity of a node. Split at each node occurs only when it can generate the greatest improvement in classification accuracy;

- Step 4:** When selecting predictor variables, determine the best breakpoint so that the dependent variable can be best divided into two categories, which are characterized by the maximum internal uniqueness and external difference;
- Step 5:** For each split, the predictor variable with the best score of improvement is selected;
- Step 6:** Apply the algorithm recursively for all the subsets in D until all items or samples in a node have the same class, i.e. split is no longer possible;
- Step 7:** The process repeats recursively until one of the stopping rules is fulfilled:
- a) If a node becomes pure, i.e. all cases in the node have identical values of the dependent variable;
 - b) If the current tree depth reaches the user-specified maximum limit;
 - c) If the size of a node is less than the user-specified minimum size;
 - d) If the split of a node results in a child node whose node size is less than the user-specified minimum size.
- Step 8:** Framing of ‘If-Then’ decision rules, i.e. rule: (condition) $\rightarrow Y$, where the condition is a combination of predictor variables and Y is the class label (decision).

In this algorithm, the measure of the importance of independent variables (X) in relation to a decision tree is defined as the sum of improvements that X has across all the splits in the tree when it is used as a primary or surrogate splitter. The importance of X is expressed in terms of a normalized quantity relative to the variable having the largest measure of importance. It ranges from 0 to 100, with the variable having the maximum importance score of 100. Thus, the variable importance plot is a good indicator that measures the importance of independent variables (which have already appeared in the decision tree).

3.2. CHAID algorithm

This algorithm, developed by Kass [40], is a decision tree development approach, based on the Chi-squared test, generated by repeatedly splitting the subsets into two or more child nodes starting with the initial dataset. Particularly, the predictor having the strongest relationship with the dependent variable based on p -value is utilized as the split node. To

determine the best split at a particular node, any allowable pair of categories of the predictor variables is merged until there is no statistically significant difference within the pair with respect to the target variable. It is an exploratory data analysis method used to study the strongest association between a dependent variable and a large series of possible predictor variables which themselves may interact. The dependency measure may be a qualitative (nominal or ordinal) or a quantitative indicator. For qualitative (categorical) variables, a series of Chi-squared analyses is conducted between the dependent and predictor variables. For quantitative variables (continuous), F -test is used, where intervals (splits) are optimally determined for the predictor variables so as to maximize the ability to explain a dependent measure with respect to variance components [41]. This algorithm uses the following steps [40,41]:

- Step 1:** The first step is to create categorical predictor variables from continuous variables by dividing the respective continuous distributions into a given number of categories;
- Step 2:** In the merging stage, for each dependent variable, merge non-significant categories. It determines the pair of categories that is least significantly different (i.e., most similar) with respect to dependent variables. The most similar pair is the pair whose test statistic provides the largest p -value with respect to the dependent variable;
- Step 3:** If the statistical test for the given pair of predictor categories is not statistically significant, it will merge the respective predictor categories and find the next pair of categories which may now include the previously merged categories;
- Step 4:** If the statistical test for the given pair of predictor categories is statistically significant, the adjusted p -value is computed for the merged categories by applying the Bonferroni adjustments;
- Step 5:** In the splitting stage, the independent or predictor variable with the lowest significant p -value (calculated above) is selected as the best and the group is split on this predictor (i.e., each of the optimally merged categories of the predictor is used to define a subdivision of the parent group into a new subgroup). If no predictor has a significant p -value, the group is not split;
- Step 6:** The above-mentioned steps are repeated until all subgroups have either been analyzed or contain too few observations or cases.

The stopping rules are basically the same as described in the CART algorithm.

4. Decision trees for a CNC turning process

Based on Taguchi orthogonal array design plan, Gupta et al. [4] conducted 27 experiments considering cutting speed, feed rate, depth of cut, tool nose radius, and machining environment as the controllable parameters. On the other hand, Tool Life (TL) (in min), Power Consumption (PC) (in W), SR (in μm), and Cutting Force (CF) (in N) were the responses. It is worthwhile to mention here that among the considered responses, TL is the only ‘Larger-The-Better’ (LTB) type of quality characteristic, while, the remaining three are ‘Smaller-The-Better’ (STB) types. During experimentation, each of the CNC turning parameters was set at three different operating levels, as shown in Table 1. A high-speed CNC machining centre was utilized for conducting the experiments and AISI P20 tool steel bars (having a diameter of 65 mm and length 275 mm) were chosen as the work material. The results of the experimental study are provided in Table 2. In this table, the minimum, maximum, and median values for each of the responses are also shown.

For parametric analysis of the considered CNC turning process and investigating the effects of various input parameters on the process outputs (responses), the corresponding decision trees are developed using CART and CHAID algorithms in SPSS 16.0 software. For arriving at the best possible solutions, various parameters of the adopted decision tree algorithms are fine-tuned as follows.

For CART algorithm

- Growing method: CART;
- Categorical dependent variables: TL, PC, SR, and CF;
- Categorical independent variables: CS, FR, DOC, NR, and E;
- Validation: Cross-validation;

- Number of sample folds: 3;
- Growth limit: Maximum tree depth = 5;
- Minimum number of cases: Parent node = 3, Child node = 2;
- Impurity measure: Gini;
- Minimum change in improvement: 0.0001.

For CHAID algorithm

- Growing method: CHAID;
- Categorical dependent variables: TL, PC, SR and CF;
- Categorical independent variables: CS, FR, DOC, NR and E;
- Validation: Cross validation;
- Number of sample folds: 3;
- Growth limit: Maximum tree depth = 5;
- Minimum number of cases: Parent node = 3, Child node = 2.

Significance level for

- a) Splitting node = 0.03;
 - b) Merging categories = 0.05;
 - c) Chi-square statistic = Pearson;
- Model estimation:
 - a) Maximum number of iterations = 100,
 - b) Minimum change in expected cell frequencies = 0.001;
 - c) Adjust significance values using the Bonferroni method.

Figure 1 exhibits the decision tree in the form of a classification tree diagram developed using the CART algorithm for tool life. In this diagram, tool life is represented as a dependent variable in the root node. As tool life is a continuous variable, its median value (27.66 min) calculated from the experimental dataset of Table 2 is adopted here for the splitting purpose. For tool life, which is an LTB quality characteristic, its values lower than or equal to 27.66 min are termed

Table 1. Computer Numerical Control (CNC) turning parameters and their levels [4].

CNC parameter	Symbol	Unit	Level		
			Low	Medium	High
Cutting Speed	CS	m/min	120	160	200
Feed Rate	FR	mm/rev	0.10	0.12	0.14
Depth Of Cut	DOC	mm	0.20	0.35	0.50
Nose Radius	NR	mm	0.40	0.80	1.20
Environment	E	—	Dry	Wet	Cryogenic

