



## Providing A Model for Predicting and Detecting Destructive Processes to Prevent the Production of Waste and Defective Products: A Data Mining Approach

N. Safaie\*, M. R. Saadatmand, S. A. Nasri

Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

### P A P E R I N F O

#### Paper history:

Received 14 February 2021

Accepted in revised form 01 September 2021

#### Keywords:

Data mining

Decision tree

Defective production

Process control

Statistical quality control

### A B S T R A C T

Today, most industries use statistical quality control tools to improve quality and reduce the defective products and waste, but the high volume of data requires the help of a powerful tool to control processes. One of the objectives of the present study is to predict defective products and prevent their production using data mining tools due to the high power in data analysis and its predictive nature, which is less used in the industry. In this study, the statistical population of all parts produced in 2017 by Shabrun Company. The statistical sample is 2400 pieces of radiators that were randomly selected from the production line. In the operational phases of data mining, three decision tree algorithms were used: C&R Tree, Quest Tree and Chaid Tree. Using these algorithms, the most important criteria affecting quality control and rules leading to the quality of parts were determined. Comparative results showed that despite the validity of all three algorithms, the C&R Tree algorithm had the highest accuracy. Adherence to the rules resulting from the implementation of these algorithms has led to the detection and prevention of waste generation, which has increased efficiency and prevented the loss of time and cost in this production unit.

doi: 10.5829/ijee.2021.12.03.09

## INTRODUCTION

The necessity for mass production in industrial factories is that the products have a stable and desirable quality which is based on the following three principles:

1. Technical knowledge in the preparation of what is required for production.
2. Engineering activities for the implementation of the information obtained from the first stage, namely, the collection of scientific information, that is, the correct selection of systems and machinery for the production of each product in the best possible terms.
3. Correct management of production affairs, which plays the most important role, and its new systems, especially in the industries, have been designed and implemented in many advanced countries [1].

These three principles are closely related to each others, and the lack or weakness of one affects other efficiency. By the use of these factors, it is possible to

determine the quantitative and qualitative characteristics of a product or material, or service, or instrument, and adapt them with the desired features. When choosing raw materials, machines, services and services, they can use these features as criteria for the assessment and selection of contracts. In case of non-receipt of features, they are subject to the rejection or modification of facilities and products [2].

Statistical quality control has many tools and techniques that enable a firm to enhance its productivity and product or service quality [3]. The achievement of the integrity of the desired characteristics of a product or service, which is the secret of the sustainability of the survival of any phenomenon. The quality and control of it bring the enjoyment of a healthy, faultless, and reliable product, service [4]. In the opinion of the Standards Institute, the conformity of product characteristics to national and international standards, and above all, from the Office of Food, Drinking, Cosmetic and Hygiene

\*Corresponding Author Email: [nsafaie@kntu.ac.ir](mailto:nsafaie@kntu.ac.ir) (N. Safaie)

Supervision, the proportion of the product to ensure the public health and to be free from any harmful and undesirable factor, quality concepts from the viewpoint of different groups and organizations in terms of the manufacturer, the competitiveness of the product in the consumer market and its profitability. Therefore, products can not be launched into a market without controlling their quality [5]. Today, statistical quality control is an inseparable part of the production chain, which is usually controlled at the last stage of the product, and if it is defected, it will be sent to the waste sector and, if it is sound correct, it will be sent to the sales department. Thus, it is necessary that quality control is done by using strong tools. On the other hand, given the high data mining capabilities in data analysis and the use of raw data and the predictive value of data mining, it can be helpful in controlling the quality of statistical data. The decision tree is an efficient and special way to create categorizers or classifiers of the production data of a decision tree [6]. The representation tree of the decision uses widely the rational method. There are a large number of decision tree inductive algorithms that are mainly described in machine learning and applied statistical literature [7]. They are supervisory learning methods to create decision trees from a set of input-output samples which will be investigated in this paper. In this work, using the decision tree in Clementine 12 software, a method for reducing waste and improving the quality of the manufactured components will be used and the results of this research will be used to improve the manufacturing processes. Reducing waste has played an important role in improving the productivity of the manufacturing system [8]. If the waste is reduced, the proportion of the product to the raw material will increase, which will increase the circulation of production and profits of the production system [9]. The presented method is used in this paper, which results in improved manufacturing processes. The method in this study will identify processes that, if it is implemented, will increase the quality of the final parts, and the most effective processes on the quality of the components will be identified. In this research, a decision tree algorithm is implemented on the quality control database of Shabrun company would result in the extraction of effective rules on the quality of the manufactured components and increase the efficiency and prevent waste of time and cost in this unit.

There are various graphs and processes in statistical quality control, but none of them are predictive. Accordingly, in this research, with a new perspective and with the help of decision tree algorithms in data mining, by modeling the quality of parts in Shabrun Company, a new approach has been taken to this issue and an attempt has been made to predict and discover processes leading to waste and defective productions. Based on the achieved rules, the decision trees have been drawn. Reached from these decision trees, the processes that reduce the quality of the part produce, which is less than

30% of the quality according to industrial experts, are removed from the production line. Therefore, in the production system, by controlling the production line, the poor quality of parts can be prevented. In order to manufacture productions, the proper processes can be selected that increase the quality of products. As a result, identifying, discovering and eliminating processes that reduce quality is the main objective of this study.

## LITERATURE REVIEW

Ashrafi et al. [10] investigated on integration of reliability and Six Sigma as a quality control tool to improve the status of manufactured products. The method improved and enhanced the quality of products. In this paper, the main goal is to measure the level of product sigma in the construction phase using product trustworthiness. This method simultaneously improves and increases the level of product quality and product reliability during production phases. In this regard, the decision tree is one of the most widely used method. The algorithms which can produce understandable human descriptions of relationships in a dataset [11]. In this research, the decision tree is intended to improve the production process and product quality [12]. The obtained results were compared with the results of the Six Sigma method [13]. In other words, It is stated that the association rules, decision trees and data mining techniques are well-known in finding predictive rules.

Bagheri et al. [14] presented a hybrid data mining model using real data that are responsible for the assembling and sells of ATMs to banks and financial institutions. The purpose of this paper is to provide a decision support system for company managers to use extracted knowledge from raw data for decision making and designing strategies. The stages of conducting this study were based on the Crisp method and the Clementine software used in modeling. The results showed that the company technical experts have different functions repairing duration of the similar devices, which depends on the experience of experts, their training, and so on. In addition, similar devices in different provinces have different problem codes due to weather conditions, the culture of using the device, and so on. The company uses the output of this model in its management decisions and its determination of strategies. Smith et al. [15] have presented a method for reducing air pollution and making electro-static filters. In their research, data mining was, the most important features that play major role in the quality improvement and the waste reduction of air purifier filters manufacturing. In the next step, they predicted the quality of the produced filters by using simulation. Their results also indicated that the die-cast phase was the most important part in the process of filter manufacturing. Azyazov et al. [16] identified errors in the construction of wind turbines by using video data mining.

