Petroleum and Coal

Article **Open Access**

Enhancing Porosity Prediction in Reservoir Characterization through Ensemble Learning: A Comparative Study between Stacking, Bayesian Model Optimization, Boosting, and Random Forest

*Mohamed Riad Youcefi[1,2] *, Ayman Inamat Alshokri[3], Walid Boussebci[1], Khaled Ghalem[1], Asma Hadjadj[1]*

[1] *Department of Process Engineering, Faculty of Technology, University Amar Telidji-Laghouat, BP37G 03000 Laghouat, Algeria*

[2] *Laboratory of Petroleum Equipment's Reliability and Materials, Faculty of Hydrocarbons and Chemistry, University M'hamed Bougara of Boumerdes, Avenue de l'Indépendance, 35000 Boumerdes, Algeria*

[3] *Sirte University, Petroleum Engineering Department, Libya*

Received June 6, 2024; Accepted September 4, 2024

Abstract

Accurate estimation of porosity is a critical factor in reservoir characterization. This study aims to enhance porosity prediction through the implementation and comparison of various stacking ensemble learning strategies. A dataset comprising 273 points, which consists of well logs and core measurements, was collected from two wells for model development. Four base learners, including Support Vector Regression (SVR), Multi-Layer Perceptron (MLP), Random Forest Regression (RFR), and XGBoost, were trained on this dataset. These models were then integrated using multiple stacking ensemble techniques, such as weighted averaging, Bayesian model averaging, and RFR as a meta-learner. Meta-learners were trained on predictions from the base learners, generated through cross-validation on leave-out data. Performance evaluations of both base and meta learners were conducted on a separate testing dataset using statistical and graphical error analysis. Results indicate that all learners demonstrated robust performance, with weighted averaging outperforming other strategies on testing data. The stacking ensemble approach, particularly through weighted averaging, effectively improved base learner performance on testing data by leveraging individual model strengths and mitigating weaknesses. The findings of this study are valuable for geoscientists and reservoir engineers in achieving accurate reservoir characterization and facilitating exploration activities.

Keywords: *Reservoir porosity; Stacking ensemble learning; Reservoir characterization; Machine learning; boosting; Random forest regression.*

1. Introduction

Reservoir characterization is essential in developing a comprehensive geoscience model that helps predict the behavior of reservoir fluids under various conditions and identify the best production strategies to maximize the hydrocarbons recovery. This process involves a detailed analysis of hydrocarbon-bearing reservoirs by integrating precise measurements from various sources to optimize performance over time. As a multidisciplinary field, it aims to accurately describe petroleum deposits and the geological properties of the rocks containing hydrocarbons [1-2].

One of the critical factors in reservoir characterization is the accurate estimation of porosity. Porosity is the ratio of the volume of void spaces in a rock to its total bulk volume. This measure indicates the rock's capacity to store and transmit fluids, such as oil, water, and gas, which is crucial for reserve estimation [3-4]. Accurate porosity assessment plays a significant

role in making informed decisions during the exploration, development, and production phases, ultimately impacting field development strategies.

Determining rock porosity can be achieved through direct and indirect methods. Direct measurements typically involve laboratory testing of core samples taken from subsurface hydrocarbon zones [5-6]. This approach is considered highly accurate but is often costly, time-consuming, and limited to specific intervals within the reservoir [7]. Indirect methods include well logging, which is more economical and routinely performed during drilling and well-completion operations. Various well logs, such as bulk density, neutron porosity, sonic, and nuclear magnetic resonance logs, are used to estimate porosity [5,8]. However, these logs can introduce errors due to tool calibration issues, borehole conditions, and the impact of drilling mud [9–11].

Numerous theoretical and empirical correlations for reservoir properties prediction have been proposed in the literature [12–14]. These models aim to provide quick and cost-effective estimates of porosity by investigating the relationships between various rock properties such as porosity, permeability, resistivity, pore throat size, and water saturation. However, their applicability is often limited by the specific conditions and parameters for which they were developed, making it challenging to apply them universally across different reservoir types, particularly in complex systems like naturally fractured carbonate rocks with dual porosity [15]. Given the limitations of both direct core measurements and indirect well log methods, there is a pressing need to develop and refine porosity estimation techniques. Such methods must be reliable, cost-effective, and time-efficient, providing accurate results even in the face of reservoir heterogeneity and incomplete data sets.

Recently, the Machine Learning (ML) techniques have been merged as a reliable tool for predicting various parameters and properties in the oil and gas industry. This wide implementation of ML algorithms contributed to solving several technical issues such as predicting reservoir fluid properties [16–19], rock properties [20–22], and geo-mechanical properties [23]. This successful and widespread application of ML models stems from their ability to analyze large datasets, extract insights from field data, and identify patterns and correlations that might be difficult or impossible to discern through traditional methods.

Given the promising capabilities of ML in addressing these challenges, a growing body of research has focused on leveraging ML techniques for predicting rock porosity in hydrocarbon reservoirs. Zhang *et al.* [24] implemented hybrid models such as Support Vector Machine (SVM) optimized with Particle Swarm Optimization (PSO), to predict the permeability and the porosity in a heterogeneous dolomite reservoir. The authors revealed that combining ML approaches with the flow zone indicator method provides a more accurate continuous estimation of permeability and porosity. Ahmadi and Chen [25] employed hybridization of multiple algorithms methodologies, including Conventional Neural Network (CNN), Genetic Algorithm (GA), fuzzy decision tree, the Imperialist Competitive Algorithm (ICA), and PSO for predicting permeability and porosity. Okon *et al.* [26] applied Artificial Neural Network (ANN) to predict porosity, permeability, and water saturation based on well logging data. Gamal *et al.* [27] proposed a new methodology for predicting the porosity in real time based on drilling data by employing Random Forest (RF) and Decision Tree (DT) algorithms. Abdulaziz *et al.* [28] implemented Probabilistic Neural Network (PNN) to predict total porosity using seismic attributes. Chen *et al.* [29] developed a multilayer Long Short-Term Memory (LSTM) model to predict the porosity based on well log data. The authors reported that compared with conventional neural networks, LSTM reduce the porosity prediction errors when data is insufficient and logs have varied depths. Ifrene *et al.* [30] proposed a hybrid model that combines SVM classification and ANN regression for an improved fracture porosity prediction. Other recent studies were introduced by [31–33] investigated the application of various machine learning for predicting reservoir porosity based on well logging data.

