231 (2023) 212231

Contents lists available at ScienceDirect

Geoenergy Science and Engineering

journal homepage: www.sciencedirect.com/journal/geoenergy-science-and-engineering

An explainable ensemble machine learning model to elucidate the influential drilling parameters based on rate of penetration prediction

Zhipeng Feng [a], Hamdan Gani [b,c,*], Annisa Dwi Damayanti [d], Helmy Gani [e]

[a] School of Culture Creativity and Media, Hangzhou Normal University, Hangzhou, Zhejiang, China
[b] Department of Machinery Automation System, ATI Makassar Polytechnic, Makassar, Indonesia
[c] Graduate School of Computer Systems, Makassar Handayani University, Makassar, Indonesia
[d] Department of Environmental Engineering, Faculty of Engineering, Hasanuddin University, Makassar, Indonesia
[e] Department of Industrial Hygiene, Faculty of Public Health, Occupational Health and Safety, Makassar College of Health Sciences, Indonesia

ARTICLE INFO

Keywords:
Geothermal
Prediction
Rate of penetration
Drilling parameters
Explainable ML.

ABSTRACT

Many researchers have examined the benefits of machine learning (ML) algorithms in geothermal drilling, especially for predicting the rate of penetration (ROP) of drilling. However, little attention has been given to enhancing the explainability of ML algorithms. Motivated by this, this study examines the feasibility of an explainable ML approach for elucidating the behavior of the drilling parameters to generate explainable insights regarding the surface-controllable drilling parameters that could maximize the ROP for drilling effectiveness (DE). First, this study presents a performance analysis using hyper-parameter technique of the ensemble supervised machine learning (ESML) for ROP predictions. Secondly, this study uses an explainable ML method to explain the drilling parameters' behavior based on the ESML prediction. The results show that the ensemble CatBoost model shows the best performance for ROP prediction. Then, the explainable ML successfully revealed several influential drilling parameters that could maximize the ROP for DE. In earlier drilling (0–4,000 m in depth), *temperature* and *pressure* contribute to a lower ROP, and the drilling fluid *flow rate* and *rotary speed* contribute to a higher ROP. On the other hand, as the drills go deeper into the earth (4,000–7,536 m in depth), maximum ROP can be obtained by maximizing the drilling *pressure*, *weight on bit (WOB)*, *surface torque*, and *rotary speed (RPM)* of the drill bit, while drilling fluid *flow rate* does not significantly contribute to ROP. The explainable ML also highlighted the drilling parameters' inflection points when a switch from low to high rates of ROP occurs. Overall, the findings show the proposed model's strength in generating explainable insights that can improve the selection of appropriate ROP settings and support future decision-making processes in geothermal drilling.

1. Background

Geothermal energy is commonly considered a renewable and sustainable form of energy, as opposed to fossil fuels. Geothermal energy uses the earth's natural internal heat; typical fluids in the geothermal reservoir include water and steam at high temperatures. However, its exploitation requires deep drilling, which can be difficult and costly. The development of geothermal resources has many challenges; this is primarily caused by a lack of funding and insufficient technology to explore sites of such depth and high temperature (Muther et al., 2022). When drilling for geothermal resources, many factors can impact the ROP (Wang et al., 2022, 2023). Those factors, which include rock strength properties, drilling fluid properties, and drilling bit wear, are some of the parameters of concern that increase the uncertainty level of the drilling operation or ROP (He et al., 2021; Wang and He, 2023). Therefore, researchers estimate that geothermal drilling accounts for around 30%–50% of the total project cost (Allahvirdizadeh, 2020; G. Li et al., 2022). The cost will expand even more when dealing with complex oil and gas resources like offshore and ultra-deep reservoirs (Okoroafor et al., 2022).

ROP is the increase in speed or progression of the drill bit when the formation rock breaks (Sobhi et al., 2022). It is determined by measuring the required amount of time to drill a given unit of depth, and it is generally reported in feet per hour, with the typical average ROP

* Corresponding author. Department of Machinery Automation System, ATI Makassar Polytechnic, Makassar, Indonesia.
E-mail addresses: 20200058@hznu.edu.cn (Z. Feng), hamdangani@atim.ac.id (H. Gani), annisa.dd@unhas.ac.id (A.D. Damayanti), ganihelmy@gmail.com (H. Gani).

https://doi.org/10.1016/j.geoen.2023.212231
Received 2 June 2023; Received in revised form 22 July 2023; Accepted 3 August 2023
Available online 9 August 2023

between 6 and 13 feet per hour (Rosberg and Erlström, 2021). Drilling at the highest ROP possible is necessary to lower the drilling expense per foot (Ben Aoun and Madarász, 2022; Okoroafor et al., 2022). However, drilling at the highest ROP possible should be balanced with safety considerations and risk management. The optimization of drilling parameters and equipment should always prioritize safety and the preservation of wellbore stability (Gomez et al., 2021). Therefore, drilling for geothermal resources have many serious challenges, requiring technological innovation to model or simulate the drill bit's penetration rate, thus decreasing the risks and costs.

Nowadays, ML technologies have become a solution to the challenges of drilling for geothermal resources. The integration of ML technology in geothermal drilling has rapidly increased over the last two decades because it provides the solution for technological innovation which can significantly improve DE and reduce the risks and project costs (Alsaba et al., 2020; G. Li et al., 2022; Okoroafor et al., 2022). ML has been applied as a tool for drilling simulation that has been proven to reduce the risks and cost of geothermal drilling development (Ben Aoun and Madarász, 2022). ML technologies for ROP modeling have always been the primary concern of researchers and companies in the industry since ROP is directly related to the drilling cost (Sobhi et al., 2022). Therefore, many researchers have implemented those technologies in many drilling areas, such as intelligent optimization and prediction, drilling trajectory optimization, drilling risk management, drilling cementing evaluation, fracturing optimization, and completion management through automated equipment to reduce risks and drilling costs (He et al., 2022; G. Li et al., 2022).

Mostly, ML technologies for geothermal drilling studies can be grouped into intelligent algorithms and intelligent equipment. Intelligent algorithms apply ML algorithms to solve complex nonlinear problems, such as intelligent drilling prediction, which uses drilling parameters as input for the ML model to predict the ROP. Conversely, intelligent equipment supplies data sources and hardware support to determine and confirm intelligent models (G. Li et al., 2022). Many researchers have utilized these intelligent drilling predictions for many years, for example, Hegde et al. (2017). Their experimental work compared three physics-based or conventional models (i.e., Bingham, Motahhari, and Hareland) with ML models (i.e., linear regression and random forest regressor) for ROP prediction. Their ML models utilize controllable parameters (such as weight on bit, rotary speed, and flow rate) as input features. Both models were compared based on the coefficient of determination (R^2). Based on the results, ML models were obtained better performance than physics-based models. Similarly, Shahdi et al. (2021) and Han et al. (2019) compared ML models to physics-based models and found that ML achieved the best performance.

Despite the extensive research regarding intelligent drilling prediction, little attention has been given to enhancing the explainability of ML algorithms. To our knowledge, no studies explain which types of drilling parameters may potentially lead to maximizing ROP for DE. It is still challenging to dynamically manipulate the controllable parameters of drilling to ensure an effective ROP (Sobhi et al., 2022). Finding the right balance and optimizing drilling parameters based on the particular drilling circumstances and formation characteristics is crucial for maximizing ROP (Bani Mustafa et al., 2021). Therefore, understanding the controllable drilling parameters is essential, as it can provide knowledge to evaluate the geothermal project risks and costs. The knowledge gained from this study is beneficial because it (a) helps engineers choose the optimal drilling parameters to maximize ROP for future drilling projects (Ben Aoun and Madarász, 2022), (b) allows engineers to estimate the drilling risks (G. Li et al., 2022), and (c) allows researchers to optimize algorithms for maximizing ROP (Muther et al., 2022).

Explainable ML is a technique that describes the underlying processes in ML models in a way humans can understand (Barredo Arrieta et al., 2020). Numerous studies have utilized this technique. In experimental studies, Chakraborty et al. (2021) used the explainable ML technique to explain the compressive behavior of high-performance concrete. Zhipeng and Gani (Zhipeng and Gani, 2022) employed the explainable ML technique to elucidate the harmful content in video games, and it has also been utilized in the transportation field, such as to explain road traffic accidents (K. Li et al., 2022). Although previous studies have confirmed the promise of explainable ML regarding its interpretability, no study has used it to examine drilling parameters. Current research studies focus on intelligent drilling prediction and optimization to obtain metrics with the highest prediction accuracy of ROP (Ben Aoun and Madarász, 2022; Han et al., 2019; Hegde et al., 2017; G. Li et al., 2022) but lack a detailed analysis of the ML used for the purpose. Previous review studies have also raised the fact that no drilling studies focus on explainability (G. Li et al., 2022). Our study tackles these research omissions identified in previous experimental and review papers.

Therefore, the goal of this paper is to examine the feasibility of explainable ML models for elucidating the behavior of drilling parameters based on ROP prediction to generate explainable insights regarding the drilling parameters that could lead to maximizing the ROP for DE. This paper suggests two main methods to enhance the explainability of ML models. First, this study utilized hyperparameter optimization and compared six advanced ESML models to determine the best model for ROP prediction. The models were Random Forest Regressor [RFG], Gradient Boosting Regressor [GBR], CatBoost Regressor [CBR], AdaBoost Regressor [ABR], XGBoost Regressor [XGB], and LightGBM Regressor [LGR]. Then, this study applied an explainable ML technique for explaining the controllable drilling parameters based on the ESML prediction. This study proposes several explanations to increase the model's explainability using global and local interpretability analyses based on SHapley Additive exPlanation (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) framework.

This study first provides an overview of the application of ML in geothermal ROP prediction. Next, this study explains the materials, method, and development process of the ESML model and explainable ML in this work. Further, this study presents a performance analysis of ESML to determine the best model for interpretable analysis. Then, this study applies an explainable ML approach to elucidate the drilling parameters based on the ESML prediction. This approach generates new knowledge regarding the influences of the drilling parameters on the ROP and enables practitioners to identify appropriate ROP settings. Finally, this study discusses potential applications and future research directions in geothermal drilling.

1.1. Why Ensemble Supervised Machine Learning (ESML)?

The ESML integrates numerous base models to create a single best predictive model. Many researchers claimed that the ESML model performs better than the non-ESML model because it (a) has the potential to achieve higher predictive accuracy (Mohammed and Kora, 2023), (b) can decrease specific errors generated by individual models, resulting in a more accurate overall prediction by combining the predictions of numerous models (S. Zhang et al., 2021), (c) can minimize the impact of the outliers and make more reliable predictions (Lu et al., 2022), (d) can capture diverse features of the dataset and create more thorough predictions. Different models may have distinct biases, assumptions, or perspectives, and by integrating them, ESML methods can utilize their accumulate intelligence (Dong et al., 2020) and (e) can potentially provide new data insights and model predictions. It is feasible to understand better the underlying patterns and factors impacting predictions by studying the contributions of individual models within the ESML (Chen et al., 2020).

Over the years, the ESML model has successfully been implemented in numerous areas, such as software defect prediction (Sharma et al., 2023), bioinformatics (Cao et al., 2020), cancer prognosis and diagnosis (Zolfaghari et al., 2023), and solar irradiation (Alam et al., 2023). Further, many researchers reported that the ESML model outperformed non-ESML models regarding their predicted accuracy. For example, Akano and James (Akano and James, 2022) reported that the ESML model (i.e., gradient boosting, XGB, and ABR) outperformed the non-ESML model (i.e., artificial neural network, support vector machine, decision tree, and linear regression) in predicting crude oil viscosity. Amin et al. (2022) identified that ABR and RFG models performed better than decision tree and support vector regression in predicting the mechanical properties of fly ash/slag-based geopolymer concrete. Additionally, Y. Li and Chen (Y. Li and Chen, 2020) compared the performance of the ESML model (i.e., RF, ABR, XGB, and LGR) and non-ESML model (i.e., neural network, decision tree, logistic regression, naïve bayes, and support vector machine). Their experimental findings reveal that the performance of the ESML model is better than the non-ESML model. In short, previous research has demonstrated the supremacy of the ESML model over the non-ESML model. In addition, to our knowledge, the ESML models for ROP prediction are not yet researched by previous works. Therefore, studying ESML models for ROP prediction in the more recent observation time will be beneficial for the current understanding.

1.2. Why interpretability analysis using SHAP and LIME?

SHAP is a game theory-based method that provides a powerful and insightful evaluation of a feature's importance in a model (Lundberg and Lee, 2017). LIME is a technique that uses a local interpretable model to approximate any black box ML model and explain each prediction (Ribeiro et al., 2016). LIME generates a modified dataset to fit an explainable model, whereas SHAP requires a whole sample to derive SHAP values. This means that LIME requires only one observation, whereas SHAP requires many observations. Both LIME and SHAP are model interpretation techniques. LIME focuses on local explanations via perturbations and simpler models, whereas SHAP provides global explanations using Shapley values.

In the recent years, SHAP and LIME have successfully been implemented in numerous areas, such as anomaly detection (Antwarg et al., 2021), suicide attempt risk (Nordin et al., 2023), compressive strength predictions (Ekanayake et al., 2022), drug concentrations (Ogami et al., 2021), NO2 forecasting (Vega García and Aznarte, 2020), nasopharyngeal cancer survival (Alabi et al., 2023), Classification of Cardiac Arrhythmia (Abdullah et al., 2023), and prediction of tight gas (Ma et al., 2023). The LIME and SHAP approaches are by far the most complete and prevalent across the literature methods for showing feature interactions and feature relevance when explaining any black-box model (Linardatos et al., 2020). The LIME and SHAP approaches are model-agnostic and have also been shown to work with any form of data (Saranya and Subhashini, 2023). In brief, past research has demonstrated the strength of the SHAP and LIME techniques to explain the predictions of ML models. Therefore, considering the research gaps revealed by the prior studies discussed in the previous section, this study utilizes the SHAP and LIME framework to explain the behavior of drilling parameters based on ROP prediction.