In their work, artificial intelligence algorithms was applied; the method reduced the defected quality of turbines. In the following, the rules lead to waste generation extracted by using the decision tree algorithm. In their study, by discovering the rules of causing waste, an appropriate mechanism was developed to improve the quality of wind turbines. Lu and Ma [17] presented two hybrid decision tree-based models (CEEMDAN-XGBoost and 377 CEEMDAN-RF) for the water quality prediction. The CEEMDAN decomposed the raw data with large fluctuations to enhance the prediction performance of XGBoost and 379 RF and analyzed the water quality of the Gales Creek site of Tualatin River in 380 Oregon, USA. Vivancos et al. [18] presented application of a Change Point Detection method based on statistical quality control charts for changing in activities a family home by using typical monitoring data. This technique could enhance the the quality of information provided in energy feedback significantly. Hasheminezhad et al. [19], Istoto et al. [20] and Bayu et al. [21] focused on enhancing the conversion efficiency of oil and petroleum production waste. Ataee et al. [22] proposed the novel approach Based on a Deep Convolutional Neural Network to decrease human errors and enhanced accuracy of waste detection.

**PROBLEM STATEMENT**

Predicting product quality in any process requires modeling and discovering the relationship between factors that can affect the quality of process output. The popularity of data mining can be attributed to the ability of model and predict especially decision tree algorithms in complex processes. High reliability in detecting and eliminating non-random fluctuations in data, the ability to detect interactions between variables are among the features that have made this data mining superior to other classical methods of modeling and forecasting, including regression. Given the breadth of data mining algorithms and its ability to discover the rules, the question arises as to how this tool can be used to improve the statistical quality control process. Accordingly, in this paper, the independent variables are data related to the statistical quality control processes of Shabrun radiator production, including items such as pump pressure, suction pressure, oil pressure, working day, shift work, and variables related to database rules and tree diagrams. Therefore, the main research question is: How can a quality control approach present with a focus on reducing waste and improving the production process? Also, the sub-questions of the research are:

1. What are the most effective criteria of production processes on the quality of products?
2. How is it possible to control and predict destructive processes in the production system and prevent their occurrence?

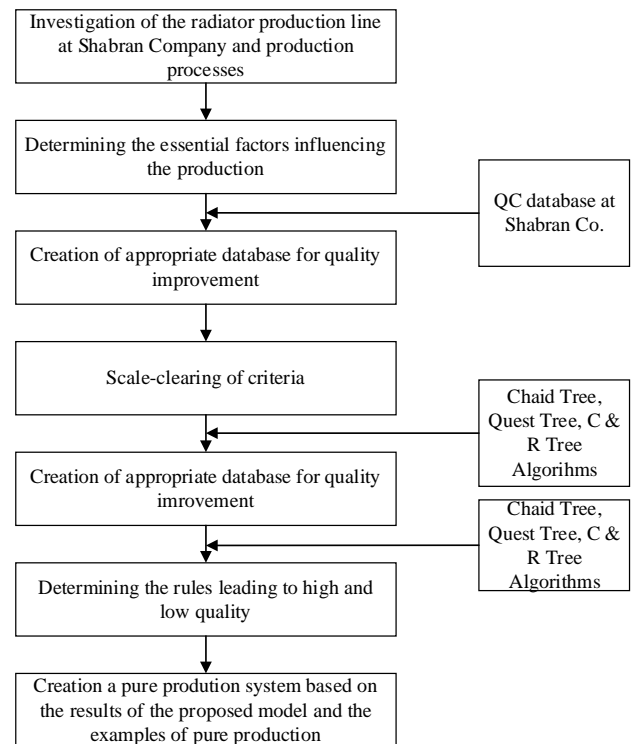
**INFORMATION GATHERING METHOD**

First, using library studies, the necessary information was collected and selected which were examined on the basis of these data-mining models and statistical quality control methods. Field methods are also used to collect necessary information for the different stages of data mining.

**Statistical population and statistical sample**

Accordingly, in this paper, the statistical population is all components manufactured in 2017. Furthermore, the statistical sample of this research are products that have been statistically controlled in the production line. In this research, simple random sampling was used to select the sample pieces from the corresponding statistical population. The data analysis method in this study is shown in Figure 1.

As shown in Figure 1, firstly, the monitoring data of the manufactured components at Shabrun plant was collected (the data recognition phase), then the quality of production line for radiator product was investigated. The most important measurements that affect the quality of the product has been identified. (Data Preparation Phase), and based on this, the QC database compiled and categorized according to the quality of the piece in different categories (modeling phase). Using tree decision algorithms, the most important features mentioned will be determined (modeling phase). To extract the rules affecting the final product quality (radiator), decision tree



**Figure 1.** Information analysis method

algorithms was used and the database rules was extracted. Besides, the algorithms used were compared with each other (evaluation phase), and the most suitable algorithms for the development phase and implementation of the proposed method was recommended (Development phase).

**RESULTS**

The spatial domain of this research is Shabrun, a part of Automobile manufacture suppliers, and the time domain of the research is related to the data of the company's quality assurance unit in 2017. The subjective domain is related to the field of quality assurance and data mining. The records consist of 2500 pieces that have been studied in terms of quality control and are classified into qualitative control in four categories, accordingly. The characteristics include 9 criteria that affect the quality of the product and are determined by quality control experts, which are 1. Diecast machine pressure, 2. Suction pressure of air compressor, 3. Oil pressure of degreasing tub, 4. Working day, 5. Shift work, 6. Cooling time in the dryer, 7. Number of welds, 8. Pistol pump pressure and 9. Welding quality.

**Create a database**

In this section, the database explain based on the following:

Records: Includes 2500 pieces which have been investigated in terms of quality control and have been classified according to the quality of the piece in different categories.

Characteristics: Characteristics include criteria that affect the quality of goods and are determined using the views of the quality control experts, which are shown below:

*a.* Pressure pumps (between 80 and 100 N/m)

The die casting machine has done all the casting operations and the die casting quality depends on the pressure applied by the device at the time of application, and the level of pressure applied at casting time.

*b.* Compressor suction pressure (120 to 180 heads)

In the brizing stage, the compressor has been used, and if the compressor compresses the air with proper pressure, the quality of the parts will be higher.

*c.* Oil pressure (between 30 and 45 torrs)

Before the detached parts are placed in special fixtures and the grid is assembled, the parts will be inserted into the fattening bath. At this stage, the parts are cleaned using a rotary shaft. This will cause to create a stainless cover. Since one of the factors affecting the radiator failure is corrosion at the radiator surface and the stainless coating is effective on the quality of the radiator, this feature was investigated in the database.

