Petroleum Science 23 (2026) 3805–3833

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

A novel interpretable machine learning framework for predicting gas-bearing properties of tight sandstone reservoirs

Liu Cao [a,b,c], Fu-Jie Jiang [b,d,*], Zhang-Xing Chen [b,c,e], Li-Na Huo [b,d], Run-Hai Feng [f], Di Chen [b,d], Meng-Yang Wang [b,d], Jian Li [c], Yang Gao [g], Ben-Jie-Ming Liu [b,c,h], Yong Ma [b,d], Xiao-Juan Wang [i], Zhi-Min Jin [i], Ao-Bo Zhang [i]

[a] College of Artificial Intelligence, China University of Petroleum (Beijing), Beijing, 102249, China
[b] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[c] Ningbo Key Laboratory of Low-Carbon Hydrogen Energy, Eastern Institute of Technology, Ningbo, 315200, Zhejiang, China
[d] College of Geosciences, China University of Petroleum (Beijing), Beijing, 102249, China
[e] Chemical and Petroleum Engineering, Schulich School of Engineering, University of Calgary, Calgary, T2N 1N4, Canada
[f] Aramco Research Center-Beijing, Aramco Asia, Beijing, 100020, China
[g] Key Laboratory of Orogenic Belts and Crustal Evolution, Ministry of Education, School of Earth and Space Sciences, Peking University, Beijing, 100871, China
[h] College of Safety and Ocean Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[i] Research Institute of Exploration and Development, PetroChina Southwest Oil & Gas Field Company, Chengdu, 610041, Sichuan, China

ARTICLE INFO

Article history:
Received 15 August 2025
Received in revised form
1 February 2026
Accepted 13 May 2026
Available online 20 May 2026

Edited by Xiu-Fang Hu

Keywords:
Tight sandstone
Gas-bearing property prediction
Interpretable machine learning framework
Semi-quantitative prediction

ABSTRACT

Predicting gas-bearing properties in tight sandstone reservoirs presents a global challenge. Traditional methods based on well log interpretation rely heavily on individual experience, which can introduce significant unknown errors. Prediction methods using seismic data and logging labels often fail to capture complex interactions between geological features, resulting in low accuracy. Furthermore, these methods typically determine only gas presence without providing quantitative results. To address these limitations, this study proposes a novel interpretable machine learning (ML) framework. Its novelty lies in: (1) directly linking well testing conclusions to logging data to provide high-resolution, semi-quantitative gas-bearing labels, eliminating intermediate interpretation errors; (2) a systematic comparison of 19 ML algorithms across different paradigms (traditional ML, deep learning, and ensemble learning) using five tailored evaluation metrics, identifying LightGBM as the optimal model for this task (Accuracy = 99.76%); and (3) integrating interpretability directly into the prediction workflow based on cooperative game theory to provide global and local explanations that align with petroleum geological knowledge, significantly enhancing the model's transparency and credibility. Applied to the Xujiahe Formation in the Sichuan Basin, this framework achieves decimeter-level accuracy and demonstrates strong generalization capability. This work proposes a novel framework that enables semi-quantitative gas-bearing property predictions with the potential for basin-scale application, directly identifying sweet spots and offering a more streamlined and interpretable high-accuracy artificial intelligence method for oil and gas resource exploration and development.

* Corresponding author.
E-mail address: jiangfj@cup.edu.cn (F.-J. Jiang).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.05.018

1. Introduction

1.1. Research background

Natural gas is a clean, low-carbon fossil fuel that plays a crucial role in environmental protection, climate change mitigation, as well as the transition to a green, low-carbon economy. Accelerating the development of the natural gas industry is also important in the process for achieving carbon neutrality (Zou et al., 2015, 2021, 2024). Tight sandstone gas, as one of the important types of unconventional natural gas, is widely distributed in major oil and gas basins around the world. Its total resource volume is approximately 210×10^{12} m^3, accounting for 75% of unconventional natural gas (Jia et al., 2022). In the process of oil and gas exploration and development, gas-bearing properties are the most important parameters of a reservoir. Currently, the primary recovery rates in major tight gas fields worldwide are between 20% and 35%, which are relatively low (Dong et al., 2012; Wang and Bing, 2017; Jia et al., 2022). Thus, a qualitative or quantitative prediction of gas-bearing properties is essential for reservoir evaluation, which directly influences the assessment of oil and gas reserves (Qin et al., 2005).

1.2. Key challenges

In the sedimentary context in China, tight sandstone reservoirs are defined as reservoirs with porosity <10% and permeability <1 mD (Zou et al., 2012). Tight sandstone reservoirs often undergo complex depositional, diagenetic, and tectonic evolution processes. They are characterized by low physical properties, complex gas-water distributions, and high clay content. The reservoir and gas-bearing properties usually exhibit strong heterogeneity both laterally and vertically. Therefore, it is challenging to predict gas-bearing properties in tight reservoirs based on traditional exploration or production techniques using seismic or logging data—a global challenge, particularly in the case of China's tight sandstone reservoirs (Zou et al., 2012; Wang et al., 2020; Zhang et al., 2021; Song et al., 2022; Zhao et al., 2022a, 2022b; Xiang et al., 2024).

1.3. Literature review

1.3.1. Seismic prediction technologies

Seismic prediction technologies have undergone long-term development (Hammond, 1974; Seydoux et al., 2020). Since Ostrander's (1984) discovery of the Amplitude Versus Offset (AVO) phenomenon corresponding to gas-bearing sandstone reservoirs in 1984, and with the resurgence of Artificial Intelligence (AI) technology and the vigorous advancement of AI for Science (Wang et al., 2023), the use of seismic data and machine learning (ML) algorithms for predicting gas-bearing properties has gained widespread application. Jiang (2022) used the Zoeppritz equations to simulate pre-stack seismic data, considering deep reservoir characteristics, and optimized and processed data from three wells with high resolution. Finally, based on the AVO theory and neural networks, gas-bearing characteristics were predicted. However, the method could only answer whether gas was present and could not provide more detailed recommendations for exploration and development from a production value perspective. Additionally, the data was derived from simulations, which showed significant differences from the actual geological characteristics of reservoirs. Similar work was done by Yuan (2017) and Zhang (2018).

Júnior et al. (2023) manually labeled the regions of interest where gas accumulations appeared in seismic images, and then input the labeled seismic data into a Long Short-Term Memory (LSTM) neural network for learning and training. Song et al. (2022) used well log interpretation results (well log labels) to mark seismic data for providing gas-bearing labels and employed transfer learning (TL) to generate more labeled data samples to compensate for the lack of data volume. Finally, they predicted gas-bearing properties in tight sandstone reservoirs based on a Semi-Supervised Learning (SSL) method. Gao et al. (2022) also used TL to generate seismic data based on well log labels and then used Convolutional Neural Networks (CNNs) for predicting gas-bearing properties. Wang et al. (2010) explored the effects of different fluids (oil, gas, and water) within a reservoir on seismic profile interpretation using physical simulation experiments.

1.3.2. Well log interpretation

Well log interpretation relies on the subjective judgment and personal experience of geologists and geophysicists, who identify formations, lithology, and fluid types by analyzing well log curves and other data. Therefore, using well log interpretation results to label seismic data could introduce unknown errors (Ellis, 2007; Bjorlykke, 2010). Additionally, gas-bearing properties do not have a simple linear relationship with parameters; changes in parameters can also be caused by other non-gaseous factors, which are difficult to capture using empirical formulas or numerical simulations. Manual extraction is time-consuming and labor-intensive, relying on the researchers' expertise. There may also be varying degrees of interaction information between parameters, which can have significant positive or negative impacts on gas-bearing predictions (Yuan et al., 2015; Wang et al., 2024). Moreover, an ability to correctly interpret a prediction result of an ML model is extremely important (Rudin, 2019). Sometimes, simple models (e.g., linear models) are preferred for their interpretability, even in the era of big data where their accuracy may not match that of more complex models (Lundberg and Lee, 2017). The relationships between parameters and gas-bearing properties (i.e., how each parameter affects a model's prediction results) are rarely investigated, which directly affects the confidence of exploration and development personnel in the model's prediction results (Aras and Hanifi Van, 2022). Therefore, a compromise between the accuracy and interpretability of complex models is crucial for enhancing the credibility of predictions and for directing improvement. More importantly, the well log labels from a few wells cannot provide high-precision gas-bearing labels, nor can they establish a gas-bearing prediction model with broad industrial application value (usually $R^2 \geq 80\%$) at the block or basin scale (Karpatne et al., 2019).

1.3.3. Data comparison

Usually, seismic data have low vertical resolution (usually on the order of meters to tens of meters), making it difficult to identify fine geological features, especially when dealing with deep reservoirs. Their lateral resolution is limited by the wavelength of seismic waves and acquisition parameters. Additionally, the volume of seismic data is enormous, and its processing workflow is complex, requiring specialized software and hardware, as well as large-scale storage and computing resources. This results in high initial exploration costs for seismic data (Yilmaz, 2001; Ellis, 2007; Bjorlykke, 2010; Gao et al., 2022).

Different from seismic data, well logging data provide extremely high vertical resolution (usually on the order of centimeters to decimeters), allowing for detailed characterization of reservoir lithology, physical properties, electrical properties, and hydrocarbon saturation. However, the lateral resolution of well logging data is limited, reflecting only the geological information near a borehole. Broad geological structures require a widely distributed network of wells. Furthermore, well logging data are abundant, standardized in format, and easier to process, with

relatively low logging costs per well (Yilmaz, 2001; Ellis, 2007; Bjorlykke, 2010; Liu et al., 2022).

1.4. Our study

In this study, we propose a novel interpretable ML framework for semi-quantitative prediction of gas content in tight sandstone reservoirs via logging data and actual welling test data from the Xujiahe Formation in the Sichuan Basin, China. This framework includes new methods for dataset construction and preprocessing, model development and evaluation, model validation, model interpretation, and model visualization. We introduce five evaluation metrics to accurately assess the predictive performance of the ML models. Through comparative experiments with 19 different ML algorithms (including traditional ML, deep learning, and ensemble learning), we provide an applicability analysis of various algorithms and establish a gas-bearing property prediction model with the potential for basin-scale application. Additionally, blind well tests conducted in multiple blocks in the Sichuan Basin comprehensively validate the model's generalization capability and robustness. Using the Shapley Additive explanation (SHAP) algorithm based on the cooperative game theory, we conduct a thorough analysis of the model's predictions, providing detailed interpretability reports and visualizing the prediction process and results. This interpretable analysis, combined with knowledge from the petroleum geology field, significantly enhances the model's transparency and the credibility of its predictions. The final results demonstrate that this ML framework exhibits high-resolution prediction capabilities. The proposed framework is designed for semi-quantitative gas-bearing property predictions at a scale that can be expanded to basin-level analyses. This technology holds significant value in the petroleum geology field, as it effectively reduces costs, improves efficiency, accurately identifies sweet spots, and enables rapid resource assessment, thereby providing strong support for the exploration and development of oil and gas resources.

2. Geological setting

The Sichuan Basin is located to the west of the Yangtze Plate and is a large superimposed basin developed on a craton foundation, covering an area of approximately 18×10^4 km^2 (Dai et al., 2021). The Sichuan Basin is the basin in China with the most discovered industrial oil and gas reservoirs to date, with total natural gas resources estimated at approximately 40×10^{12} m^3, However, the proven rate is only 18.8%, indicating significant exploration and development potential (Dai et al., 2021; Yang et al., 2023).

The Sichuan Basin has experienced three significant evolutionary stages: the early marine craton basin, the mid-foreland basin, and the late fold-thrust reformation, alongside multiple tectonic movements (Dai et al., 2021). Since the Late Triassic, the sedimentary environment has gradually shifted from marine to continental, with frequent alternation between sandstone and mudstone, forming a unique sand-mud "sandwich" structure of interbedded layers (Jiang et al., 2023). The Xujiahe Formation features a sequence of coal-bearing clastic deposits in an inland river-lake setting. It can be divided into six members (Xu1 to Xu6) from bottom to top (Fig. 1), with each member showing a trend toward arid conditions, primarily characterized by delta-lake sedimentation (Deng et al., 2022).

