



Prediction of surface roughness in duplex stainless steel top milling using machine learning techniques

Guilherme Augusto Vilas Boas Vasconcelos¹ · Matheus Brendon Francisco² · Carlos Henrique de Oliveira² · Elioenai Levi Barbedo¹ · Luiz Gustavo Paes de Souza² · Mirian de Lourdes Noronha Motta Melo¹

Received: 3 July 2024 / Accepted: 12 August 2024 / Published online: 28 August 2024
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

Abstract

This study presents an innovative approach to predicting the quality of finished parts in top milling processes by integrating robust parameter design with artificial intelligence techniques. A central composite design was used to combine controllable variables (cutting speed, tooth advance, milled width, and cutting depth) with noise variables (tool flank wear, fluid flow, and cantilevered length). Duplex stainless steel was milled under each experimental setup, and roughness data were collected. These data were used to train three machine learning models: random forest, decision tree, and support vector machine. The models predicted surface roughness, and their predictions were validated through experimental tests. The root means square error values were 0.031 for the random forest, 0.038 for the decision tree, and 0.066 for the support vector machine, indicating that the random forest model performed the best. This innovative study highlights the importance of including noise variables along with controllable factors in machine learning models, significantly improving prediction accuracy and making them more reflective of real-world results. Including these variables is crucial, as neglecting them can lead to inaccurate predictions.

Keywords Design of experiments · Artificial Intelligence · Surface quality · Robust parameter design

1 Introduction

Duplex stainless steels are widely used in various industries due to their balanced combination of ferrite and austenite, giving them desirable properties such as ductility, toughness, and resistance to stress corrosion [24]. However, machining these alloys is challenging due to their low machinability, which results in problems such as crunching, low thermal conductivity, and high strength, leading to tool wear and compromise of workpiece integrity. These stainless-steel alloys are widely used in various industrial fields such as automotive, aerospace, medical, chemical, and petrochemical [37].

Milling is a subtractive machining process that uses a rotating tool to cut flat surfaces. There are two main milling

methods: up milling and down milling. In down milling, the cutting tool and the workpiece rotate in the same direction, which eliminates the possibility of chatter, as the tool's teeth cut into the workpiece and deposit the cut chips behind the cutter. The chip thickness starts at its maximum and decreases as the cut progresses. Due to the complexity of machining duplex stainless steels, down milling methods often do not provide the required level of finish. CNC machining (up milling) stands out as a superior alternative, offering much higher surface quality and finish compared to traditional methods and is widely used in modern industry for producing mechanical parts with high precision and quality [26].

In the context of milling, surface quality monitoring is essential due to the significant investment of time and money in this process. The average surface roughness, measured by the R_a parameter, plays a crucial role in the esthetics of the products, as well as in the resistance to corrosion, fatigue, and tribological properties. The R_a value is influenced by several factors, such as feed rate, cutting speed, cutting depth, tool geometry, tool wear, temperature, and false edge formation.

✉ Guilherme Augusto Vilas Boas Vasconcelos
guilhermevasconcelos01@yahoo.com.br

¹ Mechanical Engineering Institute, Federal University of Itajubá – UNIFEI, Itajubá, Brazil

² Production and Management Institute, Federal University of Itajubá – UNIFEI, Itajubá, Brazil

There has been a growing interest in predicting surface roughness through artificial intelligence techniques, as this would bring significant benefits to the industry, including reducing waste, production costs, and increasing accuracy [22]. Recent studies have indicated that the application of machine learning techniques has demonstrated high capabilities to make accurate predictions as to surface roughness during the milling process [19].

Pimenov et al. [28] tested different machine learning models, such as random forest, standard multilayer perceptron (MLP), decision trees, and radial-based functions, for real-time prediction of surface roughness deviations from faceting machining processes. The results show that the random forest has the highest comparative precision followed by the decision trees, exhibiting higher precision than the standard MLP and radial-base function.

[34] applied predictive models of trees with reinforced gradient for surface roughness in high-speed milling in the steel and aluminum metallurgical industry. The results show accuracy ranging from 61.54% to 88.51% in the datasets, which are competitive results when compared to the other approaches, the axial cutting depth is the most influential feature for the slot datasets, and the hardness and diameter of the cutting tool are the most influential features for the geometry datasets. [16] studied the effect of machinability, microstructure, and hardness of deep cryogenic treatment on the hard turning of AISI D2 steel with ceramic cutting. The artificial intelligence method known as artificial neural networks (ANN) was used to estimate surface roughness based on cutting speed, cutting tool, part, cutting depth, and feed rate.

[7] used three machine learning algorithms, namely polynomial regression (PR), SVR, and Gaussian process regression (GPR), to predict cutting force and cutting power in the milling of AISI 1045 under minimal quantity lubrication and cutting environments with high pressure cooling. For the development of the predictive models, the machining parameters cutting speed, cutting depth, and feed rate were considered as control factors. The results showed that the SVR and GPR models have better performances when compared to the PR model. The results revealed that the ANN-based model has slightly better results in terms of accuracy than the other machine learning methods studied when estimating the quality trait. The authors further reinforce that the selected machine learning techniques produce adequate results when compared to the experimental results. Thus, using developed models, acceptable results can be estimated rather than obtained experimentally, which consequently reduces the cost and time of testing. The application of artificial intelligence algorithms to milling processes is significant.

[12] estimated the surface roughness of AA6061 alloy in milling using artificial neural networks and response

surface methodology. For these models, cutting speed, cutting depth, and feed rate were evaluated as input parameters for the experimental design. The results of neural networks showed that R^2 , MAE, and RMSE were calculated as 92.7%, 28.11, and 0.185, respectively. According to the results of the response surface methodology modeling, R^2 , MAE, and RMSE were calculated as 99.9%, 2.17, and 0.016, respectively. Karthik et al. (2021) performed the prediction and optimization of surface roughness and flank wear during high-speed milling of aluminum alloy 6061. Four factors were examined: feed rate, cutting speed, cutting depth, and cutting length. For the prediction, four algorithms were used: linear regression (LIN), support vector machine regression, a gradient augmentation tree (GBR), and an artificial neural network (ANN). The results showed that support vector machine and neural networks presented the best predictive performance surface roughness and maximum flank wear.

[13] performed a multiobjective optimization of surface roughness and tool wear in high-speed milling of AA6061 by machine learning techniques and NSGA-II. SVR and ANN showed the best predictive performance for surface roughness and maximum flank wear. [15] studied machine learning-based quality prediction for milling processes using internal machine tool data. The results show that ensemble methods such as Random Forest and Extra Trees, as well as the deep learning inception time and ResNet algorithms achieve the best performances for the data-driven quality prediction. [31] studied tool wear prediction in stainless steel faceting using singular generative adversarial network and LSTM deep learning models. The authors portray that the stacked LSTM model better predicts tool wear compared to other LSTM models.

[36] conducted a study on cutting forces in the milling process of functionally classified materials using machine learning. The authors developed Gaussian process regression models to perform the prediction of the main cutting force and its components in three directions (F_x , F_y , and F_z) in a coordinate system. This prediction was based on two predictors: the cutting depth (a_p) and the feed rate (f), in milling processes of functionally classified materials. The model results demonstrated high precision and stability, indicating that these models have the potential to quickly, economically, and reliably estimate cutting force and its components.

It is also possible to note that there are many published works using machine learning to predict the characteristics of pieces in the literature. [23] carried out a modeling and tagging of time sequence signals in the milling process based on an improved hidden semi-Markov model. [35] evaluated ensemble learning with a genetic algorithm for surface roughness prediction in multi-jet polishing. [10] studied the online monitoring model of micro-milling force incorporating tool wear prediction process.

Based on literature reviews, it is evident that the use of machine learning algorithms has been shown to be effective in predicting important characteristics in milling machining operations, such as surface roughness. These approaches have been widely applied in several studies, bringing competitive and promising results in relation to other traditional modeling techniques.

