ORIGINAL PAPER-EXPLORATION GEOLOGY

Robust fracture intensity estimation from petrophysical logs and mud loss data: a multi-level ensemble modeling approach

Ahmad Azadivash[1] · Hosseinali Soleymani[1] · Atrina Seifirad[1] · Amirali Sandani[2] · Farshid Yahyaee[1] · Ali Kadkhodaie[3]

Abstract
This study presents a pioneering machine learning approach to continuously model fracture intensity in hydrocarbon reservoirs using solely conventional well logs and mud loss data. While machine learning has previously been applied to predict discrete fracture properties, this is among the first attempts to leverage well logs for continuous fracture intensity modeling leveraging advanced ensemble techniques. A multi-level stacked ensemble methodology systematically combines the strengths of diverse algorithms like gradient boosting, random forest and XGBoost through a tiered approach, enhancing predictive performance beyond individual models. Nine base machine learning algorithms generate initial fracture intensity predictions which are combined through linear regression meta-models and further stacked using ridge regression into an integrated super-learner model. This approach achieves significant improvements over individual base models, with the super-learner attaining a mean absolute error of 0.083 and R^2 of 0.980 on test data. By quantifying the crucial fracture intensity parameter continuously as a function of depth, this data-driven methodology enables more accurate reservoir characterization compared to traditional methods. The ability to forecast fracture intensity solely from conventional well logs opens new opportunities for rapid, low-cost quantification of this parameter along new wells without requiring advanced logging tools. When incorporated into reservoir simulators, these machine learning fracture intensity models can help optimize production strategies and recovery management. This systematic stacked ensemble framework advances continuous fracture intensity modeling exclusively from well logs, overcoming limitations of prior techniques. Novel insights gained via rigorous model evaluation deepen the understanding of naturally fractured reservoirs.

Keywords Fracture intensity · Machine learning · Ensemble modeling · Well log · Kopeh Dagh

Abbreviations

ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
BN	Bayesian belief network
CMIS	Committee machine intelligent system
CTODC	Critical crack tip opening displacement
FVDC	Fracture volume per unit depth of core
GA	Genetic algorithm
KICS	Critical stress intensity factor
LSSVM	Least squares support vector machine
MAE	Mean absolute error
MELM	Multiple extreme learning machines
MLP	Multilayer perceptron
MSE	Mean squared error
R^2	Coefficient of determination
RBF	Radial basis function
RF	Random forest
rmse	Root mean squared error
svm	Support vector machine

List of symbols

n	Total number of data points (–)
Var(ytrue)	Variance of actual fracture intensity (m^4/m^6)
ypred	Predicted fracture intensity (m^2/m^3)
ytrue	Actual fracture intensity (m^2/m^3)

✉ Ali Kadkhodaie
Kadkhodaie.ali@gmail.com

1 Department of Petroleum Engineering, Amirkabir University of Technology, Tehran, Iran

2 Department of Petroleum and Chemical Engineering, Islamic Azad University Science and Research Branch, Tehran, Iran

3 Earth Science Department, Faculty of Natural Science, University of Tabriz, Tabriz, Iran

Introduction

For over five decades, the petroleum engineering community has grappled with fluid flow challenges through fractured reservoirs, characterized by subterranean labyrinths of cracks and fissures. These fractures hold the key to unlocking the full potential of oil and gas extraction, yet their inherent complexities have defied precise quantification. Recent groundbreaking advancements in machine learning offer a promising solution to extract hidden insights from conventional well logs and drilling data (Cao and Sharma 2023; Hawez et al. 2021; Kim and Durlofsky 2023).

The establishment of correlations between well-log data and fracture properties has ignited a competition to refine predictive models, heralding a new era for petroleum engineers. As machine learning algorithms are fine-tuned, they provide a powerful tool to peer into the enigmatic realms of fractured reservoirs. This paper explores the role of machine learning in illuminating the dynamics of fractures, enhancing reservoir modeling, and meeting the global energy demands more efficiently.

Fracture intensity, defined as the total fracture surface area per bulk reservoir volume (m^2/m^3), stands as a pivotal parameter in this pursuit (Zhan et al. 2017; Wang et al. 2023). It quantifies the degree of fracturing within a reservoir, with higher values indicating more extensive fracture networks and enhanced production potential. Accurate quantification of fracture intensity is crucial for reservoir behavior forecasting, well optimization, and maximizing hydrocarbon recovery (Abbasi et al. 2020; Wang et al. 2023).

Fracture intensity plays a crucial role in influencing fluid flow within fractured reservoirs. The presence of denser and interconnected fractures, characterized by high intensity, establishes pathways with increased permeability for the smooth movement of hydrocarbons. Conversely, tight zones with low fracture intensity impose restrictions on fluid flow (Questiaux et al. 2010). Efficiently mapping fracture intensity allows for the targeted identification of high-intensity sweet spots, optimizing strategies for hydrocarbon recovery. Moreover, the integration of fracture intensity models into reservoir simulators not only enhances production forecasts but also contributes to the development of more effective strategies (Questiaux et al. 2010; Abbasi et al. 2020). In this context, the application of the faulted model in simulation models proves valuable, as it enables a detailed exploration of the sensitivity of fluid movements across faults and their potential impact in proximity to fault fracture zones (Al-Dujaili et al. 2023).

Traditional methods for estimating fracture intensity have limitations, such as limited wellbore coverage in image logs and low recovery rates in fractured zones during core analysis (Zaree et al. 2016; Azadivash et al. 2023a; Wang et al. 2023). Leveraging machine learning algorithms trained on continuous conventional well logs and mud loss data offers a promising solution. Conventional logs provide extensive formation evaluation coverage, making them a cost-effective data source (Darling 2005; Sen et al. 2021; Azadivash et al. 2023b). Machine learning models can correlate log trends with fracture intensity from core and images, allowing continuous intensity prediction along boreholes and between wells. The advantages include ubiquitous log data utilization, independent estimation, and cost-effectiveness, although challenges in data volume and natural fracture variability remain (Ouenes 2000; Boerner et al. 2003).

Accurate quantification of fracture intensity through machine learning facilitates optimized reservoir development and production strategies, harnessing high-intensity areas efficiently. Integrating machine-learning fracture models into reservoir simulators enables scenario forecasting to optimize hydrocarbon recovery (He et al. 2020; Fathi et al. 2022; Ng et al. 2023). However, addressing the demands for training data volume and natural fracture heterogeneity complexities remains an ongoing research challenge. With continued progress, machine learning promises a low-cost, continuous solution for optimizing performance in naturally fractured reservoirs through precise fracture intensity characterization.

Recent studies have demonstrated the potential of machine learning techniques to estimate discrete fracture properties from well logs. These studies have made significant strides in fracture prediction, utilizing various machine learning algorithms and datasets (Boadu 1998; Ince 2004; Sarkheil et al. 2009; Ja'fari et al. 2012; Zazoun 2013; Nouri-Taleghani et al. 2015; Li et al. 2018; Bhattacharya and Mishra 2018; Rajabi et al. 2021; Tabasi et al. 2022; Pei and Zhang 2022; Delavar 2022; Gao et al. 2023).

While machine learning has been successfully applied to predict discrete fracture properties, a critical gap remains in leveraging these techniques to model the continuous parameter of fracture intensity. Conventional well logs and mud loss data offer abundant and low-cost subsurface information. However, developing advanced machine learning models to correlate this data with continuous fracture intensity measurements has been an unmet need. This study addresses this knowledge gap by introducing a pioneering multi-tiered stacked ensemble methodology that systematically combines diverse algorithms like gradient boosting, random forest, and XGBoost to construct a precise predictive model for continuous fracture intensity solely from conventional well logs and mud loss data. The ability to forecast this crucial parameter continuously as a function of depth using readily available data opens new opportunities for rapid, low-cost fracture

intensity quantification along new wells without requiring advanced logging tools. When incorporated into reservoir simulators, these data-driven fracture intensity models can enable more accurate reservoir characterization and optimization of production strategies compared to traditional methods.

The methodology employs a diverse set of algorithms, including Gradient Boosting, Random Forest, Extra Trees, XGBoost, LightGBM, CatBoost, Decision Trees, Bagging, and Histogram-based Gradient Boosting. These algorithms provide diverse perspectives to capture intricate data patterns related to fracture intensity. The predictions from these base learners are ensembled using linear regression to create two meta-learner models. Finally, ridge regression is utilized to stack the outputs of the two meta-learners, resulting in an integrated super-learner model. This multi-tiered approach capitalizes on the unique strengths of each algorithm, enhancing accuracy, diversity, and robustness in fracture intensity prediction.

Geological setting

The Kopeh Dagh Basin, located in northeastern Iran, is an integral part of the Amu Darya Basin, extending southeastward into Turkmenistan and Uzbekistan (Kavoosi et al. 2009; Robert et al. 2014). Spanning over 300 km from the Turkmenistan border to the Mashhad area, it is bounded to the north by the Kopeh Dagh mountain range, formed due to the convergence of the Eurasian and Iranian plates (Taghizadeh-Farahmand et al. 2013; Ruh et al. 2019).

Sharing a basement of deformed Paleozoic rocks from the Hercynian accreted terrane, the Kopeh Dagh Basin hosts significant gas reserves in Upper Jurassic carbonates (Mozduran Formation) and Lower Cretaceous sandstones (Shurijeh Formation), particularly in the Khangiran Field (Robert et al. 2014). The closure of the Neotethys Ocean during the Late Cretaceous led to uplift and folding, followed by episodes of shortening and thickening during the Paleocene and Eocene (Brunet et al. 2003; Zanchi et al. 2006). Figure 1 provides an overview of the topography of the Kopeh Dagh range and the critical gas field locations.

Major basin development occurred during the Oligocene and Miocene epochs, driven by the collision of the Iranian and Eurasian plates, resulting in flexural subsidence and the accumulation of crucial hydrocarbon source rocks and reservoirs (Lyberis and Manby 1999; Golonka 2004; Robert et al. 2014). With sediment thickness exceeding 12 km in the depocenter, the basin offers substantial hydrocarbon potential (Golonka 2004; Robert et al. 2014).

The intricate folding and thrust faulting within the Kopeh Dagh Basin create diverse traps for hydrocarbons, making it a world-class petroleum province in the Amu Darya Basin region (Arian 2012; Nouri and Arian 2017). Ongoing exploration efforts are expected to yield further discoveries in this geologically significant area. Figure 2 presents a stratigraphic chart of the Kopeh Dagh belt,
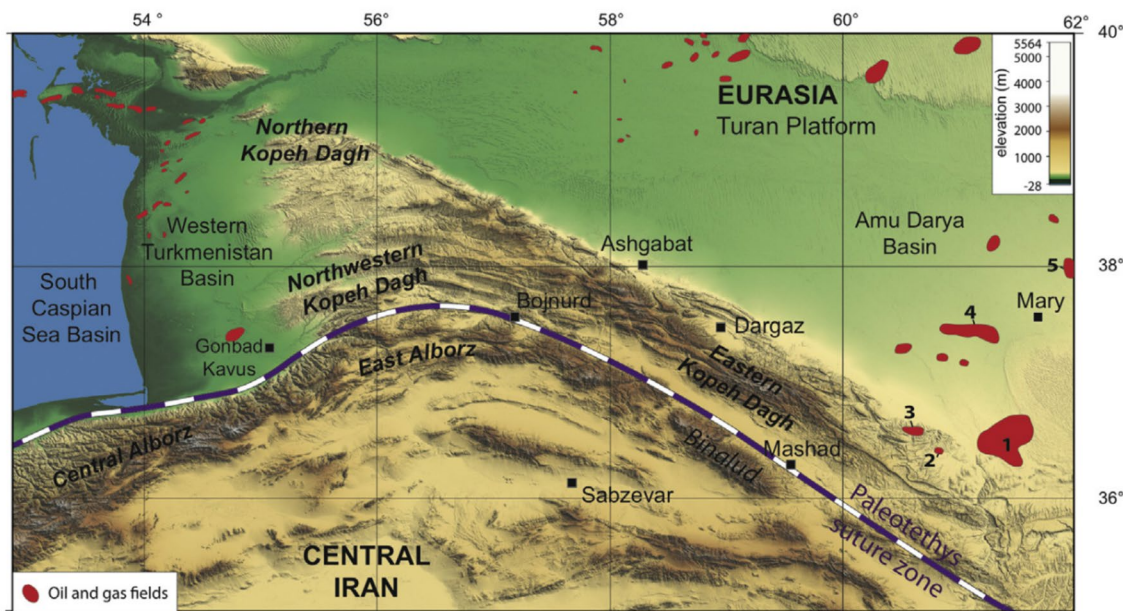