2. Literature review

In recent years, many researchers have implemented ML models in the prediction of ROP. Based on the review papers (G. Li et al., 2022; Okoroafor et al., 2022). ML application in intelligent drilling prediction is divided into two groups. The first is the application of ML in the intelligent prediction of ROP, and the second is intelligent drilling optimization for ROP prediction.

2.1. Intelligent drilling prediction of ROP

ML prediction models usually use a single model or hybrid models to improve the accuracy of ROP predictions. For example, Diaz et al. (2018) implemented a Fourier transform filter to smoothen the collected data and utilized multiple linear regression analysis and a shallow neural network to predict ROP. Both linear regression and shallow neural network models improved the accuracy of ROP prediction, with the smoothened data influencing the model's high performance. Diaz et al. (2019) implemented a shallow neural network algorithm on datasets such as pore pressure gradient, equivalent circulating density, and surface drilling data to predict ROP. These training datasets were then divided into two scenarios which comprised (1) a scenario that applied accumulative data and (2) a scenario that applied data resampling. Although the shallow neural network model performed well in the scenario of accumulating training data, it was discovered that accuracy decreased when the well's depth was increased. The strategy involving data resampling had the lowest mean percentage error. Thus, the authors suggested combining accumulative and resampling data for ROP prediction. Next, Diaz and Kim (Diaz and Kim, 2020) investigated the possibility of predicting ROP using neighboring wells. The research used a filtering procedure to choose high-quality data using a statistical method, and it fitted those data sets into a deep neural network with three hidden layers. They found that deep neural networks accurately predicted ROP.

Other researchers have implemented various ML models in the prediction of ROP, such as Hegde et al. (2017), who used an integrated random forest regressor, an artificial neural network, and linear regression; Han et al. (2019), who used an artificial neural network and a long short-term memory model; and Soares and Gray (Soares and Gray, 2019), who used random forest, support vector machines, and an artificial neural network. Oyedere and Gray (Oyedere and Gray, 2020) used logistic regression, linear discriminant analysis, quadratic discriminant analysis, support vector machines, and random forest. Al-Sahlanee et al. (2023) applied random forest, gradient boosting, and extreme gradient boost. Leines Artieda et al. (2022) used a combination of a probabilistic algorithm, a gaussian mixture model, and a deep feedforward neural network. Lastly, Gong et al. (2019) used k-means clustering, artificial neural network, support vector machine, multivariate linear regression, and multivariate adaptive regression spine.

In summary, previous works show that researchers generally use a single or combination of non-ESML models for ROP predictions. Then, the ML model is found to be more popular than physics-based models in discovering the complicated mapping relationship between the ROP and their affecting factors, such as formation characteristics, bit characteristics, and drilling parameters. In addition, to our knowledge, no previous research has investigated ESML models for ROP prediction, and no study focuses on explainability in intelligent prediction of ROP area. Thus, the explainability using ESML models is one of the most essential issues in intelligent drilling. Interpretable models for intelligent prediction should be investigated in conjunction with actual drilling scenarios.

2.2. Intelligent drilling optimization for ROP prediction

The optimization prediction of the ROP is an extension of ROP prediction, whereby optimal drilling parameters (e.g., WOB, RPM, and flow rate) are achieved using optimization algorithms. Several related studies regarding drilling optimization for ROP prediction are described below.

Liao et al. (2020) explored the effects of proper feature selection using intelligent techniques on the prediction and optimization of ROP. Six hundred and eighteen data sets, including RPM, flushing media, and compressive strength parameters, were measured and gathered for this purpose. Their study implemented an artificial bee colony to better optimize drilling parameters and utilized an artificial neural network to predict ROP. The collected results demonstrated that intelligent techniques could effectively improve the accuracy of ROP prediction. Mehrad et al. (2020) studied a drilling rate prediction model using hybrid ML. They compared the combination of least-squares support-vector machines with the cuckoo optimization algorithm, least-squares support-vector machines with particle swarm optimization, and least-squares support-vector machines with genetic algorithms for ROP prediction. Their hybrid algorithms for ROP prediction utilized a dataset from southwestern Iran. For this goal, the study collected mud logging parameters, including depth, mud density, torque, standpipe pressure, equivalent circulating density, weight on bit, revolutions per minute, flow rate, and ROP. These were combined with geo-mechanical parameters, including gamma ray logging, porosity, density, and uniaxial compressive strength. The outcomes demonstrated superior accuracy of the least-squares support-vector machines with the cuckoo optimization algorithm of all the examined models. Gan et al. (Gan et al., 2019a, 2019b) proposed the prediction of ROP using hybrid support vector regression. They utilized a hybrid bat algorithm to optimize the hyper-parameters of the support vector regression model. Data from a drilling location in the Shennongjia region of central China was used to validate their proposed model. Their findings demonstrated that the suggested method was superior to eight well-known models in prediction accuracy.

Other researchers in the field of geothermal drilling have used optimization methods and ML models in the prediction of ROP. Anemangely et al. (2018) utilized a hybrid model composed of a multilayer perceptron neural network together with either a particle swarm optimization algorithm or a cuckoo optimization algorithm. Abbas et al. (2019) applied the FSCARET optimization method and an artificial neural network. Sabah et al. (2019) applied decision trees, a random forest regressor, support vector regression, a multilayer perceptron, a radial basis function, and multilayer perceptron-particle swarm optimization. Zhang et al. (H. Zhang et al., 2020) applied a random forest regressor and particle swarm optimization. Jiang and Samuel (Jiang and Samuel, 2016) used a Bayesian regularization neural network and ant colony optimization. Hegde et al. (2019) utilized a random forest regressor and gradient ascent optimization algorithms. Gan et al. (Gan et al., 2019a, 2019b) used the support vector regression with bath algorithm and support vector regression with particle swarm optimization. Bataee and Mohseni (Bataee and Mohseni, 2011) applied artificial neural network-particle swarm optimization, and Arabjamaloei and Shadizadeh (Arabjamaloei and Shadizadeh, 2011) implemented applied artificial neural network-genetic algorithm. Moazzeni and Khamehchi (Moazzeni and Khamehchi, 2020) used the rain optimization algorithm, and Hegde and Gray (Hegde and Gray, 2018) applied random forest regressor-particle swarm optimization. Lastly, Li, Cheng, and Cheng (C. Li et al., 2023) used artificial neural networks and achieved the highest accuracy after genetics algorithms optimization.

In summary, prediction is an extension of optimization that uses the optimal drilling parameters from the results of optimization algorithms. Previous works show that optimization algorithms and single or hybrid non-ESML models are the primary methods for improving prediction accuracy. The optimization results are an essential reference for prediction. The optimal drilling parameters (e.g., WOB, rotational speed, and pump press) are obtained using optimization algorithms. The optimization techniques make it easier to choose the best drilling parameters to obtain the best accuracy of ROP prediction. However, the explainability of ML models is research omissions identified in previous papers in the area of intelligent drilling optimization. Therefore, in combination with real drilling scenarios, integration development of optimization algorithms with a well-tuned ESML model and an explainable ML technique should be explored in the current observation time.

3. Materials and method

The data set was sourced from the US Department of Energy's Frontier Observatory for Research in Geothermal Energy (UTAH FORGE logs and data from the Deep Well 58-32 dataset) (Utah-EGI, 2018). The diagnostic drilling dataset is gathered using drill bits equipped with sensors, computers, and an entire networking infrastructure to collect the drilling data in real time. The dataset is a drilling dataset that covers from 0 to 7,536,25 m in depth. The collected data contains observations of the drilling rig's performance from 27 parameters.

For simplicity, this study chose five drilling-controllable parameters and two reservoir parameters: the ROP (feet/hour) as the output, and *Weight on Bit* (k-lbs), *Temp In* (F), *Temp Out* (F), *Pump Press* (psi), *Rotary Speed (rpm)*, *Surface Torque (psi), Flow in* (gpm), and *Flow Out* (%) as the input parameters. The surface-controllable parameters are the parameters that can be adjusted at the surface to affect ROP (Ben Aoun and Madarász, 2022), while the reservoir parameters (i.e., *Temp In*, *Temp Out*) are temperature changes that occur while drilling (Podgorney et al., 2021). The detailed definition of the drilling parameters can be seen in the Drilling Parameters Definition section at the end of this document. Then, this study divided the dataset into data covering 0–4,000 m in depth and data covering 4,000–7,536 m in depth to differentiate between the early stage of drilling and drilling in deeper areas with high pressure and temperature. The final dataset characteristics are summarized in Table 1.

3.1. Pre-processing techniques

Pre-processing techniques were applied to the data training to improve predictive model performance. The dataset was fixed using several pre-processing procedures such as (1) feature selection (e.g., feature selection based on domain knowledge), (2) correlation measurement (e.g., Pearson correlation coefficient), (3) outlier removal (e.g., based on data visualization using both pair plot and boxplot), (4) data scaling (e.g., min-max normalization is used when standardization cannot be applied because the predictors do not follow a Gaussian distribution), and (5) evaluation of ESML models.

1) Feature Selection Based on Domain Knowledge

This study selects the following features or predictors using domain knowledge for drilling engineering:

Target variable Y: *ROP* (feet/hour).

Predictors Xi: *Weight on Bit* (k-lbs), *Temp In* (F), *Temp Out* (F), *Pump Press* (psi), *Rotary Speed* (rpm), *Surface Torque* (psi), *Flow in* (gpm), and *Flow Out* (%).

2) Correlation Measurement

The Pearson correlation coefficient, which evaluates the linear relationship between two variables, calculates the correlation. The correlation data are summarized in Table 2.

The correlation between ROP and predictors is moderate, showing that predictors are six negatively correlated and two positively correlated. None of the predictors have strongly correlated (either positively or negatively) with the ROP. It means that the relationship between the predictors and the response variable (i.e., ROP) may be more complex or nonlinear. Weak correlations can make it difficult for individual ML models to capture the relationship accurately. In such instances, ESML models can be helpful. By leveraging diverse models, ESML models can collectively capture various aspects of the data and improve the overall predictive performance, and they also offer opportunities for exploring complex relationships and uncovering hidden patterns in the data (Abbasi et al., 2022).

3) Outlier Removal

A dataset may contain extreme values that are out of the expected range and differ from the rest. These anomalies limit the capacity of the ML algorithm to generalize. Getting rid of these outliers boosts the model's performance. There is no general approach for outlier reduction. However, data visualization is commonly used. For this purpose, both the pair plot and the boxplot are used. The pair plot findings are shown in Fig. 1.

The pair plot displays the distribution and scatter plots for each pair

Table 1
Final drilling dataset statistics.

	ROP	Weight on Bit	Temp In	Temp Out	Pump Press	Rotary Speed	Surface Torque	Flow In	Flow Out
Count	7,302,000	7,302,000	7,302,000	7,302,000	7,302,000	7,302,000	7,302,000	7,302,000	7,302,000
Mean	41,945	3,839,660	118,330	126,059	1,266,986	54,934	131,047	716,025	79,682
Std	63,950	2,144,794	11,882	12,259	486,517	24,039	45,749	125,278	11,112
Min	1469	94,590	91,107	84,070	305,081	0,000	3580	0,000	46,217
25%	11,666	1,977,285	109,025	116,130	667,125	38,613	115,954	618,227	73,146
50%	18,510	3,856,360	117,197	124,860	1,441,351	50,094	140,457	698,632	79,830
75%	44,085	5,693,857	126,819	136,490	1,660,128	78,335	1660,128	825,379	88,298
Max	818,840	7,536,250	145,636	151,7000	1,948,669	188,080	1948,669	1,103,256	104,981

Table 2
Pearson correlation results.

Predictors	Pearson Correlation between ROP and Predictors
Weight on Bit (k-lbs)	−0,523443
Temp In (F)	−0,221723
Temp Out (F)	−0,426404
Pump Press (psi)	−0,493190
Rotary Speed (rpm)	0,289070
Surface Torque (psi)	−0,380202
Flow in (gpm)	0,481611
Flow Out (%)	−0,116062

of features. Outliers were discovered in the *Flow In*, *Flow Out, Rotary Speed*, and *Surface Torque* features. The boxplot was used to explore the detail of the outlier in depth. The ROP boxplot below clearly shows that *Flow In* greater than 2500 (gpm) is an outlier (Fig. 2 (c)). *Flow Out* less than 40 (%) is considered an outlier (Fig. 2 (d)). *Rotary Speed* above 150 (rpm) is considered an outlier (Fig. 2 (e)). Finally, the *Surface Torque* of more than 250 psi is considered an outlier (Fig. 2 (g)).

4) Data Scaling

ML methods rely on minimizing the cost function. As a result, data normalization is used to speed up the gradient descent calculation. This entails utilizing a standard scale for the predictor values. Standardization cannot be used based on the probability distributions in Fig. 1 since our predictors do not have a Gaussian distribution. Instead, as described in Equation (1), min-max normalization is utilized.

$$X_{MinMax} = \frac{X - Min\ X}{Max\ X - Min\ X} \tag{1}$$

where X is a predictor value, Min X denotes the minimum value of the analyzed predictor, and Max X denotes the maximum value of the predictor.

