Prediction of turning parameters of EN AW-1350 aluminum alloy using machine learning and regression models

Fakhreddine Khrouf
Mechanics Laboratory of Constantine (LMC), Faculty of Technology Sciences, Mentouri Constantine 1 University, Chaab Ersas Campus, 25000 Constantine, Algeria.

Article Info

Article history:

Received 07/12/2023
Revised 10/06/2024
Accepted 11/06/2024

Keyword:

Aluminum alloy; Machine learning; Turning; Surface roughness; Regression methods.

ABSTRACT

The paper proposes the use of machine learning (ML) models to predict the cutting properties when turning EN-AW-1350 aluminum alloy under dry cutting conditions. cutting speeds (m/min), feed rates (mm/rev) and depths of cut (mm) are the main input control parameters selected for the present investigation. The established models can predict the surface roughness values (Ra) during the aluminum hard turning operation, which can guide the direction of experiments and eliminate the need for time-consuming traditional experimental procedures. The prediction performance of six regression methods (Decision Tree (DT), linear regression (LR), eXtreme Gradient Boosting (XGBoost), AdaBoost, Support Vector Machine (SVM) and Random Forest (RF)) were evaluated with a test set/training set ratio of 8 to 2. Among the six regression methods, XGBoost had the best prediction effect on the surface roughness of EN AW-1350 aluminum alloy with a low mean root mean square error (RMSE) and coefficient of determination (R^2) close to 1.

Corresponding Author: Email: fakhreddine.khrouf@umc.edu.dz

1. INTRODUCTION

Artificial Intelligence (AI) is playing a major role in the subsequent fourth industrial revolution. It is a process of training a computer model using complex and large data sets. The model learns from this data in a tutoring process to build its ability to make decisions or predict outcomes when presented with new data. AI techniques are widely used by the researcher to solve a whole range of hitherto intractable problems. Artificial intelligence (AI) and machine learning (ML) are really ubiquitous and exciting technologies, and we really view as another really important tool for mechanical researchers. We build the physical devices you that interact with, whether or not it's a car, it's a medical device that your surgeon's using. Moreover, what AI is going to let mechanical engineers do is to take it to the next level to develop a better device, to better understand that physical phenomenon. AI utilizes mathematical and statistical principles fundamentally to delineate a multitude of physical phenomena, such as fluid mechanics [1,2], stress and strain analysis [3,4], control problems [5,6].
Khrouf et al. [7] investigated the effects of cutting parameters on surface roughness and material removal rate when turning EN AW-1350 aluminum alloy under dry cutting conditions. The study compares the accuracy and reliability of the Artificial Neural Network (ANN) and Response Surface Methodology (RSM) in predicting and detecting the non-linearity of surface roughness and material removal rate mathematical models.

Gabsi et al. [8] examined and predicts surface roughness during the AA7075 milling process using machine learning algorithms. The study conducted experiments with a CNC milling machine using a workpiece made of an aluminum alloy (AA7075). Various cutting parameters were studied in relation to roughness average (Ra) values. The study compares thirteen machine learning algorithms, including basic and ensemble models, and identifies the voting regression model as the best performing model. The results show that the voting regression model has high performance metrics and can be employed by manufacturing companies to predict surface roughness and enhance manufacturing efficiency.

Adizue Ugonna et al. [9] purposed to develop a predictive surface roughness model and optimize process parameters for ultra-precision hard-turning finishing operation. Machine learning models such as support vector machine (SVM), Gaussian process relation (GPR), adaptive neuro-fuzzy inference system (ANFIS), and artificial neural network (ANN) were developed for surface roughness prediction. The models showed excellent predictive accuracy, with ANFIS and ANN having the lowest MAPE values. The response surface method (RSM) analysis revealed that the feed parameter had the most significant influence on minimizing surface roughness. The optimal cutting conditions were determined to be cutting speed = 100 m/min, feed = 0.025

©UBMA - 2024

mm/rev, and depth of cut = 0.09 mm. These findings can aid decision-making in the precision machining industry.

Pimenov et al. [10] used of new sensors and artificial intelligence (AI) methods for tool condition monitoring in machining operations, such as turning, milling, drilling, and grinding. It also highlights the advantages, disadvantages, and prospects of using various AI methods for tool wear monitoring, as well as the challenges in implementing machine learning techniques in the manufacturing industry.

The paper is structured as follows: in section two, the experimental procedure is detailed. The machine learning methods is presented in section three, the results are discussed is given in section four. Last of all, the conclusion shows the effect of cutting parameters on turning EN AW-1350 aluminum alloy using machine learning approaches.

2. EXPERIMENTAL PROCEDURE

Turning operations are carried out on a CNC SPINNER-TC65 lathe that develops a maximum spindle speed of 4500 rpm and a spindle power of 16.5/562-4500-kW under dry conditions. The measurements of arithmetic surface roughness for each cutting condition were obtained from a PCE-RT 1200 Roughness Tester with a cut-off length of 0.8 mm and sampling length of 4 mm. The measurements were repeated at five equally spaced locations around the circumference of the workpiece and the result is an average of these values for a given machining pass.

The present work is an attempt to examine the effect of cutting parameters on surface roughness of EN AW-1350 aluminum alloy using machine learning approaches. The experimental setup is shown in Figure 1.

