A Case Study of Artificial Neural Networks Usage for Rate of Penetration Prediction with Offset Wells Data

Antonio P. A. Ferro[1], Erasmo A. B. Silva[1], Francisco A. V. Binas Jr.[1], Lucas P. Gouveia[1], Aline S. R. Barboza[1]

[1]*Laboratory of Scientific Computing and Visualization, Federal University of Alagoas*
Avenida Lourival Melo Mota, S/N, Tabuleiro dos Martins, 57072-970, Maceió/Alagoas, Brasil
antonio.ferro@lccv.ufal.br, erasmo.bezerra@lccv.ufal.br, francisco.junior@lccv.ufal.br, lucasgouveia@lccv.ufal.br, aline@lccv.ufal.br

Abstract. The oil and gas industry is constantly motivated to implement strategies focusing on cost management and operational optimization to maximize productivity. Rate of Penetration (ROP) serves as a key performance metric, reflecting the speed at which a drilling bit penetrates subsurface formations. Increasing ROP can minimize costs by reducing drilling time. ROP prediction models aid drilling optimization plans since accurate models are required to find ideal controllable parameters to increase ROP using optimization algorithms. Drilling characteristics that significantly influence ROP include bit type, formation properties, and operation parameters, such as Weight on Bit (WOB) and Rotations per Minute (RPM) of the drill bit. The number of significant parameters makes it difficult to develop analytical expressions for predicting ROP. The present work investigates the applicability of Artificial Neural Networks (ANN) for ROP prediction using public data from three wells extracted from the Volve oil field in the North Sea. The adopted strategy is to learn from offset wells, which means using data from two wells to predict the ROP for the third. This approach helps to address a practical scenario where historical data are used to make predictions for a new well in a close region. This study uses MLP (Multilayer Perceptron) networks with WOB, RPM, Torque, Fluid Flow rate, and Delta-T Compressional as features, and the training process is conducted using the GridSearchCV method for hyperparameter tuning, comparing different model architectures to improve results. To assess the final performance of the model, metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are applied to the tested well. The results show that the developed model can capture ROP patterns without requiring any data stratification of the two wells regarding lithology information.

Keywords: Rate of Penetration, Artificial Neural Network, Drilling Data Analysis

1 Introduction

In the oil and gas industry, the pursuit of efficiency and safety in well construction is necessary to overcome the high investments and risks involved in the exploration and extraction of oil. The well drilling process is complex and, according to Elmgerbi et al. [1], the effective "bit on bottom" action accounts for approximately 20% to 30% of the total time and cost of well construction. Therefore, any advancement in improving the efficiency of this process can lead to significant cost reductions for the industry.

The Rate of Penetration (ROP) is one of the main indicators considered in the drilling efficiency analysis. The ROP is defined as the speed at which the drill bit removes rock material and penetrates the formation. Operational parameters such as Weight on Bit (WOB), Rotations per Minute (RPM) of the drill bit, and drilling fluid Flow Rate (FR) are some of the controllable parameters during the operation that most affect the ROP [2] [3]. As mentioned by Hossain and Al-Mejed [2], among the uncontrollable parameters in the operation, the most important corresponds to the properties of the rock formation. The industry is interested in implementing methodologies to optimize controllable parameters, allowing to safely reach a technical limit for increasing ROP. Strategies for ROP modeling often involve using traditional regression techniques and empirical expressions to relate the main drilling parameters to the ROP. Models capable of reasonably predicting the ROP enable the simulation of scenarios and testing different operational parameters aiming drilling optimization. New strategies incorporating machine learning have been applied to this problem, taking into account the increasing amount of data collected from monitoring systems during drilling operations. These data significantly contribute to studies on future wells.

Barbosa et al. [4] and Najjarpour et al. [3] conducted literature reviews on a large number of empirical models and those employing traditional techniques, as well as machine learning models applied to ROP prediction and optimization. Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forests (RF) are some of the main strategies employed in this problem [5]. The use of machine learning strategies with data from offset wells (existing wells in the same oil field) allows learning the complex ROP pattern using historical data, enabling practical application for the planning of future wells.

In this work, a Multilayer Perceptron neural network (MLP) is used to predict the ROP of a well using data from two wells in the same region. This case study addressed the use of techniques such as the appropriate selection of well data, representative predictors, neural network regularization, and hyperparameter tuning to reduce the prediction error of ROP in the test well. The errors obtained are compared with those from other works that employed different techniques to predict the ROP in this same well. The analyzed wells belong to the Volve field, located in the North Sea, and the data were made public by Equinor [6]. It is hoped that the techniques employed in this work can be used to improve ROP prediction using the learning strategy with data from offset wells.

