

Prediction of MRR & Surface Roughness in Wire EDM Machining using Decision Tree and Naive Bayes Algorithm

Yogesh L

Postgraduate Student
 Department of Mechanical Engineering
 Dayananda Sagar College of Engineering
 Bangalore, India
 yoginaidu944@gmail.com

M Arunadevi

Department of Mechanical Engineering
 Dayananda Sagar College of Engineering
 Bangalore, India
 arunadevi.dsce@gmail.com

C P S Prakash

Department of Mechanical Engineering
 Dayananda Sagar College of Engineering
 Bangalore, India
 drcpsprakash@gmail.com

Abstract— Manufacturing industries facing problem in optimal selection of process parameters in machining process. Finding optimum process parameters for achieving maximum Material Removal Rate and minimum Surface Roughness is a challenging task and it requires lot of time and energy for experimentation trails or experience. It wastes lot of resources and money, sometimes ends up with negative results. To overcome the above issue, this paper presents an algorithm for prediction of Surface Roughness and Material Removal rate using Decision Tree Algorithm and Naive Bayes Algorithm without experimentation. Lot of resources and time can be saved using these machine learning algorithms. In this paper, Material removal rate and Surface roughness of EDM machining of Aluminum composites is predicted using Decision tree algorithm and Naive Bayes algorithm. Then the model can be used to predict the Material Removal Rate and Surface finish of any combination process parameters before machining process.

Keywords—Decision Tree Algorithm, EDM, Optimization, Naive Bayes Algorithm, Surface Roughness, Machine Learning

I. INTRODUCTION

Now days, the conventional machines are replaced by the nonconventional machines because of complexity involved in conventional machines. Wire EDM is one of the nonconventional machines, which plays important role in cutting of Aluminum composites with high accuracy [1]. In most of manufacturing, aeronautical and automotive industries, the main structural components are made using Aluminum with wire EDM process [2]. In [3], optimization of Wear rate of electrode and MRR is obtained with input variables such as discharge current, duty cycle, pulse duration and voltage. ANN is used for modeling and Non-dominated Sorting Genetic Algorithm -II is used to solve the multi-objective optimization. In [4, 5, 7] the MRR and SR the objective functions and model is generated using Design of experiments, Different techniques such as Gray Rational Analysis[4], Particle swarm algorithm[5] are used to solve the multi objectives. In [6], Flushing pressure, servo voltage, pulse on time, wire feed rate, pulse off time and wire tension are the input parameters for measure the outputs such as MRR, kerf width and SR in wire electric discharge machining process. Gradient descent method is used for optimizing the parameters for improvement.

In recent research papers, ANN [3, 5, 8, 10, and 11] is used for developing the model and accurate predictions are made to analyze the each parameter on the responses. In [9, 11], minimizing the surface roughness and [12] maximize

the MRR is considered as the single objective in wire EDM machining process. In this paper, decision tree algorithm is proposed for the prediction of Surface roughness and Material Removal Rate. Wire-EDM process is having non linear characteristics which is highly complex in nature[13]. Due to more number of input parameters, the machine learning approach of data science is proposed in this study. The properly conducted experimental data was used for training the model. The model can be tested with extra data. It is found that the model built using data mining provides results with desired accuracy. In [14], the machine learning and the decision rules are derived to find the significance of process parameters of electric current in wire EDM machining.

II. PROBLEM DESCRIPTION

In manufacturing industries selection of process parameters plays very important role in achieving Maximum Material removal rate and minimum Surface roughness. In olden days, selection of parameters is done by the experience of operator. In case the operator is new, lot experimentation trails are required for the proper selection of machining parameters. Lot of energy and time are involved in the experimentation for all possible combination of process parameters. It wastes lot of resources.

To avoid the above problem, levels of parameters are introduced and design of experiments is done using Taguchi method. Again taguchi method will not explore all possible combination.

To solve the above problem in this study output data is converted from continuous to categorical and solved using supervised classifier techniques. Among various classifiers, Decision tree classifier and Naive bayes classifier are selected because of the following reasons

- Needs less data for training
- Not sensitive to missing or irrelevant data
- Normalization and scaling of data is not required
- Fast accurate classifiers and can be used for real time predictions
- Very intuitive and easy to implement
- Less computational complexity

A. Objectives of this study

- To conduct the experimentation using Taguchi method.
- Train the Decision tree algorithm and Naïve base algorithm using experimental data.
- To predict response values of all possible combination of all input parameters.
- To reduce the experimental cost using Decision tree algorithm and Naïve Bayes algorithm.
- To reduce the time of experimentation using above machine learning techniques.

