

## Research Article

# Research on the Rate of Penetration Prediction Method Based on Stacking Ensemble Learning

Yangfeng Ren <sup>1</sup>, Baoping Lu,<sup>2</sup> Shuangjin Zheng <sup>1</sup>, Kai Bai,<sup>3</sup> Lin Cheng,<sup>1</sup> Hao Yan,<sup>1</sup> and Gan Wang<sup>1</sup>

<sup>1</sup>School of Petroleum Engineering, Yangtze University, Wuhan, China

<sup>2</sup>State Key Laboratory of Shale Oil and Gas Enrichment Mechanism and Effective Development, Beijing, China

<sup>3</sup>School of Computer Science, Yangtze University, Jingzhou, China

Correspondence should be addressed to Shuangjin Zheng; zhengshuangjin@yangtzeu.edu.cn

Received 1 January 2023; Revised 14 February 2023; Accepted 27 February 2023; Published 14 March 2023

Academic Editor: Zhiyuan Wang

Copyright © 2023 Yangfeng Ren et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ROP is an important index to evaluate the efficiency of oil and gas drilling. In order to accurately predict the ROP of an oilfield in Xinjiang working area, a ROP prediction model based on the historical drilling data of this working area was established based on stacking ensemble learning. This model integrates the  $K$ -nearest neighbor algorithm and support vector machine algorithm by stacking ensemble strategy and uses genetic algorithm to optimize model parameters, forming a new method of ROP prediction suitable for this oilfield. The prediction results show that the accuracy of ROP prediction by this method is up to 92.5%, and the performance is stable, which can provide reference for the optimization of drilling parameters in this oilfield and has specific guiding significance for improving the efficiency of drilling operations.

## 1. Introduction

Rate of penetration (ROP) is an essential technical index to evaluate the efficiency of drilling operations. Predicting ROP can provide a basis for drilling decisions and optimization of construction parameters to reduce cost and increase efficiency. For ROP prediction, most scholars at home and abroad use traditional methods to establish equations for prediction [1]. For example, Bahari and Seyed [2] established the W.C. Maaurer equation, but the consideration factors need to be more comprehensive. Adebayo and Akande [3] established an empirical formula between rock characteristics, physical properties and ROP, but did not consider the influence of drilling parameters on ROP. Changsheng [4] improved the B-Y ROP equation based on the multiple regression method so that the coefficients of different formations could be changed, but this equation only applied to mud drilling. Kumar and Murthy [5] obtained a more accurate ROP prediction model, but the model parameters were significantly affected by lithology and were relatively complex. Hung et al. [6] summarized the ROP prediction model

for rotary percussion bits, but this model mainly applies to high-hardness formations. Jing et al. [7] established a ROP prediction model using analytic hierarchy process and feed-forward neural network, which provided help for optimizing bit parameters. Wang et al. [8] established a ROP prediction model based on neural network algorithm, which can improve development efficiency and reduce development costs. Chiranth et al. [9] preliminarily established the ROP prediction model using machine learning algorithm and optimized it using random search method, particle swarm optimization algorithm, and eyeball method. Based on the drilling data of a directional well in Changqing Oilfield, Liu et al. [10] established the ROP prediction model using BP neural network with high calculation accuracy. Xinghua et al. [11] established the ROP prediction model using four machine learning algorithms, such as the  $K$ -nearest neighbor algorithm and support vector machine. The prediction results showed that the model based on the lifting tree algorithm had the highest accuracy. Liu et al. [12] established a dynamic prediction model of gray-weighted Markov mechanical drilling rate, which is based on the grey system

theory and combined with Markov theory, to realize dynamic real-time prediction.

Through investigation and analysis, it is found that the existing ROP prediction mainly relies on experience or improves on the equations established by predecessors or tries to find the main control factors affecting ROP by means of controlling variables, and most of the machine learning algorithm models for predicting ROP described above use a single model or a single model with optimized parameters, which will inevitably lead to large errors and cannot meet the current drilling construction needs. In order to solve the above problems, this paper proposes a combined prediction model which combines the  $K$ -nearest neighbor algorithm and support vector machine algorithm with stacking integration strategy to predict ROP and optimizes it with genetic algorithm. Among them, the stacking ensemble learning algorithm is a classic algorithm among heterogeneous ensemble learning algorithms. The higher the difference of primary learners in the model, the better the performance of the model. Therefore, before modeling, we should not only analyze the prediction effect of each primary learner but also consider the differences between primary learners. The  $K$ -nearest neighbor algorithm is an inert learning mechanism, which is widely used because of its efficient training method and good classification effect. The training time complexity of support vector machine algorithm is high, but its classification accuracy is high, which makes it achieve good results in practical application. Combined with the historical drilling data of an oil field in Xinjiang working area, the prediction accuracy can reach 92.5%, which meets the engineering design requirements. It is helpful to guide the optimization of drilling parameters and improve drilling efficiency.

