Energy Consumption Prediction of CNC Milling based on Random Forest

Bingying Wu*

Department of Industrial Engineering, Shandong University of Science and Technology, Qingdao 266590, China

*Corresponding author: 202283010075@sdust.edu.cn

Abstract

Mechanical processing systems are mainly based on machine tools. In industries such as production and manufacturing, the use of machine tools is huge. Therefore, reducing machine tool energy consumption and improving energy efficiency are of great significance. Accurately calculating the energy consumption during the processing is the primary task. In order to accurately predict the energy consumption of CNC milling, this paper establishes a power model and a time model for CNC milling. The power and time data obtained from the experiment in this article are labeled data. Since the features are both continuous variables, this article uses random forests to regress and predict the energy consumption of CNC milling.

Keywords

Random Forest; Energy Consumption; CNC Milling; Prediction.

1. Introduction

The development of industrial manufacturing industry is so rapid, and machine tools are essential tools in industrial manufacturing. Throughout their entire life cycle of production, their operation and use have the greatest impact on the environment, accounting for 94% -99% of the total impact. Among them, the power consumption of machine tools and the greenhouse gases such as carbon dioxide and water vapor generated by fuel consumption account for a relatively large proportion. Among the various functions of the machine tool, the spindle system, cooling unit, coolant, servo drive system, and hydraulic system are the five main subsystems that consume 88% of the electricity, while only 25% of the electricity consumed is directly used for cutting processing. From this, it can be seen that machine tools consume a huge amount of energy during mechanical processing, but their energy efficiency is not high, and there is great potential for energy conservation and emission reduction. Predicting energy consumption during mechanical processing is a prerequisite and a key focus for energy conservation and emission reduction of machine tools. Therefore, it is necessary to explore more accurate predictions of mechanical processing energy consumption.

2. Literature References

Some scholars have conducted research on energy consumption prediction from a mechanistic perspective. Diaz [1] and Kara [2] respectively established specific energy consumption models under constant and non constant material removal rates to predict the energy consumed in producing parts. Li [3] conducted extensive data analysis and research on cutting experiments of different materials, and summarized an empirical formula for the specific energy consumption of machine tool cutting under different material removal rates.

Some scholars have conducted extensive research on using a single intelligent algorithm to model and predict machine tool machining energy consumption. Kant [4] developed an artificial neural network model, which uses data from 27 experiments to predict the cutting energy of carbon steel during the machining process through process parameters such as spindle speed, cutting depth, cutting width, and feed rate. He [5] proposed a data-driven energy prediction method based on convolutional neural networks, and conducted experiments on milling and grinding machines to establish a supervised prediction model between extracted features and machine energy consumption. Awan [6] proposed a machine learning based prediction method for cutting and grinding specific energy consumption, using three supervised learning techniques: Gaussian process regression, regression tree, and artificial neural network for energy consumption prediction. Brillinger [7] used decision trees to predict the energy consumption of CNC machine tools based on actual production data.

The mechanism of the machining process of CNC machine tools is complex, and each process requires the establishment of complex mathematical expressions. The mathematical methods used are diverse, and the universality for different machine tools is poor. When applied to different machine tools, corresponding models may need to be re established. But using data-driven modeling only requires establishing intelligent input-output relationships for modeling

3. Method

3.1 Random Forest

Random Forest uses a technique called Bagging to generate multiple training datasets. Assuming the original training set is $\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$, multiple subsets of size m are generated through random sampling with replacement $\{D_1, D_2, \cdots, D_B\}$, where B is the number of decision trees, each subset is used to train a decision tree, each tree is trained independently, and because the dataset is randomly sampled, each tree is different. This method effectively reduces the variance of the model.

Each decision tree was constructed using random feature selection. For each node split, randomly select a subset of features and then choose the best feature for splitting. This method improves the integration effect by reducing the correlation between trees. Assuming the feature space is $\mathcal{X}_{sub} \subseteq \mathcal{X}$, at each node, randomly select a subset of features for splitting. This process can be expressed as:

$$j^* = \arg\min_{j \in \mathcal{X}} \min_s \left[\text{Impurity}\left(D_{\text{left}}\right) + \text{Impurity}\left(D_{\text{right}}\right) \right] \quad (1)$$

For each subset, the decision tree recursively splits until it can no longer split. Predict the new data point x through each tree and determine the final prediction result based on the average value (regression).

3.2 Model based on Energy Consumption of Random Forest CNC Milling

Divide the data into a training set and a testing set, with a ratio of 7:3. The features include spindle speed N, feed rate F_z, cutting width a_e, cutting depth a_p, and the target variable is the total milling energy consumption, which includes the process from machine start-up to complete milling standby. Input the data into the default model of the random forest and optimize the hyperparameters, taking into account max_depth, min_samples_leaf, n_estimators. Max_depth is used to control the maximum depth of each tree. Larger values can capture complex patterns in the data, but if the value is set too high, the tree may become too complex, suitable for training data, and have poor generalization ability to new data. Generally, adjustments are needed to balance the bias and variance of the model. Min_Samples_leaf controls the minimum number of samples required in the leaf nodes. Larger values help smooth the decision boundaries of the model and reduce overfitting to the training data. However, setting it too high can bias the model towards underfitting, while setting it too low can lead to overfitting. N_estimators determine the number of trees in a random forest, and increasing

ISSN: 2414-1895
DOI: 10.6919/ICJE.202504_11(4).0017

the value usually improves the performance of the model because more trees can reduce the variance of the model and enhance prediction stability. Although increasing the number of trees can improve accuracy, it also increases computational costs and training time. Usually, it is necessary to find the appropriate quantity through cross validation.