Xu3 member is the focus of this study. It develops both source-internal shale gas and near-source tight gas, primarily characterized by interbedded thick shale and thin tight sandstone (Wen et al., 2024; Jin et al., 2025). This structure leads to complex gas-bearing displays that amalgamate signals from both shale gas and tight sandstone gas reservoirs. Consequently, accurately distinguishing lithology (e.g., sandstone and mudstone) becomes a critical prerequisite for reliable gas-bearing property prediction. This necessity underscores the importance of employing interpretable ML models capable of revealing which logging features (e.g., CAL, which reflects borehole conditions and indirectly indicates lithology) are pivotal for such discrimination, as will be elaborated in the model interpretability analysis (Section 5.4).

3. Data and methodology

The fundamental geological data comprise 553 actual well testing conclusions from 222 wells of Xu3 member, sourced from the Southwest Exploration and Development Research Institute of PetroChina, along with logging data from 72 wells. The logging data consist of six conventional logging curves: gamma ray (GR, API), caliper (CAL, cm), compensated neutron log (CNL, %), density log (DEN, g/cm^3), acoustic log (AC, μs/m), and resistivity log (RT, Ω·m).

3.1. Data preparation and preprocessing

3.1.1. Data preparation

Based on the Specification for Well Test in Exploration issued by the National Energy Administration of China (2021), we categorized the 553 well testing conclusions into six types: commercial gas reservoir, low-yield gas reservoir, gas-bearing water reservoir, water-bearing gas reservoir, water reservoir, and dry reservoir (Tables 1 and 2).

Well testing data are decisive for accurately reflecting the type of tested formations, while logging data provide high-resolution vertical data that reflect the physical properties and geological structure of subsurface formations. Therefore, we innovatively propose using well testing data to provide gas-bearing labels for marking logging data, thereby supplying ML models with sufficient data volume to learn the potential relationship between logging data and gas-bearing characteristics. This task is tedious and time-consuming; therefore, we develop a software named DMTL (Data Matching of Well Testing Data and Logging Data) (Fig. 2) to perform the matching between well testing data and logging data. DMTL uses well names and depths as indices to traverse all well locations' well testing data and corresponding logging data, determining whether the depth record points of the logging data fall within the depth interval of the well testing conclusions. If so, the well testing conclusion is assigned to the corresponding logging data; otherwise, the process continues with the next data point until all data have been processed. Finally, we construct a gas-bearing dataset comprising 38667 records from 72 wells (Table 1; Figs. 3 and 4).

As shown in Fig. 3, the sample distribution across categories in this dataset is notably uneven. The "commercial gas reservoir" category substantially outnumbers the others, while categories such as "water reservoir" and "gas-bearing water reservoir" have relatively few samples. Thus, it can be concluded that this dataset is moderately imbalanced.

The pairwise correlation plot (Fig. 4) intuitively reveals the mathematical association characteristics between various features. The off-diagonal regions display the distribution patterns of feature pairs, including linear and nonlinear associations; the diagonal kernel density plots present the distribution morphology of individual features, including statistical properties such as kurtosis and skewness. Taking GR-CAL and GR-CNL as examples, GR and CAL exhibit a vertical band-like distribution parallel to the y-axis, indicating no strong linear or nonlinear relationship between

Table 1
A statistical summary of the logging data from 72 wells of Xu3 member in the Sichuan Basin.

Features	Depth, m	GR, API	CAL, cm	CNL, %	DEN, g/cm^3	AC, μs/m	RT, Ω·m
Count	38667	38667	38667	38667	38667	38667	38667
Mean	3697.29	65.60	8.44	9.27	2.62	60.79	2175.73
Std	666.95	25.30	2.54	5.44	0.13	25.06	10653.06
Min	2124.00	4.05	5.78	−0.27	1.58	−3976.26	2.79
25%	3121.19	51.25	6.40	5.81	2.54	55.60	25.87
50%	3844.25	65.11	8.60	8.10	2.62	60.11	45.17
75%	4255.88	78.15	9.18	11.75	2.71	63.41	127.80
Max	4782.00	233.67	34.13	47.22	3.00	242.85	99990.00

Note: Count is the total number of data, mean is the average value, std is the standard deviation, min is the minimum, 25% is the first quartile, 50% is the median, 75% is the third quartile, and max is the maximum.

Table 2
Specifications for well testing in exploration (National Energy Administration of China, 2021).

Depth, m	Middle oblique depth of reservoir, m	Production		Well testing conclusion
		Gas production, m^3/d	Water production, m^3/d	
≤500	≤2000	≥500	≤0.25	Commercial gas reservoir
500–1000	≤2000	≥1000	≤0.25	Commercial gas reservoir
1000–2000	≤2000	≥3000	≤0.25	Commercial gas reservoir
2000–3000	2000–3000	≥5000	≤0.40	Commercial gas reservoir
3000–4000	3000–4000	≥10,000	≤0.50	Commercial gas reservoir
>4000	>4000	≥20,000	≤0.60	Commercial gas reservoir
≤500	≤2000	200–500	≤0.25	Low-yield gas reservoir
500–1000	≤2000	200–1000	≤0.25	Low-yield gas reservoir
1000–2000	≤2000	200–3000	≤0.25	Low-yield gas reservoir
2000–3000	2000–3000	400–5000	≤0.40	Low-yield gas reservoir
3000–4000	3000–4000	600–10000	≤0.50	Low-yield gas reservoir
>4000	>4000	800–20000	≤0.60	Low-yield gas reservoir
≤500	≤2000	≥500	≥0.25	Water-bearing gas reservoir
500–1000	≤2000	≥1000	≥0.25	Water-bearing gas reservoir
1000–2000	≤2000	≥3000	≥0.25	Water-bearing gas reservoir
2000–3000	2000–3000	≥5000	≥0.40	Water-bearing gas reservoir
3000–4000	3000–4000	≥10,000	≥0.50	Water-bearing gas reservoir
>4000	>4000	≥20,000	≥0.60	Water-bearing gas reservoir
≤500	≤2000	200–500	≥0.25	Gas-bearing water reservoir
500–1000	≤2000	200–1000	≥0.25	Gas-bearing water reservoir
1000–2000	≤2000	200–3000	≥0.25	Gas-bearing water reservoir
2000–3000	2000–3000	400–5000	≥0.40	Gas-bearing water reservoir
3000–4000	3000–4000	600–10000	≥0.50	Gas-bearing water reservoir
>4000	>4000	800–20000	≥0.60	Gas-bearing water reservoir
≤500	≤2000	≤200	≥0.25	Water reservoir
500–1000	≤2000	≤200	≥0.25	Water reservoir
1000–2000	≤2000	≤200	≥0.25	Water reservoir
2000–3000	2000–3000	≤400	≥0.40	Water reservoir
3000–4000	3000–4000	≤600	≥0.50	Water reservoir
>4000	>4000	≤800	≥0.60	Water reservoir
≤500	≤2000	≤200	≤0.25	Dry reservoir
500–1000	≤2000	≤200	≤0.25	Dry reservoir
1000–2000	≤2000	≤200	≤0.25	Dry reservoir
2000–3000	2000–3000	≤400	≤0.40	Dry reservoir
3000–4000	3000–4000	≤600	≤0.50	Dry reservoir
>4000	>4000	≤800	≤0.60	Dry reservoir

them, whereas GR and CNL demonstrate a distinct nonlinear distribution relationship. Consequently, the mathematical relationship between GR and CAL necessitates analysis with more refined mathematical approaches.

3.1.2. Data preprocessing

After *Z*-score standardization of the entire dataset, the mean and standard deviation are transformed to 0 and 1, respectively, thereby eliminating the impact of different scales across various features and normalizing data across different dimensions to a unified scale:

$$Z = \frac{(X - \mu)}{\sigma}. \tag{1}$$

Here,

Z: The standardized score.

X: The original data point.

μ: The mean of the dataset.

σ: The standard deviation of the dataset.

All data were randomly and rigorously partitioned into independent training and test sets at a 7:3 ratio. As the model's default hyperparameter settings achieved adequate predictive accuracy, a validation set was not utilized for hyperparameter tuning.

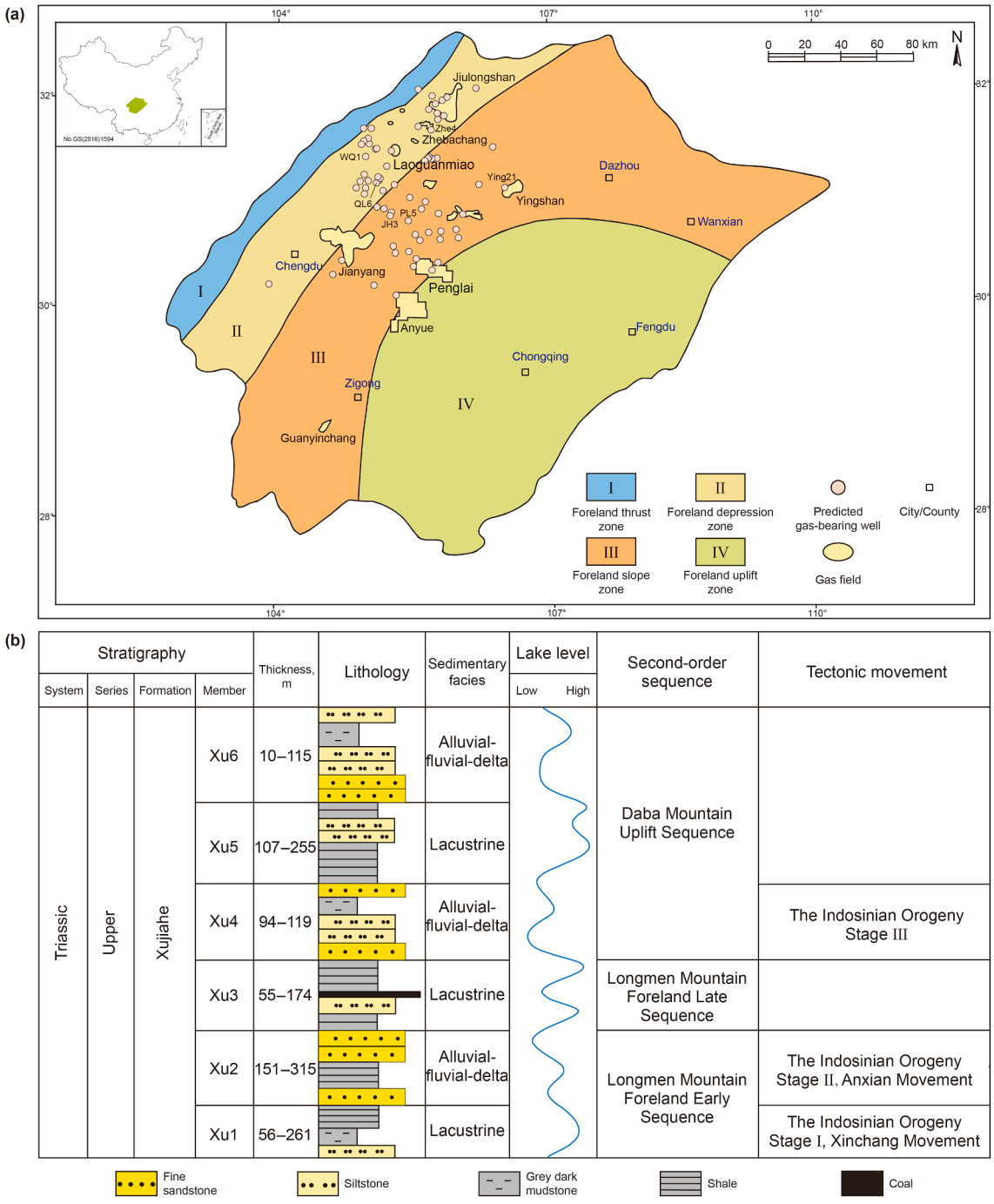

Fig. 1. Tectonic location of Sichuan Basin and lithologic column of Xujiahe Formation.

