Journal of Mining and Environment (JME) Published online

Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Effects of Input Log Selection on Permeability Accuracy in Heterogeneous Sirt Basin Reservoirs, Libya

Mohammed A. Amir[1]*, Hamzah S. Amir[2], and Mokhtar Farkash[3]

1 Department Earth Science, Faculty of Science, University of Benghazi, Benghazi, Libya.
2 Department Petroleum Engineering, Faculty of Engineering, Bright Star University, Elbrega, Libya.
3 Arabian Gulf Oil Company (AGOCO), Benghazi, Libya.

Article Info

Received 8 May 2025

Received in Revised form 24 September 2025

Accepted 19 November 2025

Published online 19 November 2025

DOI: 10.22044/jme.2025.16192.3128

Keywords

Machine learning

Reservoir characterization

Permeability

Well logging

Heterogeneity

Abstract

Permeability estimation is an essential phase in assessing the hydrocarbon potential within porous media and designing reservoir management methods. Recently, machine learning (ML) methodologies have gained prominence in the prediction of permeability. The initial stage in constructing highly reliable ML models is to identify the optimum combinations of input logs, as permeability is a highly sensitive parameter; this step is essential and can influence model accuracy. While feature engineering methods provide valuable insights in selecting suitable input logs, the effectiveness of these logs or their combinations remains underexplored, particularly in the context of high-heterogeneity reservoirs. The current study intends to save time by evaluating the effectiveness of twelve distinct models, each constructed using a Multi-Layer Perceptron (MLP), based on various combinations of input logs using data from the Nubian reservoir, Sirt Basin, Libya. The methodology involved several steps, including preprocessing, splitting, optimization, and validation. The findings demonstrate that single-input logs, mainly the Gamma-ray (GR), bulk density (RHOB), and sonic logs (DT), exhibited higher correlation coefficients compared to the multiple log combinations. The GR model attained the best R^2 of 0.994, indicating its sensitivity in capturing non-linear relationships. On the other hand, multi-log models achieved variable accuracy, resulting in increased learning complexity. The study highlights the efficiency of selecting the optimal combination of input logs, providing practical guidance for ML-based permeability prediction in heterogeneous reservoirs.

1. Introduction

Permeability is a crucial physical property of porous media that measures the ability to transmit fluid [1]. Furthermore, it plays a vital role in evaluating reservoirs, developing oil and gas fields, establishing geological models, optimizing reservoir characterization, and improving hydrocarbon recovery[2, 3 , 4, 5, 6]. Nevertheless, the conventional approaches commonly employed for determining absolute permeability require a substantial investment of time, resources, and labor. These demands can often exceed the capabilities of developing nations [7, 8 ,9] .Several empirical equations have been presented by researchers to forecast permeability in numerous basins because of the absence of core data. These equations are based on traditional equations such as [10, 11, 12]. Nevertheless, these correlation formulae do not have general applicability while is important to assess their suitability by comparing them to field studies. Providing precise findings using empirical equations is challenging due to the sensitivity of permeability to effects by various factors such as the dynamic viscosity of the fluid, applied pressure difference, and rock reservoir properties like grain size, sorting, and pore throats [7,13 ,14] . Furthermore, these approaches exhibit distinct disparities in terms of data prerequisites and parameter selection and face an issue of poor precision in the mathematical regression technique. Another method to estimate the permeability is to

Corresponding author: Itzamir924@gmail.com (M A. Amir)

find a correlation between permeability and different logging curves, which is a challenge due to the sensitivity of permeability to be affected by several factors, such as porosity type, cement, and diagenesis processes[2,15,16,17]. In recent research, other scholars have used Digital Rock Technology (DRP) alongside advanced imaging methods like X-ray computed tomography and focused on beam–scanning electron microscopes (FIB-SEM). The emergence of digital rock technology enables the acquisition of the three-dimensional architecture of porous rock formations [1]. While this technique yields highly accurate outcomes, it also presents several limitations, such as the need for high-resolution imaging and computationally intensive modeling of fluid flow within the rock matrix. The complexity of the computing infrastructure and the duration of simulation time may render these simulations unfeasible for several applications [18] . Furthermore, the challenge of collecting samples for DRP is that these samples may not be available in all intervals of reservoirs [19]. Additionally, predicting permeability in heterogeneous reservoirs requires small-scale samples, which are not always present in all reservoir intervals [20]. Recently, Machine learning tools (ML) have become popular for solving various geological tasks, such as predicting bulk density [21], Shear wave velocity [22 , 23] , and Rock Fracture Toughness [24]. Various machine learning algorithms have been used with different input logs to predict the permeability in several reservoirs, providing valuable insights into these approaches. Artificial Neural Networks (ANN) have been employed in many studies to forecast the permeability in missed intervals, for example, [25] who successfully built a high-efficiency model with a low root-mean-square prediction error (RMSE) by utilizing gamma ray, density, and neutron porosity as input logs to predict permeability in carbonate reservoirs. The study was also able to manage and capture nonlinearity between input logs and permeability.

Moreover, sonic, density, and neutron porosity logs were utilized as input logs to build a support vector machine (SVM) model to predict the permeability in heterogeneity reservoirs according to research conducted by [26]. The study generated highly precise and accurate findings. Further study by [2] used Extreme Gradient Boosting (XG Boost) to estimate permeability in tight sandstone reservoirs based on six logging curves—the gamma-ray (GR) curve, the acoustic curve (AC), the spontaneous potential (SP) curve, the caliper (CAL) curve, the deep lateral resistivity (RILD) curve, and the eight lateral resistivity (RFOC) curve—showed that the XGBoost model had better outcomes in permeability prediction along with the lowest root mean square error, proving the efficacy and viability of this approach. The researchers [27] developed an adaptive neuro-fuzzy inference system (ANFIS) to forecast permeability in heterogeneous carbonate reservoirs. They utilized multiple input logs, including gamma ray, neutron porosity, bulk density, resistivity, sonic, spontaneous potential, hole size, depths, and other logs. The study demonstrated a strong correlation coefficient of 0.95 for both training and testing, indicating a high level of accuracy. Additionally, the root-mean-square error was found to be low at 0.28. Various scholars have used Gaussian Process Regression (GPR), fuzzy logic, Bayesian model averaging (BMA), and least absolute shrinkage and selection operator regression (LASSO) methods to forecast permeability utilizing various input logs. [28 , 29 , 30 , 31].Prior studies have utilized various input logs, but the effectiveness of these logs or their combinations has not been thoroughly investigated, particularly in the context of high-heterogeneity reservoirs. Although traditional methods, such as feature engineering methods, provide valuable insights but it is having some constraints. In such reservoirs, the connections between the input logs and permeability are frequently nonlinear and intricate. Gaining insight into the logs that promote forecast accuracy may strengthen the prediction modeling, providing further information on permeability parameters and saving time. Thus, the current study intends to fill the gap by evaluating the effectiveness of twelve distinct models, each constructed using Multi-Layer Perceptron (MLP), based on various combinations of input logs using data from the Nubian reservoir, Sirt Basin, Libya. Furthermore, sensitivity analysis methods (SHAP and permutation importance), and Partial dependence plot (PDP) were employed to provide deeper interpretability of model performance.

2. Methods.

This study aims to assess and detect the optimal input logs combination to improve the accuracy of permeability predictions in intervals where the core data are absent in the Sarrir field, located in the Sirt Basin of Libya. The Sirt Basin, situated in the north-central region of Libya, is a relatively new sedimentary basin within the African Craton [32 , 33] .It has an inshore area of over 375,000 square kilometers and has been estimated to have a sedimentary volume of 1.3 million cubic kilometers [34]. The Sirt basin is known for its numerous potential fields, with the Sarir oil field being recognized as the most significant oil resource in the region. This study focused on collecting data from four wells located within the primary reservoir of the Sarir oil field, which is the Nubian sandstone. The Nubian sandstone is a formation composed of interbedded sandstone and shales, and it produces oil from the Late Jurassic-Early Cretaceous period [35]. **Figure.1** illustrates the main lithostratigraphy column of the Sirt Basin. The dataset comprises over 3200 data points of core permeability, as well as wireline logs such as Neutron Porosity (NPHI), Density (RHOB), Induction resistivity (ILD), Sonic log (DT), and Gamma-ray (GR). All data, including both wireline logs and core data, were processed and corrected to enhance the database's quality and yield valuable insights. **Figures 2 and 3** provide several log views and histograms. All statistical analyses and descriptions are reported in **Table.1**. The permeability readings in the Nubian reservoir exhibit significant and extensive fluctuation with depth, mainly due to an increase in heterogeneity and variety in the pores' throats. This results in it more difficult to establish a linear relationship between the input and output variables. **Figure.4** is a graph showing the relationship between logs and depth, highlighting a significant variance in permeability as depth increases. The methodology workflow for permeability prediction utilizing multiple input log combinations is shown in **Figure. 5** and consists of the following steps:

Table 1. Statistical analysis and details of petrophysical well log data and permeability, including central tendencies (Mean, Median), dispersion (Standard Deviation, Variance), and distribution characteristics (Skewness, Kurtosis).