The present study aims to leverage core measurements, well logging data, and the power of ML to identify patterns and correlations, ultimately building a robust model that can predict porosity with high accuracy based on well logging data. Unlike previous studies that applied various machine learning models and then selected the best models based on prediction errors, this study implements stacking ensemble learning. This approach combines base model

predictions to capitalize on the strengths of each individual model while mitigating their weaknesses, thereby enhancing the overall accuracy of porosity prediction. Additionally, this work conducts a comparative study of different ensemble learning techniques, including Weighted Averaging, XGBoost, Bayesian Model Averaging (BMA), and using Random Forest as a meta-learner.

2. Material and methods

2.1. Data acquisition and analysis

To develop high-performance machine learning (ML) models for predicting formation porosity based on well logging data, comprehensive data were collected from Well-A and Well-B. Core samples extracted from these wells were subjected to routine core analysis in a laboratory, where they were cleaned, dried, and their porosity were measured using a gas porosimeter based on Boyle's law. Concurrently, corresponding measurements of the formation's physical properties, namely, gamma ray (GR), neutron porosity (NPHI), photoelectric factor (PEFZ), bulk density (RHOZ), and shallow resistivity (RXOZ), were recorded as functions of depth via well logging operation. The well logs were calibrated to core data through depth-shifting to ensure that the physical measurements from the well log data were accurately aligned with the experimentally measured porosities. Fig. 1 illustrates the well logs and core porosity against depth.

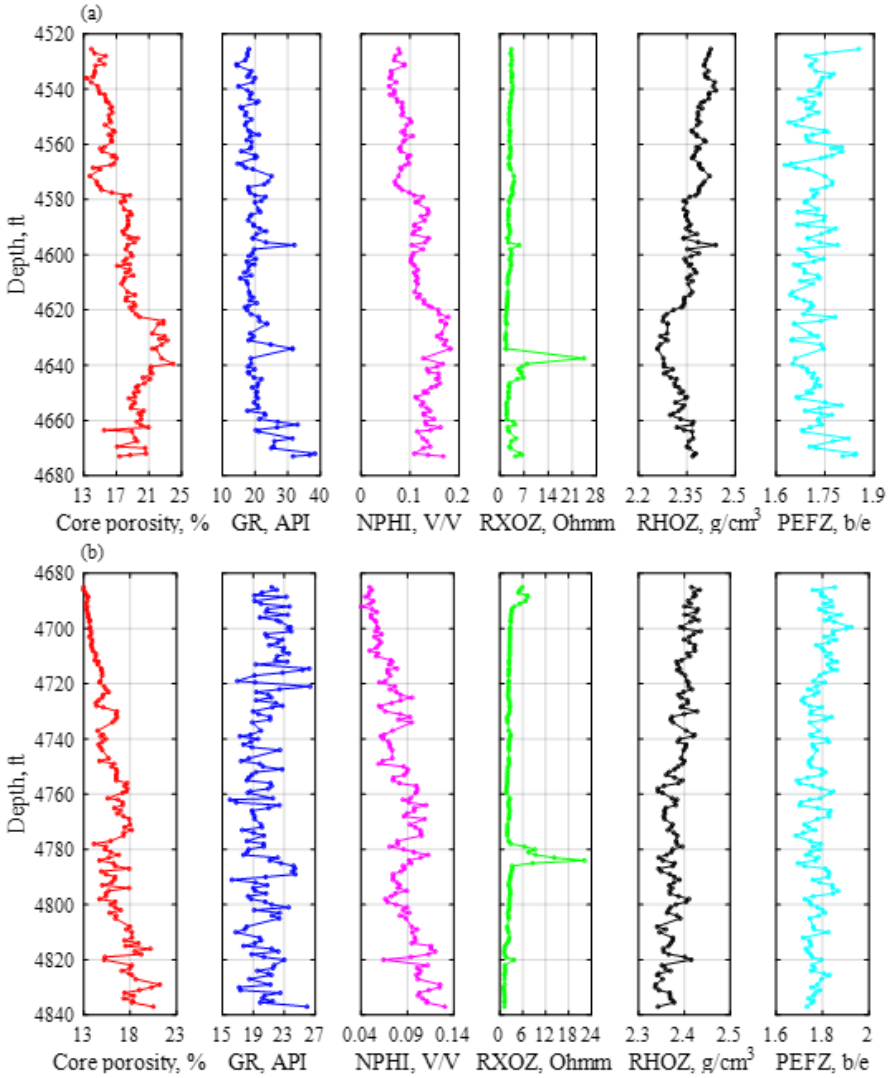

Fig. 1. Coring and well logging raw data: (a) well-A, (b) well-B.

After data processing, which included discarding outliers and unrealistic values such as null and negative readings caused by sensor malfunctions, a dataset comprising 273 points was assembled to construct and test the accuracy and reliability of the proposed ML models. This dataset includes porosity as the output variable and the aforementioned physical properties as input variables. Table 1 presents the statistical indices of the input and output parameters.

Table 1. Statistical description of the input/output data.

Statistical parameter	GR	NPHI	RHOZ	PEFZ	RXOZ	Porosity
Unit	API	V/V	g/cm^3	b/e	Ohm	%
Maximum	38.3388	0.1819	2.4401	1.9250	24.3868	23.97
Minimum	14.2955	0.0401	2.2584	1.6271	0.9673	13
Mean	20.4561	0.0976	2.3711	1.7517	3.0369	16.9678
Standard deviation	3.2574	0.0313	0.0374	0.0557	2.2964	2.3362

2.2. Stacking ensemble

Stacking is an ensemble learning technique introduced by Wolpert [34]. This method has been extensively applied across various domains due to its ability to enhance model performance by reducing both bias and variance, effectively mitigating overfitting. Stacking operates on a hierarchical framework consisting of two levels: base learners and a meta-learner. The first level, or level 0, comprises diverse machine learning models which are trained on the original dataset [35]. These base learners generate predictions that form a metadata. The second level, or level 1, features the meta-learner, which should be a relatively simple model to prevent overfitting and to handle correlated inputs effectively, is trained using the metadata.