Fig. 1 A topographic chart depicting the Kopeh Dagh mountain range with marked positions of six prominent gas fields (labeled as follows: 1—Dauletabad, 2—Gonbadli, 3—Khangiran, 4—Shaltyk, 5—Bayram-Ali, 6—Achak) is displayed in red close to the Paleotethys suture zone. Image from Robert et al. (2014)
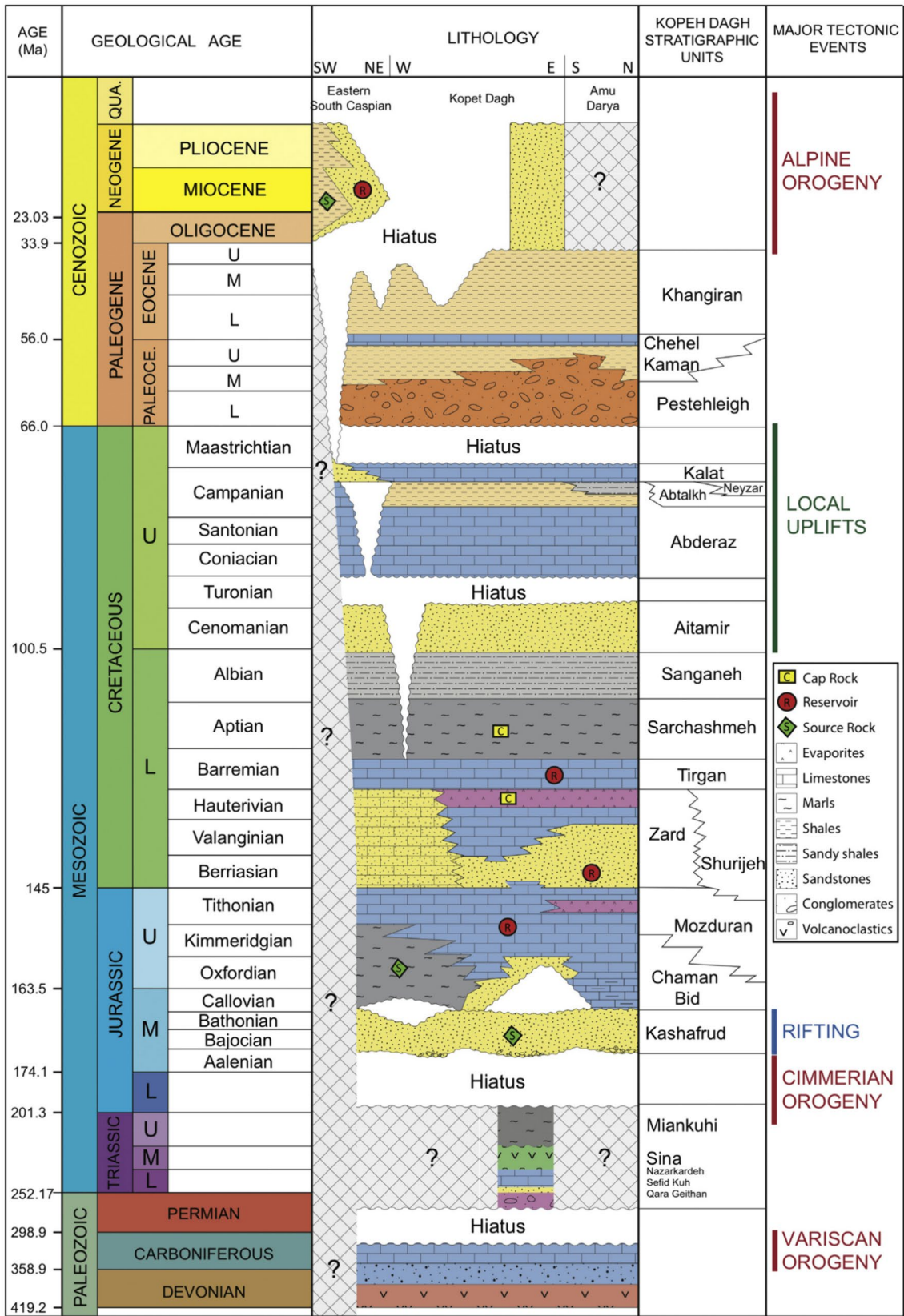
Fig. 2 Stratigraphic chart of the Kopeh Dagh belt, Amu Darya and South Caspian Sea basins showing units, major unconformities, and correlation with tectonic events. Image from Robert et al. (2014)

Amu Darya, and South Caspian, illustrating correlations between units and tectonic events.

Material and methods

Material

This study employed well log and mud loss data from three wells labeled as Wells A, B, and C, situated in the Kopeh Dagh basin. The dataset encompassed a wide range of log measurements, including bulk density (RHOB, kg/m^3), photoelectric factor (PEF, b/e), neutron porosity (NPHI, v/v), lost circulation (MUDLOSS, bph), shallow resistivity (LLS, ohm.m), deep resistivity (LLD, ohm.m), spectral gamma-ray (SGR, gAPI), acoustic travel time (DT, us/m), computed gamma-ray (CGR, gAPI), caliper (CALI, mm), bad hole (BH, mm), depth (DEPTH, m), and fracture intensity (INTENSITY, m^2/m^3) derived from electrical image logs.

The dataset comprised 7590 data points, with contributions from Well A (2674 data points), Well B (2367 data points), and Well C (2549 data points). Fracture intensity interpretation was based on electrical image logs, while lost circulation data were extracted from drilling records. Table 1 provides a statistical summary of the dataset, including metrics such as sample count, average, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum values.

To ensure data integrity, rigorous quality control and preprocessing measures were implemented. This included identifying and removing erroneous data, outliers, and missing log values to prevent potential bias in the subsequent modeling.

The log data underwent standardization, involving mean centering and scaling to unit variance, facilitating direct comparisons among log properties with varying units and magnitudes. This standardization also addressed numerical stability issues and expedited model convergence. Log depths were resampled to a consistent step size using spline interpolation to ensure uniformity in depth measurements for model compatibility. The resulting dataset, characterized by standardized, cleaned, aligned, and resampled logs, served as input for machine learning modeling.

Figure 3 presents a cross-correlation matrix plot of the 12 input parameters and fracture intensity. This figure reveals distinct relationships between fracture intensity and various logging measurements. Depth showed a positive correlation due to increased overburden pressure with depth, leading to more intense natural fracturing in deeper formations. Lost circulation directly indicated fractured zones. Lower bulk density (RHOB) and higher neutron porosity (NPHI) readings suggested increased porosity resulting from fracturing. Longer acoustic travel times (DT) indicated fractures disrupting acoustic wave propagation. Decreased gamma ray (GR) and deep resistivity (LLD) corresponded to fractured carbonate lithology and saturated vertical fractures, respectively. Irregular borehole diameter measured by caliper (CALI) occurred more often in fractured intervals (Pei and Zhang 2022).

While some parameters displayed stronger individual correlations with fracture intensity, all 12 well-log inputs were utilized for machine learning prediction modeling. This comprehensive approach accounted for potential interdependencies between variables that may not be evident through direct correlation analysis. Weaker individual correlations did not discount their potential contributions to prediction through complex nonlinear multivariate interactions. Employing the full suite of diverse well-log measurements allowed machine learning models to determine optimal input combinations for maximizing prediction accuracy. This holistic modeling approach could uncover subtle relationships between

Table 1 Statistical summary of the dataset

| Data	Count	Mean	Std	Min	25%	50%	75%	Max
DEPTH (m)	7590	3178.87	333.08	2682.39	2920.44	3121.23	3382.63	3953.10
BH (in)	7590	22.31	98.77	−111.44	−94.03	81.39	104.33	268.34
CALI (in)	7590	286.97	65.94	199.71	217.03	315.60	322.65	579.49
CGR (gAPI)	7590	34.62	33.67	0.24	4.48	24.69	61.31	138.28
DT (us/m)	7590	190.32	22.74	145.67	170.78	188.78	206.68	365.16
SGR (gAPI)	7590	43.65	36.84	1.77	10.84	34.65	72.46	149.99
LLD (ohm.m)	7590	373.83	699.55	3.28	24.52	66.30	414.88	11,797.35
LLS (ohm.m)	7590	245.26	503.97	3.65	21.61	44.83	142.06	1950.00
MUDLOSS (bph)	7590	1.12	6.21	0.00	0.00	0.00	0.00	50.00
NPHI (v/v)	7590	0.08	0.07	−0.02	0.03	0.07	0.12	0.31
PEF (b/e)	7590	4.65	1.63	2.35	3.58	4.11	5.20	10.00
RHOB (kg/m3)	7590	2667.64	123.90	1951.90	2613.72	2706.95	2748.88	2976.23
INTENSITY (m2/m3)	7590	0.80	1.54	0.00	0.00	0.00	1.02	12.56
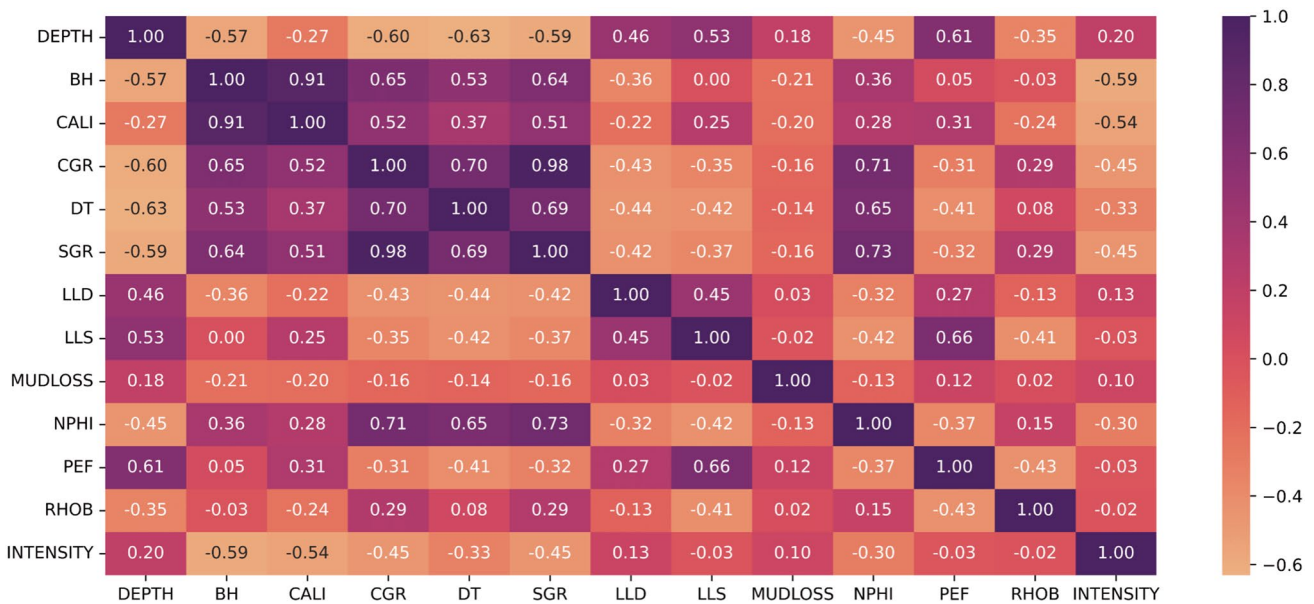
Fig. 3 Cross-correlation matrix plot of 12 input parameters and fracture intensity

various parameters and fractures, even those not detectable through standalone correlations.

Methods

The methodology for predicting fracture intensity strategically employs various machine learning techniques, offering a comprehensive and systematic approach. Nine base learners, including Gradient Boosting, Random Forest, Extra Trees, Xgboost, LightGBM, CatBoost, Decision Trees, Bagging, and Histogram-based Gradient Boosting models, generate initial predictions. These diverse models bring unique strengths to the table, from capturing intricate data relationships to handling large datasets and categorical features efficiently.