3.2. Various classification algorithms

In petroleum geology, predicting gas-bearing properties in tight sandstone involves a complex relationship between well logging data—encompassing rock properties, porosity, and fluid characteristics—and the subsurface reservoir's microstructure, making it a challenging multivariate problem. Therefore, conducting a comprehensive comparison of various traditional ML, deep learning, and ensemble learning algorithms can reveal the applicability and effectiveness of different approaches in handling this intricate relationship. Such a comparative analysis not only enhances the understanding of how different mathematical models perform in processing well logging data but also provides a more scientific basis for model selection in predicting gas-bearing properties under complex geological conditions, thereby enhancing the accuracy and reliability of predictions.

3.2.1. Traditional machine learning

Traditional ML grounded in statistics and linear algebra, achieves data pattern learning and prediction by constructing explicit mathematical models. Its core mathematical principles encompass minimizing loss functions (e.g., mean squared error, cross-entropy) and regularization techniques (e.g., L_1/L_2 norms) to prevent overfitting. At the algorithmic level, traditional methods rely

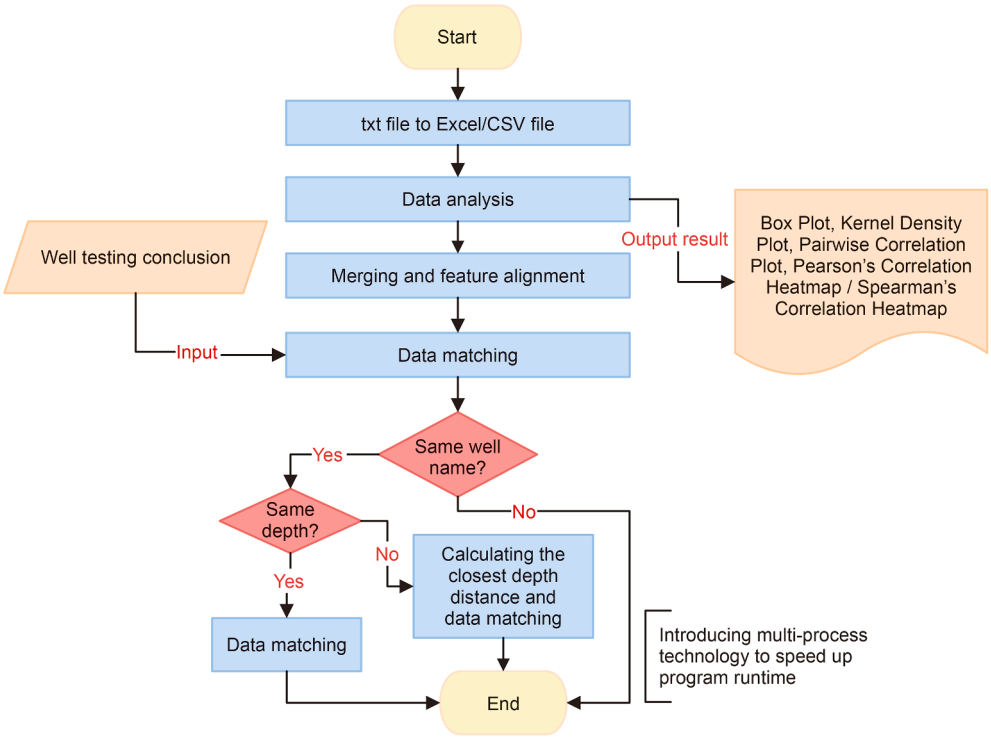

Fig. 2. The workflow of DMTL.

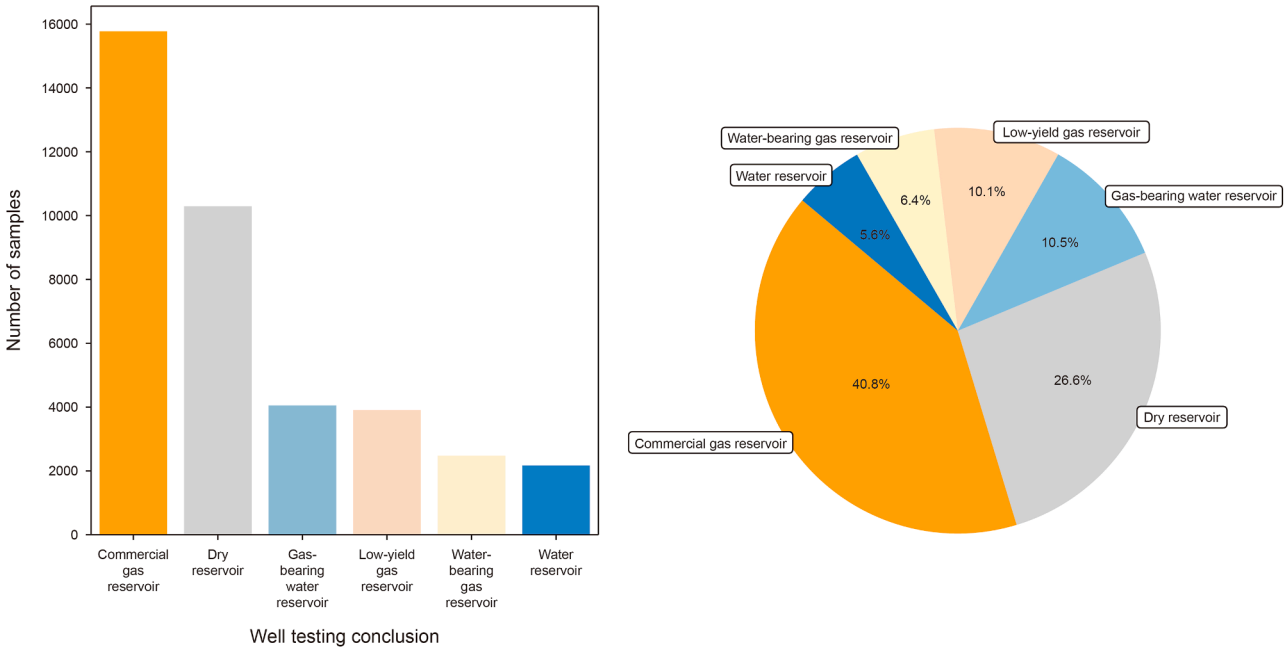

Fig. 3. Bar chart and pie chart of gas-bearing classification data.

on manually engineered feature extraction, such as dimensionality reduction via Principal Component Analysis (PCA) or kernel functions for mapping nonlinear relationships. Their strengths lie in low computational resource demands, strong model interpretability (e.g., linear regression coefficients directly reflect feature importance), and robust performance on small-scale datasets. However, traditional ML methods exhibit high sensitivity to feature quality, limited generalization capabilities constrained by handcrafted feature spaces, and a significantly declining performance when processing high-dimensional nonlinear data (e.g., images, audio).

3.2.1.1. Logistic. The Logistic algorithm is a linear model widely used for binary classification problems. It linearly weights the input features and maps them to probability values using the sigmoid function (Cox, 1958):

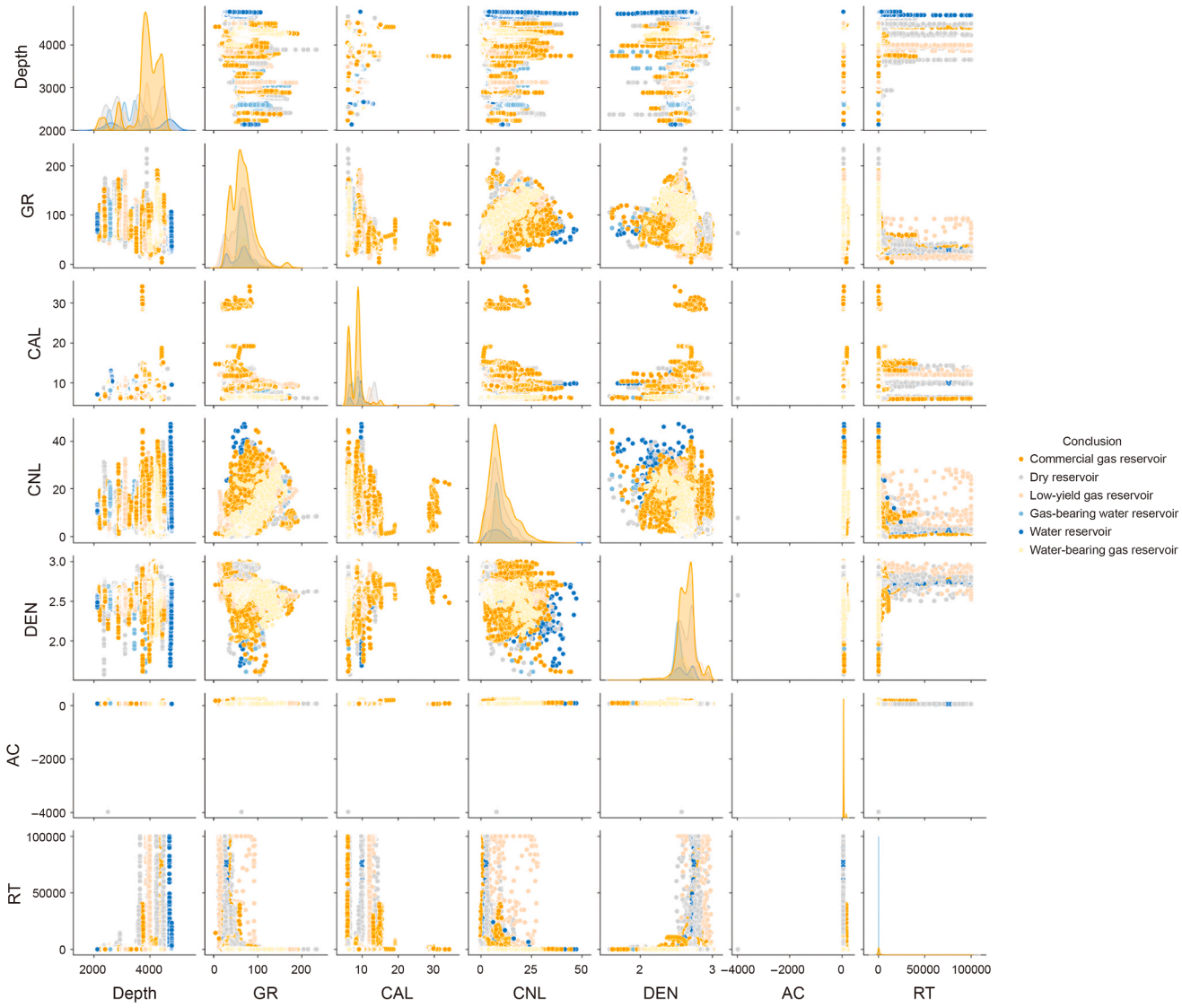

Fig. 4. Pairwise correlation plot of 72 wells between well logging features and well testing conclusions (units are the same as that in Table 1).

$$P(y=1|x)=\frac{1}{1+e^{-(\boldsymbol{w}\cdot x+b)}}, \tag{2}$$

here,

$P(y=1|x)$: The probability of the output being classified as the positive class, given the input x.

$\boldsymbol{w}$: The vector of weights.

b: The bias term.

e: The base of the natural logarithm.

3.2.1.2. Decision tree. The Decision Tree (DT) creates a tree structure via feature splitting strategies (Table A1), used for both classification and regression tasks. It has the advantages of being easy to interpret and not requiring feature standardization (Chou, 1991; Quinlan, 1986; Salzberg, 1994).

3.2.1.3. Support Vector Machine. Support Vector Machine (SVM) finds the optimal hyperplane in a high-dimensional space to maximize the margin between different classes of data, making it suitable for classification tasks (Cortes and Vapnik, 1995):

$$f(x)=\sum_{i=1}^{m}(\widehat{\alpha}_i-\alpha_i)\kappa\left(x_i^{\mathrm{T}}x\right)+b. \tag{3}$$

Here,

$\kappa\left(x_i^{\mathrm{T}}x\right)$: The kernel function.

$\widehat{\alpha}_i$: The predicted value.

α_i: The true value.

b: The biased term.