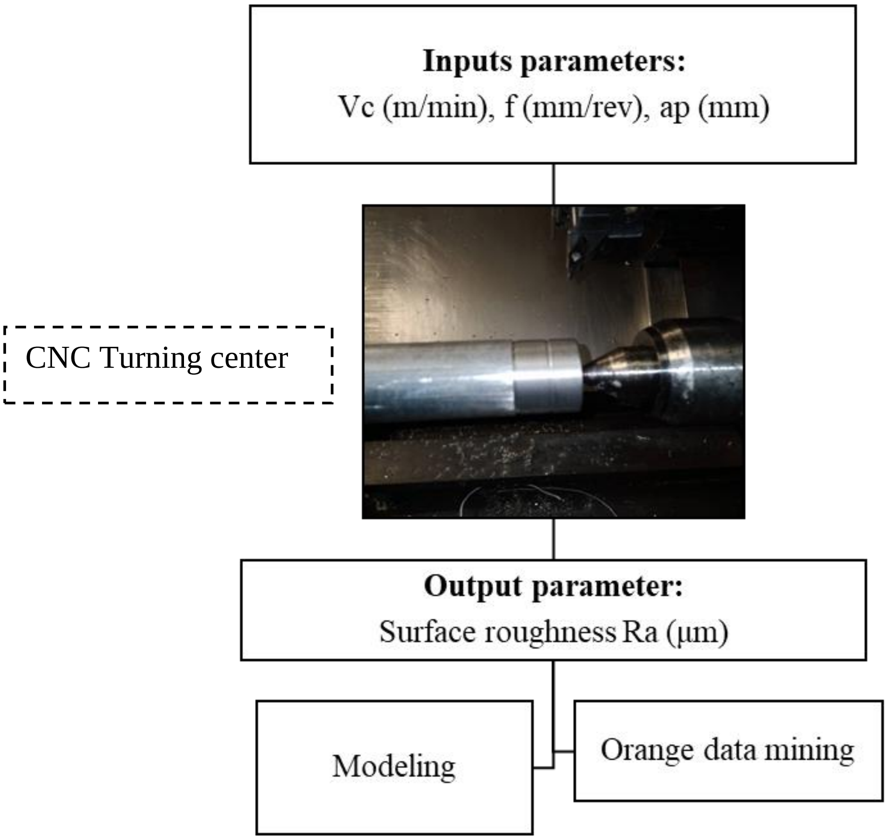


Figure.1. Experimental methodology.

2.1. Workpiece material and cutting tool

Due to its properties, lightweight aluminum alloys have been used in various applications. The 8xxx series is used in air conditioning, while the 7xxx series is commonly used in aircraft wing manufacturing and the aerospace and marine industries due to its good weldability. All remaining series (from 6xxx to 1xxx) are used in a wide range of applications, from automotive and aerospace to construction and packaging. Chemical composition of work material is as follows: 0.19 Fe, 0.11 Si, 0.023 Cu, 0.038 Zn, 0.010 Cr, 0.018 Ti, 0.021 Mg, 0.09 Mn. Specimens of 500 mm in length and 50 mm in diameter are adopted for the turning process using a

©UBMA - 2024

coated carbide tool. Cutting tool type is SVJBL 2020 K16 and its geometry is specified by a 93° principal cutting edge angle, a 5° cutting edge angle and a nose radius of 0.4 mm.

2.2. Data set

The present work examines the effect of cutting parameters on mechanical properties of EN AW-1350 aluminum alloy using ML approaches. Outlined experimental details and setups are schematized in Figure 1. As shown in Table 1, the tests are performed based on three variables, each with three appropriately chosen levels: *(i)* cutting speeds (600; 700 and 800 m/min), *(ii)* feed rates (0.8; 0.12 and 0.16 mm/rev) and *(iii)* depths of cut (0.15; 0.30 and 0.45 mm). The levels of the cutting parameters are selected from the intervals recommended by the cutting tools manufacturer.

Table 1. Experimental data for EN AW-1350 aluminum alloy.

Run	*Vc* (m/min)	*f* (mm/rev)	*ap* (mm)	*Ra (μm)*
1	600	0.08	0.15	1.35
2	600	0.08	0.30	1.35
3	600	0.08	0.45	1.33
4	600	0.12	0.15	1.61
5	600	0.12	0.30	1.60
6	600	0.12	0.45	1.62
7	600	0.16	0.15	1.75
8	600	0.16	0.30	1.73
9	600	0.16	0.45	1.71
10	700	0.08	0.15	1.21
11	700	0.08	0.30	1.22
12	700	0.08	0.45	1.23
13	700	0.12	0.15	1.39
14	700	0.12	0.30	1.41
15	700	0.12	0.45	1.42
16	700	0.16	0.15	1.52
17	700	0.16	0.30	1.53
18	700	0.16	0.45	1.54
19	800	0.08	0.15	1.17
20	800	0.08	0.30	1.18
21	800	0.08	0.45	1.10
22	800	0.12	0.15	1.34
23	800	0.12	0.30	1.35
24	800	0.12	0.45	1.37
25	800	0.16	0.15	1.43
26	800	0.16	0.30	1.45
27	800	0.16	0.45	1.46

3. MACHINE LEARNING METHODS

Supervised learning models require a lot of upfront work from data scientists. Input data sets need to be labeled, and output parameters and expected results need to be specified. During the learning process, accuracy also needs to be adjusted. Various machine learning methods, namely Decision tree, Linear regression, eXtreme Gradient Boosting, Random Forest, AdaBoost, Support Vector Machine were applied to this work to predict the surface roughness of EN AW-1350 aluminum alloy. These machine learning methods are briefly described below.