However, the differentiator of this study is the addition of uncontrollable variables (noise) to the controllable variables of the process for creating machine learning models. Noise represents external or random factors that can influence the quality characteristics of the machining process, but which are outside the direct control of the operators or the pre-established settings. The consideration of these uncontrollable variables in creating machine learning models is extremely relevant, as it more accurately reflects the real operating conditions in manufacturing industries, where there are multiple sources of variation and uncertainty.

This article is divided in 5 sections: Sect. 2 presents a literature review. Section 3 then discusses the research methodology adopted, offering valuable insights into the process and procedures used. Section 4 presents the results and discussions, providing a critical and enlightening analysis of the collected data. Finally, Sect. 5 brings the conclusions of this work, offering an engaging synthesis of all the elements addressed and culminating in a satisfactory and significant outcome.

2 Theoretical framework

2.1 Design of experiments

Design of experiments (DOE) is a statistical methodology used to plan, conduct, and analyze experiments efficiently and systematically. The goal of DOE is to gain valuable and relevant information about how specific variables affect a process or system. With this methodology, researchers can optimize processes, improve products, and identify critical factors that influence the outcome of the experiment 25.

In the context of design of experiments, the factorial design is a very useful strategy for investigating the effect of multiple factors on an experimental system. The factorial design involves testing all possible combinations of the levels of each factor, which allows you to analyze the main effects of each factor and also the interactions between them. This method is particularly useful when you want to identify which factors have the greatest influence on the outcome of the experiment and how they might interact with each other.

The fractional factorial design is an extension of the concept of factorial design that is used when the number of combinations of factor levels is too large to be tested in a complete design. In some situations, testing every possible

combination can be expensive, time-consuming, or impractical. In this scenario, the researchers use a fractional factorial design, which consists of selecting only a subset of the total combinations to be tested. This selection is made strategically, using a fraction plan of the full design.

For example, consider an experiment with three factors, each with two levels (high and low). In the full factorial design, $2^3 = 8$ combinations would be required, but a fractional plan could select only 4 combinations, saving time and resources. The fractional factorial design can provide important information about the main effects of the factors, but some interactions may not be completely identified due to the lack of testing with all possible combinations. The choice of fractional plan depends on the nature of the experiment, the number of factors, and the purpose of the research.

In summary, both the full factorial design and the fractional factorial design are design of experiments techniques that allow to study the effect of multiple factors on a system. The first covers all possible combinations, while the second selects a strategic subset of these combinations to save resources but still provide relevant information about the factors and their interactions.

Going further, the combined design is used to study how controllable and noise factors affect the response or variable of interest in the experiment. The main idea is to create an experimental plan that allows you to control and measure both types of factors to analyze their impact on the outcome. A practical example would be a manufacturing study in which the goal is to optimize roughness in machining processes. In this case, the controllable factors can be cutting speed, advance, depth, etc. Noise factors can include variations in raw material quality, tool wear, etc.

The combined design would allow to design an experiment that controls the selected (controllable) factors while recording and considering the random variations (noise). This would provide a more complete analysis of the effects of the factors on the roughness of the product and would help identify the best configuration to achieve the least roughness.

In summary, the combined design is a powerful approach to design experiments that involves the combination of controllable and noise factors. This technique allows you to understand how these factors influence the outcome of the experiment and helps you make more informed decisions to optimize the process or product under study.

2.2 Decision tree regression

Decision trees are classic machine learning algorithms that can be used for both classification and regression. Although their learning ability is not excellent when considered individually, they are known for their ability to generalize and filter features. When applied to regression tasks, they are called regression trees 27. Compared to other algorithms

that require data preprocessing, the decision tree algorithm requires less effort in this process. There is no need to perform normalization or scaling of the data. In addition, decision trees can handle incomplete data. In terms of disadvantages, a small change in the data may require several changes to the structure of the decision tree, which can cause instability. The computational time required to train a decision tree is often high.

The process of algorithm training begins with the entire dataset at the root node of the tree. The algorithm searches for the feature and the corresponding value that can best split the data into two or more subsets in a way that minimizes the variance of the target variable within each subset. After finding the best split, the algorithm creates child nodes, each representing a subset of the data based on the split criterion. It continues this process recursively for each child node until a stopping condition is met, such as reaching a maximum tree depth or a minimum number of samples required to create a node. At each terminal node (leaf), instead of returning a class label as in classification tasks, a regression decision tree returns the predicted value for the target variable. This value is usually the mean or median value of the target variable in that subset.

To make predictions, new data points are passed down the tree, and at each node, the algorithm evaluates the corresponding feature's value. It follows the branches based on the feature values until it reaches a leaf node. The value at the leaf node becomes the prediction for the target variable for that particular data point.

Like other decision trees, regression decision trees are easy to understand and interpret. The tree structure can be visualized, allowing users to see how the decision-making process flows. Regression decision trees can capture non-linear relationships between features and the target variable, which can be challenging for linear regression models. Decision trees can naturally handle feature interactions without requiring feature engineering. Regression decision trees are less affected by outliers compared to some other regression techniques, as the splitting process is based on data percentiles.

Just like classification decision trees, regression decision trees are prone to overfitting when the tree becomes too complex and captures noise in the training data. Small changes in the data can lead to significantly different trees, making the model unstable and sensitive to data variations. Regression decision trees cannot accurately predict values outside the range of the training data, as they make predictions based on the mean or median values within the training subsets.

To address the limitations of a single regression decision tree, ensemble techniques like Random Forests is often used. These methods combine the predictions of multiple decision trees to improve the model's accuracy, generalization, and

robustness. Ensemble techniques help reduce overfitting, provide better stability, and enable improved predictions for both the data points within the training range and those slightly beyond it.

2.3 Random Forest

The Random Forest algorithm is a powerful approach to the field of machine learning that has proven extremely efficient for solving complex problems in several areas of engineering. It belongs to the category of ensemble methods, which combine several individual predictions of simpler models to achieve more reliable and accurate results [6].

Random Forest is built on decision trees, which are hierarchical decision structures that divide the dataset into more homogeneous subsets. Each tree is trained on a random subset of the dataset, through replacement sampling [32]. The decision nodes and leaves represent the structure of the decision tree, where the leaves represent the final results and the decision nodes are points where the data is divided. This randomness and diversity of trees ensure that the ensemble has less tendency to overfitting, making it more generalizable. This model is widely used due to its simplicity and diversity, being applied to both regression and classification problems [4].

The Random Forest training process is divided into two main stages: creating the set of trees and combining the predictions. In the first step, several decision trees are constructed using different samples from the training dataset. For each tree, recursive division is applied iteratively until the nodes are purely homogeneous (for classification tasks) or a maximum depth is reached. In the second step, the predictions of all trees are combined through voting (for classification) or mean (for regression) to produce the final answer [32].

Random Forest is widely used in engineering in various applications, such as equipment failure prediction, industrial process optimization, time series analysis, and maintenance decision-making, among others. Its ability to handle complex, multidimensional data sets makes it especially useful for problems with uncertainties and variations. Considered one of the most powerful and versatile machine learning techniques, Random Forest has advantages such as reduced overfitting, high precision, missing data handling, and high dimensionality, as well as providing measures of importance of variables. Its applicability in large-scale and complex problems makes it an essential tool in modern engineering [14].

2.4 Support vector machine

Based on the theory of statistical learning, the support vector machine (SVM) was developed by [9], in order

to solve problems of classification and regression. The support vector regression is a specific technique within SVM developed by Drucker et. al. (1997). Its main goal is to find the best fit line represented by a hyperplane with the most points. To determine this hyperplane, the support vector regression selects extreme points/vectors, known as support vectors, that justify the nomenclature of the technique. Support vector regression aims to fit the ideal line within a limit value range, which is the distance between the boundary line and the hyperplane [20].