3.2.1.4. K-Nearest Neighbor. K-Nearest Neighbor (KNN) is a non-parametric method (Table A2) that classifies or regresses based on the proximity of samples in the feature space, relying on distance metrics (Cover and Hart, 1967).

3.2.1.5. Naive Bayes. Naive Bayes (NB) is a probabilistic model based on Bayes' theorem, assuming conditional independence among features, and is suitable for tasks such as text classification (Bayes, 1958).

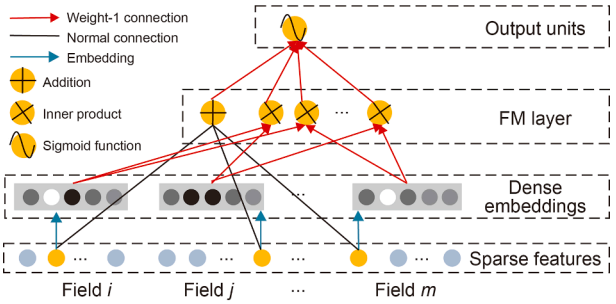

Fig. 5. Architecture diagram of FM (Rendle, 2010).

$$P(y|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|y)\cdot P(y)}{P(\boldsymbol{x})}, \tag{4}$$

$$P(\boldsymbol{x}|y) = \prod_{i=1}^{n} P(x_i|y). \tag{5}$$

Here,

$P(y|\boldsymbol{x})$: The posterior probability, which is the probability of a sample belonging to class y given the feature x.

$P(\boldsymbol{x}|y)$: The likelihood of observing the feature $\boldsymbol{x}$ given the class y.
$P(y)$: The prior probability of class y.
$P(\boldsymbol{x})$: The marginal probability of feature $\boldsymbol{x}$.
n: The total number of features.
x_i: The i-th feature of the feature vector $\boldsymbol{x}$.

3.2.1.6. Factorization Machine. Factorization Machine (FM) models the second-order interactions between features and is used for recommendation systems and classification tasks in sparse data scenarios (Rendle, 2010) (Fig. 5). The prediction function of FM is:

$$\begin{cases} y_{\mathrm{FM}} = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=j+1}^{n} <\boldsymbol{v}_i, \boldsymbol{v}_j> x_i x_j, \\ <\boldsymbol{v}_i, \boldsymbol{v}_j> = \sum_{f=1}^{k} \boldsymbol{v}_{i,f} \cdot \boldsymbol{v}_{j,f}, k \in N_0^+, \\ w_0 \in R, w \in R, V \in R^{n\times k}. \end{cases} \tag{6}$$

Here,

$<\cdot, \cdot>$: The dot product of two vectors of size k.
w_0: The global bias.

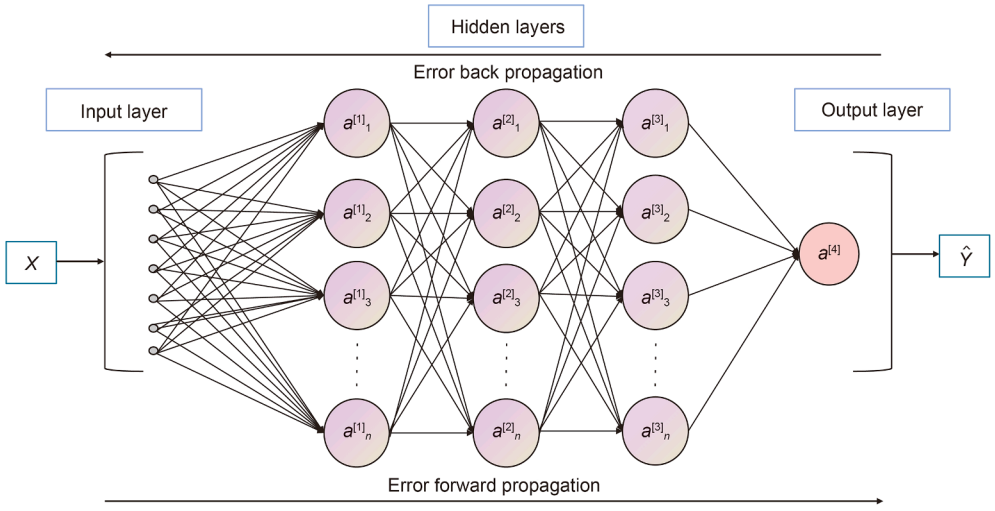

Fig. 6. Architecture diagram of BPNN.

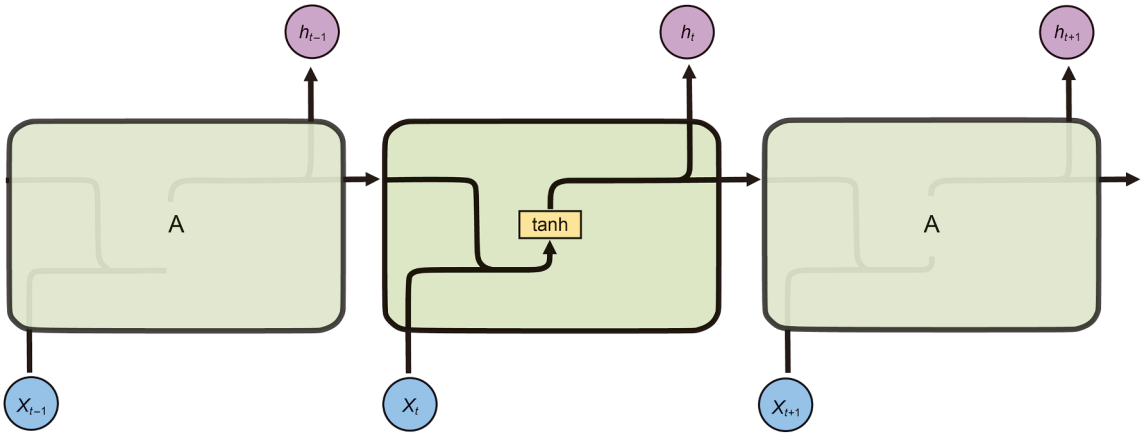

Fig. 7. Architecture diagram of RNN (Elman, 1990).

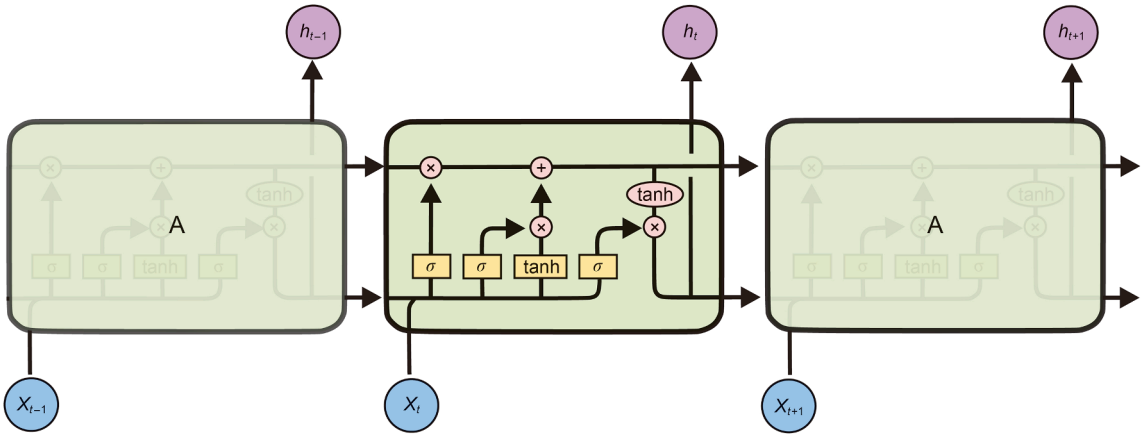

Fig. 8. Architecture diagram of LSTM (Hochreiter and Schmidhuber, 1997).

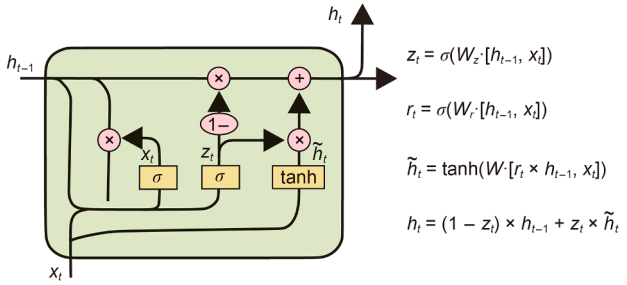

Fig. 9. Architecture diagram of GRU (Cho et al., 2014).

w_i: Modeling the strength of the i-th variable.
$\boldsymbol{v}_i$: The hidden factor vector of x_i feature.
$\boldsymbol{v}_j$: The hidden factor vector of x_j feature.
$< \boldsymbol{v}_i, \boldsymbol{v}_j >$: The inner product of eigenvector $\boldsymbol{v}_i$ and eigenvector $\boldsymbol{v}_j$.

Thus, the FM algorithm introduces latent vector representations to capture the potential interactions between features and calculates the interactions among these latent vectors through dot product operations, thereby achieving the capture of second-order interaction information (Rendle, 2010).

3.2.2. Deep learning

Deep learning simulates the hierarchical feature extraction mechanisms of the human brain through multilayer neural networks. Its mathematical foundation includes the backpropagation algorithm driven by the chain rule, tensor operations, and gradient update strategies of optimizers. The strengths of deep learning lie in its end-to-end feature learning, enabling automatic capture of complex nonlinear relationships in data, and breaking performance bottlenecks in domains such as image classification and natural language processing. However, its models involve massive parameter scales (up to hundreds of millions), training relies on vast labeled datasets and high-performance computing resources, and its "black-box" nature compromises decision interpretability. In oil and gas geological exploration scenarios, this may induce critical decision risks—for instance, low interpretability of reservoir prediction models could lead to biases in well placement selection or resource assessment distortions due to errors in identifying hydrocarbon reservoir boundaries, ultimately resulting in significant economic losses and exploration strategy miscalculations.

3.2.2.1. Back Propagation Neural Network. Back Propagation Neural Network (BPNN) is a feedforward neural network that uses the backpropagation algorithm to minimize the loss function, thereby optimizing the model parameters (Rumelhart et al., 1986) (Fig. 6).

3.2.2.2. Recurrent Neural Network. Recurrent Neural Network (RNN) is a neural network suited for sequential data, capable of passing information between time steps, but it suffers from the vanishing gradient problem (Elman, 1990) (Fig. 7).

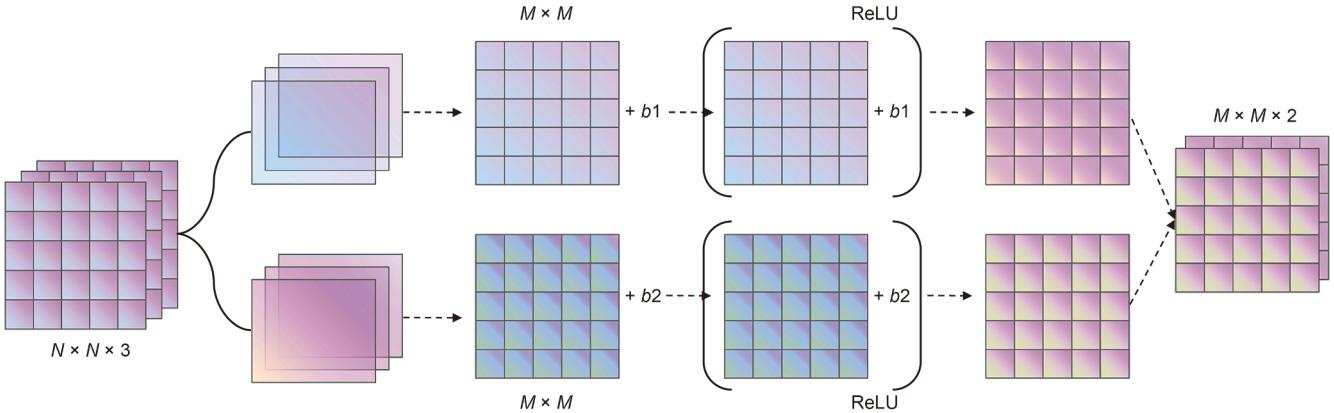

Fig. 10. Architecture diagram for the convolutional layers of a CNN.