In the second stage, linear regression amalgamates the first-level predictions, forming two meta-learners designed to capture nuanced data patterns. The third level employs ridge regression to stack the outputs of these meta-learners, creating a unified super-learner model. This rigorous selection process ensures the ensemble model's adaptability across various fracture intensity scenarios.

The methodology leverages the interactions and synergies among different methods, enhancing overall accuracy, diversity, and robustness. Figure 4 illustrates the systematic approach employed, showcasing the orchestration of diverse machine learning techniques to predict fracture intensity.

Base algorithms employed

This section covers various machine learning techniques employed in this research.

Gradient Boosting is an ensemble technique that combines weak learners sequentially to reduce errors, utilizing gradient descent optimization and regularization to prevent overfitting (Dorogush et al. 2018; Bentéjac et al. 2021). Gradient Boosting is favored for its ability to model complex data patterns and flexibility in hyperparameter tuning, making it a robust approach for regression, classification, and ranking problems (Natekin and Knoll 2013; Ayyadevara 2018).

Random Forest leverages multiple decorrelated decision trees trained on random data samples to improve generalizability and avoid overfitting. Tuning hyperparameters like number of trees allows optimizing performance (Breiman 2001; Speiser et al. 2019). Random Forest is renowned for its versatility, scalability and robustness in classification and regression tasks (Smith et al. 2013; Rodriguez-Galiano et al. 2015).

Extra Trees introduces additional randomness in the training process as a defense against overfitting, making it well-suited for noisy, high-dimensional data (Geurts et al. 2006; Ahmad et al. 2018). The combination of ensemble learning and extreme randomness yields a reliable model for challenging datasets (Goetz et al. 2014).

XGBoost utilizes regularization, tree pruning and custom loss functions for optimal performance, renowned for its speed, accuracy and flexibility via hyperparameter
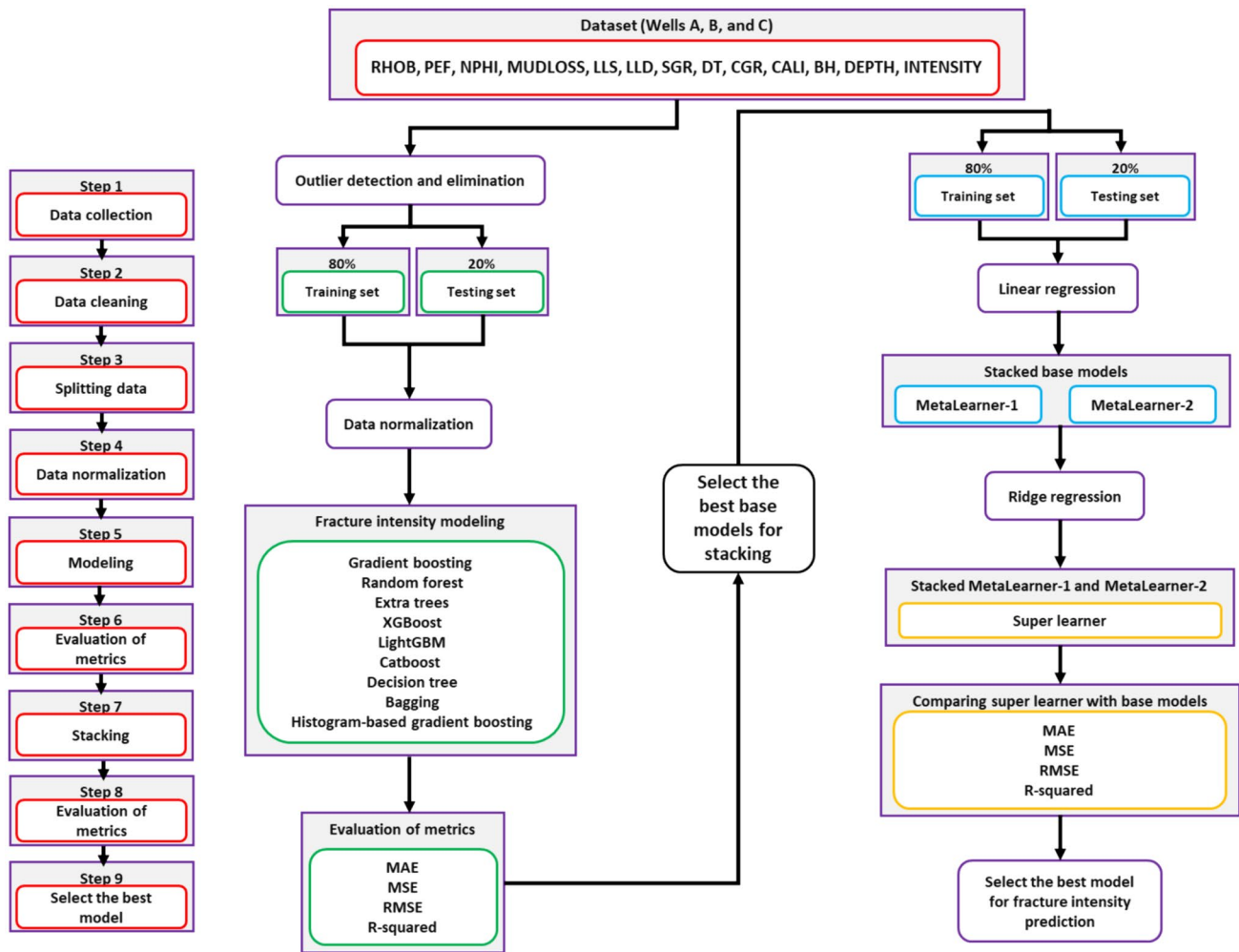
Fig. 4 Workflow of the present study

tuning (Chen et al. 2015). It is popular across academia and industry for diverse regression, classification and ranking problems (Hastie et al. 2009).

LightGBM utilizes a histogram-based algorithm for fast, low memory tree construction, enabling high performance for large datasets (Ke et al. 2017; Daoud 2019). Its computational efficiency and scalability make it a preferred gradient boosting tool.

CatBoost handles categorical features seamlessly, improving resistance to overfitting and reducing preprocessing needs (Dorogush et al. 2018; Prokhorenkova et al. 2018). Its capabilities for automatic feature encoding and predictive accuracy make CatBoost effective for modeling tabular data.

Decision trees offer intuitive hierarchical structure for classification/regression, providing transparency into feature importance (Kotsiantis 2013). They benefit from pruning and ensemble methods to improve performance and combat overfitting (Myles et al. 2004; Robnik-Šikonja 2004). Their

interpretability makes them a versatile supervised learning technique.

Bagging combines independently trained models on resampled data to improve stability and accuracy through model averaging (Breiman 1996; Bauer and Kohavi 1999; Galar et al. 2012; Lee et al. 2020). Training base models on random subsets avoids overfitting.

Histogram-based Gradient Boosting utilizes histograms to accelerate tree construction, enhancing scalability without losing accuracy (Guryanov 2019; Nhat-Duc and Van-Duc 2023). This allows gradient boosting to handle large datasets and real-time applications efficiently (Table 2).

Stacking algorithm

Stacking is an ensemble technique that combines multiple diverse base learners, with each capturing distinct data patterns and relationships. A meta-learner is then trained on the predictions of the base learners to effectively fuse their

Table 2 Overview of hyperparameters employed in base models

Model	Hyperparameters
Gradient boosting	alpha: 0.9, criterion: friedman_mse, learning_rate: 0.1, loss: squared_error, max_depth: 3, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 100, subsample: 1.0, tol: 0.0001, validation_fraction: 0.1
Random forest	bootstrap: True, criterion: squared_error, max_features: 1.0, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 100
Extra trees	criterion: squared_error, max_features: 1.0, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 100
XGBoost	objective: reg:squarederror, missing: nan, n_estimators: 100
LightGBM	boosting_type: gbdt, colsample_bytree: 1.0, importance_type: split, learning_rate: 0.1, max_depth: -1, min_child_samples: 20, min_child_weight: 0.001, n_estimators: 100, num_leaves: 31, subsample: 1.0, subsample_for_bin: 200,000
CatBoost	loss_function: RMSE
Decision tree	criterion: squared_error, min_samples_leaf: 1, min_samples_split: 2, splitter: best
Bagging	base_estimator: deprecated, bootstrap: True, max_features: 1.0, max_samples: 1.0, n_estimators: 10
Histogram-based Gradient boosting	early_stopping: auto, learning_rate: 0.1, loss: squared_error, max_bins: 255, max_iter: 100, max_leaf_nodes: 31, min_samples_leaf: 20, n_iter_no_change: 10, scoring: loss, tol: 1e-07, validation_fraction: 0.1

collective knowledge (Džeroski and Ženko 2004; Sikora and Al-Laymoun 2015). By learning to optimize the blend of diverse predictions, the meta-learner acts as a supervisor over the ensemble. This allows stacking to improve accuracy and robustness compared to individual models, leveraging the wisdom of the crowd (Džeroski and Ženko 2004; Goliatt et al. 2023).

A key advantage of stacking is its versatility in accommodating different base learner types like decision trees, neural networks, etc. This makes it adaptable to diverse problem domains with complex data where singular models may falter (Goliatt et al. 2023). Researchers have explored numerous extensions like incorporating feature engineering and advanced meta-learner optimization to further enhance utility (Džeroski and Ženko 2004; Rahman et al. 2021). With its potential to significantly boost predictive performance through model collaboration, stacking is a pivotal technique for pushing the frontiers of machine learning.

Results

This section presents the results of fracture intensity prediction, employing a range of machine learning models. The log data underwent division into training and testing datasets, facilitating an assessment of model performance. Furthermore, an evaluation of the models was conducted utilizing diverse metrics, complemented by generating pertinent plots to acquire insights into their predictive capabilities.

Data partitioning

To enable reliable evaluation of model performance, the full dataset was split into an 80% training set and 20% testing set. Specifically, 6072 data points were used to train the base learners. The remaining 1,518 data points were held out as an independent test set to assess generalization capability. This 80–20 split was consistently applied when evaluating the various models.

The base learners were trained on 80% of the full data and tested on the 20% holdout set. The meta-learners were trained on the base learner outputs and evaluated on the same holdout set. Finally, the super-learner was trained on the meta-learner outputs and tested on the holdout set.

The rationale for the 80–20 split is two-fold. First, the large training set enables robust model fitting. Second, the 20% test set includes varied data not seen during training, allowing true assessment of generalization performance. Moreover, using one fixed test set for all models enables fair comparison of performance between the different models.

Model evaluation metrics

A collection of widely recognized evaluation metrics was employed to gauge the performance of each machine learning model in predicting fracture intensity. These metrics act as quantitative measures of model effectiveness, allowing for the evaluation of accuracy and reliability. The equations utilized in this study are presented as Eqs. (1, 2, 3, 4) in the upcoming section:

The Mean Absolute Error (MAE) is determined by computing the average absolute difference between the predicted ($ypred$) and actual ($ytrue$) fracture intensity values across all data points:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |ytrue - ypred| \tag{1}$$

where:

n is the total number of data points.

The Mean Squared Error (MSE) is calculated as the average of the squared differences between the predicted and actual fracture intensity values:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (ytrue - ypred)^2 \qquad (2)$$

The Root Mean Squared Error (RMSE) is derived as the square root of the MSE and offers a measure of the average error in the same units as the target variable:

$$RMSE = \sqrt{MSE} \qquad (3)$$

The Coefficient of Determination (R^2) quantifies the proportion of variance in the target variable explained by the model. Its calculation is as follows:

$$R^2 = 1 - \frac{MSE}{Var(ytrue)} \qquad (4)$$

where:

Var(ytrue) is the variance of the actual fracture intensity values.

These metrics provide valuable insights into the overall performance of each machine-learning model. Superior predictive accuracy is indicated by lower values of MAE, MSE, and RMSE, while higher R^2 values suggest a better fit to the data. In subsequent analyses, the implications of these metrics and their importance in selecting the most suitable model for fracture intensity prediction will be explored.

