

Study On Rate of Penetration Prediction Model of Offshore Drilling Machinery and Structure Optimization Base on Deep Learning

Tong Jiao

College of Petroleum Engineering, Xi'an Shiyou University, Xi'an 710065, China

Abstract: With the development of modern downhole monitoring technologies such as measurement while drilling, the data obtained during drilling has the characteristics of huge volume, rich variety, and strong timeliness, which greatly promotes big data and artificial intelligence in drilling rate prediction from the data level. Progress. How to further effectively excavate and use drilling big data, improve drilling efficiency, and reduce drilling risks is still a hot research topic. This article explores the real-time drilling data of an exploratory well in the South China Sea, and establishes a reasonable machine learning model to accurately predict the ROP and its changing trend. The research first cleans the real-time drilling data, and improves the data quality used for modeling through standardized data processing; then establishes a deep neural network, optimizes input and output parameters, trains the model, and validates the prediction results; finally, optimizes the model structure to improve prediction The accuracy and efficiency of the model. The research results show that when the amount of data is sufficient and the missing values are few, the neural network can more accurately predict the ROP and its change trend, and at the same time, reasonable optimization of the model can also improve the accuracy and calculation efficiency of the drilling rate prediction.

Keywords: Drilling real-time data, Deep neural network, ROP prediction, Model structure optimization.

1. Introduction

Raising drilling speed is one of the goals pursued continuously in drilling engineering and operations. However, China's current drilling efficiency still lags behind the advanced level abroad. According to statistics, at present, a domestic drilling rig can only drill 2-3 horizontal wells a year, while a drilling rig in North America can complete 15-20 horizontal wells a year. With the large-scale development of domestic offshore oil and gas and unconventional oil and gas resources, the contradiction of drilling rig shortage has become increasingly prominent. Drilling efficiency has become a bottleneck problem that seriously restricts the rapid scale production of domestic oil and gas resources. Therefore, during the drilling process, how to effectively reduce the drilling cycle, improve the drilling efficiency and reduce the drilling cost has become a key hot issue in oil and gas exploration and development [1]. The penetration rate is the key factor to measure the drilling efficiency [2], so it is one of the hot spots in the drilling field to determine the penetration rate by modeling the influence of various variables.

With the continuous emergence of new drilling tools, sensors and downhole acquisition tools in recent years, downhole intelligent measurement systems are becoming more and more complex, and the type and scale of data obtained are growing exponentially; However, these data are not completely independent, and there are very complex internal links between them. In addition, the complexity, diversity and uncertainty of underground environment and geological parameters. Among these huge data, the factors that affect the penetration rate are also very complex and difficult to analyze, making the traditional data processing model unable to effectively process and analyze some data. The traditional physical model determines the importance of drilling parameters through analytical method, but it cannot

cover the influence between parameters [3]. Soares [4] et al. revealed several limitations of using analytical equations to model the penetration rate of machinery. Therefore, people begin to use data-driven models in the field of artificial intelligence. Taking offshore drilling as an example, due to the complexity of conditions and the huge amount of drilling data, the penetration rate equation constructed by a single factor has been difficult to meet the needs of modern offshore drilling, and the development of artificial intelligence has made long-term progress in exploring drilling intelligence. Scholars at home and abroad are exploring the application of big data and deep learning in petroleum engineering, and building real-time multi factor intelligent drilling equations based on data drive. Abdulmalek [5] and others demonstrated the feasibility of artificial neural network for drilling data prediction and analyzed the impact of multiple factors on drilling speed. Li Qi and others optimized the back-propagation network with the Tenebrus algorithm, particle swarm optimization algorithm and genetic algorithm, which greatly improved the accuracy of the mechanical penetration rate prediction model. A comparative study of the differences between machine learning methods and physics based models proves that machine learning methods have higher accuracy [6-8]. Hadi et al. evaluated the role of multiple regression analysis and artificial neural network in ROP prediction [9]. Aliyev and Paul also used the artificial neural network model, but they integrated both downhole and surface parameters in their research [10].

At present, scholars at home and abroad have established a variety of models for ROP prediction, but most of the models remain in a single middle layer network structure, which can not effectively represent the nonlinear law of the drilling process; At the same time, the existing research lacks of research on model structure and model parameter optimization, which makes most of the drilling speed prediction models have a single structure and low efficiency.