This article compromises by choosing 5-fold cross validation , which randomly divides the dataset into 5 subsets, usually with each subset being nearly equal in size. In each iteration, one subset is selected as the validation set, and the remaining four subsets are merged as the training set. Train the model and evaluate its performance on the validation set. This process is repeated 5 times, each time selecting a different subset as the validation set and the other 4 subsets as the training set. At each compromise, an evaluation metric is obtained to calculate the average of the evaluation results, which serves as the final model evaluation metric for cross validation. Based on this cross validation metric, the optimal hyperparameters are obtained.

Table 1. Model hyperparameter values

Hyperparameter name	Hyperparameter range	Optimal hyperparameters
min_samples_leaf	(1,2,3)	2
n_estimators	(200,300,400)	300
max_depth	(1,2,3)	3

3.3 Forecast Results

Based on the optimal parameters, establish the final RF power and time model, and evaluate the model using evaluation metrics such as R^2, MAE, RMSE, MAPE, etc. The results of the training and testing sets are shown in Table 2.

Table 2. Evaluation metrics for training and testing sets

Evaluation metrics	Training set	Testing set
R^2	0.993	0.925
MAE	2307.14	2861.94
RMSE	3071.05	3834.78

The R^2 value of the training set is 0.993, and the R^2 value of the testing set is 0.925. This indicates that the model can explain 99.3% of the variance on the training set and 92.5% of the variance on the test set, indicating that the model has a good fitting effect on the data and strong generalization ability. The MAE of the training set is 2307.14, and the MAE of the testing set is 2861.94. The MAE of the test set is slightly higher than that of the training set, indicating an increase in the model's prediction error on the test set, but the overall error is still within an acceptable range. The RMSE of the training set is 3071.05, and the RMSE of the testing set is 3834.78. The addition of RMSE on the test set further validated the performance difference between the training and testing sets of the model, but the difference was not significant, indicating good stability of the model.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0017

4. Summary

This article focuses on the energy consumption prediction problem in CNC milling process, and establishes a power model and time model based on random forest, aiming to improve machine tool energy efficiency and reduce energy consumption. Through the analysis of experimental data, this article draws the following conclusions:

Firstly, the random forest model exhibits high accuracy and stability in predicting energy consumption in CNC milling. The R^2 value of the training set is 0.993, and the R^2 value of the test set is 0.925, indicating that the model can explain the variance in the data well and has strong generalization ability. Although the MAE (2861.94) and RMSE (3834.78) of the test set are slightly higher than those of the training set (MAE of 2307.14, RMSE of 3071.05), the overall error is still within an acceptable range, indicating that the predictive performance of the model on the test set is still reliable.

Secondly, the hyperparameters of the model were optimized through 5-fold cross validation, and the optimal hyperparameter combination (min_stamples_leaf =2, n_estimators =300, max_depth =3) was determined, effectively balancing the bias and variance of the model and avoiding overfitting and underfitting problems.

Finally, this study provides an effective data-driven method for predicting energy consumption in CNC milling. Compared to traditional mechanism modeling methods, the random forest model has stronger universality and flexibility, and can adapt to the needs of different machine tools and processing conditions. Future research can further explore other machine learning algorithms or combine multiple methods to further improve the accuracy and applicability of energy consumption prediction.

References

[1] Diaz N. Process Parameter Selection for Energy Consumption Reduction in Machining. M.S. Project Report from the Dept of Mech Eng at University of California, Berkeley 2010.

[2] Kara S, Li W. Unit Process Energy Consumption Models for Material Removal Process. Annals of the CIRP 2011; 60:1:37-40.

[3] Li W, Kara S. An empirical model for predicting energy consumption of manufacturing processes: a case of turning process. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture. 2011;225(9):1636-1646.

[4] Kant G, Sangwan K S. Predictive modelling for energy consumption in machining using artificial neural network. Procedia Cirp, 2015, 37: 205-210.

[5] He Y, Wu P, Li Y, et al. A generic energy prediction model of machine tools using deep learning algorithms. Applied Energy, 2020, 275: 115402.

[6] Awan M R, González Rojas H A, Hameed S, et al. Machine learning-based prediction of specific energy consumption for cut-off grinding. Sensors, 2022, 22(19): 7152.

[7] Brillinger M, Wuwer M, Hadi M A, et al. Energy prediction for CNC machining with machine learning[J]. CIRP Journal of Manufacturing Science and Technology, 2021, 35: 715-723.