The use of diverse algorithms at the base level enhances the model's generalization ability by leveraging the complementary strengths of each algorithm. A key aspect of stacking is the use of cross-validation to prevent overfitting [36-37]. The predictions from the base models are not directly learned by the meta-learner; instead, techniques such as leave-one-out cross-validation or K-fold cross-validation are employed to generate a new dataset for the meta-learner. This strategy ensures that all original datasets participate in the training process, enhancing the robustness of the final model. The meta-learner, often a multiple linear regression model, combines the outputs of the base learners to make the final prediction. Stacking differs from other ensemble methods like bagging and boosting, which typically involve combining models of the same algorithm. Instead, stacking combines various regressors, such random forests, support vector regressors, artificial neural networks, naive Bayes, and logistic regression, to improve modeling accuracy [36]. By integrating the outputs from diverse base models through a meta-learner, stacking capitalizes on the mutual complementarity among the base models to achieve superior predictive performance.

2.3. Extreme gradient boosting (xgboost)

Extreme Gradient Boosting (XGBoost) developed by Chen and Guestrin [38], is a supervised ensemble machine learning approach based on an ensemble of boosted trees. It represents a scalable and advanced application of the gradient boosting technique. It gradually combines weak base learners into a more powerful model by fitting the training data to the base model and subsequently fitting a second model to the residuals to enhance learning [38]. This process of iteratively addressing residuals continues until the specified criteria are met, with the final result being an aggregation of all base models' outputs. To prevent overfitting, XGBoost incorporates a regularization term in its objective function [38-39]. Unlike gradient boosting's negative loss criterion for tree splitting, XGBoost uses a depth-first approach, pruning trees in reverse using the maximum depth option. Sequential tree construction in XGBoost is achieved through parallel implementation, with the algorithm's inner loop computing tree characteristics and the outer loop enumerating leaf nodes, thereby enhancing algorithm efficiency.

2.4. Random forest regressor

Random Forest Regressors (RFR) is an ensemble-based machine learning technique used for regression task. It consists of multiple decision trees, each trained on randomly selected subsets of features and data samples [40]. The training process begins with creating a bootstrap sample from the original dataset by randomly selecting samples with replacement. Each bootstrap sample is used to build an unpruned decision tree, considering a random subset of features at each node to identify the optimal split. This process is repeated to construct multiple decision trees. The advantages of random forests include their ability to handle high-dimensional data without the need for feature selection, robustness to outliers and noise, and efficient processing of large-scale datasets. A key characteristic is the use of out-of-bag (OOB) error estimation, which provides an unbiased error estimate during training [41]. The predictions from individual trees are averaged to produce the final prediction, enhancing the model's accuracy and performance. The combination of bootstrap sampling, random feature selection, and full-depth tree construction ensures that random forests are versatile and powerful for various machine learning applications.

2.5. Multilayer perceptron

A Multilayer Perceptron (MLP), a type of ANN, is inspired by the neural architecture of the human brain [42-43]. This feedforward neural network consists of an input layer, one or more hidden layers, and an output layer. The input layer receives training data, which is processed through the hidden layers before producing the final output. Each hidden layer consists of multiple neurons, computational units, that take inputs from other interconnected neurons. These inputs are weighted and then subjected to a nonlinear transformation via an activation function, such as Sigmoid, ReLU, or Tanh. The activation function processes the linear combination of inputs, facilitating the nonlinear transformations necessary for complex problem solving [43-44]. During the training process, the weights of the connections between neurons are adjusted using a backpropagation algorithm to minimize the model's prediction error [43–45]. The MLP model has been widely implemented in various fields due to its versatility and capability to analyze complex problems [46–48].

2.6. Support vector regression

Support Vector Regression (SVR), a branch of the SVM framework, is designed for regression tasks [49]. Unlike SVM, which focuses on classification by separating data points into distinct classes, SVR aims to identify a hyperplane that minimizes the global deviation of sample points from it. This optimal hyperplane maximizes the margins by including support vectors, which are the data points closest to the hyperplane that defines the margin. In addition, SVR achieves its functionality by nonlinearly mapping the input features into a high-dimensional feature space using a kernel function. This mapping allows SVR to handle complex, nonlinear relationships in the data. The Radial Basis Function (RBF) kernel is particularly effective in SVR applications, as it enables SVR to achieve superior performance in estimating complex regression tasks. Further details of the SVR model refer to [49-50].

3. Methodology

To enhance robustness in predicting porosity based on well logging, a stacking ensemble learning was employed in the present work. The proposed stacking framework consists of four base models: SVR, MLP, RFR, and XGBoost. Fig 2. illustrates the implemented stacking framework utilized in this paper. Initially, the dataset was split into two subsets: training and testing. To ensure a representative distribution of data points into these subsets, Latin Hypercube Sampling (LHS) was utilized. 80% of the collected data was dedicated to building both the base and meta-learners, while 20% was reserved as unseen data for the meta-learner to make predictions and undergo evaluation. Each base model was trained on the training dataset using the k-fold cross-validation approach, and their predictions were used to generate a new dataset for the meta-learner.