Therefore, the depth neural network model is selected for full demonstration in this paper, and the model structure is analyzed at multiple levels and nodes. By comparing the operation effect and operation efficiency of different structural models, the highest efficiency penetration rate prediction model is finally selected to provide support for on-site testing and application.

2. Introduction to Deep Neural Network

Deep neural network (DNN) is one of the most commonly used methods for deep learning. Its most typical implementation is a multi-layer perceptron, which is generally composed of an input layer, multiple hidden layers and an output layer through full connection. A typical four layer neural network structure is shown in Figure 1.

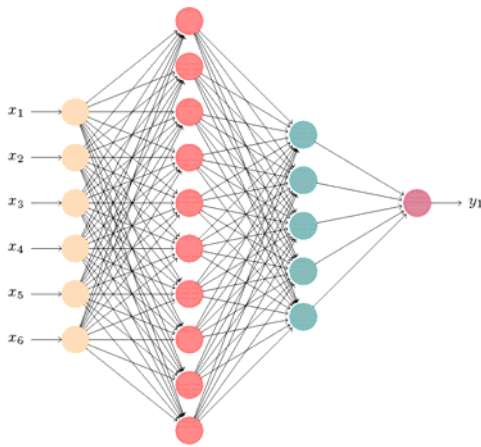


Figure 1. Schematic Diagram of Deep Neural Network Structure

After the input layer obtains the data, there are weights w and offsets b in each layer. $F(wX+b)$ is calculated through the

activation function and input to the next layer. During each training, MLP chooses to use Adam algorithm to optimize w and b to meet all training data as much as possible. The initial w and b are randomly generated groups. As shown in Figure 1, the output result of the four layer perceptron is

$$H(x) = f\left(b_3 + W_3 f_2\left(b_2 + W_2\left(f_1(w_1 x + b_1)\right)\right)\right)$$

The application paradigm of deep learning usually includes three parts: data collection and preprocessing, model training, testing and verification, and model deployment and application. Although BP neural network with single hidden layer is widely used, it is difficult to ensure the effect of feature extraction when there are many input parameters. Therefore, for complex drilling real-time data, this paper selects the depth neural network model (DNN) to study data set automatic processing and mechanical rate of penetration (ROP) depth learning model training, and optimize the model structure.

3. Modeling Method

3.1. Real time drilling data processing

The training of neural network needs regular and high-quality data, so data preprocessing is required before model training, including data cleaning and data format processing. The data in this paper is directly read from the real-time drilling data of an offshore oilfield in a certain area, that is, the las file of its well. The Las file is a data file that is recorded and processed by the data feedback of real-time measurement of downhole sensors and the ground computer during the drilling process. It records 144 parameters of real-time drilling at different sampling depths, including zero data and unused parameters. There are 29610 pieces of data in this las file. After the data is accessed, only 15 key parameters are retained according to the traditional ROP calculation formula, as shown in the table below:

Table 1. Key Parameters Affecting ROP

Parameters	Full name	Unit
DEPTH	Depth	m
WOBA	Weight On Bit	tonne
HKLA	Weight On Hook	tonne
RPMA	Rotation Per Minute	1/min
TQA	Torque	kN.m
SPPA	Standpipe Pressure	psi
MDIA	Mud Density In	g/cm ³
MFOA	Mud Flow Out	L/min
BDTI	Bit Drilled Time	h
BDDI	Bit Drilled Distance	m
MFIA	Mud Flow In	L/min
TVA	Tank Volume	M ³
ECDT	ECD At Total Depth	g/cm ³
ROP	Rate Of Penetration	m/h

After all the data are read in, the data is cleaned according to the following three rules:

1. If the drilling speed in the sampling data of a certain depth is less than or equal to 0, the data of that depth shall be eliminated.
2. If the number of 0 data under a parameter accounts for 80% of the total data of the parameter, remove this parameter.
3. If all parameters in the data sampled at a certain depth

are 0, the data at that depth will be rejected.