In real-world situations, it is rare to find applications where the data is linearly separable. This is due to various factors, such as the presence of noise and outliers in the data or the inherently non-linear nature of the problem. To address these challenges, it is allowed for some data to violate the constraint $y_i(w \cdot x_i + b) - 1 \geq 0, \forall i = 1, \dots, n$. This is done by introducing slack variables ξ_i , for all $i = 1, \dots, n$ [1]. These variables relax the constraints imposed on the primal optimization problem, resulting in a more flexible approach as shown in Eq. (1) 3:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, \forall i = 1, \dots, n \tag{1}$$

The application of this procedure smooths the margins of the linear classifier, allowing some data to remain between the hyperplanes $H1$ and $H2$ and also permitting some classification errors. For this reason, the SVMs obtained in this case can also be referred to as SVMs with soft margins. An error in the training set is indicated by a value of ξ_i may than 1. Thus, the sum of the ξ_i represents a limit on the number of training errors [8]. To account for this term, thereby minimizing the error on the training data, we have Eq. (2):

$$Min_{w, b, \xi} \frac{1}{2} \|W\|^2 + C \left(\sum_{i=1}^n \xi_i \right) \tag{2}$$

The constant C is a regularization term that weights the minimization of errors in the training set relative to the minimization of model complexity [2]. The presence of this term $\sum_{i=1}^n \xi_i$ in the optimization problem can also be seen as a minimization of marginal errors, as a value of $\xi_i \in (0,1)$ indicates data between the margins.

SVM training involves choosing the optimal hyperplane that best separates the training data in your classes. For this, SVM uses an optimization process that seeks to maximize the margin between the points of each class and the separation hyperplane. For a nonlinear regression, the concept of Kernel function is introduced, which serves to implicitly map the feature vectors to a larger space where it is assumed that the data are linearly separable. By mapping the vectors in a nonlinear way and applying the technique, a nonlinear function results in the original space, even if it is linear in another space.

SVM is widely applied in engineering in various scenarios, such as fault diagnosis in complex systems, pattern classification in images and signal processing, event prediction and decision-making in control, and automation engineering, among others. Its ability to handle high-dimensional data and the ability to adapt to nonlinear problems make it especially useful in engineering applications [26].

In terms of advantages, SVM can work very well with high-dimensional input space and is relatively memory efficient. In terms of disadvantages, SVM is not suitable for large data sets; it does not work well with any data type (e.g., dataset with more noise) [30].

2.5 Performance indicators

Three distinct performance indicators were chosen to evaluate the accuracy of the models in predicting surface roughness values. These indicators are the coefficient of determination (R^2), the mean square error (MAE), and the root of the mean square error (RMSE), as shown in Eqs. (3) to (5), respectively:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \tag{3}$$

$$MAE = \frac{1}{n} \sum_{i=0}^n |Y_i - \hat{Y}_i| \tag{4}$$

$$RMSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y})^2 \tag{5}$$

where n is the number of data points, Y_i represents observed values, \hat{Y} represents predicted values, and \bar{Y} means the mean value of Y .

When comparing the values of the metrics, priority will be given to the evaluation criteria chosen for the RMSE, since it is a more appropriate method than the MAE when the errors of the model follow a normal distribution. In addition, RMSE has a significant advantage over MAE as it avoids the use of absolute values, which may be undesirable in many mathematical calculations [33]. Therefore, when comparing the accuracy of various regression models, RMSE is a more suitable choice because it is easy to calculate and has differentiability. In addition, a higher value for R^2 is considered desirable.

Importantly, we will do an initial analysis on the data before using it in machine learning models. Outliers, for example, can have a significant impact on the accuracy of machine learning models. They can skew the results and negatively affect the model's ability to correctly generalize

patterns in the data. Outliers can violate these assumptions, compromising the validity of statistical analyses and the results obtained. Therefore, it is important to identify and deal with outliers to ensure that model assumptions are met [29].

Some algorithms are sensitive to outliers, meaning their performance can be severely affected by the presence of these outliers. Outliers may arise due to measurement errors or data corruption. Identifying and correcting these outliers is essential to ensuring the quality and integrity of the data used in model training. Therefore, doing the analysis of outliers in the data before applying machine learning algorithms is fundamental to obtain more accurate, robust and reliable models, as well as ensuring the validity of the statistical analyses and the quality of the data used.

In model performance evaluation, overfitting occurs when a model overfits the training data, capturing even the noise and outliers present in them. This results in a model that does not generalize well to new data. By treating outliers, it is possible to reduce the risk of overfitting and improve the model's ability to make accurate predictions on unseen data 5.

Finally, ML model optimization is the main challenge to achieve an effective machine learning solution. Hyperparameter optimization aims to find the optimal values for the model parameters, resulting in the best performance as assessed by the validation set, within a given machine learning algorithm. These hyper parameters are responsible for controlling the learning process and have a significant impact on predictive performance. In addition, an appropriate selection can avoid problems of overfitting and underfitting, thus increasing the accuracy of predictions (Nguyen et al., 2020). In the present study, several hyperparameters were analyzed and the best ones were chosen to be used.

3 Methodology

The top milling operation was performed in a ROMI D600 machining center, as shown in Fig. 1, with a power of 15 kW and a maximum rotation of 10,000 rpm. The part to be machined is duplex stainless steel, which has low machinability due to its low thermal conductivity. The chemical structure of duplex stainless steel UNS S32205 is mentioned in Table 1. The insert used in the cutting operation was the CoroMill R390-11T308M-MM 2030, made of carbide and with double layer of titanium nitride (TiN) and aluminum titanium nitride (TiAlN), coated by the process of physical vapor deposition (PVD), fixed in the CoroMill® R390-025A25-11 M support, with a diameter of 25 mm, position angle $\chi_r = 90^\circ$, cylindrical rod, with 3 inserts and mechanical fixation by tweezers. Both the inserts and the tool holder were provided by Sandvik Coromant.



Fig. 1 ROMI® D 600 machining center

The data that will be used as the database for the creation of machine learning algorithms will be obtained through an experimental design, in which a combined arrangement was chosen. A half-fraction factorial design was created, containing both control and noise variables.

The controllable factors of the process were cutting speed, tooth advance, cutting width, and cutting depth, and the selected levels are shown in Table 2. The uncontrollable parameters were the cantilevered length of the tool, the flow of the cutting fluid, and the flank wear, as shown in Table 3, while surface roughness was considered the response parameter. The experimental execution was carried out randomly, in order to minimize errors from other variables that were not considered, due to the impossibility of being measured and/or because they are unknown. Table 4 shows the experimental matrix used to collect data on surface roughness.

To control the overhang length (l_{t0}) during the experimental tests, a set of clamping devices was used, as shown in Fig. 2. The value of l_{t0} was verified using a Digimes® analog caliper with a resolution of 0.05 mm.

Regarding the amount of fluid (Q), two regulating valves (1 and 2) were used to control the flow during the top milling of duplex stainless steel UNS S32205. To ensure minimal flow in the machine tool, a small opening was made in valve 1, and the flow rate was measured using a graduated beaker. For maximum flow, both valves were fully opened. In the case of “dry” machining, the valves were closed to prevent the fluid from being directed to the cutting area. The valves used to control the fluid quantity in the process can be observed as shown in Fig. 3.

During the execution of the experiments, the measurements of tool flank wear (v_b) were obtained using the image analyzer (Global Image Analyzer), the Global Lab 97 Image software, and the stereoscopic microscope model SZ 61 (with 45 times magnification), as shown in Fig. 4.

Surface roughness measurements were obtained using a calibrated Mitutoyo Surftest 201 portable roughness tester before the start of measurements, as shown in Fig. 5. The

Table 1 Chemical composition (% by weight) of duplex stainless steel UNS S32205

C	Si	Mn	P	S	N	Al	Cr	Mo	Ni	Cu	W	Co
0.013	0.47	1.22	0.019	0.01	0.19	0.008	22.24	3.14	5.62	0.19	0.02	0.05

Table 2 Control variables and their respective levels

Control variables	Levels for controllable variables				
	−2.83	−1.00	0.00	+1.00	+2.83
Cutting speed [m/min]	32.57	60	75.00	90.00	117.43
Tooth feed [mm/tooth]	0.04	0.10	0.13	0.16	0.21
Cutting width [mm]	12.26	15.00	16.50	18.00	20.74
Cutting depth [mm]	0.43	0.80	1.00	1.20	1.57

Table 3 Levels defined for noise variables

Noise variables	Levels for noise variables			
	−1	0		+1
Flank wear [mm]	0	0.15		0.30
Fluid flow [l/min]	0		0.20	20
Cantilevered length [mm]	30		40	50

cutoff parameter was set to 0.8 mm for all measurements, as for this sampling length, roughness values of Ra are expected to vary between 0.1 and 2 μ m. The measurements were taken perpendicular to the machining groove. Measurements were made at the beginning, in the middle, and at the end. Table 4 displays the experimental matrix used for collecting surface roughness data. The axial points of the noise were excluded from this matrix, as machining them is physically impossible.

