Revista de Chimie
https://revistadechimie.ro
https://doi.org/10.37358/Rev.Chim.1949

Analysis of Octane Retention Prediction Model for Catalytic Cracked Gasoline Based on Ridge Regression Model and Gradient Descent Optimization

FENG LYU[1,2], XIAOJUN YANG[1,2*], LONG LYU[3]
[1]Wuhan Textile University, School of Environmental Engineering, No.1, Sunshine Avenue, Jiangxia District, Wuhan City, Hubei Province, China
[2]Engineering Research Center for Clean Production of Textile Dyeing and Printing, Ministry of Education, No.1, Sunshine Avenue, Jiangxia District, Wuhan City, Hubei Province, China
[3]Central South University, School of Business, No.932, South Lushan Road, Yuelu District, Changsha City, Hunan Province, China

Abstract: *On the basis of the given material, in order to increase the RON retention of the catalytic cracking unit, the prediction model of gasoline octane retention and the best operation variable inversion model were established based on the Ridge regression model and Gradient descent method. First, based on the Ridge regression model, the leave-one method is used to obtain the relative importance of the operational variables, and select the most important variables, so as to reduce the characteristic dimension of the model; Then, the RON retention prediction model is trained based on the Ridge regression model; Finally, based on the trained Ridge regression model and its weight parameters, the optimal operating variables were optimized separately using the gradient when the operation variable has a range or no range of value. The experimental results show that when 146 are selected from 361 operating variables, the model loss value stabilizes; when α is 0.6, the test set R^2 is 0.9882, test set MSE is 0.0193, and the comprehensive performance is better than the random forest, support vector machine model; When the operation variable has two categories of value range and no value range, 2,000 times, the best inversion value of the operation variable makes the RON retention prediction value of the test sample similar to the expected value, and the MAE drops to 2.89999×10^{-3} and 7.62939×10^{-6}, respectively. In conclusion, the RON retention prediction model proposed in this study has good results, and the best operating variable can be reversed, based on the given material parameters, making the optimal RON retention quantity.*

Keywords*: RON retention prediction, Ridge regression model, Leave-One-Out, Gradient descent algorithm, inversion, machine learning*

1. Introduction

In China, there were 297 million vehicles owned by September 2021, gasoline vehicles account for the largest proportion among them, that is, the number of gasoline vehicles in China cannot be ignored, resulting in the use of gasoline and the research work cannot be stopped [1]. More than 70% of the gasoline used in China is produced by catalytic cracking, [2-4] and the process will also produce 95% of the sulfur and olefin, which will inevitably produce a large amount of air pollution after the gasoline is used. Therefore, in order to reduce the proportion of harmful substances (such as sulfur element and olefin) or the material content of harmful gas, gasoline merchants must make further refining treatment for catalytic cracking gasoline to meet the quality requirements of gasoline. In the refining process, it is found that the Research Octane Number (RON) after refining is always less than the RON phenomenon before refining, [5, 6] but the RON is one of the most important measures to evaluate the quality of gasoline. When the RON is increased, the car can have a stronger anti-explosive performance, and indirectly reduce a small amount of fuel consumption, at the same time, protect the engine. Each increase of RON can reduce fuel consumption by up to 1.4%, and the annual total oil saving of global gasoline vehicles can also reach amazing values. Therefore, improving RON also indirectly contributes to carbon

**email: yangdavis@vip.163.com*

peak and carbon neutrality, which is of great significance to fuel resource conservation and environ – mental protection [7].

Therefore, now the oil refining enterprises focuses on exploring the chemical industry model, which can be both in the catalytic cracking process, the process and technology of gasoline catalysis and hydrogenation conducted by oil refining enterprises, to ensure that the harmful substances in gasoline are lower than the highest concentration allowed by the government; to improve the accuracy of predicting RON, operational parameter scheme, and the RON content in gasoline [8-11]. Therefore, the model can also reduce the pollution for the society and improve the economic benefits for enterprises. Because the oil production process involves more than 300 related parameters, it is complicated to judge whether each operation parameter is beneficial or affects the oil production; to judge its influence mode and influence degree; to forecast the optimal matching parameter value and the highest possible the retained RON. It is necessary to find the mathematical expression that can reflect the operation parameters and retained RON as true as possible through the machine learning model and programming code. Traditional RON prediction methods not only consider relatively few variables, but also lack the process variable analysis, and do not consider the correlation between the variables [4].

Han Qingjue and Zou Min et al. used the grey association model to screen the factors mainly affecting the RON loss from the processed data, and used them as input variables in the model [12] and then established a prediction model based on the BP neural network. Finally, more than 85% of the absolute error was less than 0.2 units. Zhao Lin and Ling Xi et al. proposed an adaptive variable-weighted RON prediction method, [4] which captures the correlation between variable data by using a new variable weighting module. The variable weight is automatically generated according to the importance of RON. In order to analyse the influence of sulfur content on RON, the predicted value of sulfur content and RON are output together with the adaptive weighted variable, and show high prediction accuracy and model performance. Jiang Wei et al. used the RFR model to predict the loss of RON of catalytic cracked gasoline, [13] and used the prediction effect as a benchmark, and found that the prediction accuracy, R^2, and the Root Mean Square Error (RMSE) of the modified PCA-RFR constructed RON loss prediction model were 99.13%, 0.983, and 3.2169×10^{-4}, respectively. Qin Qingtao and Gu HNA considered the nonlinearity and mutual strong coupling between the variables [14] and used the multivariate auto-regression log-linear method to select the main variables to establish the RON prediction model. Chen Chan and Hu et al. firstly selected 25 feature variables using the Pearson-MIC-random forest method, [7] then predicted and optimized the XGBoost and optimization model of retained RON loss value. The RMSE, MAE (which equals to the total absolute value of the difference between the target value and the predicted value [15]) and the coefficient of determination of the model were 1.3197, 0.3581, and 0.9981 respectively. The coefficient of determination is also known as the degree of information interpretation, and as R^2. It indicates the amount of information the data contains: The closer the value is to one, the more sufficient the information expression is, represents that the actual problem situation is relatively more suitable for the model; If staying away from one, the more information is lost. Inevitably, the model is relatively not applicable for current practical problems [16]. Among these algorithms that select the features, only a small part of the features is retained to simplify the model fitting procedure while satisfying the minimum prediction accuracy requirements. In practical engineering, the accuracy should be improved as the primary goal to screen features, and the retention method retains more interactions between features in the selection process to improve the accuracy of model prediction, so the literature chooses the retention method to screen features. Although the above model has also achieved good prediction accuracy, if the appropriate iterative algorithm can be further used to reverse push and optimize the operation parameter scheme, both the prediction model can be fully utilized and improve the retained RON.

Zhang Fengyu et al. developed the machine learning model, [17] which used the maximum information coefficient to elaborate high nonlinear and coupling relationships and screened 35 significant variables from 353 variables for modelling. In the process of optimizing the back propagation neural network and logistics regression, the addition of the dragonfly algorithm makes this hybrid model

balance between local search and global search. The mean squared errors of the training and test sets were 0.0241 and 0.0413 respectively, and the mean absolute errors were 0.0982 and 0.1505 respectively, indicating the high accuracy and strong generalization ability of the integrated model. After optimizing the optimized process, 163 samples reduced RON loss by 70%, and another 128 by 50% to 70%, while all samples had optimized SC to below 5 μg·g-1. The key operating variables proposed by Huo Haoling et al., were identified by both the grey association analysis method and the partial least squares regression analysis method [18]. They then used the BP neural network to build a prediction model, and use genetic GA to optimize the BP neural network. Its average error, maximum absolute error, and prediction accuracy respectively are 0.051 units, 0.121 units, and 99.16%. Finally, they propose the optimal operation scheme, which reduces RON loss by 30%. Zhang Zhongyang et al. used a catalytic cracking reaction-regeneration system in a refinery to establish a BP neural network with 6-11-1 structure [19]. The neural network was optimized by genetic algorithm and successfully reduced the Mean Square Error (MSE), which is the sum of the square of the distance between the predicted value and the real value, [15] from 5.16 to 4.92%. Wang Jie and Chen Bo et al. selected 21 input variables and 1 output variable from 273 variables based on MIC and Pearson correlation coefficient method, and then used these variables to establish the BP neural network product RON optimization model [20]. The model structure is 21-14-1, and since the predicted MAE and R^2 are respectively 0.1163 and 0.9601 with the actual values, there is a good fit and generalization ability. The established prediction model was combined with the genetic algorithm GA algorithm, and then optimized operation variables, which reduced the product RON loss by 25% on the premise of ensuring the desulfurization effect. The algorithm mainly predicts the data through complex artificial intelligence algorithms such as neural network and genetic algorithm. The neural network has the essential characteristics of poor interpretation, and the genetic algorithm model will appear emergencies and fall into local optimal values when predicting complex problems. The principles of these models are complicated and difficult to explain. In practice, the Ridge regression models are relatively easy to explain. The purpose of parameter optimization is to ensure that the best parameter scheme can adapt to practical engineering, but currently few researchers distinguish between the range of operating parameters and not the range of operating parameters, so that the search for the 'best solution' may be meaningless.

In order to ensure that the model is simple and the prediction is accurate enough, the Leave-One-Out based on Ridge regression model is adopted to select some operating parameters, and fit the data with Ridge regression model to select the best hyper-parameter α and the weight values of the operating parameters, and finally find the functional relationship expression for predicting retained RON. The Ridge regression model was contrasted with other regression models to verify its relative applicability.

In order to make full use of the established model, the model can not only predict the new samples, but also use the Gradient descent method to select the best operating parameter scheme and the corresponding best retained RON. With or without the operating parameters bounds, the inversion of the selected operating parameters is performed to find the best operating parameters and the highest retained RON in both cases. The retained RON in the operating parameter range is more in line with the actual engineering requirements, while the no parameter limit can observe the maximum possible retained RON, which can help enterprises to analyse the difference between the two and explore the possibility and necessity of expanding the parameter range. It is worth noting that in the establishment of Ridge regression model, the raw material, adsorbent, and product performance were used as factors affecting the refining effect and as parametric variables of the model. However, because the three parameters are determined by the process before the catalytic cracking process, these parameters are not considered when performing the inversion of the operation parameters.

For the actual oil refining enterprises or other chemical production enterprises, more specific modelling thinking, and Python code are provided. And shares some thinking, such as the method of solving the number of necessary iterations, finding the optimal operating parameters and target values using the Gradient descent method, and also shares the direction that can be improved for the study.