*d.* Working day (every 7 days a week)

Occasionally, manpower performance affects the quality of parts. Therefore, the quality of the parts varies in

different working days depending on the level of human labor's accuracy.

*e.* Work shift (two shifts of day and night)

Like the previous feature, working shift may affect the quality of the parts. For example, manpower working at different times has shown different levels of performance.

*f.* Cooling time in the dryer (between 120 and 240 seconds)

Before the welding stage, the parts must be cooled and the cooling time will have a great impact on the quality of the welding stage and will affect the final quality of the piece.

*g.* Number of welding (between 50 and 60)  
To connect parts of the grid and the hose, welding has been used, the number of welds will cause better sealing and prevent the detachment of parts and the damage to the produced radiators. Thus, this feature will affect the level of quality of the radiator.

*h.* Suction pressure of paint pistol (145 to 170 N/m)

In one of the last steps, the radiators are painted using a pistol, this creates another stainless layer on the radiator surface, resulting in a high resistance to corrosion of the radiator. To do this, you need to use high-pressure pistols to fill up all the pores and bring them a higher quality.

*i.* Weld quality (between 70 and 85 hardnesses)

Description of the characteristics of the number of welding has shown the importance of welding. In this part, weld quality is discussed. In the basics of welding inspection, based on criteria such as arc voltage, progressive speed, sedimentation rate and so on, numerical value is specified as weld quality based on the unit of hardness. To simplify, the symbols for these characteristics used are summarized in Table 1.

Class (label): This criterion reflects the quality and usability of the product, which after being approved by the quality control unit, is placed in four categories as follows:

A. High quality: Highly-used, used in a new car.

B. Medium quality: Usable but usually used as spare parts.

**Table 1.** Characteristics and symbols used for naming

Criterion	Symbol
Pressure of die-cast machine	a
Compressor suction pressure	b
oil pressure of fattening bath	c
Working day	d
Working Shift	e
Cooling time in the dryer	f
Number of welding	g
Suction pressure of paint pistol	h
Weld quality	i

C. Low quality: not usable, but with rebuilding it can regain average quality.

D. Very low quality: It is considered as production waste.

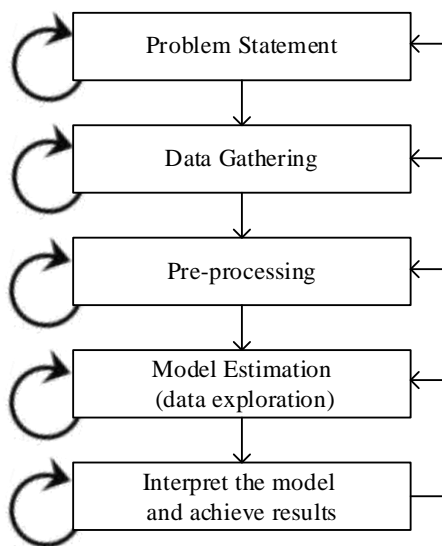
**Data pre-processing**

All the steps individually and the data mining process as a whole are very repetitive. As shown in Figure 2, if the data is not collected and pre-processed properly, or if the problem formula is not meaningful, the resulted model is not valid.

In data preprocessing, data collection is usually done from existing databases. The data preprocessing process usually involves at least two common steps. Detecting (and deleting) unusual data; Abnormalities are in fact inconsistent and unusual data values that are not consistent with most observations. Abnormal data is usually the result of error measurements, coding, and error logging, and sometimes normal and abnormal values. Such unpredictable examples can completely affect the created models later stage [7].

The stages of data preprocessing should not be considered completely independent of other stages of data mining. Each time the data mining process is repeated, all activities together can define a serious and improved data set for subsequent iterations. In general, a suitable preprocessing method provides an optimal representation of a data mining technique through the integration of prior knowledge in the form of specific application scaling and encryption.

Pre-processing is divided into two subgroups in terms of functionality and its corresponding techniques: data preparation and data size reduction. Accordingly, the statistical sample of this research are radiators that have been statistically controlled in the production line; which goes through its construction step.



**Figure 2.** Data mining process

At the end of the work of the radiators, after all the necessary steps, the product quality control unit is checked and in case of any problem or non-compliance with the plan, the radiator is referred to the previous section for possible correction.

At this stage, given the data, it is preprocessed. For this purpose, by examining each of the characteristics which values are numerical, we place them in different categories, which are visible in Table A1 (see Appendix). Thus, using Table A1, quantitative data is converted to qualitative data. It should be noted that the tolerances presented have been gathered according to the devices in the table, which is collected using the device catalogs and the help of technical expert in each section.

**Descriptive analysis of data**

In this step, we describe the data using any of the tolerances in Table A1 as well as in other criteria, the results of which are shown in Tables 2, 3 and 4. In this

**Table 2.** Description of quantitative data

No. of tolerance	No. of low tolerance	No. of medium tolerance	No. of high tolerance	Criterion
1462	545	315	177	Pressure of die-cast machine
1477	604	297	153	Compressor suction pressure
1239	714	324	223	oil pressure of fattening bath
869	784	449	398	Cooling time in the dryer
1071	617	505	308	Number of welding
1137	786	295	283	Suction pressure of paint pistol
937	956	375	233	Weld quality

**Table 3.** Description of the working shift of the sampled parts

Working nights	Normal	Working shift
589	1911	Number

**Table 4.** Description of the production day of the sampled parts

Closed	Normal	Working day
235	2265	Number

section, the implementation of one of the stages of CRISP-DM, which was identified as phase, has been addressed. Qualitative data was used in this research, and in the previous section, it is referred to the way of qualifying data.

**Determining database reliability**

In data mining, before the implementation of the algorithms, the accuracy of the database are usually analyzed, which is shown in Table 5 of the accuracy of the tree algorithm. In this section, a ten-point crediting algorithm was used. In explaining how this algorithm works, the data is divided into ten, and ten percent of the data is used as the test, and the remaining 90 percent is used as training, and this is repeated for ten times, in each step, the error rate is calculated and the learning would be improved. Finally, the number of predicted true and false values which is determined by the algorithm is shown in Table 5. In this research, CLEMENTINE 12 software was used for ten-crediting and implementing decision tree algorithms.

As shown in Table 5 the algorithm has a precision of more than 94 percent. Therefore, it can be concluded that the data are suitable for data mining. The basis for obtaining the accuracy is that the database is divided into 10 equal parts and each time a part is considered as a test, and each time, by comparing the actual value with the predicted error percentage, and finally the percentage error is summed up and the total accuracy is achieved.

**Determining decision trees of database**

In this study, we use three decision tree algorithms for implementing and determining rules that describe the results in each of these algorithms. In this section, implementation of the modeling phase, which is one of the stages of CRISP-DM, has been dealt with and the most important features were extracted.

