International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 21, Issue 4 (April 2025), PP 217-221

Research on the Application and Progress of Intelligent Algorithms in the Prediction of Drilling Rate of Penetration

Congling Li, Kaijie Deng, Zhongwen Zhang, Haofeng Gu
(School of Petroleum Engineering, Yangtze University, Wuhan 40100, China)

Abstract: *In drilling engineering, the rate of penetration (ROP) is an important indicator reflecting drilling efficiency, which is directly related to the mining cost and construction period. Accurately predicting the ROP is of great engineering significance for optimizing resource allocation and reducing the construction cost of drilling operations, and it helps to promote the digital and intelligent development of drilling operations. With the continuous development and improvement of the intelligent drilling theoretical system, intelligent prediction of the drilling ROP is gradually becoming a research hotspot in the industry. This paper reviews the application and development status of intelligent algorithms in ROP prediction, and combines the development trends of artificial intelligence technology and advanced intelligent algorithms to explore the future development direction of ROP prediction, aiming to provide strong theoretical support and technical reference for the basic theoretical research of ROP prediction and the development of its engineering practice.*
Key words: *ROP prediction; machine learning; intelligent algorithms; integrated model*

Date of Submission: 08-04-2025 Date of acceptance: 19-04-2025

I. Introduction

The reservoirs of oil and gas resources in deep formations have poor drillability and strong abrasiveness. The rocks in the penetrated formations are hard and soft alternately, the rock-breaking process is complex, and it is difficult to improve the speed and optimize, which poses a severe challenge to on-site drilling construction. The accurate prediction of the ROP is not only an important factor in optimizing drilling operations, but also plays a crucial role in improving operation efficiency and reducing economic costs. Especially under complex geological conditions, the accurate prediction of the ROP can assist in the scientific allocation of resources before drilling and the formulation of more reasonable drilling operation plans, which is of great significance for improving drilling efficiency and reducing costs.

The development of research on drilling ROP prediction can be divided into three stages: (1) Statistical regression equations based on on-site data. The coefficients in the statistical regression equations of the ROP are closely related to drilling parameters or formations, and it is difficult to determine these coefficients, so their application has great limitations. (2) Physical ROP models based on drilling knowledge theory. Such models mainly rely on the physical principles in the drilling process and the internal relationships between geological engineering parameters, and describe the relationship between the ROP and various influencing factors by constructing mathematical equations. (3) ROP prediction methods using intelligent algorithms. The non-linear relationship among drilling parameters is strong and complex, and the application of traditional models is easily limited. With the rapid development of artificial intelligence and big data technology, intelligent algorithms mainly based on machine learning can effectively obtain the complex non-linear and multi - variable relationships in the data by learning a large amount of data, and update and train according to the data. Machine learning is a branch of artificial intelligence, mainly involving the research and construction of algorithms that can learn from data, enabling computer systems to automatically improve and adapt from experience. Common intelligent algorithms include artificial neural networks, support vector machines, random forests, multi-layer perceptrons, etc. For example, (Liu H. et al.,2022) established an ROP prediction model based on LSTM-FNN for ultra - deep wells. The results show that the evaluation index of the LSTM-FNN model is better than that of the FNN model and the LSTM model, and it has a better prediction effect. (Ashrafil S. et al.,2019) used multiple regression errors and trained through optimization algorithms such as genetic algorithms and particle swarm optimization algorithms to develop and train a variety of hybrid neural networks. The results show that the particle swarm optimization multi-layer perceptron shows better performance and accuracy compared with other models. Ahmed(Ahmed O.et al.,2019) established an ROP prediction model based on an artificial neural network and optimized the neural network structure using an adaptive differential evolution algorithm. (Cheng Z. et al.,2024) established an ROP prediction model based on a long short-term memory (LSTM) neural network with an attention mechanism and deeply analyzed the role of sequence features in ROP prediction. Tang Ming(Ming T. et al.,2023) proposed a new ROP prediction model based on the principal component analysis algorithm to optimize the BP neural network. The results show that the PCA-BP model has higher prediction accuracy and faster convergence speed. Song Xianzhi (Xianzhi S.et al.,2022) established an ROP prediction model using support vector machine regression based on real-time logging data obtained on - site. This model has better stability than other models.

Intelligent algorithms have a wide range of application scenarios and also have good application effects in ROP prediction. Studying and summarizing the applications of different intelligent algorithms in ROP prediction helps to deeply understand the advantages and disadvantages, as well as the applicable conditions of various intelligent algorithms in terms of prediction accuracy, model stability, and generalization ability. This paper elaborates on the application of intelligent algorithms in ROP prediction and the process of using intelligent algorithms to predict the ROP, comprehensively reviews the application status of single algorithms and hybrid algorithms in ROP prediction, and looks ahead to the future development of ROP prediction.