5) Evaluation of ESML models

Finally, the performance of the ESML models was evaluated using four metrics: the coefficient of determination (R2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The coefficient of determination, or R2, is a statistical measurement that looks at how variations in one variable may be explained by the difference in another when predicting the result of an event. This measurement examines how strong the linear relationship is between two variables (predictors and response variables). The closeness of R2 values to 1 indicates how well a statistical model predicts an outcome. RMSE is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far the data points are from the regression line, and RMSE is a measure of how dispersed these residuals are. In other words, it reveals the degree to which the data is centered on the line of best fit. MAE provides the average absolute difference or error between model prediction and target value in the same dataset. MAPE averages all absolute percentage errors between the predicted and target values. It shows the mean of the absolute percentage errors of each entry in a dataset to determine how accurate the predicted quantities were compared to the actual quantities. The formulations are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \overline{x}_i)^2} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2} \tag{3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - \widehat{x}_i| \tag{4}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|x_i - \widehat{x}_i| \times 100\% \tag{5}$$

Where *xi* is the ROP compliance of the $i - th$ sample point in the dataset; $\widehat{x}_i$ Is the prediction value made by ESML models for the $i - th$ sample point; and x is the average value of the ROP compliance. MAE and RMSE explicitly characterize the residual error at each sample point and can accurately evaluate model performance. In comparison, R^2 normalizes the squared residual error with the variance of the dataset and creates a dimensionless score ranging from 0 to 1.

The experiments were conducted on a Windows 10 platform with a 16 GB graphics processing unit, a 1.80 GHz Intel Core i7 processor, 8 GB of RAM, and 256 GB of SSD storage. The Python environment (version 3.7.6), Scikit-learn, and the Keras library were used to create algorithms. Six ESML models were performed with the configuration of 80% for training data and 20% for validation. Finally, the SHAP and LIME frameworks were used for the interpretability analysis.

3.2. Development of the ESML Model and an Explainable ML Model

The development process of the ESML and explainable ML models are displayed in Fig. 3. (1) The first step was a pre-processing step, as explained in the section on materials and methods. (2) The next step was model creation. This step developed six ESML models, including RFG, GBR, CBR, ABR, XGB, and LGR. (3) Then, the hyper-parameter tuning was performed using the GridSearchCV technique and 10-fold cross-validation to determine the best model (Prusty et al., 2022). GridSearchCV adjusts hyperparameters to determine the best values for a particular model. Hyperparameter tuning used the scikit-learn GridSearchCV library. GridSearchCV goes through all the parameters provided into the parameter grid and creates the optimum combination of parameters based on scoring. The exhaustive search identified the best parameters for each different parameter of the ESML model (Agrawal, 2021). Generally, it takes four arguments: *estimator, parameter grid, cross-validation (CV)*, and *scoring*. The following is a description of the arguments: an *estimator* is the ESML model being used; a *parameter grid* is a dictionary that contains all of the parameters to try; *scoring* is an evaluation metric to use when ranking results; and *cross-validation* or CV is the number of cv folds for each combination of parameters. There is no rule for determining the appropriate parameters for each model. This is

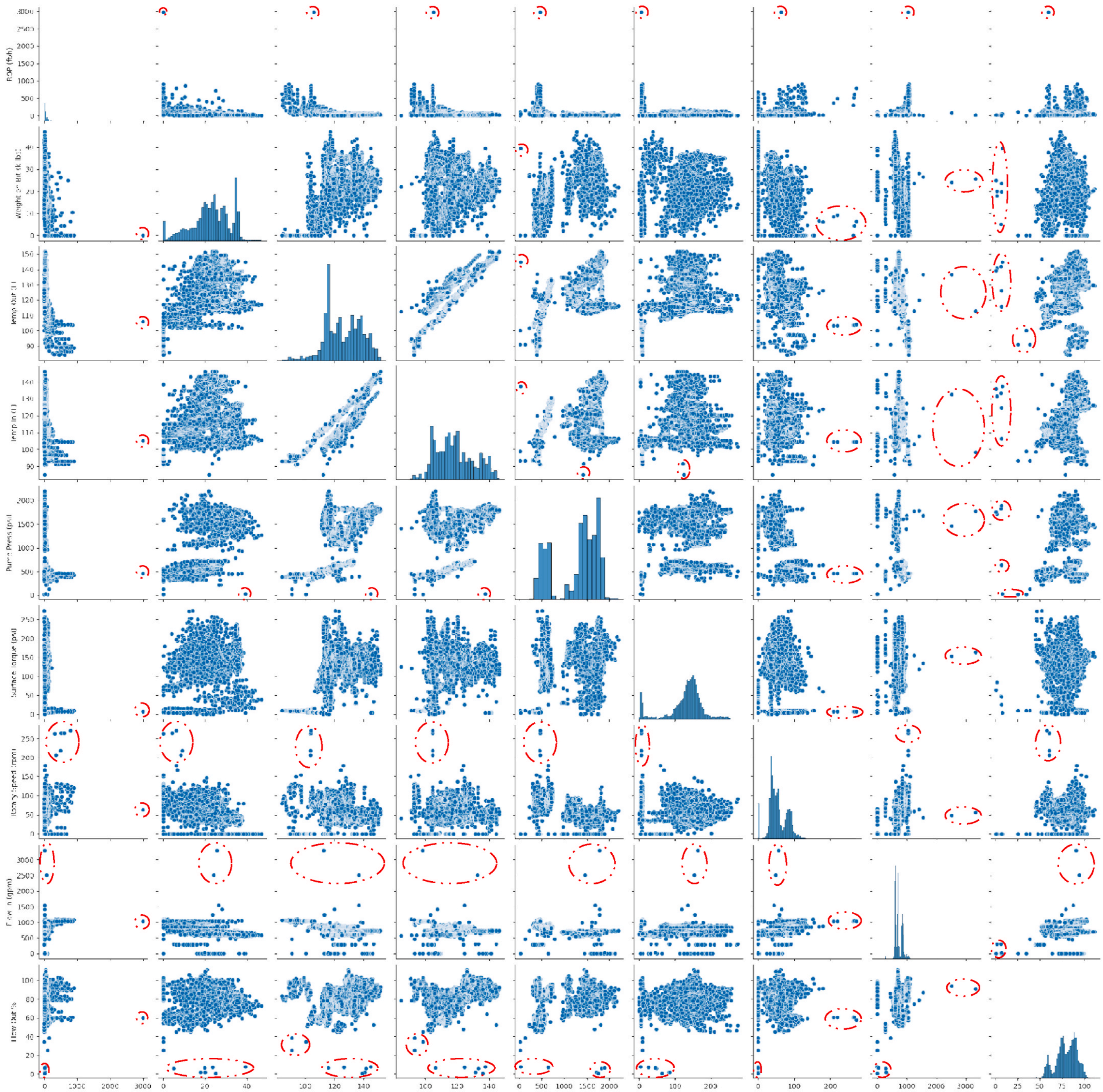


Fig. 1. Pair plot of drilling dataset with outliers encircled.

because the search can only test the parameters in the parameter grid, which can test further by adding new values and restarting the search. The parameter search range of GridSearchCV of all models is depicted in Table 3.

In Table 3, each value added to the parameter grid dictionary can significantly increase the total runtime for the best parameter search. After tuning the parameters using GridSearchCV, the best parameters of each ESML model are depicted in Table 3, column best hyper-parameters. The configuration of GridSearchCV of all models is depicted in Fig. 4. The appendix displays the detailed best parameters of each ESML model. (4) Each model was executed using a 10-fold cross-validation to obtain the best model for the interpretability analysis step. Then, four metrics, R^2, RMSE, MAE, and MAPE, were used to evaluate the performance of all ESML models.

Finally, based on the well-tuned models, an explainable ML technique using SHAP and LIME was applied to explain the underlying pattern of the prediction.

4. Results

In this section, this study compares the performance of our six ESML models and benchmarks their performance to the previous works using the same dataset. Then, this study presents the interpretability analysis for elucidating the behavior of the drilling parameters.

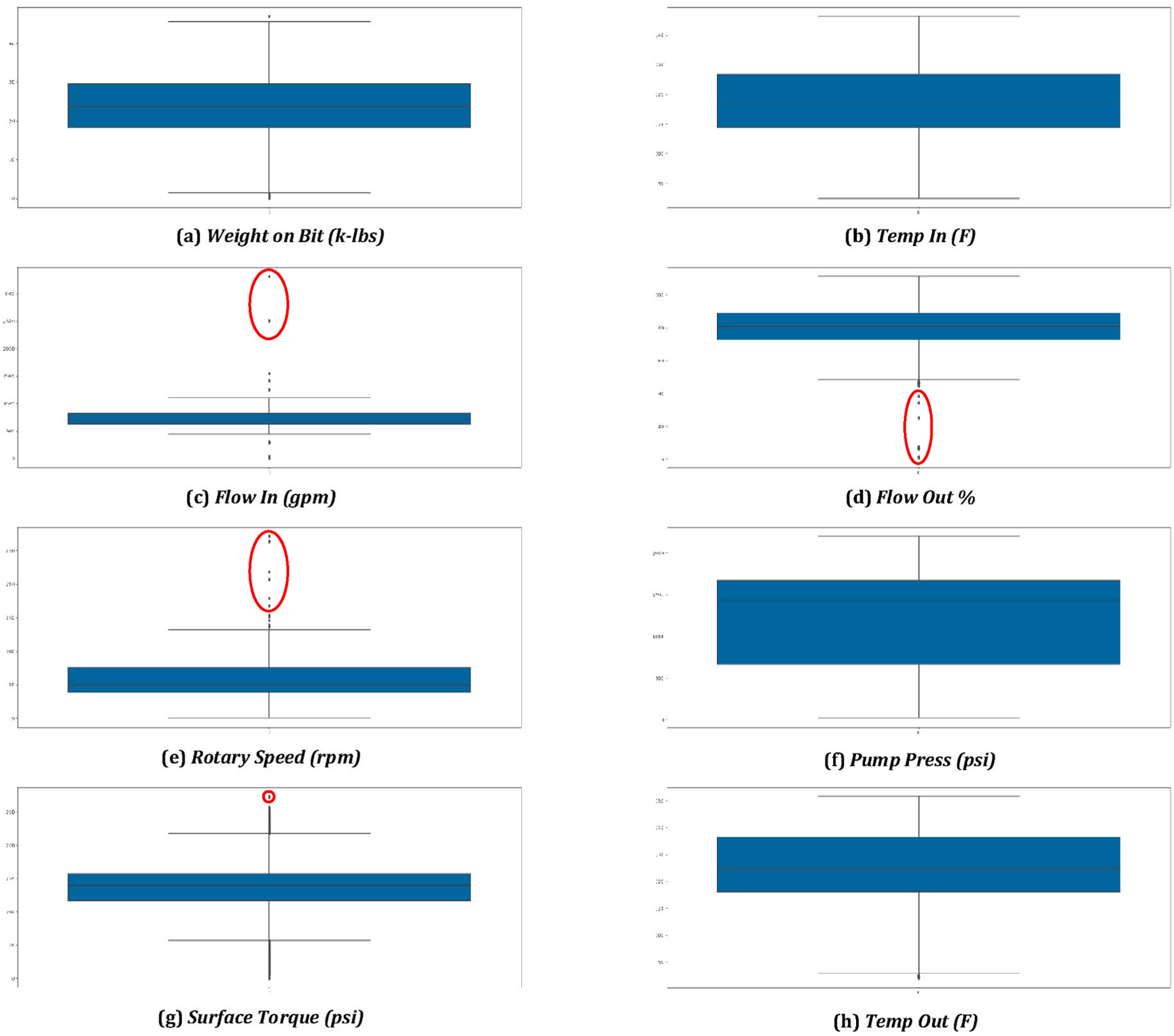


Fig. 2. Predictors boxplot.

4.1. Performance analysis

This section presents the comparative performance of six ESML models. All models were evaluated using four metrics on the test data, as depicted in Table 4. This study compares six ESMLs, namely RFG, GBR, CBR, ABR, XGB, and LGR, to determine the best prediction accuracy model. In order to choose the best model, this study examined these models using four evaluation metrics: R^2, RMSE, MAE, and MAPE.

As shown in Table 4, the R^2 of CBR is better than any other model, reaching 0,98. Besides R^2, CBR has optimal performance in RMSE and MAE. In comparison, ABR has the worst performance of these six ESML models. Thus, this study used the CBR model as a prediction model for interpretability analysis.

The results demonstrated that the best-performing models, CBR, GBR, and XGB, achieve comparable performance. Even though the findings showed the comparable performance of the models, the performance of algorithms is affected by numerous factors, such as the model's complexity, configuration, and data quality.

Our results outperformed the previous results using the same dataset (Ben Aoun and Madarász, 2022). The previous works achieved an R^2 score of 0,84 using the random forest regressor model. In contrast, our results achieved an R^2 score of 0,98, 14% higher than previous results, showing that the model explained 98% of field ROP variability. Through pre-processing and hyperparameter optimization, four ESML models (i.e., CBR, GBR, XGB, and LGR) can obtain high accuracy in the testing set ($R^2 > 0{,}90$), which vastly outperforms the previous models' R^2. The coefficient of determination of CBR, GBR, XGB, and LGR are 0,98, 0,97, 0,97, and 0,94, respectively.