**C & R Tree algorithm**

*The most important features of the C & R Tree algorithm*  
 With the implementation of C & R Tree algorithm, the most important characteristics are shown in Figure 3. It should be noted that with the implementation of C & R tree decision tree algorithm, the most important characteristics are extracted from the database, as shown in Figure 2, in the Clementine 12 software. It should be noted that Clementine 12 software uses the effect of each

attribute on the Classifier to determine the most important attribute. Thus, first, the attribute to be determined is its importance, and then it checks that to what extent Classifier has changed. Accordingly, "the characteristic by whose removal, the greatest change is created in the Classifier, is the most important characteristic.

*C & R decision tree algorithm*

With the implementation of C & R Tree algorithm, the data decision tree is formed, as shown in Figures 4 to 8.

As shown in Figure 5, if the cooling time is carried out in a dryer with a high or medium quality, at the end of the process, the quality of the piece will reach from 79.2 to 87.6, and if this process is carried out at a low and very low level, the quality of the piece after the implementation of this process will reach from 79.2 to 61.2.

As presented in Figure A1 (see Appendix), if the cooling time is carried out in a high or medium quality dryer, at the end of the process, the quality of the piece will be from 79.2 to 87.6, then if the piece is produced in

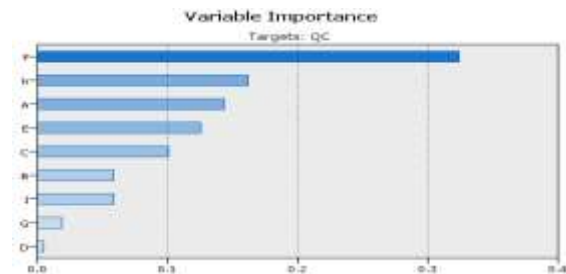


Figure 3. Ratings of characteristics using the C & R Tree algorithm

Node 0		
Category	%	n
hi	79.240	1981
low	6.480	162
medium	11.520	288
very low	2.760	69
Total	100.000	2500

Figure 4. A tree with a depth of 1 C & R Tree algorithm

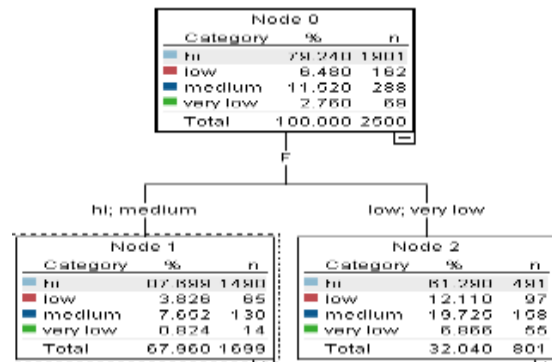


Figure 5. A tree with a depth of 2 C & R Tree algorithms

Table 5. Precision of decision tree algorithm

Type of prediction	Percent	Number
Correct	94.72%	2386
Wrong	5.28%	132
Total data	100%	2500

the working day shift, the quality to 93%, and if the piece is produced in the working night shift, the quality of the piece will reach 71.2; as well as if the cooling time in the dryer is carried out in a low and very low quality, and then Suction pressure of paint pistol process the piece with high or medium quality, the quality of the piece will increase from 61.2% to 70%, as well, if Suction pressure of paint pistol process the piece in a low or very low quality, the quality of the piece after the execution of this process will reach from 61.2 to 28.7 (Figure A1).

With regard to Figure A2 (see Appendix) and interpretation of processes, if the cooling time is high or moderate quality in the dryer, the working day shift and high oil pressure of fattening bath, the quality of the piece will increase from 79.2 to 90%. So this process is a useful process, and it will increase the final quality. On the other hand, if the cooling time is low or very low quality in a dryer, the pressure of paint pistol is low or very low and the number of welds is very low, then the quality of the piece will fall from 79% to 16%. This is a destructive process and has led to a loss of quality and should be avoided. The remaining processes are also shown clearly in the figure and will be interpreted accordingly.

**Quest Tree algorithm**

*The most important features of the Quest Tree algorithm*  
By implementing the Quest Tree algorithm, the most important characteristics are determined, which are shown in Figure 7.

*The decision tree of the Quest Tree algorithm (based on the third question of the research)*

By implementing the Quest Tree algorithm, the decision tree is formed which is shown in Figures 8 to 12.

As displayed in Figure 9, if oil pressure of fattening bath is carried out with a high or medium quality, at the end of the process, the quality of the piece will reach from 79.25 to 85.5, and if this process is received at a low and very low quality, the quality of the piece after the implementation of this process will reach from 2.79 to 59%.

As exposed in Figure 10, if oil pressure of fattening bath is carried out with a high or medium quality, at the end of the process, the quality of the piece will reach from 79.25

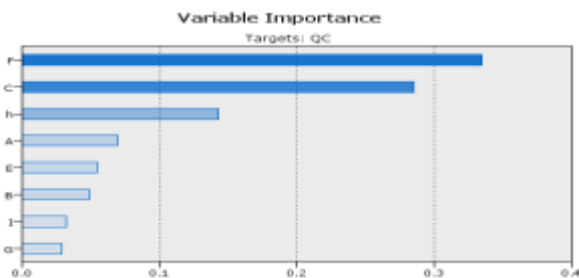


Figure 7. Ranking characteristics using Quest Tree algorithm

to 85.5, further if cooling time in the dryer with a high or medium quality will reach to 92.6%, and if the cooling time in the dryer is carried out with a high or medium quality, the quality of the piece will reach 96.4; as well, if the oil pressure of fattening bath is carried out with a low and very low quality, and then Suction pressure of paint pistol process the piece with a high, medium or low quality, quality of the piece will reach from 59 to 63.1%. In addition, if suction pressures of paint pistol process the piece with a very low quality, the quality of the piece after the implementation of this process will reach from 59 to 33.7 (Figure A3, see Appendix).

According to Figure A4 (see Appendix) and interpretation of processes, if oil pressure of fattening bath, the cooling time in the dryer, and the pressure of die-cast machine are carried out with a high or medium quality, the quality of the piece will increase from 79.2% to 96.2%, so this process is a beneficial process and will increase the final quality. On the other hand, if oil pressure of fattening bath, and suction pressure of paint pistol, and compressor suction pressure are carried out with a low and very low quality, then the quality of the piece will reach from 79.2 to 0%. The process is destructive and has led to a loss of quality and should be avoided. The rest of the processes have been shown clearly in the figure and will be interpreted accordingly.

**Chaid Tree algorithm**

*The most important features of the Chaid Tree algorithm*  
By implementing the Chaid Tree algorithm, the most important characteristics are shown in Figure 12.