After data preprocessing, there are 25224 pieces of data, with a data size of 25224 * 12. Data normalization was not used in this experiment because the difference in data magnitude was not very large. For the processed dataset, 90% of the training data and 10% of the validation data are allocated. Some data are shown in the figure below:

Table 2. Partial Data Display

	DEPTH	WOBA	HKLA	RPMA	TQA	SPPA	MDIA	MFOA	BDTI	MFIA	TVA	ECDT
1	154.89	1.4	73.59	0	0.67	3.09	9.18	5.44	0	4087.53	539.6	1.03
2	155.04	1.4	73.59	0	0.69	3.1	9.18	5.44	0	4093.71	539.52	1.03
3	155.19	1.2	73.8	0	0.69	3.08	9.18	5.44	0	4093.57	539.52	1.03
4	155.35	1.1	73.88	0	0.88	3.09	9.18	5.44	0	4093.41	539.41	1.03
5	155.52	1	74.03	0	0.59	3.08	9.18	5.44	0	4093.59	539.55	1.03
6	155.68	1	74	0	0.67	3.06	9.18	5.44	0	4093.84	539.63	1.03
7	155.79	0.2	73.7	10	0.29	6.35	9.18	5.44	0	4093.41	539.14	1.03
8	155.95	0.1	73.75	10	0.55	6.36	9.18	5.44	0	4093.56	538.96	1.03
9	156.11	0.1	73.73	10	0.26	6.42	9.18	5.45	0	4093.41	539.1	1.03
10	156.25	0.2	73.67	10	0.25	6.36	9.18	5.44	0	4093.41	539.02	1.03

3.2. Construction of deep learning model

In all the descriptions in this article, the model depth does not include the input layer. The model of deep learning includes input layer, hidden layer and output layer. According to the actual requirements of this paper, the number of input parameters is 12, and the prediction result is a result of drilling speed. Therefore, this model has 12 input nodes in the input layer and 1 node in the output layer.

According to a theory found in Lippmann's paper[11], MLP with two hidden layers is enough to see any required shape of classification area. Further research shows that MLP is a universal pusher. In the case of a hidden layer, as long as there are enough nodes in the hidden layer, MLP can approximate any function we need. In other words, a two-layer MLP can theoretically solve any problem. In single hidden layer MLP, the biggest problem is how to configure enough nodes and allocate reasonable weights. However, many counterexamples were found in the subsequent study of practical problems, such as: single-layer perceptron cannot solve XOR and nonlinear problems [12]. Some functions cannot be directly learned through the MLP of a single hidden layer or require an unlimited number of nodes [13]. Moreover, for functions that can be learned through MLP of a large enough single hidden layer, the efficiency of a single hidden layer is very low, and it will be more effective to use two or more hidden layers to learn it [14].

In addition to the normal hidden layer settings, deep learning also includes a large number of super parameters in the process of model construction. Hyperparameters are parameters that do not need data to drive, but are considered to be adjusted before or during training. Hyperparameters are usually divided into three categories: network parameters, optimization parameters, and regularization parameters. Network parameters refer to the interaction mode (addition, multiplication or concatenation) between network layers, the number and size of convolution cores, the number of network layers (also called depth) and activation function.

Optimization parameters refer to learning rate, batch size, parameters of different optimizers and adjustable parameters of some loss functions. The regularization parameter refers to the weight attenuation coefficient and the drop out ratio.

In the process of establishing the model, it is necessary to select a reasonable activation function. The activation function is used to add nonlinear features, so as to overcome the defect of insufficient representation ability of linear models. In the process of establishing this model, ReLU function is selected as the activation function. ReLU function does not have the problem of gradient disappearance. Secondly, it can maximize the median value of each feature and has fast operation speed. Adam is selected as the optimizer. The optimizer can quickly obtain good results for dealing with various problems, because it uses the same learning rate for each parameter and adapts independently with the learning process, and the optimizer is very consistent with the deep learning algorithm. The Loss function selects the mean absolute error (MAE) first. For the optimization parameters such as the learning rate and the number of batch samples, the neural network will automatically initialize them to reasonable values and adjust them according to each batch of learning. Among them,

$$ReLU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

In general, how to build a reasonable deep learning model has always been a complex problem. In most cases, the experience and experiments of workers are used to build appropriate deep learning models for specific practical problems. In this paper, we first establish a three-layer model with 32 and 16 input nodes. The R^2 correlation coefficient is used as the evaluation standard of the model, and the training batch is 200 times. The loss function loss result is shown in Figure 3. It can be seen from the results that the model has reached the convergence state in the case of 200 training, and no over fitting has occurred. The verification set performs well, with R^2 of 0.8425 and processing time of 101 seconds. The results are shown in Figure 4.