©UBMA - 2024

3.1. Decision tree (DT)

A decision tree algorithm is indeed a type of machine learning algorithm. It's used in supervised learning for both classification and regression tasks [11]. The algorithm constructs a model of decisions based on actual values of attributes in the data. Decision rules are formed by these decisions, which then predict the target or outcome variable. The 'tree' is formed by these decision rules. It's a popular algorithm due to its interpretability, as the rules formed can be visualized and understood.

3.2 Linear regression (LR)

Linear regression is a type of supervised machine learning algorithm [12]. It's used to model the relationship between a dependent variable and one or more independent variables. The relationship is modeled using a linear predictor function, and the unknown parameters of this function are estimated from the data.
In simple linear regression, there is one independent variable and one dependent variable. If there are more than one independent variables, the process is called multiple linear regression. Linear regression is widely used in practical applications because models which depend linearly on their unknown parameters are easier to fit than models which are non-linearly related to their parameters.
It's important to note that linear regression makes certain assumptions about the data, such as homogeneity of variance, independence of observations, and normality. If these assumptions are not met, other types of regression or non-parametric methods may be more appropriate.

3.3 Extreme gradient boosting (XGBoost)

XGBoost, which stands for Extreme Gradient Boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library [13]. It's an implementation of gradient-boosted decision trees designed for speed and performance. XGBoost is particularly popular due to its efficiency and effectiveness in predictive modeling tasks. It has been used extensively in many winning solutions in machine learning competitions. Some of its key features include regularization to prevent overfitting, parallel processing for faster execution, handling of missing values, tree pruning, and built-in cross-validation. However, it's worth noting that while XGBoost often achieves higher accuracy, it sacrifices the interpretability that comes with single decision trees.

3.4 Random forest (RF)

Random Forest is a supervised learning algorithm [14]. It's an ensemble learning method that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems. It's particularly effective when the individual trees are uncorrelated with each other.

3.5 Adaptive boosting (adaboost)

AdaBoost, short for Adaptive Boosting, is a statistical classification meta-algorithm formulated by Yoav Freund and Robert Schapire in 19951. It's a type of ensemble machine learning algorithm that can be used in conjunction with many other types of learning algorithms to improve their performance [15]. The AdaBoost algorithm works by combining multiple weak learners to create a strong learner. A weak learner is a classifier that performs poorly, but better than random guessing. AdaBoost works by fitting a sequence of weak learners on repeatedly modified versions of the data. The predictions from all of them are then combined through a weighted majority vote (or sum) to produce the final prediction.

3.6 Support vector machine (SVM)

Support Vector Machine (SVM) is a powerful and flexible supervised learning algorithm that is used for both classification and regression [16].
The main objective of the SVM algorithm is to find the optimal hyperplane in an N-dimensional space that can separate the data points in different classes in the feature space. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane.
SVMs can be used for a variety of tasks, such as text classification, image classification, spam detection, handwriting identification, gene expression analysis, face detection, and anomaly detection. SVMs are adaptable and efficient in a variety of applications because they can manage high-dimensional data and nonlinear relationships.

©UBMA - 2024

3.7. Regression models and accuracy measures

To evaluate a regression model, one can calculate the distance between predicted and actual values. The linear regression model which represents the relationship of the data analysed is shown in eq. (1):

$$y = b_0 + b_1 x \tag{1}$$

In this study, six different ML methods were used to investigate the cutting conditions of Aluminum Alloy and to predict surface roughness (Ra) value by Orange data mining software. The data set was randomly divided into a training set and a test set in the ratio of 8:2. The goodness of fit and performance of the six ML models were assessed and compared using three statistical indicators, namely coefficient of determination (R^2), mean squared error (MSE), and root mean square error (RMSE). In the dataset, each observed value Xi is associated with a modeled value Yi, also known as a predicted value. The general definition of RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f\ (\mathrm{x}_i) - y_i)^2} \tag{2}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f\ (\mathrm{x}_i) - y_i)^2 \tag{3}$$

The lower the RMSE and MSE value, the better the performance of the model. And the general definition of the coefficient of determination is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(xi - y_i)^2}{\sum_{i=1}^{n}(xi - \bar{t}\,)^2} \tag{4}$$

where $\bar{x}$ is the mean of the n observed data, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} xi$. If the value of R^2 is closer to 1, it follows that the regression curve can fit the data perfectly.

4. REGRESSION RESULTS AND DISCUSSIONS

Orange Data Mining is an open-source software package for data visualization, machine learning, and data mining. It features a visual programming front-end for explorative qualitative data analysis and interactive data visualization. The workflow is illustrated in figure 2. The comparison of the performances of the Orange data mining-based Decision Tree, Linear Regression, the Random Forest, AdaBoost, Support Vector Machine and the Gradient Boosting learners is presented as a summary in Tables 2.