III. METHODOLOGY

The machine learning algorithm involves the following three steps are

- i. Data Processing
- ii. Modeling
- iii. Validation

A. Data Processing

Processing of data is the main step in machine learning, because the accuracy of machine learning model depends on data processing.

Data processing involves two steps:

- Data selection
- Feature engineering.

1) *Data Selection*: From the paper[15], 18 datasets were obtained from wire EDM machining process which includes different combination of input parameters such as Voltage, Pulse-ON, Pulse-OFF, Current and Bed speed and response variables such as Material removal rate and Surface roughness is considered which is shown in table 1. Material removal rate and surface roughness values are converted from continuous to categorical.

2) *Feature Engineering*: Feature engineering is the process of selecting the suitable input variables for output prediction. In this study combination of process parameters are used as input parameters for predicting Material removal rate and Surface roughness.

B. Modeling of Decision Tree Algorithm

Decision Tree algorithm is a type of supervised learning algorithms which can be used to solve both Regression and Classification Problem. The aim of creating Decision Tree is to train a model that can be used to predict the class or value of the target variable by learning simple decision rules inferred from the training data. In Decision Trees, predictions will start from the root of the node and then compare the values of the root attribute with the other attribute. On the basis of comparison, the branch moves from root to corresponding to that value and jump to the next node. Steps involved in decision tree algorithm.

- It starts with original data set as S and start with root node.
- During every iteration, it searches the next attribute from unselected attributes of set S and calculates Entropy $E(P)$ and Information gain $G(I)$ of the attributes using the followings formulas

$$E(P) = -\sum_{i=1}^n p_i * \log_2 p_i$$

$$G(I) = E(P) - \sum_{i=1}^n \frac{n_i}{n} E(i)$$

- Attribute which has smallest Entropy or Largest Information gain is selected.
- Then the set S is splitted by the selected attribute and subset of the data is created.
- The algorithm continues to recur on each subset, and consider only the attributes which never selected before.

C. Modeling of Naïve Bayes Algorithm

Naive Bayes algorithm is an algorithm which works based on Bayes theorem and used for solving classification problems. It is a type of supervised learning algorithm and a probabilistic classifier, where prediction will be done based on the basis of the probability of occurrence. Bayes theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability. Conditional probability can be calculated using the following formula.

$$P(A|B) = P(B|A) * P(A) / P(B)$$

$P(A|B)$ - Posterior probability

$P(B|A)$ - Likelihood probability:

$P(A)$ - Prior Probability

$P(B)$ - Marginal Probability

Steps involved in naive Bayes classification

1. Conversion of data set in to frequency tables.
2. Generation of Likelihood table by finding the probabilities of given features.
3. Calculation of posterior probability using Bayes theorem.

D. Validation of model

Assessment on performance and the accuracy of trained model can be achieved by the validation of a model. In every machine learning technique, the actual data will be divided in to two set as a training set and test set. The training set of data will be used for training of model and the test set will be used for validation of model.

TABLE I. WIRE EDM PROCESS PARAMETERS AND ITS RESPONSE VARIABLES

Data No.	Voltage	Pulse ON	Pulse OFF	Current	Bed speed	MRR	MRR	SR	SR
No.	volt	μs	μs	amps	$\mu m/s$	mm^3/min	class	(μm)	class
D1	75	40	9	2	50	5.641	C1	1.784	C3
D2	75	40	12	4	150	16.500	C3	1.528	C2
D3	75	40	15	6	250	17.984	C3	1.572	C2
D4	75	30	9	2	150	13.895	C2	1.737	C3
D5	75	30	12	4	250	17.671	C3	1.739	C3
D6	75	30	15	6	50	5.906	C1	1.387	C1
D7	75	20	9	4	50	5.901	C1	1.645	C2
D8	75	20	12	6	150	17.886	C3	1.567	C2
D9	75	20	15	2	250	8.800	C2	1.749	C3
D10	100	40	9	6	250	30.137	C3	1.405	C1
D11	100	40	12	2	50	5.911	C1	1.576	C2
D12	100	40	15	4	150	17.483	C3	1.595	C2
D13	100	30	9	4	250	28.884	C3	1.759	C3
D14	100	30	12	6	50	5.880	C1	1.617	C2
D15	100	30	15	2	150	11.681	C2	1.768	C3
D16	100	20	9	6	150	17.862	C3	1.458	C1
D17	100	20	12	2	250	14.270	C2	1.729	C3
D18	100	20	15	4	50	5.877	C1	1.501	C2

IV. RESULTS AND DISCUSSIONS

In this study all the input parameters are categorical values and output also converted in to categorical values by some conditions which is shown in table 2.