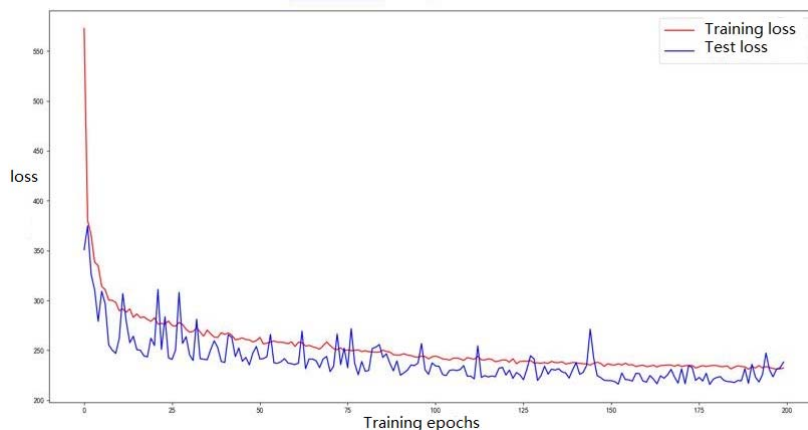


Figure 3. Loss function of three-layer model in training set and verification set

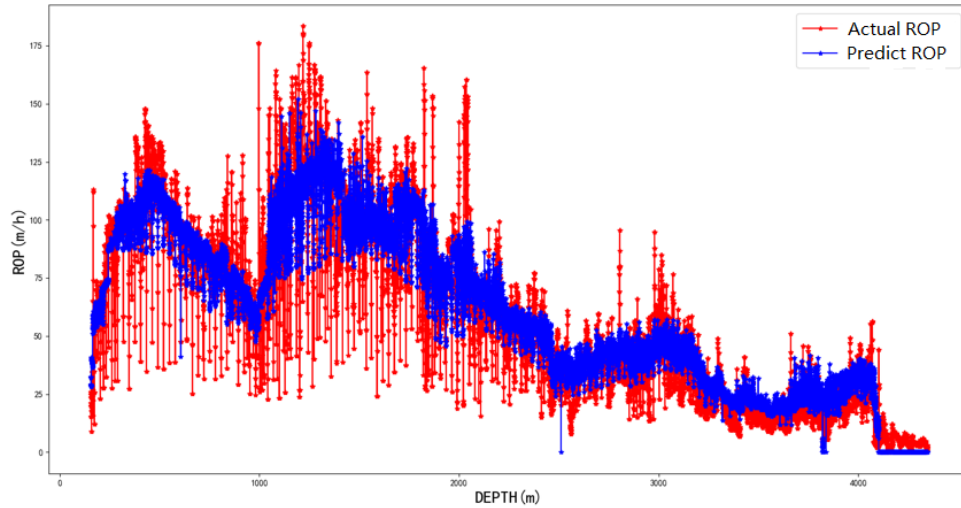


Figure 4. Comparison between predicted results and actual results

3.3. Structure optimization of deep learning model

It can be seen from the above experiments that the original three-layer model can accurately predict the results of drilling speed, but there is a lot of room for optimization in the model structure, so we optimized the model structure to ensure the best fitting effect.

Generally, if the data set has too few samples, the model will perform well in the training set, but poorly in the test set. This phenomenon is called overfitting. The basic method to solve the over fitting problem is to increase the number of samples. In the process of model training, there are still 25224 pieces of data in the data set after data cleaning, so the over fitting problem is reasonably avoided in the process of model establishment.

Table 2. Training results of different model structures

Training No	Number of model layers	Hidden layer structure	Convergence or not	R^2	Processing time (seconds)
1	3	32/16	Yes	0.8425	101
2	3	32/32	Yes	0.8487	111
3	3	64/32	Yes	0.8505	106
4	3	64/64	Yes	0.8391	111
5	3	64/16	Yes	0.8455	108
6	3	128/64	Yes	0.8490	118
N/A	3	128/128	No	N/A	N/A
7	4	16/16/16	Yes	0.8509	101
8	4	32/32/32	Yes	0.8518	110
9	4	64/32/16	Yes	0.8547	126
10	4	64/32/32	Yes	0.8506	129
11	4	64/64/64	Yes	0.8549	122
12	4	128/64/32	Yes	0.8533	118
13	5	128/64/32/16	Yes	0.8615	142
14	5	128/64/32/32	Yes	0.8575	141
15	5	128/32/32/32	Yes	0.8524	141
16	6	256/128/64/32/16	Yes	0.8536	171

It can be seen from the training results that, first of all, when the number of layers is the same, increasing the number of nodes has little effect on the correlation coefficient of the model, tends to be stable in general, and has little impact on efficiency. When the node structure is designed as a decreasing structure, the model convergence effect is good, while when the node structure is set as a structure with the same number, the model prediction effect is poor and the efficiency is low. In some cases, the total number of runs

exceeds 500, and no obvious convergence effect is obtained. Secondly, when the depth of the model structure is increased, the correlation coefficient is significantly improved, but the time is increased. Finally, when the number of layers is increased to a certain level, the correlation coefficient is reduced compared with the previous level, indicating that there are some thresholds for the impact of increasing the number of model layers on the accuracy.