In another domain, the CBR model reportedly outperformed XGB in predicting fiber-reinforced polymer (FRP) composite materials (Kim et al., 2022). Similarly, after hyperparameters tuning, CatBoost outperformed other models (i.e., ABR, LGR, XGB, and RFG) in predicting photovoltaic maximum current and achieved great accuracy above 99% in performance when applied to a testing dataset (Omer and Shareef, 2022). Further, CBR reportedly outperformed XGB, RFG, ABR, and GBR in predicting loan approvals and staff promotion (Ibrahim et al., 2020). However, in another paper, XGB and LGR exceeded CBR in predicting loan of home credit (Al Daoud, 2019). Kolesnikov et al. (2019) also found that XGB surpassed CBR in predicting the spread of dengue fever. These results confirm that the pre-processing and cross-validated hyperparameter tuning improves the model's predictability. Although the findings showed the comparable performance of ESML (i.e., CBR, GBR, XGB, and LGR), the authors noticed that numerous factors influence algorithm performance, such as the model's complexity, the quality and size of the training data, feature selection and engineering, hyperparameter tuning, model evaluation and validation, and computing resources.

4.2. Interpretability analysis

In this section, this study provides the explainable ML results for global explanations and local explanations from the SHAP and LIME

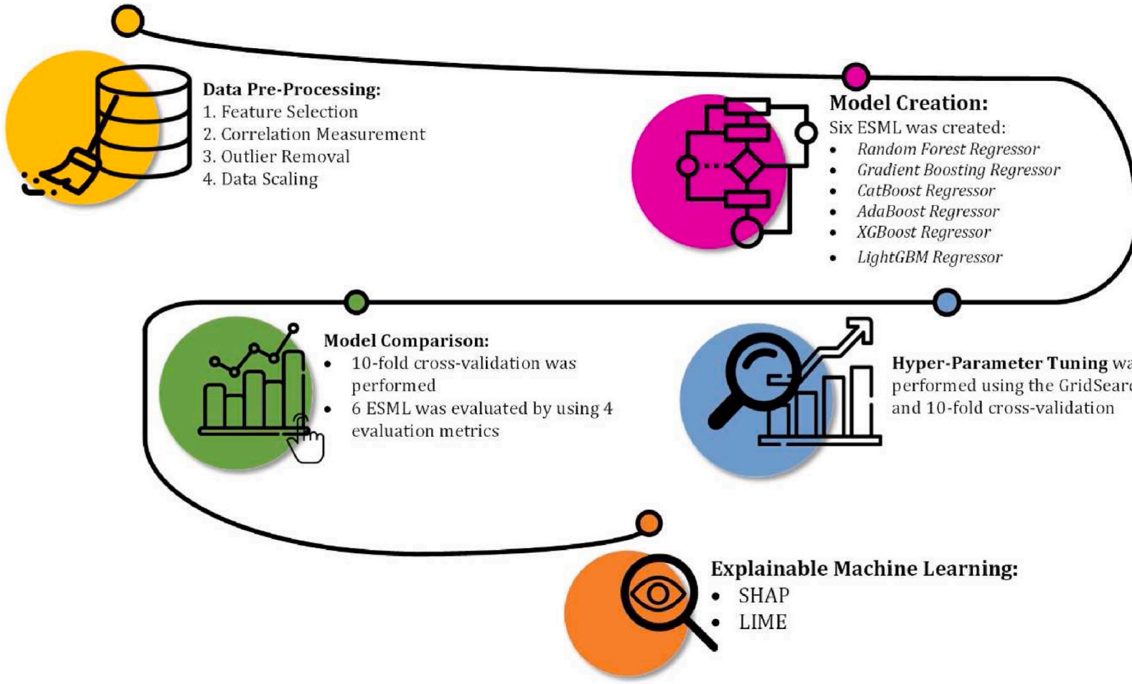


Fig. 3. The development process of ESML and an explainable ML.

explainers.

4.2.1. *Global interpretability analysis*

Global interpretability analysis aims to describe the expected behavior of a machine learning model concerning the entire distribution of values for its input variables. Averaging the SHAP values for all instances over the population is accomplished using SHAP. SHAP examines each feature's mean absolute SHAP value across all data. This measures the average contribution of each feature to the ROP prediction, whether positive or negative.

Based on the calculated SHAP values, the results show the influence of various features on ROP prediction. Fig. 5 (a) and (b) shows a SHAP summary plot in which each dot represents an individual data point in the dataset. The position of the dots on the x-axis indicates the impact of the individual feature's value on the ROP prediction. The y-axis displays the variable names in ascending order of importance from top to bottom. The red dots correspond to higher SHAP values, which indicate high ROP values. The blue dots correspond to lower SHAP values, which indicate lower ROP values. A positive SHAP value impacts prediction, leading the model to a higher ROP prediction. A negative SHAP value means a negative impact, leading the model to a lower ROP prediction. For example, Fig. 5 (a) shows that higher *Flow In* (red dots) correspond to higher SHAP values, which indicate high ROP values.

Fig. 5 (a) shows that the top four features for 0–4,000 m in depth that primarily influence the global ROP predictions, from the highest to the lowest order, are *Temp In*, *Pump Press, Temp Out,* and *Flow In*. High values of *Temp In*, *Pump Press,* and *Temp Out* have a high negative contribution (represented by blue dots) on the prediction, while low values have a positive contribution (red dots). *Flow In* and *Rotary Speed* have a high positive contribution when their values are high, represented by red dots, which means the feature pushes the ROP prediction higher (i.e., higher Shapley values on the x-axis). Based on the interpretability of the SHAP framework, this study can hypothesize that the relative importance of the drilling parameters on ROP prediction is in the order of *Temp in* > *Pump Press* > *Temp Out* > *Flow In* > *Weight on Bit* > *Rotary Speed* > *Flow Out* > *Surface Torque*.

Fig. 5 (b) shows that the top four features for 4,000–7,536 m in depth that primarily influence the global ROP predictions, from the highest to the lowest order, are *Temp Out*, *Pump Press, Weight on Bit,* and *Temp In*. High values of *Temp Out*, *Pump Press, Weight on Bit,* and *Temp In* have a high positive contribution (represented by red dots) on the prediction, while low values have a negative contribution (blue dots). This shows that higher *Temp Out*, *Pump Press, Weight on Bit,* and *Temp In* (the dot extending towards the right is increasingly red) lead to higher ROP predictions. *Flow In* and *Flow Out* have SHAP values (x-axis) on the positive side, represented by a blue dot (the dot extending towards the right is increasingly blue). This indicates that *Flow In* and *Flow Out* have lower predicted ROP. Based on the interpretability of the SHAP framework, this study can hypothesize that the relative importance of the drilling parameters on ROP prediction is in the order of *Temp Out* > *Pump Press* > *Weight on Bit* > *Temp In* > *Surface Torque* > *Rotary Speed* > *Flow In* > *Flow Out*.

4.2.2. *Local interpretability analysis*

In this section, the local interpretability analysis is described. Local interpretability aims to explain each sample's prediction and elaborate on how the model predictions differ from the feature values and interactions. Each input parameter's SHAP value can describe each feature's contribution.

4.2.2.1. Dependency and interaction analysis. As shown in Fig. 6, the value of features and corresponding SHAP values are plotted on the x-axes (the value of individual features) and y-axes (the SHAP value), respectively. The dependence plots are further highlighted by feature interactions (through color bars), which indicate the combined effects of features on prediction. It should be noted that the SHAP values do not represent causal linkages but rather characterize the model's behavior.

Fig. 6 (a) shows the impact of features on ROP for 0–4,000 m in depth. The dependence and interaction effects indicate lower ROP values (red dots) occurred when *Temp In* $\geq$ 114 °F and *Pump Press* $\geq$ 536 psi. High values of *Temp In* and *Pump Press* drive down the SHAP values, corresponding to lower ROP values. Similarly, high *Temp In* and *Temp Out* values drive down ROP values. Furthermore, Fig. 6 (a) demonstrates a moderate positive relationship between *Flow In* and *Rotary Speed*. Higher ROP values (red dots) can be obtained when *Flow In* $\geq$ 859 gpm and *Rotary Speed* $\geq$ 85 rpm.

Table 3
The best parameters for each ESML model based on the GridSearchCV.

Model	Proposed	Search Range	Best Hyper-Parameters
RFG	Max Depth	[700, 800, 900]	900
	Number of estimators	[500, 600, 700, 800, 900]	600
	Minimum sample leaf	[1, 2, 3, 4, 5]	1
	Minimum sample split	[1, 2, 3, 4, 5]	2
GBR	Max Depth	[3, 4, 5, 6, 7, 8]	3
	Number of estimators	[10, 50, 100, 300]	100
	Learning rate	[0,0001, 0,001, 0,01, 0,1]	0,1
	Subsample	[0,5, 0,7, 1,0]	1,0
	Minimum sample leaf	[1, 2, 3, 4, 5]	1
	Minimum sample split	[1, 2, 3, 4, 5]	2
CBR	Max Depth	[5, 6, 7, 8]	5
	Number of estimators	[400, 500, 600]	600
	Learning rate	[0,01, 0,05, 0,1]	0,01
	Minimum sample split	[1, 2, 3, 4, 5]	3
ABR	Number of estimators	[30, 40, 50]	50
	Learning rate	[0,01, 0,1, 1]	1
XGB	Max Depth	[4, 5, 6]	6
	Number of estimators	[100, 200, 300]	100
	Learning rate	[0,1, 0,2, 0,25, 0,300000012, 0,4]	0,300
	Minimum child weight	[1, 5, 10]	1
LGR	Max Depth	[3, 4, 5]	5
	Number of estimators	[777, 888, 999]	999
	Learning rate	[0,1, 0,03, 0,05]	0,05
	Number of leaves	[7, 14, 21, 28, 31, 50]	31
	Minimum child weight	[0,001, 0,01, 0,1]	0,001
	Minimum child sample	[10, 20, 30, 40, 50]	20

Fig. 6 (b) shows the impact of features on ROP for 0–7,536 m in depth. The dependence and interaction effects indicate higher ROP values (red dots) when *Temp Out* $\geq$ 123 °F and *Pump Press* $\geq$ 1,387 psi. Also, it reveals that higher values of *Weight on Bit*, *Temp In*, *Surface Torque,* and *Rotary Speed* drive up the SHAP values, corresponding to higher ROP values. Higher ROP values (red dots) can be obtained when *Weight on Bit* $\geq$ 29 k-lbs, *Temp In* $\geq$ 111 °F, *Surface Torque* $\geq$ 137gpm, and higher *Rotary Speed* $\geq$ 35 rpm.

4.2.2.2. The inflection points. To further investigate the critical inflection points that drive the high and low ROP rates, this study performed a detailed analysis with the LIME framework, as shown in Fig. 7. These plots explain the predictions of 18 instances from the ESML models for the ROP prediction to enhance their explainability. The inflection points are a perspective that explains the parameter at which the change from low to high ROP occurs. This analysis explains the critical inflection points of the predictor variables that trigger the transition between high

Table 4
Comparison of ESML models.

Model	R^2	RMSE	MAPE	MAE
RFG	0,88	34,92	1,42	26,72
GBR	0,97	0,17	0,42	0,13
CBR	0,98	0,14	0,55	0,11
ABR	0,78	48,97	1,50	38,66
XGB	0,97	10,43	0,11	3,75
LGR	0,94	23,93	0,94	17,13

*The value of each metric was repeated ten times, and the average of the value was used for Table 4.

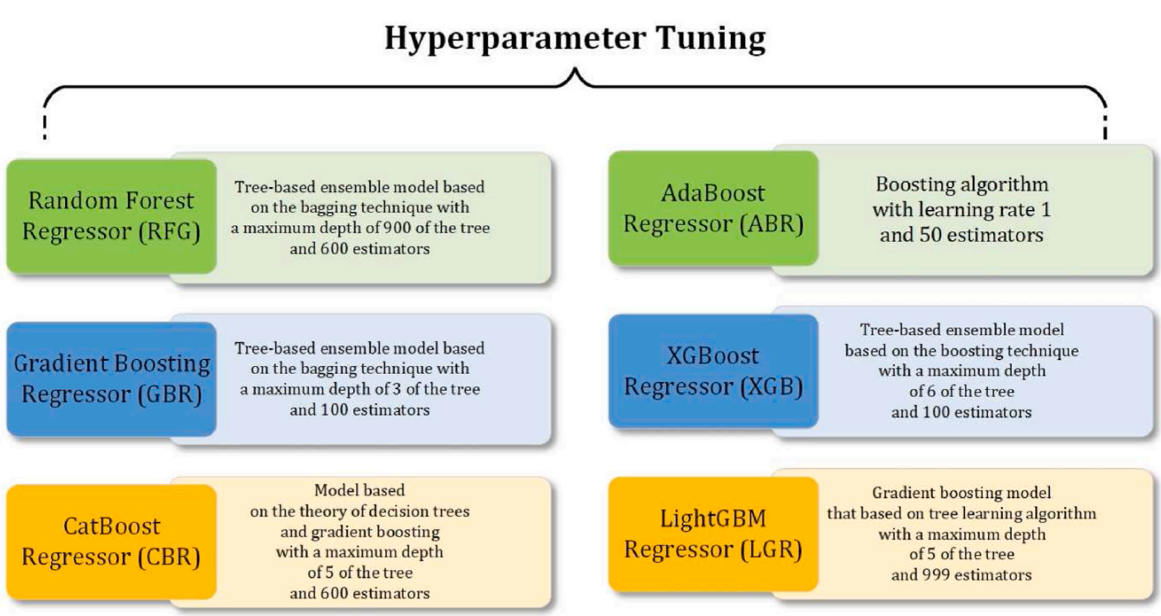
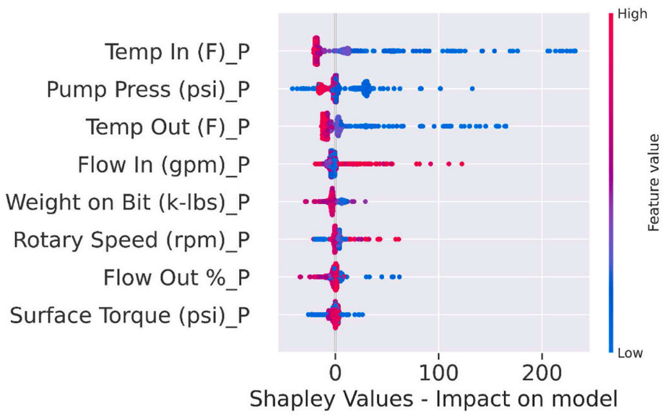
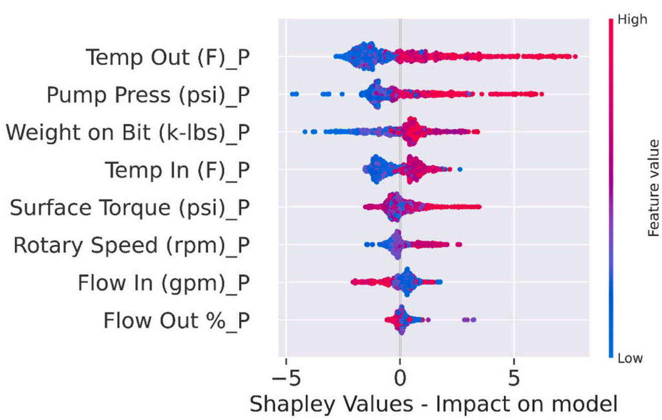


Fig. 4. Details of the ensemble ML models.