	Mean	Median	Standard Deviation	Variance	Minimum	25th Percentile	50th Percentile (Median)	75th Percentile	Skewness	Kurtosis
DEPTH	9023.75	9023.75	468.01	219037.62	8213.50	8618.62	9023.75	9428.88	0.00	-1.20
DT	81.67	75.75	16.45	270.59	48.06	70.98	75.75	86.01	1.27	0.63
GR	40.97	37.07	20.95	438.82	11.68	23.16	37.07	56.60	0.73	0.35
K	117.39	2.42	298.15	88891.65	0.024	0.20	2.42	40.25	3.42	12.45
NPHI	0.22	0.17	0.12	0.02	0.01	0.13	0.17	0.29	1.03	-0.12
RHOB	2.35	2.38	0.16	0.03	1.70	2.25	2.38	2.47	-0.63	0.05
ILD	1.36	0.84	2.54	6.44	0.15	0.51	0.84	1.33	7.72	71.36

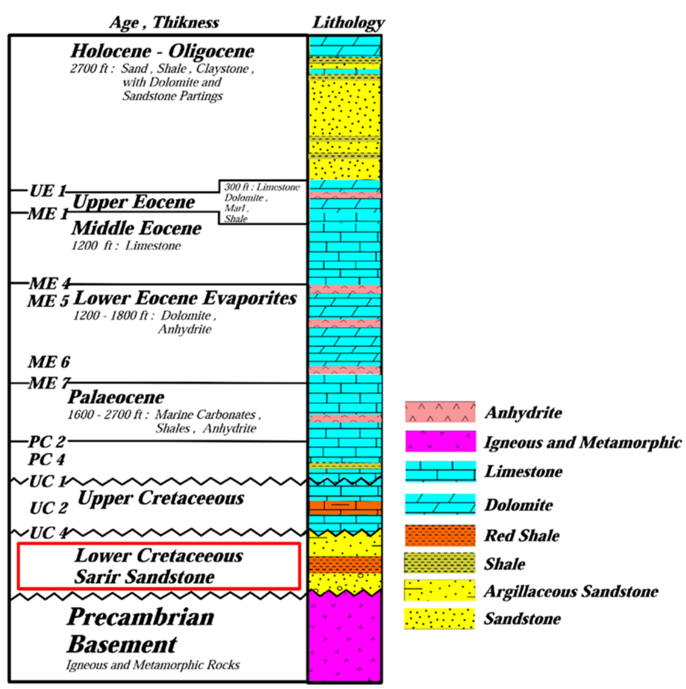


Figure 1. General lithostratigraphy of the Sirt basin, the red box illustrates the target reservoir. Sarir Sandstone consists of sandstone interbedded with shales and deposited in the lower Cretaceous.

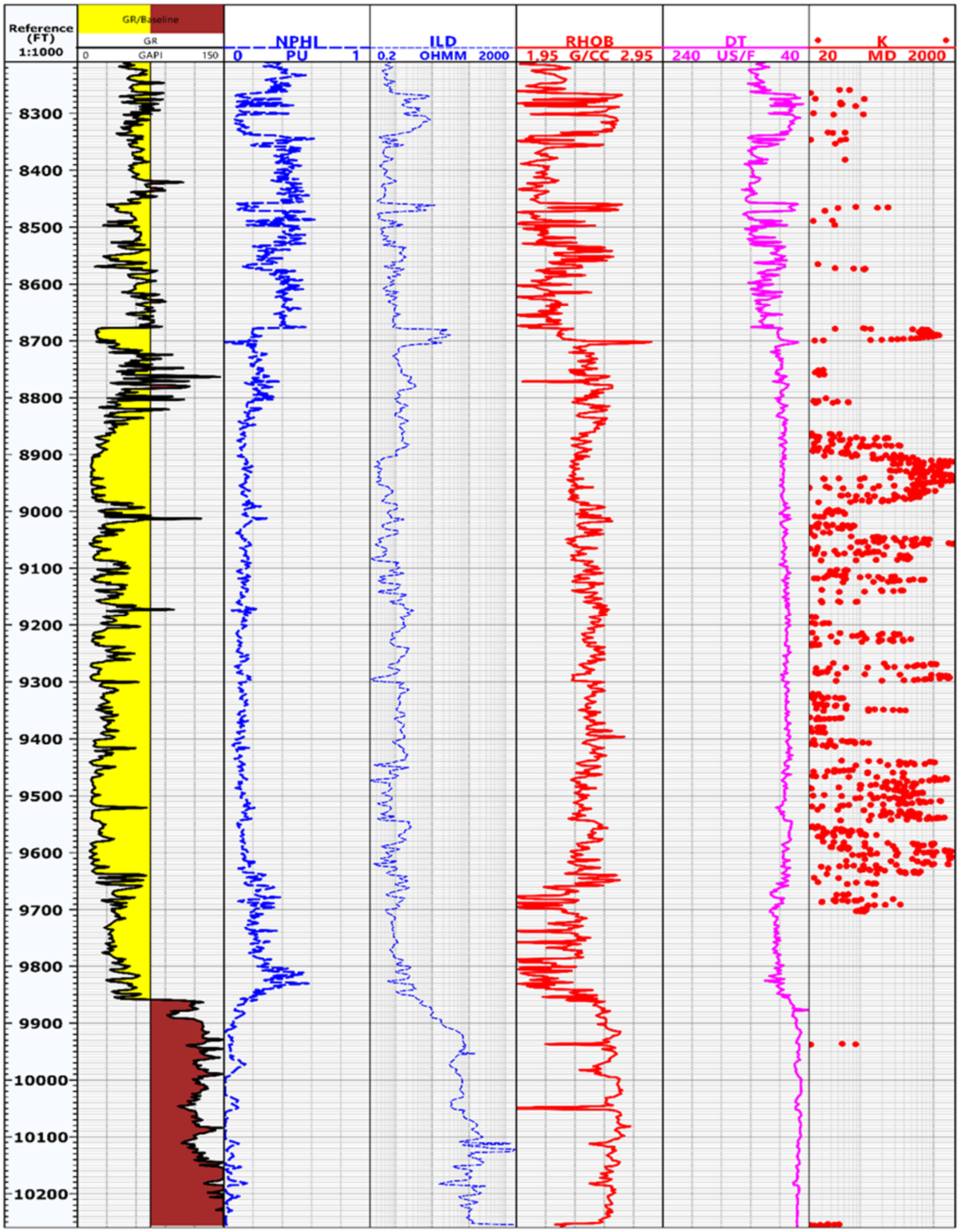


Figure 2. Wireline log profiles and core permeability data for one of the target wells. From left to right, the tracks display DEPTH (reference), Gamma Ray (GR), Neutron Porosity (NPHI), Induction Resistivity (ILD), Bulk Density (RHOB), Sonic (DT), and Core Permeability (K) data.

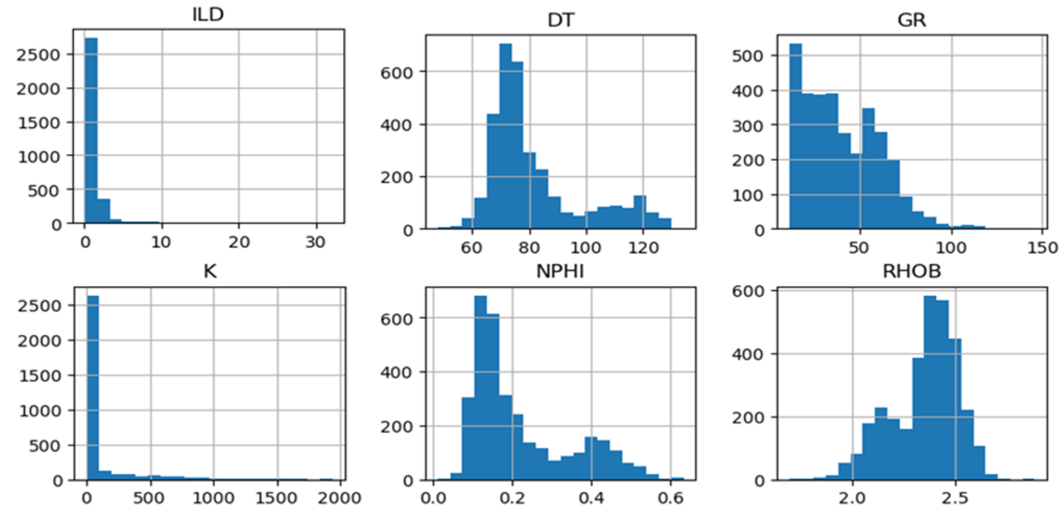


Figure 3. Histograms displaying the statistical distributions of the petrophysical variables: Gamma Ray (GR), Neutron Porosity (NPHI), Bulk Density (RHOB), Sonic (DT), induction resistivity (ILD) and Core Permeability (K).

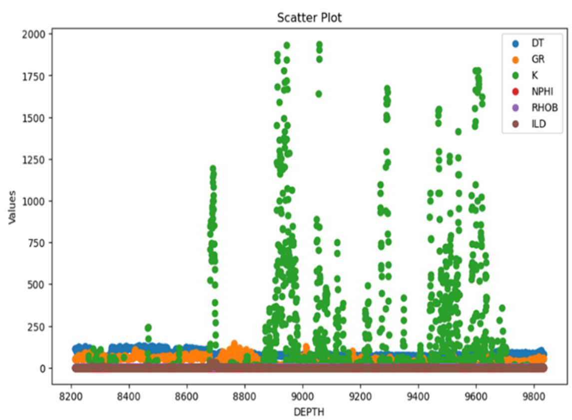


Figure 4. A graph displaying the variation of the dataset with respect to depth. The scatter plot highlights a significant variance in permeability as depth increases.