After performing the experiments, we move on to the construction part of the machine learning model. The experimental data were divided into training and test sets, representing respectively 70% and 30% of the total number of experiments performed, which corresponds to 50 training attempts and 22 test attempts. All models were built using the Python language. The resulting dataset was normalized to ensure a consistent scale and distribution of all variables.

The training data were used to train three regression algorithms: support vector machine, Random Forest, and decision tree regression. The test data were used to validate the performance of the model. To evaluate the accuracy of the proposed model, error metrics such as root mean square error (RMSE), mean absolute error (MAE), and R^2 parameter were used. Figure 6 shows the flowchart of the methodology applied in this article.

There are several common strategies for optimizing hyper parameters, including manual fit, grid search, random search, Bayesian optimization, gradient-based optimization, and evolutionary optimization [18]. In this study, we used

grid search using the GridSearch CV method, a traditional technique for adjusting hyper parameters. This approach allows to find the best hyperparameters through a grid of combinations in each order [17]. Several hyperparameters were tested for the algorithms, and the best grid values found for the models are presented in Table 5.

Finally, it is important to mention that the method of cross-validation of ten times repeated 10 times was used to train and validate the models. The original dataset is randomly divided into 10 sets of equal size, called folds. The regressor is trained nine times using nine folds as the training set and the last fold as the validation set. The accuracy of the regressor is evaluated in instances that were not used for training. This process is repeated 10 times, using a different fold for validation each time, and the mean RMSE (mean quadratic error) of the validation folds in the 10 repetitions is calculated. In this way, the variance of the regressor's precision is reduced, and its prediction results can be generalized.

4 Results and discussion

4.1 Outliers precision in machine learning models

The analyses of the outliers of the controllable variables of this study can be seen in Fig. 7. It is possible to notice that there are no outliers in the controllable variables. It is worth mentioning that in noise variables, outlier analysis is usually not done because these variables are usually considered to be random and uncontrollable. Noise in a dataset is a source of variation not explained by the independent variables and the model itself. Outliers in noise variables are treated differently from outliers in variables of interest. Generally, outliers in noise variables are not considered outliers that need to be corrected or removed. They are seen as a natural part of random variation and do not have a significant influence on model interpretation or performance.

4.2 Correlation analysis

Figure 8 analyzes the correlations between roughness and other variables (flank wear, tooth advance, cutting speed, cutting depth, cantilevered length, and milled width). By looking at the graph, it is possible to identify the intensity of the correlations and their directions. The result shows that the variable that has the highest positive correlation with roughness is flank wear (v_b). This indicates that with increased flank wear, the cutting edge of the tool becomes

Table 4 Experimental matrix

EXP	Controllable variables				Noise variables			Response
	v_c [m/min]	f_z [mm/Tooth]	a_e [mm]	a_p [mm]	V_B [mm]	Q [l/min]	It_0 [mm]	R_a [μ m]
1	60,00	0,10	15,0	0,80	0,00	0,00	50,0	0,520
2	90,00	0,10	15,0	0,80	0,00	0,00	30,0	0,347
3	60,00	0,16	15,0	0,80	0,00	0,00	30,0	0,630
4	90,00	0,16	15,0	0,80	0,00	0,00	50,0	0,717
5	60,00	0,10	18,0	0,80	0,00	0,00	30,0	0,503
6	90,00	0,10	18,0	0,80	0,00	0,00	50,0	0,360
7	60,00	0,16	18,0	0,80	0,00	0,00	50,0	0,657
8	90,00	0,16	18,0	0,80	0,00	0,00	30,0	0,747
9	60,00	0,10	15,0	1,20	0,00	0,00	30,0	0,503
10	90,00	0,10	15,0	1,20	0,00	0,00	50,0	0,303
11	60,00	0,16	15,0	1,20	0,00	0,00	50,0	0,653
12	90,00	0,16	15,0	1,20	0,00	0,00	30,0	0,700
13	60,00	0,10	18,0	1,20	0,00	0,00	50,0	0,487
14	90,00	0,10	18,0	1,20	0,00	0,00	30,0	0,317
15	60,00	0,16	18,0	1,20	0,00	0,00	30,0	0,653
16	90,00	0,16	18,0	1,20	0,00	0,00	50,0	0,737
17	60,00	0,10	15,0	0,80	0,30	0,00	30,0	0,837
18	90,00	0,10	15,0	0,80	0,30	0,00	50,0	0,867
19	60,00	0,16	15,0	0,80	0,30	0,00	50,0	0,637
20	90,00	0,16	15,0	0,80	0,30	0,00	30,0	0,960
21	60,00	0,10	18,0	0,80	0,30	0,00	50,0	0,930
22	90,00	0,10	18,0	0,80	0,30	0,00	30,0	0,813
23	60,00	0,16	18,0	0,80	0,30	0,00	30,0	0,667
24	90,00	0,16	18,0	0,80	0,30	0,00	50,0	1,003
25	60,00	0,10	15,0	1,20	0,30	0,00	50,0	0,958
26	90,00	0,10	15,0	1,20	0,30	0,00	30,0	0,858
27	60,00	0,16	15,0	1,20	0,30	0,00	30,0	0,683
28	90,00	0,16	15,0	1,20	0,30	0,00	50,0	1,033
29	60,00	0,10	18,0	1,20	0,30	0,00	30,0	0,971
30	90,00	0,10	18,0	1,20	0,30	0,00	50,0	0,719
31	60,00	0,16	18,0	1,20	0,30	0,00	50,0	0,670
32	90,00	0,16	18,0	1,20	0,30	0,00	30,0	1,097
33	60,00	0,10	15,0	0,80	0,00	20,0	30,0	0,453
34	90,00	0,10	15,0	0,80	0,00	20,0	50,0	0,277
35	60,00	0,16	15,0	0,80	0,00	20,0	50,0	0,570
36	90,00	0,16	15,0	0,80	0,00	20,0	30,0	0,630
37	60,00	0,10	18,0	0,80	0,00	20,0	50,0	0,403
38	90,00	0,10	18,0	0,80	0,00	20,0	30,0	0,270
39	60,00	0,16	18,0	0,80	0,00	20,0	30,0	0,487
40	90,00	0,16	18,0	0,80	0,00	20,0	50,0	0,657
41	60,00	0,10	15,0	1,20	0,00	20,0	50,0	0,440
42	90,00	0,10	15,0	1,20	0,00	20,0	30,0	0,287
43	60,00	0,16	15,0	1,20	0,00	20,0	30,0	0,523
44	90,00	0,16	15,0	1,20	0,00	20,0	50,0	0,693
45	60,00	0,10	18,0	1,20	0,00	20,0	30,0	0,430
46	90,00	0,10	18,0	1,20	0,00	20,0	50,0	0,243
47	60,00	0,16	18,0	1,20	0,00	20,0	50,0	0,540
48	90,00	0,16	18,0	1,20	0,00	20,0	30,0	0,667

Table 4 (continued)