The various input features of the surface roughness in Table 1 were selected for training the ML models, and all the ML models include DT, LR, XGBoost, AdaBoost, SVM and RF. Each ML model was optimized using the same input conditions, and the root mean square (RMSE), mean absolute error (MAE), mean squared error (MSE) and coefficient of determination (R^2) values for training of each ML model are shown in Table 2.

According to different goodness-of-fit indicators, it can be clearly shown that the prediction performance of the SVM model is unsatisfactory compared to XGBoost, AdaBoost, DT, Random forest and LR. The reason for this phenomenon may be that there are few sample data affecting the surface roughness values of EN AW-1350 Aluminum Alloy. Considering the goodness of fit and performance, the model based on the XGBoost algorithm can be used as an ideal model for predicting the surface roughness values of EN AW-1350 Aluminum Alloy (RMSE =0.003, R2= 1.000). Meanwhile, the AdaBoost model has the smallest error among all models and the R2 is almost close to 1 (RMSE =0.004, R2= 0.999). Compared with the above models, XGBoost can be the most ideal model for predicting surface roughness of EN AW-1350 Aluminum Alloy.

Fig. 4 shows a comparison of the predicted and actual surface roughness values for each ML model. The solid blue line in the figure represents the measured values, and the solid red line represents the predicted value of the proposed ML model.

©UBMA - 2024

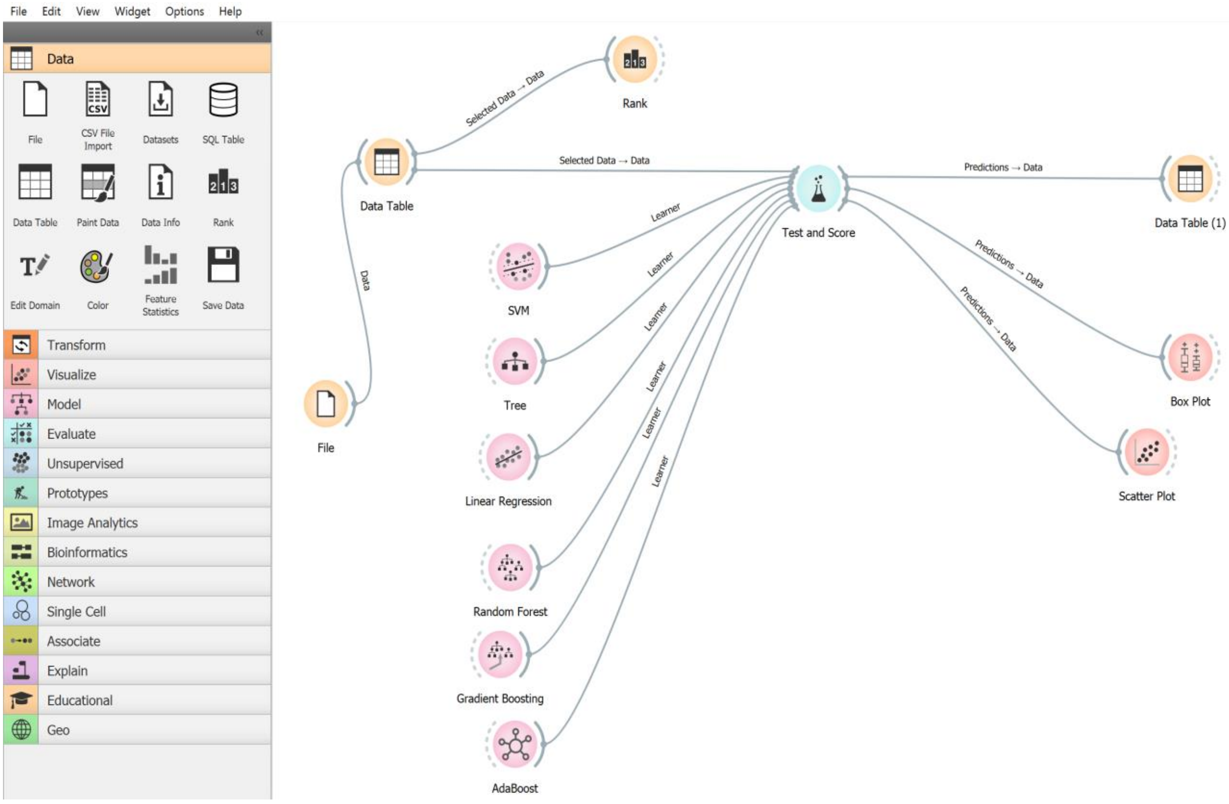


Figure. 2. The user interface of Orange data mining® (The workflow is built in the center of the screen, on the right there are the node and workflow repository views).

Table 2. Comparative evaluation of data mining algorithms performance using Orange software®

Algorithm	MSE	RMSE	MAE	R^2
Decision Tree	0.000	0.016	0.012	0.992
Random Forest	0.001	0.031	0.025	0.967
Linear Regression	0.002	0.043	0.038	0.939
AdaBoost	0.000	0.004	0.002	0.999
Support Vector Machine	0.005	0.071	0.058	0.834
Gradient Boosting	0.000	0.003	0.002	1.000

The comparison between actual and predicted response for Ra is illustrated in Fig. 3. According to those figures, it can be seen that points split is evenly by the 45-degree line. This reflects the good agreement between experimental values illustrated in Table 2 and predicted values obtained with machine learning algorithms.