2. Materials and methods

2.1. Data collection and pre-treatment

2.1.1. Data collection

The data of the literature comes from the 2020s Data Modelling Competition, which can be found by the browser, and contains the petrochemical real-time data of Sinopec Gaoqiao. The operating variables were collected from April 2017 to May 2020, including the data from 354 sites on the engineering equipment and devices, so the data from these 354 operating parameters (including temperature, pressure, device flow, etc.) will be used later in the study. The study focuses on the RON contained in the pre-refined and post-refined oil in the 325-samples data of the question bank. RON serves as the dependent variable of the mathematical model and operational parameters as the independent variable of the model.

Because the operation parameters can be adjusted in the catalytic optimization process, they are used as the inversion object when performing the parameter inversion. In the question bank, there are 14 raw material properties (e.g., sulfur content, RON, aromatic hydrocarbon, etc.), refined product properties (e.g., sulfur content and RON), and properties of raw and regeneration adsorbent parameters (e.g., coke content before and after adsorption), which do not belong to the operating parameters controlled by catalytic cracking engineering.

2.1.2. Data pre-treatment

In all kinds of projects, there are always some original data that are incomplete, which is a very common phenomenon in actual engineering, and the project also has this phenomenon. There are also two problems with the data generated at different sites of each device: some operation parameters contain only partial time period data; some operation parameters are all null or some data is null. Therefore, before building the simulation of the data, the following is necessary to pre-process the original data to facilitate the subsequent model fitting and analysis. Data pre-processing method is performed as follows:

Removal of duplicate values. Due to no completely sound management ability and reliable calculation method, accidental sample data duplication errors cannot be avoided in the process of data collection, which need to be removed manually.

Deletion missing. For sites containing only some time points, if the data missing (%) is more than 20%, and it cannot be supplemented, so that the index and row data will be deleted.

Fill the empty value. For some sites where the data is null, the data at the null value is replaced by the average of the two firstly and second hours.

Manual removal of the outliers. There are some original data beyond the reasonable range of process and operation experience, which can be convenient manually, but the code is inconvenient to find out, thus it is more suitable to reduce the code complexity through manual elimination.

The 3σ criteria removed outliers. Equal precision measures the variables to be measured, yielding x_1, x_2, ..., x_n, and the mean value (x) and the residual error ($v_i = x_i - x$ (i= 1, 2, ..., n)) can be calculated, thus the standard error (σ) can be obtained by using the Bessel formula. When The residual error (v_b, $1 \le b \le n$) of a measurement meet the condition ($| v_b | = | x_b - x | > 3\sigma$, take xb as an invalid value with a coarse error value and delete it. Equation 1 is a Bessel formula (Zhao et al., 2022; Jin et al., 2021) [21, 22].

$$\sigma = [\frac{1}{n-1}\sum_{i=1}^{n} v_i^2]^{1/2} = \{[\sum_{i=1}^{n} x_i^{\,2} - (\sum_{i=1}^{n} x_i)^2 / n] / (n-1)\}^{1/2} \quad (1)$$

where Σ is standard error; vi is remainder error; vb is the remaining error of xb; xiis measured value; n is sample number.

Use Bessel's formula to filter repeatedly, in other words, calculate the mean and find outliers and then it around.

2.2. Feature processing

2.2.1. Feature screening (Leave-One-Out)

There are 354 operation variables in the database, and the number of operation parameters already belongs to the number of large-scale operation parameters. If all the operation parameters are used for model calculation, it will cause unnecessary computer card machine, excessive waste of memory resources, and too long calculation time to affect the work efficiency. At same time, the modelling process may be overfitting due to containing too many useless indicators, and even the calculation results will appear large error, so that the work results cannot be used. Therefore, the operation parameters need to be screened by the Python code and the Leave-One-Out method based on Ridge regression model before modelling. The Leave-One-Out method retains more interactions between features in the selection process, so that the goal of various feature selection can be uniformly optimized [23].

2.2.2. Data partitioning

The target object of the study is the retained RON after catalytic cracking performed by the S Zorb device (catalytic gasoline adsorption and desulfurization device). In order to solve the problem of constant lag of octane measurement, and of the possibility of overfitting by large lack of operating data leading, the average of the operating variable data for the first two hours of octane data measurement was taken as the operating variable data at this time. The qualified 325 samples were divided into training and test sets. Among them, 176 sets of data from 17 April 2017 to 31 December 2018 were divided into training sets, and 149 sets from 2 January 2019 to 26 May 2020 were classified as the test set.

2.2.3. Standardization / normalization

There is a phenomenon in the common data problems in life, that is, manydifferent variables often contain partial variables with different dimensions. The normalization can eliminate the effect of the dimension on the result, and make the different variables comparable [24]. For example, compare the performance strengths of the two people. The sum of Chinese (150 points in total) and sports (10 points in total) as the total score. A's Chinese and PE scores are 135 and 9 respectively, while B's Chinese and PE scores are 140 and 6 respectively. Although B's total score (146 points) is greater than that of A (144 points), A's Chinese and sports performance are superior, and B's sports is worse. From the perspective of data, B is better than A since its total score is higher, but from the perspective of social cognition, since both scores of A’s are close to full marks, A is better. Therefore, standard normalization needs to be done before data modelling, like the statistical comprehensive performance method. The new parameter data after normalization or standardization method has the characteristics of mean value of zero and standard deviation of one. The transformation formula is shown in equation 2. Finally, standardized regression processing.

$$X' = (x - \mu) / \sigma \tag{2}$$

where X' is the normalized data; μ is the mean of the sample; σ is the standard deviation of the sample.

2.3. Data collection and pre-treatment

2.3.1. Feature selection (Leave-One-Out method）

In the literature, 0.5 of α was assumed in advance, and an operation parameter was used to find the predictive value, and compared with the actual value and find the MSE. Then another operation parameter was added to fit again to find the predicted value and MSE. And so on, one operation parameter was added at a time and fitted once with a Ridge regression model, until all 354 parameters were added to the model, seeking the predictive value and MSE. Therefore, the data of the corresponding test set MSE, training set MSE, weight value MSE and training R^2 for different feature numbers can be displayed and compared with a graph, to facilitate visual observation, analysis and select the appropriate number of features. When selecting the number of features, the specific selected feature indicators can also be selected.

According to assuming α as 0.5, and using the Ridge regression linear model algorithm, the code can find the best number of parameters, and the feature combination. The process takes the comprehensive error Mz as the measure feature standard, and the best number is found according to the number of parameters-mean variance map. Mz algorithm is shown in equation 3.

$$M_Z = (0.5 * \sigma_{train} + 0.5 * \sigma_{test}) \quad (3)$$

where Mz is the integrated error; σtrain is the mean variance of the training set; σtest is the test-set mean variance. In order to intuitively feel the size of the RON and the MSE for different numbers of features, the number of features was extracted in three contrasting cases. For example, when the number of features feature_num equals 1,146 and 361. Then, the training set and the test set were fitted with the Ridge regression model, and the degree of coincidence between the actual values and the predicted values was used to judge the fitting effect of different feature quantities.

2.3.2. Ridge regression

Prediction continuous data methods often use least squares and Ridge regression linear model algorithms, but the two face different situations. Ridge regression is a biased estimation regression method to deal with the presence of collinearity data [25, 26]. Fundamentally, it is a modified model of the least squares' method. Partial information and accuracy are sacrificed to fit a more realistic and reliable regression model, and it is more applicable than the least squares method when the data is sick [27]. Many characteristic variables often have a high collinearity relationship in engineering, including the gasoline refining project in the literature. If the project uses the least squares method, the generalization ability of the fitting model will decrease. In other words, the accuracy of predicting the new sample is less than that of the Ridge regression model; Although the R^2 of the Ridge regression equation is often slightly lower than that of the ordinary regression analysis, the significance of the Ridge regression is often better than the ordinary regression analysis, and it has more practical value when there are collinearity problems in the actual data and more pathological data. Therefore, the Ridge regression linear model algorithm is used in the study.

Select the best hyper-parameter α. When selecting the appropriate number of features in the previous step, the pre-set α value is 0.5 (often applicable α value), however, in the actual modelling process, the model needs to verify different α s and select more appropriate α. The study separately fitted different α values in the range of 10^{-4} to 10^4 and 10^{-2} to 10^2, and then the suitable α was selected by analysing theconvergence of the weight value, the MSE size of the test set, and the change trend of the R^2 in the training set. Because the more the α is larger than the appropriate minimum α, the more information is lost, [28] it is necessary to find the minimum hyper-parameter α value in the appropriate range and make it serve as the best hyper-parameter α in the Ridge regression model.

Use the loss function of the Ridge regression to find the appropriate hyper parameter α of the normalized coefficient stability period with the corresponding coefficient vector. The complete expression of the loss function of the Ridge regression model is shown in equation 4.

$$J(\theta) = \min \| Xw - y \|_2\text{^}2 + \alpha \| w \|_2\text{^}2 \quad (4)$$

where J(θ) is loss value; α is complexity parameter controlling the amount of shrinkage, collectively referred to as the hyper-parameter; X is an independent variable matrix; w is weight; y is dependent variable.

2.3.3. Model comparison

In order to verify whether the Ridge regression model is more suitable for the project compared than other models, multiple commonly used applicable models (e.g., decision tree, linear regression, random forest, Bayes, etc.) are established. Fit with training set data firstly, and then verified with test set data,

finally the predicted value of each model is compared with the actual value and the MSE of each model is calculated. For easy visual observation, the MSE and the training set R^2 of the training set and the test set of each model are shown in the broken line chart and compared. When choosing an appropriate model, the researcher should not only consider R^2, but also consider the test set MSE and generalization ability comprehensively, and finally select the optimal model for modelling.

2.3.4. Gradient descent algorithm

Gradient descent method is an iterative inversion method with the gradient as the search direction. The loss function of the Ridge regression model is used to evaluate the accuracy of the model. Normally, the smaller the value of the loss function, the higher the accuracy of the model. The Gradient descent method is used to find the minimum loss function. Exploring this lowest point is like looking for the lowest point in a valley; By using the derivative in calculus and finding the derivative of the function at every step, can find the downward direction of the function in the valley or the lowest point / extreme point.

Through the gradient of continuous decline, the weight and deviation in the loss function are constantly adjusted, and the loss function is constantly reduced. When the loss value is the minimum or the local minimum is the same as the gradient is zero, the corresponding operation parameters are the local optimal or the global optimal scheme. However, if the step length is too large, the loss value may be larger. In general, the gradient shows a decreasing trend with the increasing number of iterations. In order to explore the appropriate number of iterations, the number of iterations and the effect of iterations are shown with charts.