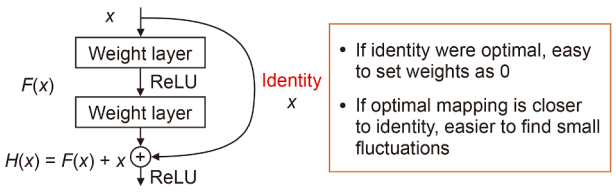

Fig. 11. Architecture diagram for the residual block of ResNet (He et al., 2016).

3.2.2.3. Long Short-Term Memory Neural Network. Long Short-Term Memory Neural Network (LSTM) is an improved version of RNN that features specialized memory cells, allowing it to better capture dependencies in long sequences (Hochreiter and Schmidhuber, 1997) (Fig. 8).

3.2.2.4. Gated Recurrent Unit Neural Network. Gated Recurrent Unit Neural Network (GRU) is a simplified version of LSTM, featuring similar memory functions but with fewer parameters and higher computational efficiency (Cho et al., 2014) (Fig. 9).

3.2.2.5. Convolutional Neural Networks. Convolutional Neural Networks (CNNs) extract local features through convolutional layers and are widely used in image processing and computer vision tasks. The mathematical principles underlying CNNs, particularly the local connectivity and weight-sharing characteristics of their convolutional layers (Fig. 10), predispose them to capturing spatial or local relationships between adjacent features in the input data (LeCun et al., 1998).

3.2.2.6. Residual Neural Network. Residual Neural Network (ResNet) addresses the vanishing gradient problem in deep networks by introducing residual connections (Fig. 11), effectively increasing model depth and performance (He et al., 2016).

3.2.2.7. Deep Factorization Machine. Deep Factorization Machine (DeepFM) consists of an FM component and a Deep Neural Network (DNN) component, which share the same embedding layer input and are responsible for extracting low- and high-order feature interactions, respectively (Fig. 12). The FM component uses latent vectors to capture second-order interactions between features, effectively learning even those that are infrequently observed. The DNN component is a multi-layer feedforward neural network designed to extract high-order feature interactions, enabling the learning of more complex patterns. Consequently, DeepFM can automatically perform feature cross-interactions without the need for manual feature combination design, and it possesses the capability for both linear and nonlinear feature interactions. This confers on DeepFM strong interpretability and positions it as a state-of-the-art (SOTA) model in the field of recommendation systems (Guo et al., 2017). The prediction function of DeepFM is:

$$\hat{y} = \text{Sigmoid}(y_{\text{FM}} + y_{\text{DNN}}). \tag{7}$$

Here,

$\hat{y}$: The prediction result.
Sigmoid: The sigmoid function.
y_{FM}: The output of FM component.
y_{DNN}: The output of deep component.

3.2.3. Ensemble learning

Ensemble learning enhances overall performance by combining multiple base learners (e.g., decision trees, neural networks). Its mathematical foundation lies in the bias-variance decomposition theory: it reduces the variance or bias of individual models through averaging (Bagging) or weighted fusion (Boosting). Representative algorithms include Random Forests, which construct low-correlation decision trees via Bootstrap sampling and random feature selection, integrating results through voting mechanisms; and eXtreme Gradient Boosting (XGBoost) dynamically adjusts sample weights via a gradient boosting framework to progressively optimize residuals. The strengths of ensemble learning include significantly improved generalization capability and robustness to outliers and noise. However, its computational complexity increases exponentially with the number of base learners, and the overall model interpretability faces dual challenges: while a single decision tree possesses inherent interpretability (e.g., visualization of feature splitting paths), interpreting the group decision mechanism of an ensemble requires post-hoc analysis tools such as SHAP values; moreover, designing diversity among base learners still relies on domain-specific knowledge for tuning—otherwise, excessively high model coupling may diminish ensemble benefits.

3.2.3.1. Adaptive Boosting. Adaptive Boosting (AdaBoost) improves the model's ability to identify hard-to-classify samples by combining multiple weak classifiers, making it suitable for classification tasks (Freund and Schapire, 1997) (Table A3).

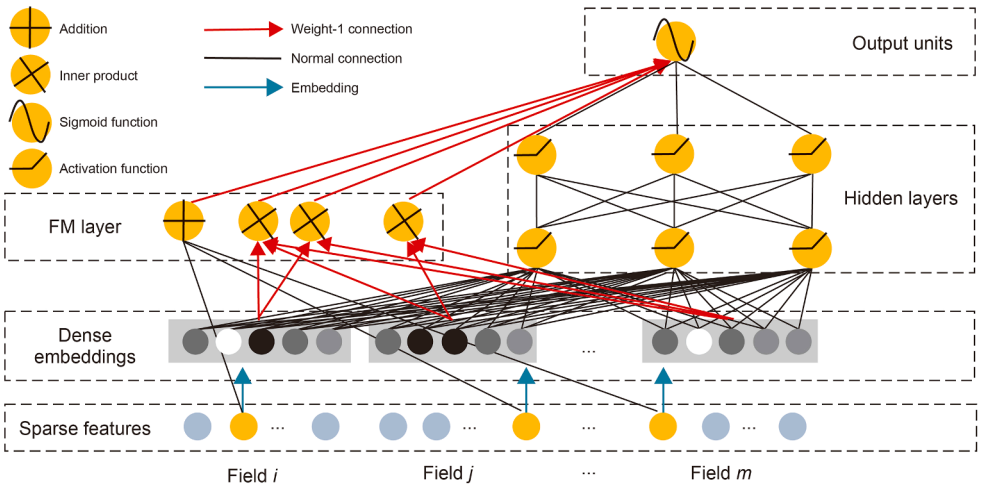

Fig. 12. Architecture diagram of DeepFM (Guo et al., 2017).

3.2.3.2. Gradient Boosting Decision Tree. Gradient Boosting Decision Tree (GBDT) minimizes the loss function by iteratively building multiple decision trees, and it is widely used in regression and classification tasks (Friedman, 2001) (Table A4).

3.2.3.3. Random Forest. Random Forest (RF) is an ensemble method based on multiple decision trees, which improves the model's generalization ability and stability by introducing randomness in data sampling and feature selection (Breiman, 2001a) (Table A5).

3.2.3.4. eXtreme gradient boosting. eXtreme Gradient Boosting (XGBoost) is an improved GBDT algorithm that enhances model performance and computational efficiency through efficient tree structure construction and regularization techniques (Chen and Guestrin, 2016) (Table A6).

3.2.3.5. Light Gradient Boosting Machine. Light Gradient Boosting Machine (LightGBM) is an optimized GBDT algorithm that uses a histogram-based decision tree growth strategy, making it suitable for large-scale datasets (Ke et al., 2017) (Table A7).

3.2.3.6. Categorical Boosting. Categorical Boosting (CatBoost) is a GBDT algorithm specifically designed to handle categorical features, excelling in classification tasks by addressing data biases and optimizing training efficiency (Prokhorenkova et al., 2018) (Table A8).

3.3. Model evaluation metrics

Evaluation metrics are crucial in interpreting the performance of ML models. They provide quantitative analysis tools to help assess how well a model performs on specific tasks and datasets. By selecting and applying appropriate evaluation metrics, one can reveal the strengths and weaknesses of a model, especially when dealing with challenges like data imbalance and bias. These metrics guide the process of model selection, tuning, and improvement, ensuring that ML models possess reliability and generalization ability, which are key to achieving effectiveness in real-world applications.

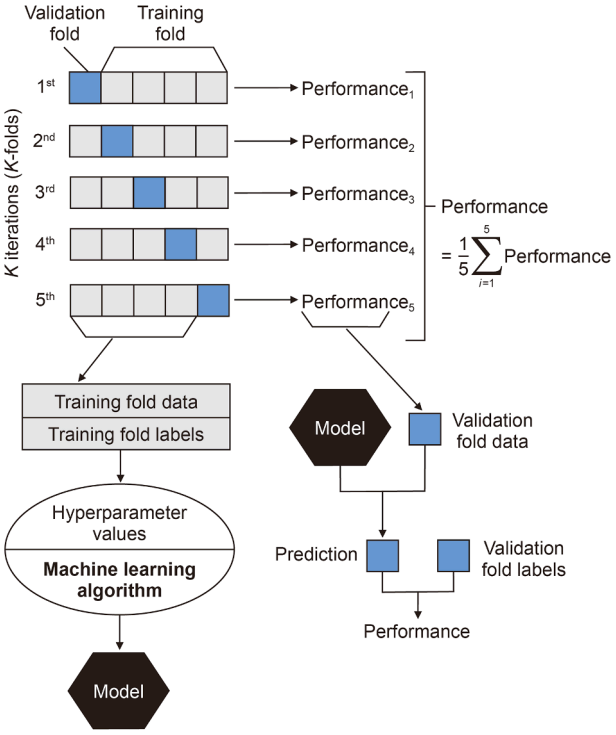

Fig. 13. K-fold cross-validation (take 5-fold cross-validation as an example).

3.3.1. Accuracy

Accuracy measures the proportion of correctly classified samples over the total samples:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \text{Accuracy} \in [0, 1], \tag{8}$$

here,

TP (True Positives): The number of samples correctly predicted as positive.

TN (True Negatives): The number of samples correctly predicted as negative.

FP (False Positives): The number of negative samples incorrectly predicted as positive.

FN (False Negatives): The number of positive samples incorrectly predicted as negative.

When the model is completely incorrect, the Accuracy is 0; when the model is completely correct, the Accuracy is 1. While it is simple and easy to understand, Accuracy can lead to misleading results in scenarios with imbalanced classes.

3.3.2. F1-score

The F1-score is the harmonic mean of Precision and Recall. It is calculated as:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \text{F1} \in [0, 1], \tag{9}$$

here,

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \text{Precision} \in [0, 1], \tag{10}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \text{Recall} \in [0, 1]. \tag{11}$$

The F1-score ranges from 0 to 1. It reaches its maximum value of 1 when both Precision and Recall are high.

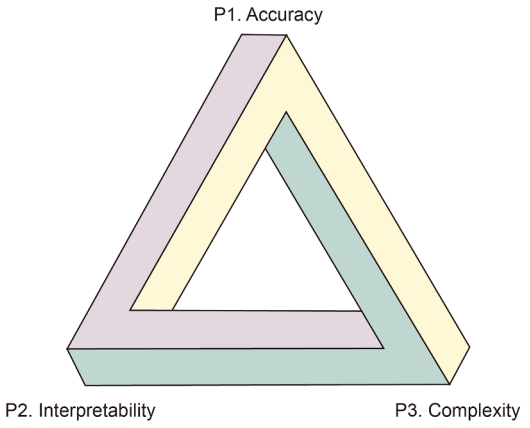

Fig. 14. Impossible triangle of model accuracy, model interpretability and model complexity.

3.3.3. Cross-Entropy Loss

Cross-Entropy Loss (CEL) quantifies the discrepancy between the true label distribution and the predicted probabilities. For multi-class classification, it is defined as:

$$\mathrm{CEL} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(p_{i,c}\right), \mathrm{CEL} \in \left[0, +\infty\right), \tag{12}$$

here,

N: The number of samples.

C: The number of classes.

$y_{i,c}$: The true label for sample i in class c, which is 1 if sample i belongs to class c, otherwise 0.

$p_{i,c}$: The probability predicted by the model that sample i belongs to class c.

A smaller CEL value indicates that the model's predicted probabilities are closer to the true labels. While it performs well on balanced datasets, it may be biased towards the classes with a larger number of samples in imbalanced datasets (Shannon, 1948).

3.3.4. Weighted cross- entropy loss

Weighted Cross-Entropy Loss (WCEL) builds on CEL by assigning weights to each class, making it suitable for moderately imbalanced datasets. By adjusting class weights, WCEL reduces the model's bias towards the majority class, enhances focus on the minority class, and balances the model's attention across all classes. This ensures that the model does not disproportionately favor the class with more samples due to class imbalance (Krizhevsky et al., 2017):

$$\mathrm{WCEL} = -\sum_{i=1}^{N}\sum_{c=1}^{C} w_c \cdot y_{i,c}\log\left(p_{i,c}\right), \mathrm{WCEL} \in \left[0, +\infty\right), \tag{13}$$

here,

w_c: The weight for class c is typically set based on the degree of class imbalance.