©UBMA - 2024

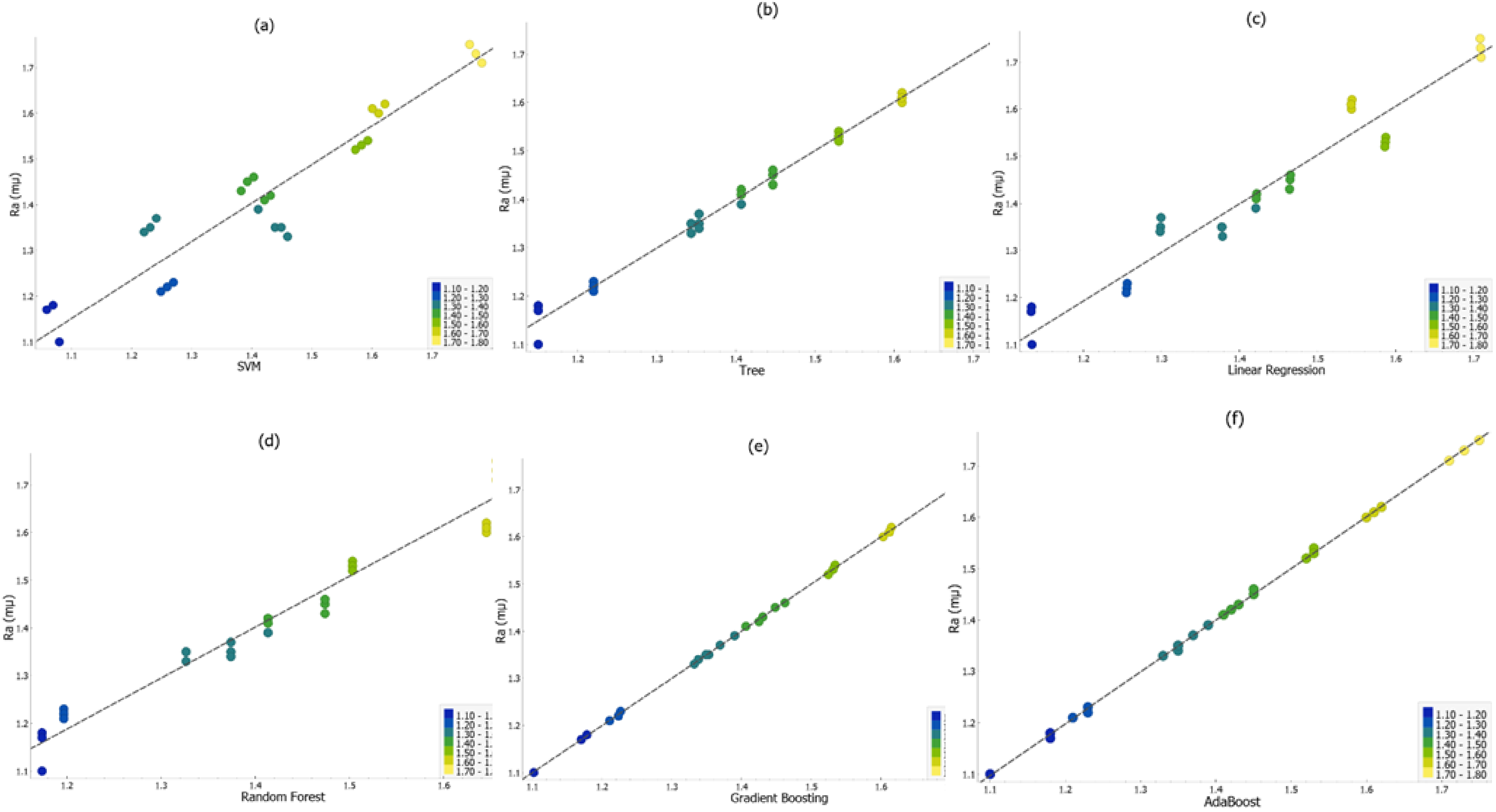
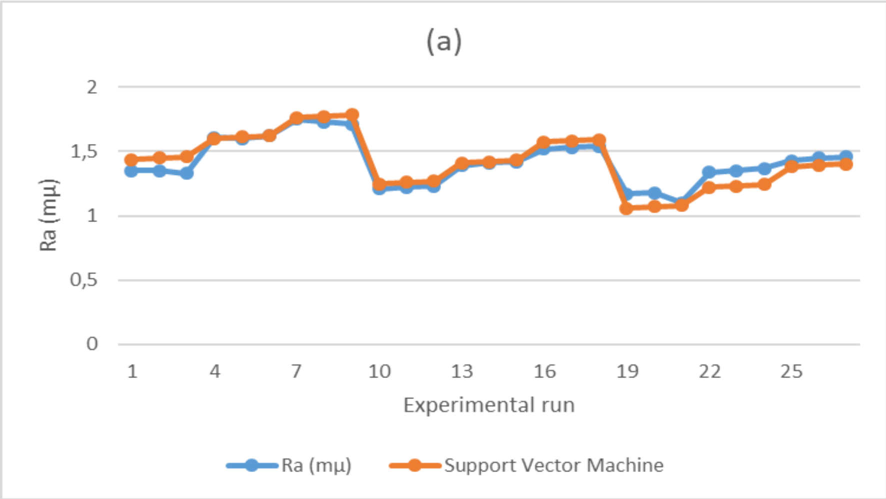