2 Artificial Neural Network

Artificial Neural Networks (ANNs) are a class of machine learning algorithms initially developed with the motivation of replicating the biological functioning of neural networks in the brain. The so-called perceptron, formally described by Rosenblatt [7], defines a mathematical model of an artificial neuron with a well-defined learning process. The network architeture known as MLP consists of coupling perceptrons in one or more hidden layers. This model is a nonlinear estimator that can be applied to learn complex patterns [8].

The goal of training an MLP is to optimize the weights in order to minimize the cost function (typically involving absolute and quadratic metrics). The most commonly used approach for training models is the application of the traditional gradient descent method combined with a backpropagation strategy, which is an efficient method developed to allow the estimation of the components of the gradient vector. The neurons in an MLP are the processing units responsible for receiving multiple input signals, summing them, and computing an output based on a nonlinear transformation. This transformation is achieved through the activation function, which allows generalization for estimating complex patterns in the network by adding nonlinearity to the model. Figure 1 represents the scheme of a two-layer MLP architecture, indicating two possible activation functions used in neural networks.

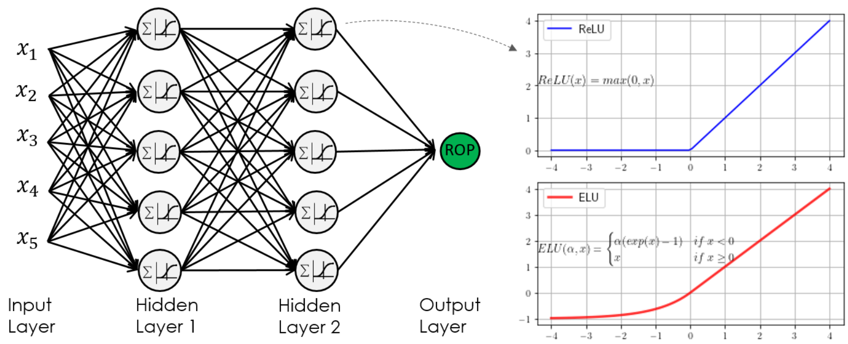

Figure 1. A MLP with two layers and two possible activation functions: ReLU and ELU

As mentioned in Ashena [9], this is the most widely applied network model in the oil industry. The number of neurons, the number of layers, the learning rate, and the activation function are common hyperparameters, meaning they control the model's architecture and are not adjusted during training. The ideal hyperparameters depend on each problem, so the effort to fine-tune hyperparameters is essential. As noted by Najjarpour et al. [3] and Barbosa et al. [4], the most commonly used machine learning model for the ROP prediction problem is the MLP with one layer and using backpropagation. Furthermore, there is a trend to simplify predictive models for ROP by using fewer predictors, with the most common number being between 5 and 8. As also highlighted by Najjarpour et al. [3], some studies mention that including petrophysical properties that characterize the formation, along with operational predictors, improves model performance.

3 Methods and Methodology

This section provides details on the data employed in this study, the selected predictors, the preprocessing methods applied, the training process, and the details of using MLP networks for ROP prediction.

3.1 The Dataset

The public data released by Equinor [6] contain data from various wells in the Volve oil field. Given the availability of operational data from the so-called geophysical logs, which include data on the properties of the drilled rock formation, the highlighted wells in red, shown in Fig. 2, are investigated in this work.

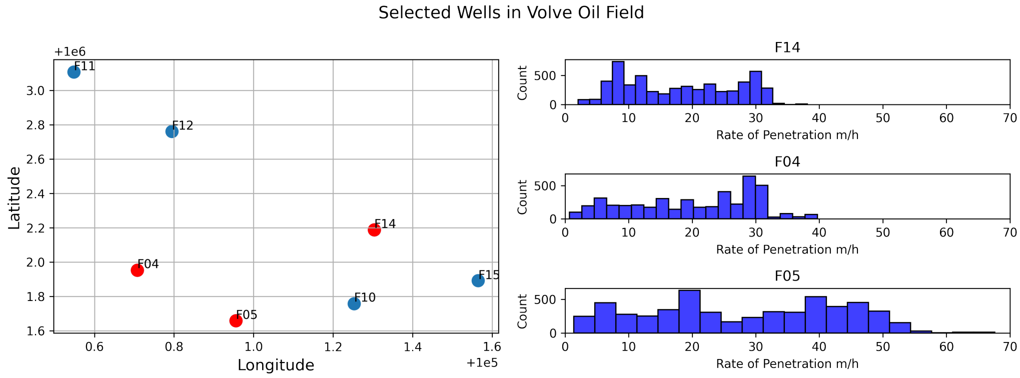

Figure 2. Location of the three analyzed wells in the Volve oil field and ROP histogram