Base models performance

In this section, an examination of the performance evaluation of a variety of machine learning models is undertaken. These models were individually trained to predict fracture intensity using a comprehensive set of 12 key input features: Bulk density, Photoelectric factor, Neutron porosity, Lost circulation, Shallow resistivity, Deep resistivity, Spectral gamma-ray, Acoustic travel time, Computed gamma-ray, Caliper, Badhole, and Depth. The model selection process involved nine contenders: Gradient Boosting, Random Forest, Extra Trees, XGBoost, LightGBM, CatBoost, Decision Tree, Bagging, and Histogram-based Gradient Boosting. Each model underwent thorough training and testing.

The results of the model evaluation are showcased in Table 3 and Fig. 5. Table 3 provides a concise overview of the principal performance metrics, encompassing Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2), for both the training and test datasets. Figure 5 complements this by providing a visual representation of model performance through four distinct subfigures, each illustrating a different performance metric: (a) MAE, (b) MSE, (c) RMSE, and (d) R^2. These visual aids facilitate a comprehensive comparison of the nine models across these critical metrics.

Upon close examination of these base models, two models, Random Forest and Extra Trees, are identified as clear frontrunners. Exceptional accuracy during training is exhibited by Random Forest, with a near-perfect R^2 score of 0.996, accompanied by impressively low error metrics (MAE: 0.037, MSE: 0.010, RMSE: 0.100). Crucially, robust generalization capabilities are maintained by it, as evidenced by the high accuracy on the test data, with an MAE of 0.092 and an R^2 of 0.974. On the other hand, Extra Trees achieves perfect scores on all training metrics and continues to excel on the test data with an MAE of 0.083 and an R^2 of 0.978. XGBoost follows closely behind, showcasing remarkable training and test dataset accuracy, with an R^2 of 0.998 and an MAE of 0.109. These models are regarded as optimal choices for tasks requiring precision and generalization.

Gradient Boosting, LightGBM, and CatBoost present a competitive and balanced performance profile. While their training data accuracy falls slightly short of the top performers, commendable R^2 scores (0.889, 0.985, and 0.988, respectively) and relatively low error metrics (MAE and RMSE) are still achieved. These models strike an attractive

Table 3 Performance metrics of the base models on train and test data

| Model	Train				Test			
	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2
Gradient boosting	0.256	0.268	0.518	0.889	0.263	0.291	0.539	0.871
Random forest	0.037	0.010	0.100	0.996	0.092	0.057	0.240	0.974
Extra trees	0.000	0.000	0.000	1.000	0.083	0.052	0.228	0.978
XGBoost	0.031	0.004	0.065	0.998	0.109	0.061	0.248	0.973
LightGBM	0.088	0.036	0.190	0.985	0.130	0.090	0.300	0.960
CatBoost	0.085	0.028	0.168	0.988	0.136	0.089	0.299	0.960
Decision tree	0.000	0.000	0.000	1.000	0.082	0.085	0.292	0.964
Bagging	0.041	0.015	0.121	0.991	0.100	0.067	0.260	0.966
Histogram-based gradient boosting	0.092	0.039	0.198	0.984	0.136	0.103	0.321	0.954
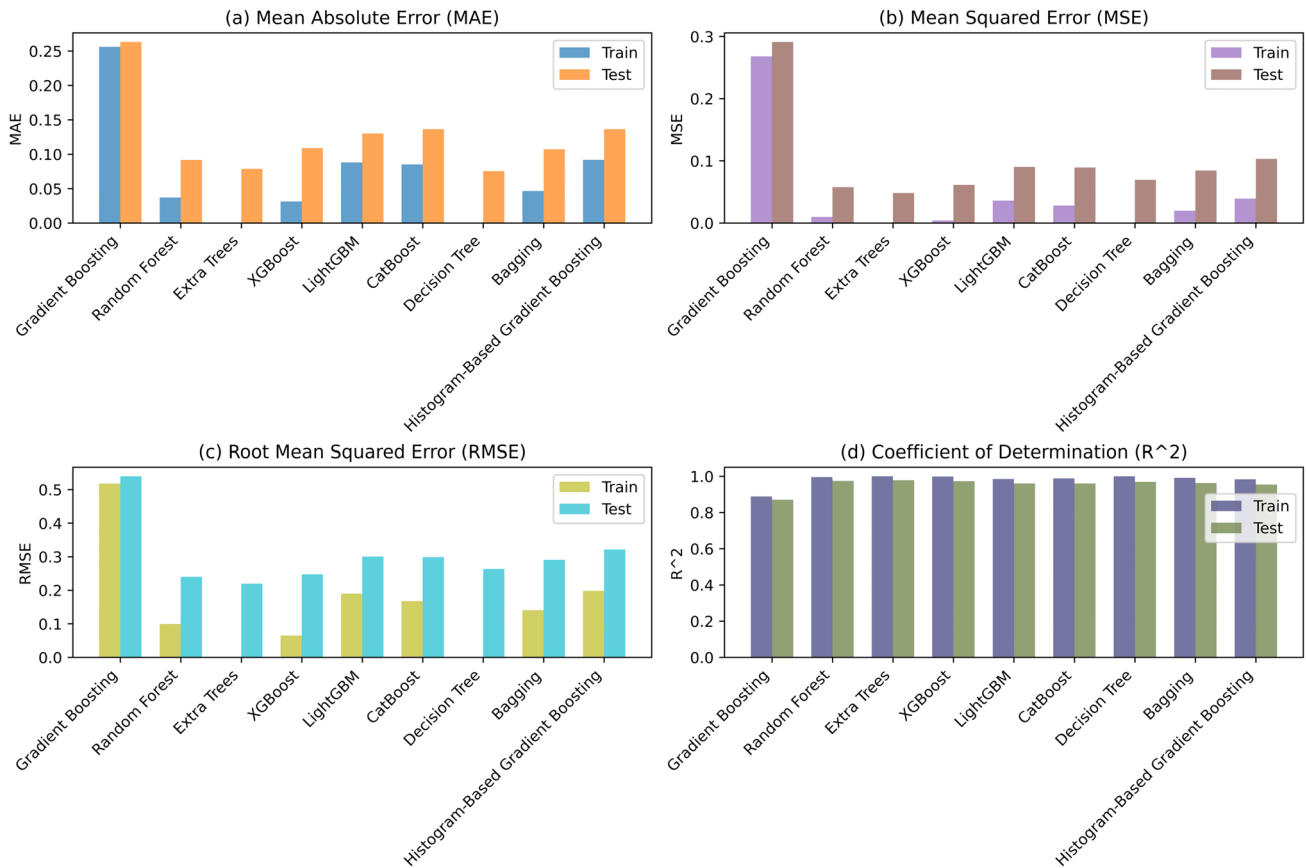
Fig. 5 A visual overview of model performance metrics **a** MAE, **b** MSE, **c** RMSE, and **d** R^2

balance between accuracy and generalization, rendering them suitable for a wide range of applications where model interpretability and generalizability are paramount. For instance, LightGBM maintains a reasonable MAE of 0.130 and an R^2 of 0.960 on the test data.

Decision Tree and Histogram-based Gradient Boosting models display substantial accuracy on the training data, indicating their ability to fit the data exceptionally well, as evidenced by their perfect training metrics. However, susceptibility to overfitting may be a concern, as reflected in their slightly lower generalization performance on the test data, such as Histogram-based Gradient Boosting's MAE of 0.136 and an R^2 of 0.954. In contrast, a robust ensemble approach is provided by Bagging, maintaining excellent generalization while achieving respectable accuracy on both training and test datasets (MAE: 0.100, R^2: 0.966 on the test data).

In fracture intensity prediction, the standout model is undeniably Extra Trees, showcasing remarkable accuracy and strong generalization ability. In sharp contrast, Gradient Boosting is the least proficient model within the assessed pool. In contrast, the remaining models deliver

commendable performance, offering dependable choices for accurate fracture intensity predictions.

Feature importance analysis

The Feature Importance plot is a potent tool for gaining profound insights into the influential input features that affect fracture intensity predictions within the ensemble of machine learning models. The feature importance plot of the base models is displayed in Fig. 6. "DEPTH" consistently emerges as a paramount factor among the prominent features, exerting substantial influence across most models. Notably, "DEPTH" is particularly emphasized by models such as XGBoost and Decision Tree, underscoring its indispensable role in accurately forecasting fracture intensity.

The results show that "DEPTH" is assigned the highest importance in most models, especially in Gradient Boosting, Random Forest, Decision Tree, and Histogram-Based Gradient Boosting. This consistent emphasis on "DEPTH" reinforces the notion that the depth of the geological formation plays a critical role as an input feature in predicting fracture intensity across various modeling techniques.
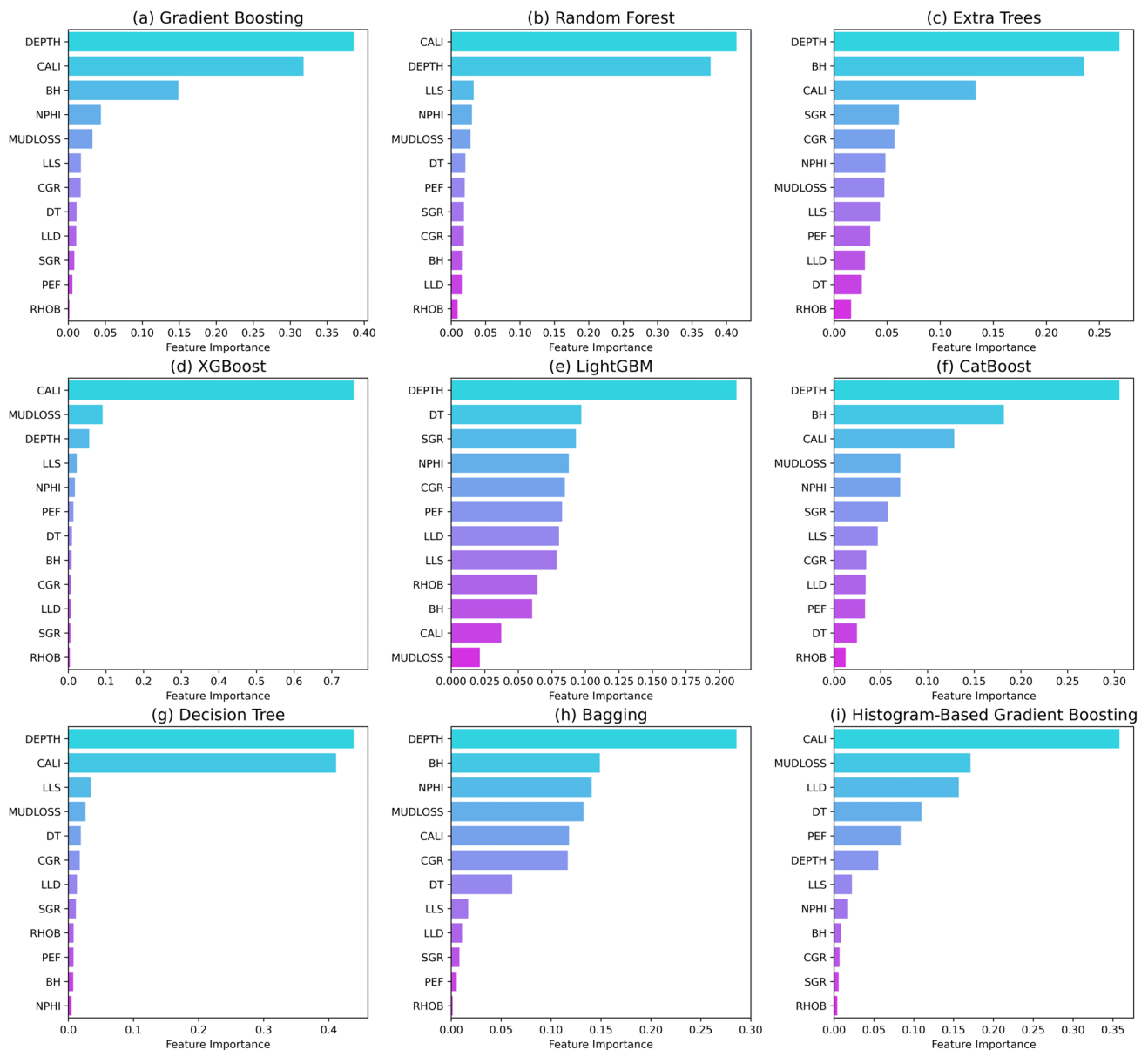
Fig. 6 Feature importance plot of the base models: **a** gradient boosting, **b** random forest, **c** extra trees, **d** XGBoost, **e** LightGBM, **f** CatBoost, **g** decision tree, **h** bagging, **i** histogram-based gradient boosting

In contrast, features such as "CGR" (Computed gamma-ray) and "DT" (Acoustic travel time) tend to exhibit relatively lower importance across the entire spectrum of models. This suggests that these features may not wield as much influence in facilitating accurate predictions as "DEPTH."