II. The Process of Predicting the Rate of Penetration with Intelligent Algorithms

The implementation of predicting the ROP with intelligent algorithms is a multi - stage and systematic project. Its goal is to establish a mathematical model that reflects the physical and geological characteristics of drilling through the real-time collection and processing of a large amount of multi - dimensional data on the drilling site, so as to achieve accurate ROP prediction and process optimization. The framework diagram of ROP prediction is shown in Figure 1(Yuan Q. et al.,2025).

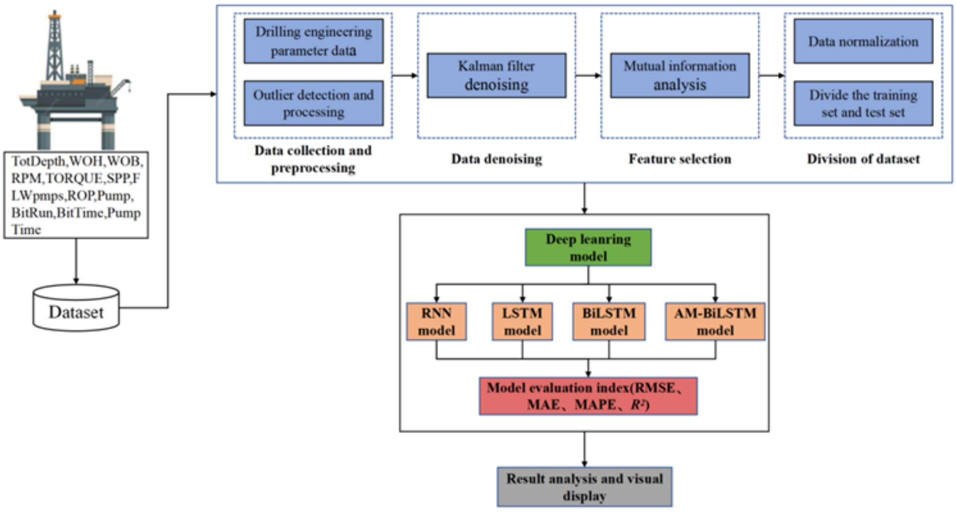

Figure 1 Framework Diagram of Rate of Penetration Prediction

During the drilling process, parameters such as rotational speed, torque, weight on bit, and displacement are monitored in real-time through multiple sensors, and geological data related to the drilled formation, such as rock density and formation porosity, are collected. These original data often contain noise, missing values, or outliers, and need to be subjected to data cleaning, denoising processing, etc., to ensure the accuracy and stability of the data. After data preprocessing, feature extraction is carried out on the original data by combining professional knowledge in the field of drilling engineering. At the same time, methods such as statistical analysis and correlation analysis are used to screen out variables that are highly correlated with the ROP, which helps to reduce the data dimension, reduce the computational complexity of the model, and improve the prediction accuracy. The preprocessed data is normalized to eliminate the influence of different magnitudes among features.

In model construction, an appropriate algorithm model is selected to describe the complex non-linear relationship between the ROP and various factors. Algorithms such as support vector machines, random forests, and decision tree regression perform well in processing small - and medium - scale datasets and have strong interpretability. multi-layer neural networks, long-short-term memory networks (LSTM), and other models are suitable for scenarios with a large amount of data and can capture more complex non-linear dynamic change characteristics.

After preliminary training, the model needs to be verified by comparing with historical drilling data and on - site real-time data. During the model training process, means such as cross - validation and hyperparameter tuning are used to ensure that the model has good generalization ability. At the same time, evaluation metrics such as mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) are used to evaluate the prediction performance of the model.

III. Applications of Intelligent Algorithms in Rate of Penetration Prediction

3.1 Applications of Single Algorithms

In recent years, with the booming development of big data and artificial intelligence technology, many scholars have established ROP prediction models using intelligent algorithms such as machine learning and artificial neural networks based on a large amount of real-time drilling data. The distribution of algorithm selection in ROP prediction is shown in Figure 2(Li Q.et al.,2024).