2.1 Preprocessing

To ensure the model's success, the dataset was preprocessed. The preprocessing step in data science is essential to accomplish accurate data analytics. Considering the inherent complexity of building operations and the deficiencies in data quality, it is crucial to include operational data analysis as an essential stage [36 , 37].Data preprocessing encompasses several procedures aimed at improving the quality of raw data, including the elimination of outliers and the imputation of missing values [38 , 39].The initial step in data preprocessing involves addressing missing data for permeability. To address this issue, the missing permeability data points were imputed using the K-nearest neighbor (KNN) method. Prior studies have demonstrated that the KNN method can yield satisfactory results, even when dealing with a relatively high proportion of missing data (i.e., 5-15%) [40]. Thereafter, the process of eliminating outliers was carried out, since these outliers might have a substantial impact on the outcome. Therefore, it is essential to address or eliminate them before the analysis to prevent the distortion of results [38 , 41]. A box plot was utilized to display the distribution of data and detect any observed outliers. According to **Figure.6**, the box plot reveals significant outliers in the permeability data. These outliers were removed using Tukey's approach, which involves identifying the first and third quartiles and calculating the interquartile range (IQR) using equation (1) [41]. The Lower Inner Fence (LIF) and Upper Inner Fence (UIF) are computed using equations (2 and 3). Accordingly, Tukey's technique identifies outliers as values that fall outside of the upper and lower fences and recommends their elimination [41, 42 ,43].

$$IQR = 1.5 \times (Q_3 - Q_1) \quad (1)$$

$$LIF = Q_1 - IQR \quad (2)$$

$$UIF = Q_3 + IQR \quad (3)$$

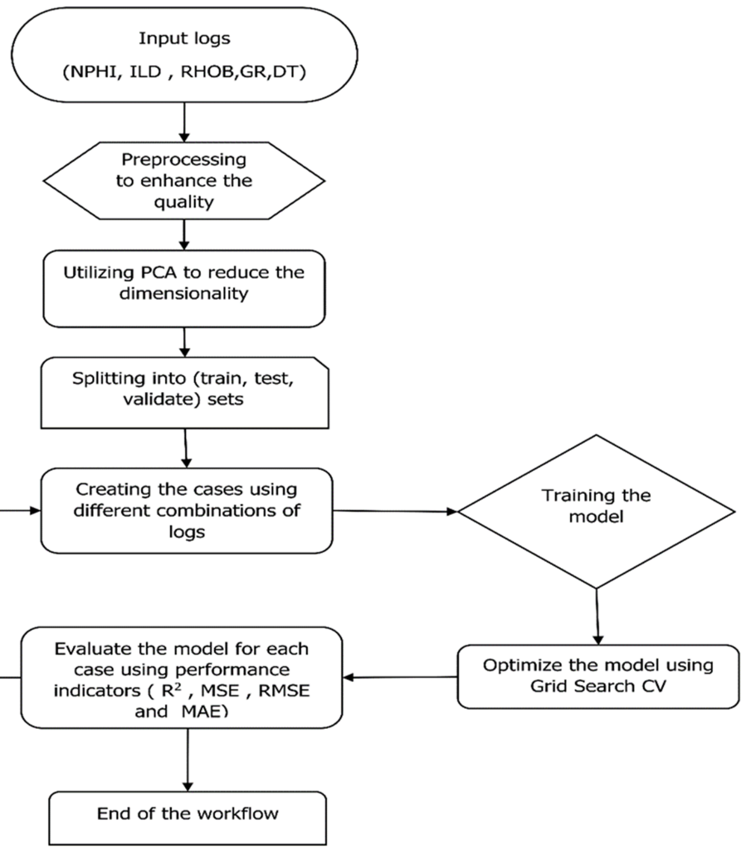


Figure 5. Details the workflow of prediction processes using different cases, displaying each step to achieve the goal of the current study.

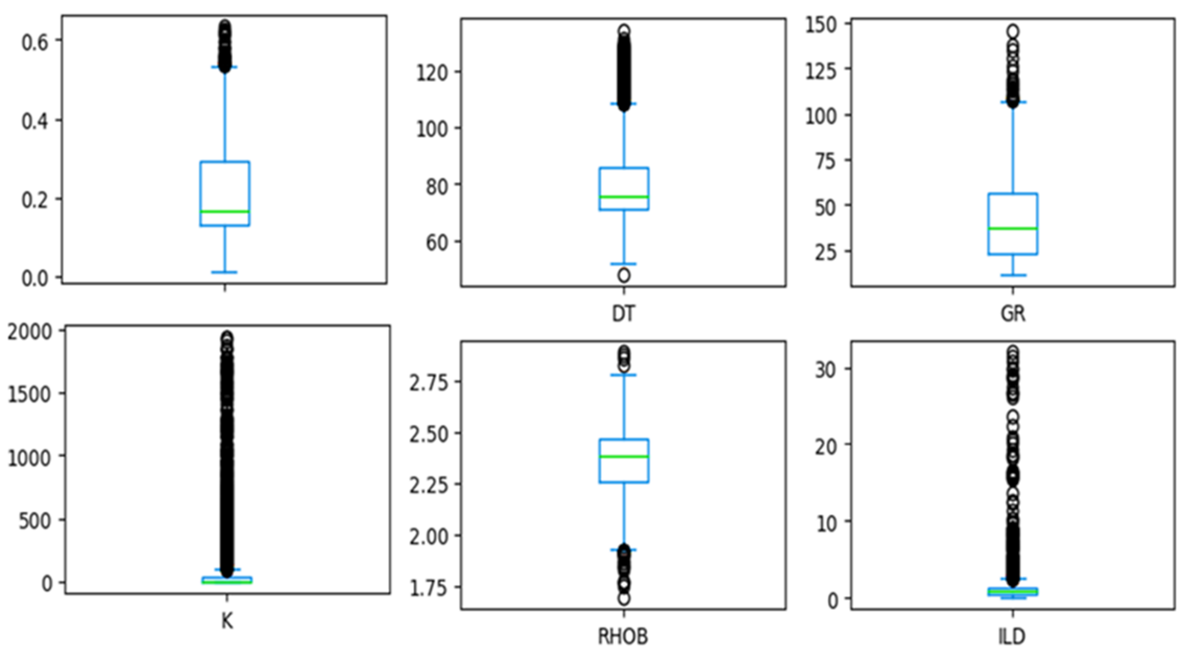


Figure 6. Box plots illustrating the distributions and outliers for each variable.

To expedite the training process and minimize the influence on network accuracy, it is essential to normalize both the logging data and core data, since there exist varying degrees of differences between them [44]. The present study employed the Min-Max Scaler approach to normalize the datasets. Min-Max Scaler can be described as follows:[44].

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4}$$

Consider X to be an input parameter, X_{min} and X_{max} represent the minimum and maximum values of the input parameter, respectively, and X_{norm} denotes the normalized value of the parameter within the range of 0 to 1.

After confirming the data quality, heat maps and simple scatter plots were used to identify the input parameter that has the most influence on th

e output parameter, as shown in **Figures 7 and 8**. Both methods demonstrated that an increase in heterogeneity and variability within the reservoir results in fluctuations in observed values, which impacts the identification of a linear relationship between the logs and permeability. This study investigates multiple cases, each case utilizing different input logs, to determine the most effective approach for accurately determining permeability. The aim is to save time and enhance accuracy in future exploration efforts. The input logs cases are shown in **Table.2.**

Subsequently, Principal Component Analysis (PCA) to reduce the dimensionality of a dataset while preserving and enhancing variance [45, 46] . Principal Component Analysis (PCA) is a method used for reducing the dimensionality of a dataset while preserving and enhancing the variance to improve the efficiency of data analysis methods, including data exploration, data visualization, and machine learning [47,48]. The covariance matrix was determined for each case as indicated in the **Table.3**. The results demonstrate that PCA successfully reduced the multi input logs cases while preserving their essential features for instance in case (NPHI–GR–DT), three input logs were reduced to one effective log that explain (81 %) of variance, in addition in case 9 (DT–GR–RHOB) were reduced to one effective log that explain (85 %) of variance. These results show that PCA is an effective approach to reducing the redundancy in highly correlated logs that will enable the creation of simple model structures that may enhance the linear relationship between the input logs and the permeability target the linearity relashinship between input logs and target, hence allowing more efficient learning by the predictive model [23]. Therefore, the dates were divided into three datasets: 50% for training, 20% for testing the models, and 30% for validating the model using unseen data. The training set illustrates the discovery of concealed intricate patterns, while the testing set can be considered a preliminary stage to evaluate the model's robustness and a validation set to confirm the consistency of the model and ensure hyperparameter selection. The present strategy was applied in various studies [21,23],provided an optimal balance between different sets and ensured the reliability of predictions.

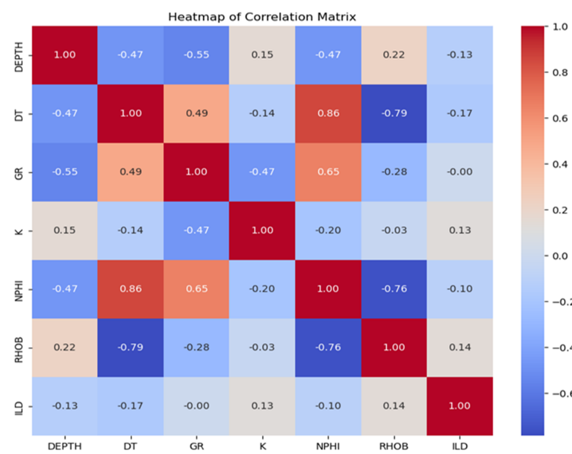


Figure 7. Heat map displaying the correlation coefficient between the variables. The color intensity and color bar indicate the strength and direction of the linear relationship between each pair of variables, ranging from −1 to +1.