In summary, the results of the feature importance analysis highlight that "DEPTH" serves as a critical driver of fracture intensity predictions in the ensemble of machine learning models. Models like XGBoost and Decision Tree mainly rely on this feature for accurate forecasting, while features like "CGR" and "DT" play a relatively minor role in this prediction task.

Actual versus predicted analysis

The analysis of Actual vs. Predicted Performance, an essential tool for evaluating machine learning model efficacy, is applied in this study. This visualization comprises a scatter plot where the x-axis represents actual values derived from the target variable in the train and test dataset. At the same time, the y-axis illustrates the corresponding predicted values generated by each model. Figure 7 presents the Actual versus Predicted Plots for various base models; all underwent comprehensive evaluation using the coefficient of determination (R^2) metric applied to both
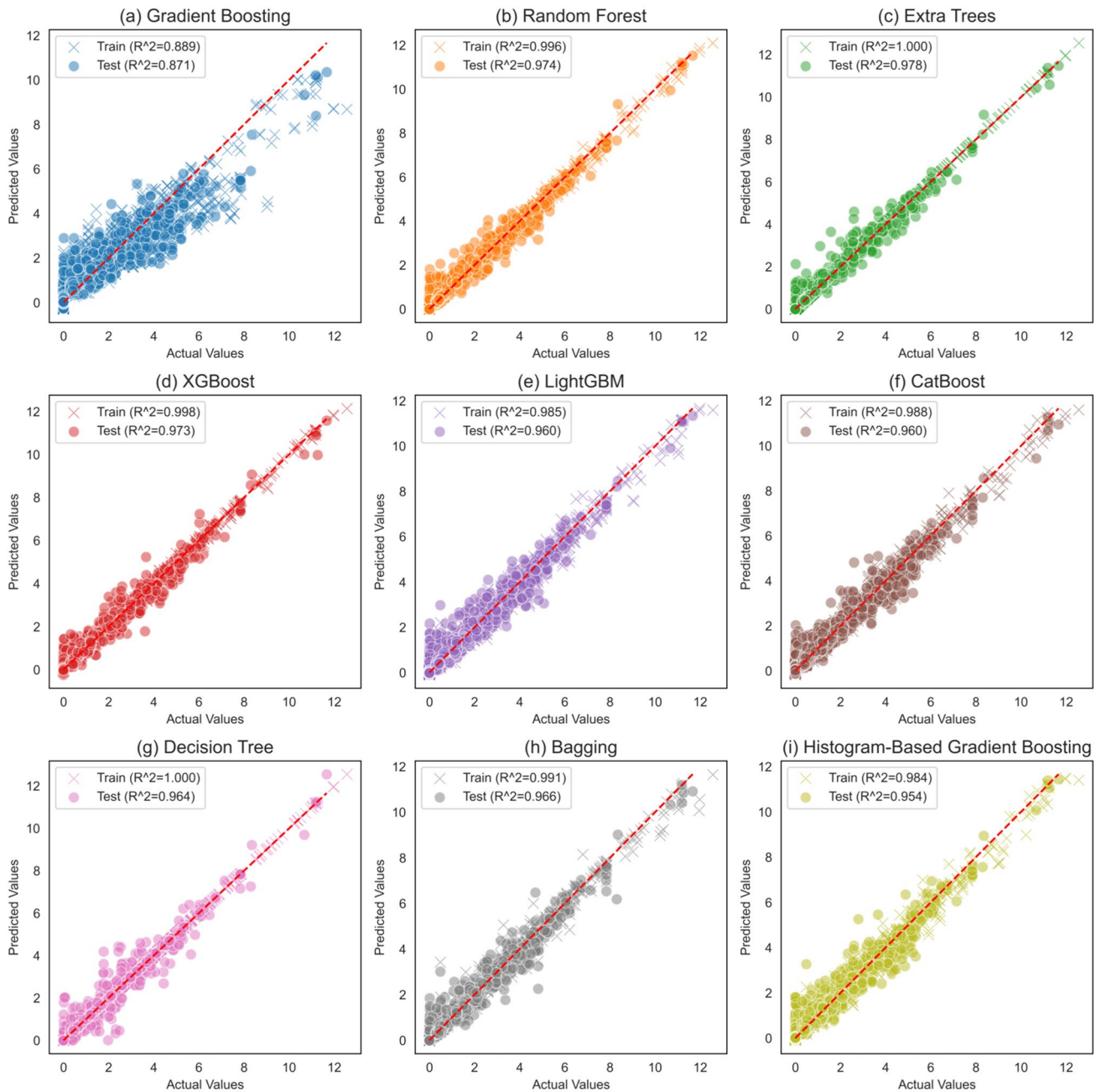
Fig. 7 A comparative analysis of actual vs. predicted plots for various base models: **a** gradient boosting, **b** random forest, **c** extra trees, **d** XGBoost, **e** LightGBM, **f** CatBoost, **g** decision tree, **h** bagging, **i** histogram-based gradient boosting

the training and test datasets to assess their predictive capabilities.

Within the plot, a diagonal line, commonly referred to as the "ideal line" or the "line of perfect predictions" $(y = x)$, illustrates a scenario in which the model's predictions align perfectly with the actual values. The proximity of data points to this line signifies precision and accuracy in predictions, whereas deviations from the line reveal disparities between the model's predictions and the actual ground truth.

The analysis revealed varying levels of model performance. Notably, the Random Forest model exhibited exceptional performance, achieving R^2 scores of 0.996 on the training data and 0.974 on the test data, demonstrating its remarkable ability to generalize. Similarly, the Gradient Boosting model showcased substantial predictive capabilities, attaining R^2 scores of 0.889 on the training data and 0.871 on the test data.

Furthermore, the Extra Trees and Decision Tree models attained flawless R^2 scores of 1.000 on the training data, indicating an exceptional fit to the training set. Extra Trees maintained a high R^2 score of 0.978 on the test data, exemplifying its robust generalization ability. In contrast, a slightly lower R^2 score of 0.964 for the Decision Tree indicated a minor reduction in predictive accuracy.

Additionally, on the test data, strong predictive capabilities were demonstrated by XGBoost, CatBoost, and Bagging, with R^2 scores of 0.973, 0.960, and 0.966, respectively. While still performing effectively, LightGBM and Histogram-based Gradient Boosting exhibited slightly more variability in their predictions, recording R^2 scores of 0.960 and 0.954 on the test data.

Residual analysis

In this research, residual plot analysis is a crucial diagnostic tool for evaluating the performance of different machine-learning models. Residuals, representing the differences between actual and predicted values, provide valuable insights into a model's ability to capture inherent data patterns. Figure 8 illustrates the residual plots for different base models, while Table 4 presents statistical data related to these residuals. The data analysis reveals distinct characteristics in terms of residuals for different models.

Upon closer examination of the mean residuals, it is evident that residuals close to zero are consistently maintained by Gradient Boosting, Random Forest, and Decision Tree models. This suggests their capacity to provide reasonably accurate predictions, with mean residual values of approximately -0.006, -0.011, and -0.005, respectively. These models demonstrate an overall solid fit to the data.

When the standard deviation of residuals is scrutinized, the highest spread at 0.539 is exhibited by Gradient Boosting, indicating potential challenges with specific data points and variability in prediction accuracy. In contrast, the lowest standard deviation values of 0.220 and 0.256 are showcased by Extra Trees and Bagging models, respectively, indicating their consistent performance across the dataset.

An analysis of the minimum and maximum residuals reveals significant differences among the models. For example, Gradient Boosting displays a minimum residual of -2.884 and a maximum residual of 2.853, occasionally leading to substantial prediction errors. In contrast, models like Bagging and Decision Tree feature a narrower range of residuals, with Bagging demonstrating a minimum residual of -1.591 and a maximum residual of 1.463. These findings highlight the potential for substantial errors in specific situations with Gradient Boosting, while other models offer more stability.

The comparison of residual characteristics among the various machine learning models allows the identification of their strengths and weaknesses. Among the models evaluated, Extra Trees and Bagging consistently emerge as the top performers in terms of residuals. Mean residuals close to zero, indicating overall solid performance, are consistently yielded by these models, and they exhibit relatively low standard deviations of residuals, signifying consistent predictions. Additionally, the minimum and maximum residuals for Extra Trees and Bagging fall within an acceptable range, suggesting a reduced likelihood of producing substantial errors compared to other models.

Conversely, Gradient Boosting stands out as the model with the most variable predictions. While it also boasts a mean residual close to zero, indicating reasonable accuracy on average, its high standard deviation of residuals (0.539) raises concerns about its ability to handle specific data points. Moreover, Gradient Boosting displays the most expansive range between minimum and maximum residuals, suggesting a propensity for significant errors in specific cases.

Based on the residual analysis and the statistical information presented in Table 4, Extra Trees and Bagging can be considered the best-performing models among those evaluated in this study. They offer consistency and relatively accurate predictions. In contrast, Gradient Boosting is the least robust model, potentially struggling with prediction accuracy and occasionally producing more significant errors compared to its counterparts.

Multi-level model stacking

This section explores advanced model stacking to improve predictive performance by acknowledging that not all base models contribute equally. Gradient boosting, identified as the weakest base model, is excluded.

The stacking process begins with two groups of base models. The first group comprises Random Forest, Extra Trees, Decision Tree, and Bagging, combined via linear regression to form MetaLearner-1. Simultaneously, the second group, consisting of XGBoost, CatBoost, Light-GBM, and Histogram-based Gradient Boosting, is integrated into MetaLearner-2.