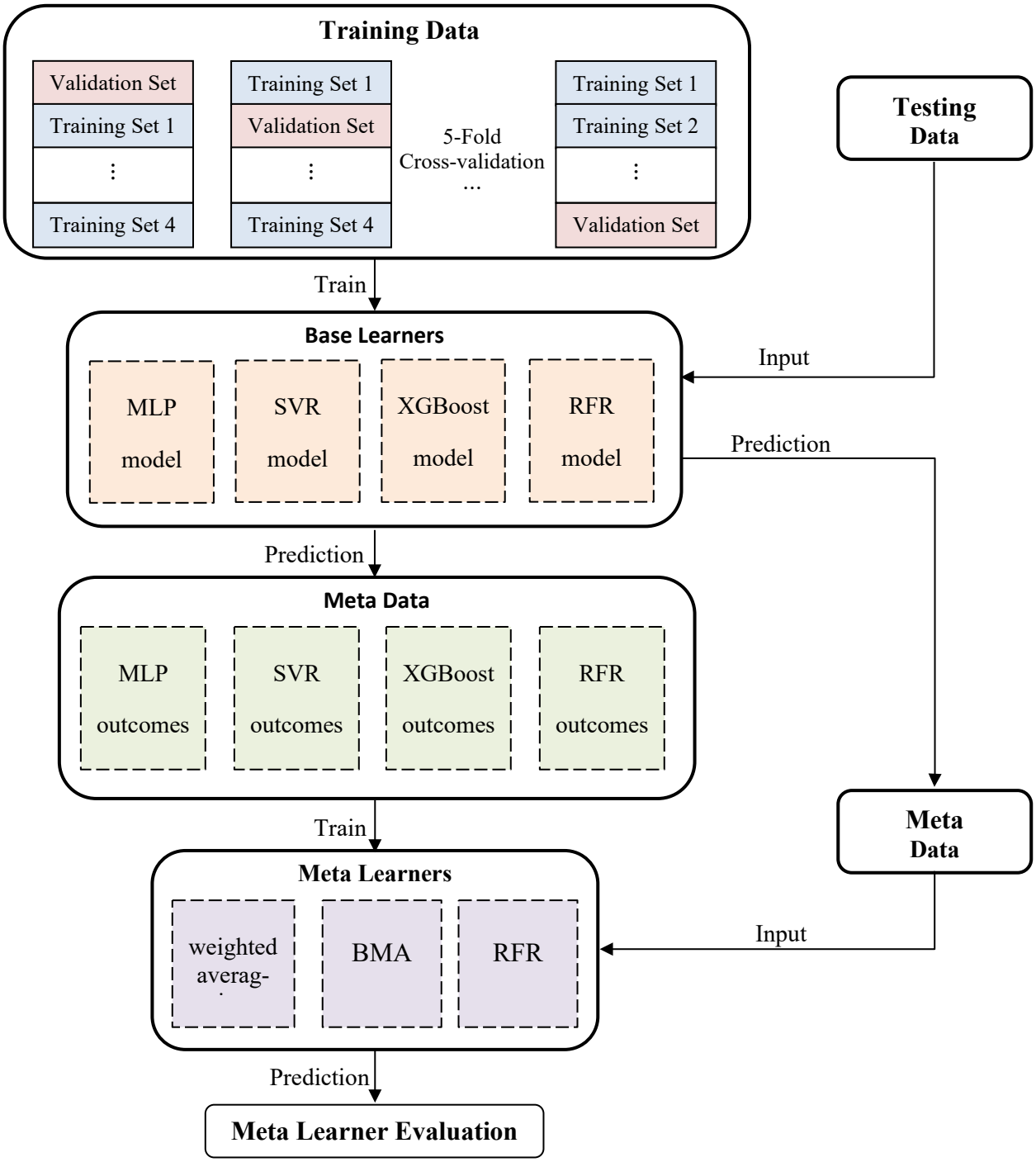

Fig. 2. Flowchart of the stacking ensemble methodology used in this study.

3.1. Hyperparameter tuning

To optimize the performance of the machine learning models, we employed Bayesian optimization for hyperparameter tuning using the Optuna library [51]. Unlike traditional random search, Bayesian optimization is more efficient tool as it employs a probabilistic model to choose the most promising hyperparameters, reducing the number of evaluations needed [52]. For the SVR model, we tuned parameters such as the regularization parameter C, the epsilon parameter ε, the kernel function, and the kernel coefficient. The MLP model's hyperparameters included the number of hidden layers and neurons, the activation function, the initial learning

rate, and the solver for weight optimization. The XGBoost model's hyperparameters consists of defining the number of estimators, the maximum depth of a tree, the learning rate, the minimum loss reduction, and regularization terms. For the RFR model, we adjusted the number of estimators, the maximum depth of trees, the minimum number of samples required to split a node, and the minimum number of samples required at each leaf node. Optuna's "create_study" function was used to define the optimization process, with each trial representing a different combination of hyperparameters. The objective function evaluated the performance of each set of hyperparameters using cross-validation to ensure robust and unbiased estimates of model performance.

3.2. Cross-validation for training base learners

To ensure the reliability and generalizability of our base models, we utilized k-fold cross-validation during their training process. In k-fold cross-validation, the training data is partitioned into k equal-sized folds. Each model is trained on k-1 folds and validated on the remaining fold, rotating through all k folds. This technique helps in assessing the model's performance across different subsets of data, thereby reducing the risk of overfitting and ensuring that the model performs well on unseen data.

For each base learner, including SVR, MLP, RF, and XGBoost, the following steps were performed:

- Data splitting: The training dataset was divided into 5 folds.
- Training and validation: Each model was trained on 4 folds and validated on the remaining fold. This process was repeated 5 times, with each fold serving as the validation set once.
- Performance aggregation: The performance metrics from each fold were averaged to obtain an overall performance estimate for the model.

This approach provided a more robust evaluation of the base models, ensuring that the selected hyperparameters and the trained models were not biased towards a particular subset of the data.

3.3. Training the meta-learner

In the process of constructing the meta-learner, predictions generated by the base models -SVR, MLP, RF, and XGBoost- based on the leave-out data during cross-validation were harnessed. After that, the meta-learner, also known as second-level model, was trained on the predictions made by the base models. In the present study, diverse set of approaches were investigated to combine the outputs of the base models. The goal was to enhance the overall prediction accuracy while avoid overfitting during the stacking ensemble. These approaches included:

- Weighted averaging**:** In this approach distinct weights are assigned to the predictions of each base model to emphasize the contributions of certain models while mitigating the influence of others. The determination of these weights was achieved through quadratic programming, optimizing them to minimize the mean squared error between the meta-learner predictions and the measered values.
- Bayesian Model Averaging (BMA): Employing BMA, we computed the posterior probabilities of each base model's predictions based on their respective marginal likelihoods [53]. By incorporating these probabilities into the prediction aggregation process, BMA facilitated the optimal combination of base model outputs, accounting for their respective uncertainties [53].
- Random Forest as Meta-Learner: In a distinct approach, we utilized a RFR as the meta-learner. Trained on the predictions generated by the base models, the Random Forest meta-learner learned to predict the target variable more accurately by capturing the complex interactions between the base models' outputs.

4. Results and discussion

In the present section, the performance of the base models, and the mate learners were assessed. Predictions are made based on testing data that were not seen by the models during

the training process. Three statistical indices, namely root mean square error (RMSE), mean absolute percentage error (MAPE), and the correlation coefficient R^2, were employed in this study to quantify the performance achieved by each model on both training and validation data. Table 3 highlights the model's performance. The expressions for these statistical indices are shown below:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\emptyset_{i,exp} - \emptyset_{i,pred})^2} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\emptyset_{i,exp} - \emptyset_{i,pred})^2}{\sum_{i=1}^{n}(\emptyset_{i,exp} - \emptyset_{i,pred})^2} \quad (2)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\emptyset_{i,exp} - \emptyset_{i,pred}}{\emptyset_{i,exp}}\right| \times 100 \quad (3)$$

where $\emptyset_{i,exp}$ and $\emptyset_{i,pred}$ represent the experimental and predicted values of porosity, respectively, while n stands for the number of data.