Table 2 . Cases and their Corresponding Input Logs Utilized in this study

Cases	Input logs
1	ILD
2	NPHI
3	RHOB
4	GR
5	DT
6	NPHI-GR-DT
7	GR-DT
8	NPHI-DT
9	DT-GR-RHOB
10	ILD-RHOB
11	ILD-GR
12	DT-RHOB

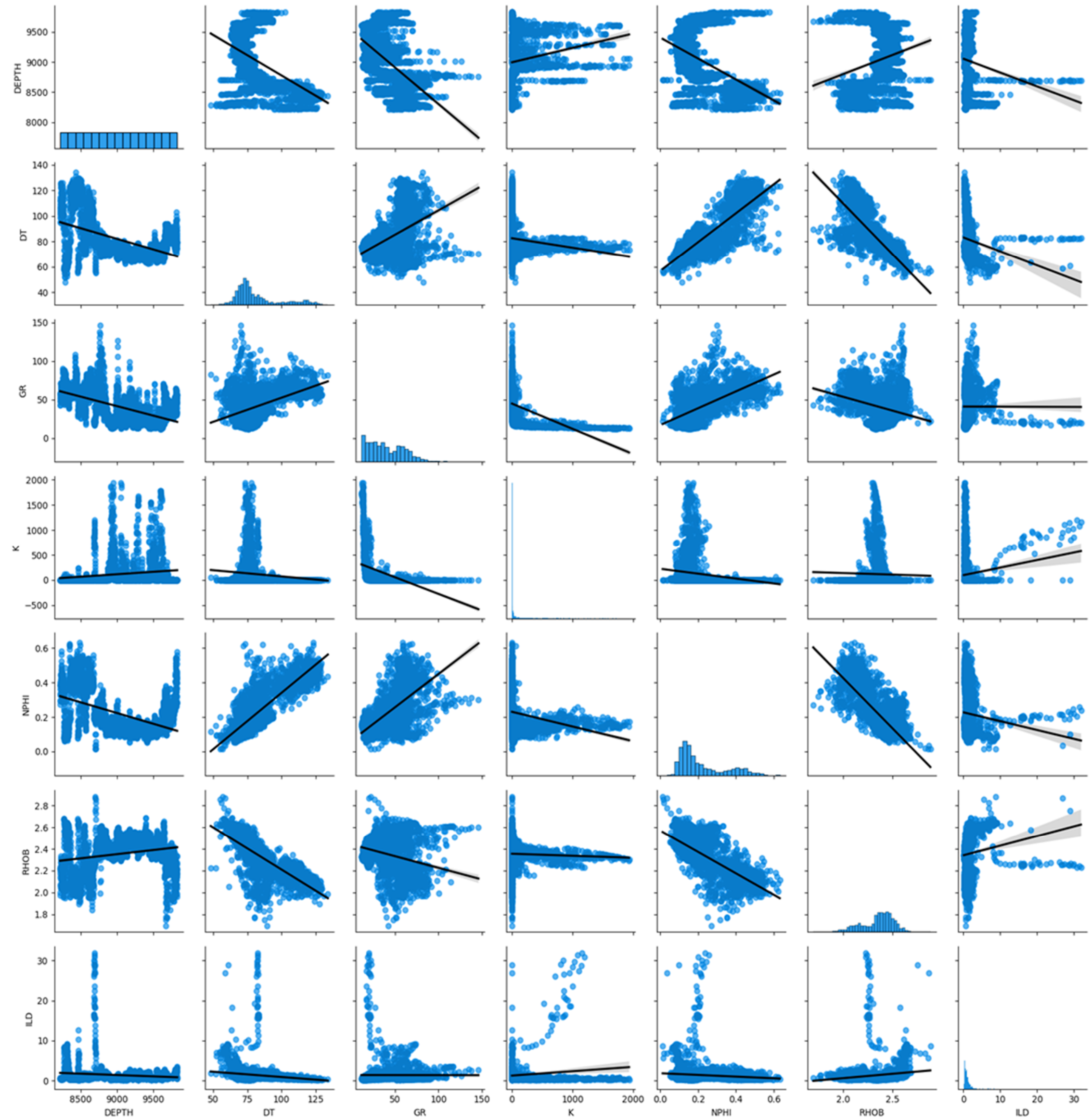


Figure 8. Multi-scatter plots between different variables display the fitted regression line of these relationships, which may be strong, weak, positive, negative, or neutral.

Table 3. Findings of explained variance for each case, explained variance reflects the percentage of essential features preserved while reducing the dimensionality for each case.

Cases	Input logs	No.of components	explained variance
1	ILD	1	89%
2	NPHI	1	88%
3	RHOB	1	94%
4	GR	1	95%
5	DT	1	93.1%
6	NPHI-GR-DT	1	81%
7	GR-DT	1	86%
8	NPHI-DT	1	79%
9	DT-GR-RHOB	1	85%
10	ILD-RHOB	1	81%
11	ILD-GR	1	82%
12	DT-RHOB	1	87%

2.2. Model Development.

An artificial neural network (ANN) is a mathematical model of a biological neural network according to [49]. In this work, the Multi-Layer Perceptron (MLP) was employed to examine all cases. MLP is widely recognized as the primary ML technique for forecasting petrophysical properties based on wireline logs [50]. MLP is considered the most basic configuration of an ANN which is made up of a fully connected structure containing an input layer, a hidden layer, and an output layer [51,52]. Each layer consists of neurons, also known as nodes. The nodes in the input layer are interconnected with the nodes in the hidden layer, which in turn are interconnected with the nodes in the output layer [21,53,54].The input layer nodes are linked to the hidden layer nodes, which are then linked to the output layer nodes. A weight assigned to a node-to-node connection is regulated through back-propagation during training iterations as explained by [55]. The general expression of MLP is

$$x_i^r = \Phi\left(\sum_{j=1}^{n} z_i^{r-1} \times G_{ji}^r + S_i^j\right) \tag{5}$$

The symbol z_i^{r-1} indicates the output of the previous layer. G_{ji}^r represents the weight connecting the neuron to the j--th input. S_i^j signifies the bias value of the i-th neuron. Φ (.) represents the transfer function. The subsequent loss function is used to execute the backpropagation operation at the output layer:

$$L = L(z_1^r, \dots, z_h^r) = \sum_{i=1}^{h} (z_i^r - e_i)^2 \tag{6}$$

The values r -th and e_i denote the output layer and the anticipated output of the i-th output neuron, respectively.

The implementation was carried out in the Python programming language using the Jupyter Notebook editor. The scikit-learn library was used for the machine learning method.

2.3. Hyperparameters Optimization.

To improve model consistency and accuracy, optimization of the model is necessary. MLP is constructed using various hyperparameters like optimizer, activation function, and the number of hidden layers and neurons, which need to be carefully tuned to create a high-performance model. In this study, Grid search CV was used to optimize all MLP models. Grid search CV is a commonly used method for hyperparameter tuning [56], which is based on systematically exploring all possible combinations of parameter values to find the optimal values that can improve the performance of a training algorithm. This process involves searching through predefined ranges for each hyperparameter and determining the values that result in the best model [57].

2.4. Performance Metrics

Several measures are employed to assess the effectiveness of machine learning models. The study utilizes performance metrics such as Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Squared Error (MSE), which are defined by the following equations.

$$R^2 = \frac{\sum_{i=1}^{m} (\tilde{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{m} (y_i - \bar{y}_i)^2} \tag{7}$$

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=0}^{n-1} y_i - \tilde{y}_i\right)} \tag{8}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m} |(y_i - \tilde{y}_i)|, \tag{9}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N} (\tilde{y}_i - y_i)^2 \tag{10}$$

The variables in the equation are defined as follows: m indicates the number of samples, $y_i$denote the real value, $\tilde{y}_i$ indicate the predicted value, and $\bar{y}_i$ represents the mean of samples.

3. Results

Forecasting permeability in heterogeneous reservoirs poses significant challenges because of the non-linear relationships among different geological parameters. Various elements, including differences in mineral makeup, cracks, and changes caused by compression or fluid-filled pores, can result in the unevenness of reservoirs. These elements augment the complexity of the reservoir regarding permeability prediction, making it more difficult to capture all these characteristics using conventional logs, which are tailored to specific purposes. To develop high-performance machine learning models, it is essential to select the optimal input log combination, as this decision profoundly affects the model's accuracy. The research analyzes the efficacy of using various input logs across distinct scenarios to forecast permeability in the Nubian sandstone of the Sirt basin. Over 950 permeability data points were utilized as unseen data to assess the efficacy of all models. Twelve cases were assessed, with each case constructed from distinct sets of input logs, and all models were executed using the optimally determined hyperparameters, as seen in **Table 4**, which demonstrates that each scenario used different hyperparameters to achieve the model's objectives. This disparity pertains to the characterization of input logs and their correlation with permeability.

Overall, the results indicate that models using single-input logs demonstrate greater accuracy compared to those utilizing multiple log inputs. The GR model exhibits the greatest R^2 value of 0.994 across all scenarios, featuring an MSE of 301 and an RMSE of 17.7. The results demonstrate the effectiveness of GR in identifying the minerals and lithological differences within the target reservoir. Similarly, instances when RHOB and DT logs are used as input cases (3 & 5) demonstrated considerable prediction accuracy, shown by R^2 values of 0.982 and 0.966, respectively. For cases (1 and 2), which use ILD and NPHI as input logs, the findings demonstrated a moderate correlation coefficient between anticipated and actual data, with R^2 values of 0.82 and 0.80, respectively.

In contrast, the various input log scenarios (cases 6–12) display a wide range of accuracy levels. Cases 6-8 showed lower to moderate correlation coefficients, featuring R^2 values of 0.50, 0.81, and 0.43, respectively. The findings indicate that the selected models are unable to attain a high degree of accuracy in forecasting permeability. Nonetheless, models constructed from examples (9-12) provide robust and satisfactory results for permeability prediction. The DT-GR-RHOB model (case 9) had a high R^2 of 0.902 and a low RMSE of 18.15, suggesting that the selection of an appropriate log combination might enhance prediction accuracy. Furthermore, the ILD-GR (case 11) and DT-RHOB (case 12) models exhibited exceptional performance, shown by high R^2 values of 0.910 and 0.906, respectively, and minimal RMSE and MAE values. In conclusion, the results indicate that single-input logs, particularly GR, DT, and RHOB, yield greater prediction accuracy compared to scenarios with multiple input logs. All selected models' performances are compared in Figure.9.