2.3.5. Parameter inversion based on the gradient backpropagation algorithm

Operation parameter range is not set: The operational data in the original data source must not be the best operating parameter value. To allow the established Ridge regression model work to play more roles, it can be used to increase retained RON when refining gasoline in practical engineering. In other words, the previous established Ridge regression model was iterated, and the appropriate number of iterations was found by the Python code of the Gradient descent method. Then, the best operation parameters corresponding to the iteration number are brought into the Ridge regression model to find the predicted RON value after the inversion. The obtained predicted retained RON value is compared with the actual retained octane value before the inversion of the operational parameters, so as to understand the optimization degree of the proposed modelling method. The optimized predicted retained RON is compared to the expected retained RON, the amount of all retained RON in oil, to know the room for continued iterative inversion. Through this programming thinking, people can understand the method of finding the best number of necessary inversions, and the increasing degree of RON retention after parameter inversion.

Set the value range of the actual operation parameters: In actual engineering operation, many operating parameters cannot pursue the highest retained RON due for cost, safety or other factors. Therefore, basing on the above step, the Gradient descent method is still used to manually add the value range of the parameters in the Python code. The specific value range data is still derived from the question bank of the 2020 mathematical modelling competition. From the theoretical analysis, the optimal inversion number of the limited parameter value should be different from that of the unlimited parameter range, and the optimal operation parameters obtained when limiting the value range of the operating parameters are more feasible.

3. Results and discussions

3.1. Feature selection (Leave-One-Out method based on Ridge regression)

The Python code of the Ridge regression model is run to show the change trend of the MSE of the training set and the test set with the number of operation parameters. The best number is calculated according to the number of parameters-mean variance map, as shown in Figure 1.

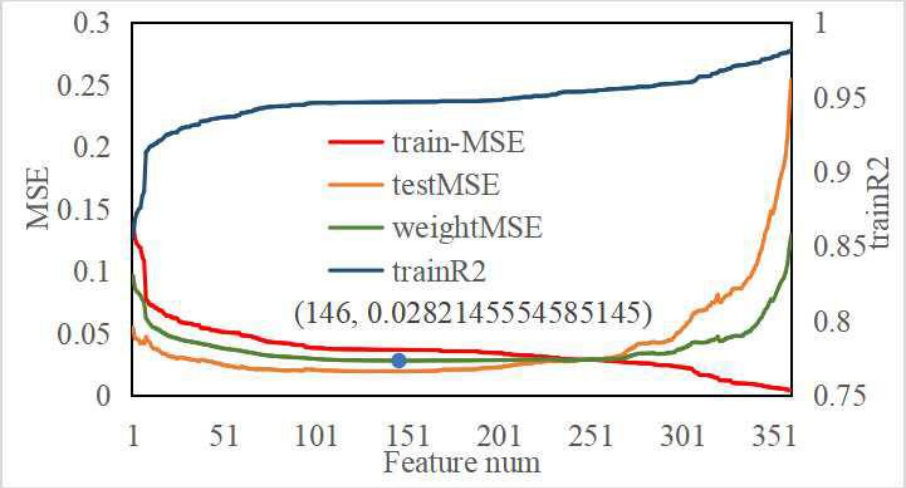
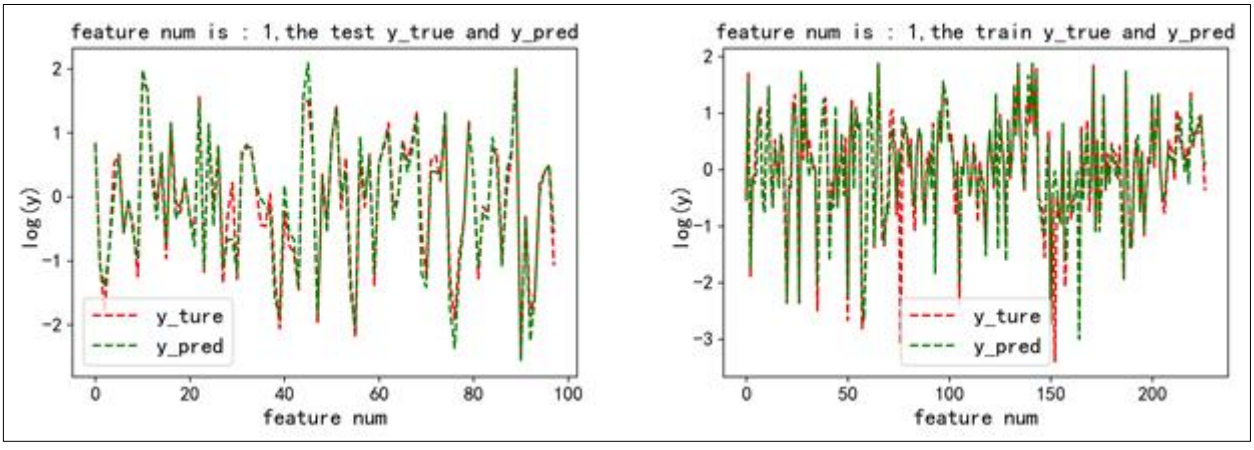


Figure 1. Training-training set / test set mean variance plot for different feature quantities

As can be seen from Figure 1, with the increase of the number of features, the MSE of the training set firstly drops rapidly and then stabilizes around the error of 0.04 units. At this time, the number of features is about 150, and then the MSE gradually decreases to 0.01, indicating that the model fitting effect is getting better and better; When the number of features varies from 0 to 150, the MSE of the test set firstly decreases and then stabilizes. After more than 150, it grows from slowly to fast, indicating that the model later appears overfitting phenomenon, and lead to a decline in the generalization ability; Training set R^2 firstly increased rapidly and then stabilized to around 95.2%, and finally rose to 98%.

The number of features can be selected when the MSE of the training set and the minimum MSE of the test set are stable. Because this condition is satisfied when the number of features is 146, 146 is the best number of features, and the 146 features are also used in the subsequent Ridge regression model training in the study.

In order to distinguish between the fitting effect of the training set and test set in the three cases of too few, suitable and too features, the broken line diagram shows the curve coincidence degree of the predicted value and the actual value when the feature_num is 1, 146, and 361 respectively, as shown in Figure 2.



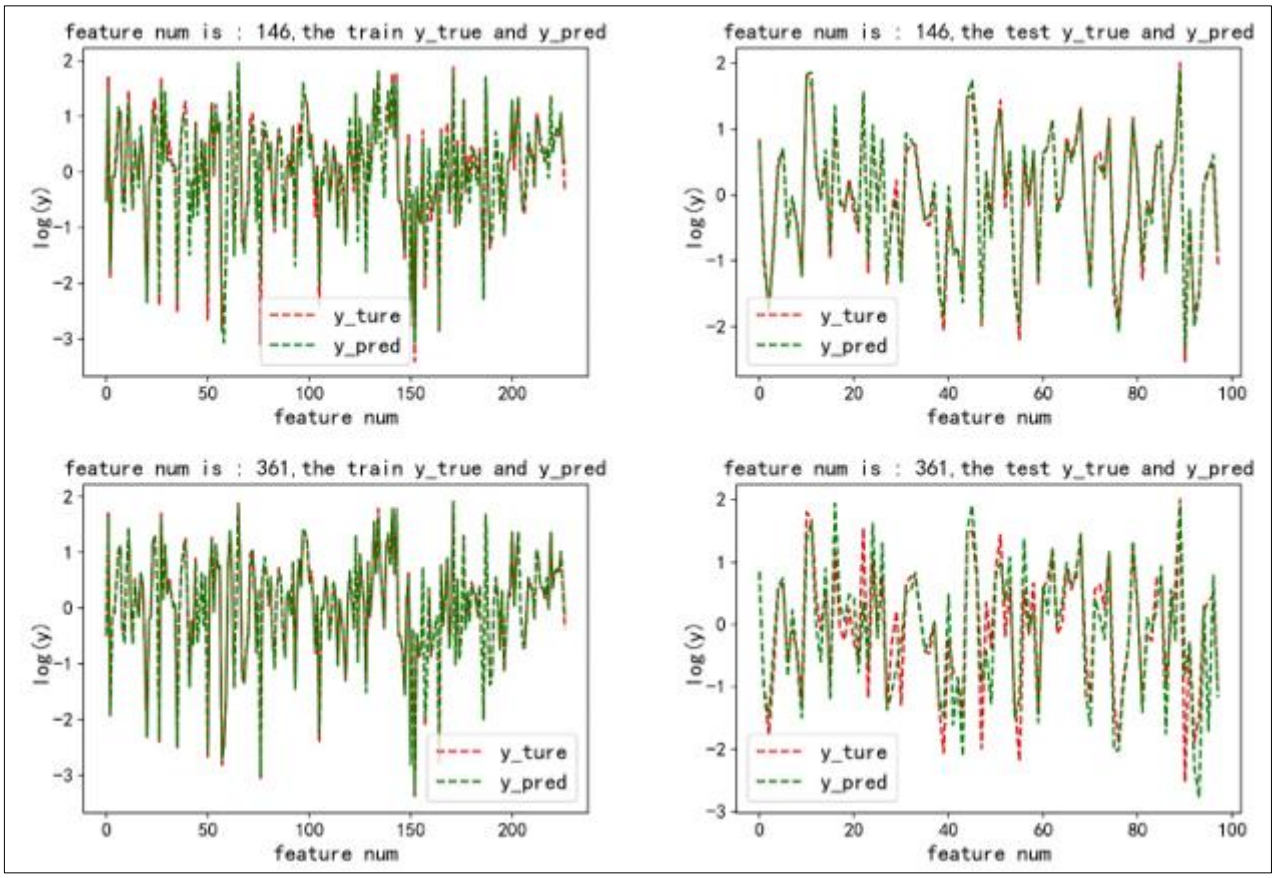


Figure 2. umber of features-the fit effect of the training set / test set.
(a) feature_num = 1; (b) feature_num = 146; (c) feature_num = 361

By observing Figure 2, it is easily seen that, when the number of features is feature_num is 1 or 361, both the actual value curves of the training set and the test set are worse than the two curves when the feature number is 146. Interpreted from the professional perspective of the numerical analysis, the fit of the training set and the test set show an under-fitting effect when the number of features is one.