EXP	Controllable variables				Noise variables			Response
	v_c [m/min]	f_z [mm/Tooth]	a_c [mm]	a_p [mm]	V_B [mm]	Q [l/min]	l_{t0} [mm]	R_a [μ m]
49	60,00	0,10	15,0	0,80	0,30	20,0	50,0	0,716
50	90,00	0,10	15,0	0,80	0,30	20,0	30,0	0,707
51	60,00	0,16	15,0	0,80	0,30	20,0	30,0	0,447
52	90,00	0,16	15,0	0,80	0,30	20,0	50,0	0,840
53	60,00	0,10	18,0	0,80	0,30	20,00	30,0	0,820
54	90,00	0,10	18,0	0,80	0,30	20,0	50,0	0,773
55	60,00	0,16	18,0	0,80	0,30	20,0	50,0	0,510
56	90,00	0,16	18,0	0,80	0,30	20,0	30,0	0,800
57	60,00	0,10	15,0	1,20	0,30	20,0	30,0	0,823
58	90,00	0,10	15,0	1,20	0,30	20,0	50,0	0,723
59	60,00	0,16	15,0	1,20	0,30	20,0	50,0	0,593
60	90,00	0,16	15,0	1,20	0,30	20,0	30,0	0,830
61	60,00	0,10	18,0	1,20	0,30	20,0	50,0	0,873
62	90,00	0,10	18,0	1,20	0,30	20,0	30,0	0,603
63	60,00	0,16	18,0	1,20	0,30	20,0	30,0	0,620
64	90,00	0,16	18,0	1,20	0,30	20,0	50,0	0,870
65	32,50	0,13	16,5	1,00	0,15	0,20	40,00	0,754
66	117,40	0,13	16,50	1,00	0,15	0,20	40,00	0,767
67	75,00	0,05	16,50	1,00	0,15	0,20	40,00	0,692
68	75,00	0,21	16,50	1,00	0,15	0,20	40,00	1,071
69	75,00	0,13	12,26	1,00	0,15	0,20	40,00	0,933
70	75,00	0,13	20,74	1,00	0,15	0,20	40,00	0,940
71	75,00	0,13	16,50	0,43	0,15	0,20	40,00	0,743
72	75,00	0,13	16,50	1,57	0,15	0,20	40,00	0,784



Fig. 2 Tool holder cone and cutting tool

damaged or worn, which results in lower cutting efficiency, with the tool failing to cut the material properly, leading to a less uniform and rougher machined surface.

Next, the variable that presents the second highest correlation with roughness is feed rate (f_z), again indicating a positive relationship. This can be explained by the fact that

when the feed rate increases, a greater load is applied to the cutting tool, especially when machining harder materials. This additional load can cause vibrations and deformations in the tool, resulting in imperfections in the machined surface and, consequently, greater roughness.

Similarly, the other variables, such as cutting speed (v_c), cutting depth (a_p), cantilevered length (l_{t0}), and milled width (a_c), also show some degree of correlation with roughness, but at lower intensities.

4.3 Predictive performance of models

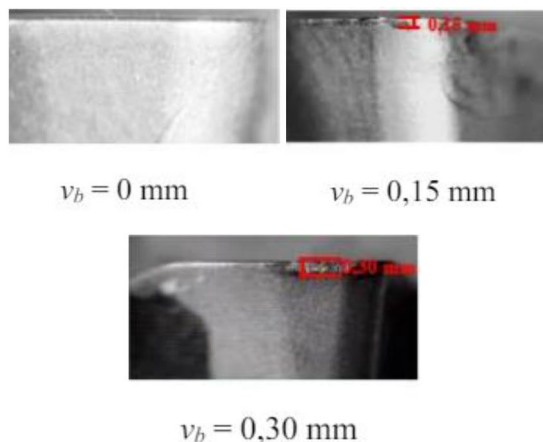
Table 6 presents the performance of the three machine learning models in predicting R_a for the test suites. Based on the results presented and considering the criterion of choosing the best algorithm such as RMSE, we observed that Random Forest obtained the lowest RMSE value, with a result of 0.031. In addition, Random Forest also had the lowest MAE value, with a result of 0.001. In terms of R^2 , both the Decision Tree and the Random Forest obtained high results, with values of 0.96 and 0.97, respectively.

We can conclude that Random Forest presented the best overall performance in relation to the other algorithms (SVM



Fig. 3 Fluid quantity control

Fig. 4 Flank wear of cutting inserts



and Decision Tree) for the prediction of surface roughness in the duplex stainless steel top milling process. Figures 9, 10, and 11 illustrate the line and scatterplots of the predicted and observed values of R_a for the test sets.

4.4 Validation of predicted results

To assess the efficiency of the predictions made by the machine learning algorithms, confirmatory experiments were conducted. The objective of these experiments was to verify the machine learning model’s capability to predict surface roughness in the top milling process of duplex stainless steel UNS S32205 concerning noise variables. The first step of the optimization was to establish the objective function, as shown in Eq. 6.

$$\begin{aligned}
 R_a(x, z) = & 0,425 + 0,015v_c + 0,055f_z + 0,001a_e + 0,008a_p + \\
 & 0,138v_b - 0,053Q + 0,004lt_0 + 0,045vc^2 + 0,060f_z^2 + 0,067a_e^2 \\
 & + 0,045a_p^2 + 0,088v_bf_z - 0,005v_c a_e - 0,011v_c a_p + 0,035v_c v_b \\
 & + 0,000v_c Q + 0,001v_c lt_0 + 0,005f_z a_e + 0,010f_z a_p - 0,075f_z v_b \\
 & - 0,008f_z Q + 0,002f_z lt_0 - 0,005a_e a_p + 0,004a_e v_b - 0,001a_e Q \\
 & + 0,005a_e lt_0 + 0,010a_p v_b + 0,003a_p Q - 0,005a_p lt_0
 \end{aligned}
 \tag{6}$$

Applying the concept of robust parameter design (RPD) through a combined arrangement, the optimization of the variables involved two functions, requiring the application of a dual method. In this study, the concept of mean squared error was used, which aims to jointly minimize the bias and variance of the factors. Therefore, the mean

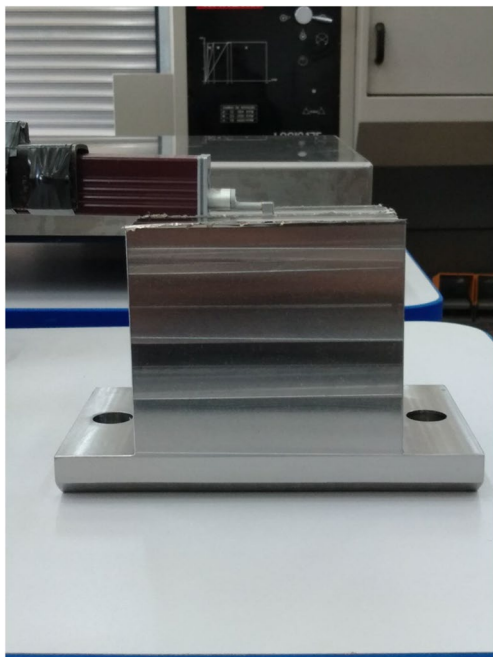


Fig. 5 Roughness measurement

$$\begin{aligned} \mu_{R_a} = & 0,426 + 0,16v_c + 0,056f_z + 0,002a_e + 0,008a_p + 0,045v_c^2 \\ & + 0,060f_z^2 + 0,067a_e^2 + 0,046a_p^2 + 0,088v_c f_z - 0,005v_c a_e \\ & - 0,011v_c a_p + 0,005f_z a_e + 0,010f_z a_p - 0,005a_e a_p \end{aligned} \quad (7)$$

$$\begin{aligned} Var[R_a] = & 0,025 + 0,010v_c - 0,020f_z + 0,002a_e + 0,002a_p \\ & + 0,001v_c^2 + 0,006f_z^2 + 0,0001a_e^2 + 0,0001a_p^2 \\ & - 0,005v_c f_z + 0,0003v_c a_e + 0,001v_c a_p - 0,001f_z a_e \\ & - 0,002f_z a_p + 0,0001a_e a_p \end{aligned} \quad (8)$$

Equations (7) and (8) are written solely in terms of the control variables, although the noise variables are tested at different levels during the experiments. The target (T) for the mean of R_a was estimated by performing individual optimization using the system of Eq. (9) through the Generalized Reduced Gradient (GRG) algorithm, available in the Excel® package. According to 21, GRG is one of the most robust and efficient methods for constrained nonlinear optimization.