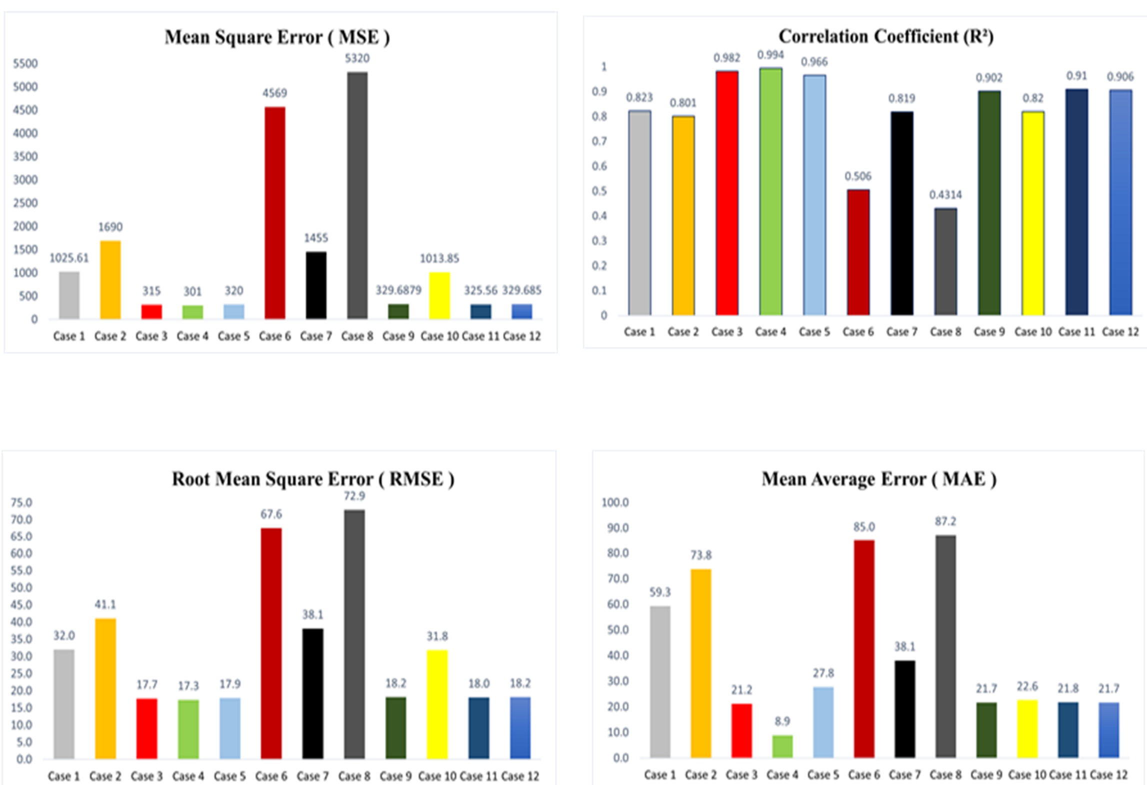
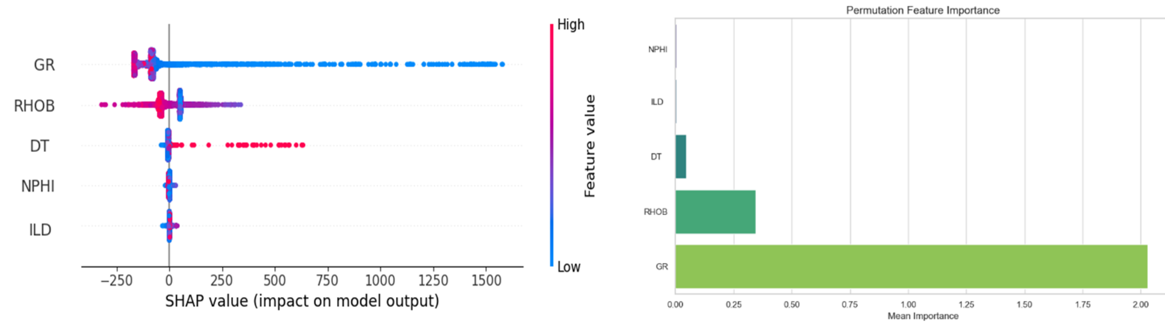


Figure 9. Performance indicators for each model, the plot presents a comprehensive performance metrics for each model and its corresponding case study.

3.1. Further validation.

SHAP (Shapley Additive Explanations) was employed to validate the study's results by illustrating the most significant input log affecting permeability prediction based on MLP. SHAP assesses the significance of each characteristic by evaluating the influence of each input feature on permeability estimates. Visualizing SHAP values improves model interpretability for end-users and provides insights into selecting the optimal combination of input features [58 ,59].Furthermore, Permutation importance (PI) was employed to assess feature importance. Permutation Importance (PI) is an interpretability method employed to ascertain the significance of features based on their influence on the predictions of a trained machine learning model [22 , 59 , 60] The results, as seen in the Figures.10a and 10b, demonstrate that GR is the most significant characteristic, followed by RHOB and DT, respectively. The findings align with the study's findings, indicating that the GR exhibited superior accuracy compared to other input logs.



(a) **(b)**

Figure 10. (a) SHAP (Shapley Additive Explanations) (b) Permutation feature importance, both plots demonstrate that GR, RHOB, and DT are the most influential input logs on permeability predictions

Table 4. Selected hyperparameters by Grid search CV for each predictive model.

Cases	Best Hyperparameters
1	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.01', 'number of neurons' : 64, 'max_iter': 160, 'power_t': 0.5, 'optimizer': Adam,
2	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 64, 'max_iter': 280, 'power_t': 0.5, 'optimizer': Adam,
3	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.1', 'number of neurons' : 32, 'max_iter': 125, 'power_t': 0.5, 'optimizer': Adam,
4	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.01', 'number of neurons' : 64, 'max_iter': 300, 'power_t': 0.5, 'optimizer': Adam,
5	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 52, 'max_iter': 100, 'power_t': 0.5, 'optimizer': Adam,
6	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.01', 'number of neurons' : 220, 'max_iter': 360, 'power_t': 0.5, 'optimizer': Adam,
7	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 32, 'max_iter': 200, 'power_t': 0.5, 'optimizer': Adam,
8	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 325, 'max_iter': 300, 'power_t': 0.5, 'optimizer': Adam,
9	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 200, 'max_iter': 450, 'power_t': 0.5, 'optimizer': Adam,
10	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.01', 'number of neurons' : 112, 'max_iter': 300, 'power_t': 0.5, 'optimizer': Adam,
11	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.001', 'number of neurons' : 300, 'max_iter': 500, 'power_t': 0.5, 'optimizer': Adam,
12	'activation': 'relu', 'hidden_layer_sizes': 2, 'learning_rate': '0.1', 'number of neurons' : 180, 'max_iter': 300, 'power_t': 0.5, 'optimizer': Adam,

4. Discussions

Machine learning techniques, particularly neural networks, have shown great promise in handling complex, non-linear relationships in data. These non-linear relationships are widely demanded in heterogeneous reservoirs, which are characterized by their complex geological structures and variations in rock properties [59]. This complexity makes it difficult to establish straightforward relationships between input logs and permeability [60,61].Using the appropriate input logs leads to better and more accurate results. Multiple methods were used to choose the best input logs, such as Heat maps and linear cross plots. These methods sometimes provide some limitations to building a relation with the permeability, especially in heterogeneous and unconventional reservoirs. In addition, there is a limited discussion of the comparative efficiency of different logs and combinations thereof. Understanding which logs contribute most effectively to prediction accuracy can optimize the modeling process. This study aims to explore the effect of different input wireline logs on the accuracy of permeability predictions by conducting multiple cases with varying input combinations. The finding of the study indicates that using single-input logs would provide more precise results, probably because the model focuses on the strong signal from one type of measurement. The finding indicates that GR logs provides the highest R^2, which probably related to the ability of GR in capture the clay minerals that invaded the pores which is has a primary impact on permeability especially in Nubian reservoirs where the sandstone is dominated by shale, but that's maybe not work in others different reservoirs, so investigation previous researches on the reservoirs may lead to select the appropriate input logs based on geological analysis. Several prior studies concluded that using a single feature as input logs provides more accurate results because they capture dominant characteristics without increasing unnecessary noise or multicollinearity associated with additional features, which leads to reduced overfitting, increasing the generalization, and improving the models' performance [62,63,64].On the other hand, the decreased accuracy in multiple log cases can probably be related to an increase in the complexity of the model and complications in the learning process, requiring more sophisticated techniques to handle potential multicollinearity and noise [65,66]. For instance, generating a linear relationship between porosity and permeability requires a high degree of homogeneity, and that's not available in a complex paleoenvironment, where the rock is formed using different mineralogy, lithology, and various pore throats, which leads to an increase the complexity. This complexity may be reduced by using PCA, but in some cases, with more features than samples, the covariance matrix used in PCA becomes singular, making the computation of principal components problematic [67]. Furthermore, the complex models are more susceptible to overfitting, becoming overly tailored to the training data and performing poorly on unseen data[67,68]. In general, the finding of this study suggests that single logs are strong predictors of permeability compared to multi-log cases. Using multiple logs can generate a more complex model since each log provides different information and knowledge, which increases model complexity, reduces the generalization, and necessitates careful consideration to prevent overfitting.

