Advances in Educational Technology and Psychology (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/aetp.2025.090510
ISSN 2371-9400 Vol. 9 Num. 5

Comparative Analysis of Prediction Algorithms for Surface Roughness in AI-oriented Courses

Chunhua Feng[1,a,*], Ming Li[1,a]

[1]*School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China*

[a]*fengchunhua333@163.com*, [b]*liming5976@163.com*

**Corresponding author: Chunhua Feng*

Keywords: AI-oriented courses, surface roughness prediction, ensemble learning

Abstract: The integration of artificial intelligence (AI) into mechanical engineering curricula has garnered significant interest. A critical challenge lies in effectively incorporating AI technologies into foundational courses, which is essential for advancing the practical implementation of AI. In the field of machining, surface roughness serves as a crucial parameter for assessing the quality of manufactured components, influencing properties such as wear resistance, fatigue strength, and dimensional accuracy. Conventional empirical approaches struggle to accurately model the complex nonlinear dynamics involved in machining processes. As a result, data-driven intelligent prediction methods have emerged as a prominent area of research in this domain. This paper aims to investigate the effectiveness of machine learning algorithms in predicting surface roughness. Prediction models are developed using the Exponential Function (EF), Ridge Regression (RR), Gradient Boosting Regression (GBR), eXtreme Gradient Boosting (XGBoost), and Ensemble Learning based on a Genetic Algorithm (ELGA). Through the training and testing based on experimental data, the prediction accuracy and stability of various algorithms were evaluated and compared. The results demonstrate that the ELGA algorithm proposed in this paper further reduces prediction error by employing a unique global optimization strategy. Specifically, the root mean square error (RMSE) is 0.035, the mean absolute percentage error (MAPE) is 0.027, and the coefficient of determination (R^2) reaches 0.955. Overall, ELGA outperforms individual machine learning models, significantly enhancing both the accuracy and robustness of predictive performance. This advancement provides an effective solution and a valuable reference for algorithm selection in high-precision surface roughness prediction.

1. Introduction

AI+ technology has established a pivotal role in the teaching of mechanical engineering courses. To more effectively integrate AI technology into specialized curricula and advance the integration of AI+ with industry and education, this paper uses machining as a case study to explore the application scenarios of various AI technologies. As widely recognized, the rapid advancement of industries such as aerospace, automotive manufacturing, semiconductors, and microelectronics has significantly raised the demands for both the quantity and quality of ultra-precision machined components. In particular, surface quality directly influences the performance and service life of mechanical workpieces, with surface roughness serving as the key metric for evaluating surface quality [1]. Therefore, it is particularly important to construct an accurate surface roughness prediction model. Through this model, accurate prediction of surface roughness can be achieved, which is of crucial significance for precise control of machining quality and significant reduction of the rejection rate. The accuracy of surface roughness model prediction is affected by various factors, including geometric parameters, cutting parameters, machine tool accuracy, and tool parameters [2]. In order to obtain the best surface quality of the processed parts, the optimal process parameters were previously found through experiments. This method not only requires a large amount of manpower and material resources, but also usually has a long cycle, high costs, and poor portability. It has become a significant bottleneck restricting high-quality and high-efficiency production [3].

With the rapid development of artificial intelligence technology, especially the maturity of machine learning and deep learning algorithms, applying machine learning algorithms to the prediction of surface roughness has increasingly become a research hotspot in academia and industry. The core advantage of using machine learning algorithms to predict surface roughness lies in their ability to bypass complex physical processes that are difficult to accurately model and directly learn patterns from historical processing data [4]. This method analyzes cutting parameters and automatically learns the complex non-linear mapping relationship between them and surface roughness. This data-based method not only significantly reduces the dependence on prior physical knowledge, but also features high prediction accuracy and strong generalization ability. It can effectively reduce the cost and time of repeated trial cutting, providing key technical support for the online optimization and intelligent control of machining parameters and promoting the transformation of intelligent manufacturing.

By applying machine learning algorithms, a high-precision surface roughness prediction model can be constructed. Among them, machine learning methods can be modeled based on common algorithms. In addition, this paper proposes an ensemble learning method based on the genetic algorithm (ELGA). By combining the advantages of each individual prediction model, overall optimization is achieved, thus constructing a more accurate surface roughness prediction model.