## 2. Rate of Penetration Prediction Algorithm

**2.1.  $K$ -Nearest Neighbor Algorithm.** The  $K$ -nearest neighbor [13–15](KNN) is a semisupervised machine learning algorithm with mature theoretical basis and good application stability. The core of the algorithm is first to measure the distance between the new sample and the classified samples in the training data set and then classify the new sample according to most of the  $K$  neighbors closest to the new sample. The basic process is as follows: (1) using the historical data set, build a database with large capacity and particular representative; (2) set the state vector; (3) calculate the distance between the new state vector and the historical data state vector; (4) determine the value of the number of the nearest neighbors  $K$ ; (5) determine the nearest neighbor weight; and (6) the predictive value of the algorithm can be obtained by weighting  $K$  neighbors. In Figure 1, “?” represents samples of unknown categories, “triangle” and “circle” are marked as samples of known categories, and  $K$  of the KNN algorithm is set to 5; then, the cluster needs to contain samples of 5 known categories. Obviously, the number of “triangle” categories is greater than the number of “circle” categories, so it can judge the “?” is the “triangle” category by voting. The actual application of the algorithm is often

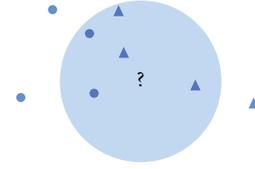


FIGURE 1: Example of  $K$ -nearest neighbor.

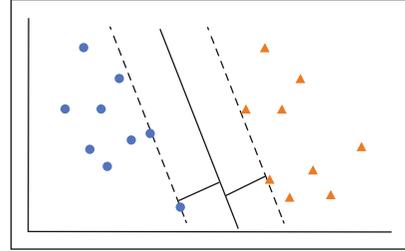


FIGURE 2: Optimal “hyperplane” in two-dimensional feature space.

more complex than this example, and the dimensions of the data will be higher.

**2.2. Support Vector Machine Algorithm.** The main ideas of support vector machine (SVM) algorithm [16–19] are structural risk minimization principle and VC dimension theory of statistical learning theory. The SVM algorithm has many unique advantages in solving small sample and nonlinear regression models. It overcomes the problems of “dimension disaster” and “over-learning” to a large extent and can be used to classify discrete dependent variables and predict continuous dependent variables.

The principle of the algorithm is shown in Figure 2, where triangles are labeled as category 1, origin is labeled as category 2, and support vectors are all on the dashed lines. SVM uses kernel function to transform linear inseparable problems into linear separable problems. In the case of linear inseparable problems, nonlinear transformation is needed. After transformation, the equation for dividing the hyperplane is

$$f(x) = \omega^T \varphi(x) + b, \quad (1)$$

where  $\omega$  is the normal vector,  $b$  is the displacement term, and  $\varphi(x)$  is the mapping transformation of  $x$ .

In order to maximize the distance between the found hyperplane and different categories, the following requirements should be met:

$$\begin{aligned} \min_{\omega, b} \quad & \frac{1}{2} \|\omega\|^2 \\ \text{s.t.} \quad & y_i (\omega^T \varphi(x) + b) \geq 1 \quad i = 1, 2, \dots, m. \end{aligned} \quad (2)$$