Table 2. Performance analysis of the meta and base learners on both training and testing data.

Model		Training data			Testing data		
		RMSE	MAPE	R^2	RMSE	MAPE	R^2
Base learners	XGBoost	0.4805	1.263	0.979	0.9628	4.2115	0.9118
	MLP	0.8987	4.0285	0.9242	1.039	4.0429	0.8908
	RFR	0.8284	3.7564	0.93703	1.0008	4.138	0.9005
	SVR	0.84	3.8292	0.9344	1.0441	4.2896	0.8898
Meta learners	Weighted averaging	0.5631	2.2911	0.9722	0.9401	3.9373	0.9128
	BMA	0.4793	1.2625	0.9791	0.9618	4.2088	0.912
	RFR	0.5306	2.4882	0.9950	0.9744	4.0927	0.9087

Table 3. Probability distribution of MAPE for meta and base learners.

Error	Stacking	SVR	MLP	RFR	XGBoost
MAPE <1	0.2407	0.1666	0.1481	0.2222	0.1851
MAPE <2.00	0.4259	0.3703	0.3703	0.3703	0.3703
MAPE <3.00	0.5555	0.574	0.4814	0.5	0.5370
MAPE <4.00	0.7037	0.6851	0.6296	0.6111	0.6111
MAPE <5.00	0.7592	0.7592	0.6481	0.6851	0.6851
MAPE <10.00	0.9444	0.9444	0.9259	0.9259	0.9074
MAPE <15.00	1	1	1	1	1

Table 2 highlights the performance of the models. Based on the analysis of Table 2, it is evident that all models demonstrate robust performance when evaluated on unseen testing data. The coefficient of determination R^2 spans from 0. 8898 to 0. 9128, while the MAPE and RMSE range from 3.9373 to 4.2896 and from 0.9401 to 1.0441, respectively. Notably, these models showcase consistent performance on data not encountered during the training phase, underscoring their capacity for generalization. Among the base learners, MLP emerges as the top performer in terms of MAPE, with a value of 4.0429. XGBoost stands out as the best performer in terms of RMSE and R^2, with values of 0.9628 and 0.9118, respectively. While SVR exhibits reasonably competitive metrics, its performance lags behind that of RFR, MLP, and XGBoost. Regarding the methodologies for combining base learners, weighted averaging emerges as the most effective approach, yielding superior prediction accuracy when employed to combine base learners and make predictions on unseen data. Although BMA and RFR as a meta-learner performs well on training data, they struggles to deal with unseen data and fails

to keep superiority on testing data. Further analysis comparing the performance of base learners and the meta-learner reveals that the stacking model, achieved through weighted averaging, performed best on the testing data. This indicates that by combining machine learning models of diverse structures, the stacking model can outperform all its base models. This approach leverages the strengths of each individual model while mitigating their weaknesses, leading to enhanced overall prediction accuracy and avoiding overfitting.

To visually assess the performance results, Fig. 3 illustrates the porosity values predicted by various models, including SVR, RFR, XGBoost, MLP, and the stacking model created through the weighted averaging approach, plotted against the core porosity values in parity plots. As can be obviously seen, the predictions made by the developed models demonstrate strong concordance with the measured values. Additionally, it is notable that for the that the outcomes of the XGBoost model are more closely aligned with the unit slope for the training data, whereas the stacking model's outcomes perform exceptionally well on testing data. This observation confirms the superior performance of the meta learner on unseen data compared to the others.

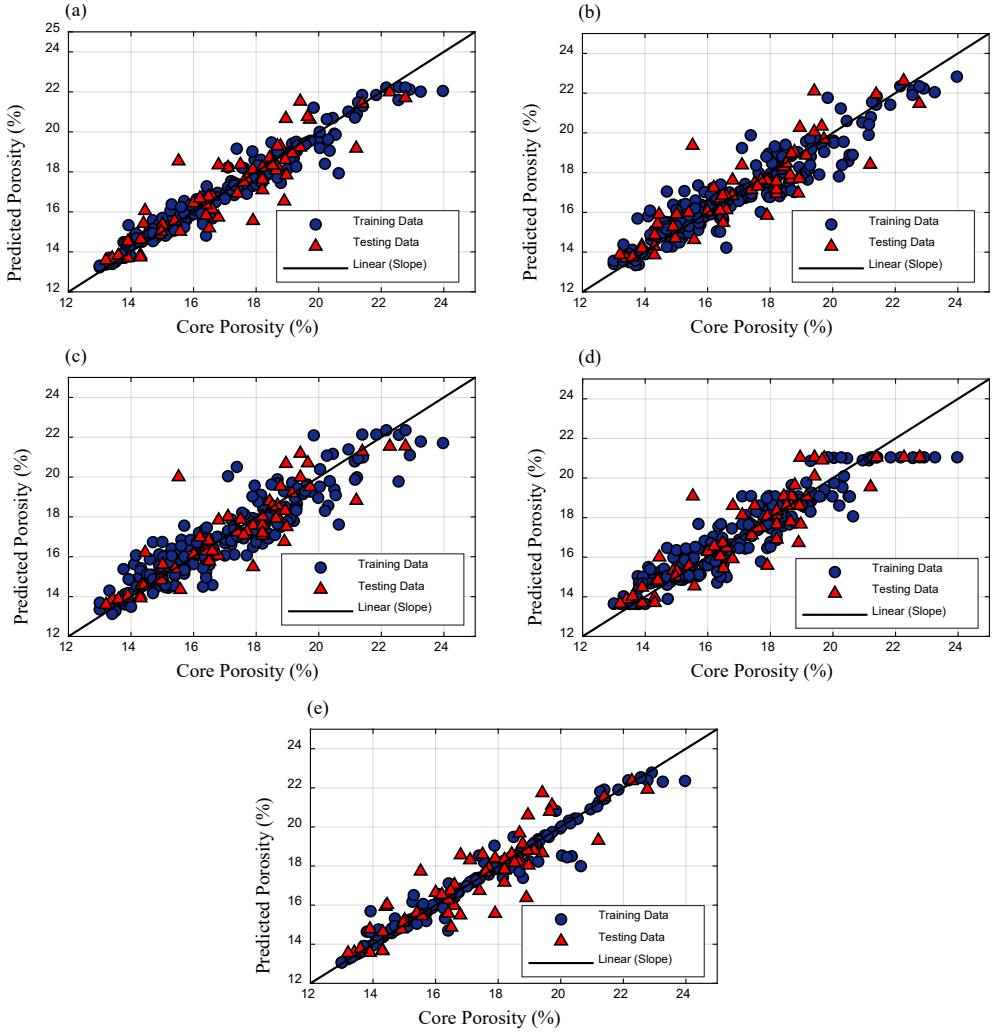

Fig. 3. Parity plot of core porosity and porosity predicted by meta and base learners: (a) Stacking model achieved through weighted averaging, (b) SVR model, (c) MLP model, (d) RFR model, (e) XGBoost model.