Node 0		
Category	%	n
hi	79.240	1981
low	6.480	162
medium	11.520	288
very low	2.760	69
Total	100.000	2500

Figure 8. A tree with a depth of 1 Quest Tree algorithm

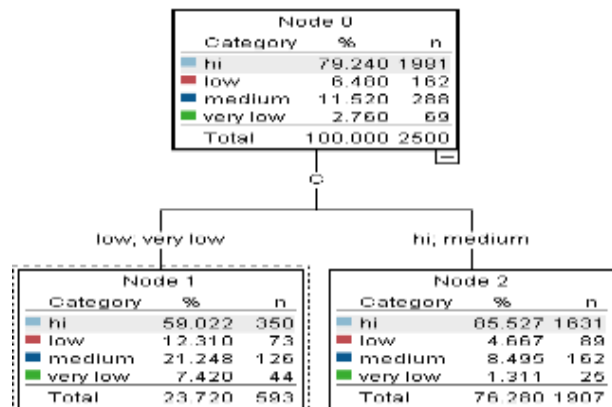


Figure 9. A tree with a depth of 2 Quest Tree algorithms

*The decision tree of the Chaid Tree algorithm (based on the third question of the research)*

By implementing the Chaid Tree algorithm, the data decision tree is formed, as shown in Figures 13 and A5 (see Appendix).

As delineated in Figure 14, if the oil pressure of fattening bath is carried out with a high, medium, low and very low quality, at the end of the process, the quality of the piece will reach from 79.29 to 89.9, 87.7, 62.7% and 53% respectively.

As shown in Figure A5, if the quality of the oil pressure of fattening bath is with a high or medium and the weld quality is with a high or medium quality, the final quality will reach from 79.2 to 95.3. So, this process is a beneficial process and will increase the quality of the final quality. On the other hand, if the oil pressure of fattening bath is with a very low quality and cooling time in the dryer is low or very low, the quality will reach from 79.2 to 30.7. This process is destructive and has led to a loss of quality and should be avoided. The rest of the processes are also shown clearly in the figure and will be interpreted accordingly.

**Practical implications of computational results**

In this section, the last step of CRISP-DM is examined, which will investigate the methods and algorithms used.

Using Clementine 12 software and analysis tool, the accuracy for each algorithm was calculated, that we calculated the highest precision for C & R Tree algorithm and the accuracy of the algorithm was 82.2. Moreover, Chaid Tree and Tree Quest algorithms have a precision of 81.72 and 80.88, respectively.

One of the consequences of this research as practical implications is that according to industry experts, if the final quality of the part is less than 30%, that part is considered waste, so all processes that produce products with less than 30% quality lead to waste. Therefore, according to the results of this research, the destructive process is identified and removed from the production line. After implementing the proposed method, the final quality of production parts has increased by 20% and the volume of waste has been reduced by 15%. The quality control unit has announced the quality of construction before the implementation of the proposed method as

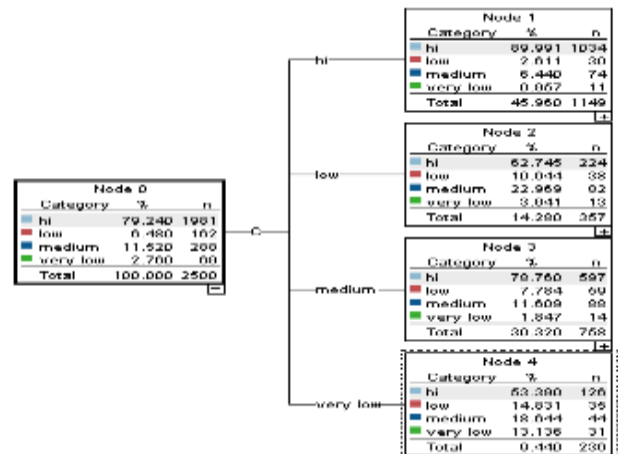


Figure 13. A tree with a depth of 2 Chaid Tree algorithms

70%, but after the implementation of the method presented in this article, the quality control unit has acknowledged 90% of the quality.

**CONCLUSION**

Quality assurance means inspection and review of project processes and products to ensure they are in compliance with standards and procedures. The goal of quality assurance activities is to cooperate with the project team in order to achieve quality. For this purpose, the quality groups submit the inspection and review results to the project managers and the employer in the form of periodic and case reports. Quality assurance activities are considered as one of the project management activities so that, by carrying out these activities, the Quality Assurance team cooperates in planning and compliance with the project standards, as an independent group, inspects the procedure for carrying out the activities and products, and ensures their accuracy. Quality assurance activities are carried out according to a specific methodology. In this methodology, methods, processes, scope of activities, interaction protocols between project stakeholders in relation to quality assurance, and specific methods and tools for carrying out quality assurance activities are presented. Quality assurance methodology in the company studied in order to meet the goals set. In this methodology, a set of activities is required to carry out specific quality assurance activities and tools for doing these activities. It also describes how to use these tools to carry out quality assurance activities.

This methodology includes all activities of the quality groups. The activities of the quality groups are described in the following sections separately. The quality assurance methodology of the company is based on the experience gained from various quality assurance and monitoring projects carried out at this company. Supervisory and quality assurance projects carried out by

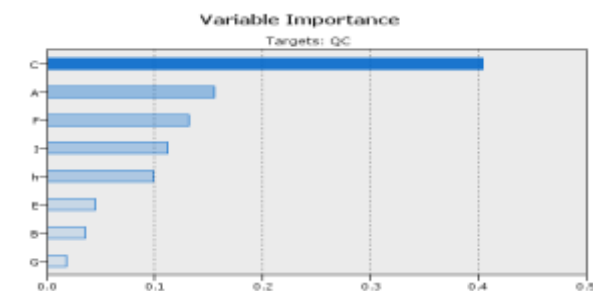


Figure 12. Ranking the characteristics using the Chaid Tree algorithm



the corporation. There is no company that can succeed without controlling the quality of its products in the business. In the quality assurance section, there are a variety of processes, but the most important process that deals with the customer, the product, and the entire production process is product quality analysis. In this process, defective components are identified and should be taken to preventative measures. In the discussion of statistical quality control, there are various diagrams and processes, but all of these processes and charts have a descriptive state and none of them has predictive value. Accordingly, in this dissertation, we looked at this issue with a new perspective and tried to predict product defects. According to the data presented in the first chapter of the data mining tool, due to its predictive nature, it has a good ability in various fields. After implementation of the proposed method, the final quality of the manufactured components has increased by 20%, and the volume of waste has decreased by 15%. It should be noted that the quality control unit has announced the quality of construction quality before implementing the proposed method of 70%, but after implementing the method presented in this paper, the quality control unit has acknowledged 90% of the quality. In line with improving and completing the results of this research, other researchers are recommended to take steps in the following ways:

- Using the Rough Set Theory to determine the most important characteristics and comparing selected characteristics by this method with the method presented in this paper.
- Using Feature Ranking algorithms to determine the weight of each characteristic and create mathematical models to optimize the final quality of the components in terms of cost, applicability and production time using new and improved methods of monitoring and quality control for more accurate recording of component quality.