The hyperplane equation is obtained by quadratic programming and kernel function:

$$f(x) = \sum_{i=1}^m \alpha_i y_i k(x, x_i) + b, \quad (3)$$

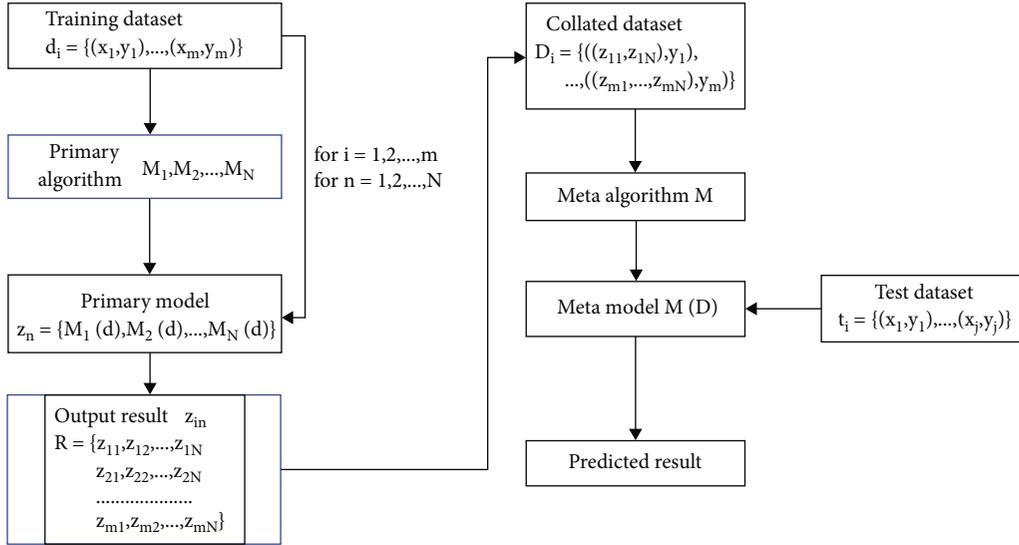


FIGURE 3: Operating principle of the stacking model.

where  $\alpha_i$  is the weight coefficient,  $y_i$  is the truth value,  $k$  is the kernel function,  $k(x, x_i)$  represents  $\phi(x) \cdot \phi(x_i)$ , and  $\phi(x)$  represents the mapping transformation with respect to  $x$ .

**2.3. Stacking Algorithm.** The stacking algorithm [20–23] is a combination strategy algorithm in ensemble learning. The algorithm requires multiple learners to support its training process. These learners are divided into primary learners and meta-learners, the former training individual data, and the latter training combined data. As a heterogeneous integrated learning technology, the stacking algorithm can grasp the advantages of different prediction models, infinitely superpose primary learners through superposition, and use meta-learners to optimize the combination of different primary learners. Therefore, compared with a single prediction model, the stacking algorithm has better performance, stronger robustness, and generalization. This paper uses KNN and support vector machine algorithm as primary learners in stacking ensemble learning. The specific operating principle of stacking algorithm is shown in Figure 3.

It can be seen from Figure 3 that the new training data is obtained based on the training of the primary model. Overfitting will occur if the original training data is directly used to train the meta-model. Therefore,  $K$ -fold cross-validation method should be used to generate the training data needed for the subsequent establishment of the meta-model with the samples not used by the primary model.

The main steps of ROP prediction based on Stacking ensemble learning are as follows: (1) the original data of drilling blocks are cleaned and normalized. Then the normalized data set is divided into training set and test set. (2) The training set is divided into  $K$  parts, and the  $K$ -fold cross-validation method is used for cross-validation of each primary learner. In each cross-validation, 1 part is used as the validation set of the model, and the remaining ( $K-1$ ) parts serve as the training set. After each cross-validation, the trained learner predicts the test set and the validation set. (3) The predicted values of each validation set obtained

by a single primary learner after  $K$  cross-validation are integrated into matrix  $M$ , and matrix  $Q$  is obtained by averaging the  $K$  predicted values of the test set in rows. (4) The output feature matrix  $(M_1, M_2, \dots, M_n)$  of the training set can be obtained after the training of the primary learner in the first part, and this matrix is used as the training set of the meta-learner in the second part. (5) In the second part, after the training of the metalearner, the matrix  $(Q_1, Q_2, \dots, Q_n)$  is taken as the test set, and the predicted value of the model is finally output.

As a model fusion algorithm, stacking ensemble learning can achieve better performance in many practical problems than a single algorithm or other simple ensemble learning algorithms, especially in processing classification problems. Stacking ensemble learning has been widely applied in classification, recognition, prediction, and other problems [21–24].