Under-fitting can be considered that, when the model fits the training set data, the data points used in the model are too far away from the actual data points of the project, and make the error between the prediction value of the training set and the actual value is too large. Inevitably, the prediction must not achieve the test accuracy usually required by the test set or the new sample set data. Therefore, it is also believed that the model is insufficient to "learn" the "general law" in the data set [29].

When feature_num is 146, the fitting effect of training set and test set is relatively best. The fit of the training set/test set was over-fitted at a feature_num of 361. Overfitting is that the model goes through each actual data point of the training set as much as possible during training, in order to avoid the under-fitting phenomenon. Despite the model is well suited to the data prediction effect of the training set, it usually causes the model to be more distant from the actual value when predicting the test set or the new sample. Therefore, it can be considered that the model's "learning" ability is too strong, or the generalization ability is too weak [29].

Generalization ability is the ability of machine learning algorithms to adapt to fresh samples [30]. The purpose of learning is to use the model to collect the rules behind the data that is not easy to be detected by human beings, and to require that the prediction value error of the training set is within a certain range. And the prediction error of the model with good generalization ability when predicting other data sets with the same law should also be within the appropriate or very small range.

3.2. Equations

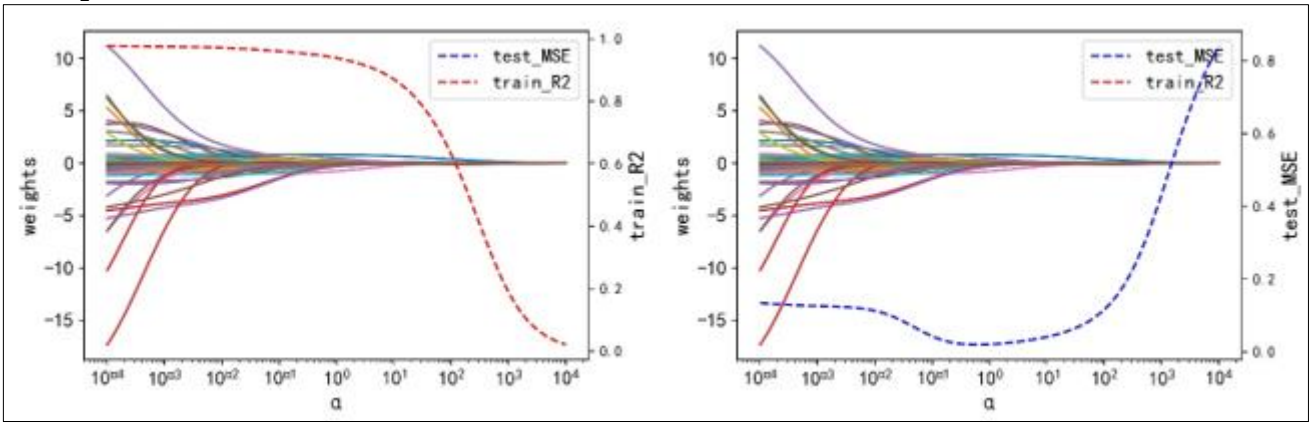


Figure 3. Ridge coefficients as a function of the regularization

On Figure 3 when α is around one, the weight coefficient uniformly levelled off. Moreover, when the α is greater than one, the R^2 of the training set begins to drop sharply, indicating that the degree of agreement between the experimental data and the fitting function decreases sharply when the α is larger than one, so the α cannot be larger than one. Meanwhile, the MSE of the test set begins to rise rapidly, which further indicates that the value of α cannot be larger than one, so it can be considered that the α optimal value is around one.

According to Figure 3, it is not possible to clearly observe the weight value trend of α around 0.6, so that the graph is enlarged in the range from10^{-2} to10^{2} to observe the convergence of most of the weights when α is 0.6, as shown in Figure 4.

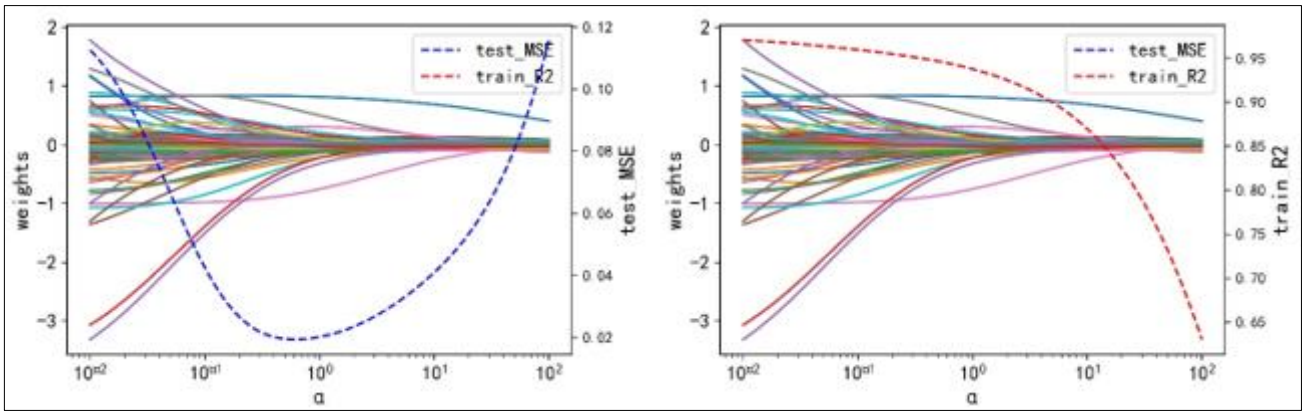


Figure 4. Convergence of most parameters with α of 0.6

From Figure 4, when α is 0.6, almost the ownership weights tend to converge; The test set MSE also tends to the minimum; And the training set R^2 begins to drop sharply. Therefore, it is reasonable to use the α of 0.6 as the hyper-parameter value for this Ridge regression, and the literature also uses the α of 0.6 for the analysis study when using the Ridge regression model. Finally, the relationship expression between operation parameters and RON is obtained, and the expression intercept is 2.74524575722225 $\times 10^{-3}$. For the selected operation parameters, and the value ranges and weights of each parameter, see Table 1. Filtered Operation Parameters, Parameter Value range, and Weight Values.

Table 1. Filtered operation parameters, parameter value range, and weight values

No.	Item	Name	Span	Weight
1	-	(Raw material properties)RON	-	0.830015597
2	-	(Raw material properties)arene, v%	-	-0.043006862
3	-	(Raw material properties) bromine number, gBr/100g	-	-0.007754015
4	-	(Raw material properties)density (20℃), kg/m³	-	0.043257492
5	-	(Nature of standby adsorbent)coke, wt%	-	0.00499942
6	-	(Nature of standby adsorbent)S, wt%	-	0.041182071
7	-	(Nature of the regenerative adsorbent)coke, wt%	-	-0.081841565
8	-	(Nature of the regenerative adsorbent)S, wt%	-	0.017127962
9	S-ZORB.FT_1002.PV	1#Catalytic gasoline intake unit flow rate	0 - 140	-0.273538408
10	S-ZORB.TC_2801.PV	Reducer temperature	200 - 350	0.12514685
11	S-ZORB.TE_9002.DACA	D-203 - Top outlet pipe temperature	10 - 50	0.002198672
12	S-ZORB.FT_5201.PV	Flow rate of the product gasoline outlet device	10 - 155	0.10860839
13	S-ZORB.TE_5002.DACA	C-201 - Lower feed pipe temperature	100 - 150	-0.837236397
14	S-ZORB.TE_1201.PV	D104 - Temperature	100 - 150	0.605531538
15	S-ZORB.PT_6009.DACA	Preheater air outlet pressure	0 - 1.5	-0.028822969
16	S-ZORB.TE_5102.PV	Dry gas outlet unit temperature	20 - 40	-0.048319417
17	S-ZORB.FT_1003.PV	2#, Catalytic gasoline intake unit flow rate	0 - 75	-0.026973779
18	S-ZORB.PT_9401.PV	Purify the air inlet device pressure	0.35 - 0.60	-0.043687896
19	S-ZORB.TE_1601.PV	Heating furnace inlet temperature	350 - 400	-0.085858854
20	S-ZORB.FT_9401.PV	Purify the air inlet device flow rate	25 - 900	0.015678794
21	S-ZORB.FC_5202.PV	Flow rate of refined gasoline outlet unit	90 - 160	-0.106637214
22	S-ZORB.FT_9102.PV	Flare gas discharge flow rate	20 - 20000000	0.065487634
23	S-ZORB.BS_LT_2401.PV	Closed lock material bucket liquid level	2.5 - 62.5	0.089349976
24	S-ZORB.FC_2601.PV	R102 - The regenerator boosts the nitrogen flow rate	2 - 100	0.092819321
25	S-ZORB.TE_5006.DACA	Stabilize the tower bottom outlet temperature	100 - 150	-0.050861749
26	S-ZORB.TE_2004.DACA	R - 101 - Lower bed temperature	400 - 450	0.125808018
27	S-ZORB.PT_9301.PV	Steam inlet unit pressure	0.5 - 1.3	0.034073516
28	S-ZORB.TE_5101.DACA	A - 201 - Outlet main pipe temperature	20 - 80	-0.070532788
29	S-ZORB.TE_2002.DACA	R - 101 - Central temperature of bed layer	400 - 450	-0.167388951
30	S-ZORB.TE_5001.DACA	E - 206 - Shell process outlet tube temperature	50 - 150	0.016314492
31	S-ZORB.LI_2107.DACA	DI - 2107	- 3 - 9	-0.064221614
32	S-ZORB.FC_5001.DACA	E203 - Condensate water flow of the reboiler pipe process outlet	500 - 3800	0.145043338
33	S-ZORB.SIS_FT_3202.PV	EH - 103, Entrance flow	120 - 350	0.027007018
34	S-ZORB.PDI_2801.DACA	Regenerator - LH, differential pressure	0 - 2.5	1 月 0 日
35	S-ZORB.PDT_3602.DACA	Cold nitrogen filter - ME - 114, differential pressure	0 - 1	0.031373324
36	S-ZORB.FT_3001.DACA	D - 113 - Top empty line flow	15 - 250	-0.038813053
37	S-ZORB.PDT_3503.DACA	ME - 112 - Filter pressure difference	- 0.5 - 1.5	-0.044598031
38	S-ZORB.PT_1501.PV	-	0 - 2.5	0.062982478
39	S-ZORB.TE_1603.DACA	F - 101 - Export branch pipe#2, temperature	400 - 420	0.069820638
40	S-ZORB.FT_9101.PV	Sewage oil outlet device	0 - 85	-0.048698784
41	S-ZORB.FT_9302.PV	0.3MPa - Flow rate of the condensate water outlet device	3 - 6500	0.106140608
42	S-ZORB.TC_2201.OP	EH101 - export	44701	-0.063692404
43	S-ZORB.FC_2301.PV	D105 - Fluid hydrogen flow	0 - 350	-0.070339509
44	S-ZORB.SIS_TEX_3103B.PV	EH - 102 - Heating element / B beam temperature	80 - 150	0.096650382
45	S-ZORB.PC_6001.PV	Radiation chamber outlet pressure	(-0.20) - 0	-0.020598089
46	S-ZORB.FT_1006.DACA.PV	Hydrocracking light naphtha inlet device flow rate	0 - 12000	-0.055953368
47	S-ZORB.RXL_0001.AUXCALCA.PV	Heating furnace efficiency	90 - 98	0.083802416
48	S-ZORB.AT-0001.DACA.PV	S_ZORB AT - 0001	0 - 10	0.060429601
49	S-ZORB.TE_2501.DACA	D - 107 - temperature	100 - 250	0.022922527