## REFERENCES

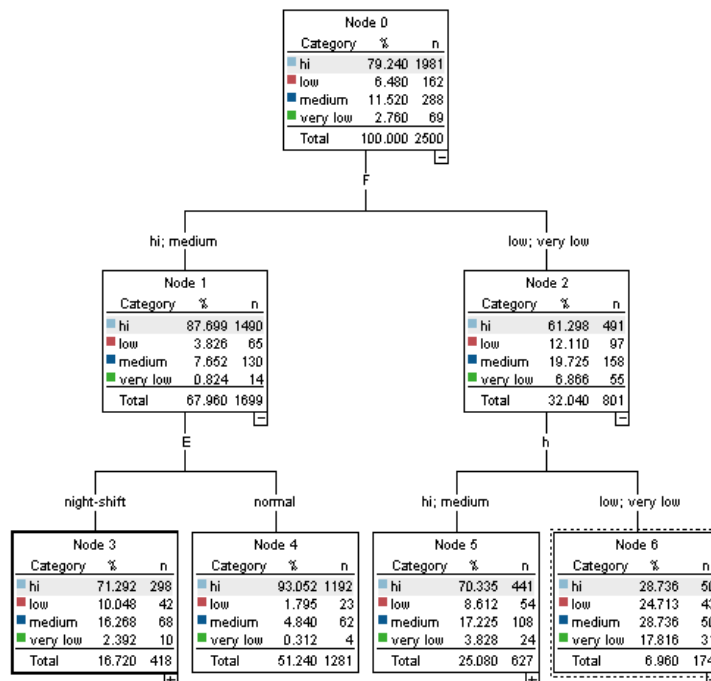
1. Lindsey, T. C., 2011. Sustainable Principles: Common Values for Achieving Sustainability. *Journal of Cleaner Production*, 19(5), pp.561-565. <https://doi.org/10.1016/j.jclepro.2010.10.014>
2. Homesi, S. A. H., Nikomaram, H. and Seyyed Mohammadi, N., 2006. Study on the Application of Statistical Quality Control Process in Sima-E-Sork Co. *Agricultural Sciences*, 12(2), pp.463-478 [In Persian].
3. Lotfi, S., Ghasemzadeh, M., Mohsenzadeh, M. and Mirzarezaee, M., 2021. The Construction of Scalable Decision Tree Based on Fast Splitting and J-Max Pre-Pruning on Large Datasets. *International Journal of Engineering, Transactions B: Applications*, 34(8), <https://dx.doi.org/10.5829/ije.2021.34.08b01>
4. Hossain, M. M., Prybutok, V. R., Abdullah, A. and Talukder, M., 2010. The Development and Research Tradition of Statistical Quality Control. *International Journal of Productivity and Quality Management*, 5(1), pp.21-37. <https://doi.org/10.1504/IJPMQ.2010.029505>
5. Cheng, G. and Li, L., 2020. Joint Optimization of Production, Quality Control and Maintenance for Serial-Parallel Multistage Production Systems. *Reliability Engineering and System Safety*, 204pp.107146. <https://doi.org/10.1016/j.res.2020.107146>
6. Chen, S., Li, X., Liu, R. and Zeng, S., 2019. Extension Data Mining Method for Improving Product Manufacturing Quality. *Procedia Computer Science*, 162, pp.146-155. <https://doi.org/10.1016/j.procs.2019.11.270>
7. Koulinas, G., Paraschos, P. and Koulouriotis, D., 2020. A Decision Trees-Based Knowledge Mining Approach for Controlling a Complex Production System. *Procedia Manufacturing*, 51, pp.1439-1445. <https://doi.org/10.1016/j.promfg.2020.10.200>
8. Sarkar, M. and Sarkar, B., 2020. How Does an Industry Reduce Waste and Consumed Energy within a Multi-Stage Smart Sustainable Biofuel Production System? *Journal of Cleaner Production*, 262, pp.121200. <https://doi.org/10.1016/j.jclepro.2020.121200>
9. Schmitt, T., Wolf, C., Lennerfors, T. T. and Okwir, S., 2021. Beyond "Leaner" Production: A Multi-Level Approach for Achieving Circularity in a Lean Manufacturing Context. *Journal of Cleaner Production*, 318, pp.128531. <https://doi.org/10.1016/j.jclepro.2021.128531>
10. Ashrafi, H., Raissi Ardali, G. and Zolfaghari, H., 2017. Integration of Reliability and Six Sigma as a Quality Control Tool to Improve the Situation of Manufacturing Products. *Journal of Science and Engineering Elites*, (5), pp.101-109 [In Persian].
11. Mousavi, S., 2012. A Multi-Criteria Decision-Making Approach with Interval Numbers for Evaluating Project Risk Responses. *International Journal of Engineering, Transactions B: Applications*, 25(2), pp.121-130. <https://doi.org/10.5829/idosi.ije.2012.25.02b.05>
12. Hamidi, H. and Qaribpour, F., 2017. An Efficient Predictive Model for Probability of Genetic Diseases Transmission Using a Combined Model. *International Journal of Engineering, Transactions B: Applications*, 30(8), pp.1152-1159. <https://doi.org/10.5829/ije.2017.30.08b.06>
13. Saniga, E. M., 1993. Decision Support and Statistical Quality Control. *International Journal of Quality and Reliability Management*, 10(2), pp.9-17. <https://doi.org/10.1108/02656719310027948>
14. Bagheri, P., Qousi, R. and Haeri, A., A Composite Data Modeling Model in an Atm Manufacturing Company, in Second International Conference on Knowledge Based Research in Computer Engineering and Information Technology. 2017: Majlisi University, Tehran, Iran [In Persian].
15. Smith, H. D., Megahed, F. M., Jones- Farmer, L. A. and Clark, M., 2014. Using Visual Data Mining to Enhance the Simple Tools in Statistical Process Control: A Case Study. *Quality and Reliability Engineering International*, 30(6), pp.905-917. <https://doi.org/10.1002/qre.1706>
16. Azyazov, V. N., Torbin, A. P., Mebel, A. M., Bresler, S. M. and Heaven, M. C., 2017. Product Channels of the Reactions of Rb (62p) with H<sub>2</sub>, CH<sub>4</sub> and C<sub>2</sub>H<sub>6</sub>. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 196, pp.46-52. <https://doi.org/10.1016/j.jqsrt.2017.03.032>
17. Lu, H. and Ma, X., 2020. Hybrid Decision Tree-Based Machine Learning Models for Short-Term Water Quality Prediction. *Chemosphere*, 249, pp.126169. <https://doi.org/10.1016/j.chemosphere.2020.126169>
18. Vivancos, J.-L., Buswell, R. A., Cosar-Jorda, P. and Aparicio-Fernandez, C., 2020. The Application of Quality Control Charts for Identifying Changes in Time-Series Home Energy Data. *Energy & Buildings*, 215, pp.109841. <https://doi.org/10.1016/j.enbuild.2020.109841>