### 3. Model Establishment

**3.1. Data Preprocessing.** The modeling data used in this paper are all from an oil field in Xinjiang working area. In order to avoid the high dimension of the original data set affecting the prediction results, the max–min method in the standardized method is used to process the 400 groups of original data, and the max–min method is shown in

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad (4)$$

where  $x'_i$  is the normalized data,  $x_i$  is the original sample data before normalization,  $x_{\min}$  is the minimum eigenvalue of the original sample data, and  $x_{\max}$  is the maximum eigenvalue of the original sample data.

**3.2. Rate of Penetration Classification.** Analysis of ROP data from this field shows that ROP varies significantly due to different bit sizes. According to field experience, the ROP

TABLE 1: Classification of ROP corresponding to different bit sizes in an oil field in Xinjiang working area.

Bit diameter/mm	Low ROP range/(m/h)	Medium ROP range/(m/h)	High ROP range/(m/h)
149.2	<1.5	1.5-2.5	>2.5
215.9	<4	4-6	>6
311.2	<10	10-15	>15
444.5	<30	30-40	>40

TABLE 2: Confusion matrix of results of  $K$ -nearest neighbor algorithm for predicting ROP grade.

Confusion matrix (KNN)	True value		
	1	2	3
Predictive value 1	18	2	3
Predictive value 2	2	2	0
Predictive value 3	3	4	4

TABLE 3: Confusion matrix of support vector machine algorithm for predicting ROP grade.

Confusion matrix (SVM)	True value		
	1	2	3
Predictive value 1	21	3	2
Predictive value 2	1	3	1
Predictive value 3	1	2	4

TABLE 4: Confusion matrix of prediction results of the stacking integrated model ROP grade.

Confusion matrix (stacking)	True value		
	1	2	3
Predictive value 1	22	2	1
Predictive value 2	0	3	1
Predictive value 3	2	1	6

TABLE 5: Default model prediction results.

Model	Accuracy of classification
KNN	63.2%
SVM	73.7%
Stacking	81.6%

of different sizes of drill bits is divided into three grades: low ROP, medium ROP, and high ROP. The ROP classification makes the value of ROP independent of the size of the drill bit and transforms the modeling task from regression problem to classification problem, which is beneficial to improve the accuracy of model prediction. The rate of penetration classification standard of this field is shown in Table 1.

3.3. *Establish the Prediction Model of ROP.* In order to avoid the problem of insufficient generalization ability of the prediction model and ensure the accuracy and stability of the

TABLE 6: Confusion matrix of optimized  $K$ -nearest neighbor algorithm for predicting ROP level.

Confusion matrix (KNN)	True value		
	1	2	3
Predictive value 1	21	2	3
Predictive value 2	0	3	1
Predictive value 3	2	3	4

TABLE 7: Confusion matrix of optimized support vector machine algorithm for predicting ROP level.

Confusion matrix (SVM)	True value		
	1	2	3
Predictive value 1	22	3	3
Predictive value 2	0	4	1
Predictive value 3	1	1	4

TABLE 8: Confusion matrix of optimized Stacking integration model for predicting ROP level.

Confusion matrix (stacking)	True value		
	1	2	3
Predictive value 1	23	0	1
Predictive value 2	0	7	1
Predictive value 3	0	1	5

TABLE 9: Model prediction results optimized by genetic algorithm.

Model	Accuracy of classification
KNN	73.7%
SVM	78.9%
Stacking	92.5%

prediction results, 380 sets of original data obtained in the field were randomly divided into training set and test set according to the ratio of 9:1, among which 342 training samples were used to construct the model and 38 test samples were used to evaluate the model.

The rate of penetration prediction model with default parameters was established based on the KNN algorithm, SVM algorithm, and stacking algorithm, with the input values of drilling weight, drilling speed, rotation speed, drilling fluid displacement, density, funnel viscosity, bit pressure drop, and bit diameter in the sample data and the output

TABLE 10: Comparison of predicted ROP data and actual ROP data.