With MetaLearner-1 and MetaLearner-2 established, ridge regression merges their outputs to create a super learner. This strategic fusion of base models and their predictions via advanced regression techniques results in a potent ensemble model, harnessing individual strengths while mitigating weaknesses. The outcome is a robust predictive tool outperforming any single base model. Figure 9 illustrates the stacking algorithm for enhancing fracture intensity prediction.
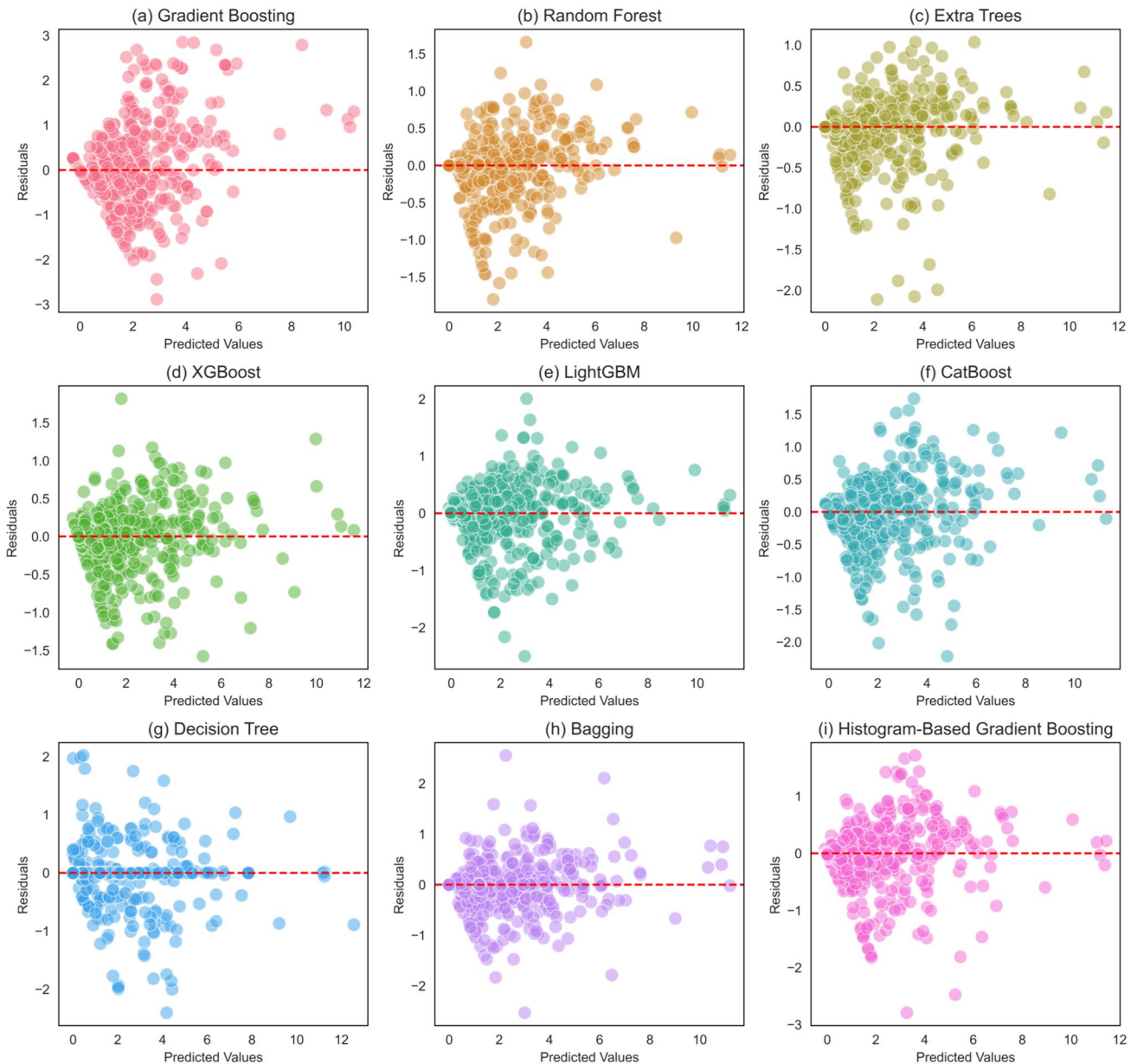
Fig. 8 Residual plots of various base models: **a** gradient boosting, **b** random forest, **c** extra trees, **d** XGBoost, **e** LightGBM, **f** CatBoost, **g** decision tree, **h** bagging, **i** histogram-based gradient boosting

Stacking performance

The efficacy of the stacking method in enhancing predictive performance is showcased by evaluating performance metrics for three models: MetaLearner-1, MetaLearner-2, and the Super Learner. Performance metrics for testing are presented in Table 5, while a visual comparison of the performance of these models in predicting fracture intensity is provided in Fig. 10. Strong predictive capabilities are exhibited by both MetaLearner-1 and MetaLearner-2, which are constructed from different sets of base models.

A Mean Absolute Error (MAE) of 0.095, a Mean Squared Error (MSE) of 0.059, a Root Mean Squared Error (RMSE) of 0.242, and a high R^2 value of 0.974 are demonstrated by MetaLearner-1. Similarly, MetaLearner-2 displays an MAE of 0.109, MSE of 0.061, RMSE of 0.248, and an impressive R^2 value of 0.973. These metrics indicate that the initial choice of base models was sound and contributed positively to the ensemble's performance.

Nonetheless, the Super Learner stands out, with the best performance across all metrics. An outstanding MAE of 0.083, an MSE of 0.039, an RMSE of 0.198, and the highest

Table 4 Statistical summary of residuals for base models

Model	Mean residual	Median residual	Std deviation of residuals	Min residual	Max residual	25th percentile	75th percentile
Gradient boosting	−0.006	−0.005	0.539	−2.884	2.853	−0.088	0.002
Random forest	−0.011	0.000	0.255	−1.810	1.920	0.000	0.000
Extra trees	−0.016	0.000	0.220	−2.235	1.067	0.000	0.000
XGBoost	−0.009	0.000	0.247	−1.575	1.819	−0.011	0.010
LightGBM	−0.015	−0.002	0.300	−2.500	2.010	−0.005	0.008
CatBoost	−0.010	0.000	0.299	−2.215	1.743	−0.021	0.020
Decision tree	−0.005	0.000	0.280	−2.403	2.020	0.000	0.000
Bagging	−0.002	0.000	0.256	−1.591	1.463	0.000	0.000
Histogram-based Gradient boosting	−0.012	−0.001	0.321	−2.790	1.716	−0.006	0.004
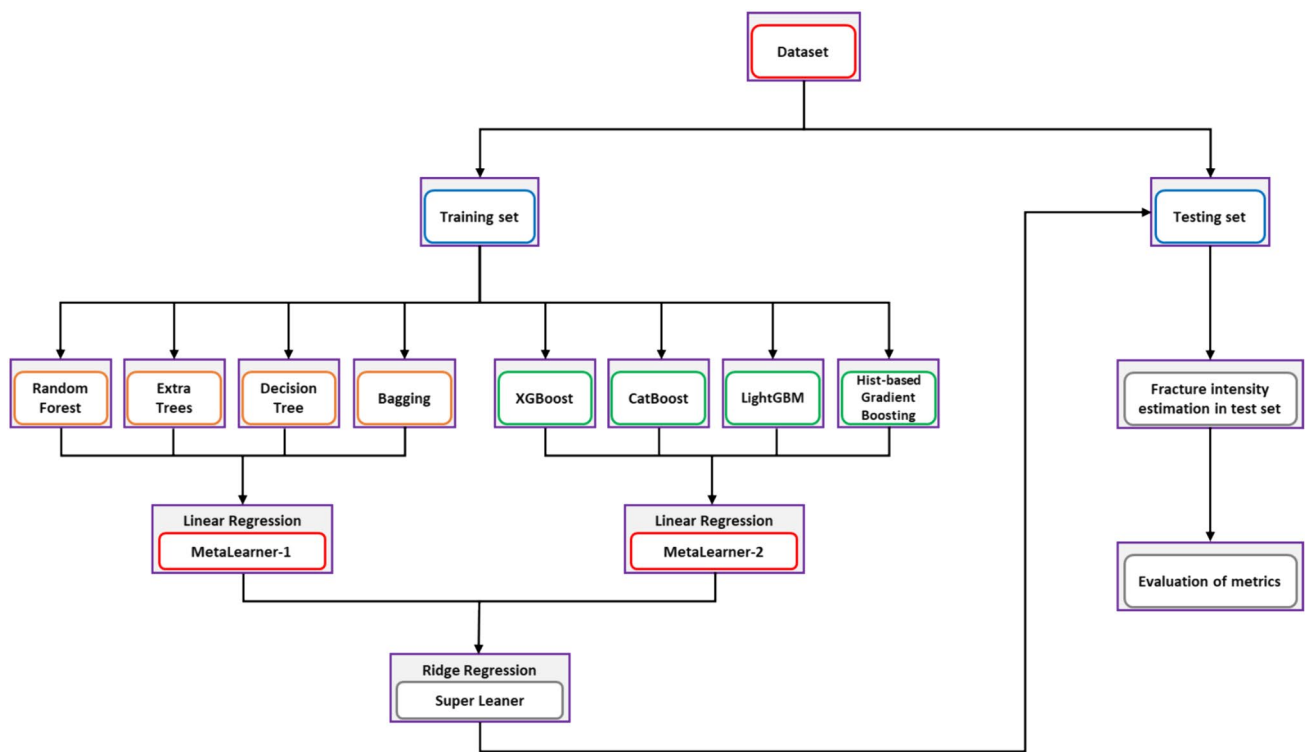
Fig. 9 Stacking algorithm for enhanced fracture intensity prediction

Table 5 Performance metrics for MetaLearner-1, MetaLearner-2, and the super learner in fracture intensity prediction

Model	MAE	MSE	RMSE	R^2
MetaLearner-1	0.095	0.059	0.242	0.974
MetaLearner-2	0.109	0.061	0.248	0.973
Super learner	0.083	0.039	0.198	0.980

R^2 value of 0.980 among the three models are boasted by the Super Learner. These numerical findings unequivocally confirm that the collective strength of the base models is effectively harnessed by stacking, resulting in a more accurate and robust ensemble. Figure 11 displays the actual vs. predicted plot for MetaLearner-1, MetaLearner-2, and the Super Learner.

The stacked Super Learner underscores the potential of ensemble methods in machine learning, emphasizing their practical significance. By amalgamating the predictions of multiple models, the ensemble harnesses the diversity of the
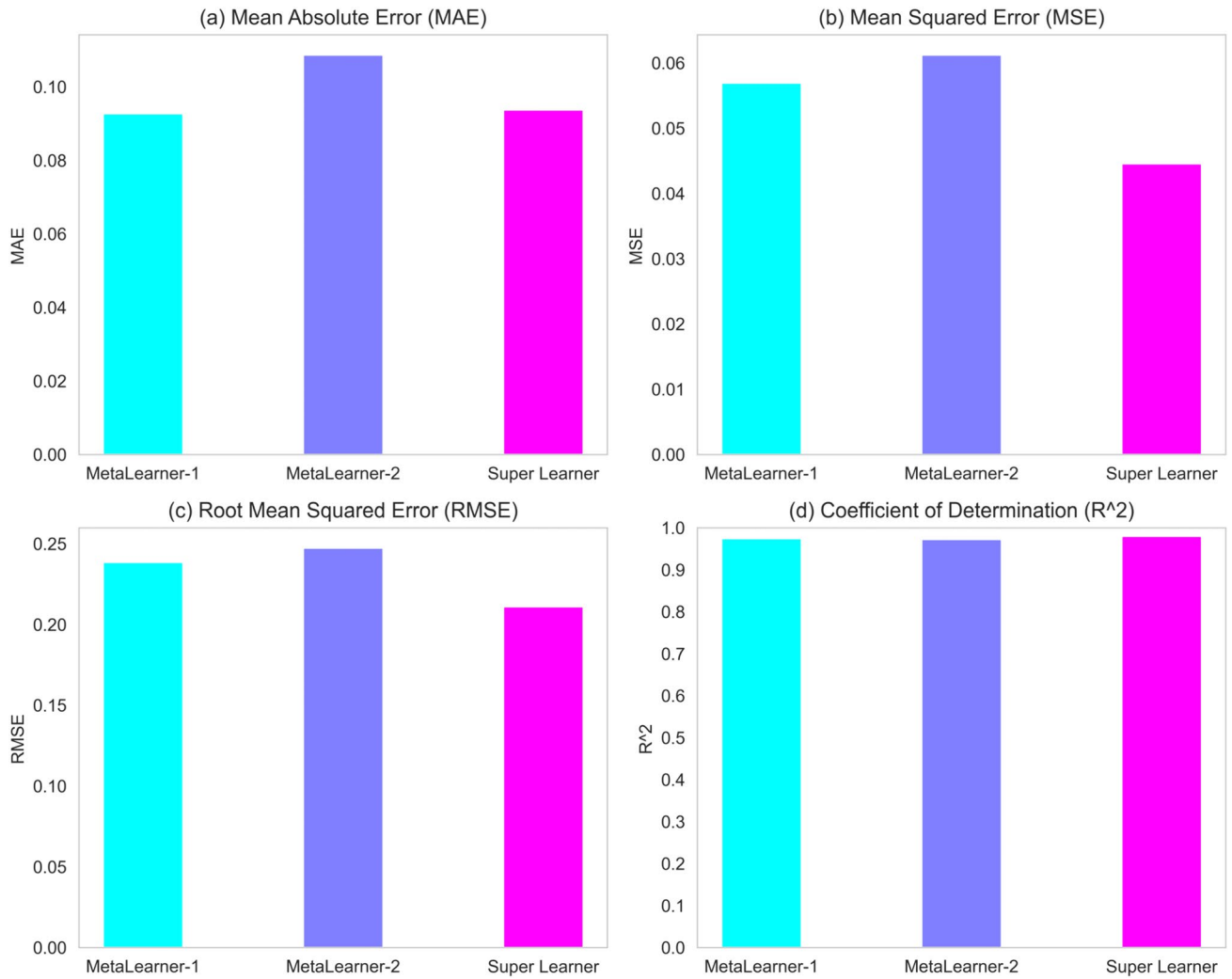
Fig. 10 Comparing the performance of MetaLearner-1, MetaLearner-2, and the super learner in predicting fracture intensity: **a** MAE, **b** MSE, **c**) RMSE, and **d** R^2
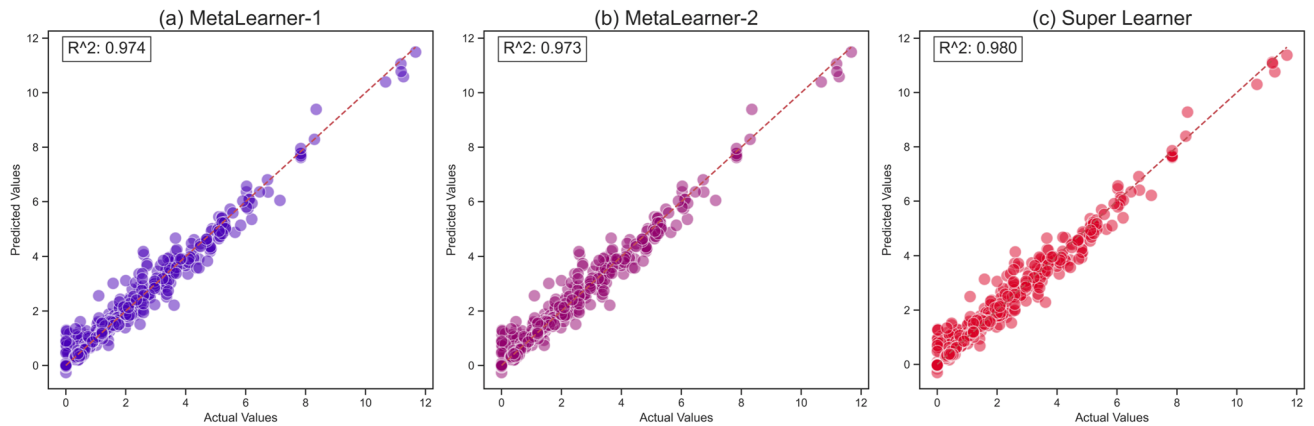