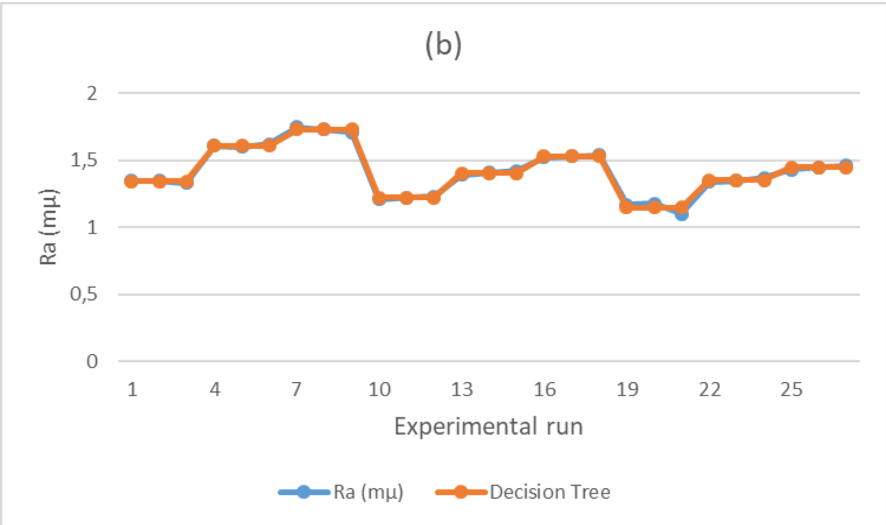
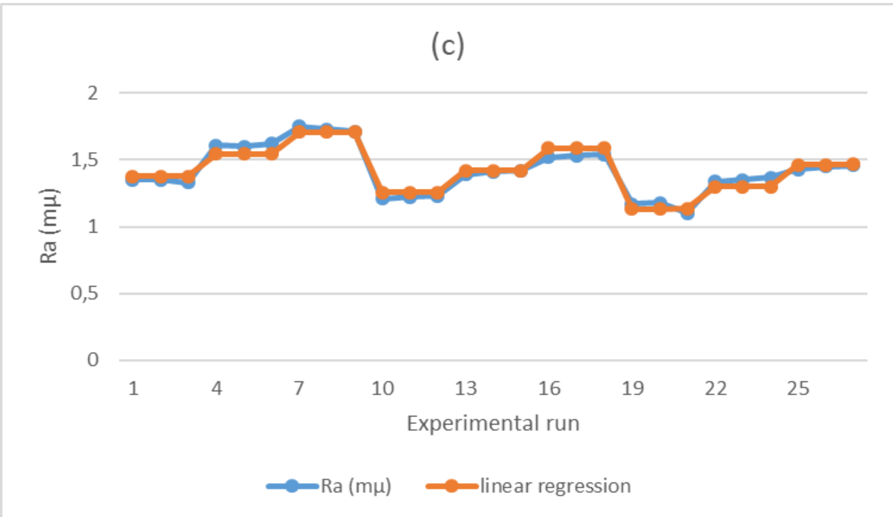