(a)

Bit tooth penetration freshness (%)	Bit tooth out freshness (%)	Bit tooth wear (%)	Bit pressure (kN)	Speed (r/min)	Drilling fluid circulation rate (L/s)
100	50	50	80	30	24
100	90	10	40	50	50
100	95	5	30	63	16

(b)

Drilling fluid density (g/cm <sup>3</sup> )	Funnel viscosity of drilling fluid (s)	The pressure drop of the bit (MPa)	The bit diameter (mm)	Actual average ROP (m/h)	Predicted average ROP (m/h)
1.6	57	0.53	215.9	3.8 (low)	<4 m/h
1.12	51	1.31	311.2	23.42 (high)	>15 m/h
1.2	51	0.13	149.2	2.16 (medium)	>1.5 m/h and <2.5 m/h

values of ROP grade. The default parameters of each model are as follows: the number  $K$  of KNN neighbor samples is 5, the voting weight of neighbor samples is the same, and the distance standard adopts Euclidean distance. The penalty coefficient  $C$  of SVM is set to 1, the kernel function  $g$  is set to  $1/n\_features$ , and the kernel function adopts Gaussian kernel function. Building the stacking integrated learning model is divided into two stages. First, the primary learner (primary model) is trained by the stacking algorithm using the training dataset. Then, a new training dataset is generated based on the primary learner to train the metalearner (metamodel).

Considering that this problem is a classification problem, the accuracy rate is used as the evaluation method of model classification accuracy. The specific idea is as follows: classification accuracy = number of accurate classification groups/total number of predicted groups. The total number of groups in this prediction is a fixed value of 38 groups. The prediction results of the default model are shown in Tables 2–5, where “low ROP is 1, medium ROP is 2, and high ROP is 3.”

It can be seen from Tables 2–5 that the classification accuracy of KNN, SVM, and stacking models for predicting ROP is low. The classification prediction accuracy of the integrated stacking model is the highest, but only 81.6%. Therefore, the algorithm parameters need to be optimized to improve the classification prediction accuracy of the model.

**3.4. Optimization Model.** The genetic algorithm [25–31] (GA) is an efficient random search and optimization method based on natural genetic mechanism and biological evolution. In the solving process, this algorithm simulates the evolutionary mechanism of the biological world and then obtains the optimal solution through the operation of chromosome selection, crossover, and mutation. The core of this algorithm is to evaluate the merits of individuals through the fitness function. The dominant individuals will occupy a higher proportion of the next generation population. Com-

pared with traditional optimization algorithms, the genetic algorithm has strong robustness. In addition, global search characteristics and implied parallelism are also two significant advantages.

Genetic algorithm optimization steps are as follows: (1) determine input and output variables, and transform the optimal solution problem into population reproduction problem by coding (binary coding is adopted in this paper); (2) randomly generate individuals and set populations; (3) determine the fitness function; (4) determine selection, crossover, and mutation operations; (5) calculate the fitness value and eliminate it; (6) if the iteration stop condition is met, the optimal solution can be obtained; otherwise, the operation of steps (4)–(6) can be continued.

In this paper, the parameters of the genetic algorithm are set as follows: the population number is 150, the maximum reproduction number is 70, the chromosome length is 100, the crossover probability is 0.85, the mutation probability is 0.01, and the absolute error is selected as the fitness function. The number  $K$  of KNN neighbor samples after optimization is 30. The voting weight of neighbor samples is inversely proportional to the distance, and the distance standard is Manhattan distance. The penalty coefficient of the optimized SVM is 12, the kernel function parameter is 0.1, and the kernel function adopts Gaussian kernel function.

The optimized model parameters were input into the corresponding model, and all the optimized parameters were fused with the stacking ensemble algorithm. The precision and generalization of the optimized model were verified with 40 groups of test data. The prediction results of each optimized model are shown in Tables 6–9, where “low ROP is 1, medium ROP is 2, and high ROP is 3.”

The prediction results of ROP level showed that the accuracy of each model was improved after the optimization of genetic algorithm. The classification prediction accuracy of the stacking integrated model was the highest, which was 92.5%. The stacking integrated model is the most suitable for predicting ROP in this oilfield and provides a reference for optimizing drilling construction parameters.

#### 4. Prediction of Rate of Penetration of an Oilfield in Xinjiang

The rate of penetration prediction software was prepared based on the stacking model optimized by the genetic algorithm. The construction parameters of the deep well section of an oilfield in Xinjiang working area were entered as follows: the bit entering degree is 100%, the bit out degree is 50%, the bit tooth wear is 50%, the bit weight is 80 kN, the speed is 30 r/min, the bit pressure drop is 0.53 MPa, the bit diameter is 215.9 mm, the drilling fluid density is 1.6 g/cm<sup>3</sup>, the drilling fluid circulation displacement is 24 L/s, and the drilling fluid funnel viscosity is 57 s. Software calculation results are as follows: the prediction of ROP level under this construction parameter is “1-low speed, ROP less than 4 m/h.” In the same way, ROP prediction data under the other two construction parameters can be obtained, as shown in Table 10.