Another visual representation of performance on the testing data is shown in Fig. 4, which depicts the cumulative frequency of MAPE for both the meta and base learners. Table 3 provides details of Fig. 4 using the concept of MAPE probability distributions. A thorough analysis

of the cumulative frequency distribution of MAPE in Fig. 4 reveals that the meta model created through weighted averaging achieves the highest cumulative frequency at most given MAPE levels compared to the base learners. For instance, 24.07% of the stacking model predictions have a MAPE below 1%, whereas the corresponding cumulative frequencies for SVR, MLP, RFR, and XGBoost are 16.66%, 14.81%, 22.22%, and 18.51%, respectively. This demonstrates the superior performance of the meta learner.

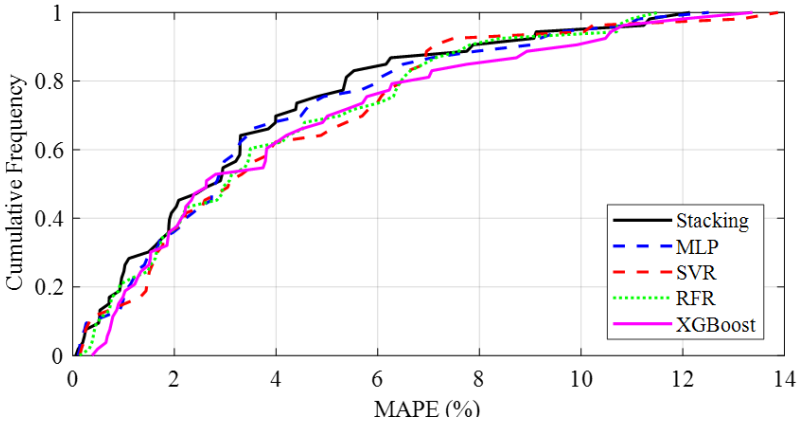

Fig. 4. Cumulative distribution of MAPE for meta and base learners.

Finally, the Leverage technique was utilized to confirm the statistical significance of the proposed stacking model and to delineate its applicability domain while identifying outliers. The Leverage approach is a dependable algorithm for detecting a group of data observations that deviate from the main body of the training dataset table . This method considers residual values, which indicate the differences between the model's predicted values and the measured data points, as well as a matrix known as the Hat matrix. The elements of this matrix are computed using the following equation [54]:

$$H = X(X^tX)^{-1}X^t \quad (4)$$

where X is a two-dimensional matrix comprising n data points and m input parameters.

The diagonal elements of the H matrix are referred to as Hat values. Williams' plot is employed to visually identify outliers by plotting the Hat values against the model's standardized residual values. The warning Leverage value H* is calculated using the formula $3(m+1)/n$.

In the present work, with n=273 and m=5, the warning Leverage value is calculated to be 0.0659. Most of the data points falling within the ranges $-3 < R < 3$ and $0 < H < H^*$ indicates that both the data employed for model development and the model's outcomes are within the applicability domain, confirming the model's statistical accuracy. Data points with $-3 < R < 3$ and $H^* < H$ are considered outside the scope of the model's applicability and are termed "Good High Leverage" points. Points within $R<-3$ or $R>3$, regardless of whether their Hat values are below or above H*, are identified as "Bad High Leverage" points or outliers [54]. Fig. 5 illustrates the Williams plot for the employed data in the stacking model.

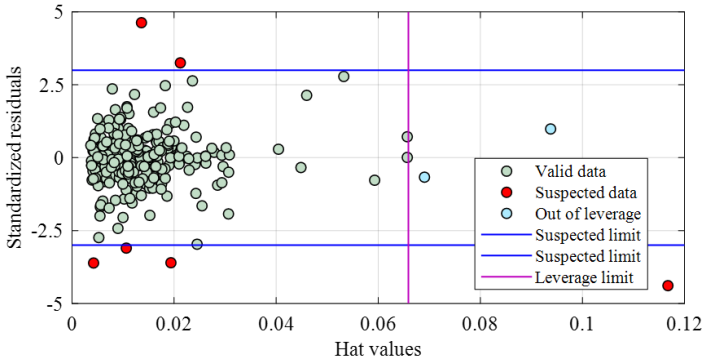

Fig. 5. The Williams plot of porosity dataset for the proposed stacking model.

This plot shows that the majority of data points fall within $0 < H < 0.0659$ and $-3 < R < 3$. Additionally, two data points (0.7326% of the data) lie outside the model's applicability range, while only six data points (2.1978% of the data) are identified as outliers. Consequently, the Leverage method demonstrates that the stacking model developed in this study for predicting the reservoir porosity and the used data are statistically reliable.

5. Conclusions

In this study, four base learners, including SVR, MLP, RFR, and XGBoost, were trained were trained using well logging and core measurements. Subsequently, these base learners were integrated through various stacking ensemble strategies, including weighted averaging, Bayesian model averaging, and RFR as a meta-learner. Different meta-learners were trained on predictions made by the base learners on leave-out data during cross-validation. The performance of both the meta and base learners was then evaluated on a separate testing dataset, unseen during the training process, and compared using statistical and graphical error analysis. The key findings are summarized as follows:

- All base and meta learners demonstrated good performance on both training and testing datasets for porosity prediction. This underscores the importance of the comprehensive methodology implemented in the current study to mitigate overfitting.
- Among the implemented stacking ensembles, the weighted averaging approach outperformed the others on the testing data, although Bayesian model averaging and RFR exhibited better results on the training data. This suggests that simpler models for combining base learner outcomes may provide better generalization capabilities compared to more complex models, which can lead to overfitting.
- The stacking ensemble learning approach, achieved through weighted averaging, effectively enhanced the performance of the base learners, particularly on the testing data, highlighting its ability to leverage the strengths of each individual model while mitigating their weaknesses.
- Overall, this study underscores the effectiveness of machine learning, particularly stacking ensembles, in accurately predicting porosity. These findings are valuable for geoscientists and reservoir engineers in achieving accurate reservoir characterization and facilitating exploration activities.