(a)



(b)

Fig. 5. (a) Global feature influences on ROP at depths from 0 to 4,000 m. (b) Global feature influences on ROP at depths from 4,000–7,536 m. E.g., in (a), *Temp In* is the most influential feature, while *Surface Torque* is the least influential feature for ROP predictions. A higher SHAP value indicates that the model predicts higher ROP values and vice-versa.

and low ROP values.

Fig. 7 (a) explains the ROP prediction models for 0–4,000 m depth. It reveals that the inflection points of *Temp In, Pump Press*, *Temp Out, Flow In,* and *Rotary Speed* are 114,07 °F, 536,04 psi, 116,25 °F, 859,31 gpm, and 85,50 rpm, respectively. *Weight on Bit, Flow Out,* and *Surface Torque* can be ignored since they have a minimal impact on ROP, according to the global interpretability analysis in Fig. 6 (a). Based on the above interpretability, this study hypothesizes:

- H1: if *Temp In* ≥ ~114,07 °F, *Pump Press* ≥ ~536,04 psi, *Temp Out* ≥ ~116,25 °F, *Flow In* ≤ ~859,31 gpm, and *Rotary Speed* ≤ ~85,50 rpm, the ROP value will be low.
- H2: if *Temp In* ≤ ~114,07 °F, *Pump Press* ≤ ~536,04 psi, *Temp Out* ≤ ~116,25 °F, *Flow In* ≥ ~859,31 gpm, and *Rotary Speed* ≥ ~85,50 rpm, the ROP value will be high.

Fig. 7 (b) explains the ROP prediction models for 4,000–7,536 m depth. The inflection points of *Temp Out, Pump Press*, *Weight on Bit, Temp In, Surface Torque,* and *Rotary Speed* are 123,64 °F, 1,387,75 psi, 29,31 k-lbs, 111,31 °F, 137,83 gpm, and 35,27 rpm, respectively. *Flow In* and *Flow Out* can be ignored since they have a minimal impact on ROP, according to the global interpretability analysis in Fig. 6 (b). Based on the above interpretability, this study hypothesizes:

- H3: if *Temp Out* ≤ ~123,64 °F, *Pump Press* ≤ ~1,387,75 psi, *Weight on Bit* ≤ ~29,31 k-lbs, *Temp In* ≤ ~111,31 °F, *Surface Torque* ≤ ~137,83 gpm and *Rotary Speed* ≤ ~35,27 rpm, the ROP value will be low.
- H4: if *Temp Out* ≥ ~123,64 °F, *Pump Press* ≥ ~1,387,75 psi, *Weight on Bit* ≥ ~29,31 k-lbs, *Temp In* ≥ ~111,31 °F, *Surface Torque* ≥ ~137,83 gpm and *Rotary Speed* ≥ ~35,27 rpm, the ROP value will be high.

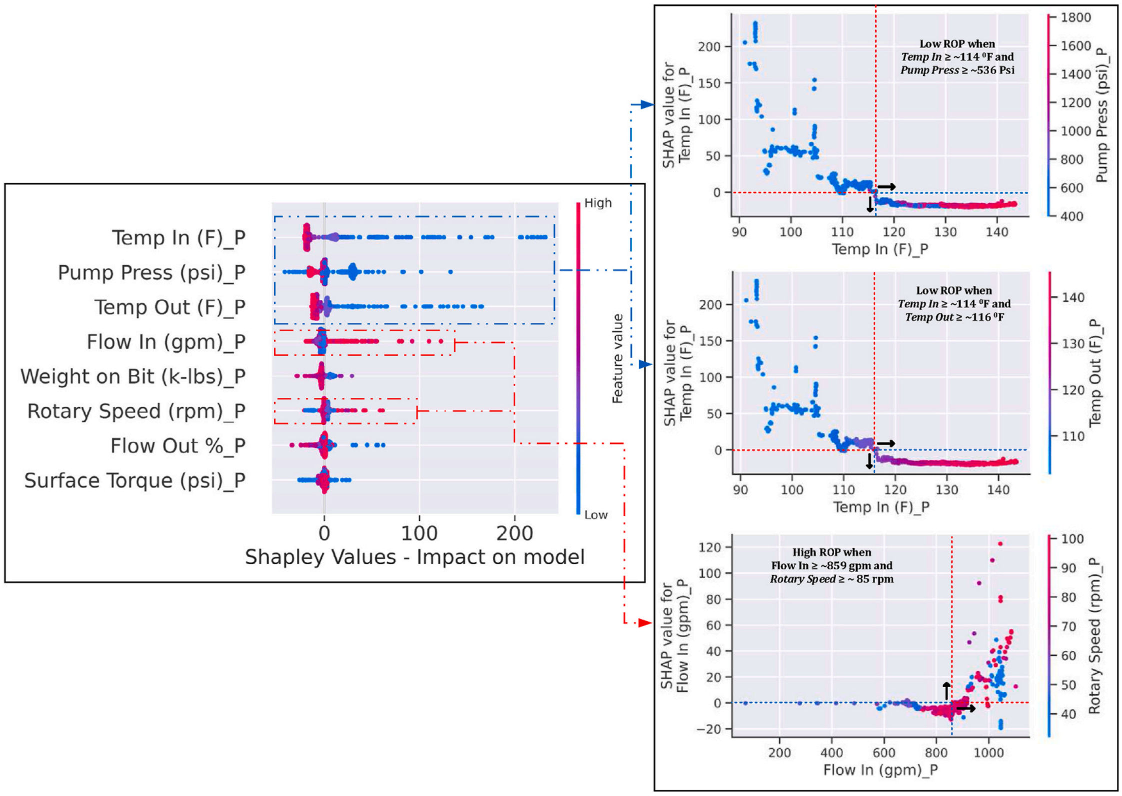
This analysis can be used for engineers or future studies on strategically controlling drilling parameters and adjusting the constraints of optimal parameters for various scenarios. With such knowledge, engineers can manipulate drilling parameters and simulate them using an ML model to ensure an optimal ROP without requiring expensive and time-consuming experimental tests. This analysis is also essential for understanding how the model reached a specific result.

5. Discussion

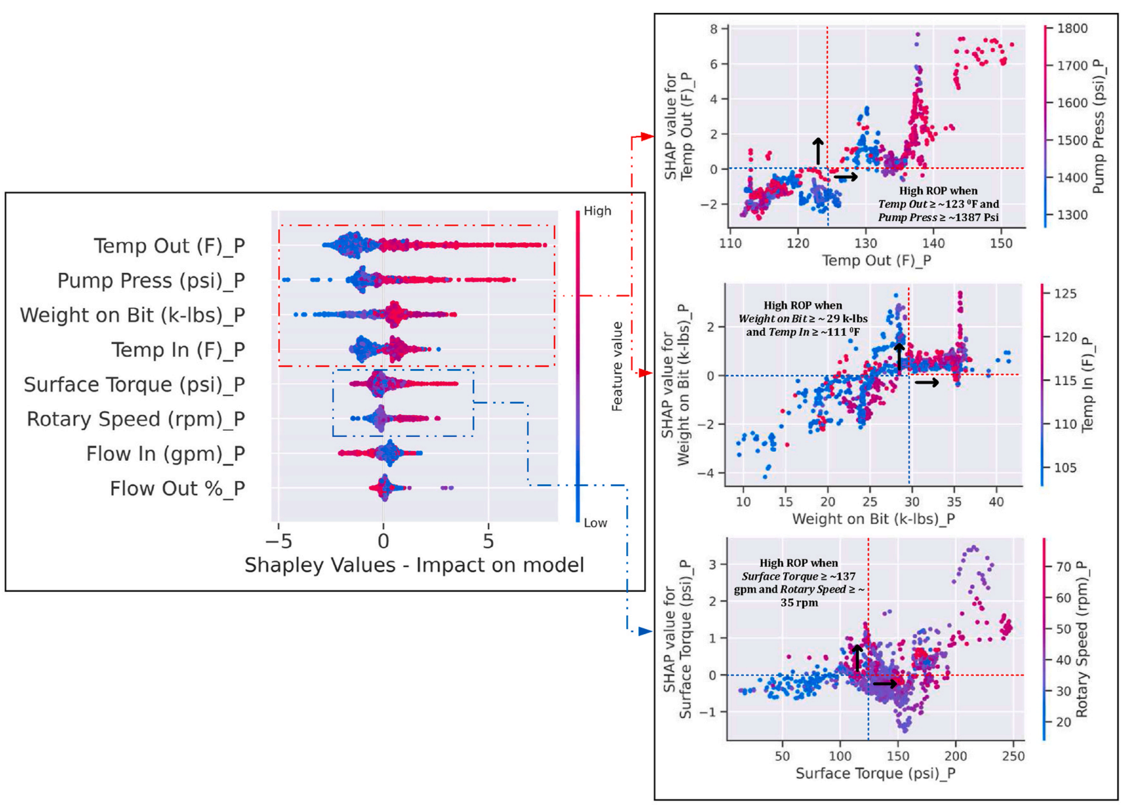
Recent previous works have presented several models to predict ROP (Ben Aoun and Madarász, 2022). Their work showed that the random forest regressor achieved the best performance for the UTAH FORGE dataset, achieving an R^2 score of 0,84. Our method outperformed those results using the same dataset. Using the pre-processing technique and hyperparameter optimization, the CBR achieved an R^2 score of 0,98, meaning that the model explained 98% of field ROP variability, which is 14% higher than previous work.

Further, the results of the global interpretability analysis for 0–4,000 m in depth show that the following features (in order of importance) have the most impact on ROP prediction: *Temp in* > *Pump Press* > *Temp Out* > *Flow In* > *Weight on Bit* > *Rotary Speed* > *Flow Out* > *Surface Torque*. Three features, *Temp In, Pump Press,* and *Temp Out,* negatively contribute to the ROP. That is, a low value (blue dots) for these features is strongly expected to lead to a low ROP. In contrast, *Flow In* and *Rotary Speed* contribute positively to the ROP. A high value (red dots) for these features drives the model output toward high ROP.

As for the local interpretability analysis, the study quantified the critical inflection points in each predictor that led to the transition from low ROP to high ROP rates: (H1) if *Temp In* ≥ ~114,07 °F, *Pump Press* ≥ ~536,04 psi, *Temp Out* ≥ ~116,25 °F, *Flow In* ≤ ~859,31 gpm, and *Rotary Speed* ≤ ~ 85,50 rpm, the ROP value will be low and (H2) if *Temp In* ≤ ~114,07 °F, *Pump Press* ≤ ~536,04 psi, *Temp Out* ≤ ~116,25 °F, *Flow In* ≥ ~859,31 gpm, and *Rotary Speed* ≥ ~ 85,50 rpm, the ROP value will be high. From a practical drilling standpoint, when the drills penetrate deeper into the earth, pressures and temperatures (i.e., *Temp In, Pump Press*, and *Temp Out*) push drilling equipment to its limits, and ROP can drop to a few feet per minute (Gautam et al., 2022; Mohamed et al., 2021). These results agree with the findings of other studies, which found that higher pressures and temperatures can exert significant mechanical stress on drilling equipment. Excessive pressure can cause structural deformation, leading to equipment failure. Similarly, elevated temperatures can weaken the materials and decrease their strength, making them more prone to damage. Thus, it can drop the ROP (He et al., 2021). Previous study also reported that higher pressures and temperatures can result in accelerated wear of the drill bit, and increased energy requirements, which can lower the ROP (Wang et al., 2023). The high pressures and temperatures can influence the properties of drilling fluids, such as viscosity and density. High pressures can compress the drilling fluid, making it denser and potentially more challenging to pump, which can hinder the circulation and cooling of the drilling fluid. Raised temperatures can also affect the viscosity, reducing its lubrication properties and potentially increasing friction between the drill string and the wellbore walls (Mao et al., 2020; Palaoro et al., 2022). Lastly, the high pressures and temperatures can lead to rock instability, and elevated temperatures can cause rock formations to become more plastic or prone to sloughing. This instability can increase the risk of stuck pipes or other drilling problems, impacting the ROP (Sygała et al., 2013). Thus, it is reasonable to argue that higher pressures