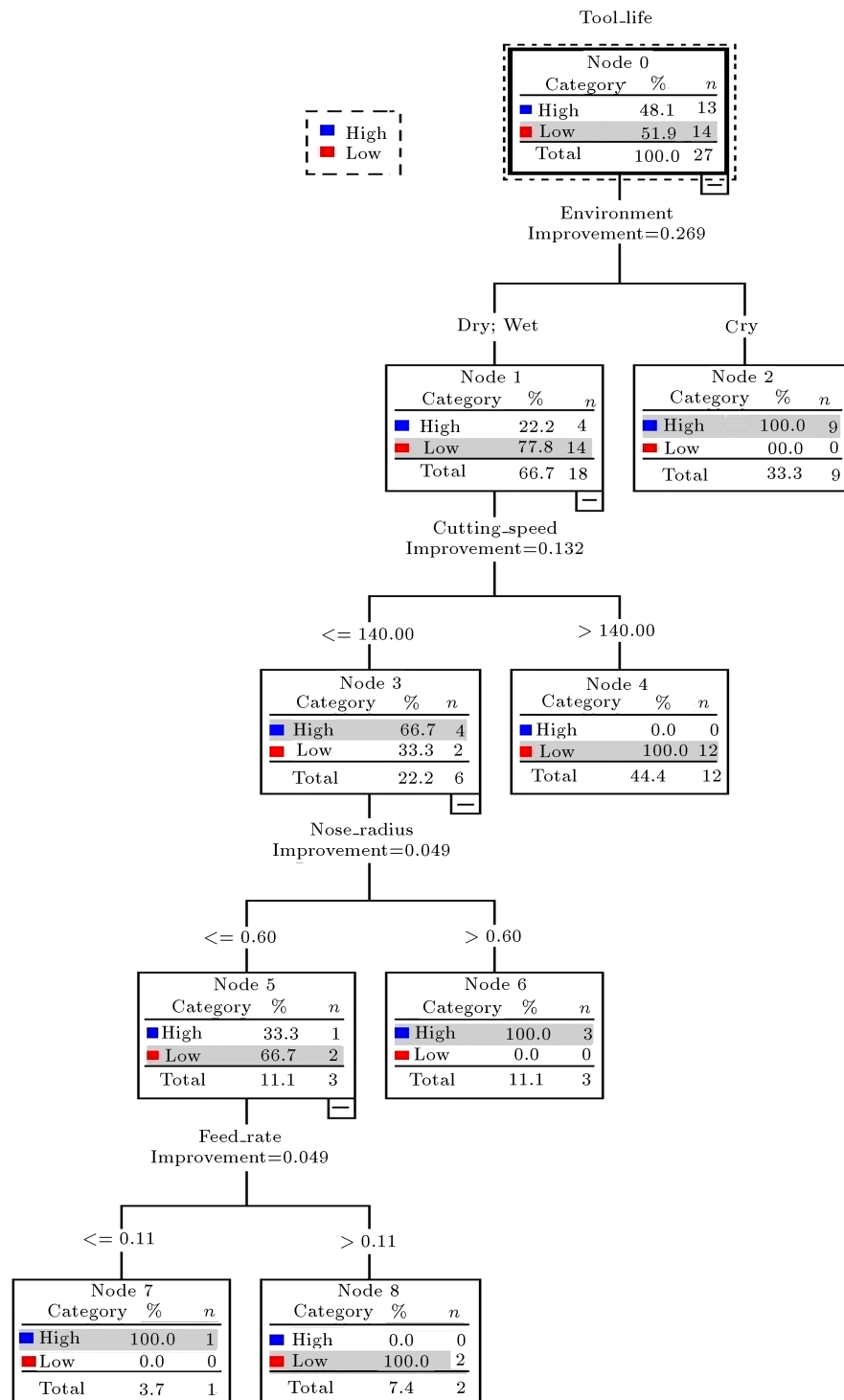


Figure 1. Classification tree for tool life using Classification And Regression Tree (CART) algorithm.

as 'low', whereas, values higher than 27.66 min are designated as 'high'. From the root node of the developed decision tree, it can be noticed that in the initial dataset, there are 13 experimental observations with high tool life and 14 observations have low tool life values. The first splitting is performed while taking the machining environment as the most important

predictor variable. Between the two formed child nodes, node 2 appears to be a terminal node from where no further splitting can be possible. It is also identified as a pure node with no misclassification error. From node 1, taking cutting speed as the next important predictor variable, another classification is performed with the formation of node 4 as a terminal and pure

Table 2. Experimental data for the Computer Numerical Control (CNC) turning operation [4].

Exp. no.	CNC parameter					Response			
	CS	FR	DOC	NR	E	TL	PC	SR	CF
1	120	0.10	0.20	0.40	Dry	29.00	1066	1.41	171.30
2	120	0.10	0.35	0.80	Wet	34.00	1560	0.71	147.50
3	120	0.10	0.50	1.20	CRYO	54.67	866	0.60	111.74
4	120	0.12	0.20	0.80	Wet	34.67	1493	0.47	120.30
5	120	0.12	0.35	1.20	CRYO	51.66	987	0.19	180.60
6	120	0.12	0.50	0.40	Dry	27.00	1187	1.18	236.20
7	120	0.14	0.20	1.20	CRYO	50.00	960	0.67	157.70
8	120	0.14	0.35	0.40	Dry	24.66	1134	1.16	214.40
9	120	0.14	0.50	0.80	Wet	28.33	1813	0.92	286.90
10	160	0.10	0.20	1.20	Wet	27.66	1586	0.18	116.37
11	160	0.10	0.35	0.40	CRYO	47.66	1013	0.45	133.33
12	160	0.10	0.50	0.80	Dry	21.66	1240	0.43	191.23
13	160	0.12	0.20	0.40	CRYO	45.66	893	0.58	125.40
14	160	0.12	0.35	0.80	Dry	20.33	1253	0.72	149.43
15	160	0.12	0.50	1.20	Wet	25.66	1773	0.31	212.46
16	160	0.14	0.20	0.80	Dry	20.00	1107	0.66	162.93
17	160	0.14	0.35	1.20	Wet	22.33	1533	0.64	190.23
18	160	0.14	0.50	0.40	CRYO	41.33	1373	0.75	177.76
19	200	0.10	0.20	0.80	CRYO	40.00	1053	0.16	106.23
20	200	0.10	0.35	1.20	Dry	15.67	1373	0.23	208.50
21	200	0.10	0.50	0.40	Wet	21.67	2094	0.67	209.80
22	200	0.12	0.20	1.20	Dry	14.67	1286	0.40	200.20
23	200	0.12	0.35	0.40	Wet	20.33	1866	0.50	178.80
24	200	0.12	0.50	0.80	CRYO	37.66	1613	0.18	168.70
25	200	0.14	0.20	0.40	Wet	18.00	1573	0.64	162.00
26	200	0.14	0.35	0.80	CRYO	34.33	1453	0.31	162.00
27	200	0.14	0.50	1.20	Dry	16.66	1667	0.48	276.16
					Minimum	14.67	866	0.16	106.23
					Maximum	54.67	2094	1.41	286.90
					Median	27.66	1373	0.58	171.30

node. Two nodes, i.e. 5 and 6 now emerge out from node 3 using tool nose radius as the predictor variable. Finally, based on feed rate, the last two terminal nodes are constructed. This entire classification process along with the related characteristics is provided in Table 3. The percentages of correct classification at all the nodes are presented in Table 4. It can be observed that for tool life, there is no misclassification error in the

decision tree developed using the CART algorithm. When the decision tree of Figure 1 and classification characteristics are analyzed in detail, several decision rules in the form of ‘If-Then’ statements are generated.

CART-based rules for tool life

Rule 1: If environment = cryogenic Then *TL* is

Table 3. Classification of tool life based on Classification And Regression Tree (CART).

Classification	Node	Characteristics
First	2	Cryogenic environment provides higher tool life
Second	1, 4	Dry or wet environment and $CS > 140$ m/min provide lower tool life
Third	1, 3, 6	Dry or wet environment, $CS \leq 140$ m/min and $NR > 0.60$ mm are responsible for higher tool life
Fourth	1, 3, 5, 7	Dry or wet environment, $CS \leq 140$ m/min, $NR \leq 0.60$ mm and $FR \leq 0.11$ mm/rev lead to higher tool life
Fifth	1, 3, 5, 8	Dry or wet environment, $CS \leq 140$ m/min, $NR \leq 0.60$ mm and $FR > 0.11$ mm/rev provide lower tool life

Table 4. Percentages of correct classification of tool life based on Classification And Regression Tree (CART).

Classification	Tool life			
	Low (≤ 27.66 min)		High (> 27.66 min)	
	Number of observations	Percentage	Number of observations	Percentage
1	0	0%	9	100%
2	12	100%	0	0%
3	0	0%	3	100%
4	0	0%	1	100%
5	2	100%	0	0%

(27.66–54.69]:

$$[P = 100\%, Q = 69.23\%, C = 33.33\%, QTY = 9]$$

$$[T = 202.56\%].$$

Rule 2: If environment = dry or wet and $CS > 140$ m/min Then TL is [14.67–27.66]:

$$[P = 100\%, Q = 85.71\%, C = 44.44\%, QTY = 12]$$

$$[T = 230.15\%].$$

Rule 3: If environment = dry or wet, $CS \leq 140$ m/min and $NR > 0.60$ mm, Then TL is (27.66–54.69]:

$$[P = 100\%, Q = 23.09\%, C = 11.11\%, QTY = 3]$$

$$[T = 134.20\%].$$

Rule 4: If environment = dry or wet, $CS \leq 140$ m/min, $NR \leq 0.60$ mm and $FR \leq 0.11$ mm/rev Then TL is (27.66–54.69]:

$$[P = 100\%, Q = 7.70\%, C = 3.70\%, QTY = 1]$$

$$[T = 111.40\%].$$

Rule 5: If environment = dry or wet, $CS \leq 140$ m/min, $NR \leq 0.60$ mm and $FR > 0.11$ mm/rev Then TL is [14.67–27.66]:

$$[P = 100\%, Q = 14.29\%, C = 7.41\%, QTY = 2]$$

$$[T = 121.70\%],$$

where P is the percentage of objects in the condition attribute set that corresponds to a rule (a measure of rule confidence), Q is the percentage of objects in the decision attribute set that corresponds to a rule, C is the percentage of objects that correspond to a rule (a measure of rule support) and QTY is the number of objects satisfying a particular rule. In this algorithm, $T(T = P + Q + C)$ represents the total strength (relative importance) of a rule [42].