Table 10 shows that ROP predicted by the stacking integrated model in the oilfield is consistent with the actual ROP. Confirming that the ROP prediction method based on stacking ensemble learning can be used to predict the field of ROP level and scope. Field operation personnel can adjust construction parameters based on the prediction results and the dynamic optimization of the drilling construction plan to improve the rate of penetration and drilling efficiency.

#### 5. Conclusion

- (1) According to the original data of an oil field in Xinjiang working area, data cleaning and processing were carried out. KNN, SVM, and stacking algorithm were used to establish the ROP prediction model, and the prediction accuracy was as follows: 63.2%, 73.7%, and 81.6%. The highest accuracy of the stacking ensemble model was 81.6%, which needs to be further optimized
- (2) The parameters of each model are optimized by the genetic algorithm. The optimized accuracies of KNN, SVM, and stacking models for predicting ROP are as follows: 73.7%, 78.9%, and 92.5%. The prediction accuracy of the stacking integrated model is the highest, which can be used for ROP prediction and drilling construction parameter optimization in an oilfield in Xinjiang working area
- (3) The ROP prediction software suitable for an oil field in Xinjiang was developed based on the integrated learning model optimized by genetic algorithm. The field operators can adjust the construction parameters based on the prediction results and dynamically optimize the drilling construction scheme
- (4) Due to the limited data available, only a few features mentioned in this paper can be considered factors affecting the ROP. More detailed features such as the specific composition of drilling fluid and formation parameters can be considered in the future

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### Acknowledgments

This work was supported by grants from the Fund “Research on the Method of Rate of Penetration Prediction and Construction Parameter Optimization Based on Machine Learning,” the Open Fund “Hubei Key Laboratory of Drilling and Production Engineering for Oil and Gas (Yangtze University),” and the Open Foundation of Cooperative Innovation Center of Unconventional Oil and Gas, Yangtze University (Ministry of Education & Hubei Province) (No. UOG2022-06). Here, I would like to express my deep gratitude.

#### References

- [1] A. Bahari and A. B. Seyed, “Trust-Region Approach To Find Constants of Bourgoyne and Young Penetration Rate Model in Khangiran Iranian Gas Field,” in *SPE Latin American and Caribbean Petroleum Engineering Conference (LACPEC 2007)*. Society of Petroleum Engineers (SPE), pp. 1–6, Buenos Aires, Argentina, April 2007.
- [2] Y. Lin, Y. Zong, Z. Liang, T. Shi, and R. Li, “Research progress on ROP of petroleum drilling machines,” *Petroleum Drilling Technology*, vol. 32, no. 1, pp. 10–13, 2004.
- [3] B. Adebayo and J. M. Akande, “Textural properties of rock for penetration rate prediction,” *Journal of Science and Technology*, vol. 6, no. 1, pp. 1–8, 2011.
- [4] L. Changsheng, “Study of method for predict rate of penetration based of multiple regression analysis,” *Science Technology and Engineering*, vol. 13, no. 7, pp. 1740–1744, 2013.
- [5] S. Kumar and V. M. Murthy, “Experimental studies on drill penetration rate prediction in coal measure rocks through Cerchar hardness index tests,” *Journal of Mines Metals & Fuels*, vol. 62, pp. 88–95, 2014.
- [6] N. V. Hung, L. Gerbaud, R. Souchal, C. Urbanczyk, and C. Fouchard, “Penetration rate prediction for percussive drilling with rotary in very hard rock,” *Journal of Science and Technology*, vol. 54, no. 1, p. 133, 2016.
- [7] N. Jing, H. Fan, R. Ji, Y. Zhai, and T. Liu, “Data mining technology-based research on the prediction method of deep-well ROP,” *China Petroleum Machinery*, vol. 40, no. 7, pp. 17–20, 2012.
- [8] W. Wang, X. Liu, P. Dou, H. Lin, and L. Tao, “A ROP prediction method based on neutral network for the deep layers,” *Oil Drilling & Production Technology*, vol. 40, no. S1, pp. 121–124, 2018.
- [9] D. H. ChiranthH and K. E. Gray, “Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models [J],” *SPE Journal*, vol. 23, no. 5, pp. 1706–1722, 2018.
- [10] S. Liu, J. Sun, X. Gao, and M. Wang, “Analysis and establishment of drilling speed prediction model for drilling machinery