19. Hasheminezhad, A., Hashemi, S. and Tabatabaie, R., 2018. Evaluation of Operative Factors on Conversion Efficiency of Biodiesel Production from Waste Cooking Oil. *Iranian Journal of Energy and Environment*, 9(2), pp.100-104. <https://doi.org/10.5829/ijee.2018.09.02.04>
20. Istoto, E. H., Widayat, W. and Saptadi, S., 2019. Production of Fuels from High Density Polyethylene and Low Density Polyethylene Plastic Wastes Via Pyrolysis Methods (Research Note). *Iranian Journal of Energy and Environment*, 10(3), pp.185-189. <https://doi.org/10.5829/ijee.2019.10.03.04>
21. Bayu, A., Amibo, T. and Akuma, D., 2020. Conversion of Degradable Municipal Solid Waste into Fuel Briquette: Case of Jimma City Municipal Solid Waste. *Iranian Journal of Energy and Environment*, 11(2), pp.122-129. <https://doi.org/10.5829/ijee.2020.11.02.05>
22. Ataee, A., Kazemitabar, J. and Najafi, M., 2020. A Framework for Dry Waste Detection Based on a Deep Convolutional Neural Network. *Iranian Journal of Energy and Environment*, 11(4), pp.248-252. <https://doi.org/10.5829/ijee.2020.11.04.01>

**APPENDIX**

**Table A1.** Data qualifying

Weld quality	Suction pressure of paint pistol	Number of welding	Cooling time in the dryer	Oil pressure of fattening bath	Compressor suction pressure	Pressure of die-cast machine	Criterion
More than 84 and less than 71	More than 168 and less than 147	More than 59 and less than 51	More than 230 and less than 135	More than 44 and less than 31	More than 177 and less than 122	100 and 80	High Tolerance
More than 83 and less than 73	More than 166 and less than 149	More than 58 and less than 53	More than 220 and less than 150	More than 43 and less than 33	More than 172 and less than 127	99, 98, 81 and 82	Medium Tolerance
More than 79 and less than 74	More than 161 and less than 153	More than 56 and less than 54	More than 200 and less than 165	More than 40 and less than 35	More than 165 and less than 135	96, 97, 83, 84 and 85	Low Tolerance
Residual Domain	Residual Domain	Residual Domain	Residual Domain	Residual Domain	Residual Domain	Residual Domain	No Tolerance
85-70	170-145	60-50	240-120	45-30	180-120	100-80	Domain