Among these decision rules, Rule 2, having the maximum total strength of 230.15, states that when the said CNC turning operation is performed under a dry or wet environment and the cutting speed is greater than 140 m/min, the corresponding tool life would be low. On the other hand, Rule 1 with a total strength of 202.56 depicts that higher tool life can only be achievable under a cryogenic (CRYO) machining environment. It can also be noticed from these rules that low feed rate and low tool nose radius lead to higher tool life. The importance plot for tool life, as depicted in Figure 2, identifies the machining environment as the

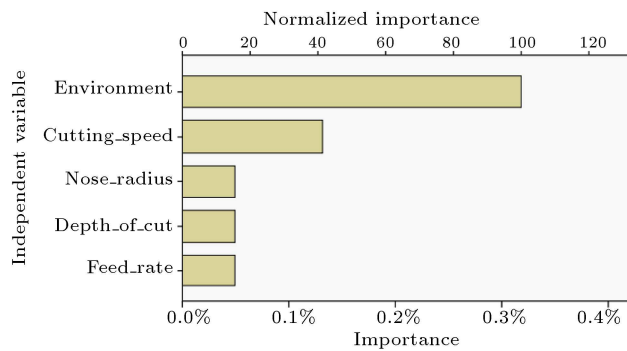


Figure 2. Importance plot for tool life .

most important CNC turning parameter, followed by cutting speed. Nose radius, depth of cut, and feed rate are observed to have the least importance on tool life. Similarly, the related decision tree for tool life is also generated using the CHAID algorithm, as exhibited in Figure 3. The classification characteristics and

percentages of accurate classification at the identified nodes are provided in Tables 5 and 6 respectively. As compared to five classifications in the CART algorithm for tool life, there are only four classifications in the CHAID algorithm. Here, at the third classification in node 5, there is a 33.33% misclassification error. The corresponding rules generated using this algorithm are quite similar to those as developed by the CART algorithm. Machining environment and cutting speed are observed to be the two most important CNC turning parameters affecting tool life. While analyzing both these sets of decision rules, it can be concluded that for attaining higher tool life, cryogenic environment, and low values of cutting speed, tool nose radius and feed rate are always preferred. It is interestingly revealed that depth of cut appears to be an insignificant parameter having no effect on tool life. Based on analysis of variance (ANOVA) results for tool life, Gupta et al. [4] also identified machining

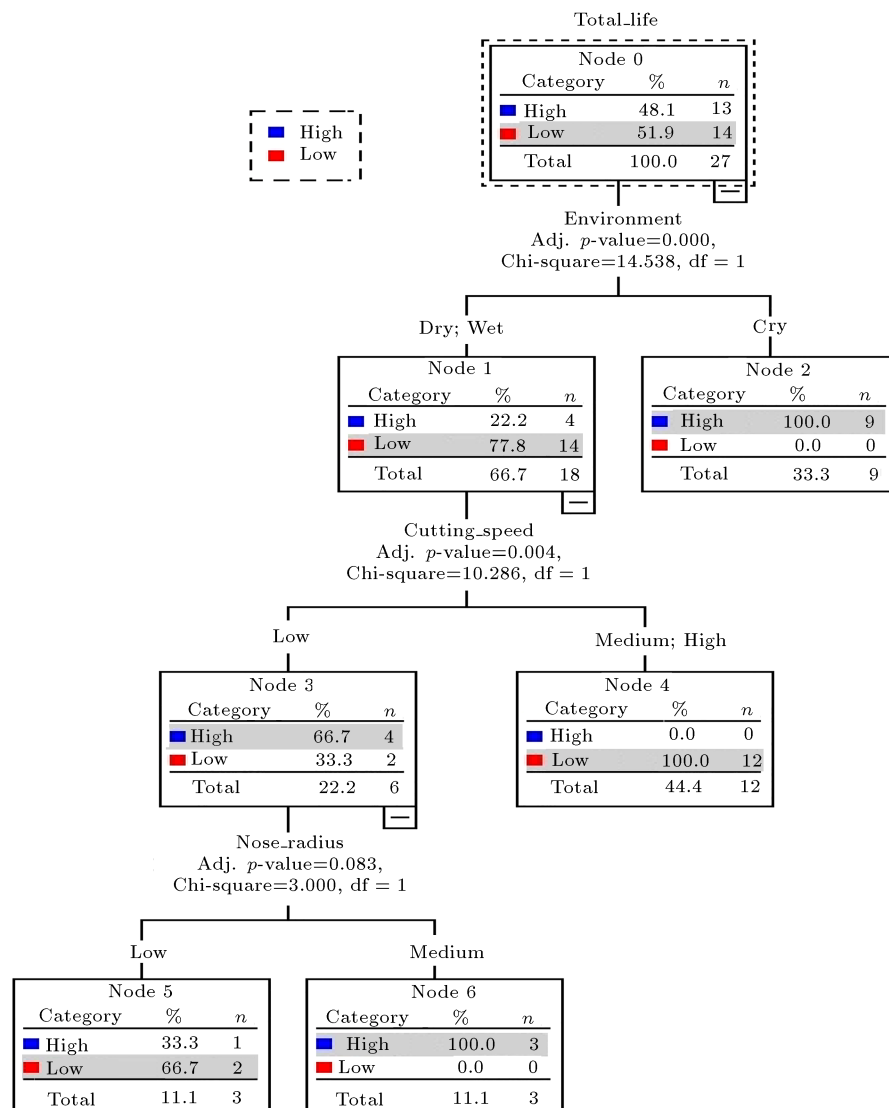


Figure 3. Classification tree for tool life using Chi-squared Automatic Interaction Detection (CHAID) algorithm.

Table 5. Classification of tool life based on Chi-squared Automatic Interaction Detection (CHAID).

Classification	Node	Characteristics
First	2	Cryogenic environment provides higher tool life.
Second	1,4	Dry or wet environment and medium or high CS (160 m/min or 200 m/min) lead to lower tool life.
Third	1,3,5	Dry or wet environment, low CS (120 m/min) and low NR (0.40 mm) are responsible for lower tool life.
Fourth	1,3,6	Dry or wet environment, low CS (120 m/min) and medium NR (0.80 mm) provide higher tool life.

Table 6. Percentages of accurate classification of tool life based on Chi-squared Automatic Interaction Detection (CHAID).

Classification	Tool life			
	Low (≤ 27.66 min)		High (> 27.66 min)	
	Number of observations	Percentage	Number of observations	Percentage
1	0	0%	9	100%
2	12	100%	0	0%
3	2	66.7%	1	33.3%
4	0	0%	3	100%

environment as the most important CNC turning parameter (69.27% contribution), followed by cutting speed (24.36% contribution). Depth of cut had almost no contribution (0.19%) on tool life. Using Signal-to-Noise (S/N) ratio values, it is recommended that the optimal combination of parameters for achieving a higher tool life are low cutting speed, low feed rate, low depth of cut, medium nose radius, and cryogenic environment which almost matches the combination proposed by a decision tree.

CHAID-based rules for tool life

Rule 1: If environment = cryogenic Then TL is (27.66–54.69]:

$$[P = 100\%, Q = 69.23\%, C = 33.33\%, QTY = 9]$$

$$[T = 202.56\%].$$

Rule 2: If environment = dry or wet and CS = medium or high Then TL is [14.67–27.66]:

$$[P = 100\%, Q = 85.71\%, C = 44.44\%, QTY = 12]$$

$$[T = 230.15\%].$$

Rule 3: If environment = dry or wet, CS = low and NR = low, Then TL is (27.66–54.69]:

$$[P = 66.70\%, Q = 14.28\%, C = 7.40\%, QTY = 2]$$

$$[T = 88.38\%].$$

Rule 4: If environment = dry or wet, CS = low and NR = medium, Then TL is [14.67–27.66]:

$$[P = 100\%, Q = 23.07\%, C = 11.11\%, QTY = 3]$$

$$[T = 134.18\%].$$

The decision tree for power consumption, which is developed using the CART algorithm, is shown in Figure 4. When the power consumption is less than or equal to 1373 W, it is designated as 'low' and when it is greater than 1373 W, its value is 'high'. As it is an STB type of response, its 'low' values are always preferred. The 'If-Then' rules extracted from the decision tree of Figure 4 highlight that dry or cryogenic machining environment and cutting speed less than or equal to 180 m/min always lead to lower power consumption (Rule 2 with total strength 224.44). Thus, a wet environment is responsible for higher power consumption (Rule 1 with total strength 208.33). Low feed rate and low depth of cut cause lower power consumption. The important plot of Figure 5 identifies the machining environment as the most critical CNC turning parameter affecting power consumption, followed by cutting speed. Interestingly, tool nose radius plays no significant role