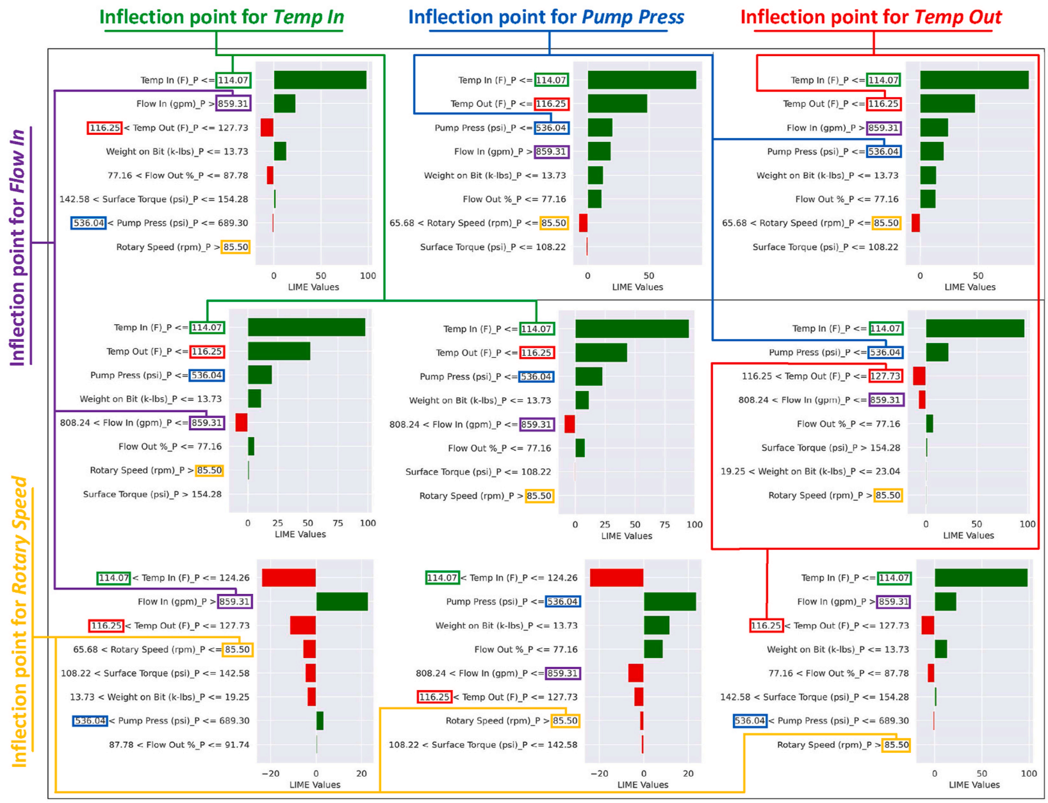


(a)

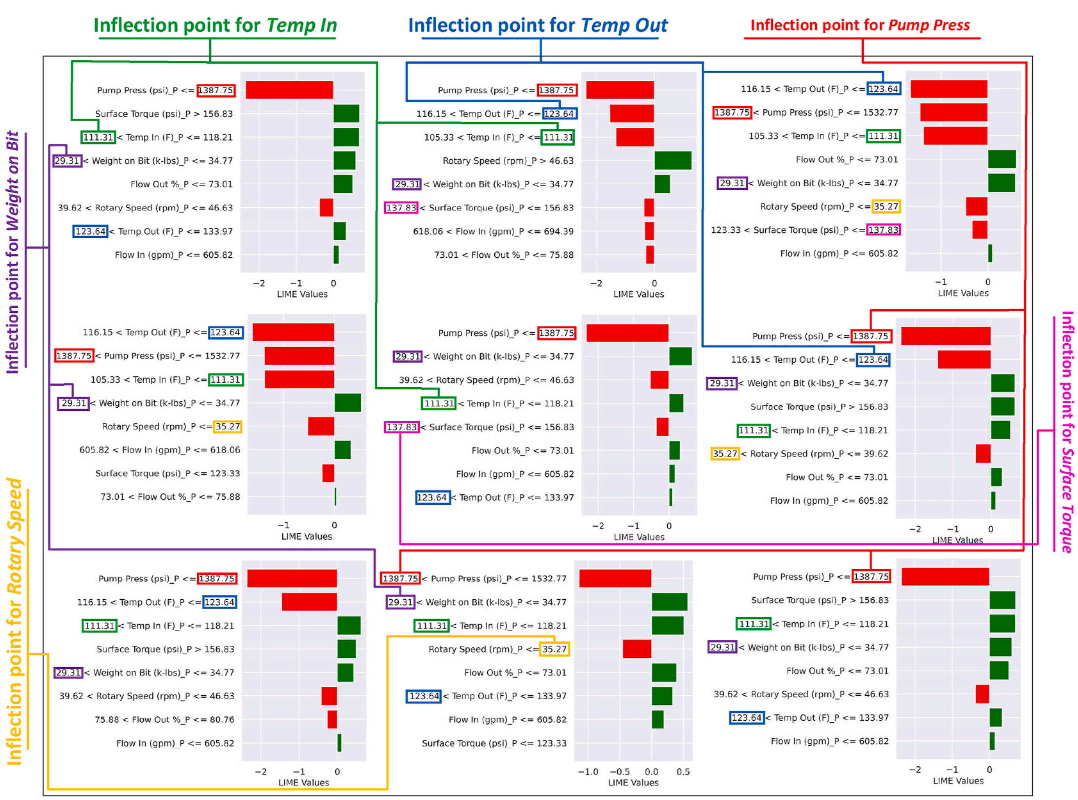


(b)

Fig. 6. (a) The impact of features on ROP for 0–4,000 m in depth; (b) The impact of features on ROP for 4,000–7,536 m in depth.



(a)



(b)

Fig. 7. (a) The inflection points for 0–4,000 m in depth; (b) The inflection points for 4,000–7,536 m in depth.

and temperatures would result in lower ROP.

Flow in and *Rotary Speed* shows a positive impact on ROP, which is also reasonable, considering that an increase in flow rate also increases the hydraulic power that assists in improving the drilling performance because fluid flow through the drill string supports spins the bit, leading to higher rotation speeds and ROP (Cheraghi et al., 2022; Liu et al., 2021). Higher *Rotary Speed* can help decrease the friction and torque between the bit and the formation, making the drilling process smoother. By minimizing friction, the bit encounters less resistance while drilling, leading to an increased ROP (Bani Mustafa et al., 2021).

Meanwhile, the global interpretability analysis for 4,000–7,536 m in depth shows that the following features contribute most to the ROP prediction: *Temp Out* > *Pump Press* > *Weight on Bit* > *Temp In* > *Surface Torque* > *Rotary Speed* > *Flow In* > *Flow Out*. Three features, *Temp Out, Pump Press,* and *Weight on Bit,* positively contribute to the ROP. High values (red dots) for these features strongly influence a higher ROP. In contrast, *Flow In* and *Flow Out* (blue dots) drive the model output toward low ROP.

As for the local interpretability analysis, the study presents two critical inflection points: (H3) if *Temp Out* $\leq \sim$123,64 °F, *Pump Press* $\leq \sim$1,387,75 psi, *Weight on Bit* $\leq \sim$ 29,31 k-lbs, *Temp In* $\leq \sim$111,31 °F, *Surface Torque* $\leq \sim$137,83 gpm and *Rotary Speed* $\leq \sim$ 35,27 rpm, the ROP value will be low. (H4) if *Temp Out* $\geq \sim$123,64 °F, *Pump Press* $\geq \sim$1,387,75 psi, *Weight on Bit* $\geq \sim$ 29,31 k-lbs, *Temp In* $\geq \sim$111,31 °F, *Surface Torque* $\geq \sim$137,83 gpm and *Rotary Speed* $\geq \sim$ 35,27 rpm, the ROP value will be high. Under these more demanding drilling conditions, when the drills penetrate deeper into the earth, pressures and temperatures (i.e., *Temp Out, Pump Press*, and *Temp In*) still significantly impact ROP.

However, in contrast to 0–4,000 m in depth, the higher pressures and temperature result in a higher ROP. This is consistent with the results that an increase in *Surface Torque, Weight on Bit,* and *Rotary Speed* also increases drilling performance, which is reasonable, considering that when the drills penetrate deep, the engineers will push their drilling equipment to maximum limits. One interesting result discovered was that in contrast to 0–4,000 m in depth, the flow rate (i.e., *Flow In* and *Flow Out*), which should cause a high ROP as the drills penetrate deep, actually showed a negative contribution to ROP. It seems that ROP tends to decrease with increasing viscosity of flow rate due to mechanical friction (plastic viscosity) induced during drilling, influencing a reduction in bit performance (Alum and Egbon, 2011). Thus, it can be inferred that low values for those flow rates lead to a high ROP (see Fig. 5 (b) represented by red dots of SHAP values). This indicates that a high ROP is attainable at 4,000–7,536 m in depth while minimizing the flow rate or volume of liquid that moves through a pipe, which may lead to cost savings in terms of the quantity of fluid required to achieve the maximum ROP by controlling these parameters. Nevertheless, it is crucial to remember that the optimal drilling fluid characteristics and rotation speed may have limitations. Excessive fluid flow rates or excessively high rotational speeds might lead to issues including excessive vibration, bit damage, or fluid loss (Badrouchi et al., 2022). Finding the proper balance and improving these parameters based on the unique drilling circumstances and formation features is critical for maximizing ROP.

Thus far, ESML models showed similar results between drilling at 0–4,000 m and 4,000–7,536 m in depth regarding global interpretability analysis results. Temperature and pressure contribute to a lower ROP in earlier drilling (0–4,000 m in depth). In contrast, *Flow In* and *Rotary Speed* contributes to a higher ROP. This indicates that a high ROP is attainable while maximizing the drills' flow rate and rotation speeds (see the inflection points for 0–4,000 m in depth). On the other hand, as the drills go deeper into the earth (i.e., 4,000–7,536 m in depth), maximum ROP can be obtained by maximizing the *Pump Press, Weight on Bit, Surface Torque,* and *Rotary Speed*, while flow rate (i.e., *Flow In* and *Flow Out*) do not significantly contribute to ROP.

The dataset generally reflects a geothermal drilling operation's expected behavior, consistent with this section's global and local interpretability analysis. Many factors can influence the ROP, and a comprehensive study is required regarding the interaction between the drilling bit and the environment's properties (i.e., hard rock), bit characteristics, and drilling parameters. This study does not claim that our outcomes, the understanding of controllable drilling parameters and using them in the drilling process, will directly maximize ROP; instead, based on data analysis, this study suggests that the controllable drilling parameters can be adjusted based on the constraints of optimal, ML-derived parameters, potentially leading to maximizing ROP for drilling effectiveness. Our ESML model can simulate the ROP without requiring expensive and time-consuming experimental tests. However, more experimental investigations, such as studies using different datasets, are needed to understand the big picture of the drilling process.

The overall result confirms that our model with pre-processing and hyperparameter optimization improves predictability and outperforms the previous results using the same dataset. The explainable ML technique using a SHAP and LIME framework successfully reveals the behavior of controllable drilling parameters that can enhance the explainability and trustworthiness of the ML models.

6. Conclusions

The overarching goal of this paper is to examine the feasibility of an explainable ML approach for elucidating the behavior of the drilling parameters to generate explainable insights regarding the drilling parameters that could lead to maximizing the ROP for DE.

This study presents several pre-processing procedures and hyperparameter optimization for the ESML models to generate the best model. The results of the comparative analysis revealed that the CBR or CatBoost model produces the most accurate predictions compared to other models and to any of the other models that have been utilized in the literature.

Then, this study applies the explainable ML to elucidate the drilling parameters' behavior that could lead to maximizing the ROP for DE. This study demonstrates that the explainable ML can generate new knowledge regarding the drilling parameters' behavior on the ROP prediction. This study confirms that the explainability of ML was enhanced when the models were coupled with global and local interpretability analysis. Explainability in this context refers to fidelity (i.e., the prediction provided by an explanation should be as close to the real-world geothermal drilling operations as possible, and the explanation should be simple enough to understand). Such explanations can be used to establish testable hypotheses or demonstrate that the model resonates with real-world physical systems (Saranya and Subhashini, 2023).

Our global interpretability analysis revealed that *Temp In, Pump Press,* and *Temp Out* are the most influential controllable parameters at the 0–4,000 m depth, while *Temp Out, Pump Press, Weight on Bit*, and *Temp In* are the most influential controllable parameters in 4,000–7,536 m depth. Besides that, this study also presented a local interpretability analysis that quantified the critical inflection points in each predictor that led to the transition from low ROP to high ROP rates. This study virtually identified that if *Temp In* $\leq \sim$114,07 °F, *Pump Press* $\leq \sim$536,04 psi, *Temp Out* $\leq \sim$116,25 °F, *Flow In* $\geq \sim$859,31 gpm, and *Rotary Speed* $\geq \sim$85,50 rpm, the ROP value will be high at 0–4,000 m depth and if *Temp Out* $\geq \sim$123,64 °F, *Pump Press* $\geq \sim$1,387,75 psi, *Weight on Bit* $\geq \sim$29,31 k-lbs, *Temp In* $\geq \sim$111,31 °F, *Surface Torque* $\geq \sim$137,83 gpm and *Rotary Speed* $\geq \sim$35,27 rpm, the ROP value will be high at 4,000–7,536 m depth. These findings can be used as guidance to select the appropriate ROP settings for DE.

Furthermore, the knowledge gained from this study could inform future studies on understanding the behavior of different drilling parameters. In practical, challenging scenarios during drilling, many complex features can affect the ROP. Hence, finding the best settings for controllable drilling parameters to maximize the ROP is essential. The proposed application has shown the ability to explain the controllable drilling parameters that can be used to minimize the risks and costs of geothermal drilling. No additional time or money will be required because the chosen parameters are typically collected at the surface during drilling operations since the predictive variables are generated from the geothermal site in real time. In conclusion, the findings suggest that the strength of the proposed model can be expected to increase users' confidence and support future decision-making processes in geothermal drilling.