$$\begin{aligned} & Min \hat{y}_i \\ \text{Subject : } & x^T x \leq \rho^2 = 2,83^2 \\ & \sigma_i^2 \geq 0,001 \end{aligned} \quad (9)$$

and variance models were developed as per Eqs. (7) and (8), based on the regression model presented in Eq. (6):

where \hat{y}_i represents the model for the mean of R_a and σ_i^2 represents the model for the variance of R_a . The constraint $x^T x \leq \rho^2$ represents the set of convex constraints of the experimental region, and for the CCD adopted in this work, $\rho = \alpha$, where α corresponds to the axial distance of the experimental arrangement. Once the target (T) for roughness

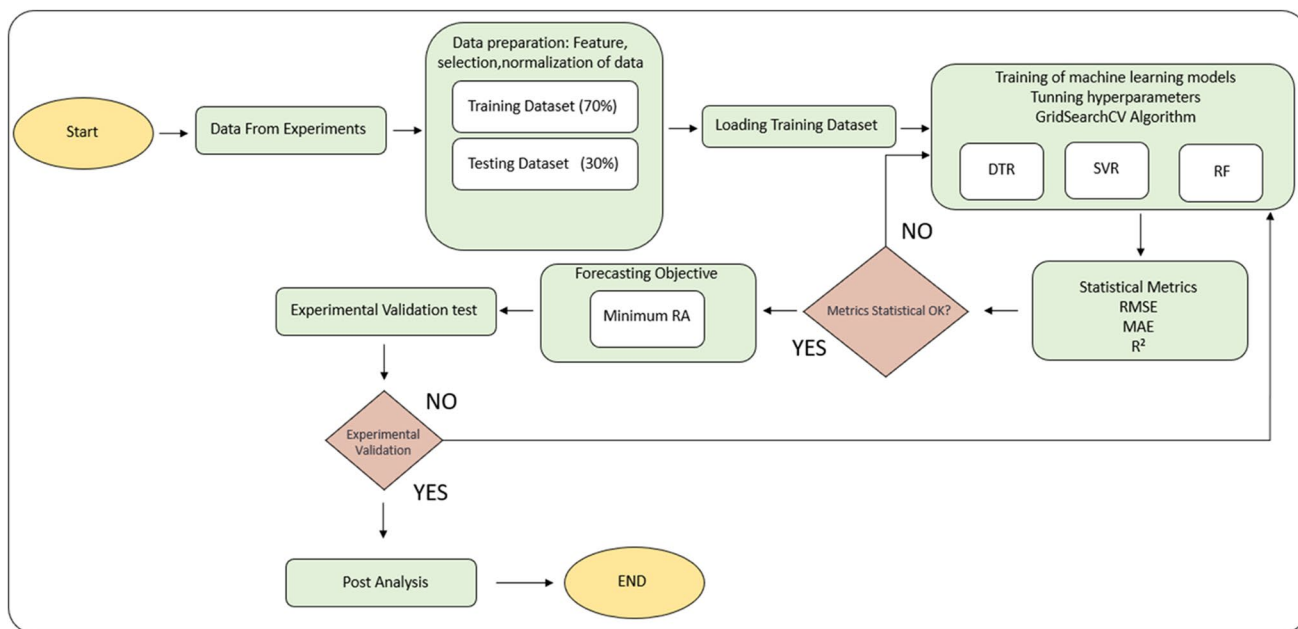
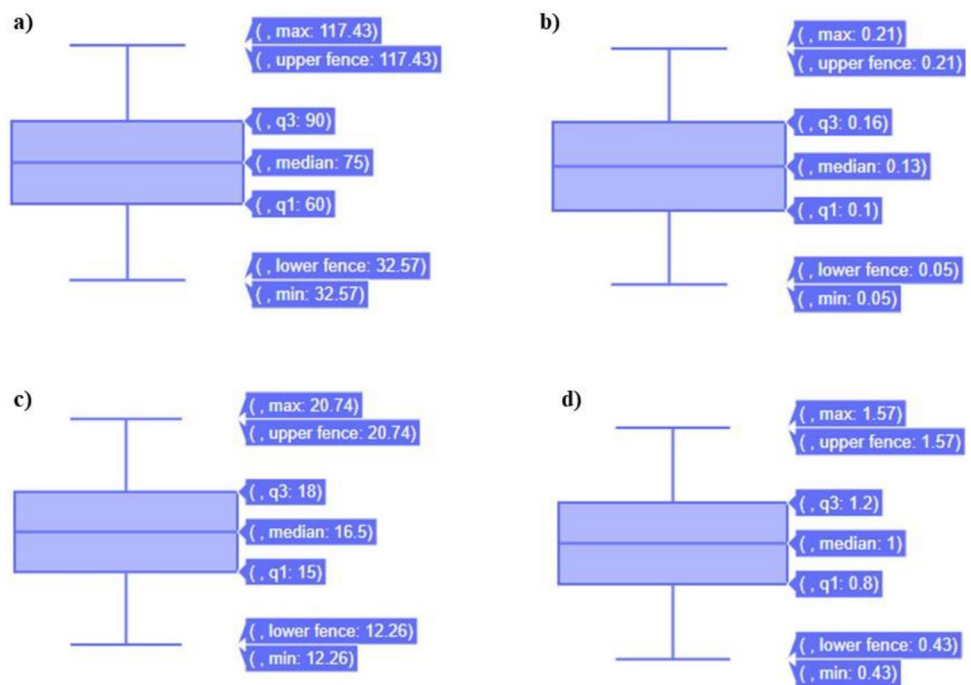


Fig. 6 General methodology used in this study

Table 5 Hyperparameters for machine learning models

Model	Hyperparameters tuned	Grid space	Results
SVR	Kernel	['rbf', 'sigmoid', 'poly']	'rbf'
	C	[80, 100, 120, 150]	100
	degree	[1e-6, 1e-5, 1e-4, 1e-3, 5e-3]	1e-6
	Gamma	[0.02, 0.04, 0.05, 0.06]	0.05
	cv	[6, 10, 15, 20]	6
DTR	max_depth	[None, 1, 2, 3,4, 5, 6, 7, 8, 9, 10]	'None'
	min_samples_split	[2, 3,4, 5, 6, 7, 8, 9, 10]	2
	min_samples_leaf	[1, 2, 3,4, 5]	1
	random_state	-	2
	cv	[6, 10, 15, 20]	6
RF	n_estimators	[10, 100, 200, 400, 500]	'rbf'
	criterion	['squared_error', 'absolute_error', 'friedman_mse', 'poisson']	'squared_error'
	max_depth	[None, 1, 2, 3,4, 5, 6, 7, 8, 9, 10]	'None'
	random_state	-	2
	cv	[6, 10, 15, 20]	6

Fig. 7 Analyze of outliers for **a** cutting speed, **b** feed per tooth, **c** cutting width, and **d** cutting depth



was defined, the mean squared error (*MSE*) equation was established as shown in Eq. (10).

$$MSE_{R_a} = (\mu_{R_a} - 0,399)^2 + \sigma^2_{R_a} \tag{10}$$

Finally, the optimization was performed using the GRG algorithm from the Solver® add-in program in the Excel® package, respecting the constraint of Eq. (10):

$$\begin{aligned} & \text{Min} && \bar{f}_1(x) \\ & \text{Subject :} && \bar{f}_1(x) - \bar{f}_2(x) + 2w - 1 = 0 \\ & && g_1(x) \leq 2,83^2 \\ & && 0 \leq w \leq 1 \end{aligned} \tag{11}$$

The optimal parameters for the control variables, as shown in Table 7, were then set in the CNC machining center command. A Taguchi L9 matrix was created for the