**Figure A1.** A tree with a depth of 3 C and R Tree algorithms

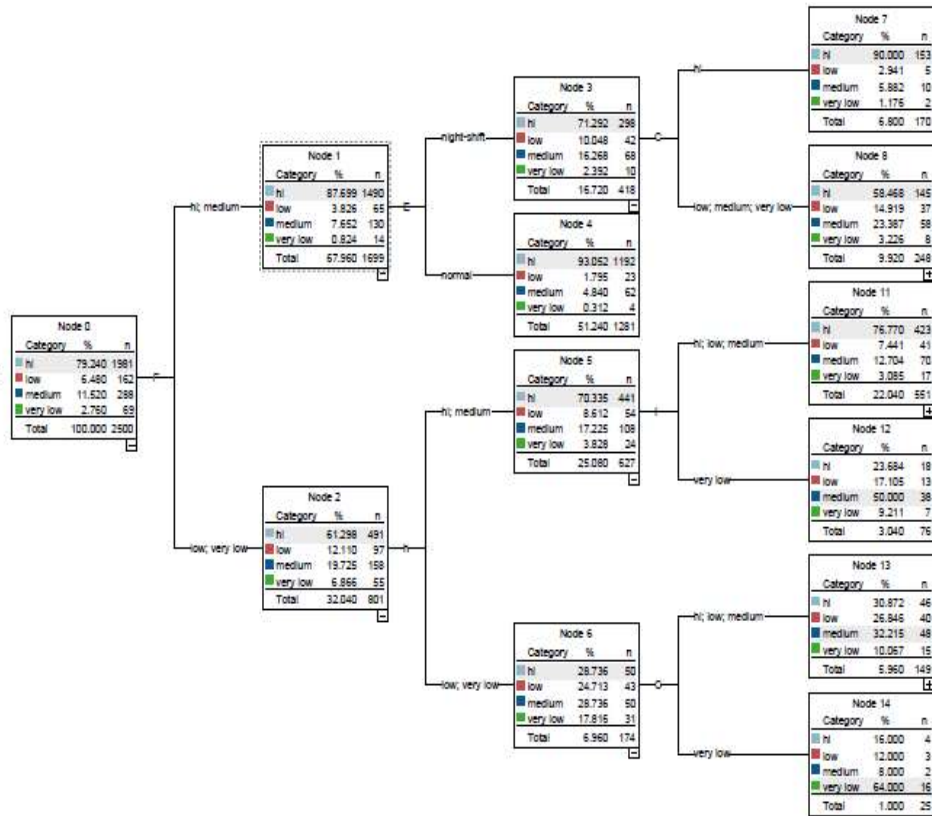


Figure A2. A tree with a depth of 4 C and R Tree algorithms

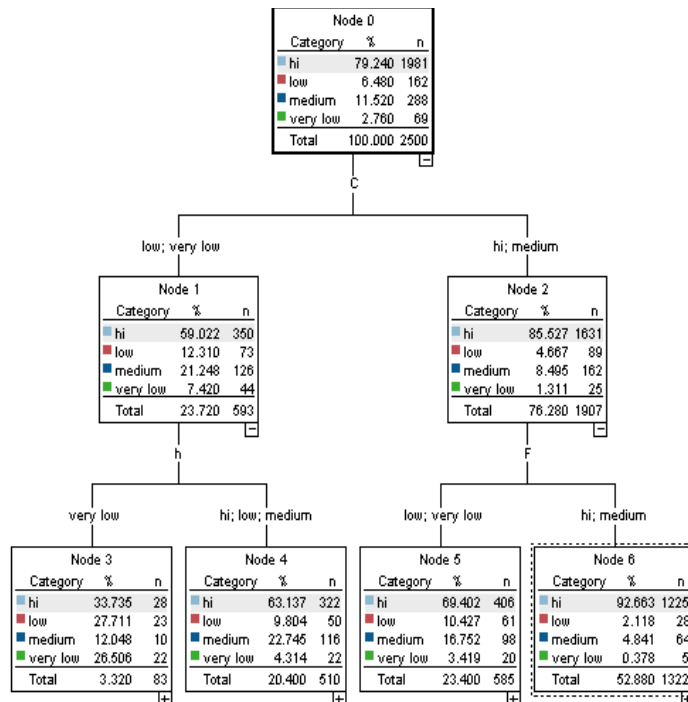


Figure A3. A tree with a depth of 4 C and R Tree algorithms



Figure A4. A tree with a depth of 4 Quest Tree algorithms

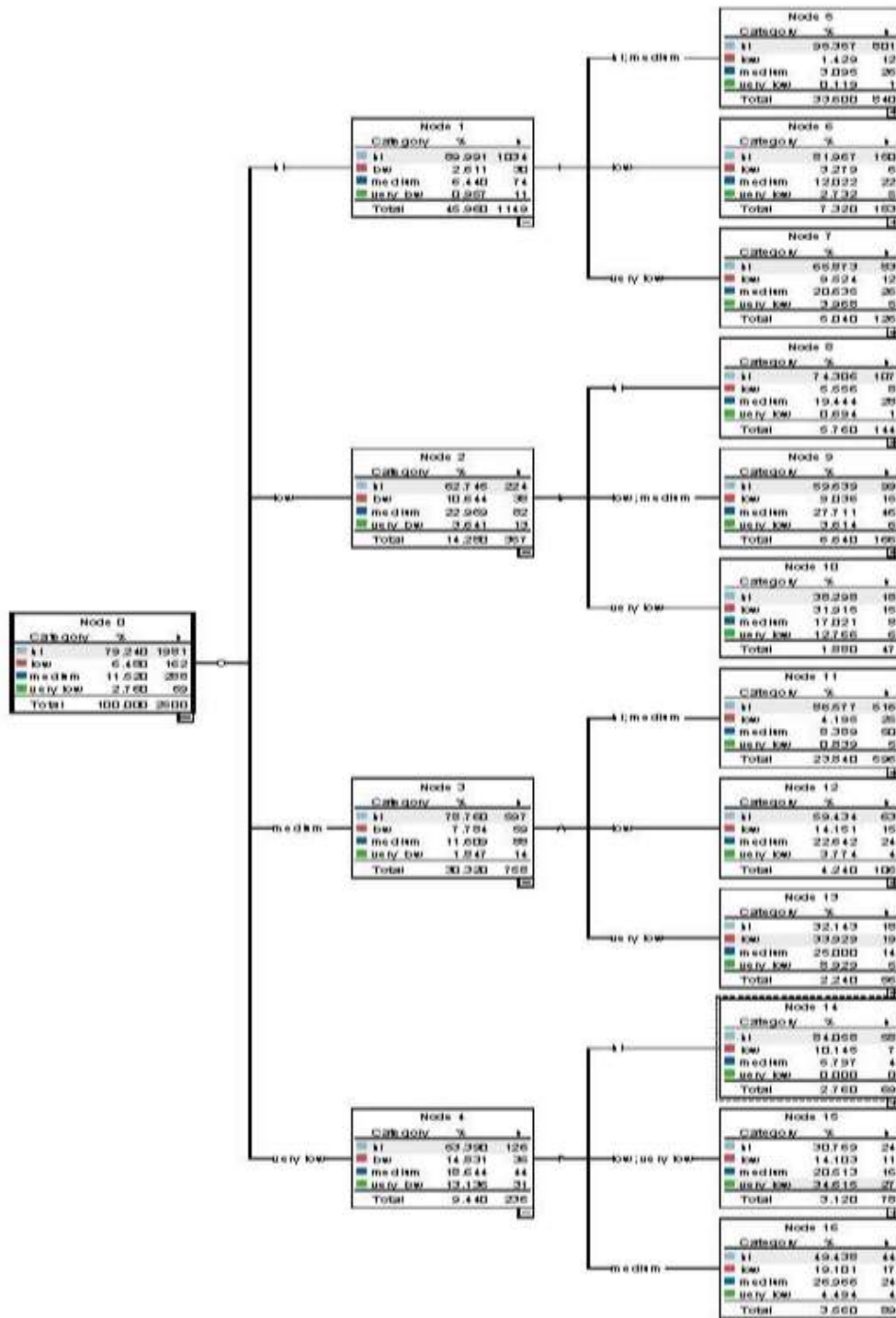


Figure A5. A tree with a depth of 2 Chaid Tree algorithms

#### COPYRIGHTS

©2021 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



---

#### Persian Abstract

---

##### چکیده

امروزه در اکثر صنایع از ابزار کنترل کیفیت آماری جهت بهبود کیفیت و عدم تولید محصولات معیوب و کاهش ضایعات استفاده می‌نمایند ولی حجم بالای داده‌ها، کمک یک ابزار قوی را جهت کنترل فرایندها می‌طلبد. یکی از اهداف مطالعه حاضر پیش‌بینی محصولات معیوب و جلوگیری از تولید آن‌ها با استفاده از ابزار داده کاوی با توجه به توان بالا در تحلیل داده‌ها و خاصیت پیش‌بینی‌کنندگی آن است که در صنعت از این رویکرد کمتر بهره گرفته شده است. در این پژوهش، جامعه آماری تمامی قطعات تولیدی شرکت شابرین در سال ۲۰۱۷ می‌باشد. نمونه آماری تعداد ۲۴۰۰ قطعه رادپاتوری است که به طور تصادفی از خط تولید انتخاب شده‌اند. در فازهای عملیاتی داده‌کاوی از سه الگوریتم درخت تصمیم استفاده شد که عبارت‌اند از: Chaid Tree و Quest Tree و C&R Tree که با استفاده از این الگوریتم‌ها مهم‌ترین معیارهای موثر بر کنترل کیفیت و قوانین منجر به میزان کیفیت قطعات تعیین شد. نتایج مقایسه‌ای نشان داد که علی‌رغم اعتبار هر سه الگوریتم، الگوریتم C&R Tree بالاترین دقت را دارا بوده است. تبعیت از قوانین حاصل از پیاده‌سازی این الگوریتم‌ها موجب کشف و جلوگیری از تولید ضایعات شده که این مهم موجب افزایش کارایی و جلوگیری از اتلاف زمان و هزینه در این واحد تولیدی گردید. نتایج نشان داد کیفیت نهایی قطعات تولیدی ۲۰ درصد افزایش یافته و از حجم ضایعات نیز به میزان ۱۵ درصد کاسته شده است. لازم به ذکر است که واحد کنترل کیفیت، میزان کیفیت ساخت را قبل از پیاده‌سازی روش پیشنهادی ۷۰ درصد اعلام نمود اما پس از پیاده‌سازی روش ارائه شده در این مقاله، واحد مذکور به میزان کیفیت ۹۰ درصد اذعان نموده است.

---