Future research directions can be proposed: (1) This study only uses a limited number of surface controllable drilling parameters as inputs for the proposed explainable ML. Other drilling parameters, such as depth, hook load, temperatures out, surface torque, standpipe pressure, iron roughneck torque, rock strength, pore pressure, mud weight, and wellbore trajectory, are also crucial to research to highlight different aspects of the drilling process. (2) Our interpretability analysis was based on the UTAH FORGE dataset, so datasets from other drilling sites were not explored. Future works should use data from other sites to obtain a broader understanding of the drilling process. (3) Finally, the proposed model using explainable ML can potentially be extended to build intelligent models that are interpretable and transferable. It can potentially be applied to other related fields, such as intelligent prediction and optimization of a well's trajectory, intelligent warning and control of drilling risks, intelligent evaluation and optimization of cementing quality, intelligent design and optimization of the fracturing process, intelligent design and optimization of drilling completion, and overall optimization and intelligent decision-making of the geothermal drilling process.

Credit author statement

Zhipeng Feng: Conceptualization, Investigation, Writing-Reviewing and Editing; Hamdan Gani: Conceptualization, Investigation, Methodology, Software, Data Curation, Visualization, Writing-Original Draft; Annisa Dwi Damayanti: Writing-Reviewing and Editing, Formal Analysis, Resources, Supervision, Project administration, and Funding acquisition; Helmy Gani: Writing- Reviewing and Editing, Investigation, and Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The shared dataset is available on this link (https://dx.doi.org/10.15121/1452735).

Acknowledgments

The authors thank the Energy and Geoscience Institute at the University of Utah for providing the data set on their webpage.

Nomenclature

Abbreviations

ABR	Adaboost Regressor
CBR	CatBoost Regressor
DE	Drilling Effectiveness
ESML	Ensemble Supervised Machine Learning
F	Fahrenheit
GBR	Gradient Boosting Regressor
GPM	Gallons Per Minute
K-LBS	Kilo Pounds
LGR	LightGBM Regressor
LIME	Local Interpretable Model-Agnostic Explanations
M	Meter
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
PSI	Pounds per Square Inch
R^2	Coefficient of Determination
RFG	Random Forest Regressor
RMSE	Root Mean Square Error
ROP	Rate of Penetration
RPM	Revolutions Per Minute
SHAP	SHapley Additive exPlanation
WOB	Weight on Bit
XGB	Extreme Gradient Boosting Regressor

Drilling Parameters Definition

Abbreviations	*definitions*
Flow In	Drilling fluid flow rate in (gallons per minute)
Flow Out	Drilling fluid flow rate out (gallons per minute)
Pump Press	Drilling pressure(psi)
ROP	Rate of penetration (feet/hour) is the speed at which a drill bit breaks the rocks
Rotary Speed	The number of complete rotations made by the drill bit (revolutions per minute)
Surface Torque	The rotational force between the drill string and the rock (psi)
Temp In	The temperature in while drilling (Fahrenheit)
Temp Out	The temperature out while drilling (Fahrenheit)
WOB	The amount of downward force exerted on the drill bit (kilo pounds)

Appendix A

https://www.dropbox.com/s/xyvshag43i9sgkj/Appendix%20Geothermal.docx?dl=0.

References

Abbas, A.K., Rushdi, S., Alsaba, M., Al Dushaishi, M.F., 2019. Drilling rate of penetration prediction of high-angled wells using artificial neural networks. J. Energy Resour. Technol. 141 (11), 1–11. https://asmedigitalcollection.asme.org/energyresources/article/doi/10.1115/1.4043699/727157/Drilling-Rate-of-Penetration-Prediction-of.

Abbasi, E., Alavi Moghaddam, M.R., Kowsari, E., 2022. A systematic and critical review on development of machine learning based-ensemble models for prediction of adsorption process efficiency. J. Clean. Prod. 379, 134588. https://www.sciencedirect.com/science/article/pii/S0959652622041609.

Abdullah, T.A.A., Zahid, M.S.M., Ali, W., Hassan, S.U., 2023. B-LIME: an improvement of LIME for interpretable deep learning classification of cardiac Arrhythmia from ECG signals. Processes 11 (2), 595. https://www.mdpi.com/2227-9717/11/2/595.

Agrawal, T., 2021. Hyperparameter optimization using scikit-learn. In: Agrawal, T. (Ed.), Hyperparameter Optimization in Machine Learning, pp. 31–51. https://doi.org/10.1007/978-1-4842-6579-6_2. Apress.

Akano, T.T., James, C.C., 2022. An assessment of ensemble learning approaches and single-based machine learning algorithms for the characterization of undersaturated oil viscosity. Beni-Suef University Journal of Basic and Applied Sciences 11 (1), 149. https://doi.org/10.1186/s43088-022-00327-8.

Al-Sahlanee, D.T., Allawi, R.H., Al-Mudhafar, W.J., Yao, C., 2023. Ensemble machine learning for data-driven predictive analytics of drilling rate of penetration (ROP) modeling: a case study in a southern Iraqi oil field. Day 4 Thu, May 25, 2023, 2023-May, D021S004R007. https://onepetro.org/SPEWRM/proceedings/23WRM/4-23WRM/D021S004R007/519629.

Al Daoud, E., 2019. Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. International Journal of Computer and Information Engineering 13 (1), 6–10. https://publications.waset.org/10009954/comparison-between-xgboost-lightgbm-and-catboost-using-a-home-credit-dataset.

Alabi, R.O., Elmusrati, M., Leivo, I., Almangush, A., Mäkitie, A.A., 2023. Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP. Sci. Rep. 13 (1), 8984.

Alam, M.S., Al-Ismail, F.S., Hossain, M.S., Rahman, S.M., 2023. Ensemble machine-learning models for accurate prediction of solar irradiation in Bangladesh. Processes 11 (3), 908. https://www.mdpi.com/2227-9717/11/3/908.

Allahvirdizadeh, P., 2020. A review on geothermal wells: well integrity issues. J. Clean. Prod. 275, 124009. https://linkinghub.elsevier.com/retrieve/pii/S0959652620340543.

Alsaba, M.T., Al Dushaishi, M.F., Abbas, A.K., 2020. A comprehensive review of nanoparticles applications in the oil and gas industry. J. Pet. Explor. Prod. Technol. 10 (4), 1389–1399. https://link.springer.com/10.1007/s13202-019-00825-z.
Alum, M.A.O., Egbon, F., 2011. Semi-Analytical Models on the Effect of Drilling Fluid Properties on Rate of Penetration (ROP). All Days, pp. 697–709. In: https://onepetro.org/SPENAIC/proceedings/11NAICE/All-11NAICE/Abuja, Nigeria/150918.
Amin, M.N., Khan, K., Javed, M.F., Aslam, F., Qadir, M.G., Faraz, M.I., 2022. Prediction of mechanical properties of fly-ash/slag-based geopolymer concrete using ensemble and non-ensemble machine-learning techniques. Materials 15 (10), 3478. https://www.mdpi.com/1996-1944/15/10/3478.
Anemangely, M., Ramezanzadeh, A., Tokhmechi, B., Molaghab, A., Mohammadian, A., 2018. Drilling rate prediction from petrophysical logs and mud logging data using an optimized multilayer perceptron neural network. J. Geophys. Eng. 15 (4), 1146–1159. https://academic.oup.com/jge/article/15/4/1146-1159/5204266.
Antwarg, L., Miller, R.M., Shapira, B., Rokach, L., 2021. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. Expert Syst. Appl. 186 (August), 115736 https://doi.org/10.1016/j.eswa.2021.115736.
Arabjamaloei, R., Shadizadeh, S., 2011. Modeling and optimizing rate of penetration using intelligent systems in an Iranian southern oil field (ahwaz oil field). Petrol. Sci. Technol. 29 (16), 1637–1648. http://www.tandfonline.com/doi/abs/10.1080/10916460902882818.
Badrouchi, F., Rasouli, V., Badrouchi, N., 2022. Impact of hole cleaning and drilling performance on the equivalent circulating density. J. Petrol. Sci. Eng. 211, 110150. https://www.sciencedirect.com/science/article/pii/S0920410522000444.
Bani Mustafa, A., Abbas, A.K., Alsaba, M., Alameen, M., 2021. Improving drilling performance through optimizing controllable drilling parameters. J. Pet. Explor. Prod. Technol. 11 (3), 1223–1232. https://doi.org/10.1007/s13202-021-01116-2.
Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012. October 2019.
Bataee, M., Mohseni, S., 2011. Application of Artificial Intelligent Systems in ROP Optimization: A Case Study in Shadegan Oil Field. All Days, vols. 13–22. In: https://onepetro.org/SPEUGM/proceedings/11UGM/All-11UGM/Muscat.
Ben Aoun, M.A., Madarász, T., 2022. Applying machine learning to predict the rate of penetration for geothermal drilling located in the Utah FORGE site. Energies 15 (12), 4288. https://www.mdpi.com/1996-1073/15/12/4288.
Cao, Y., Geddes, T.A., Yang, J.Y.H., Yang, P., 2020. Ensemble deep learning in bioinformatics. Nat. Mach. Intell. 2 (9), 500–508. https://www.nature.com/articles/s42256-020-0217-y.
Chakraborty, D., Awolusi, I., Gutierrez, L., 2021. An explainable machine learning model to predict and elucidate the compressive behavior of high-performance concrete. Results in Engineering 11, 100245. https://doi.org/10.1016/j.rineng.2021.100245.
Chen, C.-H., Tanaka, K., Kotera, M., Funatsu, K., 2020. Comparison and improvement of the predictability and interpretability with ensemble learning models in QSPR applications. J. Cheminf. 12 (1), 19. https://doi.org/10.1186/s13321-020-0417-9.
Cheraghi, M., Hajipour, M., Emamzadeh, A., 2022. Enhancement of the heat capacity of water-based drilling fluids for deep drilling applications. Braz. J. Chem. Eng. 39 (1), 77–86. https://doi.org/10.1007/s43153-021-00201-7.
Diaz, M.B., Kim, K.Y., 2020. Improving rate of penetration prediction by combining data from an adjacent well in a geothermal project. Renew. Energy 155, 1394–1400. https://doi.org/10.1016/j.renene.2020.04.029.
Diaz, M.B., Kim, K.Y., Kang, T.-H., Shin, H.-S., 2018. Drilling data from an enhanced geothermal project and its pre-processing for ROP forecasting improvement. Geothermics 72, 348–357. https://doi.org/10.1016/j.geothermics.2017.12.007. December 2017.
Diaz, M.B., Kim, K.Y., Shin, H.-S., Zhuang, L., 2019. Predicting rate of penetration during drilling of deep geothermal well in Korea using artificial neural networks and real-time data collection. J. Nat. Gas Sci. Eng. 67, 225–232. https://doi.org/10.1016/j.jngse.2019.05.004. October 2018.
Dong, X., Yu, Z., Cao, W., Shi, Y., Ma, Q., 2020. A survey on ensemble learning. Front. Comput. Sci. 14 (2), 241–258. https://doi.org/10.1007/s11704-019-8208-z.
Ekanayake, I.U., Meddage, D.P.P., Rathnayake, U., 2022. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Stud. Constr. Mater. 16, e01059. https://linkinghub.elsevier.com/retrieve/pii/S2214509522001917.
Gan, C., Cao, W.-H., Wu, M., Chen, X., Hu, Y.-L., Liu, K.-Z., Wang, F.-W., Zhang, S.-B., 2019a. Prediction of drilling rate of penetration (ROP) using hybrid support vector regression: a case study on the Shennongjia area, Central China. J. Petrol. Sci. Eng. 181, 106200. November 2018. https://linkinghub.elsevier.com/retrieve/pii/S0920410519306126.
Gan, C., Cao, W., Wu, M., Liu, K.-Z., Chen, X., Hu, Y., Ning, F., 2019b. Two-level intelligent modeling method for the rate of penetration in complex geological drilling process. Appl. Soft Comput. 80, 592–602. https://linkinghub.elsevier.com/retrieve/pii/S156849461930211X.
Gautam, S., Guria, C., Rajak, V.K., 2022. A state of the art review on the performance of high-pressure and high-temperature drilling fluids: towards understanding the structure-property relationship of drilling fluid additives. J. Petrol. Sci. Eng. 213, 110318. https://www.sciencedirect.com/science/article/pii/S0920410522002078.
Gomez, E., Ombe, E., Goodkey, B., Carvalho, R., 2021. Drilling automation: the step forward for improving safety, consistency, and performance in onshore gas drilling. Day 2 Mon, November 29, D021S002R007, 2021, 2021-Novem. https://onepetro.org/SPEMEOS/proceedings/21MEOS/2-21MEOS/D021S002R007/474698.
Gong, Y., Mehana, M., El-Monier, I., Xu, F., Xiong, F., 2019. Machine learning for estimating rock mechanical properties beyond traditional considerations. In: Proceedings of the 7th Unconventional Resources Technology Conference, D013S011R003. https://doi.org/10.15530/urtec-2019-897.
Han, J., Sun, Y., Zhang, S., 2019. March 22). A data driven approach of ROP prediction and drilling performance estimation. In: International Petroleum Technology Conference. http://www.onepetro.org/doi/10.2523/19430-MS.
He, M., Zhang, Z., Zhu, J., Li, N., Li, G., Chen, Y., 2021. Correlation between the rockburst proneness and friction characteristics of rock materials and a new method for rockburst proneness prediction: field demonstration. J. Petrol. Sci. Eng. 205, 108997. https://www.sciencedirect.com/science/article/pii/S0920410521006562.
He, M., Zhou, J., Li, P., Yang, B., Wang, H., Wang, J., 2022. Novel approach to predicting the spatial distribution of the hydraulic conductivity of a rock mass using convolutional neural networks. Q. J. Eng. Geol. Hydrogeol. 56 (1), qjegh2021-q2169 https://doi.org/10.1144/qjegh2021-169.
Hegde, C., Daigle, H., Millwater, H., Gray, K., 2017. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. J. Petrol. Sci. Eng. 159, 295–306. https://doi.org/10.1016/j.petrol.2017.09.020.
Hegde, C., Gray, K., 2018. Evaluation of coupled machine learning models for drilling optimization. J. Nat. Gas Sci. Eng. 56, 397–407. https://linkinghub.elsevier.com/retrieve/pii/S1875510018302543.
Hegde, C., Millwater, H., Pyrcz, M., Daigle, H., Gray, K., 2019. Rate of penetration (ROP) optimization in drilling with vibration control. J. Nat. Gas Sci. Eng. 67, 71–81. https://linkinghub.elsevier.com/retrieve/pii/S187551001930112X.
Ibrahim, A., A., L., R., M., M., O., R., & A., G., 2020. Comparison of the CatBoost classifier with other machine learning methods. Int. J. Adv. Comput. Sci. Appl. 11 (11), 738–748. http://thesai.org/Publications/ViewPaper?Volume=11&Issue=11&Code=IJACSA&SerialNo=90.
Jiang, W., Samuel, R., 2016. Optimization of rate of penetration in a convoluted drilling framework using ant colony optimization. In: Day 3 Thu, March 03, 2016, 2016-Janua. In: https://onepetro.org/SPEDC/proceedings/16DC/3-16DC/Fort Worth. Texas, USA/207828.
Kim, B., Lee, D.-E., Hu, G., Natarajan, Y., Preethaa, S., Rathinakumar, A.P., 2022. Ensemble machine learning-based approach for predicting of FRP–concrete interfacial bonding. Mathematics 10 (2), 231. https://www.mdpi.com/2227-7390/10/2/231.
Kolesnikov, A.A., Kikin, P.M., Portnov, A.M., 2019. Diseases spread prediction in tropical areas by machine learning methods ENSEMBLING and spatial analysis techniques. Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. 221–226. XLII-3/W8(3/W8. https://isprs-archives.copernicus.org/articles/XLII-3-W8/221/2019/.
Leines-Artieda, J.A., Fiallos-Torres, M.X., Alharthi, A., Mahmoud, S. E.din, Al Hashmi, A., Alqaydi, M., Ramsay, T., Gong, Y., Yu, W., Miao, J., Escorcia, A., Useche, F., Al Bannay, A., Fonseca, R.-M., Sepehrnoori, K., 2022. A machine learning-based data augmentation approach for unconventional reservoir characterization using microseismic data and EDFM. Day 2 Tue, November 01, 2022. https://doi.org/10.2118/210989-MS.
Li, C., Cheng, P., Cheng, C., 2023. A comparison of machine learning algorithms for rate of penetration prediction for directional wells. Day 1 Sun, February 19, D011S011R003. https://doi.org/10.2118/213321-MS, 2023.
Li, G., Song, X., Tian, S., Zhu, Z., 2022. Intelligent drilling and completion: a review. Engineering 18, 33–48. https://doi.org/10.1016/j.eng.2022.07.014.
Li, K., Xu, H., Liu, X., 2022. Analysis and visualization of accidents severity based on LightGBM-TPE. Chaos, Solit. Fractals 157, 111987. https://doi.org/10.1016/j.chaos.2022.111987.
Li, Y., Chen, W., 2020. A comparative performance assessment of ensemble learning for credit scoring. Mathematics 8 (10), 1756. https://www.mdpi.com/2227-7390/8/10/1756.
Liao, X., Khandelwal, M., Yang, H., Koopialipoor, M., Murlidhar, B.R., 2020. Effects of a proper feature selection on prediction and optimization of drilling rate using intelligent techniques. Eng. Comput. 36 (2), 499–510. https://doi.org/10.1007/s00366-019-00711-6.
Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2020. Explainable ai: a review of machine learning interpretability methods. Entropy 23 (1), 18. https://www.mdpi.com/1099-4300/23/1/18.
Liu, N., Zhang, D., Gao, H., Hu, Y., Duan, L., 2021. Real-time measurement of drilling fluid rheological properties: a review. Sensors 21 (11), 3592. https://doi.org/10.1007/s43153-021-00201-7.
Lu, Y., Bo, Y., He, W., 2022. An ensemble model for combating label noise. Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining 608–617. https://doi.org/10.1145/3488560.3498376.
Lundberg, S., Lee, S.-I., 2017. A unified approach to interpreting model predictions. ArXiv:1705.07874 ArXiv Preprint. https://dl.acm.org/doi/10.5555/3295222.3295230.
Ma, X., Hou, M., Zhan, J., Liu, Z., 2023. Interpretable predictive modeling of tight gas well productivity with SHAP and LIME techniques. Energies 16 (9), 3653.
Mao, H., Yang, Y., Zhang, H., Zheng, J., Zhong, Y., 2020. Conceptual design and methodology for rheological control of water-based drilling fluids in ultra-high temperature and ultra-high pressure drilling applications. J. Petrol. Sci. Eng. 188, 106884. https://linkinghub.elsevier.com/retrieve/pii/S0920410519312999.
Mehrad, M., Bajolvand, M., Ramezanzadeh, A., Neycharan, J.G., 2020. Developing a new rigorous drilling rate prediction model using a machine learning technique. J. Petrol. Sci. Eng. 192 (February), 107338 https://doi.org/10.1016/j.petrol.2020.107338.
Moazzeni, A.R., Khamehchi, E., 2020. Rain optimization algorithm (ROA): a new metaheuristic method for drilling optimization solutions. J. Petrol. Sci. Eng. 195, 107512. https://linkinghub.elsevier.com/retrieve/pii/S0920410520305830.
Mohamed, A., Salehi, S., Ahmed, R., 2021. Significance and complications of drilling fluid rheology in geothermal drilling: a review. Geothermics 93, 102066. https://www.sciencedirect.com/science/article/pii/S0375650521000262.