Figure. 3. Comparison between predicted and experimental results

©UBMA - 2024


(a)
2
1,5
1
0,5
0
Ra (mµ)
1
4
7
10
13
16
19
22
25
Experimental run
Ra (mµ)
Support Vector Machine


(b)
2
1,5
1
0,5
0
Ra (mµ)
1
4
7
10
13
16
19
22
25
Experimental run
Ra (mµ)
Decision Tree


(c)
2
1,5
1
0,5
0
Ra (mµ)
1
4
7
10
13
16
19
22
25
Experimental run
Ra (mµ)
linear regression

©UBMA - 2024

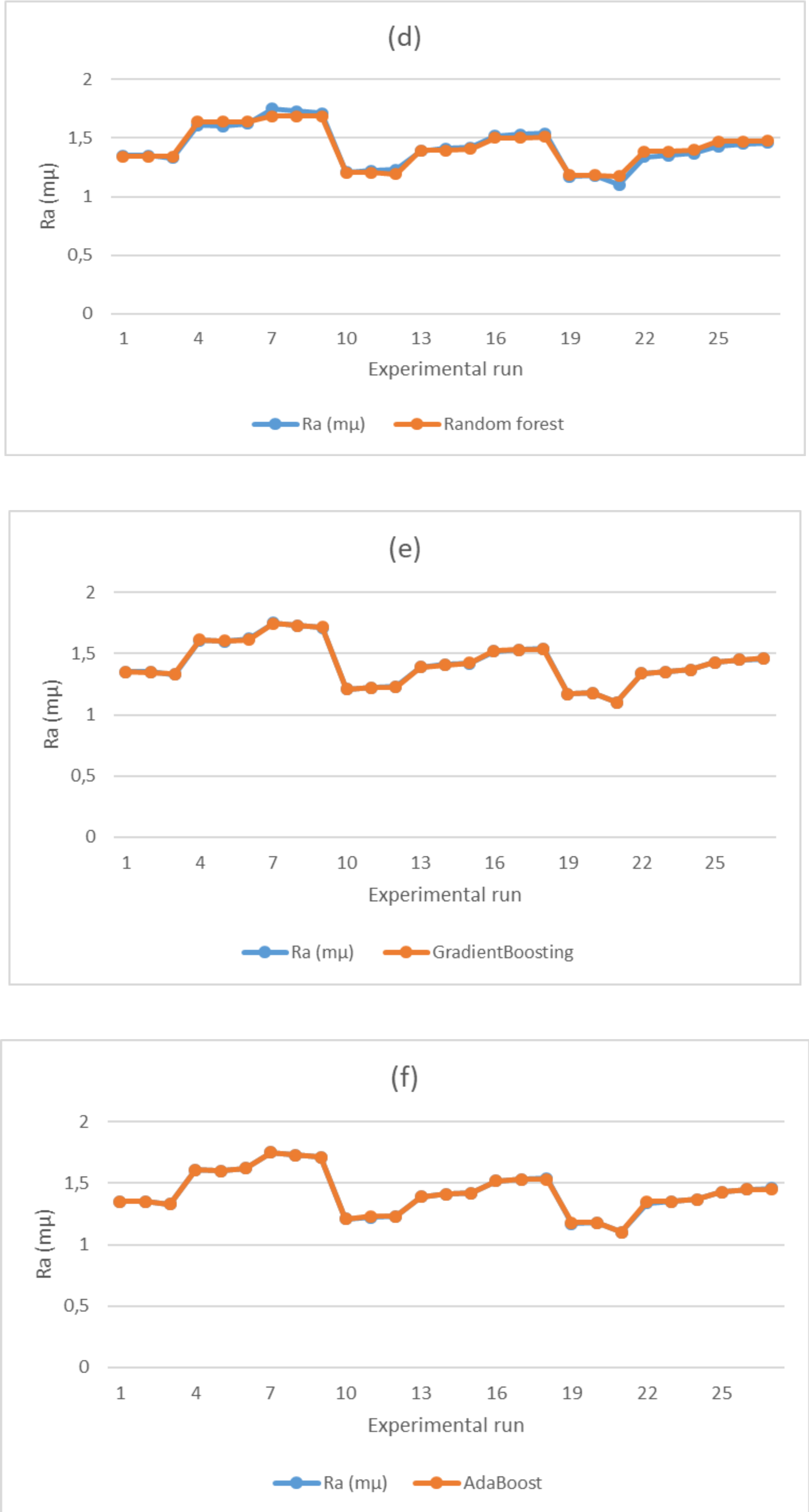


Figure. 4. Comparison of observed and predicted surface roughness based on DT, LR, XGBoost, AdaBoost, SVM and RF models: (a) Support Vector Machine -based model (R^2 = 0.834); (b) Decision Tree -based model (R^2 = 0.992); (c) Linear Regression -based model (R^2 = 0.939), (d) Random forest-based model (R^2 = 0.967), (e) XGBoost-based model (R^2 = 1.000), (f) AdaBoost-based model (R^2 = 0.999).

©UBMA - 2024

In order to better compare the performance of each model, Fig. 4 compares the predicted and observed values through data visualization. It can be seen that the predicted and observed curves of the high-performance model (XGBoost, AdaBoost and DT) almost overlap, while the predicted curves of LR, SVM and RF fluctuate more than the observed curves. In conclusion, a variety of ML models such as DT, LR, RF, AdaBoost, SVM and XGBoost were established and trained to predict the surface roughness values (Ra) of EN AW-1350 Aluminum Alloy, and all six ML models could successfully predict the surface roughness parameters. Compared to other ML methods, LR, SVM and RF models have poor model training and testing performance. According to R^2 as the evaluation index of ML (XGBoost > AdaBoost > DT > RF > LR> SVM), same as training. The results show that both the proposed XGBoost, AdaBoost and DT models can be effectively used for the prediction of the surface roughness of EN AW-1350 Aluminum Alloy without laboratory experiments. This can reduce experimental time and cost, and guide the experimental process accordingly.

5. CONCLUSION

In the present study, an attempt has been made to investigate the effect of process parameters (cutting speed, feed rate and depth of cut) on the performance characteristic (surface roughness) in finish turning of EN AW-1350 Aluminum Alloy using the regression algorithms, and the results obtained through machine learning analysis can be summarized as follows:

1. The analysis of the cutting parameters through the application of machine learning approaches allowed the exploration of the influence of each factor on the response outputs represented in the present case by surface roughness.
2. The high R squared values obtained for the regression models indicates a good fit and the same may be used to predict the result for any values of cutting speed, feed rates and depths of cut.
3. The XGBoost, AdaBoost and DT model both showed the better prediction performance than RF, LR and SVM model.
4. The performance of machine learning models is influenced by a multitude of factors, one of which is model complexity. Boosting methods such as XGBoost and AdaBoost construct their models sequentially. Each subsequent model is designed to address the inaccuracies of its predecessors. This iterative refinement enables the creation of more sophisticated models capable of discerning intricate patterns within the data.
5. The machine learning model for surface roughness indicates that feed rates has the highest contribution of 55%.