TABLE II. CONDITIONS FOR CONTINUOUS TO CATEGORY

Variable	Range	Class
Material	<8	C1
Removal rate mm^3/min	8 to 16	C2
	>16	C3
Surface Roughness (μm)	<1.5	C1
	1.5-1.7	C2
	>1.7	C3

In this work, the combination of voltage, pulse-ON, pulse-OFF, current and bed speed is considered as the input and the targets are Material removal rate and Surface Roughness. Decision tree algorithm is used for training the network.

A. Prediction of MRR using Decision Tree

To create the decision tree of Material removal rate, the entropy and Information gain is calculated using MRR class as an output in an iterative method which is given in Table 3. In iteration 1, by comparing all 5 attributes entropy and information gain values, Bed speed is having higher information gain and low entropy value. So bed speed is selected as root node for the Decision tree which is shown in Figure 1. In iteration 2, total table is divided in to 3 tables and solved as a three individual table. In iteration 2, bed speed 50 table has all the class as same (C1), it stops the tree there itself.

TABLE III. ENTROPY AND INFORMATION GAIN VALUES OF MRR

Iteration	Data set	Variable	E	GI	C
1	(D1-D18)	Voltage	1.530493	0	-
		Pulse-ON	1.36274	0.167753	
		Pulse-OFF	1.501086	0.029407	
		Current	0.918296	0.612197	
		Bed speed	0.612197	0.918296	
	Bed Speed 50 (D1,D6,D7,	Voltage	0	0	C1
		Pulse-ON	0	0	

2	D11,D14,D18)	Pulse-OFF	0	0	-
		Current	0	0	
	Bed Speed 150 (D2,D4,D8, D12,D15,D16)	Voltage	0.918296	0	
		Pulse-ON	0.389975	0.528321	
		Pulse-OFF	0.834419	0.083877	
		Current	0.389975	0.528321	
	Bed Speed 250 (D3,D5,D9, D10,D13,D17)	Voltage	0.918296	0	
		Pulse-ON	0.389975	0.528321	
		Pulse-OFF	0.834419	0.083877	
3	Bed Speed 150	Current 2			C2
	Bed Speed 150	Current 4			C3
	Bed Speed 150	Current 6			C3
	Bed Speed 250	Pulse-ON 20			C2
	Bed Speed 250	Pulse-ON 30			C3
	Bed Speed 250	Pulse-ON 40			C3

The other two tables in second iteration has the different class values, the problem continues to find the next attribute for the tree. In bed speed 150 and 250, by comparing 4 attributes, current and pulse-ON are having higher information gain and low entropy value, anyone attribute can be selected as next attribute(current or Pulse-ON). Current selected for bed speed 150 and Pulse-ON is selected for bed speed 250. In iteration 3, tree is ended because all the attributes has only one class value. The tree structure for Material removal rate is created and shown below.

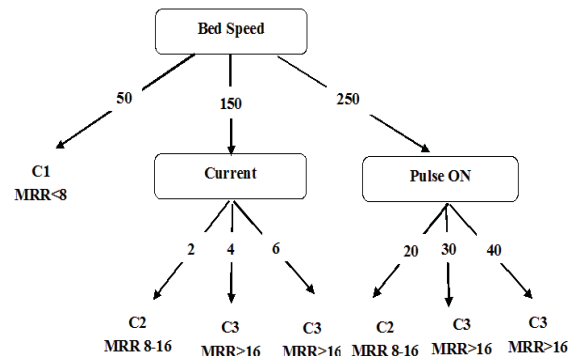


Fig. 1. MRR Decision Tree

In iteration 1, by comparing all 5 attributes entropy and information gain values, Bed speed is having higher information gain and low entropy value. So bed speed is selected as root node for the Decision tree which is shown in Figure 1.

In iteration 2, total table is divided in to 3 tables and solved as a three individual table. In second layer of tree, Pulse- ON, Current and Pulse-OFF are selected as root node for the further iterations.

Iteration 3, Voltage and Pulse OFF are selected as next node under pulse on 30 and 40. Pulse-OFF is selected under the current 6. Current and Pulse-ON is selected under Pulse-OFF 9 and 12. The tree structure for Surface roughness is created and shown below.