2. Comparative Algorithms for Predictive Modelling of Surface Roughness

2.1 Surface Roughness Prediction Process

In this study, surface roughness data samples under different technological parameters were collected. To construct a prediction model, the overall dataset was divided into a training set and a test set according to a specific ratio. The training set was used for model training and learning, while the test set was used to evaluate the generalization ability of the model. Subsequently, various machine learning algorithms (such as the GBR model, Ridge regression model, etc.) were applied to establish a non-linear mapping model between surface roughness and cutting parameters. To quantitatively evaluate the prediction accuracy and reliability of each model, root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) were used as the core evaluation indicators. Flow chart for surface roughness prediction is shown in Fig. 1.

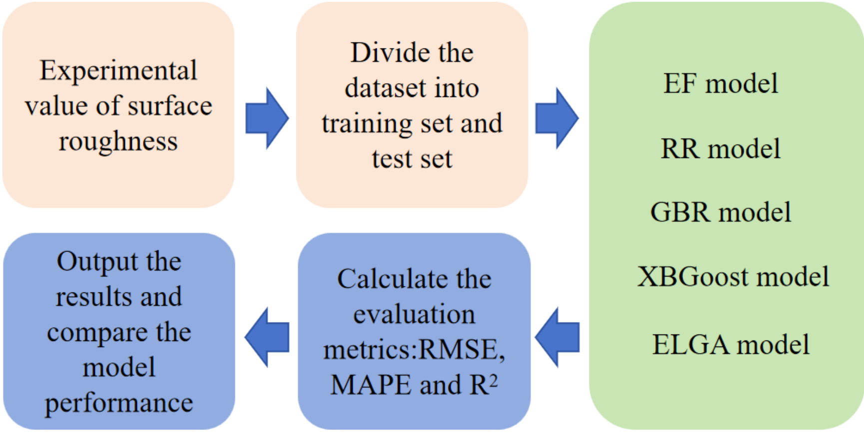

Fig.1 Flow chart for surface roughness prediction

2.2 Acquisition of Surface Roughness Data

The surface roughness data in this paper are derived from the research of Feng et al. [5]. In the experiment, the Taguchi method was used to design the experiment. Taguchi method is a widely used experimental design method, aiming to achieve the best experimental results by optimizing the combination of multiple influencing factors. The core of this method lies in selecting an appropriate number of factors and levels and arranging them orthogonally to ensure that the experimental results can accurately reflect the influence of each factor on the outcome. Four key cutting parameters were selected as control factors, namely spindle speed (n), feed rate (vf), cutting depth (ap), and cutting width (ae). For the orthogonal experimental design, it is necessary to determine different levels of each cutting parameter and ensure that the differences between levels are equal. Through this design method, errors and interferences can be minimized to obtain more accurate and reliable experimental results. Finally, the orthogonal experimental scheme with four factors and three levels described in Table 1 was adopted, and a total of 27 groups of experiments were designed.

Table 1 Orthogonal experimental design for milling surface roughness

Level	n (r/min)	v_f (mm/min)	a_p (mm)	a_e (mm)
1	8000	240	1	4
2	9000	280	1.5	6
3	10000	320	2	8

2.3 The Principles and Outcomes of Various AI Prediction Algorithms

The exponential function model, an empirical model, accurately describes in exponential form the law of how cutting parameters affect surface roughness. Based on experimental data and empirical knowledge, this model can systematically analyze the influence of cutting parameters and successfully establish the functional relationship between surface roughness and cutting parameters. The specific calculation method of the exponential function for surface roughness is as follows:

$$R_a = r_1 n^{r_2} v_f^{r_3} a_p^{r_4} a_e^{r_5} \tag{1}$$

where r1, r2, r3, r4, r5 represent empirical constants. Different materials, processing methods and cutting tools may all result in different values of these parameters. These parameters were all obtained through experiments.

Ridge Regression is an enhanced version of linear regression, specifically developed to address the issue of multicollinearity, which occurs when predictor variables are highly correlated. It falls under the category of regression modeling techniques. In situations where multicollinearity causes the data matrix to become rank-deficient, the matrix inversion required by the Ordinary Least Squares (OLS) approach becomes unfeasible, making direct parameter estimation impossible. To overcome this limitation, Ridge Regression incorporates a regularization component, often referred to as a penalty term into the loss function. This addition helps to stabilize the estimation of model parameters. As a result, Ridge Regression is particularly effective in managing strong interdependencies among variables, making it well-suited for applications such as predicting surface roughness. The inclusion of the regularization term not only improves the robustness of the model but also contributes to reducing its complexity, thereby mitigating the risk of overfitting.