Acknowledgements:

This work is supported by the General Directorate of Scientific Research and Technological Development (DGRSDT) affiliated to the Algerian Ministry of Higher Education and Scientific Research (MESRS). Fruitful discussions with members of LMC (U. Constantine-1) and LR3MI (UBM Annaba) laboratories are highly appreciated.

Nomenclature

ap	*depth of cut (mm)*
Vc	*cutting speed (m / min)*
f	*feed rate (mm / rev)*
Ra	*arithmetic average of absolute roughness (μm)*
R^2	*coefficient of determination*
MSE	*mean squared error*
RMSE	*root mean square error*

Abbreviations

DT	*Decision Trees*
XGBoost	*eXtreme Gradient Boosting*
AdaBoost	*Adaptive Boosting*
SVM	*Support Vector Machines*
RF	*Random Forest*
LR	*Linear Regression*
ML	*Machine Learning*

©UBMA - 2024

REFERENCES:

[1]. Zhang, W. W., & Noack, B. R. 2021. Artificial intelligence in fluid mechanics. *Acta Mechanica Sinica*, *37*(12), 1715-1717.

[2]. Vinuesa, R., Brunton, S. L., & McKeon, B. J. 2023. The transformative potential of machine learning for experiments in fluid mechanics. *Nature Reviews Physics*, *5*(9), 536-545.

[3]. Fakharian, P., Eidgahee, D. R., Akbari, M., Jahangir, H., & Taeb, A. A. 2023. Compressive strength prediction of hollow concrete masonry blocks using artificial intelligence algorithms. In *Structures* (Vol. 47, pp. 1790-1802). Elsevier.

[4]. Baghbani, A., Choudhury, T., Samui, P., & Costa, S. 2023. Prediction of secant shear modulus and damping ratio for an extremely dilative silica sand based on machine learning techniques. *Soil Dynamics and Earthquake Engineering*, *165*, 107708.

[5]. Olabi, A. G., Abdelghafar, A. A., Maghrabie, H. M., Sayed, E. T., Rezk, H., Al Radi, M., ... & Abdelkareem, M. A. 2023. Application of artificial intelligence for prediction, optimization, and control of thermal energy storage systems. *Thermal Science and Engineering Progress*, 101730.

[6]. Ahmed, S. K., Ali, R. M., Lashin, M. M., & Sherif, F. F. 2023. Designing a new fast solution to control isolation rooms in hospitals depending on artificial intelligence decision. *Biomedical Signal Processing and Control*, *79*, 104100.

[7]. Khrouf, F., Tebassi, H., Yallese, M. A., Chaoui, K., & Haddad, A. 2022. Modeling and optimization of cutting parameters when turning EN-AW-1350 aluminum alloy. *International Journal of Applied Mechanics and Engineering*, *27*(2), 124-142.

[8]. Gabsi, A. E. H., Ben Aissa, C., & Mathlouthi, S. 2023. A comparative study of basic and ensemble artificial intelligence models for surface roughness prediction during the AA7075 milling process. *The International Journal of Advanced Manufacturing Technology*, *126*(1-2), 1-15.

[9]. Adizue, U. L., Tura, A. D., Isaya, E. O., Farkas, B. Z., & Takács, M. 2023. Surface quality prediction by machine learning methods and process parameter optimization in ultra-precision machining of AISI D2 using CBN tool. *The International Journal of Advanced Manufacturing Technology*, 1-20.

[10]. Pimenov, D. Y., Bustillo, A., Wojciechowski, S., Sharma, V. S., Gupta, M. K., & Kuntoğlu, M. 2023. Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review. *Journal of Intelligent Manufacturing*, *34*(5), 2079-2121.

[11]. Bansal, M., Goyal, A., & Choudhary, A. 2022. A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short-term memory algorithms in machine learning. *Decision Analytics Journal*, *3*, 100071.

[12]. Mahesh, P. V., Meyyappan, S., & Alla, R. 2022. Maximum power point tracking with regression machine learning algorithms for solar PV systems. *International Journal of Renewable Energy Research*, *12*(3), 1327-1338.

[13]. Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F., & El-Shafie, A. 2021. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal*, *12*(2), 1545-1556.

[14]. Zhou, X., Ren, J., An, J., Yan, D., Shi, X., & Jin, X. 2021. Predicting open-plan office window operating behavior using the random forest algorithm. *Journal of Building Engineering*, *42*, 102514.

[15]. Taherkhani, A., Cosma, G., & McGinnity, T. M. 2020. AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning. *Neurocomputing*, *404*, 351-366.

[16]. Azarmdel, H., Jahanbakhshi, A., Mohtasebi, S. S., & Muñoz, A. R. 2020. Evaluation of image processing technique as an expert system in mulberry fruit grading based on ripeness level using artificial neural networks (ANNs) and support vector machine (SVM). *Postharvest Biology and Technology*, *166*, 111.

©UBMA - 2024