Fig. 11 Comparing actual versus predicted fracture intensity: **a** MetaLearner-1, **b** MetaLearner-2, and **c** super learner

base models, capitalizing on their complementary strengths while mitigating individual weaknesses. In conclusion, the stacking process, supported by numerical evidence, has successfully yielded a Super Learner that outperforms individual meta-learners. The value of stacking as an advanced ensemble technique for constructing powerful predictive models is underscored by its exceptional performance, characterized by reduced prediction errors and higher R^2 values. This approach emphasizes the importance of careful base model selection and demonstrates the potential to achieve superior predictive accuracy by amalgamating the strengths of multiple models through advanced regression techniques like ridge regression.

Discussion

The results presented in this study indicate compelling evidence that advanced ensemble techniques, such as stacking, can substantially improve the predictive performance of machine learning methods for fracture intensity forecasting. Among the tested models, it was consistently observed that the Super Learner, created through stacking, outperformed individual base learners across key evaluation metrics like MAE, MSE, RMSE, and R^2. This superior performance underscores the value of thoughtfully combining complementary models to leverage their strengths. The Super Learner's outstanding MAE of 0.083 and R2 of 0.980 on the test set validate the efficacy of model stacking in improving generalization ability.

A more detailed examination of the base models reveals that ensemble methods, such as Random Forest, Extra Trees, and XGBoost, emerged as the top performers among the base models, demonstrating exceptional accuracy and generalization capability. Their combination of accuracy and robustness renders them well-suited for real-world fracture intensity forecasting. Nevertheless, it was demonstrated that the Super Learner outperformed even these strong baselines, thus highlighting how ensembling can harness the predictive strengths of models like Extra Trees and Random Forest while mitigating their individual limitations. The creation of MetaLearner-1 and MetaLearner-2 also resulted in incremental improvements, further substantiating the utility of model stacking.

The model evaluation metrics demonstrate the superiority of advanced ensemble techniques over individual models for fracture intensity forecasting. Additionally, metrics like feature importance and residual analysis offer valuable diagnostic insights. For example, it consistently emerges that depth is the most influential input feature across models, emphasizing its indispensable role in accurate predictions. Concurrently, residual analysis aided in identifying

model-specific deficiencies, such as the higher variability exhibited by Gradient Boosting.

Overall, the results suggest that ensemble methods are ideally suited for predicting fracture intensity, considering the complexity of the task. By amalgamating diverse models, stacking offers a robust approach to harnessing their complementary capabilities. However, the performance metrics indicate the presence of room for improvement. For instance, the Super Learner's RMSE of 0.198 suggests some variability between predictions and actuals. Investigating whether additional tuning of the base models or a different combination of learners could further enhance accuracy is warranted. Furthermore, testing alternative stacking algorithms like blending could yield incremental improvements. Expanding the model pool to include other advanced techniques like neural networks may also enhance performance.

In terms of practical applications, this study demonstrates the potential to predict fracture intensity in new wells using only conventional well logs and mud loss data without needing advanced and costly image logs or core analysis. The proposed stacked ensemble approach can forecast fracture intensities in new wells during drilling operations by leveraging readily available conventional logs and mud loss data. Once deployed, this would enable real-time optimization of drilling design and hazard mitigation without substantial additional logging investments. This approach could empower geosteering companies to provide predictive services to operators by leveraging legacy data. However, it is imperative to conduct further real-world testing to validate the feasibility and value of large-scale deployment using only conventional logs and mud loss data across fields.

In conclusion, this study underscores the significant performance gains that can be achieved by applying state-of-the-art ensemble techniques like stacking for an intrinsically complex task such as fracture intensity prediction. Nonetheless, specific limitations exist, such as the relatively lower performance of gradient boosting and the potential overfitting of specific models like decision trees. Future work could explore techniques like regularization to enhance generalizability. The proposed approach offers a robust framework for harnessing artificial intelligence to advance fracture intensity modeling.

Conclusions

This research marks a pioneering venture into the domain of continuous fracture intensity modeling through the lens of well logs, employing a sophisticated stacked ensemble methodology. By quantifying fracture intensity as a function of depth, the study sets a new precedent for reservoir characterization accuracy. The innovative stacked ensemble method, integrating the capabilities of gradient

boosting, random forest, and XGBoost, demonstrates a significant leap in predictive performance compared to singular model applications. Furthermore, the ability to predict fracture intensity exclusively from conventional well logs heralds a new era of rapid and cost-efficient reservoir analysis, enabling enhanced recovery strategies through machine learning-informed reservoir simulators. The study culminates in several transformative insights and methodologies that promise to reshape applications within the energy sector:

(1) This study successfully demonstrates the potential of advanced machine learning techniques to predict fracture intensity in geological formations using well-log and mud loss data.

(2) Ensemble-based methods, especially Extra Trees and Random Forest, excel in accuracy and generalizability, showcasing the innovative power of ensemble architectures.

(3) While other parameters (bulk density, neutron porosity, spectral gamma-ray) are important, depth consistently ranks as the predominant factor, necessitating the integration of diverse well-log parameters.

(4) Bagging and Extra Trees are well-suited methods, offering consistency and precision, while Gradient Boosting faces challenges in handling outliers.

(5) The stacking ensemble approach, creating a Super Learner through meta-regression, outperforms individual models, showcasing the advantages of model synergy in ensemble learning for improved accuracy.

(6) The study systematically validates the capability to decode complex geological parameters within well-log data, offering profound implications for refining exploration strategies and optimizing hydrocarbon reservoir targeting.

(7) The ability to predict fracture intensity from well logs advances artificial intelligence applications in refining exploration, optimizing reservoir targeting, and enhancing development and production planning, providing a competitive edge in the oil and gas industry.

(8) To enhance model generalizability, incorporating well logs from diverse geological settings is recommended. Additionally, exploring neural networks and deep learning architectures presents opportunities for further improvement and innovation.

(9) The study lays a robust foundation for leveraging artificial intelligence to extract valuable insights from well-log data, driving innovation at the intersection of data science and the energy industry. It opens new possibilities for gaining a competitive edge in the upstream oil and gas workflow.

Author contributions Conceptualization: Ahmad Azadivash, Ali Kadkhodaie; Methodology: Ahmad Azadivash, Hosseinali Soleymani, Ali Kadkhodaie; Data curation: Atrina Seifrad, Amirali Sandani, Farshid Yahyaee; Visualization: Atrina Seifirad, Amirali Sandani, Farshid Yahyaee; Formal analysis and investigation: Ahmad Azadivash, Ali Kadkhodaie; Writing—original draft preparation: Ahmad Azadivash; Writing—review and editing: Ahmad Azadivash, Hosseinali Soleymani, Ali Kadkhodaie; Resources: Ali Kadkhodaie; Supervision: Ali Kadkhodaie.

Data availability The authors do not have permission to share the data.

Declarations

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

References

Abbasi M, Sharifi M, Kazemi A (2020) Fluid flow in fractured reservoirs: estimation of fracture intensity distribution, capillary diffusion coefficient and shape factor from saturation data. J Hydrol 582:124461. https://doi.org/10.1016/j.jhydrol.2019.124461

Ahmad MW, Reynolds J, Rezgui Y (2018) Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees, and regression trees. J Clean Prod 203:810–821. https://doi.org/10.1016/j.jclepro.2018.08.207

Al-Dujaili AN, Shabani M, Al-Jawad MS (2023) Lithofacies, deposition, and clinoforms characterization using detailed core data, nuclear magnetic resonance logs, and modular formation dynamics tests for Mishrif formation intervals in west Qurna/1 oil field, Iraq. SPE Reserv Evaluation Eng. https://doi.org/10.2118/214689-PA

Arian M (2012) Clustering of diapiric provinces in the Central Iran Basin. Carbonates Evaporites 27:9–18. https://doi.org/10.1007/s13146-011-0079-9

Ayyadevara VK (2018) Pro machine learning algorithms. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-3564-5

Azadivash A, Shabani M, Mehdipour V, Rabbani A (2023a) Deep dive into net pay layers: An in-depth study in Abadan Plain. South Iran Heliyon 9:e17204. https://doi.org/10.1016/j.heliyon.2023.e17204

Azadivash A, Soleymani H, Kadkhodaie A, Yahyaee F, Rabbani AR (2023b) Petrophysical log-driven kerogen typing: unveiling the

potential of hybrid machine learning. J Pet Explor Prod Technol. https://doi.org/10.1007/s13202-023-01688-1

Bauer E, Kohavi R (1999) An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Mach Learn 36:105–139. https://doi.org/10.1023/A:1007515423169

Bentéjac C, Csörgő A, Martínez-Muñoz G (2021) A comparative analysis of gradient boosting algorithms. Artif Intell Rev 54:1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Bhattacharya S, Mishra S (2018) Applications of machine learning for facies and fracture prediction using Bayesian network theory and random forest: case studies from the Appalachian basin. USA J Pet Sci Eng 170:1005–1017. https://doi.org/10.1016/j.petrol.2018.06.075

Boadu FK (1998) Inversion of fracture density from field seismic velocities using artificial neural networks. Geophysics 63:534–545. https://doi.org/10.1190/1.1444354

Boerner S, Gray D, Todorovic-Marinic D, Zellou A, Schnerk G (2003) Employing neural networks to integrate seismic and other data for the prediction of fracture intensity. OnePetro. https://doi.org/10.2118/84453-MS

Breiman L (1996) Bagging predictors. Mach Learn 24:123–140. https://doi.org/10.1007/BF00058655

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Brunet M-F, Korotaev MV, Ershov AV, Nikishin AM (2003) The South Caspian basin: a review of its evolution from subsidence modeling. Sediment Geol 156:119–148. https://doi.org/10.1016/S0037-0738(02)00285-3

Cao M, Sharma MM (2023) A computationally efficient model for fracture propagation and fluid flow in naturally fractured reservoirs. J Pet Sci Eng 220:111249. https://doi.org/10.1016/j.petrol.2022.111249

Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H (2015) Xgboost: extreme gradient boosting. R Package Version 04–2(1):1–4

Daoud EA (2019) Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. Int J Comput Inf Eng 13:6–10

Darling T (2005) Well logging and formation evaluation. Elsevier

Delavar MR (2022) Hybrid machine learning approaches for classification and detection of fractures in carbonate reservoir. J Pet Sci Eng 208:109327. https://doi.org/10.1016/j.petrol.2021.109327

Dorogush AV, Ershov V, Gulin A (2018) CatBoost: gradient boosting with categorical features support. https://doi.org/10.48550/arXiv.1810.11363.