When the model's prediction is completely accurate, the loss is 0; when the prediction is completely wrong, the loss approaches $+\infty$.

3.3.5. Focal Loss

Focal Loss (FL) is a loss function specifically designed for extremely imbalanced data. It introduces a modulation factor to reduce the loss contribution of easily classified samples, thereby increasing the model's focus on hard-to-classify samples and improving its ability to recognize minority classes. This method is particularly effective in identifying small-sample classes and is well-suited for datasets with a long-tailed distribution (Lin et al., 2020):

$$\mathrm{FL} = -\sum_{i=1}^{N}\sum_{c=1}^{C} \alpha_c \cdot \left(1 - p_{i,c}\right)^{\gamma} \cdot y_{i,c}\log\left(p_{i,c}\right), \mathrm{FL} \in \left[0, +\infty\right), \tag{14}$$

here,

α_c: The balancing factor for class c is used to balance the impact of positive and negative samples.

γ: The modulation factor, used to reduce the loss contribution of easily classified samples.

When the model's prediction is completely accurate, the loss is 0; when the prediction is completely wrong, the loss approaches $+\infty$.

3.4. K-fold cross-validation

K-fold cross-validation is a robust method for assessing model generalizability by partitioning the dataset into k subsets (folds). The process involves k iterations where, in each iteration, $k-1$ folds are used for training and the remaining fold is used for validation (Fig. 13). The final performance metric is the average of the metrics obtained from the k validation folds (Stone, 1974). This approach reduces the variance of performance estimation and provides a more reliable evaluation of how the model will perform on unseen data (Geisser, 1975; Stone, 1974):

$$\mathrm{CV}(M) = \frac{1}{k}\sum_{i=1}^{k} M_i, \tag{15}$$

Table 3
The results of five evaluation metrics for 19 ML prediction models before 10-fold cross-validation (values in bold indicate the best performing model for each metric).

Categories	Algorithms	Accuracy	F1-score	CEL	WCEL	FL	Time, s
Traditional ML	Logistic	49.38%	43.88%	1.3221	1.7218	1.0637	0.4643
	DT	**98.34%**	**98.33%**	0.5996	0.8024	**0.4564**	0.1989
	SVM	82.48%	82.17%	0.5527	0.7329	0.6983	36.7465
	KNN	94.82%	94.80%	**0.3946**	**0.5082**	0.5024	0.0122
	NB	30.74%	26.66%	3.8976	3.5783	1.1839	**0.0074**
	FM	43.94%	29.77%	1.3753	1.3753	3.3692	19.6470
Deep learning	BPNN	49.44%	39.64%	1.6493	1.3976	4.3233	**3.8673**
	RNN	39.26%	40.58%	1.5393	1.4527	1.0048	4.9102
	LSTM	32.47%	31.81%	1.6090	1.5205	1.0527	7.6199
	GRU	48.05%	39.51%	1.3575	1.3575	0.1872	5.8413
	CNN	90.86%	91.00%	0.2644	0.2366	0.1321	21.6440
	ResNet	93.47%	93.44%	0.2052	0.2795	0.0947	25.9179
	DeepFM	**97.54%**	**97.54%**	**0.0731**	**0.0731**	**0.0402**	286.5755
Ensemble learning	AdaBoost	43.43%	45.68%	1.6406	1.6522	1.2121	**0.9837**
	GBDT	95.73%	95.72%	0.1977	0.2258	0.5580	27.4078
	RF	99.40%	99.39%	0.0757	0.0975	0.4887	3.2426
	XGBoost	**99.74%**	**99.74%**	**0.0125**	**0.0172**	**0.4443**	4.1586
	LightGBM	99.68%	99.68%	0.0138	0.0197	0.4448	2.2886
	CatBoost	99.42%	99.42%	0.0323	0.0404	0.4592	4.7182

Note: The hardware configuration comprised an Apple M2 chip with an 8-core CPU (4 performance cores and 4 efficiency cores), a 10-core GPU, and 8 GB of unified memory. The software environment consisted of the macOS operating system and the relevant ML frameworks utilized for model training and inference.

Table 4
The results of five evaluation metrics for 19 ML prediction models after 10-fold cross-validation (values in bold indicate the best performing model for each metric).

Categories	Algorithms	Accuracy	F1-score	CEL	WCEL	FL	Time, s
Traditional ML	Logistic	42.23%	27.52%	1.4618	1.9790	1.1187	3.8968
	DT	**98.95%**	**98.95%**	**0.3775**	**0.4724**	**0.4496**	1.2284
	SVM	50.21%	43.58%	1.3230	1.7481	1.0717	953.5264
	KNN	92.38%	92.37%	0.5927	0.6087	0.5276	0.1618
	NB	26.85%	22.44%	3.7962	3.4338	1.2113	**0.0552**
	FM	45.60%	35.20%	1.3757	1.3757	3.4163	162.9770
Deep learning	BPNN	46.31%	36.85%	1.6549	1.3958	4.3460	**23.9097**
	RNN	38.94%	40.18%	1.5336	1.4452	0.9967	30.7019
	LSTM	34.05%	33.16%	1.6112	1.5247	1.0540	60.8312
	GRU	48.61%	39.95%	1.3616	1.3616	0.1883	48.1760
	CNN	91.13%	91.21%	0.2453	0.2188	0.1292	196.5983
	ResNet	93.93%	93.91%	0.2054	0.2719	0.0898	251.1108
	DeepFM	**97.79%**	**97.79%**	**0.0205**	**0.1616**	**0.0205**	2106.2677
Ensemble learning	AdaBoost	42.98%	44.82%	1.6491	1.6523	1.2139	**8.1164**
	GBDT	95.75%	95.73%	0.1956	0.2151	0.5578	237.7517
	RF	99.51%	99.51%	0.0638	0.0811	0.4809	28.9522
	XGBoost	99.73%	99.73%	0.0110	0.0152	**0.4433**	30.7700
	LightGBM	**99.76%**	**99.76%**	**0.0104**	**0.0133**	0.4436	16.8701
	CatBoost	99.58%	99.58%	0.0273	0.0319	0.4559	41.2417

Note: The higher WCEL value for DeepFM compared to its CEL can be attributed to the interaction between class weights and the model's training dynamics, highlighting the challenge of fine-tuning complex deep learning models on imbalanced datasets.

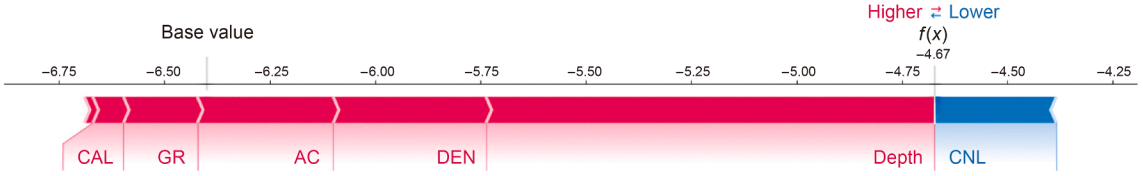

Fig. 15. Force plot of a single prediction (first prediction).

$$M_i = M(f_i,\ D_i). \tag{16}$$

Here,

D_i: The i-th subset of the original dataset D;

k: The number of subsets (folds);

M: The evaluation metric;

M_i: The model performance on the i-th fold;

CV(M): The average performance of k-fold cross-validation;

$M_i = M(f_i,\ D_i)$: The model performance calculated on the i-th validation set D_i using the model trained on the corresponding training data from the i-th fold.

3.5. Model interpretability

In the realm of ML, the pursuit of accurate predictions often leads to the deployment of complex models, such as deep learning. These models can resemble black boxes, rendering their decision-making processes opaque (Rudin, 2019). There exists a trade-off between the interpretability, accuracy, and complexity of models (Fig. 14) (Breiman, 2001b; Bishop, 2006; Rudin, 2019). Elucidating the inner workings of models to enable researchers to understand which features drive predictions is not only crucial for establishing trust in the outcomes but also vital for ensuring the models are fair and accountable (Rudin, 2019; Aras and Hanifi Van, 2022).

The SHAP algorithm is grounded in the cooperative game theory, utilizing the Shapley value to quantify the contribution of each feature to the model's prediction as a numerical value. Within the SHAP framework, a game is defined as a scenario encompassing multiple participants (or features), each striving to maximize their contribution (or utility) through interactions with other participants (Lundberg and Lee, 2017). Here, "utility" refers to the degree of influence a participant exerts on the model's predictive outcome.

Local explanation models often use simplified input $\boldsymbol{x}'$ that map to the original inputs through a mapping function $\boldsymbol{x} = h_x(\boldsymbol{x}')$, and try to ensure $g(\boldsymbol{z}') \approx f(h_x(\boldsymbol{z}'))$ whenever $\boldsymbol{z}' \approx \boldsymbol{x}'$. $h_x(\boldsymbol{x}')$ may contain less information than $\boldsymbol{x}$ because h_x is specific to the current input $\boldsymbol{x}$ (Ribeiro et al., 2016). The SHAP algorithm provides a feature contribution breakdown for model predictions by considering all possible feature combinations. This method of decomposition implicitly includes interaction information between features, encompassing second-order and higher-order interactions (Lundberg and Lee, 2017).

$$\begin{cases} \phi_i(f,\boldsymbol{x}) = \sum\limits_{\boldsymbol{z}'\subseteq\boldsymbol{x}'} \dfrac{|\boldsymbol{z}'|\,!\,(M-|\boldsymbol{z}'|-1)!}{M!}[f_x(\boldsymbol{z}') - f_x(\boldsymbol{z}'\backslash i)], \\ \boldsymbol{z}' \subseteq \boldsymbol{x}'. \end{cases} \tag{17}$$

Here,

ϕ_i: The effects of each feature (i.e., Shapely value).

$|\boldsymbol{z}'|$: The number of non-zero entries in $\boldsymbol{z}'$.

$\boldsymbol{z}' \subseteq \boldsymbol{x}'$: Representing all $\boldsymbol{z}'$ vectors where the non-zero entries are a subset of the non-zero entries in $\boldsymbol{x}'$.

4. Results

4.1. Prediction and validation

We conducted a comparative experiment with 19 algorithms: six traditional ML, seven deep learning, and six ensemble learning algorithms (Table 3). We also used five evaluation metrics tailored for various data distribution characteristics to assess each model's

performance comprehensively. This comparison aims to demonstrate the applicability and effectiveness of ML algorithms in addressing the complex geological relationship between well logging data and the gas-bearing capacity of tight sandstone. The model has already achieved excellent performance with default hyperparameters, thus no hyperparameter tuning is involved here.

Table 3 indicates that XGBoost achieved the highest overall Accuracy of 99.74% and the lowest WCEL value of 0.0172, while LightGBM followed closely behind. In contrast, DeepFM reached an Accuracy of 97.54%, excluding tree-based algorithms, and had the lowest FL value of 0.0402.

10-fold cross-validation is a common form of cross-validation where the dataset is divided into 10 subsets. In each iteration, 9 subsets are used for training, and the remaining 1 subset is used for validation. This process is repeated 10 times, with each subset serving as the validation set once. The average performance across all 10 iterations is taken as the final evaluation result. This method effectively reduces the risk of overfitting and provides a reliable estimate of the model's performance on unseen data.

Compared to Table 3, the prediction Accuracy of LightGBM and DeepFM in Table 4 has increased. LightGBM achieved the highest overall Accuracy of 99.76% and the lowest WCEL value of 0.0133. Meanwhile, DeepFM attained the highest Accuracy of 97.54% among non-tree-based algorithms and the lowest global FL value of 0.0402. This result demonstrated the predictive reliability of LightGBM and DeepFM on unseen data, highlighting their strong generalization ability within the geological context of the Xu3 member.