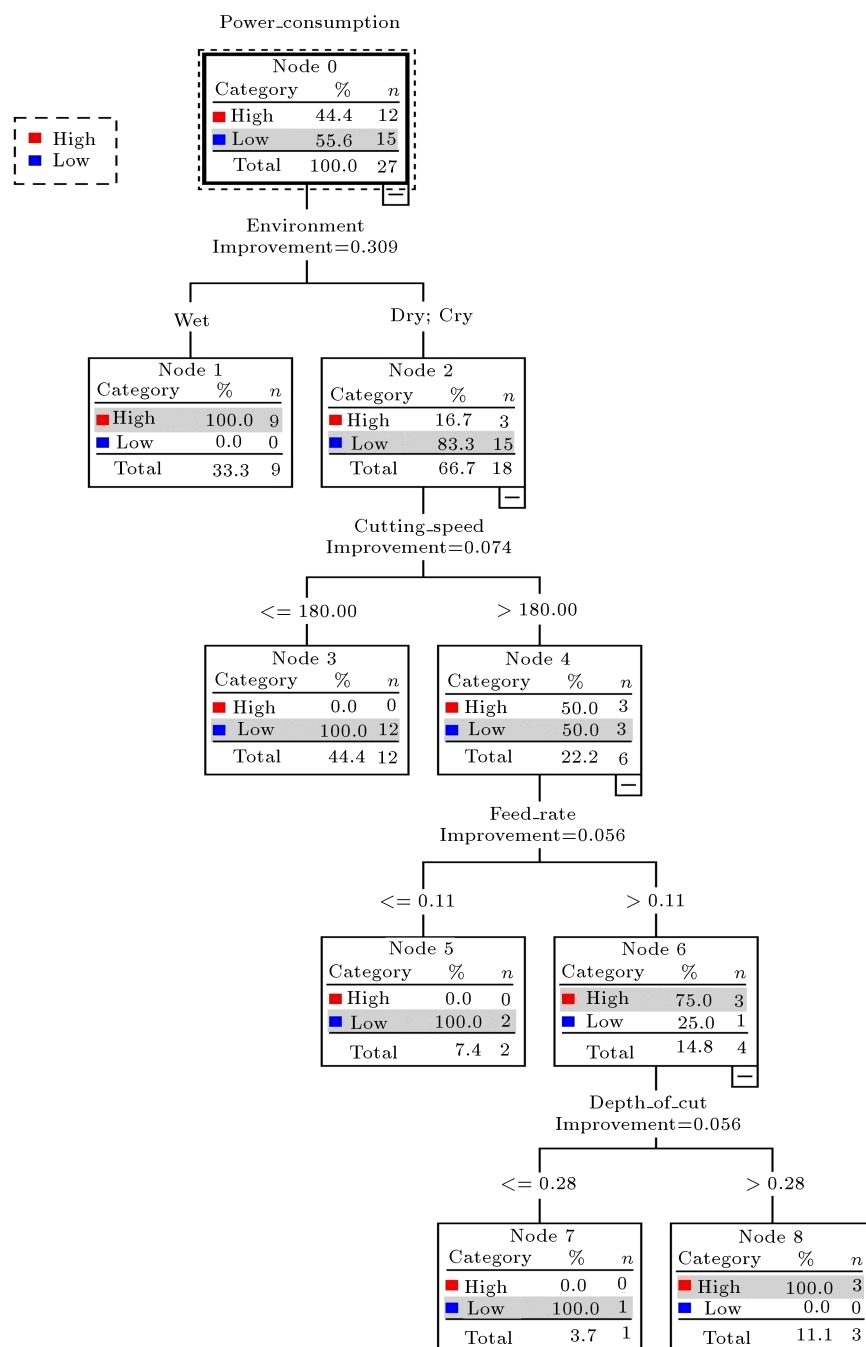


Figure 4. Classification tree for power consumption using Classification And Regression Tree (CART) algorithm.

in power consumption. The rules developed from the decision trees generated using CHAID algorithm (not shown here due to lack of space) also confirm these observations. Combining both the sets of decision rules from CART and CHAID algorithms, the optimal parametric mix of dry or cryogenic environment, low or medium cutting speed, low feed rate, and low nose radius would always lead to lower power consumption. Gupta et al. [4] also observed that machining environment = cryogenic, cutting speed = low, feed rate = low, depth of cut = low, and nose radius = medium

were responsible for attaining the most desirable value of lower power consumption in the said CNC turning centre.

CART-based rules for power consumption

Rule 1: If environment = wet Then *PC* is (1373–2094]:

$$[P = 100\%, Q = 75\%, C = 33.33\%, QTY = 9]$$

$$[T = 208.33\%].$$

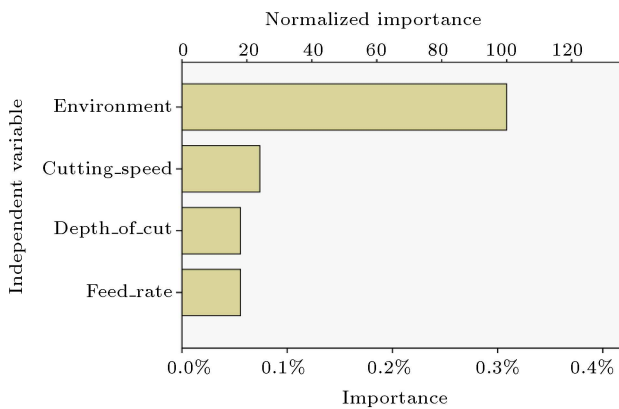


Figure 5. Importance of Computer Numerical Control (CNC) turning parameters affecting power consumption.

Rule 2: If environment = dry or cryogenic and $CS \leq 180$ m/min Then PC is [866–1373]:

$$[P = 100\%, Q = 80\%, C = 44.44\%, QTY = 12]$$

$$[T = 224.44\%].$$

Rule 3: If environment = dry or cryogenic, $CS > 180$ m/min and $FR \leq 0.11$ mm/rev Then PC is [866–1373]:

$$[P = 100\%, Q = 13.33\%, C = 7.40\%, QTY = 2]$$

$$[T = 120.73\%].$$

Rule 4: If environment = dry or cryogenic, $CS > 180$ m/min, $FR > 0.11$ mm/rev and $DOC \leq 0.28$ mm Then PC is [866–1373]:

$$[P = 100\%, Q = 6.67\%, C = 3.70\%, QTY = 1]$$

$$[T = 110.37\%].$$

Rule 5: If environment = dry or cryogenic, $CS > 180$ m/min, $FR > 0.11$ mm/rev and $DOC > 0.28$ mm Then PC is [866–1373]:

$$[P = 100\%, Q = 25\%, C = 11.11\%, QTY = 3]$$

$$[T = 136.11\%].$$

CHAID-based rules for power consumption

Rule 1: If environment = wet Then PC is (1373–2094):

$$[P = 100\%, Q = 75\%, C = 33.33\%, QTY = 9]$$

$$[T = 208.33\%].$$

Rule 2: If environment = dry or cryogenic and CS = low or medium Then PC is [866–1373]:

$$[P = 100\%, Q = 80\%, C = 44.44\%, QTY = 12]$$

$$[T = 224.44\%].$$

Rule 3: If environment = dry or cryogenic, CS = high and FR = low Then PC is [866–1373]:

$$[P = 100\%, Q = 13.33\%, C = 7.40\%, QTY = 2]$$

$$[T = 120.73\%].$$

Rule 4: If environment = dry or cryogenic, CS = high and FR = medium Then PC is [866–1373]:

$$[P = 50\%, Q = 6.67\%, C = 3.70\%, QTY = 1]$$

$$[T = 60.37\%].$$

Rule 5: If environment = dry or cryogenic, CS = high and FR = high then PC is (1373–2094):

$$[P = 100\%, Q = 16.67\%, C = 7.40\%, QTY = 2]$$

$$[T = 124.07\%].$$

In Figure 6, the decision tree for SR which is developed using the CART algorithm is exhibited. The corresponding ‘If-Then’ rules are also subsequently generated. In this case, the SR values less than or equal to $0.58 \mu\text{m}$ are termed as ‘low’ (satisfactory) and those with greater than $0.58 \mu\text{m}$ values are styled as ‘high’. Rule 1 with the maximum total strength of 148.40 reveals that when the machining environment is cryogenic and cutting speed is more than 140 m/min, the SR of the turned components would be satisfactory (low). A high nose radius also provides lower SR (Rule 3 with a total strength of 137.52). Similarly, a high feed rate is responsible for poor SR. The rules extracted from the decision tree which is developed using the CHAID algorithm (not shown here due to lack of space) prove that low or medium feed rate achieves better SR. In both the sets of rules, depth of cut appears to be an unimportant CNC turning parameter having no impact on SR. The importance of each of the turning parameters on SR is depicted in Figure 7 which clearly reveals the fact that machining environment and cutting speed are the two most significant parameters affecting SR. Depth of cut is the least important turning parameter. Gupta et al. [4] identified that a parametric combination of cutting speed = high, feed rate = medium, depth of cut = medium, nose radius = high, and machining environment = cryogenic would provide better SR of the turned components.