Mohammed, A., Kora, R., 2023. A comprehensive review on ensemble deep learning: opportunities and challenges. Journal of King Saud University - Computer and Information Sciences 35 (2), 757–774. https://www.sciencedirect.com/science/article/pii/S1319157823000228.

Muther, T., Syed, F.I., Lancaster, A.T., Salsabila, F.D., Dahaghi, A.K., Negahban, S., 2022. Geothermal 4.0: AI-enabled geothermal reservoir development- current status, potentials, limitations, and ways forward. Geothermics 100 (January), 102348. https://doi.org/10.1016/j.geothermics.2022.102348.

Nordin, N., Zainol, Z., Mohd Noor, M.H., Chan, L.F., 2023. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. Asian Journal of Psychiatry 79, 103316. https://www.sciencedirect.com/science/article/pii/S1876201822003148.

Ogami, C., Tsuji, Y., Seki, H., Kawano, H., To, H., Matsumoto, Y., Hosono, H., 2021. An artificial neural network−pharmacokinetic model and its interpretation using Shapley additive explanations. CPT Pharmacometrics Syst. Pharmacol. 10 (7), 760–768. https://onlinelibrary.wiley.com/doi/10.1002/psp4.12643.

Okoroafor, E.R., Smith, C.M., Ochie, K.I., Nwosu, C.J., Gudmundsdottir, H., , Jabs, Aljubran, M., 2022. Machine learning in subsurface geothermal energy: two decades in review. Geothermics 102 (February), 102401. https://doi.org/10.1016/j.geothermics.2022.102401.

Omer, Z.M., Shareef, H., 2022. Comparison of decision tree based ensemble methods for prediction of photovoltaic maximum current. Energy Convers. Manag. X 16, 100333. https://www.sciencedirect.com/science/article/pii/S2590174522001568.

Oyedere, M., Gray, K., 2020. ROP and TOB optimization using machine learning classification algorithms. J. Nat. Gas Sci. Eng. 77, 103230. https://linkinghub.elsevier.com/retrieve/pii/S1875510020300846.

Palaoro, G., Andrade, D.E.V., Galdino, J.F., Franco, A.T., Alves, E., Waldmann, A., 2022. Influence of pressure on the gel strength and on the solid-like behavior for an inverted emulsion drilling fluid. J. Petrol. Sci. Eng. 219, 111114. https://www.sciencedirect.com/science/article/pii/S0920410522009664.

Podgorney, R., Finnila, A., Simmons, S., McLennan, J., 2021. A reference thermal-hydrologic-mechanical native state model of the Utah FORGE enhanced geothermal site. Energies 14 (16), 4758. https://www.mdpi.com/1996-1073/14/16/4758.

Prusty, S., Patnaik, S., Dash, S.K., 2022. SKCV: stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. Frontiers in Nanotechnology 4, 972421. https://www.frontiersin.org/articles/10.3389/fnano.2022.972421/full.

Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why should I trust you?": explaining the predictions of any classifier. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, pp. 97–101. http://aclweb.org/anthology/N16-3020.

Rosberg, J.-E., Erlström, M., 2021. Evaluation of deep geothermal exploration drillings in the crystalline basement of the Fennoscandian Shield Border Zone in south Sweden. Geoth. Energy 9 (1), 20. https://doi.org/10.1186/s40517-021-00203-1.

Sabah, M., Talebkeikhah, M., Wood, D.A., Khosravanian, R., Anemangely, M., Younesi, A., 2019. A machine learning approach to predict drilling rate using petrophysical and mud logging data. Earth Science Informatics 12 (3), 319–339. http://link.springer.com/10.1007/s12145-019-00381-4.

Saranya, A., Subhashini, R., 2023. A systematic review of Explainable Artificial Intelligence models and applications: recent developments and future trends. Decision Analytics Journal 7, 100230. https://linkinghub.elsevier.com/retrieve/pii/S277266222300070X.

Shahdi, A., Lee, S., Karpatne, A., Nojabaei, B., 2021. Exploratory analysis of machine learning methods in predicting subsurface temperature and geothermal gradient of Northeastern United States. Geoth. Energy 9 (1), 18. https://geothermal-energy-journal.springeropen.com/articles/10.1186/s40517-021-00200-4.

Sharma, T., Jatain, A., Bhaskar, S., Pabreja, K., 2023. Ensemble machine learning paradigms in software defect prediction. Proc. Comput. Sci. 218, 199–209. https://linkinghub.elsevier.com/retrieve/pii/S1877050923000029.

Soares, C., Gray, K., 2019. Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models. J. Petrol. Sci. Eng. 172 (April 2018), 934–959. https://doi.org/10.1016/j.petrol.2018.08.083.

Sobhi, I., Dobbi, A., Hachana, O., 2022. Prediction and analysis of penetration rate in drilling operation using deterministic and metaheuristic optimization methods. J. Pet. Explor. Prod. Technol. 12 (5), 1341–1352. https://doi.org/10.1007/s13202-021-01394-w.

Sygała, A., Bukowska, M., Janoszek, T., 2013. High temperature versus geomechanical parameters of selected rocks – the present state of research. Journal of Sustainable Mining 12 (4), 45–51. https://www.sciencedirect.com/science/article/pii/S2300396015300689.

Utah-EGI, E. and G. I. at the U. of., 2018. In: Moore, J. (Ed.), Utah FORGE Logs and Data from Deep Well 58-32 (MU-ESW1). Energy and Geoscience Institute at the University of Utah. https://doi.org/10.15121/1452735.

Vega García, M., Aznarte, J.L., 2020. Shapley additive explanations for NO2 forecasting. Ecol. Inf. 56, 101039. https://linkinghub.elsevier.com/retrieve/pii/S1574954119303498.

Wang, H., He, M., 2023. Determining method of tensile strength of rock based on friction characteristics in the drilling process. Rock Mech. Rock Eng. 56 (6), 4211–4227. https://doi.org/10.1007/s00603-023-03276-5.

Wang, H., He, M., Zhang, Z., Zhu, J., 2022. Determination of the constant mi in the Hoek-Brown criterion of rock based on drilling parameters. Int. J. Min. Sci. Technol. 32 (4), 747–759. https://doi.org/10.1016/j.petrol.2021.108997.

Wang, H., He, M., Zhao, J., Zhang, Y., Yang, B., 2023. Cutting energy characteristics for brittleness evaluation of rock using digital drilling method. Eng. Geol. 319 (April), 107099 https://doi.org/10.1016/j.enggeo.2023.107099.

Zhang, H., Ni, H., Wang, Z., Liu, S., Liang, H., 2020. Optimization and application study on targeted formation ROP enhancement with impact drilling modes based on clustering characteristics of logging. Energy Rep. 6, 2903–2912. https://linkinghub.elsevier.com/retrieve/pii/S2352484720313731.

Zhang, S., Chen, Y., Zhang, W., Feng, R., 2021. A novel ensemble deep learning model with dynamic error correction and multi-objective ensemble pruning for time series forecasting. Inf. Sci. 544, 427–445. https://www.sciencedirect.com/science/article/pii/S0020025520308197.

Zhipeng, F., Gani, H., 2022. Interpretable models for the potentially harmful content in video games based on game rating predictions. Appl. Artif. Intell. 36 (1) https://doi.org/10.1080/08839514.2021.2008148.

Zolfaghari, B., Mirsadeghi, L., Bibak, K., Kavousi, K., 2023. Cancer prognosis and diagnosis methods based on ensemble learning. ACM Comput. Surv. 55 (12), 1–34. https://dl.acm.org/doi/10.1145/3580218.