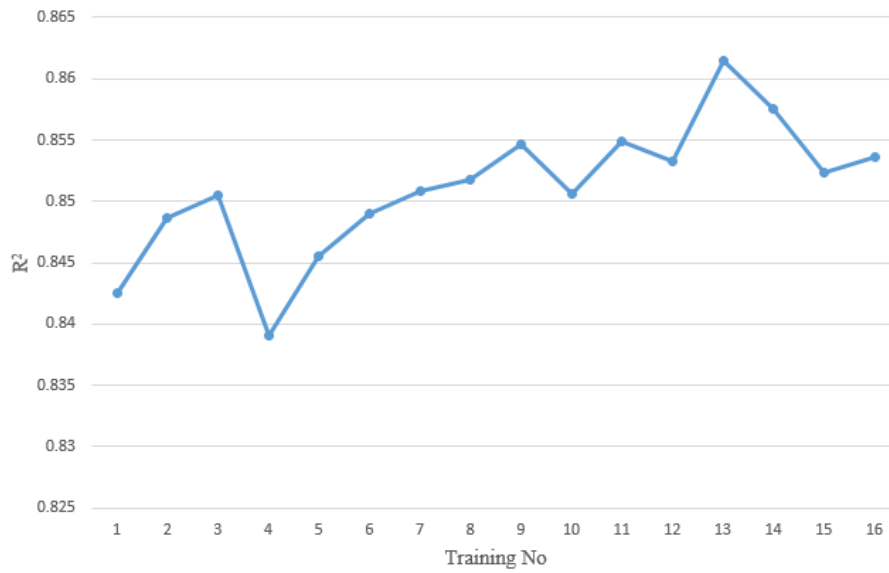


Figure 5. Accuracy evaluation of different model structures R²

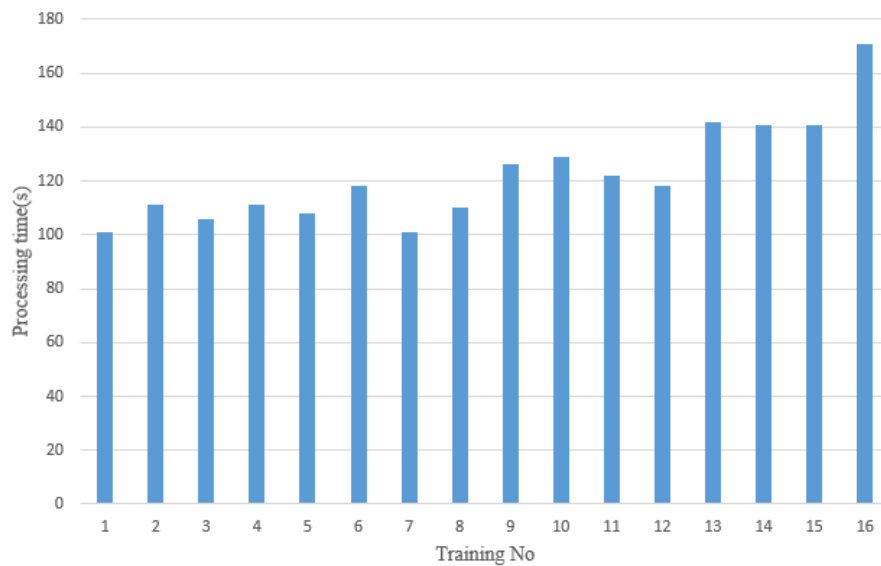


Figure 6. Training time for different model structures

After the training is completed, select the optimal model 16 to increase the training batch to 1000 and select this model to predict the penetration rate. It is found that R2 is adjusted to