- based on artificial neural networks,” *Computer Science*, vol. 46, no. 6A, pp. 605–608, 2019.
- [11] S. Xinghua, S. Junming, G. Xiang, and W. Min, “Prediction method of rate of penetration of penetration based on GBDT algorithm,” *Computer Applications and Software*, vol. 36, no. 12, pp. 87–92, 2019.
- [12] J. Liu, G. Yan, X. Guo, M. Wan, M. Luo, and H. Y. Zhu, “Dynamic prediction of drilling rate based on grey-weighted Markov,” *Journal of Systems Science and Mathematical Sciences*, vol. 42, no. 7, pp. 1727–1739, 2022.
- [13] T. M. Cover and P. E. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1953.
- [14] Z. Zhang, Y. Huang, and H. Wang, “A new KNN classification approach,” *Computer Science*, vol. 35, no. 3, pp. 170–172, 2008.
- [15] N. Zhang, Z. Jia, and Z. Shi, “Text categorization with KNN algorithm,” *Computer Engineering*, vol. 31, no. 8, 2005.
- [16] S. Ding, B. Qi, and H. Tan, “An overview on theory and algorithm of support vector machines,” *Journal of University of Electronic Science and Technology of China*, vol. 40, no. 1, pp. 2–10, 2011.
- [17] X. Zhang, “Introduction to statistical learning theory and support vector machines,” *Acta Automatica Sinica*, vol. 26, no. 1, pp. 32–42, 2000.
- [18] H. Qi, “Support vector machines and application research overview,” *Computer Engineering*, vol. 30, no. 10, pp. 6–9, 2004.
- [19] G. Feng, “Parameter optimizing for support vector machines classification,” *Computer Engineering and Applications*, vol. 47, no. 3, pp. 123–124, 2011.
- [20] H. Li, J. Zhu, and T. Yao, “Combined multiple classifiers based on a stacking algorithm and their application to Chinese text chunking,” *Journal of Computer Research and Development*, vol. 42, no. 5, pp. 844–848, 2005.
- [21] X. Huili, *The Study and Improvement of Stacking*, South China University of Technology, 2018.
- [22] X. Zhou, L. Ding, R. Wan, and Q. Ge, “Research on classifier ensemble algorithms,” *Journal of Wuhan University (Natural Science Edition)*, vol. 61, no. 6, pp. 503–508, 2015.
- [23] S. Li and R. Huang, “Chinese sentiment classification based on stacking combination method,” *Journal of Chinese Information Processing*, vol. 24, no. 5, pp. 56–61, 2010.
- [24] J. Gong, *The Study of Loan-Decision Model Based on Optimized and Fusion Stacking Algorithm*, Shanxi Normal University, 2019.
- [25] J. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, USA, 1975.
- [26] Y. Ma and W. Yun, “Research progress of genetic algorithm,” *Application Research of Computers*, vol. 29, no. 4, 2012.
- [27] X. Bian and L. Mi, “Development on genetic algorithm theory and its applications,” *Application Research of Computers*, vol. 27, no. 7, 2010.
- [28] G. Ji, “Survey on genetic algorithm,” *Computer Applications and Software*, vol. 21, no. 2, pp. 69–73, 2004.
- [29] W. Q. Fu, Z. Wang, J. Zhang, and B. Sun, “Methane hydrate formation in a water-continuous vertical flow loop with xanthan gum,” *Fuel*, vol. 265, article 116963, 2020.
- [30] W. Q. Fu, Z. Y. Wang, L. T. Chen, and B. Sun, “Experimental investigation of methane hydrate formation in the carboxymethylcellulose (CMC) aqueous solution,” *SPE Journal*, vol. 25, no. 3, pp. 1042–1056, 2020.
- [31] W. Q. Fu, J. Yu, Y. Xiao, C. Wang, B. Huang, and B. Sun, “A pressure drop prediction model for hydrate slurry based on energy dissipation under turbulent flow condition,” *Fuel*, vol. 311, article 122188, p. 122188, 2022.