Gradient Boosting Regression (GBR) is a powerful ensemble learning technique known for its efficiency. The core idea behind GBR is to iteratively train a series of weak learners, with each subsequent model aiming to correct the errors made by its predecessor. In the end, the predictions from all these weak models are combined through a weighted summation, resulting in a stronger, more accurate predictive model. This approach follows the negative gradient of the loss function to guide the learning process, enabling the model to capture intricate non-linear patterns. In surface roughness prediction tasks, there often exists a complex and highly non-linear relationship between machining parameters—such as cutting speed, feed rate, and depth of cut—and the resulting surface quality. GBR is capable of automatically detecting and modeling the non-linear dependencies and interactions among these input variables and the output roughness. Additionally, the algorithm demonstrates a certain degree of robustness against outliers in the dataset. As a result, GBR can serve as an effective foundation for constructing a high-precision predictive system that maps machining parameters to surface roughness.

XGBoost, short for eXtreme Gradient Boosting, is a machine learning technique that belongs to the family of ensemble methods based on gradient boosting. It works by sequentially building multiple decision trees, with each subsequent tree aiming to correct the residuals or errors of the previous ones. The final prediction is obtained by aggregating the outputs of all individual trees, which enhances the overall predictive accuracy and robustness of the model. During each iteration, the algorithm focuses on minimizing the difference between the predicted and actual values, progressively refining the model's performance. One of the key strengths of XGBoost is its efficiency in combining numerous weak models while applying regularization techniques to control overfitting. This enables the model to effectively capture intricate non-linear patterns within the data as well as the interactions among different features. As a result, XGBoost has proven to be highly effective in handling regression tasks, such as predicting surface roughness.

The ELGA (Ensemble Learning with Genetic Algorithm) model proposed in this paper integrates the EF model, RR model, and GBR model based on the genetic algorithm for regression prediction tasks. First, an EF model is constructed based on 22 groups of surface roughness experimental data, and the remaining 5 groups of data are used as a test set to evaluate the generalization ability of the model. Meanwhile, RR and GBR prediction models are respectively established based on the same training set. Under the framework of the genetic algorithm, the Particle Swarm Optimization (PSO) algorithm is introduced as the search strategy. The mean absolute percentage error (MAPE) between the predicted value and the experimental value of surface roughness is used as the fitness function. Through iterative optimization of the population, the optimal weight combination of each sub-model is determined. The ELGA model effectively combines the prior knowledge of the EF model and the regression advantages of the RR and GBR models. Through ensemble learning, the comprehensiveness and accuracy of surface roughness prediction are significantly improved.

By comparing the performance of XGBoost with EF, RR, GBR, and ELGA in the integrated model on various test sets, the fitting and prediction abilities of each model for surface roughness data can be comprehensively evaluated, and then the most suitable modeling method for this task can be screened out. In practical applications, further verification and tuning of the model are also required according to specific needs to achieve the optimal prediction effect. In this paper, the surface roughness under various cutting parameters is predicted to clearly compare the performance of different models on the test set. The detailed results are shown in Table 2. At the same time, the root mean square error (RMSE), MAPE, and coefficient of determination (R^2) are used as evaluation indicators to quantitatively evaluate the performance of each model. Finally, a comprehensive performance comparison chart is drawn, as shown in Fig. 2.

Table 2 The prediction results (in μm) of different surface roughness prediction models

No.	Experimental results	XGBoost	EF	RR	GBR	ELGA
1	0.897	0.826	0.838	0.873	0.851	0.868
2	0.915	0.999	0.979	1.010	0.944	0.963
3	1.328	1.287	1.319	1.321	1.299	1.332
4	1.197	1.250	1.176	1.192	1.211	1.221
5	1.092	1.021	1.085	1.078	0.989	1.090