It is observed that Tables 3 and 4 reveal that ensemble learning algorithms generally outperform Deep Learning algorithms in terms of WCEL, while Deep Learning algorithms excel in FL. Given that the dataset in this study is class-imbalanced rather than extremely imbalanced, and based on the discussion in Section 3.3, we chose WCEL as the best metric for evaluating model predictions and selected LightGBM as the final predictive model.

4.2. Model interpretability

In numerous applications, comprehending why a model makes a particular prediction can be as crucial as the accuracy of the prediction itself, especially for complex models such as ensemble learning and deep learning (Lundberg and Lee, 2017). Understanding these models can foster appropriate user trust, provide insights into how the model can be improved, and support the understanding of the processes being modeled (Lundberg and Lee, 2017; Rudin, 2019). Here, the SHAP algorithm is employed to elucidate the predictions of the highly generalizable LightGBM model, encompassing both local interpretability and global interpretability.

4.2.1. Local interpretability

Local interpretability provides granular details of predictions, focusing on explaining how individual predictions are generated. It aids decision-makers in trusting the model and elucidates the impact of each feature on the model's single decision-making instance (Lundberg and Lee, 2017).

SHAP values decompose the influence of each feature within a prediction by contrasting the prediction with the baseline value of a particular feature, thereby elucidating the impact of that feature taking on a specific value. Fig. 15 illustrates the prediction decomposition process for the first sample: the baseline prediction value is −6.401, with cumulative feature contributions elevating the prediction to −4.67. Features marked in red exert a positive driving effect on the prediction outcome, while those in blue demonstrate inhibitory effects, where the length of feature bars visually reflects their contribution intensity. Fig. 16 quantifies the cumulative contribution paths of each feature: the depth parameter constitutes the primary decision basis with a SHAP value of +1.06, followed by DEN at +0.36. This fully presents the decision chain formed by synergistic contributions of all parameters, clarifying the impact of each feature on the model's individual decision instance, thereby enhancing model trustworthiness.

4.2.2. Global interpretability

Global interpretability seeks to comprehend the overarching structure of a model, which is inherently more challenging than explaining individual predictions. This endeavor involves elucidating the general working principles of the model, rather than focusing on a single prediction.

Fig. 17 reveals the feature control mechanisms in commercial gas reservoir prediction through SHAP value distribution. The vertical axis ranks features by descending importance, while the horizontal axis indicates the contribution direction of SHAP values to predictions (right: positive, left: negative). Color gradients reflect feature value magnitude (red: high, blue: low). As the primary controlling factor, high-value samples of the Depth parameter (deep red points) cluster predominantly in the positive SHAP region (right half-axis). Dense point clusters indicate specific geological blocks, reflecting a geological regularity: sandstone gas-bearing properties exhibit similarity under specific conditions, with wells concentrated in favorable zones. Conversely, RT samples cluster near the central axis, and their variations exhibit no significant impact on the model's predictive outcomes. This

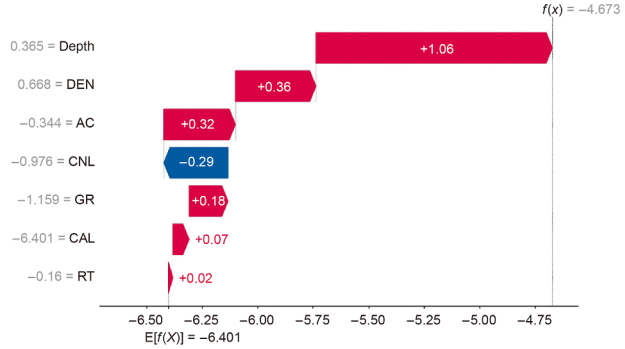

Fig. 16. Waterfall plot of a single prediction (first prediction).

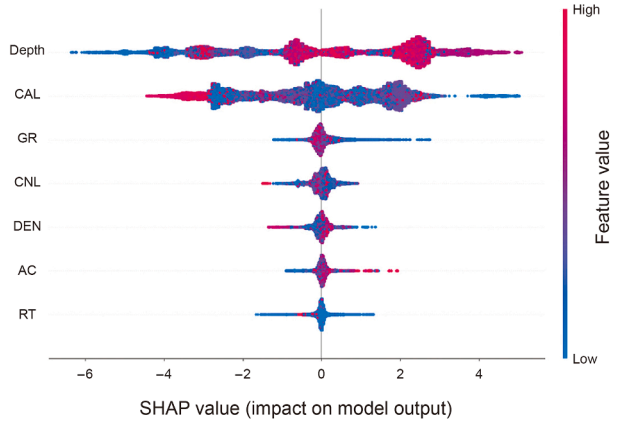

Fig. 17. SHAP value (impact on mode output for class of commercial gas reservoir).

indicates that RT may not be a critical factor in predicting the gas-bearing properties of commercial gas reservoir—consistent with the well-logging principle that RT reflects fluid property changes in formations. Since the model exclusively predicts commercial gas reservoirs without encountering fluid property variations, RT becomes less consequential in this context.

Fig. 18 displays the SHAP value heatmap matrix for the top 1000 samples in the commercial gas reservoir. The vertical axis represents distinct features, with each row indicating the SHAP values of one feature across samples; the horizontal axis denotes individual samples, with each column showing the SHAP values of all features for one sample. Red indicates positive contributions of features to model predictions (pushing predictions toward the positive direction), while blue signifies negative contributions (pushing predictions toward the negative direction). This visualization reveals the influence intensity and variation of each feature across samples. For instance, features such as Depth and CAL exhibit significant chromatic variations among samples, indicating inconsistent impacts on predictions—reflecting the complexity of geological conditions where different depths and CAL values may differentially affect gas-bearing properties. Additionally, cluster patterns in the heatmap correspond to concentrated distributions of commercial gas reservoir data points, potentially indicating geologically favorable gas-bearing zones where features exert pronounced influences. By identifying feature-specific impacts on sample clusters, this analysis indirectly reveals the distribution of favorable gas-bearing zones, thereby enhancing the interpretation of model predictions and enabling integration with geological knowledge for validation and optimization.

The cross-category comparative system based on Mean (|SHAP value|) (Fig. 19) reveals the distribution patterns of feature contribution intensity across different gas-bearing classes. Each subplot employs feature-ranked bar charts to visualize parametric influence disparities: the horizontal axis quantifies the average impact intensity of features on model output (Mean (|SHAP value|)), while the vertical axis lists features in descending order of influence. Analysis demonstrates that Depth and CAL dominate in commercial gas reservoir, potentially reflecting synergistic control of burial depth and caliper on gas reservoir identification. In low-yield gas reservoir and water-bearing gas reservoir, AC and DEN emerge as core parameters alongside Depth and CAL, likely revealing pore structure constraints on productivity. Conversely, for gas-bearing water reservoir, water reservoir, and dry reservoir,

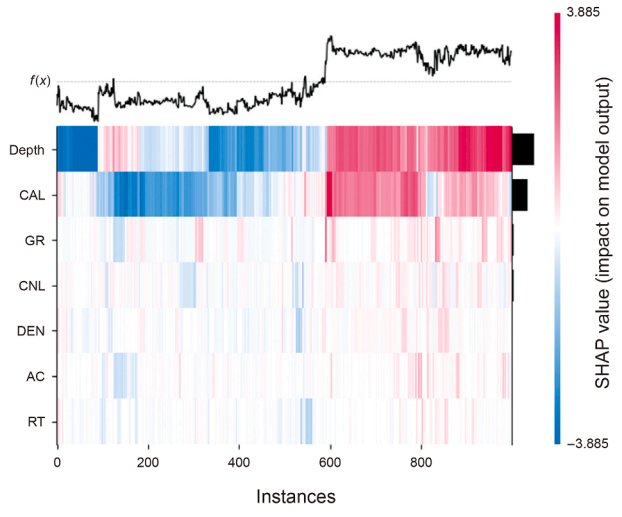

Fig. 18. Heatmap of SHAP value for the initial 1000 predictions (impact on mode output for class of commercial gas reservoir).

the importance of RT exhibits a stepwise increase, with its mean variation closely aligned with the resistivity response attenuation caused by elevated formation water saturation. This category-specific parameter contribution provides quantitative evidence for deciphering model decision-making pathways under complex geological conditions, while simultaneously validating the efficacy of well-logging parameter response mechanisms within the ML framework. In Section 5.4, we will discuss variations in feature importance across different categories to provide a model interpretability report.

The SHAP interaction value matrix (Fig. 20) systematically deciphers multi-parameter synergistic mechanisms in commercial gas reservoir prediction. This matrix decouples complex high-order interactions into quantifiable pairwise feature interaction terms through mathematical modeling: diagonal elements characterize single-parameter main effect intensity, while off-diagonal elements reveal dual-parameter synergies (e.g., a distinct GR-CAL interaction is observed, further illustrated in Fig. 21). Within the SHAP theoretical framework, the sum of interaction effects equals the global deviation between model predictions and the base value. Thus, in matrix representation, each off-diagonal interaction term is equally allocated to symmetric positions. This visualization method not only preserves the mathematical completeness of feature interactions but also intuitively displays the direction and intensity of parametric synergies through chromatic gradients (red indicates positive interactions, blue indicates negative interactions), establishing a quantitative analysis pathway for interpreting the geological decision logic of ML models.

The feature interaction SHAP values (Fig. 21) reveal synergistic mechanisms of different parameter combinations on commercial gas reservoir prediction. The horizontal axis denotes the value of the first feature, the vertical axis represents the interaction SHAP values between the first and second features, and the color indicates the value of the second feature (reflecting how interaction effects vary with the second feature). Taking the GR-CAL interaction as an example (Fig. 21(d)), increasing GR values (indicating higher clay content) correlate with significantly decreasing interaction SHAP values (blue regions with SHAP values < -0.5), exhibiting a distinct negative exponential distribution. This indicates that the coupling of high GR and stable CAL values (no caliper anomaly) suppresses the model's probability of gas reservoir identification. Such negative interaction aligns with the geological setting of the Xujiahe Formation's sand-mud interbeds: intervals with clay interlayers (high GR) typically exhibit caliper stabilization (CAL near theoretical values), suggesting low fracture density and high clay content that reduce reservoir permeability, thus decreasing gas-bearing probability predictions. Conversely, in the AC-CAL interaction (Fig. 21(f)), data points show random dispersion, indicating no significant synergistic impact of acoustic transit time (AC) and caliper (CAL) combinations on model predictions. This reflects the absence of geological causality between acoustic propagation characteristics and borehole stability in the study area. The disparity in interaction effects unveils the model's learning logic for reservoir heterogeneity—prioritizing feature combinations with geological causality (e.g., the joint response of GR-CAL to clay content) while ignoring parameter interactions with weak statistical correlations.

4.3. Model visualization

Model visualization is essential for gaining insights, making informed decisions, and clearly communicating results. It bridges the gap between the mysterious inner workings of ML models and our visual understanding of patterns. Proper model visualization

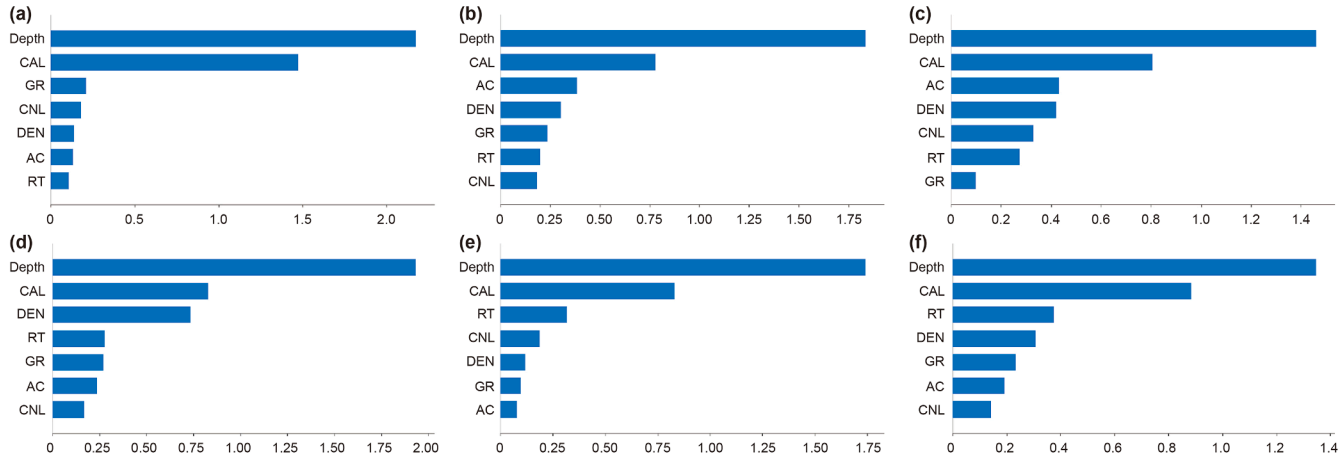

Fig. 19. Mean (|SHAP value|) (average impact on mode output magnitude for all classes): **(a)** commercial gas reservoir, **(b)** low-yield gas reservoir, **(c)** gas-bearing water reservoir, **(d)** water-bearing gas reservoir, **(e)** water reservoir, **(d)** dry reservoir.