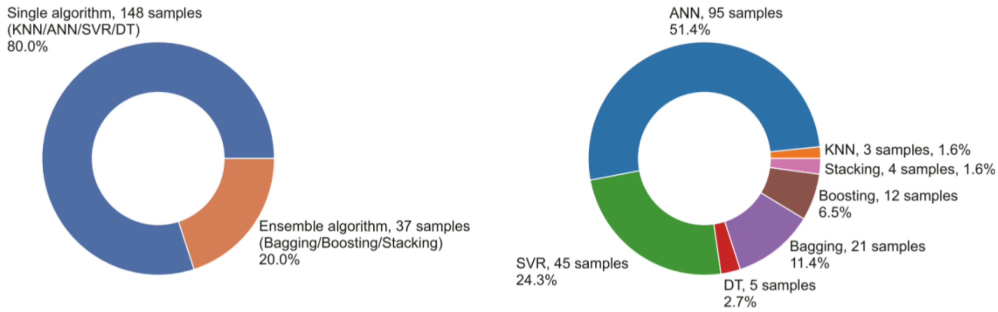

Figure 2 Distribution of Algorithm Selection in Rate of Penetration Prediction

Liu W(Liu W. et al.,2023) established four ROP prediction models using four machine learning methods, namely support vector machine regression, linear regression, regression tree, and gradient boosting regression, and compared them. Among them, the gradient boosting regression had the best prediction effect. Haodong C(Haodong C. et al.,2022) characterized the formation using acoustic transmission time and established an ROP prediction model based on a deep neural network method. The results show that the matching degree between the model and the real data can reach 82%. Ehsan B(Ehsan B. et al.,2022) used six machine learning models and two traditional prediction models to predict the ROP. The results show that machine learning models are more efficient and reliable than traditional models. Tu B(Tu B. et al.,2024) proposed a GRU-Informer model for real-time ROP prediction. The results show that GRU-Informer is superior to traditional recurrent neural networks such as LSTM, GRU neural networks, and Informer in real-time ROP prediction and has greater practical value. Junior J R B (Junior J. et al.,2025) compared three data stream algorithms (ARFAW, ARFPH, FIMTDD) with three traditional machine learning algorithms (random forest, gradient boosting tree, adaptive boosting tree). Compared with traditional machine learning methods, using appropriate data stream algorithms can reduce the average absolute error rate by 63%. Shaygan K (Shaygan K.et al.,2023) adopted a comprehensive data mining method for ROP prediction, used a multi-layer perceptron (MLP) neural network and a random forest regression model for ROP prediction, and designed its structure by adjusting the hyperparameters of the model.

3.2 Applications of Hybrid Algorithms

A hybrid model refers to a model composed of multiple basic models. These basic models can be homogeneous (such as a random forest composed of multiple decision trees) or heterogeneous (such as a combination of logistic regression, decision trees, support vector machines, etc.). The goal of a hybrid model is to improve the overall model performance and robustness by combining the prediction results of multiple models. Common integration methods include Bagging, Boosting, and Stacking, etc.

Feng Z (Feng Z. et al.,2023) conducted a performance analysis of ROP prediction using the hyperparameter technology of integrated supervised machine learning and used interpretable machine learning methods to explain the behavior of drilling parameters. The results show that the integrated model shows the best performance in ROP prediction. Gao Yunwei (Yunwei G. et al.,2024) fused the K - nearest neighbor algorithm, support vector machine algorithm, and random forest algorithm through the stacking integration strategy and optimized the parameters of each basic model using the genetic algorithm, achieving a high prediction accuracy. Alsaihati A (Alsaihati A. et al.,2022) introduced an integrated model based on a random forest, with artificial neural networks and adaptive neuro - fuzzy inference systems as basic learning models, to predict the ROP of different lithology formations. Yang Y(Yang Y. et al.,2025) combined basic learners based on machine learning algorithms such as decision tree regression, random forest, support vector regression, and backpropagation neural network to construct an integrated model. Compared with single models and other integrated models, the proposed integrated model has higher accuracy, as well as stronger robustness and generalization ability. Ahmednour O (Ahmednour O. et al.,2025) integrated a Stacking integrated machine learning model, utilized the diversity of basic learners, and used linear support vector regression (LinearSVR) as a meta - learner for comprehensive prediction. Compared with single models, it achieved higher accuracy. Xu M (Xu M. et al.,2021) considered the collaborative combination of different machine learning models and used time cost and goodness of fit as model evaluation criteria to solve the problem of low fitting accuracy of single models and achieved accurate prediction of the integrated model. The principle of Stacking ensemble learning is shown in Figure 3(Ren Y. et al.,2023).