Džeroski S, Ženko B (2004) Is combining classifiers with stacking better than selecting the best one? Mach Learn 54:255–273. https://doi.org/10.1023/B:MACH.0000015881.36452.6e

Fathi E, Carr TR, Adenan MF, Panetta B, Kumar A, Carney BJ (2022) High-quality fracture network mapping using high-frequency logging while drilling (LWD) data: MSEEL case study. Mach Learn Appl 10:100421. https://doi.org/10.1016/j.mlwa.2022.100421

Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F (2012) A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern Part C 42:463–484. https://doi.org/10.1109/TSMCC.2011.2161285

Gao G, Hazbeh O, Davoodi S, Tabasi S, Rajabi M, Ghorbani H (2023) Prediction of fracture density in a gas reservoir using robust computational approaches. Front Earth Sci. https://doi.org/10.3389/feart.2022.1023578

Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. Mach Learn 63:3–42. https://doi.org/10.1007/s10994-006-6226-1

Goetz M, Weber C, Bloecher J, Stieltjes B, Meinzer H-P, Maier-Hein K (2014) Extremely randomized trees based brain tumor segmentation

Goliatt L, Saporetti C, Pereira E (2023) Super learner approach to predict total organic carbon using stacking machine learning models based on well logs. Fuel 353:128682. https://doi.org/10.1016/j.fuel.2023.128682

Golonka J (2004) Plate tectonic evolution of the southern margin of Eurasia in the Mesozoic and Cenozoic. Tectonophysics 381:235–273. https://doi.org/10.1016/j.tecto.2002.06.004

Guryanov A (2019) Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees. In: van der Aalst WMP, Batagelj V, Ignatov DI, Khachay M, Kuskova V, Kutuzov A et al (eds) Analysis of images, social networks and texts. Springer International Publishing, Cham, pp 39–50. https://doi.org/10.1007/978-3-030-37334-4_4

Hastie T, Tibshirani R, Friedman J (2009) Boosting and Additive trees. In: Hastie T, Tibshirani R, Friedman J (eds) The elements of statistical learning: data mining, inference, and prediction. Springer, New York, NY, pp 337–87. https://doi.org/10.1007/978-0-387-84858-7_10

Hawez HK, Sanaee R, Faisal NH (2021) A critical review on coupled geomechanics and fluid flow in naturally fractured reservoirs. J Nat Gas Sci Eng 95:104150. https://doi.org/10.1016/j.jngse.2021.104150

He X, Santoso R, Hoteit H (2020) Application of machine-learning to construct equivalent continuum models from high-resolution discrete-fracture models. OnePetro. https://doi.org/10.2523/IPTC-20040-MS

Ince R (2004) Prediction of fracture parameters of concrete by artificial neural networks. Eng Fract Mech 71:2143–2159. https://doi.org/10.1016/j.engfracmech.2003.12.004

Ja'fari A, Kadkhodaie-Ilkhchi A, Sharghi Y, Ghanavati K (2012) Fracture density estimation from petrophysical log data using the adaptive neuro-fuzzy inference system. J Geophys Eng 9:105–14. https://doi.org/10.1088/1742-2132/9/1/013

Kavoosi MA, Lasemi Y, Sherkati S, Moussavi-Harami R (2009) Facies analysis and depositional sequences of the upper Jurassic Mozduran formation, a carbonate reservoir in the Kopet Dagh basin, ne Iran. J Pet Geol 32:235–259. https://doi.org/10.1111/j.1747-5457.2009.00446.x

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W (2017) LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Advances in Neural Information Processing Systems. Curran Associates, Inc. Vol 30

Kim YD, Durlofsky LJ (2023) Neural network surrogate for flow prediction and robust optimization in fractured reservoir systems. Fuel 351:128756. https://doi.org/10.1016/j.fuel.2023.128756

Kotsiantis SB (2013) Decision trees: a recent overview. Artif Intell Rev 39:261–283. https://doi.org/10.1007/s10462-011-9272-4

Lee T-H, Ullah A, Wang R (2020) Bootstrap aggregating and random forest. In: Fuleky P (ed) Macroeconomic forecasting in the era of big data: theory and practice. Springer International Publishing, Cham, pp 389–429. https://doi.org/10.1007/978-3-030-31150-6_13

Li T, Wang R, Wang Z, Zhao M, Li L (2018) Prediction of fracture density using genetic algorithm support vector machine based on acoustic logging data. Geophysics 83:D49-60. https://doi.org/10.1190/geo2017-0229.1

Lyberis N, Manby G (1999) Oblique to orthogonal convergence across the Turan block in the post-miocene. Bulletin. https://doi.org/10.1306/E4FD2E97-1732-11D7-8645000102C1865D

Myles AJ, Feudale RN, Liu Y, Woody NA, Brown SD (2004) An introduction to decision tree modeling. J Chemom 18:275–285. https://doi.org/10.1002/cem.873

Natekin A, Knoll A (2013) Gradient boosting machines, a tutorial. Front. Neurorobot. 7:21

Ng CSW, Nait Amar M, Jahanbani Ghahfarokhi A, Imsland LS (2023) A survey on the application of machine learning and metaheuristic algorithms for intelligent proxy modeling in reservoir simulation. Comput Chem Eng 170:108107. https://doi.org/10.1016/j.compchemeng.2022.108107

Nhat-Duc H, Van-Duc T (2023) Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for pavement raveling severity classification. Autom Constr 148:104767. https://doi.org/10.1016/j.autcon.2023.104767

Nouri R, Arian M (2017) Multifractal modeling of the gold mineralization in the Takab area (NW Iran). Arab J Geosci 10:105. https://doi.org/10.1007/s12517-017-2923-2

Nouri-Taleghani M, Mahmoudifar M, Shokrollahi A, Tatar A, Karimi-Khaledi M (2015) Fracture density determination using a novel hybrid computational scheme: a case study on an Iranian Marun oil field reservoir. J Geophys Eng 12:188–198. https://doi.org/10.1088/1742-2132/12/2/188

Ouenes A (2000) Practical application of fuzzy logic and neural networks to fractured reservoir characterization. Comput Geosci 26:953–962. https://doi.org/10.1016/S0098-3004(00)00031-5

Pei J, Zhang Y (2022) Prediction of reservoir fracture parameters based on the multi-layer perceptron machine-learning method: a case study of Ordovician and Cambrian carbonate rocks in Nanpu Sag, Bohai Bay Basin. China Processes 10:2445. https://doi.org/10.3390/pr10112445

Prokhorenkova L, Gusev G, Vorobev A, Dorogush A.V, Gulin A (2018) CatBoost: unbiased boosting with categorical features. Advances in neural information processing systems. Curran Associates, Inc. Vol 31

Questiaux J-M, Couples GD, Ruby N (2010) Fractured reservoirs with fracture corridors. Geophys Prospect 58:279–295. https://doi.org/10.1111/j.1365-2478.2009.00810.x

Rahman M, Chen N, Elbeltagi A, Islam MM, Alam M, Pourghasemi HR et al (2021) Application of stacking hybrid machine learning algorithms in delineating multi-type flooding in Bangladesh. J Environ Manage 295:113086. https://doi.org/10.1016/j.jenvman.2021.113086

Rajabi M, Beheshtian S, Davoodi S, Ghorbani H, Mohamadian N, Radwan AE et al (2021) Novel hybrid machine learning optimizer algorithms to prediction of fracture density by petrophysical data. J Pet Explor Prod Technol 11:4375–4397. https://doi.org/10.1007/s13202-021-01321-z

Robert AMM, Letouzey J, Kavoosi MA, Sherkati S, Müller C, Vergés J et al (2014) Structural evolution of the Kopeh Dagh fold-and-thrust belt (NE Iran) and interactions with the south Caspian sea basin and Amu Darya basin. Mar Pet Geol 57:68–87. https://doi.org/10.1016/j.marpetgeo.2014.05.002

Robnik-Šikonja M (2004) Improving random forests. In: Boulicaut J-F, Esposito F, Giannotti F, Pedreschi D (eds) Machine learning: ECML 2004. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 359–370. https://doi.org/10.1007/978-3-540-30115-8_34

Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, Chica-Rivas M (2015) Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol Rev 71:804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Ruh JB, Valero L, Aghajari L, Beamud E, Gharabeigli G (2019) Vertical-axis rotation in East Kopet Dagh, NE Iran, inferred from paleomagnetic data: oroclinal bending or complex local folding kinematics? Swiss. J Geosci 112:543–562. https://doi.org/10.1007/s00015-019-00348-z

Sarkheil H, Hassani H, Alinia F (2009) The fracture network modeling in naturally fractured reservoirs using artificial neural network based on image loges and core measurements. Aust J Basic Appl Sci 3:3297–3306

Sen S, Abioui M, Ganguli SS, Elsheikh A, Debnath A, Benssaou M et al (2021) Petrophysical heterogeneity of the early Cretaceous Alamein dolomite reservoir from North Razzak oil field, Egypt integrating well logs, core measurements, and machine learning approach. Fuel 306:121698. https://doi.org/10.1016/j.fuel.2021.121698

Sikora R, Al-Laymoun O (2015) A modified stacking ensemble machine learning algorithm using genetic algorithms. Handbook of research on organizational transformations through big data analytics. IGI Global, Hershey, pp 43–53. https://doi.org/10.4018/978-1-4666-7272-7.ch004

Smith PF, Ganesh S, Liu P (2013) A comparison of random forest regression and multiple linear regression for prediction in neuroscience. J Neurosci Methods 220:85–91. https://doi.org/10.1016/j.jneumeth.2013.08.024

Speiser JL, Miller ME, Tooze J, Ip E (2019) A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl 134:93–101. https://doi.org/10.1016/j.eswa.2019.05.028

Tabasi S, Soltani Tehrani P, Rajabi M, Wood DA, Davoodi S, Ghorbani H et al (2022) Optimized machine learning models for natural fractures prediction using conventional well logs. Fuel 326:124952. https://doi.org/10.1016/j.fuel.2022.124952

Taghizadeh-Farahmand F, Sodoudi F, Afsari N, Mohammadi N (2013) A detailed receiver function image of the lithosphere beneath the Kopeh-Dagh (Northeast Iran). J Seismol 17:1207–1221. https://doi.org/10.1007/s10950-013-9388-x

Wang Q, Narr W, Laubach SE (2023) Quantitative characterization of fracture spatial arrangement and intensity in a reservoir anticline using horizontal wellbore image logs and an outcrop analogue. Mar Pet Geol 152:106238. https://doi.org/10.1016/j.marpetgeo.2023.106238

Zanchi A, Berra F, Mattei MR, Ghassemi M, Sabouri J (2006) Inversion tectonics in central Alborz, Iran. J. Struct. Geol. 28:2023–37. https://doi.org/10.1016/j.jsg.2006.06.020

Zaree V, Riahi MA, Khoshbakht F, Hemmati HR (2016) Estimating fracture intensity in hydrocarbon reservoir: an approach using DSI data analysis. Carbonates Evaporites 31:101–107. https://doi.org/10.1007/s13146-015-0246-5

Zazoun RS (2013) Fracture density estimation from core and conventional well logs data using artificial neural networks: the Cambro-Ordovician reservoir of Mesdar oil field. Algeria J Afr Earth Sci 83:55–73. https://doi.org/10.1016/j.jafrearsci.2013.03.003

Zhan J, Chen J, Xu P, Han X, Chen Y, Ruan Y et al (2017) Computational framework for obtaining volumetric fracture intensity from 3D fracture network models using Delaunay triangulations. Comput Geotech 89:179–194. https://doi.org/10.1016/j.compgeo.2017.05.005