The F05 well is the target. In Tunkiel et al. [10] and also in Ferro et al. [11], results are presented for ROP prediction in this well, in both cases employing ensemble methods and different data sets and features. As shown in Fig. 2, wells F14 and F04 are drilled in a region close to well F05, and the data from these wells are used for training the MLP. These wells were chosen due to their proximity to well F05 and the distribution of the ROP, as they have similar amplitudes and means, which are higher than those of the other wells shown in Fig. 2. Additionally, these data correspond to drilling performed with the same type of drill bit, the Polycrystalline Diamond Compact (PDC). The dataset for well F14 contains 5457 points, for well F04, it contains 4474 points, and 5459 points for the target well F05. Although these three wells are close to each other, they have distinct lithologies profiles along depth.

3.2 Feature Selection and Preprocessing

The selection of predictors used as input data in the neural network should consider domain knowledge about their influence on the resulting ROP. ROP follows complex relationships and nonlinear interconnections with other parameters, making it typically challenging to select all relevant parameters in a predictive model. Najjarpour et al. [3] mentions WOB (Weight on Bit), RPM, Depth, Flow Rate, Mud Weight, Torque on the Bit, Bit Wear, Drilling Fluid Viscosity, and UCS (Unconfined Compressive Strength) as relevant drilling variables.

In this work, five predictors were chosen: Weight on Bit (WOB), RPM, Flow Rate (FR), Torque, and Delta-T Compressional (DTC). The first four are the main operational parameters, and DTC measurement is related to the transit time of waves transmitted to the formation by instruments in the drill string. It was chosen because it is used to estimate important formation properties such as UCS (Unconfined Compressive Strength) and pore pressure through empirical correlations. Figure 3 presents the correlation matrix between the predictors used in the training set (data from wells F14 and F04).

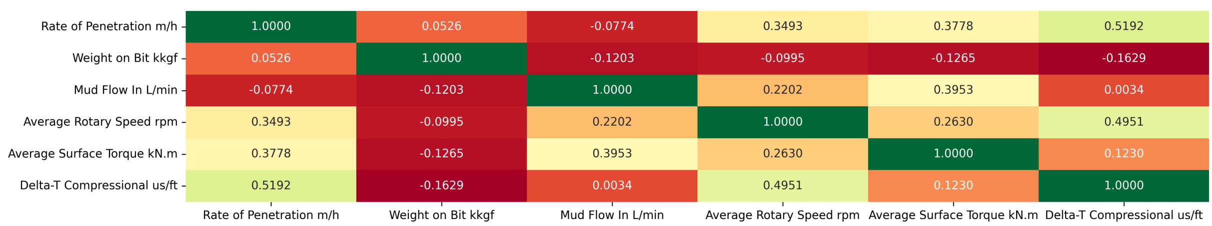

Figure 3. Heatmap showing the correlations between the drilling variables

As observed, there are no predictors with a highly linear correlation with each other (Pearson coefficient between 0.7 and 1.0), which could indicate the addition of variables that do not contribute to the model due to multicollinearity. Additionally, the relationship between the predictors and the target variable shows low positive and negative correlations, which is expected since ROP has a nonlinear relationship with these variables and exhibits both positive and negative interactions. Higher values of DTC indicate less dense formations, and the moderate positive correlation with ROP suggests that in these cases, it is associated with an increase in ROP.

As noted by Géron [12], many machine learning algorithms do not perform well when numerical attributes have very different scales, as is the case with ANNs. The convergence of the gradient descent training process, for example, is affected by the difference in scale of the predictors. Therefore, it is necessary to perform preprocessing to equalize the scale of the predictors. The Eq. 1 is the transformation employed.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Equation 1 is known as Min-Max scaling, which is a type of normalization used to ensure that the predictors have the same order of magnitude. Predictors on a scale from 0 to 1 is appropriate, since all predictors are positive values. More details about the data extraction and treatment appears in Ferro et al. [11] e Tunkiel et al. [10].