The results obtained by the MRR decision tree proved that bed speed is the most influencing parameter on the Material removal rate. Higher values of bed speed, Current and Pulse-ON lead to a better Material removal rate. In this three class problem, Bed speed , Current and Pulse-ON are the deciding Parameters for better Material Removal rate.

B. Prediction of SR using Decision Tree

The decision tree for surface roughness is also created in the same procedure and shown in Figure 2. The entropy and information gain of all the attributes are calculated using surface roughness class as an output and listed in the table 4.

The results obtained by the SR decision tree algorithm proved that bed speed is the most influencing parameter on the Surface roughness. In this tree, all five attributes are influencing the surface roughness. Comparatively voltage is have less influence on the decision tree. The selection in range of output values for different classes plays very important role in structure of decision tree. The above trees created can be used to select the process parameters to obtain the required Material removal rate and Surface roughness. Thus it shows, the decision tree approach brings out the influencing parameters much more effectively.

TABLE IV. ENTROPY AND INFORMATION GAIN VALUES OF SR

Iteration	Data set	Variable	E	GI	C
1	(D1-D18)	Voltage	1.4613	0.019362	-
		Pulse-ON	1.1908	0.289872	
		Pulse-OFF	1.1488	0.33181	
		Current	0.85610	0.624576	
		Bed speed	0.44970	1.030974	
2	Bed Speed 50 (D1,D6,D7, D11,D14,D18)	Voltage	0.792481	0.4591479	-
		Pulse-ON	0.666666	0.5849632	
		Pulse-OFF	0.666666	0.584963	
		Current	0.834419	0.417211	
	Bed Speed 150 (D2,D4,D8, D12,D15,D16)	Voltage	1.251629	0.207519	-
		Pulse-ON	0.333333	1.125815	
		Pulse-OFF	0.666666	0.792482	
		Current	0.333333	1.125815	
	Bed Speed 250 (D3,D5,D9, D10,D13,D17)	Voltage	0.918296	0.333333	-
		Pulse-ON	1	0.251629	
		Pulse-OFF	0.666666	0.584963	
		Current	1	0.251629	
3	Bed Speed 50	Pulse-ON 20			C2
	Bed Speed 50	Pulse-ON 30	Voltage 75		C1
	Bed Speed 50	Pulse-ON 30	Voltage 100		C2
	Bed Speed 50	Pulse-ON 40	Pulse-OFF 9		C3
	Bed Speed 50	Pulse-ON 40	Pulse-OFF 12		C2
	Bed Speed 150	Current 2			C3
	Bed Speed 150	Current 4			C2

Bed Speed 150	Current 6	Pulse-OFF 9	C1
Bed Speed 150	Current 6	Pulse-OFF 12	C2
Bed Speed 250	Pulse-OFF 9	Current 4	C3
Bed Speed 250	Pulse-OFF 9	Current 6	C1
Bed Speed 250	Pulse-OFF 12		C3
Bed Speed 250	Pulse-OFF 15	Pulse-ON 30	C3
Bed Speed 250	Pulse-OFF 15	Pulse-ON 40	C1

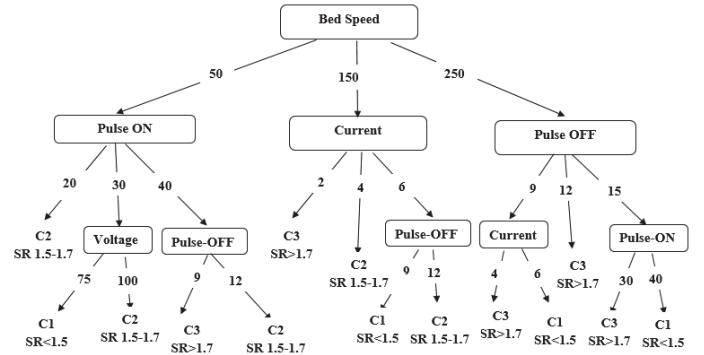


Fig. 2. Surface Roughness Decision Tree

C. Prediction of MRR using Naïve Bayes Algorithm

The likelihood table of Material Removal rate is calculated using the frequency of attributes in the experimental data set which is shown in Table 4.