4.1. Benchmarking and Broader Context.

Selecting the best combinations of input logs enhances prediction accuracy across various subsets, especially in geoscience applications. Various feature engineering methods were introduced to offer valuable insights in selecting the suitable input logs. For instance, a heat map, which is a technique of feature engineering, is commonly employed in determining the best input selection by evaluating the range of correlation coefficients from −1 to 1 [50]. Values nearing 1 suggest a strong positive correlation, while those closer to −1 reflect a strong negative correlation [23]. Feature engineering methods face constraints, especially in geological reservoirs where non-linear relationships prevail.[23,50].To overcome these limitations, this study intends to save time by evaluating the effectiveness of twelve distinct models, each constructed using various combinations of input logs. The results were confirmed by comparing them with previous studies; for instance, a study by [67] to forecast permeability in carbonate reservoirs using ANN, relying on RHOB and GR as input logs, attaining a significant correlation coefficient (R^2 = 0.98). A different study by [68] utilized data that were gathered from multiple wells drilled in the Western Desert and Gulf regions of Egypt through an ANN. The study shows that DT, RHOB, and GR are the key input logs for permeability predictions, attaining an R^2 of 98% during training and 96.5% during testing. A study effort by [50] applied different machine learning techniques to forecast permeability for the Precipice Sandstone in the Surat Basin, Eastern Australia; these ML techniques used DT, RHOB, and GR as input logs, achieving R^2 values of 0.93 and 0.87 for the training and blind data sets, respectively. In a study by [69], RHOB and GR were used to predict permeability in the Malay Basin, Malaysia; the results showed the MLP model attained R^2 = 0.99 for the training set and R^2 = 0.93 for the testing set.

The benchmarking results show that GR, RHOB, and DT consistently serve as strong predictors of permeability across different basins. The present study confirms the strength of these logs in the Nubian Sandstone Formation of the Sirt Basin, Libya, and highlights the trade-off between the simplicity of the model and the predictive accuracy in heterogeneous reservoirs [50,70,71].

4.2. Partial dependence plot (PDP)

A partial dependence plot (PDP) acts as a visual instrument for machine learning models, facilitating the analysis of the relationship between input variables and a model's aggregate predictions. PDP is instrumental in identifying scenarios where the effect of one variable on predicted outcomes is dependent on the value of another variable [70]. The findings illustrated in **Figure. 11** indicate that permeability is highly sensitive to variations in GR, particularly in zones where GR is below 18 API; this sensitivity diminishes significantly when GR exceeds 25 API due to the increase of clay minerals within the pores. For RHOB, the plot reveals that permeability becomes less sensitive when RHOB ranges from 2.4 to 2.5 G/CC, in contrast to the pronounced variability observed when RHOB is below 2.4 G/CC. Conversely, DT, NPHI, and ILD remain consistently uniform across the entire spectrum. In conclusion, GR and RHOB serve as performance indicators for permeability variation compared to other logs[58].

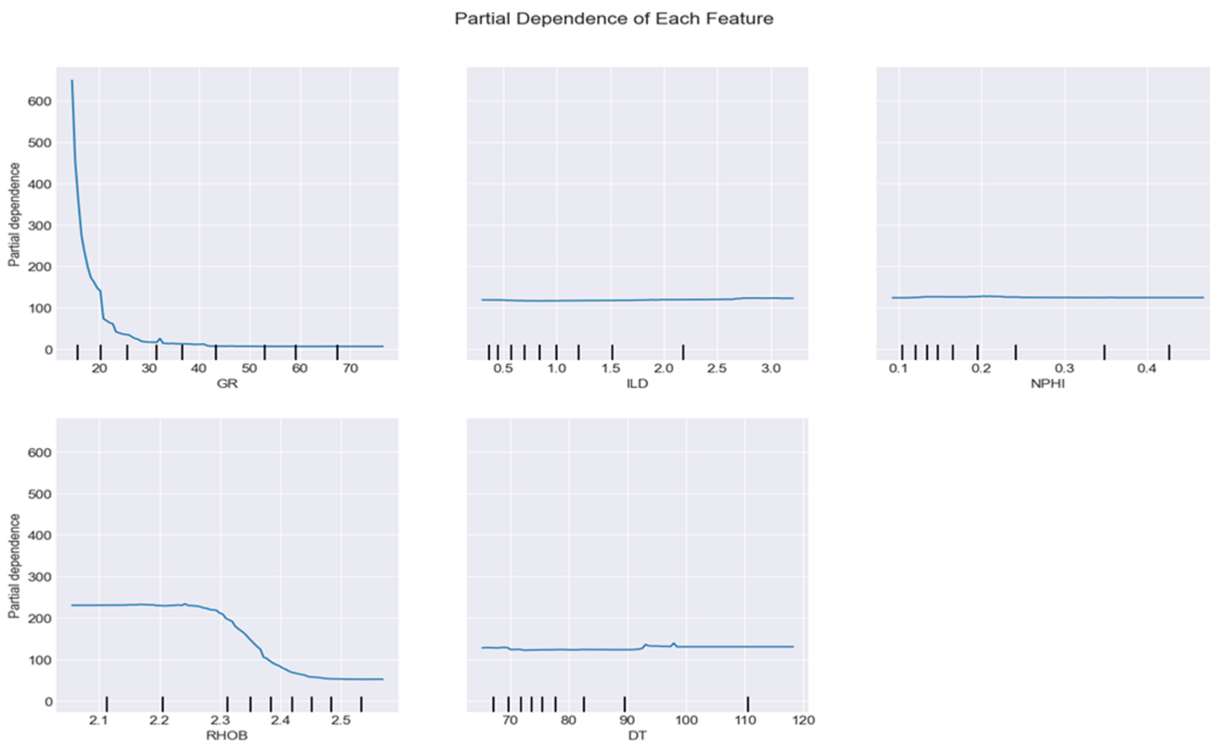


Figure 11. A partial dependence plot (PDP) illustrates the sensitivity of permeability to variations in each input log feature.

4.3. Limitations and future works.

The present work has certain limitations that were observed, particularly regarding data diversity. Utilizing conventional well logs as input for training models is prevalent; however, this approach may result in diminishing the generalization throughout the training and validation phases. Integrating seismic attributes with well logs and production data could enhance the model's reliability when used across several locations in the region. Future research must further explore the optimal balance between input data volume and model complexity, emphasizing the development of methods that can effectively integrate diverse inputs while minimizing noise and redundancy.

5. Conclusions

Permeability is an essential physical characteristic of porous materials that assesses the capacity to convey fluid. It serves a crucial function in assessing reservoirs and advancing oil and gas fields. Numerous studies have been introduced to estimate permeability using machine learning methods based on conventional well logs data. Although previous studies have attained high accuracy results, the effectiveness of these logs or their combinations on Permeability prediction accuracy has not been comprehensively explored, especially in the context of highly heterogeneous reservoirs. The present study presents practical guidance to understand the effectiveness of input logs on prediction processes and to discover which log or a log combination provides higher accuracy findings. The presented workflow is based on building twelve distinct models, each constructed using a Multi-Layer Perceptron (MLP), based on various combinations of input logs using data from the Nubian reservoir, Sirt Basin. Over 3,200 permeability core data and wireline records are being processed to improve their quality. Subsequently, Principal Component Analysis (PCA) was utilized to reduce dimensionality, and then the data were partitioned into training, testing, and validation sets. The models have been optimized by Grid Search optimization to identify the most effective hyperparameters. The innovation of the present work lies in integrating input-log testing with advanced sensitivity analysis methods (SHAP and permutation importance), and the Partial dependence plot (PDP) to provide deeper interpretability of model performance. The Main conclusions of the study are as follows:

1. The findings demonstrate that single-input logs, mainly the Gamma-ray (GR), Bulk density (RHOB), and Sonic logs (DT), exhibited higher correlation coefficients compared to the multiple log combinations.
2. The GR model achieves the highest R^2 of 0.994, although RHOB and DT demonstrate R^2 values of 0.98 and 0.96, respectively, reflecting their sensitivity to non-linear correlations.
3. The enhanced efficacy of single log models may be attributed to the lack of ambiguity of the signal derived from particular measurements, which are less susceptible to being distorted by the complexities brought by several inputs.
4. The findings from the SHAP and permutation techniques for sensitivity analysis indicate that GR is the most significant attribute, followed by RHOB and DT, respectively. The results correspond with the study's conclusions, demonstrating that the GR showed enhanced accuracy relative to other input logs.
5. The findings from PDP indicate that permeability is highly sensitive to variations in GR compared to other logs, particularly in zones where GR is below 18 API.
6. Certain constraints were identified in the current study, mainly concerning the variety of the data that might reduce generalization. Future research should focus on combining various regional datasets to ensure the reliability across different locations.

Nomenclature

PCA	Principal Component Analysis
ANN	Artificial Neural Network
MLP	Multi-Layer Perceptron
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
MSE	Mean Squared Error
SHAP	Shapley Additive Explanations
PDP	Partial Dependence Plot
NPHI	Neutron Porosity
RHOB	Bulk Density
ILD	Induction resistivity
DT	Sonic log
GR	Gamma-ray

References.