CART-based rules for SR

Rule 1: If environment = cryogenic and $CS > 140$ m/min Then SR is [0.16–0.58]:

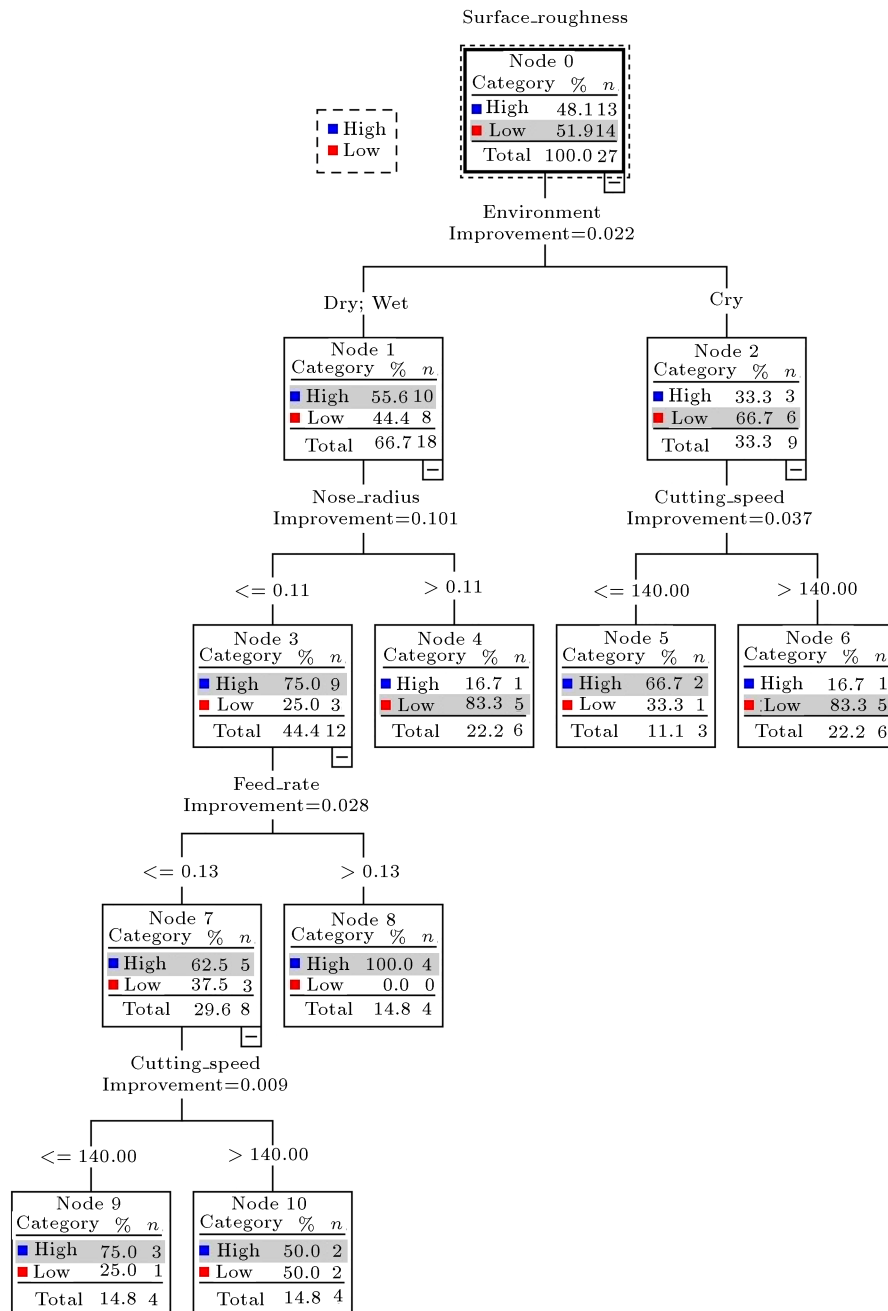


Figure 6. Classification tree for Surface Roughness (SR) using Classification And Regression Tree (CART) algorithm.

$[P = 83.33\%, Q = 42.85\%, C = 22.22\%, QTY = 6]$

$[T = 148.40\%]$.

Rule 2: If environment = cryogenic and $CS \leq 140$ m/min Then SR is (0.58–1.41]:

$[P = 66.70\%, Q = 15.38\%, C = 7.40\%, QTY = 2]$

$[T = 89.48\%]$.

Rule 3: If environment = dry or wet and $NR >$

1.00 mm Then SR is [0.16–0.58]:

$[P = 83.30\%, Q = 35.71\%, C = 18.51\%, QTY = 5]$

$[T = 137.52\%]$.

Rule 4: If environment = dry or wet, $NR \leq 1.00$ mm and $FR > 0.13$ mm/rev Then SR is (0.58–1.41]:

$[P = 100\%, Q = 30.77\%, C = 14.81\%, QTY = 4]$

$[T = 145.58\%]$.

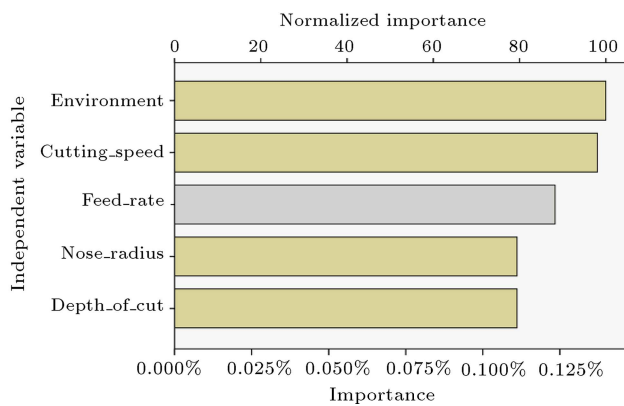


Figure 7. Importance plot for Surface Roughness (SR).

Rule 5: If environment = dry or wet, $NR \leq 1.00$ mm, $FR \leq 0.13$ mm/rev and $CS \leq 140$ m/min Then SR is (0.58–1.41]:

$$[P = 75\%, Q = 23.07\%, C = 11.11\%, QTY = 3]$$

$$[T = 109.18\%].$$

Rule 6: If environment = dry or wet, $NR \leq 1.00$ mm, $FR \leq 0.13$ mm/rev and $CS > 140$ m/min Then SR is (0.58–1.41]:

$$[P = 50\%, Q = 15.38\%, C = 7.41\%, QTY = 2]$$

$$[T = 72.79\%].$$

CHAID-based rules for SR

Rule 1: If environment = cryogenic and $CS = \text{high}$ Then SR is [0.16–0.58]:

$$[P = 100\%, Q = 21.43\%, C = 11.11\%, QTY = 3]$$

$$[T = 132.54\%].$$

Rule 2: If environment = cryogenic, $CS = \text{low or medium}$ and $FR = \text{high}$ Then SR is (0.58–1.41]:

$$[P = 100\%, Q = 15.38\%, C = 7.41\%, QTY = 2]$$

$$[T = 122.79\%].$$

Rule 3: If environment = cryogenic, $CS = \text{low or medium}$ and $FR = \text{low or medium}$ Then SR is [0.16–0.58]:

$$[P = 75\%, Q = 21.43\%, C = 11.11\%, QTY = 3]$$

$$[T = 107.54\%].$$

Rule 4: If environment = dry or wet, $NR = \text{high}$ and $CS = \text{high}$ Then SR is [0.16–0.58]:

$$[P = 100\%, Q = 21.43\%, C = 11.11\%, QTY = 3]$$

$$[T = 132.54\%].$$

Rule 5: If environment = dry or wet, $NR = \text{high}$ and $CS = \text{medium}$ Then SR is [0.16–0.58]:

$$[P = 66.70\%, Q = 14.28\%, C = 7.41\%, QTY = 2]$$

$$[T = 88.39\%].$$

Rule 6: If environment = dry or wet, $NR = \text{low or medium}$ and $FR = \text{high}$ Then SR is (0.58–1.41]:

$$[P = 100\%, Q = 30.77\%, C = 14.81\%, QTY = 4]$$

$$[T = 145.58\%].$$

Rule 7: If environment = dry or wet, $NR = \text{low or medium}$ and $FR = \text{low or medium}$ Then SR is (0.58–1.41]:

$$[P = 62.50\%, Q = 38.46\%, C = 18.52\%, QTY = 5]$$

$$[T = 119.48\%].$$

The decision tree for cutting force originated from CART algorithm is shown in Figure 8. The corresponding ‘If-Then’ rules are subsequently generated from this decision tree. When the values of cutting force are less than or equal to 171.30 N, they are denoted as ‘low’ and when its values are greater than 171.30 N, they are designated as ‘high’. An analysis of these rules reveals that when the machining environment is cryogenic, cutting speed is less than or equal to 180 m/min and feed rate is less than or equal to 0.11 mm/rev, and the achievable cutting force would be low. Similarly, a high depth of cut leads to higher cutting force. The rules extracted from the decision tree based on CHAID algorithm (not presented here) state that cryogenic environment would always provide lower cutting force. On the other hand, low depth of cut and low or medium nose radius are often responsible for attaining lower cutting force. When the relative importance of all the considered CNC turning parameters is plotted in Figure 9, it determines that the depth of cut is the most important parameter that affects the cutting force, followed by the machining environment and feed rate. Nose radius appears to be an insignificant CNC turning parameter for cutting force. An optimal parametric mix of moderate cutting speed, low feed rate, low depth of cut, moderate nose radius and cryogenic environment was identified by Gupta et al. [4] for lower cutting force, which almost corroborates with the decision trees-based observations.