50	S-ZORB.PT_6005.DACA	F - 101 - Bottom pressure of the radiation chamber	(- 2) - 0	-0.15994479
51	S-ZORB.FC_2801.PV	Reducer fluidihydrogen flow	600 - 1000	-0.055026486
52	S-ZORB.PT_2607.DACA	R - 102 - Nitrogen line pressure behind the bottom drain slide valve	0.1 - 0.2	0.019513913
53	S-ZORB.FT_3702.DACA	Lock the hopper H2 filter outlet gas flow rate	0 - 60	-0.048253188
54	S-ZORB.CAL_H2.PV	Hydrogen oil ratio	0.20 - 0.37	-0.003278637
55	S-ZORB.TE_1602.DACA	F - 101 - Exit branch # 1 temperature	400 - 420	-0.094138792
56	S-ZORB.PT_6008.DACA	F - 101 - Radiation chamber outlet pressure	(- 0.5) - 0	0.014588788
57	S-ZORB.PDI_2105.DACA	Backblowing gas accumulator / supplementary hydrogen differential pressure	2.5 - 5.5	0.040187255
58	S-ZORB.PC_5101.PV	Stabilize tower top pressure	0.60 - 0.70	0.010407133
59	S-ZORB.PDT_2704.DACA	regenerator receiver top / regenerator receiver bottom differential pressure	25 - 55	-0.063747646
60	S-ZORB.FT_9002.DACA	D203 - Export fuel gas flow rate	300 - 650	0.060782663
61	S-ZORB.TE_1001.PV	Raw material inlet device temperature	35 - 80	0.308709304
62	S-ZORB.AT_5201.PV	Sulfur content of the refined gasoline effluent unit	0 - 5	-0.027602935
63	S-ZORB.TE_5003.DACA	C - 201#37 - Tower disc temperature	50 - 100	-0.033993714
64	S-ZORB.TE_6001.DACA	Temperature of the flue gas out of the convection chamber	300 - 400	0.275584435
65	S-ZORB.SIS_TE_6009.PV	Preheater inlet flue gas temperature	10 - 350	-0.125533929
66	S-ZORB.TE_1106.DACA	E - 101A - Shell process outlet tube temperature	100 - 200	0.109074485
67	S-ZORB.PDT_2906.DACA	ME - 108 - Filter differential pressure	- 2 - 25	0.075830681
68	S-ZORB.TE_7504B.DACA	K - 102B - discharge temperature	2 - 150	0.226814549
69	S-ZORB.FT_5204.DACA.PV	Gasoline products to divide the gas flow rate	0 - 2500	0.022549194
70	S-ZORB.PDT_1002.DACA	P - 101 - AInlet filter differential pressure	- 0.5 - 12	-0.078547629
71	S-ZORB.PT_7103B.DACA	K - 101 - Badmission pressure	0 - 2.5	-0.118420612
72	S-ZORB.CAL.LINE.PV	Reactor line speed	0.2 - 0.5	0.01374251
73	S-ZORB.TE_7102B.DACA	K - 101B - inlet temperature	0 - 50	0.059876355
74	S-ZORB.TE_1504.DACA	E - 106 - Pipe line inlet pipe temperature	2 - 100	-0.336243383
75	S-ZORB.DT_2001.DACA	R - 101 - Lower bed lamination drop	0 - 120	0.064654789
76	S-ZORB.FT_1501.TOTAL	New hydrogen inlet device flow rate	0 - 55000000	-0.124560863
77	S-ZORB.TE_2902.DACA	D - 109 - bottom	3 - 60	-0.148634944
78	S-ZORB.PC_3001.DACA	D - 113 - tension	0 - 0.15	-0.005126743
79	S-ZORB.PC_1301.PV	K101 - Machine export pressure	2.55 - 3.55	0.138586938
80	S-ZORB.FT_9101.TOTAL	-	425 - 120000	-0.137636853
81	S-ZORB.CAL.CANGLIANG.PV	Reactor storage	15 - 45	-0.089547406
82	S-ZORB.FT_3304.DACA	D - 123 - Condensate water inlet flow rate	- 10000 - 2000	-0.210701158
83	S-ZORB.LI_9102.DACA	D - 204 - level	10 - 90	-0.007610577
84	S-ZORB.SIS_TE_2605.PV	Reator lower temperature	450 - 550	-0.000587591
85	S-ZORB.LC_1201.PV	D104 - liquid level	45 - 55	-0.013299277
86	S-ZORB.LC_5101.PV	Top return tank D201 level	40 - 800	-0.065658954
87	S-ZORB.FT_5102.PV	-	0 - 450	0.004255801
88	S-ZORB.TE_3112.DACA	EH - 102 - outlet pipe	65 - 150	-0.06353213
89	S-ZORB.TE_2103.PV	Reactor upper temperature	410 - 435	0.06463739
90	S-ZORB.TE_2401.DACA	D - 106 - temperature	200 - 350	-0.057708449
91	S-ZORB.LT_9101.DACA	Flare tank D - 206, liquid level	- 5 - 35	0.020043878
92	S-ZORB.TE_2608.DACA	R - 102 - Bottom cone temperature	100 - 500	0.028167303
93	S-ZORB.TXE_3202A.DACA	EH - 103 - Heating element temperature	250 - 500	0.034307998
94	S-ZORB.PDI_2301.DACA	Reactor receiver LH differential pressure	0 - 2.5	-0.078305703
95	S-ZORB.FC_2432.PIDA.SP	Step 3.0, FIC2432.SP	0 - 80	0.031592406
96	S-ZORB.LC_5002.DACA	D - 202 - level	35 - 70	-0.367534428
97	S-ZORB.PT_7505.DACA	K - 102 - Aexhaust pressure	0 - 16	-0.465409958
98	S-ZORB.PT_5201.DACA	Refined gasoline outlet line pressure	0.5 - 0.65	-0.152728079
99	S-ZORB.AT-0011.DACA.PV	S_ZORB AT - 0011	0.4 - 0.8	0.075376838
100	S-ZORB.TE_1105.PV	Inlet temperature of raw material heat exchanger pipe	40 - 80	-0.126806086
101	S-ZORB.TE_1604.DACA	F - 101Export branch pipe #3 - Temperature	400 - 450	0.055684382
102	S-ZORB.LC_3301.DACA	D123 - Condensed water tank level	45 - 55	0.014146941
103	S-ZORB.PC_9002.DACA	D - 203 - Top outlet pipe	0.35 - 0.40	0.010646481
104	S-ZORB.LT_1301.DACA	D - 103 - Bottom level	- 1.5 - 9	0.014226803
105	S-ZORB.LT_3101.DACA	D - 124 - level	- 1.8 - 7.0	0.026803878

106	S-ZORB.AT_1001.PV	Sulfur content of the inlet plant raw material	1.5 - 650	0.023353757
107	S-ZORB.FT_9401.TOTAL	-	15000 - 3500000	0.146409658
108	S-ZORB.TE_7106.DACA	K - 101A - Left exhaust temperature	5 - 65	0.008575939
109	S-ZORB.FT_2431.DACA	-	20 - 1500	0.013303107
110	S-ZORB.PT_6002.PV	Heating furnace and furnace pressure	- 0.60 - (- 0.15)	0.088056027
111	S-ZORB.PDT_2605.DACA	R - 102 - Bottom spray head pressure difference	- 0.5 - 30	0.019988004
112	S-ZORB.PDI_2703A.PV	D110 - Top bottom pressure difference	25 - 60	0.030106471
113	S-ZORB.TE_3101.DACA	D - 124 - Top outlet pipe temperature	3 - 40	0.096853099
114	S-ZORB.FT_5102.DACA.PV	D - 201 - Sulfur-containing sewage displacement	0 - 420	-0.015759223
115	S-ZORB.PC_2105.PV	Anti - blowing hydrogen pressure	4.5 - 5.85	-0.127135946
116	S-ZORB.FT_9403.TOTAL	-	120000 - 30000000	0.009042252
117	S-ZORB.SIS_TE_6010.PV	Heat the furnace and exhaust smoke outlet temperature	10 - 180	0.016036043
118	S-ZORB.LC_1203.PIDA.PV	D - 121 - Sulfur-containing sewage liquid level	35 - 55	0.007427085
119	S-ZORB.PDT_1004.DACA	ME - 104 - passageway	- 1 - 55	-0.009103655
120	S-ZORB.PT_2603.DACA	R - 102 - Lower pressure	0.1 - 0.2	0.02883449
121	S-ZORB.AT-0006.DACA.PV	S_ZORB AT - 0006	0.4 - 0.6	0.003861565
122	S-ZORB.PDT_2604.PV	Top differential pressure of regenerator	20 - 45	-0.024521834
123	S-ZORB.LI_2104.DACA	As calculated by the PDI2104	15 - 45	-0.084132764
124	S-ZORB.TE_2001.DACA	R - 101 - Central temperature of bed layer	- 243600 - 12500000	-0.035235681
125	S-ZORB.FT_9001.PV	Fuel gas inlet device flow rate	350 - 600	-0.004626248
126	S-ZORB.FT_1004.PV	3#-Catalytic gasoline intake unit flow rate	0 - 90	0.018072898
127	S-ZORB.TE_5009.DACA	E - 205 - Pipe line inlet pipe temperature	5 - 50	0.048375513
128	S-ZORB.TE_6008.DACA.PV	Preheater outlet air temperature	0 - 300	-0.05193536
129	S-ZORB.FC_1202.TOTAL	Accumulated flow rate of waste hydrogen discharge	25000 - 500000	0.157849063
130	S-ZORB.SIS_PT_6007.PV	Air preheater flue gas outlet pressure	- 1500 - (- 100)	0.000966148
131	S-ZORB.LT_3801.DACA	D - 125 - Level	- 0.85 - 2.00	-0.002986763
132	S-ZORB.PDT_2409.DACA	ME - 115 - Filter pressure difference	- 0.5 - 25	0.007626224
133	S-ZORB.PT_1602A.PV	Front pressure of the heater main fire nozzle valve	0.35 - 0.40	-0.029403749
134	S-ZORB.AI_2903.PV	Front pressure of the heater main fire nozzle valve	0.5 - 3	-0.006306914
135	S-ZORB.TE_9003.DACA	D - 203 - Fuel gas inlet pipe temperature	10 - 40	-0.0354609
136	S-ZORB.PT_1101.DACA	Remove the hydrogen mixing point pressure at the circulating hydrogen compressor outlet	2.5 - 3.5	-0.015476296
137	S-ZORB.TE_1203.PV	D121 - temperature	25 - 50	0.011109287
138	S-ZORB.FT_1204.TOTAL	-	45000 - 2500000	0.010970381
139	S-ZORB.DT_2107.DACA	R - 101 - The upper bed is laminated	- 3 - 10	-0.066159393
140	S-ZORB.TC_3102.DACA	E - 105 - Pipe exit pipe	150 - 300	0.017440465
141	S-ZORB.TE_2601.PV	Reator top flue gas temperature	200 - 400	-0.035380008
142	S-ZORB.PC_2401B.PIDA.OP	Step 9.0 - PIC2401B.OP	15 - 55	-0.025045507
143	S-ZORB.FC_1202.PV	D121 - Top the torch flow	0 - 300	0.001329927
144	S-ZORB.TE_1503.DACA	E - 106 - Pipe line outlet pipe temperature	2 - 50	-0.015407718
145	S-ZORB.TXE_3201A.DACA	EH - 103 - Heating element temperature	300 - 550	-0.01349998
146	S-ZORB.LC_1203.DACA	D - 121 - Water level	35 - 55	0.013867702