TABLE V. LIKELIHOOD TABLE FOR MATERIAL REMOVAL RATE

Material Removal Rate				
Attribute	Value	Probability (C1)	Probability (C2)	Probability (C3)
Voltage	75	3/6	2/4	4/8
	100	3/6	2/4	4/8
Pulse-ON	20	2/6	2/4	2/8
	30	2/6	2/4	2/8
	40	2/6	0/4	4/8
Pulse-OFF	9	2/6	1/4	3/8
	12	2/6	1/4	3/8
	15	2/6	2/4	2/8
Current	2	2/6	4/4	0/8
	4	2/6	0/4	4/8
	6	2/6	0/4	4/8
Bed speed	50	6/6	0/4	0/8
	150	0/6	2/4	4/8
	250	0/6	2/4	4/8
overall Probability		6/18	4/18	8/18

From the likelihood table, the Material removal rate can be predicted using Bayes theorem equation.

For example, If Voltage = 75; Pulse-ON = 40;

Pulse-OFF = 9; Current = 6; Bed speed = 150;

MRR Prediction using Table 5.

$$\text{Prob}(C1) = (3/6 * 2/6 * 2/6 * 0.1/6.3) * (6/18) = 0.0000979$$

$$\text{Prob}(C2) = (2/4 * 0.1/4.3 * 1/4 * 0.1/4.3 * 2/4) * (4/18) = 0.0000075$$

$$\text{Prob}(C3) = (4/8 * 4/8 * 3/8 * 4/8 * 4/8) * (8/18) = 0.01$$

By comparing all the three probability Prob(C3) is having higher value, so Predicted MRR is C3

D. Prediction of SR using Naïve Bayes Algorithm

The likelihood table of Surface roughness is calculated using the frequency of attributes in the experimental data set which is shown in Table 4.

TABLE VI. LIKELIHOOD TABLE FOR SURFACE ROUGHNESS

Surface Roughness				
Attribute	Value	Probability (C1)	Probability (C2)	Probability(C3)
Voltage	75	1/3	4/8	4/7
	100	2/3	4/8	3/7
Pulse-ON	20	1/3	3/8	2/7
	30	1/3	1/8	4/7
	40	1/3	4/8	1/7
Pulse-OFF	9	2/3	1/8	3/7
	12	0/3	4/8	2/7
	15	1/3	3/8	2/7
Current	2	0/3	1/8	5/7
	4	0/3	4/8	2/7
	6	3/3	3/8	0/7
Bed speed	50	1/3	4/8	1/7
	150	1/3	3/8	2/7
	250	1/3	1/8	4/7
Overall Probability		3/18	8/18	7/18

From the likelihood table, the Material removal rate can be predicted using Bayes theorem equation.

For the same example, If Voltage = 75; Pulse-ON = 40;

Pulse-OFF = 9; Current = 6; Bed speed = 150;

Surface Roughness Prediction using Table 6.

$\text{Prob}(C1) = (1/3 * 1/3 * 2/3 * 3/3 * 1/3) * (3/18) = 0.0041$

$\text{Prob}(C2) = (4/8 * 4/8 * 1/8 * 3/8 * 3/8) * (4/18) = 0.0019$

$\text{Prob}(C3) = (4/7 * 1/7 * 3/7 * 0.1/7.3 * 2/7) * (7/18) = 0.01$

By comparing all the three probability Prob(C1) is having higher value, so Predicted MRR is C1.

The same input parameters Predictions made using both Decision tree algorithm and Naïve bayes algorithm for two examples are compared and results are tabulated in Table 7.

TABLE VII. COMPARISON BETWEEN DECISION TREE AND NAÏVE BAYES

Sl. No	Attribute Value		MRR Prediction		SR Prediction	
			Decision Tree	Naïve Bayes	Decision Tree	Naïve Bayes
1	Voltage	75	C3	C3	C1	C1
	Pulse-ON	40				
	Pulse-OFF	9				
	Current	6				
	Bed speed	150				
2	Voltage	100	C1	C1	C2	C2
	Pulse-ON	30				
	Pulse-OFF	12				
	Current	4				
	Bed speed	50				

In [18], the same dataset is solved as continuous output using Linear regression and accuracy of those model is compared with Decision tree classifier and Naïve bayes classifier in this study which is shown in table Table 8.

TABLE VIII. COMPARISON WITH LINEAR REGRESSION MODEL

Sl.No	Model	Accuracy (%)	
		MRR	Surface Roughness
1	Linear Regression	86.96	62.81
2	Decision Tree Algorithm	100	100
3	Naïve Bayes Algorithm	100	100

Compared to regressor, classifier will give better results in this study, because all input values are selected based on levels using Taguchi method which is categorical and only outputs are the continuous values. In this study outputs are also converted in to categorical to solve using different classifier algorithms. It is proved that decision tree classifier and naïve bayes classifier gave better results compared to linear regression.