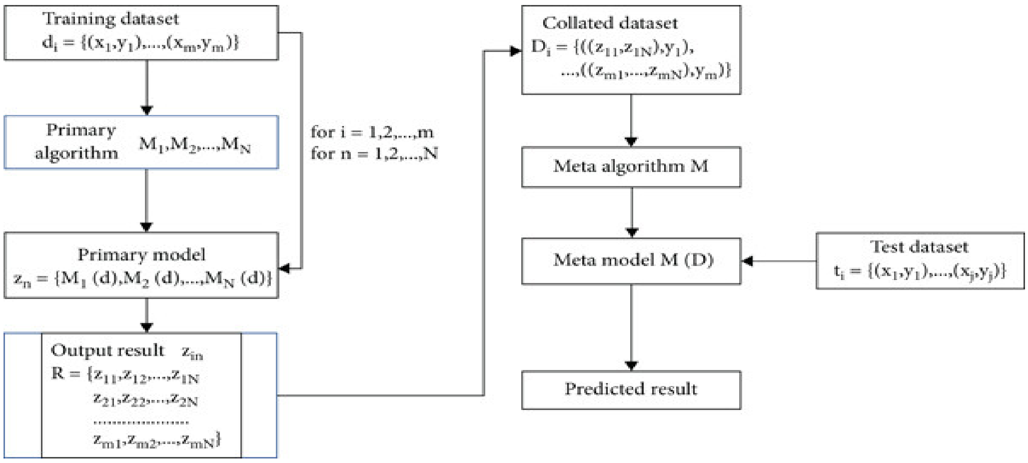

Figure 3 Schematic Diagram of Stacking ensemble learning Principle

IV. Future Development Prospects

(1) The data acquisition technology and sensor devices will be further intelligent and integrated. A higher - precision and multi - modal sensor network will be used to achieve real-time and high - frequency acquisition of key parameters in the drilling process, providing a more abundant and accurate data basis for model construction.
(2) With the continuous development of algorithms such as deep learning, reinforcement learning, and transfer learning, the ROP prediction model will be continuously optimized. It will be able to capture the complex non-linear relationships and time - varying dynamic characteristics in the drilling process, achieving higher prediction accuracy and real-time response capabilities. The introduction of model adaptation and online update mechanisms enables the system to continuously adjust itself according to changes in the drilling site environment and geological conditions, enhancing robustness and generalization ability.
(3) The integrated application based on digital twin and cloud computing helps to break through the time - space limitations of traditional data processing technologies and realize remote monitoring and intelligent decision - making support. In addition, the organic combination of interdisciplinary and multi - field collaborative innovation models will promote the transformation of ROP prediction technology from single - application to drilling intelligence, and comprehensively improve drilling efficiency, cost control, and other aspects.

V. Conclusion

(1) A variety of machine learning models represented by random forests, gradient boosting trees, support vector machine regression, extreme learning machines, etc., have been widely applied in the field of ROP prediction and have achieved remarkable results. Through learning and training on a large amount of drilling parameter data, these models can automatically extract hidden features from the data and capture the non-linear relationships and multi - dimensional interaction effects among variables, thus achieving accurate modeling of drilling parameter sequences.
(2)Compared with single models, integrated models integrate the advantages of multiple base learners. By adopting integration strategies such as Bagging, Boosting, and Stacking, the local features and rules captured by different models when processing data can complement each other. At the same time, it balances the bias and variance problems of each model caused by insufficient training data or noise interference, thereby achieving more stable and robust prediction results.
(3)With the progress of cutting edge technologies such as big data and artificial intelligence, combining these technologies with artificial intelligent systems optimization algorithms and applying them to ROP prediction and optimization research will promote the continuous development of drilling engineering towards intelligence, real - time monitoring and feedback control, integration, and automation, providing more reliable technical support for the efficient, safe, and sustainable operations of drilling.

References

[1] Barbosa L F F M, Nascimento A, Mathias M H, et al. Machine learning methods applied to drilling rate of penetration prediction and optimization-A review[J]. Journal of Petroleum Science and Engineering, 2019, 183: 106332.