3.2.1. Display of the fitting effect

The predicted and actual values of the training set and the test sets are shown to facilitate the observation and confirmation of the effect of the Ridge regression model fitting to the training set and the test set, as shown in Figure 5.

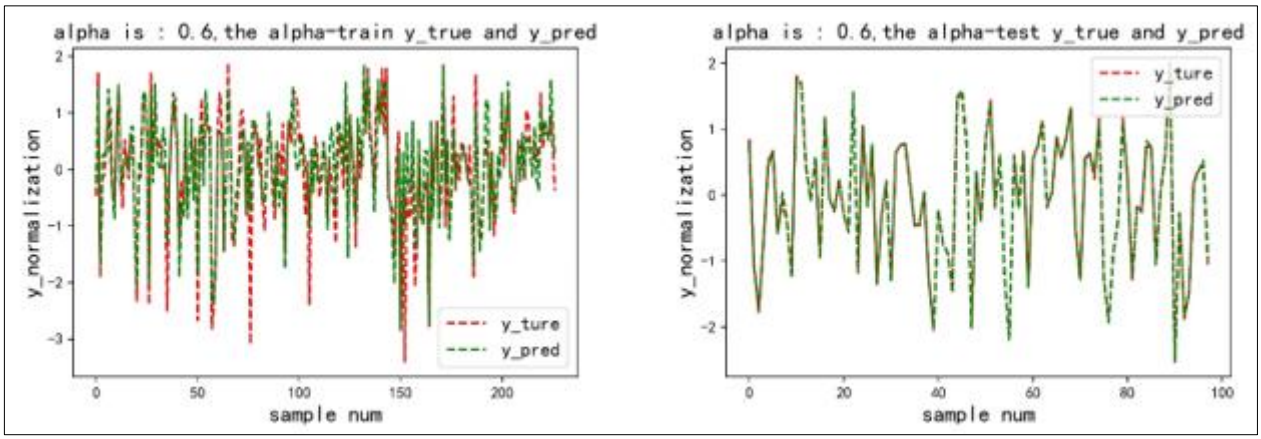


Figure 5. Fit the effects map to the training set and test sets

According to the Figure 5, the fitting effect of the training set and the test set of this database is not a poor effect. By the operation of the python code, the test set R^2 is 98.8% when the α value is 0.6, indicating that the Ridge regression model has 94.5% information interpretation, and the training set and prediction set MSE are 0.037855 and 0.019342 respectively.

3.3. Comparison of models

To evaluate the fitting effect of the target model in the literature compared with other models, should not only observe the MSE size of each model training set to judge the fitting effect of the training set, but also observe the MSE size of the test set to judge the fitting effect of the test set, and also consider R^2. The comparison results are shown in Figure 6.

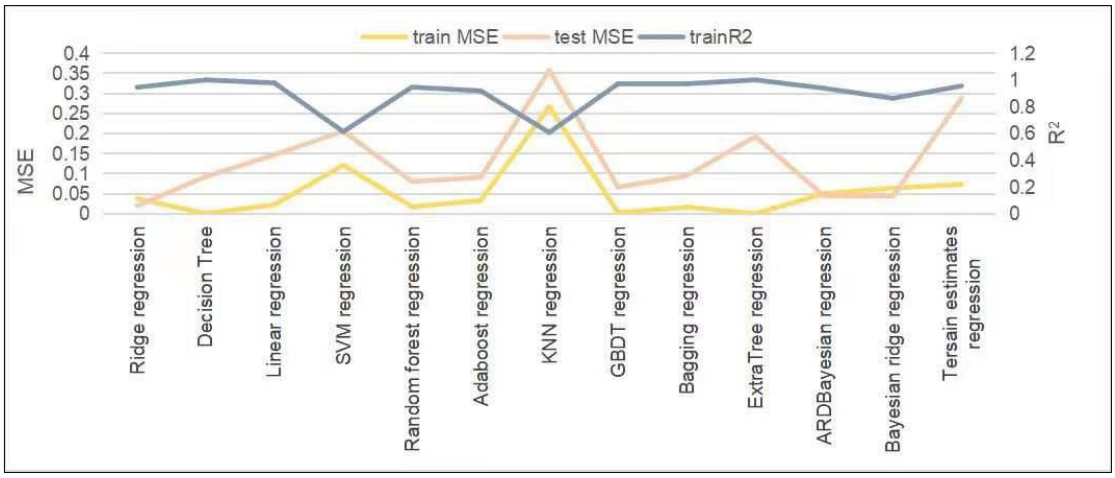


Figure 6. Comparison of the results of the different models

According to the figure above, the model "caters" to more non-characteristic data points in the training set in order to meet the sufficiently low MSE of the training set, which often leads to overfitting phenomena, for example decision tree, GBDT, etc. Although the training set of these several models has a higher R^2, the test set MSE is too large, so it can be considered that the overfitting phenomenon has occurred. It shows that despite the model fits well to the training set, the test set predicts relatively poor results, so the study does not choose these models.

Too high training set MSE of some models can lead to under-fitting, such as KNN and SVM models, so these models have a weak generalization ability. To prevent overfitting, the Ridge regression model added L2 regular term to smaller the test set MSE. In other words, the Ridge regression model has a strong generalization ability, [31] and makes the MSE of both the training set and the test set lower,

which also makes it have a higher R^2 compared to the other models. That is, it has a strong explanatory power, and the comprehensive conditions also reach the optimal state. Therefore, it can be verified that the selected Ridge regression model is better than the other models when analysing the data.

3.4. Parameter inversion

3.4.1. No operation boundary iterations-solving the optimal operation parameters

In the previous fitting work of the Ridge regression model, obtained the model that can predict the retained RON according to the operating parameters. Therefore, the Gradient descent method can be used to iterate on the Ridge regression model, and to get more data needed for the engineering, such as the operating parameter scheme after inversion. Then, by comparing the predicted values after each iteration, the number of iterations required to reach the expected predicted value. In other words, solve the number of parameter inversions when the retained RON after the parameter inversion is closest to the RON contained in the raw material, and the optimal operating parameter scheme corresponding to the best retained RON can also be solved. At this time, the optimal strategy of the model is constantly adjusted without considering the actual value range of each parameter. The relationship between the number of iterations and the iteration effect is shown in Figure 7 and 8.

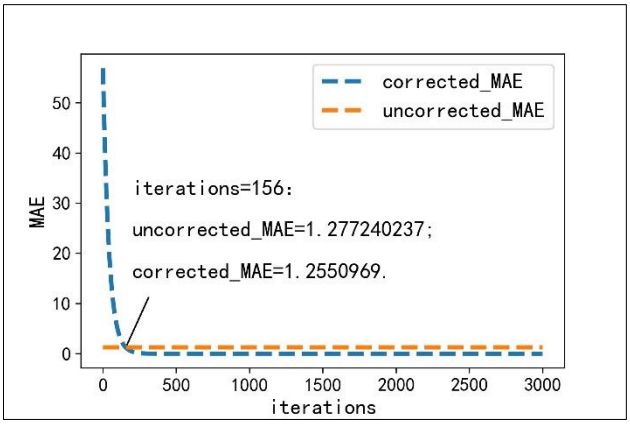


Figure 7. Relationship between the predicted loss value and the number of iterations for each case

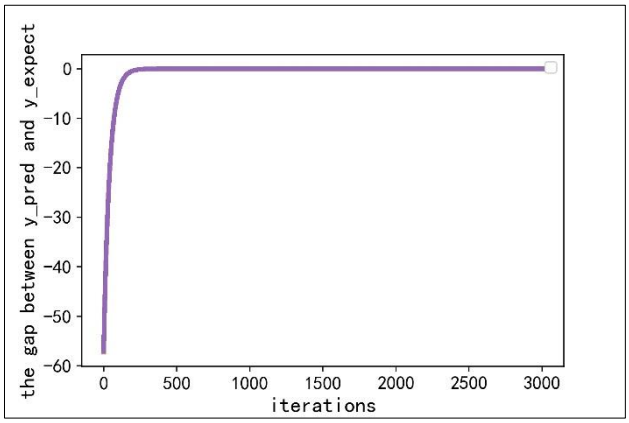


Figure 8. The difference between the expected values of the optimization strategy

Figure 7 shows that the MAE of the RON predicted value after parameter optimization is represented by the blue line, and the MAE of the actual RON of the original operation is represented by the orange line. In the absence of the value boundary of the operation parameters, when the number of iterations reaches 156 times, the MAE between the predicted value after parameter inversions and the raw material

RON starts to be less than that of the original operation. It is shown that, after 156 iterations, the RON lost during the refined gasoline process are less than that of the original operation scheme, and the former yields more retained RON. Therefore, it can be considered that the operation scheme after 156 inversion superior to the original operation scheme.