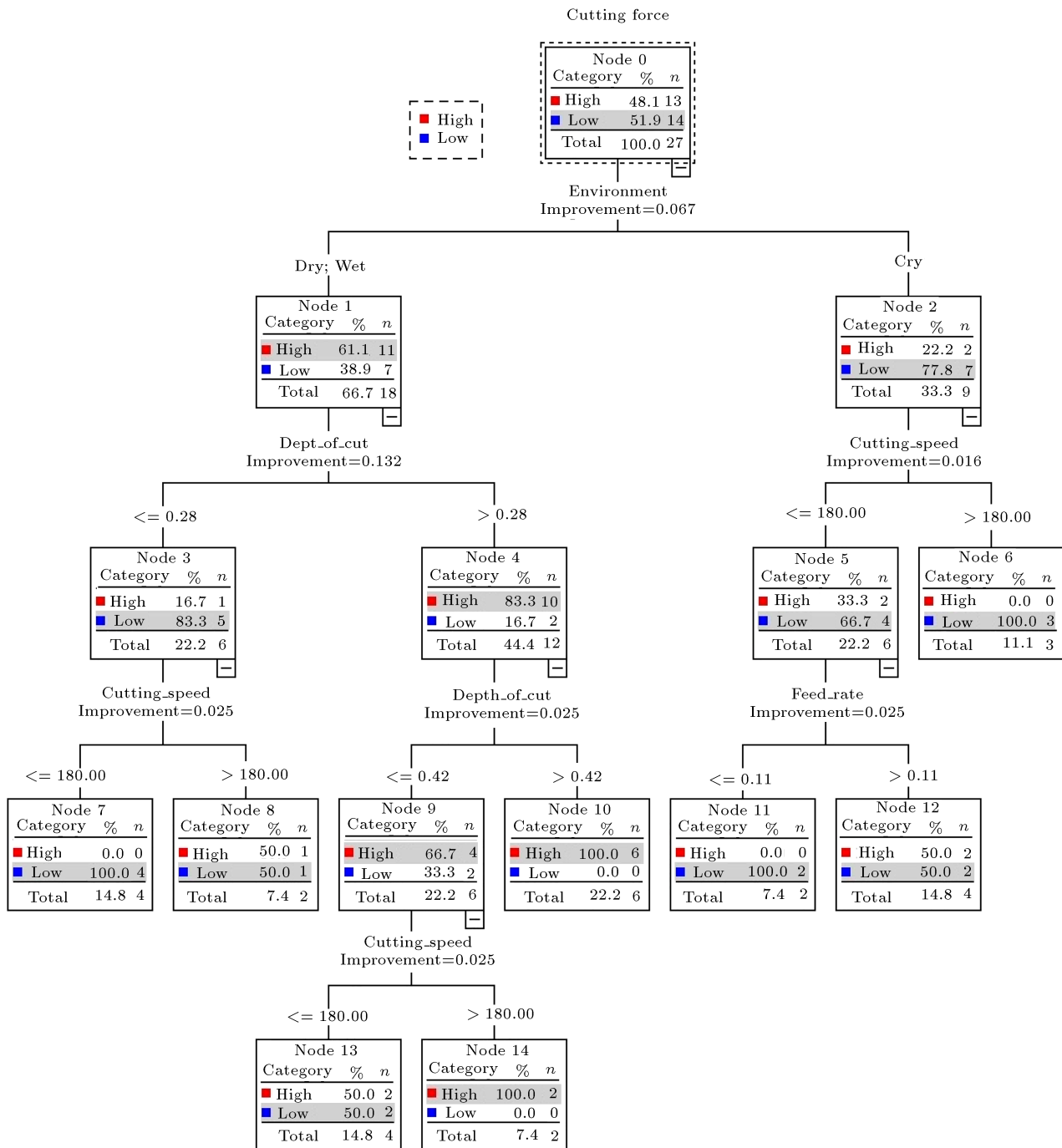
CART-based rules for cutting force

Rule 1: If environment = cryogenic and $CS > 180$ m/min Then CF is [106.23–171.30]:

$$[P = 100\%, Q = 21.43\%, C = 11.11\%, QTY = 3]$$

$$[T = 132.54\%].$$

Rule 2: If environment = cryogenic, $CS \leq 180$ m/min and $FR \leq 0.11$ mm/rev Then CF is [106.23–171.30]:



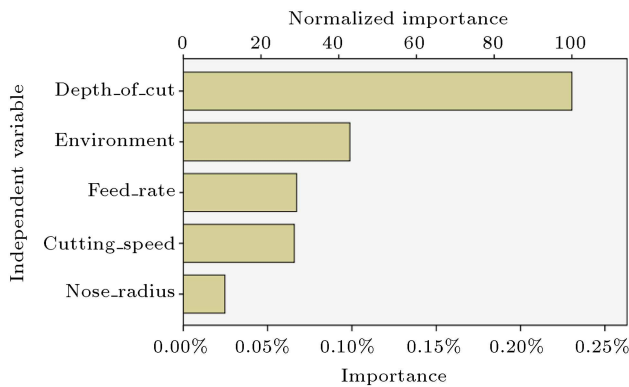


Figure 9. Importance of Computer Numerical Control (CNC) turning parameters affecting cutting force.

$$[P = 100\%, Q = 7.69\%, C = 3.70\%, QTY = 1]$$

$$[T = 111.39\%].$$

Rule 6: If environment = dry or wet, $DOC > 0.28$ mm Then CF is $[171.30\text{--}286.90]$:

$$[P = 100\%, Q = 46.15\%, C = 22.22\%, QTY = 6]$$

$$[T = 168.37\%].$$

Rule 7: If environment = dry or wet, $DOC > 0.28$ mm or ≤ 0.42 mm, and $CS \leq 180$ m/min Then CF is $[106.23\text{--}171.30]$:

$$[P = 50\%, Q = 14.28\%, C = 7.41\%, QTY = 2]$$

$$[T = 71.69\%].$$

Rule 8: If environment = dry or wet, $DOC > 0.28$ mm or ≤ 0.42 mm, and $CS > 180$ m/min Then CF is $[171.30\text{--}286.90]$:

$$[P = 100\%, Q = 15.38\%, C = 7.41\%, QTY = 2]$$

$$[T = 122.79\%].$$

CHAID-based rules for cutting force

Rule 1: If environment = cryogenic Then CF is $[106.23\text{--}171.30]$:

$$[P = 77.80\%, Q = 50\%, C = 25.92\%, QTY = 7]$$

$$[T = 153.72\%].$$

Rule 2: If environment = dry or wet and DOC = low Then CF is $[106.23\text{--}171.30]$:

$$[P = 83.30\%, Q = 35.71\%, C = 18.52\%, QTY = 5]$$

$$[T = 136.53\%].$$

Rule 3: If environment = dry or wet, DOC =

medium or high and NR = low or high Then CF is $[171.30\text{--}286.90]$:

$$[P = 100\%, Q = 61.54\%, C = 29.63\%, QTY = 8]$$

$$[T = 191.17\%].$$

Rule 4: If environment = dry or wet, DOC = medium or high and NR = medium Then CF is $[106.23\text{--}171.30]$:

$$[P = 50\%, Q = 14.29\%, C = 7.41\%, QTY = 2]$$

$$[T = 71.70\%].$$

In Table 7, a comparison of the classification accuracies for CART and CHAID algorithms for all the four responses is provided. From this table, it can be noted that for tool life and power consumption responses, CART algorithm can perfectly predict low and high tool life, and low and high power consumption values. The classification accuracies for high and low SR are 84.6% and 78.6% respectively. Similarly, using CART algorithm, high and low cutting forces can be predicted with accuracies of 85.7% and 76.9% respectively. Thus, CART algorithm can almost perfectly predict both the high and low values of all the considered responses, although it has a slightly greater tendency to accurately estimate high values of the responses. In case of CHAID algorithm, it can perfectly predict low values of tool life, power consumption and cutting force. High values of tool life and power consumption are predicted with 92.3% and 91.7% accuracies respectively. It has prediction accuracies of 84.6% and 71.4% for high and low SR values respectively. The classification accuracy for high cutting force is only 61.5%. Thus, it can be concluded that CHAID algorithm performs better in predicting low values of the considered responses. The overall classification accuracies of both these algorithms for the four responses are provided in Table 8. With respect to overall classification accuracy, CART algorithm outperforms CHAID in almost exactly predicting the responses of the CNC turning process under consideration. The corresponding values of Standard Error (SE) for CART are also comparatively low as compared to CHAID algorithm.

In order to visualize the effects of changing values of the responses of the considered CNC turning process on the prediction performance of CART and CHAID algorithms, a sensitivity analysis study is performed here. In this approach, incremental changes are made in the response values of the experimental dataset based on the equation:

$$R_N = R_O + (2 \times RAND() - 1) \times E \times R_O,$$

where R_O is the original response value, $RAND()$ is a uniform random number generator function between 0

Table 7. Classification accuracies for Tool Life (TL), Power Consumption (PC), Surface Roughness (SR), and Cutting Force (CF) using Classification And Regression Tree (CART) and Chi-squared Automatic Interaction Detection (CHAID) algorithms.