0.8647. The results show that increasing the training batch can improve the training accuracy.

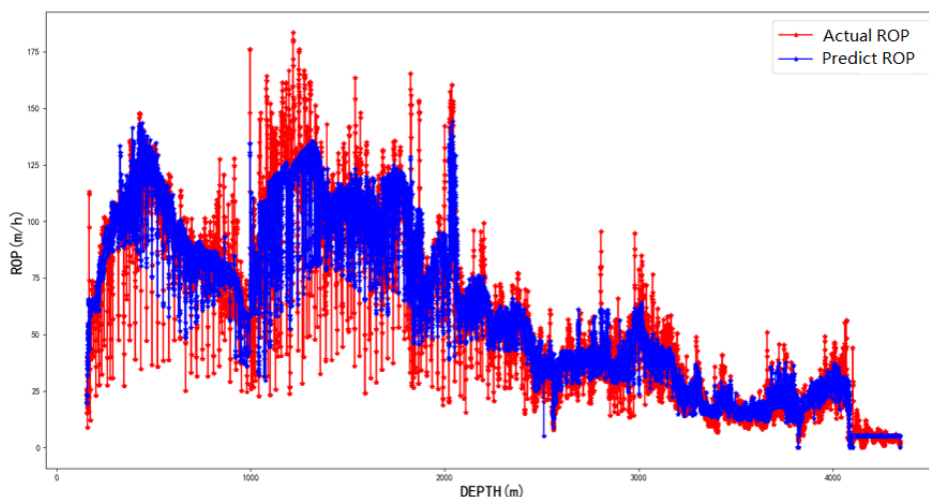


Figure 7. Comparison between predicted results and actual results of the optimal model

4. Conclusion

(1) The depth learning method is more effective and accurate in predicting the drilling speed, and the research finally improved the correlation coefficient R2 to 0.8647.

(2) Deep learning model structure optimization is more conducive to characterizing nonlinear problems, and selecting appropriate model layers and nodes can better match the penetration rate prediction problem.

(3) The depth neural network model has high efficiency and the overall operation time is less than 3 minutes, which can be effectively used for real-time drilling prediction.

References

- [1] Hegde, C., Pycrz, M., Millwater, H., Daigle, H., and Gray, K., Fully coupled end-to-end drilling optimization model using machine learning, *Journal of Petroleum Science and Engineering*, vol. 186, 2020, doi: 10.1016/j.petrol.2019.106681.
- [2] Soares C , Gray K . Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models[J]. *Journal of Petroleum Science and Engineering*, 2019, 172:934-959.
- [3] Hegde C , Daigle H , Millwater H , et al. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models[J]. *Journal of Petroleum Science and Engineering*, 2017:295-306.
- [4] Soares C , Daigle H , Gray K . Evaluation of PDC bit ROP models and the effect of rock strength on model coefficients[J]. *Journal of Natural Gas Science and Engineering*, 2016, 34:1225-1236.
- [5] Abdulmalek Ahmed, Abdulwahab Ali, Salaheldin Elkatatny, Abdulazeez Abdullaheem. *New Artificial Neural Networks Model for Predicting Rate of Penetration in Deep Shale Formation*[J]. *Sustainability*, 2019, 11(22):
- [6] Arabjamaloei R , Shadizadeh S . Modeling and Optimizing Rate of Penetration Using Intelligent Systems in an Iranian Southern Oil Field (Ahwaz Oil Field)[J]. *Petroleum Science and Technology*, 2011, 29(16):1637-1648.
- [7] Amar K , Ibrahim A . Rate of penetration prediction and optimization using advances in artificial neural networks, a comparative study. 2012.
- [8] Bataee M , Irawan S , Kamyab M . Artificial Neural Network Model for Prediction of Drilling Rate of Penetration and Optimization of Parameters[J]. *Journal of the Japan Petroleum Institute*, 2014, 57(2):65-70.
- [9] Hadi, F., Altaie, H., and AlKamil, E. 2019. Modeling Rate of Penetration Using Artificial Intelligent System and Multiple Regression Analysis. Paper presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 11–14 November.
- [10] Aliyev R , Paul D . A Novel Application of Artificial Neural Networks to Predict Rate of Penetration[C]// *SPE Western Regional Meeting*. 2019.
- [11] Lippmann R P . An introduction to computing with neural nets[J]. *IEEE Assp Magazine*, 1988, 4(2):4-22.
- [12] Cybenko G . Approximation by superpositions of a sigmoidal function[J]. *Mathematics of Control, Signals and Systems*, 1989, 2(4):303-314.
- [13] Hornik K , Stinchcombe M , White H . Multilayer feedforward networks are universal approximators[J]. *Neural Networks*, 1989, 2(5):359-366.
- [14] Liang S , Srikant R . Why Deep Neural Networks for Function Approximation?[J]. 2016.