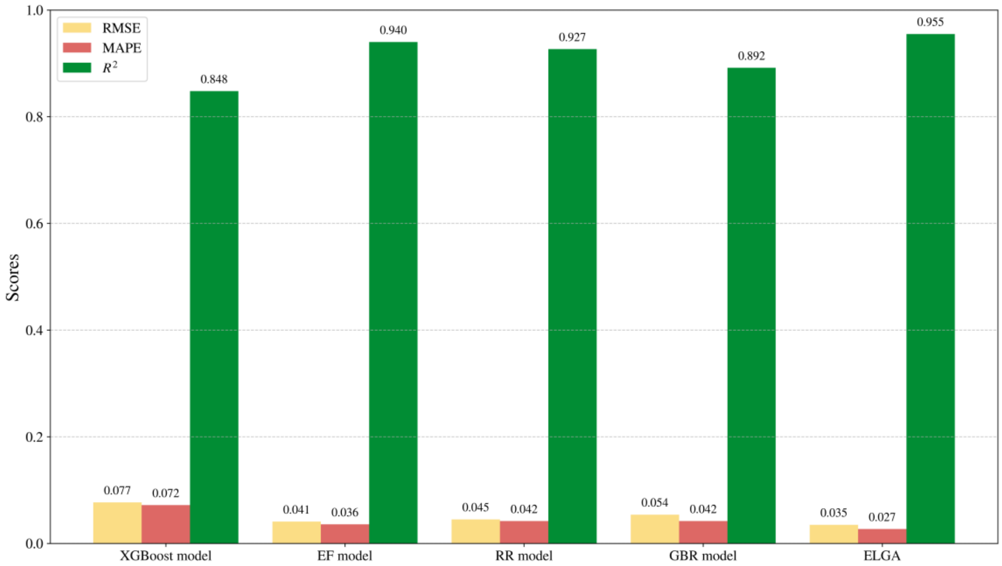

Fig.2 Comparison for different surface roughness prediction models

As can be seen from Fig. 2, the EF model, RR model, and GBR model exhibit better modeling performance on the test set data than other machine learning models. This indicates that these models can better capture the non-linear relationship between cutting parameters and surface roughness and possess stronger generalization ability. The ELGA model combines the advantages of these regression models. By using the genetic algorithm to obtain the optimal weights of each model, it not only utilizes the prediction capabilities of individual models but also enhances the overall performance of the model. Therefore, the performance of the ELGA model is significantly higher than that of a single model among the regression models. The prediction results of ELGA demonstrate an RMSE of 0.035 and a MAPE of 0.027, indicating that the model achieves a high level of predictive accuracy. These results also suggest that the ELGA model can effectively predict surface roughness values, thereby offering valuable support for optimizing cutting parameters. Furthermore, the R^2 is 0.955, which reflects a strong correlation between cutting parameters and surface roughness. This finding confirms that the surface roughness prediction model established using ELGA is highly reliable and can serve as an effective tool for surface quality control in practical production settings.

3. Conclusion

This study presents a systematic investigation into the effectiveness of various machine learning algorithms in predicting surface roughness. A particular focus is placed on the comparative analysis of predictive performance across several models: the EF model, RR, GBR, XGBoost, and a newly proposed ELGA. Using experimental data for both training and testing, and employing RMSE, MAPE, and R^2 as primary evaluation metrics, the accuracy and reliability of each model are rigorously evaluated. The findings of this research demonstrate that:

(1) Single models such as EF, RR, and GBR demonstrate strong capabilities in capturing the complex nonlinear relationships between cutting parameters and surface roughness. However, their predictive performance still exhibits certain limitations. To address this, the ELGA model proposed in this study integrates EF, RR, and GBR by optimizing their respective weights using a genetic algorithm. This approach fully leverages the strengths of each individual sub-model, thereby enhancing both prediction accuracy and model robustness. Experimental results show that ELGA achieves an RMSE of 0.035, a MAPE of 0.027, and an R^2 value of 0.955 on the test set, significantly outperforming other comparative models.

(2) The superior performance of the ELGA model demonstrates its ability to effectively integrate information from multiple models. Through a global optimization strategy, the model enhances prediction consistency and reliability, offering an efficient solution for high-precision surface roughness prediction. This study not only confirms the applicability of machine learning in predicting surface roughness but also provides solid theoretical support and methodological guidance for optimizing processing parameters and quality control in the context of intelligent manufacturing.

Through the introduction of various intelligent algorithms applied in surface roughness prediction, mechanical engineering students can deeply understand the application process of AI technology in specific industries.

Acknowledgements

This research is funded by the Ministry of Education industry-university cooperative education project No. 231104429231321.

References

[1] Yang H, Zheng H, Zhang T. A review of artificial intelligent methods for machined surface roughness prediction. Tribology International, 2024, 199:109935.
[2] Ulkir O, Akgun G. Predicting and optimising the surface roughness of additive manufactured parts using an artificial neural network model and genetic algorithm. Science and Technology of Welding and Joining, 2023, 28(7): 548-557.
[3] Abellán-Nebot J V, Vila Pastor C, Siller H R. A review of the factors influencing surface roughness in machining and their impact on sustainability. Sustainability, 2024, 16(5): 1917.
[4] Wu P, Dai H, Li Y, et al. A physics-informed machine learning model for surface roughness prediction in milling operations. The International Journal of Advanced Manufacturing Technology, 2022, 123(11): 4065-4076.
[5] Feng C, Qin E, Li W, et al. Energy consumption and surface roughness modelling for multi-objective optimisation of machining processes. The International Journal of Advanced Manufacturing Technology, 2024, 133(9): 5065-5087.