Response	Observed	CART			CHAID		
		Predicted			Predicted		
		High (> 27.66 min)	Low (≤ 27.66 min)	Percent correct	High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct
TL	High (> 27.66 min)	13	0	100%	12	1	92.3%
	Low (≤ 27.66 min)	0	14	100%	0	14	100%
	Overall percentage	48.1%	51.9 %	100%	44.4%	55.6%	96.2%
PC	Observed	Predicted			Predicted		
		High (> 1373 W)	Low (≤ 1373 W)	Percent correct	High (> 1373 W)	Low (≤ 1373 W)	Percent correct
		High (> 1373 W)	Low (≤ 1373 W)	Percent correct	High (> 1373 W)	Low (≤ 1373 W)	Percent correct
PC	High (> 1373 W)	12	0	100%	11	1	91.7%
	Low (≤ 1373 W)	0	15	100%	0	15	100%
	Overall percentage	44.4%	55.6 %	100%	40.7%	59.3%	95.8%
SR	Observed	Predicted			Predicted		
		High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct	High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct
		High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct	High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct
SR	High ($> 0.58 \mu\text{m}$)	11	2	84.6 %	11	2	84.6%
	Low ($\leq 0.58 \mu\text{m}$)	3	11	78.6 %	4	10	71.4%
	Overall percentage	51.8 %	48.1 %	81.6 %	55.5%	44.4%	78.0%
CF	Observed	Predicted			Predicted		
		High (> 171.30 N)	Low (≤ 171.30 N)	Percent correct	High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct
		High (> 171.30 N)	Low (≤ 171.30 N)	Percent correct	High ($> 0.58 \mu\text{m}$)	Low ($\leq 0.58 \mu\text{m}$)	Percent correct
CF	High (> 171.30 N)	12	1	85.7 %	8	5	61.5%
	Low (≤ 171.30 N)	4	10	76.9 %	0	14	100%
	Overall percentage	59.2 %	40.8 %	81.3 %	29.6%	70.4%	80.7%

and 1, E is the relative error level and R_N is the new perturbed response value.

The relative error levels are set here as 5, 10, 15, 20, and 25%. The classification accuracies of both the algorithms at varying errors levels are provided in Table 9. It can be clearly propounded that the prediction performance of the CART algorithm is least affected by the changing response values in the experimental dataset and it is a more robust technique as compared to the CHAID algorithm.

Over the past few decades, decision tree algo-

rithms, like CART and CHAID, have been in extensive use for solving predictive analytics problems. As they are generic models based on effective calculation procedures, they can easily arrive at the optimal solutions for a given classification/prediction problem. Decision trees generated by these algorithms are efficient managerial tools that present all the decisions/outcomes in the form of a flowchart with branches and leaves. Decision trees thus solve problems of machine learning by transforming the data into a tree representation. Each branch of the tree symbolizes a decision option.

Table 8. Risk of classifying Surface Roughness (SR), Tool Life (TL), Cutting Force (CF) and Power Consumption (PC).

Response	Method	Accuracy	Standard error
TL	CART	1.000	0.001
	CHAID	0.963	0.036
PC	CART	1.000	0.001
	CHAID	0.958	0.036
SR	CART	0.815	0.075
	CHAID	0.778	0.080
CF	CART	0.818	0.075
	CHAID	0.807	0.036

The leaves at the end of the branches show the possible outcomes. Decision trees can deal with quantitative, qualitative, or categorical attributes by assigning objects to a specific class in a classification problem. A decision tree is one of the simplest and most popular classification algorithms to learn, understand and interpret. They have several advantages, like the requirement of less computational effort for data preparation during pre-processing, no need for normalization and scaling of data, least affectability towards the missing observation in the dataset, provision of explanation about how a particular decision has been reached etc. Similarly, they also suffer from some disadvantages, like the requirement of higher time for training, a small change in data may cause a large change in the tree structure causing instability, inability to be applied for regression, and prediction of continuous variables.

5. Conclusions

This paper deals with the application of a data mining tool in the form of the development of decision trees using Classification And Regression Tree (CART) and Chi-squared Automatic Interaction Detection (CHAID) algorithms to determine the most

preferable combinations of various machining parameters in a Computer Numerical Control (CNC) turning process. The ‘If-Then’ rules extracted from both the decision trees would guide the concerned process engineers in investigating the effects of the input parameters on the considered responses. Based on the detailed analysis of the corresponding decision trees and decision rules, the following conclusions can be derived:

- For achieving higher tool life, cryogenic environment, and low values of cutting speed, tool nose radius and feed rate need to be set. Depth of cut has almost no effect on tool life;
- A combination of cryogenic environment, low or medium cutting speed, low feed rate, and low nose radius are responsible for lower power consumption;
- To attain lower surface roughness of the turned components, cryogenic environment, high cutting speed, low or medium feed rate, and high nose radius would be the recommended setting for the said CNC turning process. Depth of cut plays in significant role on surface roughness;
- A parametric mix of cryogenic environment, low cutting speed, low or medium feed rate, low depth of cut, and low or medium nose radius would provide lower cutting speed;
- The CART algorithm supersedes the CHAID algorithm with respect to higher overall classification accuracy and lower prediction risk. Although, for some of the responses, CART generates a slightly more number of decision rules as compared to CHAID, it can almost perfectly predict high as well as low values of all the responses;
- Between these two algorithms, CART has a higher capacity to predict high values of the responses, whereas, CHAID performs better for low responses values;
- Based on the sensitivity analysis study, the prediction performance of the CART algorithm is observed to be least affected by the perturbed response values in the experimental dataset as compared to the CHAID algorithm.

Table 9. Classification accuracies of Classification And Regression Tree (CART) and Chi-squared Automatic Interaction Detection (CHAID) algorithms at various error levels.

Response	Error level									
	CART					CHAID				
	5%	10%	15%	20%	25%	5%	10%	15%	20%	25%
TL	100	100	100	92.6	92.6	92.6	92.6	92.6	82.6	85.2
PC	100	100	96.3	96.3	92.6	100	100	92.6	96.3	96.3
SR	88.9	81.5	85.2	85.2	81.5	81.5	80.9	80.45	81.5	79.0
CF	92.6	82.6	92.6	92.6	88.9	88.9	77.8	88.9	88.9	85.2

The most preferable parametric settings for the considered CNC turning process are observed to be in close agreement with those derived by the past researchers based on Taguchi methodology, which proves the efficacy of the developed decision as revealed clearly by the study of the material removal mechanism. These classification algorithms can thus be applied to any machining process to investigate the effects of different input parameters on the responses and identify the best machining conditions for enhanced process monitoring and control.

As the decision rules mainly focus on classification, they often neglect predicting the interrelationships between the input parameters and responses in the form of regression. While a continuous variable is divided into intervals and turned into a classification problem, there is a high possibility of a loss of valuable information. Thus, there must be always a trade-off between predictive accuracy and computational effort to arrive at the most appropriate set of decision rules.