3.3 Model Design and Training Process

In this work, Keras, a high-level API for training and executing neural networks, was used. MLP networks with the same number of neurons in each layer and the same activation function for all neurons were chosen. As indicated in [12], this configuration performs very well in most cases and reduces the hyperparameters to be adjusted in a MLP. Regarding the architecture of the networks, hyperparameter tuning was employed to compare different combinations. The cost function in the models is defined as the Mean Squared Error (MSE), common in regression problems, which showed good results compared to other common metrics. The minimization of the cost function in the training process is performed using Stochastic Gradient Descent (SGD) using backpropagation with the technique known as Nesterov Momentum, where the weight adjustment is made iteratively following the opposite direction of the gradient, calculating a step ahead in this direction and employing a numerical value known as momentum, which adds inertia to the iterations using a momentum parameter, chosen as 0.9. These strategies help reduce oscillations and avoid local minima.

The adopted training process corresponds to the use of Cross Validation (CV) on the dataset from the two wells (F14 and F04) using 5-Fold CV. The models are trained with data from 4 folds and evaluated on the partition not used in training (7944 training instances and 1987 validation instances), with the validation partition replaced in each iteration. This division is essential to mitigate the overfitting phenomenon, where the model captures noise and trends present in the training set and performs poorly on data not used in training. The evaluation result of the model is conducted with the average of the folds.

3.4 GridSearchCV and Hyperparameter Tuning

One of the simplest strategies for finding optimal hyperparameters is the GridSearchCV. This method involves a brute-force comparison of models with different hyperparameters within the cross-validation framework. Once the model is selected, the generalization error can be determined on the test set. Table 1 presents the compared hyperparameters.

Table 1. Search options in GridSearchCV

Hiperparameters	Analyzed Values
Number of Layers	1, 2
Number of Neurons	5, 10, 15, 20
Activation Functions	ReLU, ELU
Learning Rate	10^{-3}, 10^{-4}

The ReLU activation function is one of the most commonly used in deep networks, and ELU, as indicated in Géron [12], is a strong alternative to ReLU. The initialization of neural network weights was not considered a hyperparameter in this work, as it is often recommended for ReLU and its derivatives, like ELU, the He initializers, such as the uniform one [12]. Similarly, the number of epochs, one of the most significant hyperparameters, was not treated as a hyperparameter because the early stopping technique was applied. In early stopping, a portion of the data is set aside during training as a validation set, and the validation data is periodically used to monitor the prediction error of the trained network at each epoch. If the error begins to increase, training is halted. In this work, 200 epochs were adopted, with a value of 10 for the number of epochs without improvement to stop the training. A separate validation set is used to monitor this training, with 85% of the data used for training and 15% for validation (considering the cross-validation split, this results in 6752 effective instances for network training and 1192 for validation in each training partition). The L2 regularization was also applied to reduce the occurrence of overfitting, with $\alpha = 0.01$ (based on trial and error experiment). The batch size used in training was 200.

3.5 Model Performance Evaluation

The correct evaluation of the model's performance should be done using data that were not used in the training of the models. This evaluation helps provide an idea of the generalization error. Equation 2 and 3 are the some of the main metrics for assessing the model's performance in regression problems, which share the same units as the response variable (ROP).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - h_i| \tag{2}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - h_i)^2} \tag{3}$$

Where h_i indicates the predictions, and y_i indicates the measured value (ROP), and N indicates the number of data in the test set. The Mean Absolute Error (MAE) indicates the average absolute error of the model for new data. This metric shares the same unit as the ROP and is less sensitive to the presence of outliers in the data. The Root Mean Square Error (RMSE) is derived from the Mean Squared Error (MSE) and, by initially considering the squared errors, is more influenced by outliers. The RMSE also shares the same unit as the MAE and is generally larger. The differences between the MAE and RMSE increases in the presence of extreme values [5] [12].

4 Results and Discussions

Figure 4 illustrates the result of the GridSearchCV with the training data. The network with the best performance on the validation set (obtained with MSE) is highlighted in green, corresponding to a network with two layers, 15 neurons, and a ReLU activation function with a learning rate of 10^{-3}. It is possible to identify significant performance differences related to the various architectures, and even subtle changes in hyperparameters can result in this differences. Thus, this observation emphasize the importance of including a thorough hyperparameter tuning in ANNs models in order to improve prediction performance.