Nomenclature

ANN	*Artificial Neural Network*
BMA	*Bayesian Model Averaging*
CNN	*Conventional Neural Network*
DT	*Decision Tree*
GA	*Genetic Algorithm*
GR	*Gamma Ray*
ICA	*Imperialist Competitive Algorithm*
LHS	*Latin Hypercube Sampling*
LSTM	*Long Short-Term Memory*
MAPE	*Mean Absolute Percentage Error*
ML	*Machine Learning*
MLP	*Multilayer Perceptron*
NPHI	*Neutron Porosity*
OOB	*Out-Of-Bag*
PEFZ	*Photoelectric Factor*
PNN	*Probabilistic Neural Network*
PSO	*Particle Swarm Optimization*
R^2	*Correlation Coefficient*
RBF	*Radial Basis Function*
RF	*Random Forest*
RFR	*Random Forest Regressors*
RHOZ	*Bulk Density*
RMSE	*Root Mean Square Error*

RXOZ	*Shallow Resistivity*
SVM	*Support Vector Machine*
SVR	*Support Vector Regression*
XGBoost	*Extreme Gradient Boosting*

References

[1] Aminzadeh F, Dasgupta SN. Chapter 6 - Reservoir Characterization. In: Aminzadeh F, Dasgupta SNBTD in PS, editors. Geophysics for Petroleum Engineers. Elsevier; 2013. p. 151–89.

[2] Lichun K, Lianhua H, Zhi Y, Songtao W. Key parameters and methods of lacustrine shale oil reservoir characterization. Acta Pet Sin. 2021;42(1):1.

[3] Ganat TAAO. Physical Properties of Reservoir Rocks BT - Fundamentals of Reservoir Rock Properties. In: Ganat TAAO, editor. Cham: Springer International Publishing; 2020. p. 1–4.

[4] Alyafei N. Fundamentals of Reservoir Rock Properties - 2nd edition. Hamad bin Khalifa University Press; 2021.

[5] Tiab D, Donaldson EC. Petrophysics: theory and practice of measuring reservoir rock and fluid transport properties. Elsevier; 2024.

[6] McPhee C, Reed J, Zubizarreta I. Core analysis: a best practice guide. Elsevier; 2015.

[7] Hertel SA, Rydzy M, Anger B, Berg S, Appel M, de Jong H. Upscaling of digital rock porosities by correlation with whole-core CT-scan histograms. Petrophysics. 2018;59(05):694–702.

[8] Darling T. Well logging and formation evaluation. Elsevier; 2005.

[9] Ghosh S. A review of basic well log interpretation techniques in highly deviated wells. J Pet Explor Prod Technol. 2022;12(7):1889–906.

[10] Thakur PD, Agnihotri P, Deng L, Soliman AM, Kieduppatum P, Fernandes W. The most common impacts of drilling dynamics and environments on log-while-drilling data: A study from Abu dhabi. In: Abu Dhabi International Petroleum Exhibition and Conference. SPE; 2018. p. D021S048R004.

[11] Akinsete OO, Adekoya DA. Effects of mud filtrate invasion on well log measurements. In: SPE Nigeria Annual International Conference and Exhibition. SPE; 2016. p. SPE-184308.

[12] Bloch S. Empirical prediction of porosity and permeability in sandstones (1). Am Assoc Pet Geol Bull. 1991;75(7):1145–60.

[13] Timur A. An investigation of permeability, porosity, & residual water saturation relationships for sandstone reservoirs. Log Anal. 1968;9(04).

[14] Kolodzie Jr S. Analysis of pore throat size and use of the Waxman-Smits equation to determine OOIP in Spindle Field, Colorado. In: SPE Annual Technical Conference and Exhibition? SPE; 1980. p. SPE-9382.

[15] Ajdukiewicz JM, Lander RH. Sandstone reservoir quality prediction: The state of the art. Am Assoc Pet Geol Bull. 2010;94(8):1083–91.

[16] Nait Amar M, Ghriga MA, Hemmati-Sarapardeh A. Application of gene expression programming for predicting density of binary and ternary mixtures of ionic liquids and molecular solvents. J Taiwan Inst Chem Eng. 2020;117:63–74.

[17] Menad NA, Noureddine Z, Hemmati-Sarapardeh A, Shamshirband S, Mosavi A, Chau K wing. Modeling temperature dependency of oil-water relative permeability in thermal enhanced oil recovery processes using group method of data handling and gene expression programming. Eng Appl Comput Fluid Mech. 2019;13(1):724–43.

[18] Amar MN, Zeraibi N, Redouane K. Optimization of WAG process using dynamic proxy, genetic algorithm and ant colony optimization. Arab J Sci Eng. 2018;43(11):6399–412.

[19] Amar MN, Zeraibi N, Redouane K. Pure CO2-Oil System Minimum Miscibility Pressure Prediction Using Optimized Artificial Neural Network by Differential Evolution. Pet Coal. 2018;60(2):284–93.

[20] Al Khalifah H, Glover PWJ, Lorinczi P. Permeability prediction and diagenesis in tight carbonates using machine learning techniques. Mar Pet Geol. 2020;112:104096.

[21] Erofeev A, Orlov D, Ryzhov A, Koroteev D. Prediction of porosity and permeability alteration based on machine learning algorithms. Transp Porous Media. 2019;128:677–700.

[22] Arinkoola AO, Lawal AB, Fatola JO, Alagbe SO, Akinwole IO, Fajimi LI. Development of Empirical NMR Log-Derived Permeability Correlations Using Gaussian Process Regression and Robust Support Vector Machine Technique. Pet Coal. 2023;65(2):439–48.

[23] Alloush RM, Elkatatny SM, Mahmoud MA, Moussa TM, Ali AZ, Abdulraheem A. Estimation of geomechanical failure parameters from well logs using artificial intelligence techniques. In: SPE Kuwait Oil and Gas Show and Conference. SPE; 2017. p. D031S010R002.