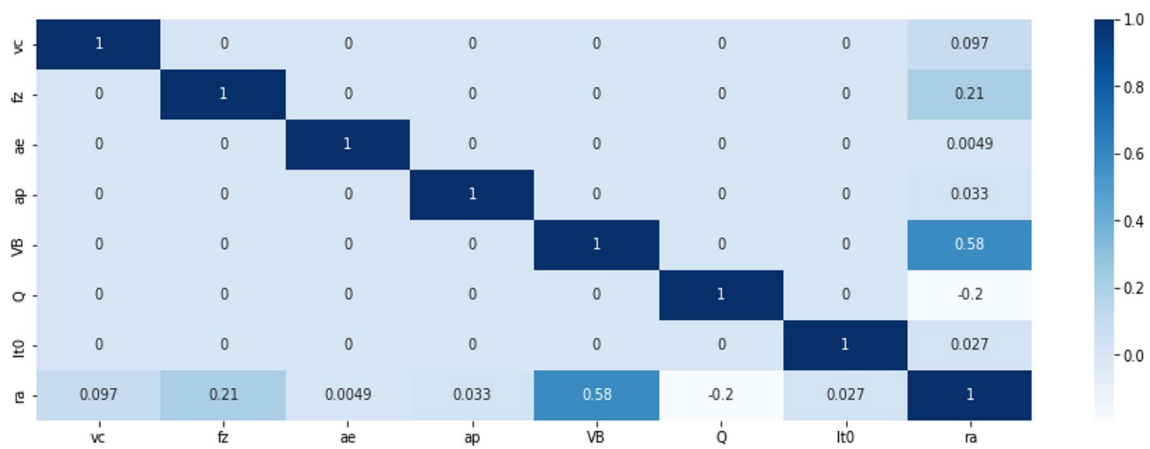


Fig. 8 Correlation analysis between controllable and uncontrollable factors

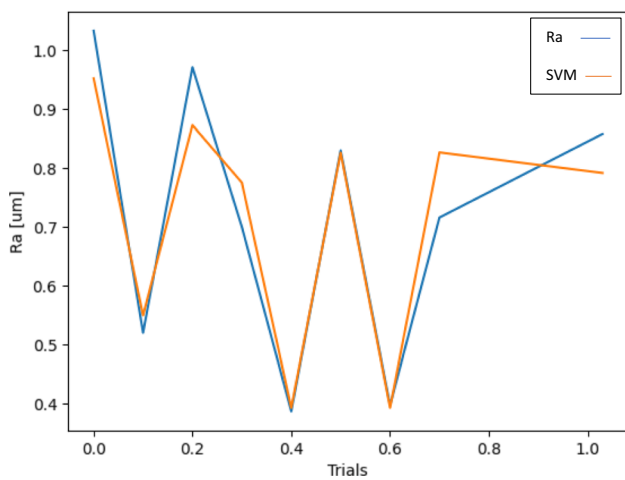


Fig. 9 Comparison of experimental measurement and prediction using $R_a \times$ SVM on the test dataset

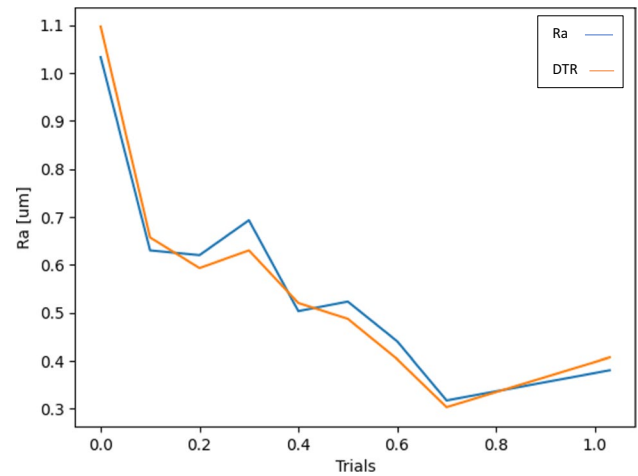


Fig. 10 Comparison of experimental measurement and prediction using $R_a \times$ DTR on the test dataset

Table 6 Accuracy statistics of ML regression models using testing datasets

Models	Testing dataset		
	RMSE	MAE	R^2 Training/testing
SVM	0.066	0.004	0.99/0.91
DTR	0.038	0.034	0.99/0.96
RF	0.031	0.001	0.91/0.97

levels of the noise variables. The results of the predictions with the experimental data are compared in Table 8.

The results indicate that the experimental data closely match the predicted values obtained through the confirmatory experiments. The largest percentage errors of R_a are 2.4% and 1.5%. Therefore, it is possible to assert that

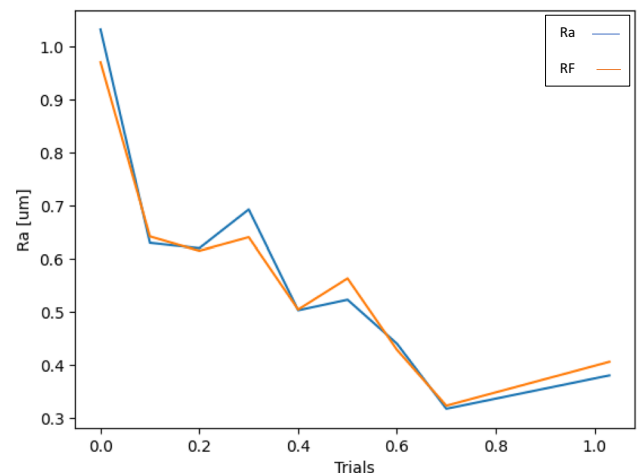


Fig. 11 Comparison of experimental measurement and prediction using $R_a \times$ RF on the test dataset

Table 7 Control parameters used in the confirmation experiments

Controllable variables			
v_c	f_z	a_e	a_p
[m/min]	[mm/Tooth]	[mm]	[mm]
69.8	0.17	15.34	0.74

machine learning techniques can be used to achieve the desired R_a in the top milling process of duplex stainless steel, as the differences found between the actual (experimental) value and the predictions are very small.

5 Conclusions

In this study, three machine learning models (SVM, DRT, and RF) were employed to predict surface roughness in the milling process of duplex stainless steel UNS S32205. The performance of these models was evaluated through experimental tests using three quality metrics: RMSE, MAE, and R^2 .

It was observed that the number of experiments was sufficient for training the models. Data normalization was performed, and although the presence of outliers was analyzed, no values were excluded on that basis. Validation tests were conducted to ensure the consistency of the models with reality, and the results demonstrated highly accurate predictions.

The analysis of the results revealed that the RMSE values of the models were quite close, suggesting that all models are significant. However, the RF model showed the best performance, with an RMSE value of 0.031, followed by the DTR (0.038) and SVM (0.066) models. Experimental verification indicated that the percentage errors between the actual (experimental) and predicted R_a values were 2.35%, 2.4%, and 1.15%, respectively. These results confirm the effectiveness of the machine learning models

in predicting surface roughness for the milling of duplex stainless steel.

In summary, this study compares the predictive accuracy of different machine learning models based on RMSE values and percentage errors. The RF model demonstrated superior precision, making it the most reliable for this application.

These results highlight the importance of considering noise during the training of machine learning models. This consideration allows artificial intelligence to better understand real-world processes, encompassing both controllable factors and the presence of noise, resulting in predictions that are more precise and aligned with reality.

Author contribution The work is divided into (i) literature review, (ii) background, (iii) statistical analysis, and (iv) writing and review. The work was divided as follows:

G.V.: literature review, background, statistical analysis, writing, and review.

M.B.: literature review, statistical analysis, writing, and review.

C.H.d.O.: background, statistical analysis, and review.

E.L.B.: background, statistical analysis, and review.

L.G.P.d.S.: background, statistical analysis, and review.

M.M.: Literature review, background, and review.