[1]. Tian, J., Qi, C., Sun, Y., Yaseen, Z. M., & Pham, B. T. (2021). Permeability prediction of porous media using a combination of computational fluid dynamics and hybrid machine learning methods. *Engineering with Computers*, *37,3455-3471.*

[2]. Liu, J. J., & Liu, J. C. (2022). Permeability predictions for tight sandstone reservoir using explainable machine learning and particle swarm optimization. *Geofluids, (1), 2263329.*

[3]. Shehata, A. A., Osman, O. A., & Nabawy, B. S. (2021). Neural network application to petrophysical and lithofacies analysis based on multi-scale data: An integrated study using conventional well log, core and borehole image data. *Journal of Natural Gas Science and Engineering*, *93,104015.*

[4]. Tiab, D., & Donaldson, E. C. (2015). Petrophysics: Theory and practice of measuring reservoir rock and fluid transport properties (4th ed.), *Elsevier.*

[5]. Amir, M. A., Abdelmalik, M. B. A., & Elorfi, H. A. (2022). Effect of uranium content on petrophysical parameters: A case study of the Mamuniyat reservoir in the Murzuq Basin, southeastern Libya. In *SAGE Record*, Benghazi, Libya, 1.

[6]. Abdollahipour, A., Fatehi Marji, M., Yarahmadi Bafghi, A., & Gholamnejad, J. (2015). Numerical investigation on the effect of crack geometrical parameters in hydraulic fracturing process of hydrocarbon reservoirs. *Journal of Mining and Environment*, *205-214.*

[7]. Musleh, D. A., Olatunji, S. O., Almajed, A. A., Alghamdi, A. S., Alamoudi, B. K., Almousa, F. S., Aleid, R. A., Alamoudi, S. K., Jan, F., Al-Mofeez, K. A., & Rahman, A. (2023). Ensemble learning based sustainable approach to carbonate reservoirs permeability prediction. *Sustainability*, *15, 1351–1373.*

[8]. Xie, B., Wang, Y., Zhou, X., Zhao, Z. A., Zhou, X., Zhao, X. R., Li, K. X., Yu, D. G., & Xie, T. (2015). Production prediction based on the heterogeneity analysis of the tight carbonate reservoir in the Sichuan Basin, China. In *SPE/IATMI Asia Pacific Oil Gas Conference and Exhibition*, *SPE-176091.*

[9]. Hulea, I. N. (2013). Capillary pressure and permeability prediction in carbonate rocks - New methods for fractures detection and accurate matrix properties prediction. In *SPE Middle East Oil Gas Show Conference Proceedings*, *SPE-164251.*

[10]. Wyllie, M. R. J., & Rose, W. D. (1950). Some theoretical considerations related to the quantitative evaluation of the physical characteristics of reservoir rock from electrical log data. *Journal of Petroleum Technology*, *2,105-118.*

[11]. Coates, G. R., & Dumanoir, J. L. (1973). A new approach to improved log-derived permeability, *In SPWLA annual logging symposium.*

[12]. Carman, P. C. (1997). Fluid flow through granular beds. *Process Safety and Environmental Protection*, 75.463-484

[13]. Chen, C. H., & Lin, Z. S. (2006). A committee machine with empirical formulas for permeability prediction. *Computers & Geosciences*, *32,485-496.*

[14]. Balan, B., Mohaghegh, S., & Ameri, S. (1995). State-of-the-art in permeability determination from well log data: Part 1 - a comparative study, model development. In *SPE Eastern Regional Conference and Exhibition Proceedings.SPE-30978.*

[15]. Abdulkhaleq, H. B., Khalil, K. A., Al-Mudhafar, W. J., & Wood, D. A. (2024). Advanced machine learning for missing petrophysical property imputation applied to improve the characterization of carbonate reservoirs. *Geoenergy Science and Engineering*, *238, 212900.*

[16]. Akkuş, E., Bal, U., Koçoğlu, F. Ö., & Beyhan, S. (2024). Hyperparameter optimization of pre-trained convolutional neural networks using adolescent identity search algorithm. *Neural Computing and Applications*, *36,1523–1537.*

[17]. Dalirnasab, A., Marji, M., Nejati, H.-R., & Mohebi, M. (2024). Effects of porosity on the strength and mechanical behaviour of porous geo-materials under cyclic loading. *Rudarsko-Gorniczy Zbiornik*, *39, 15–30.*

[18]. Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., Keehm, Y., Krzikalla, F., Lee, M., Madonna, C., Marsh, M., Mukerji, T., Saenger, E. H., Sain, R., Saxena, N., Ricker, S., & Wiegmann, A. (2013). Digital rock physics benchmarks-Part I: Imaging and segmentation. *Computers & Geosciences*, *50, 25-32.*

[19]. Keehm, Y., & Nur, A. (2003). Computational rock physics: *Transport properties in porous media and applications*, Stanford university.

[20]. Mehmani, A., & Prodanović, M. (2014). The effect of microporosity on transport properties in porous media. *Advances in Water Resources*, *63,104-119.*

[21]. Amir, M. A., & Amir, H. S. (2025). Bulk density prediction in missed intervals of Nubian reservoir using multi-machine learning and empirical methods. *Arabian Journal of Geosciences,18(3) ,58.*

[22]. Amir, M. A., & Amir, H. S. (2025). Exploring shear waves predictions in heterogeneous reservoirs using multi-machine learning models based on conventional wireline logs: A case study from the Nubian reservoir, Sirt Basin, Libya. *Arabian Journal of Science and Engineering,1-27.*

[23]. Amir, M. A., Amir, H. S., Abdel-Fattah, M. I., Elsteil, M. O., Mogren, S., Ibrahim, E., Nabawy, B. S., Ramah, M., & Reda, M. (2025). New approach to reduce the uncertainty level of the shear wave velocity prediction by utilizing ensemble model: A case study from Sirt Basin, Libya. *Earth Science Informatics*, *18, 463.*

[24]. Hussain, K., Fatehi Marji, M., Rasouli, V., & Emami Meybodi, E. (2022). Application of machine learning models for predicting the rock fracture toughness mode-I and mode-II. *Journal of Mining and Environment*, *13, 465–480.*

[25]. Al-Mudhafar, W. (2017). Integrating well log interpretations for lithofacies classification and permeability modeling through advanced machine learning algorithms. *Journal of Petroleum Exploration and Production Technology*, *7, 1023–1033.*

[26]. Al-Anazi, A., & Gates, I. D. (2010). Support-vector regression for permeability prediction in a heterogeneous reservoir: A comparative study. *SPE Reservoir Evaluation & Engineering*, *13.*

[27]. Elkatatny, S., Mahmoud, M., Tariq, Z., & Abdulraheem, A. (2018). New insights into the prediction of heterogeneous carbonate reservoir permeability from well logs using artificial intelligence network. *Neural Computing and Applications*, *30.*

[28]. Mahdaviara, M., Rostami, A., Keivanimehr, F., & Shahbazi, K. (2021). Accurate determination of permeability in carbonate reservoirs using Gaussian

process regression. *Journal of Petroleum Science and Engineering*, *196.*

[29]. Anifowose, F., & Abdulraheem, A. (2011). Fuzzy logic-driven and SVM-driven hybrid computational intelligence models applied to oil and gas reservoir characterization. *Journal of Natural Gas Science and Engineering*, *3, 505-517.*

[30]. Al-Mudhafar, W. J. (2019). Bayesian and LASSO regressions for comparative permeability modeling of sandstone reservoirs. *Natural Resources Research*, *28,47-62.*

[31]. Abdulraheem, A., Sabakhi, E., Ahmed, M., Vantala, A., Raharja, I., & Korvin, G. (2007). Estimation of permeability from wireline logs in a Middle Eastern carbonate reservoir using fuzzy logic. In *SPE Middle East Oil Gas Show Conference Proceedings,SPE-105350.*

[32]. Hallett, D., & Clark-Lowes, D. (2016). *Petroleum geology of Libya (2nd ed.). Elsevier.*

[33]. Hassan, H. S., & Kendall, C. C. G. (2014). Hydrocarbon provinces of Libya: A petroleum system study. In *AAPG Memoir*, *pp. 101–141.*

[34]. Rusk, D. C. (2001). Libya: Petroleum potential of the underexplored basin centers - A twenty-first-century challenge. *AAPG Memoir*, *74.*

[35]. Abdelmalik, M. B. A., Esteil, M., Amir, M. A., Al Brouki, M., & Amir, H. S. (2022). Pore pressure prediction from well logs: A model study in the Nubian Sandstone reservoir, Sarir Field, southeastern Sirt Basin, Libya. In *SAGE Record 071*, Benghazi, Libya, *1.*

[36]. Fan, C., Chen, M., Wang, X., Wang, J., Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, *9,652801.*

[37]. Ye, X. (2022). Optimization of multimodal Japanese teaching model using virtual reality. *Mobile Information Systems*, *652801.*

[38]. Dash, C. S. K., Behera, A. K., Dehuri, S., & Ghosh, A. (2023). An outliers detection and elimination framework in classification task of data mining. *Decision Analytics Journal*, *6, 100164.*

[39]. Fang, Q., Nguyen, H., Bui, X. N., Nguyen-Thoi, T., & Zhou, J. (2021). Modeling of rock fragmentation by firefly optimization algorithm and boosted generalized additive model. *Neural Computing and Applications*, 33(8), 3503-3519.

[40]. Kang, P. (2013). Locally linear reconstruction based missing value imputation for supervised learning. *Neurocomputing*, *118,65-78.*

[41]. Tukey, J. (1975). Mathematics and the picturing of data. *Proceedings of the International Congress of Mathematicians,523-531.*

[42]. Maletic, J., & Marcus, A. (2000). Data cleaning: Beyond integrity analysis. *IQ*.