V. CONCLUSION

In this paper, the machine learning approach of data science is proposed for the decision making process in wire EDM machining. Decision tree algorithm and Naive base algorithm are selected for this study to predict the response variables. The properly conducted experimental data was used for training the model. Experimental values of process parameters in wire-EDM machining of Aluminum composites and its response variable values are used to build a decision tree model and to create Likelihood table. Then the model can be used to predict the Material Removal Rate and Surface finish of new combination process parameters without experimentation. Then the accuracy of decision tree model and naïve bayes model are compared with linear regression model from the paper [15]. It is found that the model built using supervised classifier algorithms such as Decision Tree and Naïve bayes provides better results with desired accuracy compared to Linear regression. Future work of this paper is to build the models using other classifiers like SVM, ANN, etc... and the same methodology can be applied for different machining processes.

REFERENCES

- [1] Divya Marelli, Sateesh N, Ram Subbiah, "Review on multi objective optimization of wire cut EDM process parameters using grey relational analysis", Materials Today: Proceedings (2020)
- [2] Ashutosh Kumar Singh, Deepak Singhal, Ramanuj Kumar, "Machining of aluminum 7075 alloy using EDM process: An ANN validation", Materials Today: Proceedings (2020)
- [3] Shiba Narayan Saha, Sagar Kumar Murmub, Narayan Chandra Nayak, "Multi-objective Optimization of EDM Process with Performance Appraisal of GA based Algorithms in Neural Network Environment", Materials Today: Proceedings 18 (2019) 3982–3997
- [4] DhruvBhatta, AshishGoyalb, "Multi-objective optimization of machining parameters in wire EDM for AISI-304", Materials Today: Proceedings 18 (2019) 4227–4242
- [5] K. Naveen Babu, R. Karthikeyan, A. Punitha An integrated ANN – PSO approach to optimize the material removal rate and surface roughness of wire cut EDM on INCONEL 750, Materials Today: Proceedings (2019)
- [6] Sidharth Kumar Shukla, a Amrita Priyadarshini, "Application of Machine Learning Techniques for Multi Objective Optimization of Response Variables in Wire Cut Electro Discharge Machining Operation" Materials Science Forum, ISSN: 1662-9752, Vol. 969, pp 800-806 (2019)
- [7] M. Subrahmanyam, T. Nancharai, "Optimization of process parameters in wire-cut EDM of Inconel 625 using Taguchi's approach", Materials Today: Proceedings (2019)
- [8] Wang J, Sánchez J, Iturrioz, J.A, Ayesta, I, "Artificial Intelligence for advanced non-conventional machining processes", Procedia Manufacturing 41 (2019) 453–459

- [9] M. Fakkir Mohamed, K. Lenin, "Optimization of Wire EDM process parameters using Taguchi technique", *Materials Today: Proceedings* (2019)
- [10] S. Ganapathy, P. Balasubramanian, B. Vasanth, S. Thulasiraman, "Comparative investigation of Artificial Neural Network (ANN) and Response Surface Methodology (RSM) expectation in EDM parameters" *Materials Today: Proceedings* (2019)
- [11] Nishant K. Singh, Yashvir Singh, Sanjeev Kumar, Abhishek Sharma, "Predictive analysis of surface roughness in EDM using semi-empirical, ANN and ANFIS techniques: A comparative study", *Materials Today: Proceedings* (2019)
- [12] Indranil Ghosh, M. K. Sanyal, R. K. Jana, Pranab K. Dan, Machine Learning for Predictive Modeling in Management of Operations of EDM Equipment Product, Second International conference on Research in Computational Intelligence and Communication Networks (ICRCICN-2016)
- [13] Shajan Kuriakose, Kamal Mohan, M.S. Shunmugam "Data mining applied to wire-EDM process", *Journal of Materials Processing Technology* 142 (2003) 182–189
- [14] Josko Valentincic . Mihael Junkar , "Detection of the eroding surface in the EDM process based on the current signal in the gap" *International Journal of Advanced Manufacturing Technology*, (2007) 33: 698–705.
- [15] Kishore. G. C, Arunadevi M, C. P. S Prakash, "Parametric Optimization of Wire Electrical Discharge Machining by Taguchi Technique on Composite Material", *International Journal of Engineering Research & Technology (IJERT)* Vol. 4 Issue 09, September-2015.