[24] Zhang Z, Zhang H, Li J, Cai Z. Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: An integrated approach. J Nat Gas Sci Eng. 2021;86:103743.
[25] Ahmadi MA, Chen Z. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs. Petroleum. 2019;5(3):271–84.
[26] Okon AN, Adewole SE, Uguma EM. Artificial neural network model for reservoir petrophysical properties: porosity, permeability and water saturation prediction. Model Earth Syst Environ. 2021;7(4):2373–90.
[27] Gamal H, Elkatatny S, Alsaihati A, Abdulraheem A. Intelligent prediction for rock porosity while drilling complex lithology in real time. Comput Intell Neurosci. 2021;2021:1–12.
[28] Abdulaziz AM, Mahdi HA, Sayyouh MH. Prediction of reservoir quality using well logs and seismic attributes analysis with an artificial neural network: A case study from Farrud Reservoir, Al-Ghani Field, Libya. J Appl Geophys. 2019;161:239–54.
[29] Chen W, Yang L, Zha B, Zhang M, Chen Y. Deep learning reservoir porosity prediction based on multilayer long short-term memory network. Geophysics. 2020;85(4):WA213–25.
[30] Ifrene G, Irofti D, Ni R, Egenhoff S, Pothana P. New Insights into Fracture Porosity Estimations Using Machine Learning and Advanced Logging Tools. Fuels. 2023;4(3):333–53.
[31] Alatefi S, Abdel Azim R, Alkouh A, Hamada G. Integration of multiple bayesian optimized machine learning techniques and conventional well logs for accurate prediction of porosity in carbonate reservoirs. Processes. 2023;11(5):1339.
[32] Moosavi N, Bagheri M, Nabi-Bidhendi M, Heidari R. Porosity prediction using Fuzzy SVR and FCM SVR from well logs of an oil field in south of Iran. Acta Geophys. 2023;71(2):769–82.
[33] Ardebili PN, Jozanikohan G, Moradzadeh A. Estimation of porosity and volume of shale using artificial intelligence, case study of Kashafrud Gas Reservoir, NE Iran. J Pet Explor Prod Technol. 2024;14(2):477–94.
[34] Wolpert DH. Stacked generalization. Neural networks. 1992;5(2):241–59.
[35] Yang X, Zhang Y, Zhou D, Ji Y, Song X, Li D, et al. Drilling conditions classification based on improved stacking ensemble learning. Energies. 2023;16(15):5747.
[36] Gu J, Liu S, Zhou Z, Chalov SR, Zhuang Q. A stacking ensemble learning model for monthly rainfall prediction in the Taihu Basin, China. Water. 2022;14(3):492.
[37] Wu T, Zhang W, Jiao X, Guo W, Hamoud YA. Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. Comput Electron Agric. 2021;184:106039.
[38] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016. p. 785–94.
[39] Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. Artif Intell Rev. 2021;54:1937–67.
[40] El Mrabet Z, Sugunaraj N, Ranganathan P, Abhyankar S. Random forest regressor-based approach for detecting fault location and duration in power systems. Sensors. 2022;22(2):458.
[41] Mohandoss DP, Shi Y, Suo K. Outlier prediction using random forest classifier. In: 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC). IEEE; 2021. p. 27–33.
[42] Karimi M, Vaferi B, Hosseini SH, Olazar M, Rashidi S. Smart computing approach for design and scale-up of conical spouted beds with open-sided draft tubes. Particuology. 2021;55:179–90.
[43] Ng CSW, Djema H, Amar MN, Ghahfarokhi AJ. Modeling interfacial tension of the hydrogen-brine system using robust machine learning techniques: Implication for underground hydrogen storage. Int J Hydrogen Energy. 2022;47(93):39595–605.
[44] Youcefi MR, Hadjadj A, Bentriou A, Boukredera FS. Real-Time Prediction of Plastic Viscosity and Apparent Viscosity for Oil-Based Drilling Fluids Using a Committee Machine with Intelligent Systems. Arab J Sci Eng. 2021;1–14.
[45] Mepaiyeda EB, Oluwayomi IA, Oladipupo AO, Odutola TO. Prediction of Gas Hydrate Formation Temperature in Pipelines using Artificial Neural Network (ANN) and Firefly Algorithm (FA). Pet Coal. 2023;65(3):824–35.
[46] Heidari AA, Faris H, Mirjalili S, Aljarah I, Mafarja M. Ant lion optimizer: theory, literature review, and application in multi-layer perceptron neural networks. Nature-Inspired Optim Theor Lit Rev Appl. 2020;23–46.

[47] Wang X, Lu H, Wei X, Wei G, Behbahani SS, Iseley T. Application of artificial neural network in tunnel engineering: A systematic review. IEEE Access. 2020;8:119527–43.
[48] Doborjeh Z, Hemmington N, Doborjeh M, Kasabov N. Artificial intelligence: a systematic review of methods and applications in hospitality and tourism. Int J Contemp Hosp Manag. 2022;34(3):1154–76.
[49] Ben Seghier MEA, Kechtegar B, Amar MN, Correia JAFO, Trung NT. Simulation of the ultimate conditions of fibre-reinforced polymer confined concrete using hybrid intelligence models. Eng Fail Anal. 2021;128:105605.
[50] Vapnik V. The nature of statistical learning theory. Springer science & business media; 2013.
[51] Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019. p. 2623–31.
[52] Turner R, Eriksson D, McCourt M, Kiili J, Laaksonen E, Xu Z, et al. Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020. In: NeurIPS 2020 Competition and Demonstration Track. PMLR; 2021. p. 3–26.
[53] Hinne M, Gronau QF, van den Bergh D, Wagenmakers EJ. A conceptual introduction to Bayesian model averaging. Adv Methods Pract Psychol Sci. 2020;3(2):200–15.
[54] Rehamnia I, Benlaoukli B, Jamei M, Karbasi M, Malik A. Simulation of seepage flow through embankment dam by using a novel extended Kalman filter based neural network paradigm: Case study of Fontaine Gazelles Dam, Algeria. Measurement. 2021;176:109219.

To whom correspondence should be addressed: Dr. Mohamed Riad Youcefi, Department of Process Engineering, Faculty of Technology, University Amar Telidji-Laghouat, BP37G 03000 Laghouat, Algeria et Laboratory of Petroleum Equipment's Reliability and Materials, Faculty of Hydrocarbons and Chemistry, University M'hamed Bougara of Boumerdes, Avenue de l'Indépendance, 35000 Boumerdes, Algeria;
E-mail: mr.youcefi@lagh-univ.dz; m.youcefi@univ-boumerdes.dz