Figure 8 shows that the gap between predicted and expected RON values becomes smaller as the number of inversion increases. It is found that the gap between the predicted and expected values at the inversion number of 500 is near infinity, but is not equal to zero, indicating that the maximum retained RON can be almost be reached after 500 inversions. The optimized predicted values in the interval of 2,000 to 3,000 iterations have neither changed significantly nor fully reached the expected value. Therefore, the operation scheme at 2,000 iterations and the corresponding retained RON are taken as the optimal scheme.

In order to better observe the improvement effect after parameter optimization, the contrast effect of the actual RON of the original operation, the predicted RON after 2, 000 parameter inversions, and the product expected RON of each sample are presented on the line chart, as shown in Figure 9.

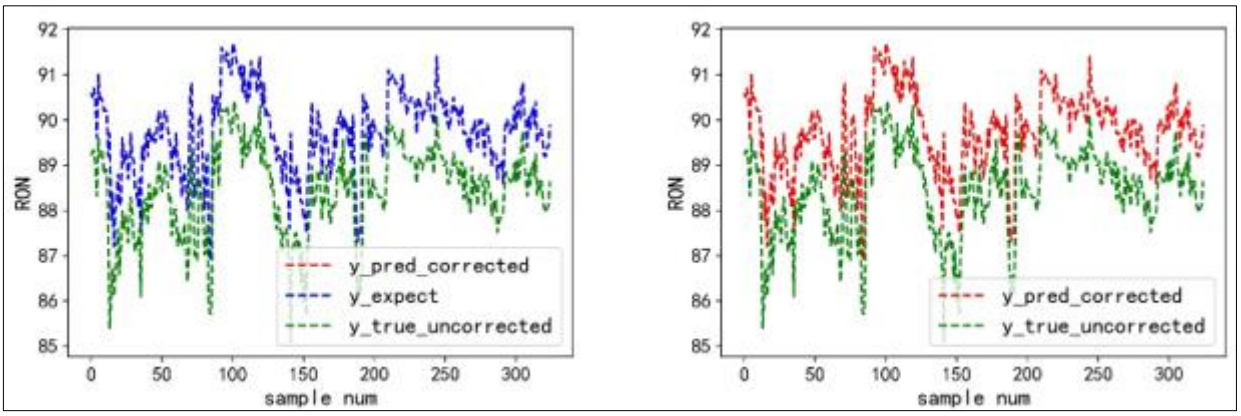


Figure 9. Comparison diagram of actual value, borderless optimization value, and expected value

From Figure 9, there is a clear difference between the green curve representing the actual RON of the original operation scheme and the blue curve representing the RON after the parameter inversion, but the blue curve almost completely overlaps with the red curve representing the predicted expected value, so it is not easy to find the red line. It shows that the predicted value after 2,000 inversions is significantly higher than the actual value of the original operating parameters, and also proves that using the operating parameters after 2,000 inversions can significantly improve the effect of the RON retention in the refined gasoline engineering. According to the python code calculation, the average RON loss of the original operation is 1.2772 units, and the average RON loss after 2,000 inversions is 0.01% of the lost RON value of the original operation scheme.

3.4.2. Limit the operation boundary to iterate-to solve the optimal operation parameters

In the previous step, the reasonable value boundary of operation parameters was not limited, only to pursue the target with the highest octane retention value. In practical engineering, the use of this thinking may bring adverse effects, such as safety risks, increased operation difficulty, and cost increase. Therefore, in the python code here, the parameter value range is added according to the actual requirements. As shown in Figure 10 and 11.

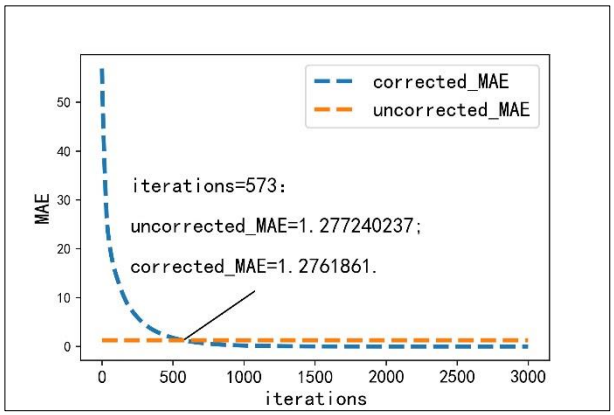


Figure 10. Relationship between the predicted loss value and the number of iterations for each case

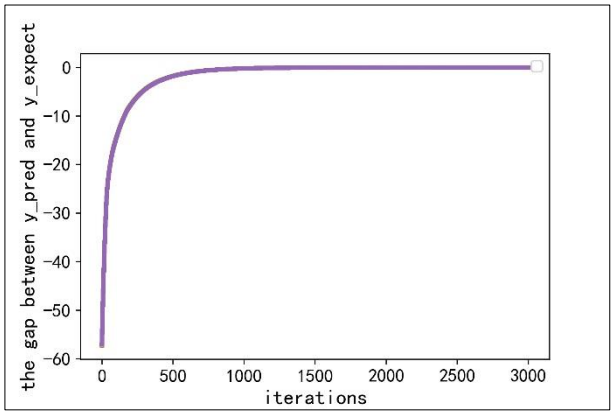


Figure 11. The difference between the predicted and expected values of the optimization strategy

Figure 10 shows that the MAE of the RON predicted value after parameter optimization is represented by the blue line, and the MAE of the actual RON of the original operation is represented by the orange line. In the absence of the value boundary of the operation parameters, when the number of iterations reaches 573 times, the MAE between the predicted value after parameter inversions and the raw material RON starts to be less than that of the original operation. It is shown that, after 573 iterations, the RON lost during the refined gasoline process are less than that of the original operation scheme, and the former yields more retained RON. Therefore, it can be considered that the operation scheme after 573 inversions superior to the original operation scheme.

It shows that, the case that limits the range of operating parameters requires nearly 400 more iterations than the case with unrestricted parameters, to reach the retained RON level of the original operation and to reach the steady state.

Figure 11 shows that the gap between predicted and expected RON values becomes smaller as the number of inversion increases. It is found that the gap between the predicted and expected values at the inversion number of 1,000 is near infinity, but is not equal to zero, indicating that the maximum retained RON can be almost be reached after 1,000 inversions. The optimized predicted values in the interval of 2,000 to 3,000 iterations have neither changed significantly nor fully reached the expected value. In order to facilitate the comparison with the unrestricted boundary case, the operation scheme at 2,000 inversions and its retained RON are still taken as the optimal scheme.

The contrast effect of the actual RON of the original operation, the predicted RON after 2, 000 policy inversions, and the product expected RON of each sample are also presented on the line chart, as shown in Figure 12.

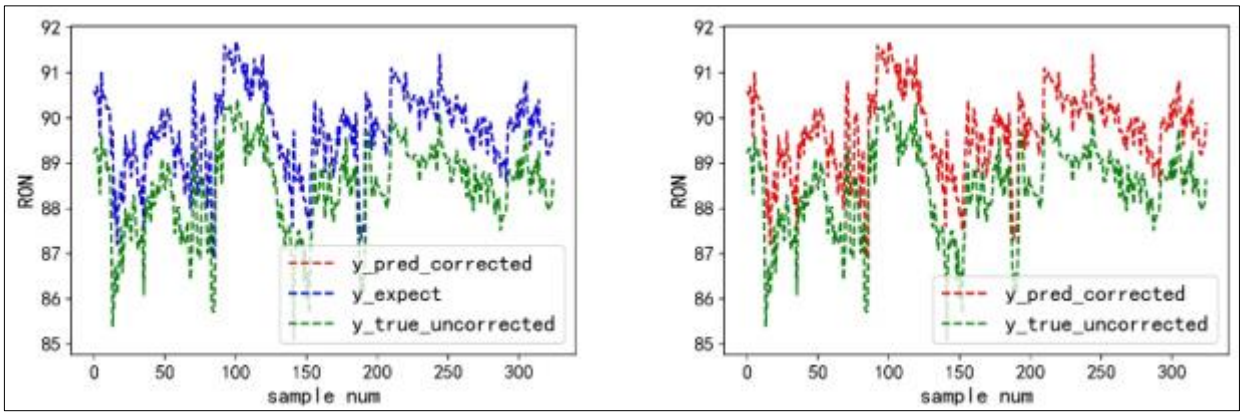


Figure 12. Comparison diagram of actual value, optimization value and expected value

From Figure 9 and 12, can find that the inversion effect of the limited operating parameter range is the same as that when the parameter range is not limited. The operating parameter scheme with 2,000 inversions can significantly improve the retention of RON in refined gasoline engineering, and its average loss of RON is only 0.01% of the lost RON of the original operation scheme.

3.4.3. Comparison of predicted values

In theory, in the presence or absence of operational boundaries, the operation parameter scheme after parameter inversion should be different, and the forecast values should be different. To further compare the iterative effect of the two cases, the predicted values are shown in the same graph, as shown in Figure 13.

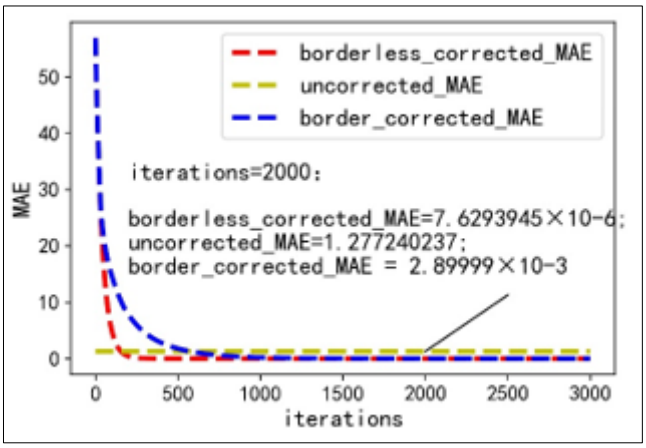


Figure 13. MAE comparison under different strategies

Figure 13, at 2,000 iterations, the MAE of the RON values for the original operating scheme, the scheme with a parameter range limit, and the scheme with no parameter range limit are 1.277240237, 2.89999×10^{-3} and 7.6293945×10^{-6}, respectively. It indicates that, the iterative RON values of these two cases are very similar, and are very close to the raw material RON values, and the previously calculated RON of the average loss can also prove these results. Therefore, the value range of the operating parameters of the project can meet the excellent iteration effect, and there is no need to expand the value range of the parameters.