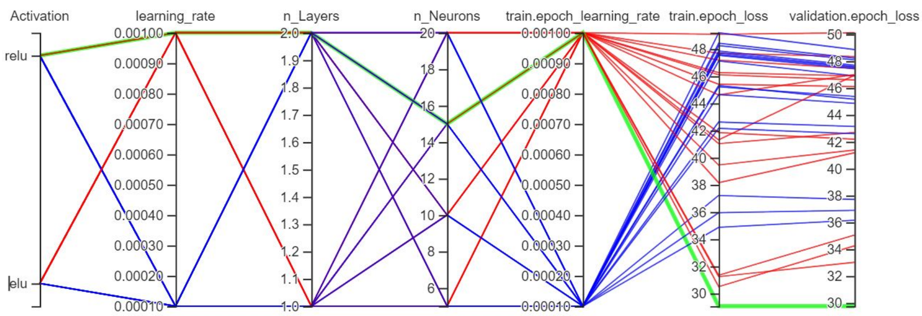

Figure 4. Architectures compared in GridSearchCV with the best architecture highlighted

Figure 5 shows the model's prediction for the test data. This allows for the evaluation of the model's results in comparison to the measured data along the depth of the well F05. It can be observed that the model captures the pattern of the ROP, indicating that the trained model successfully learned the relationship between the ROP and the chosen predictors using data from two offset wells, without making any additional data selection considering information related to lithology, for example. It is also noted that the model underestimates the higher ROP values present in this well.

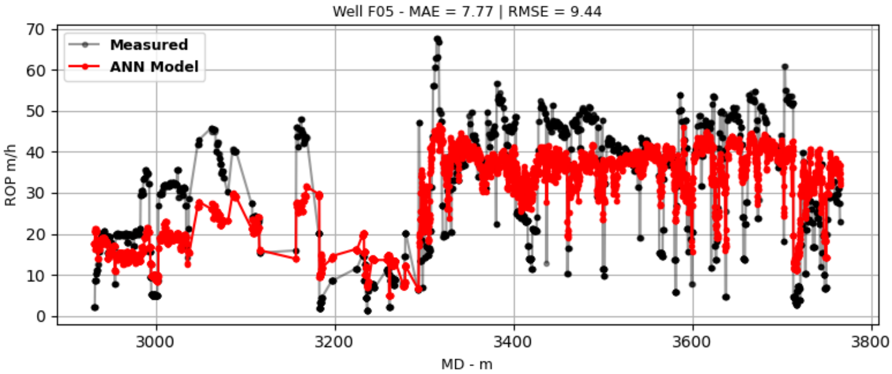

Figure 5. ROP prediction of the MLP model versus measured ROP along depth

As observed in Fig. 2, the wells used for training have a lower range, with few instances above 40 m/h, which may make it difficult to estimate higher ROP values in the F05 well. Including training data with examples of higher ROP values could help improve the results. It is worth noting that high values may be associated with specific behaviors in this well and also undesirable events, such as excessive vibration. However, capturing the ROP pattern along the depth enables the model to assist in scenario simulation and facilitate the optimization of operational parameters, allowing for better drilling control. Figure 6 shows a comparison between the predicted and actual ROP in each of the wells.