References

1. Trent, E.M., *Metal Cutting*, Woburn, Massachusetts: Butterworth-Heinemann (2010).
2. Suh, S.-H., Kang, S.K., Chung, D.-H., et al., *Theory and Design of CNC Systems*, Springer (2008).
3. Park, K.S. and Kim, S.H. "Artificial intelligent approaches to determination of CNC machining parameters in manufacturing: A review", *Artificial Intelligence in Engineering*, **12**, pp. 127–134 (1997).
4. Gupta, A., Singh, H., and Aggarwal, A. "Taguchi-fuzzy multi output optimization (MOO) in high speed CNC turning of AISI P-20 tool steel", *Expert Systems with Applications*, **38**, pp. 6822–6828 (2011).
5. Mukherjee, S., Kamal, A., and Kumar, K. "Optimization of material removal rate during turning of SAE 1020 material in CNC lathe using Taguchi technique", *Procedia Engineering*, **97**, pp. 29–35 (2014).
6. Marko, H., Simon, K., Tomaz, I., et al. "Turning parameters optimization using particle swarm optimization", *Procedia Engineering*, **69**, pp. 670–677 (2014).
7. Saini, S.K. and Pradhan, S.K. "Optimization of multi-objective response during CNC turning using Taguchi-fuzzy application", *Procedia Engineering*, **97**, pp. 141–149 (2014).
8. Vasudevan, H., Deshpande, N.C., and Rajguru, R.R. "Grey fuzzy multiobjective optimization of process parameters for CNC turning of GFRP/epoxy composites", *Procedia Engineering*, **97**, pp. 85–94 (2014).
9. Saini, S.K. and Pradhan, S.K. "Optimization of machining parameters for CNC turning of different materials", *Applied Mechanics and Materials*, **592–594**, pp. 605–609 (2014).
10. Aghdeab, S.H., Mohammed, L.A., and Ubaid, A.M. "Optimization of CNC turning for aluminum alloy using simulated annealing method", *Jordan Journal of Mechanical and Industrial Engineering*, **9**(1), pp. 39–44 (2015).
11. Camposeco-Negrete, C. "Optimization of cutting parameters using response surface method for minimizing energy consumption and maximizing cutting quality in turning of AISI 6061 T6 aluminum", *Journal of Cleaner Production*, **91**, pp. 109–117 (2015).
12. Sankaya, M. and Güllü, A. "Multi-response optimization of minimum quantity lubrication parameters using Taguchi-based grey relational analysis in turning of difficult-to-cut alloy Haynes 25", *Journal of Cleaner Production*, **91**, pp. 347–357 (2015).
13. Asiltürk, İ., Neşeli, S., and İnce, M.A. "Optimisation of parameters affecting surface roughness of Co28Cr6Mo medical material during CNC lathe machining by using the Taguchi and RSM methods", *Measurement*, **78**, pp. 120–128 (2016).
14. Kumar, U., Singh, A., and Kumar, R. "Optimization of machining parameters for tool wear rate and material removal rate in CNC turning by grey relational analysis", *International Journal of Applied Engineering Research*, **11**(4), pp. 2771–2775 (2016).
15. Maheswara Rao, C. and Venkatasubbaiah, K. "Optimization of surface roughness in CNC turning using Taguchi method and ANOVA", *International Journal of Advanced Science and Technology*, **93**, pp. 1–14 (2016).
16. Klancnik, S., Hrelja, M., Balic, J., et al. "Multi-objective optimization of the turning process using a gravitational search algorithm (GSA) and NSGA-II approach", *Advances in Production Engineering & Management*, **11**, pp. 366–376 (2016).
17. Bilga, P.S., Singh, S., and Kumar, R. "Optimization of energy consumption response parameters for turning operation using Taguchi method", *Journal of Cleaner Production*, **137**, pp. 1406–1417 (2016).
18. Kushwaha, A. and Singh, R. "Optimization of CNC process parameters on Inconel 625 using response surface methodology", *International Journal of Mechanical Engineering and Technology*, **8**(7), pp. 1830–1836 (2017).
19. Nataraj, M. and Balasubramanian, K. "Parametric optimization of CNC turning process for hybrid metal matrix composite", *International Journal Advanced Manufacturing Technology*, **93**, pp. 215–224 (2017).
20. Nayak, N.K. and Sodhi, H.S. "Optimization of CNC turning parameters for Al-6061 using response surface methodology", *International Journal of Mechanical and Production Engineering Research and Development*, **7**(4), pp. 127–138 (2017).
21. Sahoo, P., Pratap, A., and Bandyopadhyay, A. "Modeling and optimization of surface roughness and tool vibration in CNC turning of Aluminum alloy using hybrid RSM-WPCA methodology", *International*

- Journal of Industrial Engineering Computations*, **8**, pp. 385–398 (2017).
22. Mandal, N.K., Mondal, M., and Singh, N.K. “Modelling and optimisation of a sustainable manufacturing process with CNC turning centre”, *International Journal of Applied Environmental Sciences*, **12**, pp. 1101–1116 (2017).
 23. Suresh, M., Meby Selvaraj, R., Rajkumar, K., et al. “Optimisation of cutting parameters in CNC turning of EN-19 using tungsten carbide”, *International Journal of Computer Aided Engineering and Technology*, **9**(2), pp. 218–228 (2017).
 24. Akkuş, H. “Optimising the effect of cutting parameters on the average surface roughness in a turning process with the Taguchi method”, *Materials and Technology*, **52**, pp. 781–785 (2018).
 25. Bhanu Prakash, P., Brahma Raju, K., Venkata Subbaiah, K., et al. “Application of Taguchi based grey method for multi aspects optimization on CNC turning of AlSi₇ Mg”, *Materials Today: Proceedings*, **5**, pp. 14292–14301 (2018).
 26. Gadekula, R.K., Potta, M., Kamisetty, D., et al. “Investigation on parametric process optimization of HCHCR in CNC turning machine using Taguchi technique”, *Materials Today: Proceedings*, **5**, pp. 28446–28453 (2018).
 27. Palanisamy, D. and Senthil, P. “Application of grey-fuzzy approach for optimization of CNC turning process”, *Materials Today: Proceedings*, **5**, pp. 6645–6654 (2018).
 28. Sahoo, P., Satpathy, M.P., Singh, V.K., and Bandyopadhyay, A. “Performance evaluation in CNC turning of AA6063-T6 alloy using WASPAS approach”, *World Journal of Engineering*, **15**(6), pp. 700–709 (2018).
 29. Saravanakumar, A., Karthikeyan, S.C., Dhamotharan, B., et al. “Optimization of CNC turning parameters on aluminum alloy 6063 using Taguchi robust design”, *Materials Today: Proceedings*, **5**, pp. 8290–8298 (2018).
 30. Nataraj, M., Balasubramanian, K., and Palanisamy, D. “Optimization of machining parameters for CNC turning of Al/Al₂O₃ MMC using RSM approach”, *Materials Today: Proceedings*, **5**, pp. 14265–14272 (2018).
 31. Vasudevan, H., Rajguru, R., Tank, K., et al. “Optimization of multi-performance characteristics in the turning of GFRP (E) composites using principle component analysis combined with grey relational analysis”, *Materials Today: Proceedings*, **5**, pp. 5955–5967 (2018).
 32. Rao, V.D.P., Mahaboob Ali, S.R.S., Saqheed Ali, S.M.Z.M., et al. “Multi-objective optimization of cutting parameters in CNC turning of stainless steel 304 with TiAlN nano coated tool”, *Materials Today: Proceedings*, **5**, pp. 25789–25797 (2018).
 33. Vijay Kumar, M., Kiran Kumar, B.J., and Rudresha, N. “Optimization of machining parameters in CNC turning of stainless steel (EN19) by Taguchi’s orthogonal array experiments”, *Materials Today: Proceedings*, **5**, pp. 11395–11407 (2018).
 34. Arun Vikram, K., Lakshmi, V.V.K., and Venkata Praveen, A.M. “Evaluation of process parameters using GRA while machining low machinability material in dry and wet conditions”, *Materials Today: Proceedings*, **5**, pp. 25477–25485 (2018).
 35. Chau, N.L., Nguyen, M.-Q., Dao, T.-P., et al. “An effective approach of adaptive neuro-fuzzy inference system-integrated teaching learning-based optimization for use in machining optimization of S45C CNC turning”, *Optimization and Engineering*, **20**, pp. 811–832 (2019).
 36. Susanti, Y., Zukhronah, E., Pratiwi, H., et al. “Analysis of chi-square automatic interaction detection (CHAID) and classification and regression tree (CRT) for classification of corn production”, *IOP Conference Series: Journal of Physics*, **909**, pp. 1–8 (2017).
 37. Rokach, L., and Maimin, O., *Data Mining with Decision Trees: Theory and Applications*, World Scientific Publishing Co., NJ, USA (2014).
 38. Han, J., Kamber, M., and Pei, J., *Data Mining Concepts and Techniques*, Elsevier Inc., USA (2012).
 39. Breiman, L., Friedman, J.H., Olshen, R.A., et al., *Classification and Regression Tree.*, Chapman and Hall, New York (1993).
 40. Kass, G.V. “An exploratory technique for investigating large quantities of categorical data”, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **29**(2), pp. 119–127 (1980).
 41. Pitombo, C.S., de Souza, A.D., and Lindner, A. “Comparing decision tree algorithms to estimate intercity trip distribution”, *Transportation Research Part C: Emerging Technologies*, **77**, pp. 16–32 (2017).
 42. Sadoyan, H., Zakarian, A., and Mohanty, P. “Data mining algorithm for manufacturing process control”, *International Journal of Advanced Manufacturing Technology*, **28**, pp. 342–350 (2006).

Biographies

Shruti Sudhakar Dandge is currently working as a Lecturer in Mechanical Engineering at Government Polytechnic, Murtizapur, India since February 2011. Presently she is pursuing a PhD in Engineering from Jadavpur University, Kolkata, India. Born in India, she received the B Eng degree (Mechanical Engineering) in 2006 from V.Y.W.S.’s Professor Ram Meghe Institute of Technology and Research, Badnera-Amravati. She qualified for GATE (Mechanical) in 2007 and holds her M-Tech degree in Thermal Engineering from Government Engineering College, Amravati in 2009. Her current research interests are the application of computational tools including statistics, machine learning, and data mining techniques in the manufacturing domain. She also carried out an industrial project named

“Thermal and Mechanical design of shell and tube heat exchangers using fixed tube-sheets” at L&T, Mumbai (India). She is interested in areas of thermodynamics, heat transfer, and analysis of I.C. Engine.

Shankar Chakraborty is a Professor in Production Engineering Department of Jadavpur University. He had been graduated in 1986 from the University of Calcutta and had obtained his postgraduate degree from Jadavpur University in 1989. He had been

awarded PhD (Engg.) degree from Jadavpur University in 1994. His research interests include applications of different multi-criteria decision-making methods in a manufacturing environment, control chart pattern recognition, and development of MIS and ERP systems for diverse engineering applications, soft computing techniques etc. He has guided several ME and PhD (Engg.) theses and published numerous papers in international journals. He is also a regular reviewer of several journals of international repute.