[43]. Azadpour, M., Saberi, M. R., Javaherian, A., & Shabani, M. (2020). Rock physics model-based prediction of shear wave velocity utilizing machine learning technique for a carbonate reservoir, southwest Iran. *Journal of Petroleum Science and Engineering*, *195,2263329.*

[44]. Chen, T., Gao, G., Wang, P., Zhao, B., Li, Y., & Gui, Z. (2022). Prediction of shear wave velocity based on a hybrid network of two-dimensional convolutional neural network and gated recurrent unit. *Geofluids*, *9974157*

[45]. Surono, S., Goh, K. W., Onn, C. W., & Marestiani, F. (2023). Developing an optimized recurrent neural network model for air quality prediction using K-means clustering and PCA dimension reduction. *International Journal of Innovative Research in Science Studies*, *6,330-43.*

[46]. Mousa, F. A., El-Khoribi, R. A., & Shoman, M. E. (2016). A novel brain computer interface based on principle component analysis. In *Procedia Computer Science,82, 49-56.*

[47]. Salem, N., & Hussein, S. (2019). Data dimensional reduction and principal components analysis. In *Procedia Computer Science,163, 292-299.*

[48]. de Vargas, D. L., Oliva, J. T., Teixeira, M., Casanova, D., & Rosa, J. L. G. (2023). Feature extraction and selection from electroencephalogram signals for epileptic seizure diagnosis. *Neural Computing and Applications*, 35(16), *12195-12219.*

[49]. Erofeev, A., Orlov, D., Ryzhov, A., & Koroteev, D. (2019). Prediction of porosity and permeability alteration based on machine learning algorithms. *Transport in Porous Media*, *128(2), 677-700.*

[50]. Rezaee, R., & Ekundayo, J. (2022). Permeability prediction using machine learning methods for the CO2 injectivity of the Precipice Sandstone in Surat Basin, Australia. *Energies*, *15,2053.*

[51]. Liu, S., Zhao, Y., & Wang, Z. (2021). Artificial intelligence method for shear wave travel time prediction considering reservoir geological continuity. *Mathematical Problems in Engineering*, *5520428.*

[52]. Shaban, W. M., Ashraf, E., & Slama, A. E. (2024). SMP-DL: A novel stock market prediction approach based on deep learning for effective trend forecasting. *Neural Computing and Applications*, *36,1849-1873.*

[53]. Lee, H., Kim, D., & Gu, J. H. (2023). Prediction of food factory energy consumption using MLP and SVR algorithms. *Energies*, *16,1550.*

[54]. Poongodi, M., Malviya, M., Kumar, C., Hamdi, M., Vijayakumar, V., Nebhen, J., & Alyamani, H. (2022). New York City taxi trip duration prediction using MLP and XGBoost. *International Journal of*

System Assurance Engineering and Management, *13, 16–27.*

[55]. Ahmed, A., Gamal, H., Elkatatny, S., & Ali, A. (2022). Bulk density prediction while drilling vertical complex lithology using artificial intelligence. *Journal of Applied Geophysics*, *199.*

[56]. Ghate, V., & Hemalatha, C. (2023). A comprehensive comparison of machine learning approaches with hyper-parameter tuning for smartphone sensor-based human activity recognition. *Measurement: Sensors*, 30,100925.

[57]. Ogunsanya, M., Isichei, J., & Desai, S. (2023). Grid search hyperparameter tuning in additive manufacturing processes. *Manufacturing Letters*, *35,1031-1042.*

[58]. Davoodi, S., Al-Shargabi, M., Wood, D., Mehrad, M., & Rukavishnikov, V. (2025). A robust hybrid near-real-time model for prediction of drilling fluids filtration. *Engineering with Computers*, *1–25.*

[59]. Tan, Y., Chen, H., Zhang, J., Tang, R., & Liu, P. (2022). Early risk prediction of diabetes based on GA-Stacking. *Applied Sciences*, *12,632.*

[60]. Duong, H. T., & Nguyen-Thi, T. A. (2021). A review: Preprocessing techniques and data augmentation for sentiment analysis. *Computational Social Networks*, *8.*

[61]. Anemangely, M., Ramezanzadeh, A., Amiri, H., & Hoseinpour, S. A. (2019). Machine learning technique for the prediction of shear wave velocity using petrophysical logs. *Journal of Petroleum Science and Engineering*, *174, 306–327.*

[62]. Talebkeikhah, M., Sadeghtabaghi, Z., & Shabani, M. (2021). A comparison of machine learning approaches for prediction of permeability using well log data in the hydrocarbon reservoirs. *Journal of Human, Earth and Future*, *2.*

[63]. Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, *59,82-99.*

[64]. Manning, C. D., Schütze, H., & Weikurn, G. (2002). *Foundations of statistical natural language processing. SIGMOD Record*, *31,01635808.*

[65]. Zhu, P., Lin, C., Wu, P., Fan, R., Zhang, H., & Pu, W. (2015). Permeability prediction of tight sandstone reservoirs using improved BP neural network. *Open Petroleum Engineering Journal*, *8,288-292.*

[66]. Czarnowski, I. (2011). Distributed learning with data reduction. In *Lecture Notes in Computer Science*, *6660, 3–121.*

[67]. Iguyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3,1157-1182.*

[68]. Moussa, T., Elkatatny, S., Mahmoud, M., & Abdulraheem, A. (2018). Development of new permeability formulation from well log data using artificial intelligence approaches. *Journal of Energy Resources Technology, Transactions of the ASME*, *140,072903.*

[69]. Abdel Azim, R., & Aljehani, A. (2022). Neural network model for permeability prediction from reservoir well logs. *Processes*, *10,2587.*

[70]. Arafath, M. Y., Haque, A. K. M. E., Siddiqui, N. A., Venkateshwaran, B., & Ali, S. (2025). Permeability prediction and potential site assessment for CO2 storage from core data and well-log data in Malay Basin using advanced machine learning algorithms. *ACS Omega*, *10, 5430–5448.*

[71]. Reda M., Mostafa T., Abdel-Fattah M.I., Amir H.S., Amir M.A., Raef A., Fathy M., Al-Kahtany K., Farouk S. (2025). Integrated petrophysical analysis, seismic interpretation and 3D static reservoir modelling for rock-typing and geomechanical characterisation of Cretaceous reservoirs in the Arcadia and Dorra oil fields, Western Desert, Egypt. *Geological Journal*, *1–27.*

انجمن مهندسی معدن ایران
نشریه مهندسی معدن و محیط زیست
نشانی نشریه: www.jme.shahroodut.ac.ir
دانشگاه صنعتی شاهرود

اثرات انتخاب نمودار ورودی بر دقت نفوذپذیری در مخازن ناهمگن حوضه سرت، لیبی

محمد امیر[1*]، حمزه امیر[2]، و مختار فرکاش[3]

۱ .گروه علوم زمین، دانشکده علوم، دانشگاه بنغازی، بنغازی، لیبی

۲ .گروه مهندسی نفت، دانشکده مهندسی، دانشگاه ستاره درخشان، البریقه، لیبی

۳. شرکت نفت خلیج عربی (AGOCO)، بنغازی، لیبی

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۵/۰۵/۰۸

تاریخ داوری: ۲۰۲۵/۰۹/۲۴

تاریخ پذیرش: ۲۰۲۵/۱۱/۱۹

DOI: 10.22044/jme.2025.16192.3128

کلمات کلیدی

یادگیری ماشین

مشخصات مخزن

نفوذپذیری

چاه نگاری

ناهمگنی

چکیده

تخمین نفوذپذیری یک مرحله ضروری در ارزیابی پتانسیل هیدروکربن در محیط متخلخل و طراحی روش‌های مدیریت مخزن است. اخیراً، روش‌های یادگیری ماشین (ML) در پیش‌بینی نفوذپذیری اهمیت پیدا کرده‌اند. مرحله اولیه در ساخت مدل‌های ML با قابلیت اطمینان بالا، شناسایی ترکیبات بهینه از لاگ‌های ورودی است، زیرا نفوذپذیری یک پارامتر بسیار حساس است. این مرحله ضروری است و می‌تواند بر دقت مدل تأثیر بگذارد. در حالی که روش‌های مهندسی ویژگی، بینش‌های ارزشمندی در انتخاب لاگ‌های ورودی مناسب ارائه می‌دهند، اثربخشی این لاگ‌ها یا ترکیبات آنها، به ویژه در زمینه مخازن با ناهمگنی بالا، همچنان ناشناخته مانده است. مطالعه حاضر قصد دارد با ارزیابی اثربخشی دوازده مدل مجزا، که هر کدام با استفاده از یک پرسپترون چند لایه (MLP) ساخته شده‌اند، بر اساس ترکیبات مختلف لاگ‌های ورودی با استفاده از داده‌های مخزن نوبیان، حوضه سیرت، لیبی، در زمان صرفه‌جویی کند. این روش شامل چندین مرحله از جمله پیش‌پردازش، تقسیم، بهینه‌سازی و اعتبارسنجی است. یافته‌ها نشان می‌دهند که نگاره‌های تک ورودی، عمدتاً نگاره‌های پرتو گاما (GR)، چگالی حجمی (RHOB) و نگاره‌های صوتی (DT)، ضرایب همبستگی بالاتری را در مقایسه با ترکیب‌های چندگانه نگاره نشان می‌دهند. مدل GR به بهترین R^2 برابر با ۰.۹۹۴ دست یافت که نشان‌دهنده حساسیت آن در ثبت روابط غیرخطی است. از سوی دیگر، مدل‌های چند نگاره به دقت متغیری دست یافتند که منجر به افزایش پیچیدگی یادگیری شد. این مطالعه کارایی انتخاب ترکیب بهینه نگاره‌های ورودی را برجسته می‌کند و راهنمایی عملی برای پیش‌بینی نفوذپذیری مبتنی بر یادگیری ماشین در مخازن ناهمگن ارائه می‌دهد.

نویسنده مسئول مکاتبات: *Itzamir924@gmail.com* (م. امیر)