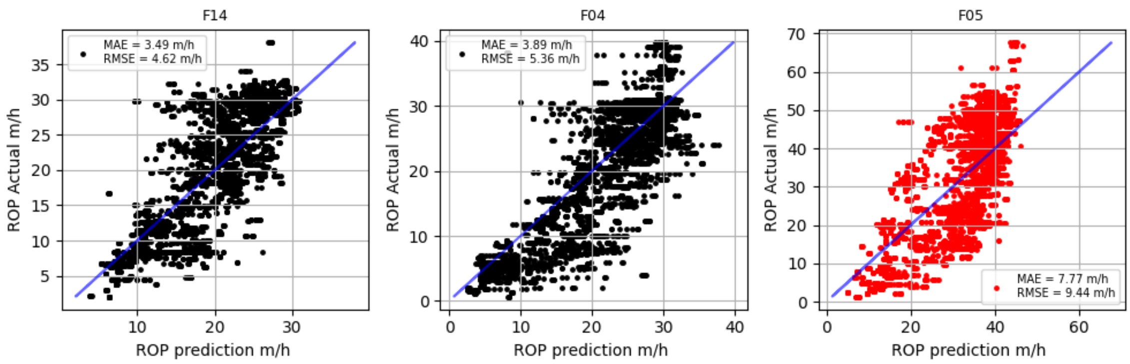

Figure 6. Comparison of measured ROP and predicted ROP for the three wells

The highlighted line at $45°$ in Fig. 6 is a reference to indicate perfect correlation. It can be seen that for wells F04 and F14, the data points aligns more closely with the slope of the line, with less tendency for the actual ROP values to be higher than those predicted by the model. The MAE obtained for wells F04 and F14 are 3.49 m/h and 3.89 m/h, respectively, while the result obtained in Fig. 5 corresponds to 7.77 m/h. The results demonstrate good model adherence for all three wells. Table 2 compares the results of different error metrics found in the literature regarding the prediction of the F05 well, with the types of models used being referenced, where Bourgoyne & Young is a traditional model that employs empirical relationships in a multivariate linear regression.

The results in Tab. 2 demonstrate that the model developed in this study performs better on the test set than models from the literature, using the same data from the F05 well as the test set. The best performance among other studies was an MAE of 10.90 m/h, achieved by Tunkiel et al. [10], who used 11 features and data from 6 wells. These findings indicate that by carefully selecting data, employing proper modeling practices, and choosing representative features, it is possible to enhance the accuracy of ROP prediction.

Table 2. Comparison of results related to ROP prediction in the F05 well from literature

Models	MAE (m/h)	RMSE (m/h)
Bourgoyne & Young (Ferro et al. [11])	18.93	22.83
Random Forest (Ferro et al. [11])	12.45	14.73
Random Forest (Tunkiel et al. [10])	19.70	-
AdaBoost (Tunkiel et al. [10])	10.90	-
Results from this work	7.77	9.44

5 Conclusions

This work presents a study using MLP for ROP prediction with data from offset wells. This type of approach enables practical applications in cases where historical drilling data from wells in an oil field are used to make predictions for a new well in that field. With an accurate model, it is possible to simulate scenarios and develop strategies for optimizing operational parameters. Thus, presenting strategies to improve the performance of predictive models in this problem is necessary, and the present study shows results with improvements in ROP prediction in a case investigated in other works in the literature. Employing hyperparameter tuning, proper selection of predictors and wells for training the neural networks are relevant approaches. Improvements in performance can be made by further exploring hyperparameter tuning and selecting more representative data from offset wells.

Acknowledgements. The authors would like to thank Petrobras for providing technical and financial support to the researcher for the development of this academic research.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] A. M. Elmgerbi, C. P. Ettinger, P. M. Tekum, and G. Thonhauser. Application of machine learning techniques for real time rate of penetration optimization. In *SPE/IADC Middle East Drilling Technology*, volume 17, pp. 25–27, 2021.

[2] M. E. Hossain and A. A. Al-Mejed. *Fundamentals of Sustainable Drilling Engineering*. Scrivener Publishing, 2015.

[3] M. Najjarpour, H. Jalalifar, and S. Norouzi-Apourvari. Half a century experience in rate of penetration management: Application of machine learning methods and optimization algorithms - a review. *Journal of Petroleum Science and Engineering*, vol. 208, pp. 109575, 2022.

[4] L. F. F. M. Barbosa, A. Nascimento, M. H. Mathias, and de J. A. Carvalho. Machine learning methods applied to drilling rate of penetration prediction and optimization - a review. *Journal of Petroleum Science and Engineering*, vol. 183, 2019.

[5] C. Soares and K. Gray. Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models. *Journal of Petroleum Science and Engineering*, vol. 172, pp. 934–959, 2019.

[6] Equinor. Disclosing all volve data. Disponível em: `https://www.equinor.com/news/archive/14jun2018-disclosing-volve-data`. Acesso em: 09 de Jul. 2022, 2018.

[7] F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, vol. 65, 1958.

[8] R. Izbicki and T. M. Santos. *Aprendizado de máquina: uma abordagem estatística*. UICLAP, 2020.

[9] R. Ashena. Optimization of drilling parameters using an innovative ga-ps artificial intelligence model. In *SPE Asia Pacific Oil & Gas Conference and Exhibition*, 2020.

[10] A. T. Tunkiel, D. Sui, and T. Wiktorski. Reference dataset for rate of penetration benchmarking. *Journal of Petroleum Science and Engineering*, vol. 196, 2021.

[11] A. P. A. Ferro, A. S. R. Barboza, and L. P. Gouveia. Performance analysis of random forest and bourgoyne and young models in rate of penetration prediction. In *ENAHPE*, Matinhos, PR, 2023.

[12] A. Géron. *Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras & TensorFlow*. Alta Books, 2 edition, 2021.