Funding I want to thank the Foundation for Research Support of the State of Minas Gerais (FAPEMIG) for the funding in projects APQ 02290–23 and APQ-03036–23.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Smola AJ, Barlett P, Schölkopf B, Schuurmans D (1999) Introduction to large margin classifiers. In A. J. Smola, P. Barlett, B. Schölkopf, and D. Schuurmans, editors, *Advances in Large Margin Classifiers*, pages 1–28. MIT Press, <https://doi.org/10.7551/mitpress/1113.001.0001>

Table 8 Validation of predicted results

EXP	v_b [mm]	Q [l/min]	l_{t0} [mm]	R_a (μm)	SVM	DTR	RF
1	0,00	0,00	30,00	0,52	0,53	0,50	0,56
2	0,00	0,20	40,00	0,57	0,55	0,55	0,54
3	0,00	20,00	50,00	0,57	0,55	0,54	0,57
4	0,15	0,00	40,00	0,57	0,54	0,56	0,60
5	0,15	0,20	50,00	0,56	0,53	0,57	0,55
6	0,15	20,00	30,00	0,56	0,55	0,55	0,52
7	0,30	0,00	50,00	0,54	0,52	0,6	0,53
8	0,30	0,20	30,00	0,54	0,6	0,66	0,6
9	0,30	20,00	40,00	0,57	0,61	0,58	0,58
Mean				0,55	0,56	0,57	0,57
Error					2,35%	2,4%	1,15%

2. Passerini Kernel A (2004) Methods, multiclass classification and applications to computational molecular biology. PhD thesis, Università Degli Studi di Firenze
3. Smola AJ, Schölkopf B (2002) Learning with Kernels. The MIT Press, Cambridge, MA
4. Azure JWA, Ayawah PEA, Kaba AGA, Kadingdi FA, Frimpong S (2021) Hydraulic shovel digging phase simulation and force prediction using machine learning techniques. *Mining Metall Explor* 38:2393–2404
5. Boukerche A, Zheng L (2020) Alfandi, Outlier detection: methods, models, and classification. *Computing ACM Sobreviver* 53:1–37
6. Bustillo A, Reis R, Machado AR, Pimenov DY (2022) Improving the accuracy of machine-learning models with data from machine test repetitions. *J Intell Manuf* 33:203–221
7. Cica D, Sredanovic B, Tesic S, Kramar D (2020) Predictive modeling of turning operations under different cooling/lubrication conditions for sustainable manufacturing using machine learning techniques. *Appl Comput Informar*
8. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining*, 2(2):1–43
9. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–297. <https://doi.org/10.1007/BF00994018>
10. Ding P, Huang X, Zhao C, Liu H, Zhang X (2023) Online monitoring model of micro-milling force incorporating tool wear prediction process. *Expert Syst Appl* 223:119886
11. Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V (1997) Support vector regression machines. *Adv Neural Inf Process Syst* 9:155–161
12. Eser E Aşkar Ayyıldız M, Ayyıldız and Kara F (2021) “Artificial intelligence-based surface roughness estimation modelling for milling of AA6061 alloy,” *Adv Mater Sci Eng* 2021:5576600. 10 pages
13. Anh-Tu Nguyen, Van-Hai Nguyen, Tien-Thinh Le, Nhu-Tung Nguyen, “Multiobjective optimization of surface roughness and tool wear in high-speed milling of AA6061 by machine learning and NSGA-II”, *Adv Mater Sci Eng* 2022:5406570, 21 pages. <https://doi.org/10.1155/2022/5406570>
14. Fernández-Delgado M, Cernadas E, Barro S, Amorin D (2014) Do we need hundreds of classifiers to solve real world classification problems. *J Mach Learn Res* 15:3133–3181
15. Fertig Weigold M, Chen Y (2022) Machine Learning based quality prediction for milling processes using internal machine tool data, *Advances in Industrial and Manufacturing Engineering*, 4:100074, ISSN 2666–9129. <https://doi.org/10.1016/j.aime.2022.100074>
16. Fuat Kara, Mustafa Karabatak, Mustafa Ayyıldız, Engin Nas, Effect of machinability, microstructure and hardness of deep cryogenic treatment in hard turning of AISI D2 steel with ceramic cutting, *Journal of Materials Research and Technology*, Volume 9, Issue 1, 2020, Pages 969–983, ISSN 2238–7854, <https://doi.org/10.1016/j.jmrt.2019.11.037>
17. Injadat, M., Moubayed, A., Nassif, A. B., & Shami, A. (2020). Systematic ensemble model selection approach for educational data mining. *Knowledge-Based Systems* 200:105992. <https://doi.org/10.1016/j.knsys.2020.105992>
18. Wu J, Chen X-Y, Zhang H, Xiong L-D, Lei H, Deng S-H (2019) Hyperparameter optimization for machine learning models based on Bayesian optimization. *J Electron Sci Technol* 17:26–40
19. Jumare AI et al (2018) Prediction model for single-point diamond tool-tip wear during machining of optical grade silicon. *Int J Adv Manuf Technol* 98(9):2519–2529
20. Jurkovic Z, Cukor G, Brezocnik M, Brajkovic T (2018) A comparison of machine learning methods for cutting parameters prediction in high-speed turning process. *J Intell Manuf* 29:1683–1693
21. KOKSOY O, (2006) Multiresponse robust design: mean square error (MSE) criterion. *Appl Mathemat Comput* 175:1716–1729
22. Kuntoğlu M, Aslan A, Pimenov DY, Usca ÜA, Salur E, Gupta MK et al (2021) A review of indirect tool condition monitoring systems and decision-making methods in turning: critical analysis and trends. *Sensors* 21(1):108. <https://doi.org/10.3390/s21010108>
23. Li K, Qiu C, Zhou X, Chen M, Lin Y, Jia X, Li B (2022) Modeling and tagging of time sequence signals in the milling process based on an improved hidden semi-Markov model. *Expert Syst Appl* 205:117758
24. Martinho RP, Silva FJG, Martins C, Lopes H (2019) Comparative study of PVD and CVD cutting tools in milling of duplex stainless steel. *Int J Adv Manuf Technol* 102(5–8):2423–2439
25. MONTGOMERY DC (2017) Designs and analysis of experiments. 9th. ed. USA: John Wiley & Sons
26. Nguyen NT, Tien DH, Tung NT, Luan ND (2021) Analysis of tool wear and surface roughness in high-speed milling process of aluminum alloy Al6061. *EUREKA: Phys Eng* (3), 71–84
27. Maimon OZ, Rokach L (2014) Data mining with decision trees: theory and applications. World Scientific, Singapore
28. Pimenov DY, Bustillo A, Mikolajczyk T (2018) Artificial intelligence for automatic prediction of required surface roughness by monitoring wear on face mill teeth. *J Intell Manuf* 29:1045–1061. <https://doi.org/10.1007/s10845-017-1381-8>
29. Ramírez-Gallego S, Krawczyk B, García S, Woźniak M, Herrera F (2017) A survey on data preprocessing for data stream mining: current status and future directions. *Neurocomputing* 239:39–57
30. Saravanamurugan S, Thiyagu S, Sakthivel NR, Nair BB (2017) Chatter prediction in boring process using machine learning technique. *Int J Manuf Res* 12(4):405–422
31. Shah M, Vakharia V, Chaudhari R et al (2022) Predicting tool wear in stainless steel face milling using a singular generative adversarial network and LSTM deep learning models. *Int J Adv Manuf Technol* 121, 723–736. <https://doi.org/10.1007/s00170-022-09356-0>
32. Speiser JL, Durkalsky VL, Lee WM (2015) Random forest classification of etiologies for an orphan disease. *Statustucs in Medicine* 34(5):887–899
33. Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE) arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7(3):1247–1250
34. Flores V, Keith B (2019) “Gradient boosted trees predictive models for surface roughness in high-speed milling in the steel and aluminum metalworking industry”, *Gradient Boosted Trees Predictive Models for Surface Roughness in High-Speed Milling in the Steel and Aluminum Metalworking Industry*
35. Wang R, Cheng MN, Loh YM, Wang C, Cheung CF (2022) Ensemble learning with a genetic algorithm for surface roughness prediction in multi-jet polishing. *Expert Syst Appl* 207:118024
36. Xu X, Zhang Y, Li Y et al (2022) Cutting forces in functionally graded material milling using machine learning. *Adv em Comp Int* 2:25. <https://doi.org/10.1007/s43674-022-00036-w>
37. Zhang L, Zhong ZQ, Qiu LC, Shi HD, Layyous A, Liu SP (2019) Coated cemented carbide tool life extension accompanied by comb cracks: the milling case of 316L stainless steel. *Wear*, 418–419, n. November 2018, p. 133–139

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.