[2] Yang X, Wu M, Lu C, et al. Prediction of rate of penetration based on drilling conditions identification for drilling process[J]. Neurocomputing, 2024, 579: 127439.
[3] Hegde C M, Wallace S P, Gray K E. Use of regression and bootstrapping in drilling inference and prediction[C]//SPE Middle East Intelligent Oil and Gas Symposium. SPE 176791, 2015.
[4] Bingham M G. How rock properties are related to drilling[J]. The Oil and Gas Journal, 1964, 62: 94-101.
[5] Bourgoyne A T, Young F S. A multiple regression approach to optimal drilling and abnormal pressure detection[J]. Society of Petroleum Engineers Journal, 1974, 14(4): 371□384.
[6] Teale R. The concept of specific energy in rock drilling[J]. International Journal of Rock Mechanics and Mining Sciences &Geomechanics Abstracts, 1965, 2(1): 57-73.
[7] Liu H, Jin Y, Song X, et al. Rate of penetration prediction method for ultra-deep wells based on LSTM–FNN[J]. Applied Sciences, 2022, 12(15): 7731.
[8] Ashrafi S B, Anemangely M, Sabah M, et al. Application of hybrid artificial neural networks for predicting rate of penetration (ROP): A case study from Marun oil field[J]. Journal of petroleum science and engineering, 2019, 175: 604-623.
[9] Ahmed O S, Adeniran A, Ariffin S. Computational intelligence based prediction of drilling rate of penetration: a comparative study[J]. Journal of Petroleum Science and Engineering, 2019, 172: 1-12.
[10] Cheng Z, Zhang F, Zhang L, et al. A sequential feature-based rate of penetration representation prediction method by attention long short-term Memory Network[J]. SPE Journal, 2024, 29(02): 681-699.
[11] Tang Ming, Wang Hanchang, He Shiming, et al. Based on PCA - BP algorithm of penetration prediction research [J]. Journal of petroleum machinery, 2023 ploidy (10) : 31 + 23-76
[12] SONG Xianzhi, PEI Zhijun, WANG Pantao, et al. Intelligent prediction for rate of penetration based on support vector machine regression[J]. Xinjiang Oil & Gas, 2022, 18(1):14-20.
[13] Yuan Q, He M, Chen Z, et al. A real-time prediction method for rate of penetration sequence in offshore deep wells drilling based on attention mechanism-enhanced BiLSTM model[J]. Ocean Engineering, 2025, 325: 120820.
[14] Li Q, Li J P, Xie L L. A systematic review of machine learning modeling processes and applications in ROP prediction in the past decade[J]. Petroleum Science, 2024.
[15] Liu W, Fu J, Tang C, et al. Real-time prediction of multivariate ROP (rate of penetration) based on machine learning regression algorithms: Algorithm comparison,model evaluation and parameter analysis[J]. Energy Exploration & Exploitation, 2023, 41(5): 1779-1801.
[16] Haodong C, Yan J, Wandong Z, et al. Deep Neural Network Prediction of Mechanical Drilling Speed[J]. Energies, 2022, 15(9): 3037-3037.
[17] Ehsan B, Ebrahim D B. Computational prediction of the drilling rate of penetration(ROP): A comparison of various machine learning approaches and traditional models[J]. Journal of Petroleum Science and Engineering, 2022, 210.
[18] Tu B, Bai K, Zhan C, et al. Real-time prediction of ROP based on GRU-Informer[J]. Scientific reports, 2024, 14(1): 2133-2133.
[19] Junior J R B, Lavi B. Enhancing rate of penetration prediction in drilling operations: A data stream framework approach[J]. Engineering Applications of Artificial Intelligence, 2025, 143: 110034.
[20] Shaygan K, Jamshidi S. Prediction of rate of penetration in directional drilling using data mining techniques[J]. Geoenergy Science and Engineering, 2023, 221: 111293.
[21] Feng Z, Gani H, Damayanti A D, et al. An explainable ensemble machine learning model to elucidate the influential drilling parameters based on rate of penetration prediction[J]. Geoenergy Science and Engineering, 2023, 231: 212231.
[22] GAO Yunwei, Luo Limin, Xue Fenglong, et al. Mechanical drilling rate prediction method based on Stacking ensemble learning [J]. Journal of petroleum machinery, 2024,52 (05) : 17-24 + 52.
[23] Alsaihati A, Elkatatny S, Gamal H. Rate of penetration prediction while drilling vertical complex lithology using an ensemble learning model[J]. Journal of Petroleum Science and Engineering, 2022, 208: 109335.
[24] Yang Y, Cen X, Ni H, et al. A highly accurate and robust prediction framework for drilling rate of penetration based on machine learning ensemble algorithm[J]. Geoenergy Science and Engineering, 2025, 244: 213423.
[25] Ahmednour O, Chen D, Liu J, et al. Comparative analysis of rate of penetration prediction and optimization in deep wells using real-time continuous stacked generalization ensemble learning: A case study in Shunbei[J]. Geoenergy Science and Engineering, 2025, 247: 213674.
[26] Xu Mingze, Wei Minghui, Deng Shuang, et al. New application of multi-model ensemble Learning in predicting penetration rate [J]. Computer Science, 2019,48(S1):619-622+657.
[27] Ren Y, Lu B, Zheng S, et al. Research on the rate of penetration prediction method based on stacking ensemble learning[J]. Geofluids, 2023, 2023(1): 6645604.