4. Conclusions

The project conducts Ridge regression model modelling and Gradient descent inversion on the data in the 2020 Data Modelling Competition. The model for predicting the retained RON after cracking catalysis was found, and the model plays an important role in predicting the retained RON and optimizing the operation parameters for actual project.

The Leave-One-Out method and python code based on the Ridge regression model were used to pick out 146 operation parameters and the best hyper-parameter α equal to 0.6, and the R^2 of the training set can reach 98.8%. The MSE of the training set and test set of the model ware also obtained, and they were 0.037855 and 0.019342, respectively.

The calculated comparative fitting effect and R^2 of the Ridge regression model with other common models verify that the Ridge regression model has better fitting effect and interpretability than other models, and the model is the optimal model with comprehensive conditions. Therefore, can prove that the Ridge regression model is relatively reasonable in the literature.

In the case of setting and not setting parameter value range, parameter optimization was conducted by Gradient descent method, and found that the expected value has stabilized infinitely after 2,000 iterations. The corresponding MAE is $2.89999\times10^{-3}$and 7.62939×10^{-6}, respectively. The average RON loss of the optimal parameter schemes in both cases are all only 0.01% of the lost RON of the original operating scheme.

The literature only pursues the goal of RON maximization. In the future, can explore the optimal operating parameters and maximum RON in pursuing the maximization of sulfur element control and regeneration adsorbent recycling, and the direction has more social significance.

In the actual project engineering, businesses need to comprehensively consider the profit, safety and environmental protection to determine the operation parameters adopted. If the researchers only pursue the optimal output value without combining the actual engineering needs, though the optimization effect is very good, it is difficult to be applied by enterprises. However, because the database contains no relevant data about cost and profit, the optimal operating parameters and the maximum profit value under the profit maximization situation cannot be calculated.

The modelling thinking, judging the appropriate number of iterations, and Python code may be used or used in many industrial engineering, especially the engineering with a large number of operating parameters, and, nonlinearity and mutually strong coupling between the variables. And even some projects only need to replace the original data, and slightly modify the range of the upper and lower boundary value conditions, so the thinking provided in the literature is highly malleable.

Acknowledgments: The author wishes to acknowledge Dr. Yang Xiaojun, associate professor of Chinese, Wuhan Textile University, for his valuable suggestions for revising this research paper. We wish to thank the timely help given by Lyu Hairong in improving timely the method of the paper format.

References

1.ZHOU, B., GUAN, D.X., LIU, G.B., WANG, H.C., Properties and optimization of gasoline blending component. Liaoning Chemical industry, **51**(9), 2022, 1320-1322.
https://www.elsevier.com/data/promis_misc/BMCL Abbreviations.pdf
2.XU, Y.H., Progress of catalytic cracking technology in china. Scientia Sinica (Chimica), **44**(1), 2014, 13-24.
3.LIU, Y. C., LI, J., Analysis of factors affecting gasoline octane number loss in S Zorb unit. Petrochemical Design, **36**(4), 2019, 12-15+5.
4.ZHAO, L., LI, X., XIE, Y.F., YI, J.W., WU, J.F., HU, W.J., Prediction method of gasoline octane number based on adaptive variable weighting, Control and Decision, **37**(10), 2022, 2738-2744.
5.ZHU, X., JIANG, J.C., PAN, Y. and WANG, R., Prediction of octane number of alkanes based on support vector machine, Natural Gas Chemical Industry, **36**(3), 2011, 54-57.
6.AMARAL, L.V., SANTOS, N.D.S.A., ROSO, V.R., SEBASTIAO, R.C.O., PUJATTI, F.J.P., Effects of gasoline composition on engine performance, exhaust gases and operational costs, Renewable Sustainable Energy Rev., **135**, 2021, 110196. https://doi.10.1016/j.rser.2020.110196.
7.CHEN, Y.Z., HU, H., REN, Z.C., CHEN, A.G., Analysis of Octane Number Loss Model of FCC Gasoline Refining Unit Based on XGBoost and Improved Grey Wolf Optimization Algorithm, Acta Petrolei Sinica (Petroleum Processing Section), **38**(1), 2022, 208-219.

8.WANG, C.X., Study on heavy oil/bio oil co catalytic cracking in riser and its micro reaction mechanism, Beijing University of Chemical Technology, Beijing, 2019.
https://doi.10.26939/d.cnki.gbhgu.2019.000209.
9.LIU, Y.X., Structural Design of MeAPO-11 Molecular Sieve and Its Application to Catalytic Cracking Catalyst, China University of Petroleum (East China), Beijing, 2017.
https://doi.10.27644/d.cnki.gsydu.2017.000031.
10.HARDING, R.H., PETERS, A.W. NEE, J.R.D., New developments in FCC catalyst technology, Applied Catalysis A, General., **221**(1), 2001, 389-396.https://doi.10.1016/S0926-860X(01)00814-6.
11.SUN, D.L., XU, J.H., WEN, H.J., WANG, Y., An optimized random forest model and its generalization ability in landslide susceptibility mapping: Application in two areas of three gorges reservoir, China, J. Earth Sci., **31**(06), 2020, 1068-1086.
12.HAN, Q.J., ZOU, M., HUO, H.L., Prediction model of octane number loss based on real-time data of gasoline catalytic cracking process, Experimental Technology and Management, **39**(01), 2022, 41-45.https://doi.10.16791/j.cnki.sjg.2022.01.008.
13.JIANG, W., TONG, G.X., Construction and analysis of gasoline octane number loss prediction model based on improved PCA-RFR algorithm, Acta Petrolei Sinica (Petroleum Processing Section), **38**(01), 2022, 220-226.
14.QIN, Q.T., GU, H.H., Research on octane number loss based on vector autoregressive model, Software Engineering, **25**(09), 2022, 34-41.https://doi.10.19644/j.cnki.issn2096-1472.2022.009.008.
15.HODSON, H.T., Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev., **15**(14), 2022, 5481-5487.https://doi.10.5194/GMD-15-5481-2022.
16.TIAN, S.L., CHEN, T., TANG, M.N., YANG, L., Research on data de duplication method based on correlation coefficient and determination coefficient, Digital Manufacture Science, **17**(03), 2019, 241-244.
17.ZHANG, F.Y., SU, X.C., TAN, A.L., YAO, J.J., LI, H.P., Prediction of research octane number loss and sulfur content in gasoline refining using machine learningPrediction of research octane number loss and sulfur content in gasoline refining using machine learning, Energy, **261**(PA), 2022.
18.YAND, Y.N., YANG, REN, Y., MAO, A.G., TIAN, H.P., Analysis of Technical Factors Affecting the Change of Gasoline Octane Number in FCC Unit, Petroleum Refinery Engineering, **49**(06), 2019, 32-35.
19.ZHANG, Z.Y., LI, Z.Q., LI, Y.L., LI, G.Q., Prediction of FCC Unit Gasoline Yield by GA Assisted BP Neural Network, Petroleum Processing and Petrochemicals, **45**(07), 2014, 91-96.
20.WANG, J., CHEN, B., LIU, S., ZHAO, M.Y., OUANG, F.S., GAO, P., Optimization Model of Octane Number of S Zorb Refined Gasoline and Its Industrial Application, Petroleum Processing and Petrochemicals, **53**(05), 2022, 88-94.
21.ZHAO, H.S., WANG, Y.Y., SUN, A.M., Based on ARIMA model and 3 σ Detection method of water intake anomaly based on criterion, Water Resources Informatization, 2022, 35-41.
https://doi.10.19364/j.1674-9405.2022.01.008.
22.JIN, Z.J., LUO, X. J., TAO, Q., ZHAN, G.P., HE, Y., RO, X.Y., PAN, D.C., Study on endpoint determination method of tablet coating based on 3σ criteria and logic regression. Zhongguo Zhongyao Zazhi. **46**(16), 2021, 4124-4130.https://doi.10.19540/J.CNKI.CJCMM.20210208.301.
23.FENG, D.C., CHEN, F., XU, W.L., Efficient leave-one-out strategy for supervised feature selection, Tsinghua Sci. Technol., **18**(6), 2013, 629-635. https://doi.10.1109/TST.2013.6678908
24.LUOR, D.C., A comparative assessment of data standardization on support vector machine for classification problems, Intelligent Data Analysis, **19** (3), 2015, 529-546.
https://doi.10.3233/IDA-150730.
25.CHOI, S.H., JUNG, H.Y., KIM, H., Ridge Fuzzy Regression Model, International Journal of Fuzzy Systems, **21**(7), 2019, 2077-2090.https://doi.10.1007/s40815-019-00692-0.

26.YASSEN, M.F., ALDUAIS, F.S., ALMAZAH, M.M.A., Ridge Regression Method and Bayesian Estimators under Composite LINEX Loss Function to Estimate the Shape Parameter in Lomax Distribution, Computational intelligence and neuroscience, 2022, 2022, 1200611-1200611. https://doi.10.1155/2022/1200611.
27.HU, L.P., Ridge regression, Sichuan Mental Health, **31**(03), 2018, 193-196.
28.CHOI, S.H., BUCKLEY, J.J., Fuzzy regression using least absolute deviation estimators, Soft Computing, **12** (3), 2008, 257-263.https://doi.10.1007/s00500-007-0198-3.
29.CZAJKOWSKI, M., KRETOWSKI, M., Decision tree under-fitting in mining of gene expression data. An evolutionary multi-test tree approach, Expert Systems with Applications, **137** (C), 2019, 392-404.https://doi.10.1016/j.eswa.2019.07.019.
30.YANG, Y.L., LI, C.X., Quantitative analysis of the generalization ability of deep feedforward neural networks, Journal of Intelligent & Fuzzy Systems, **40**(3), 2021, 4867-4876. https://doi.10.3233/JIFS-201679.
31.SHAO, Z.F., ER, M.J., Efficient Leave-One-Out Cross-Validation-based Regularized Extreme Learning Machine, Neurocomputing, **194**, 2016, 260–270.https://doi.10.1016/j.neucom.2016.02.058.

Manuscript received: 18.11.2022