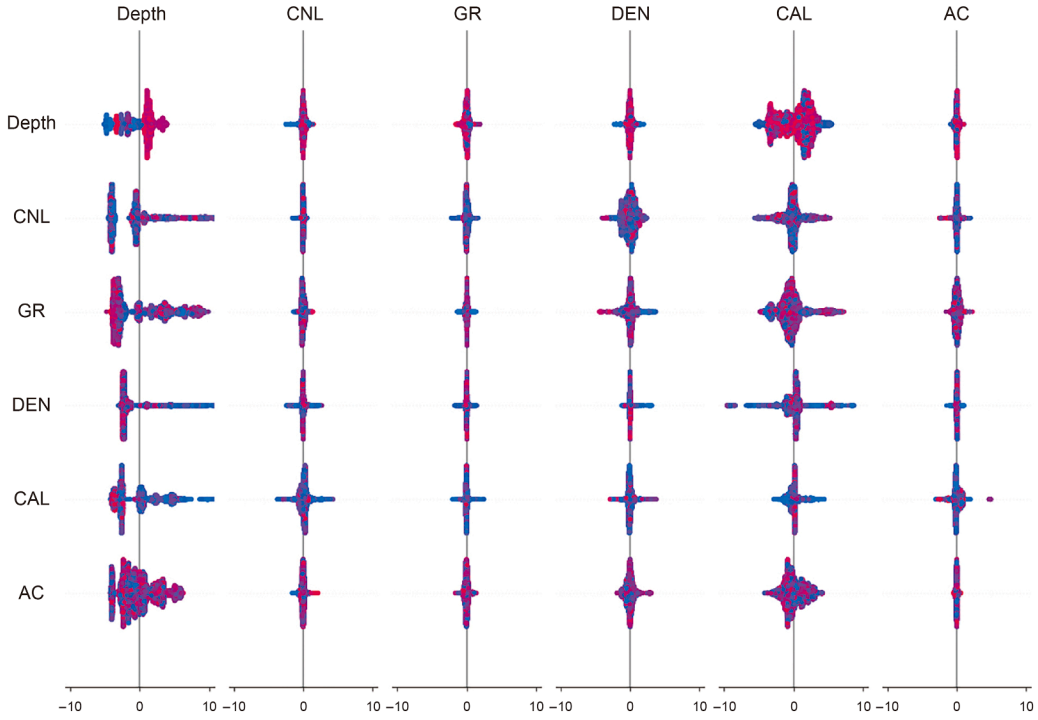

Fig. 20. SHAP interaction value matrix (class of commercial gas reservoir; RT is automatically removed due to its low impact).

not only enhances experts' trust in the model but also facilitates its acceptance and effectiveness in practical applications.

Fig. 22 reveals the learning mechanism of reservoir geological features through the topological structure of a decision tree. Decision nodes (rectangular boxes) and leaf nodes (circles) form a hierarchical decision logic, where the root node performs the initial split based on Column_0 (i.e., Depth) with a gain value of 12.7 (left child gain: 9.3, right child gain: 3.4). This indicates that the depth parameter holds the highest feature importance in data segmentation, confirming the control of vertical diagenetic evolution on reservoir heterogeneity. Non-leaf nodes direct data flow to child nodes through conditional judgments (e.g., Column_2 ≤ -0.812 corresponds to a CAL logging threshold). For instance, when CAL values meet this threshold, 231 samples are assigned to Leaf Node 0 (predicted value: -0.696). The visualization demonstrates that the model has learned complex decision rules from training data, transforming the ML black box into a traceable geological response path, thereby enabling bidirectional validation between model decision logic and geological genesis.

Figs. 23–28 validate the regional adaptability of the model through cross-regional comparisons of six randomly sampled blind well locations. The multi-well visualization demonstrates high consistency between ML predictions and actual well testing data, with prediction stability significantly superior to that of traditional manual log interpretation methods. The comparative analysis of logging curves further enhances the understanding of true formation conditions and model prediction mechanisms.

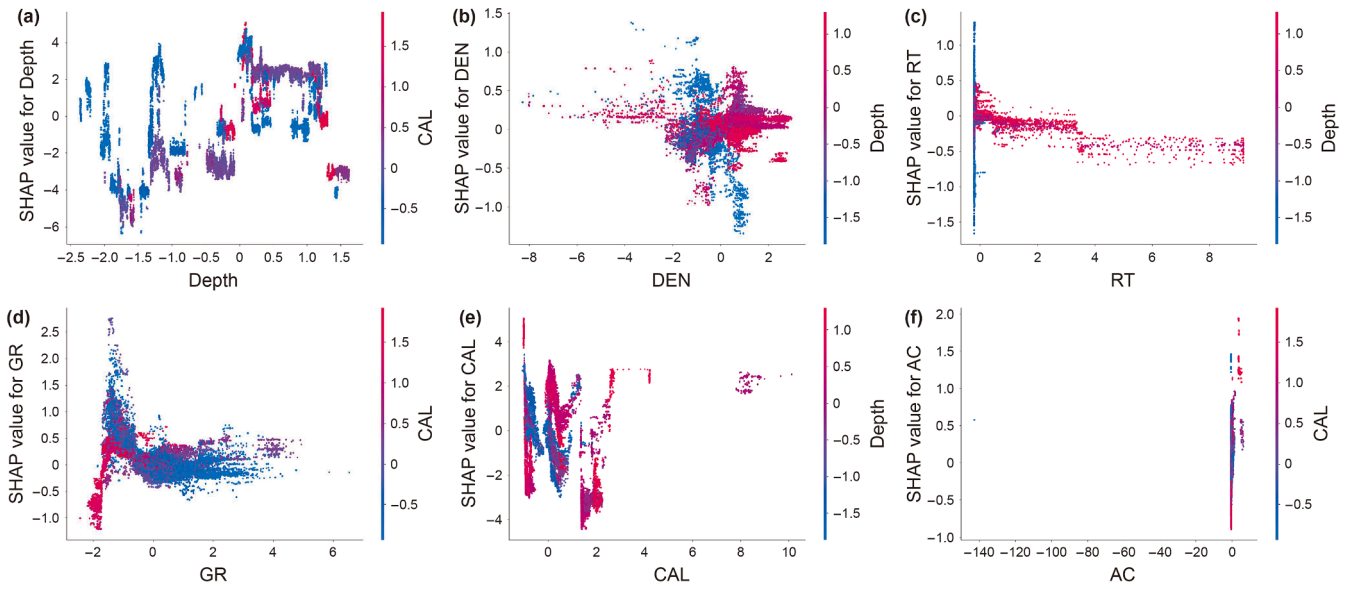

Fig. 21. SHAP value of second-order interaction (class of commercial gas reservoir): **(a)** Depth-CAL, **(b)** DEN-Depth, **(c)** RT-Depth, **(d)** GR-CAL, **(e)** CAL-Depth, **(f)** AC-CAL.

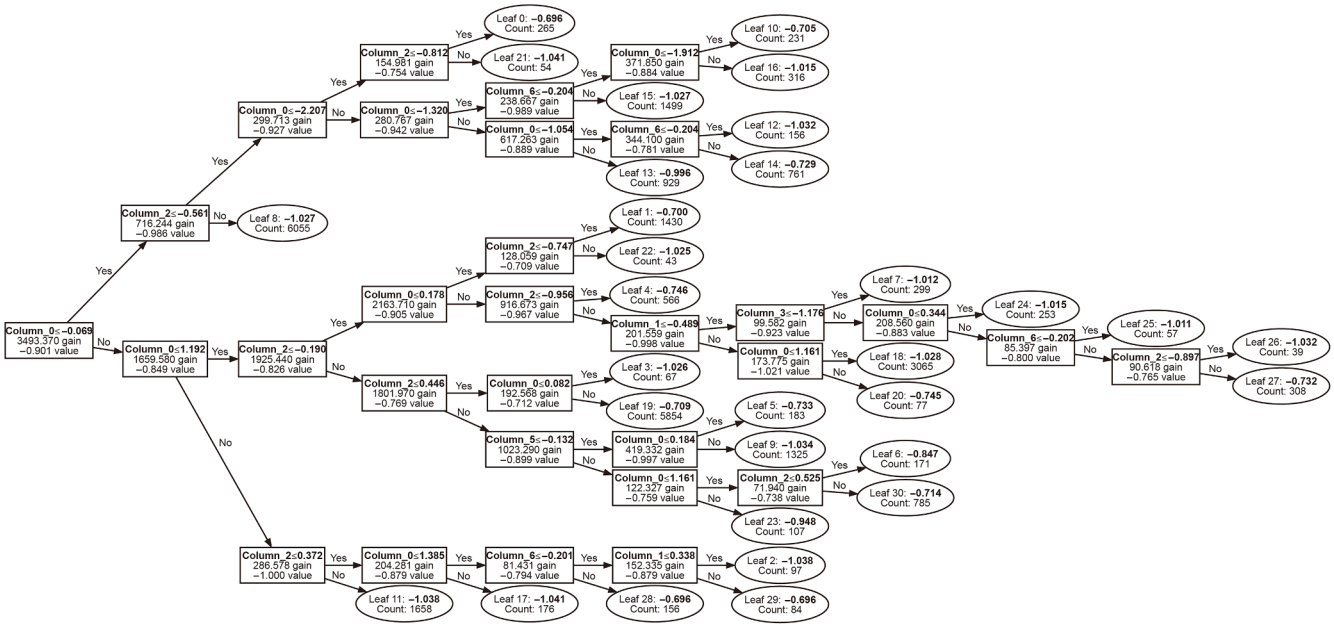

Fig. 22. Visualization of LightGBM's first decision tree.

Taking Well Zhe4 (Fig. 28) as an example, the RT log values remain elevated within the 4050–4312 m interval. Both well testing conclusions and ML predictions identify this interval as a commercial gas reservoir. At 4312 m, the RT value exhibits an abrupt decline, dropping to approximately 20 Ω·m by 4322 m, while well testing and model predictions classify the 4322–4412 m interval as a dry layer, aligning with geological identification criteria. The RT log reflects the true formation resistivity, where resistivity magnitude correlates with hydrocarbon saturation. When reservoirs contain oil or gas, resistivity increases due to the non-conductive nature of hydrocarbons. These results confirm that the ML model successfully captures complex nonlinear correlations between gas-bearing properties and logging curves, demonstrating high consistency with geological principles.

The normalized confusion matrix (Fig. 29(a)) and the unnormalized confusion matrix (Fig. 29(b)) demonstrate the predictive performance of the LightGBM model for each gas-bearing category prior to 10-fold cross-validation. Diagonal elements quantify the number and accuracy rates of correctly identified samples across six categories, including the commercial gas reservoir (Accuracy = 99.92%) and the gas-bearing water reservoir (Accuracy = 99.83%). Off-diagonal elements reveal misclassification patterns between different reservoirs. Notably, the misclassification rates between the gas-bearing water reservoir and the water-bearing gas reservoir are merely 0.17% and 1.09%, confirming the model's fine-grained resolution for transitional reservoirs.

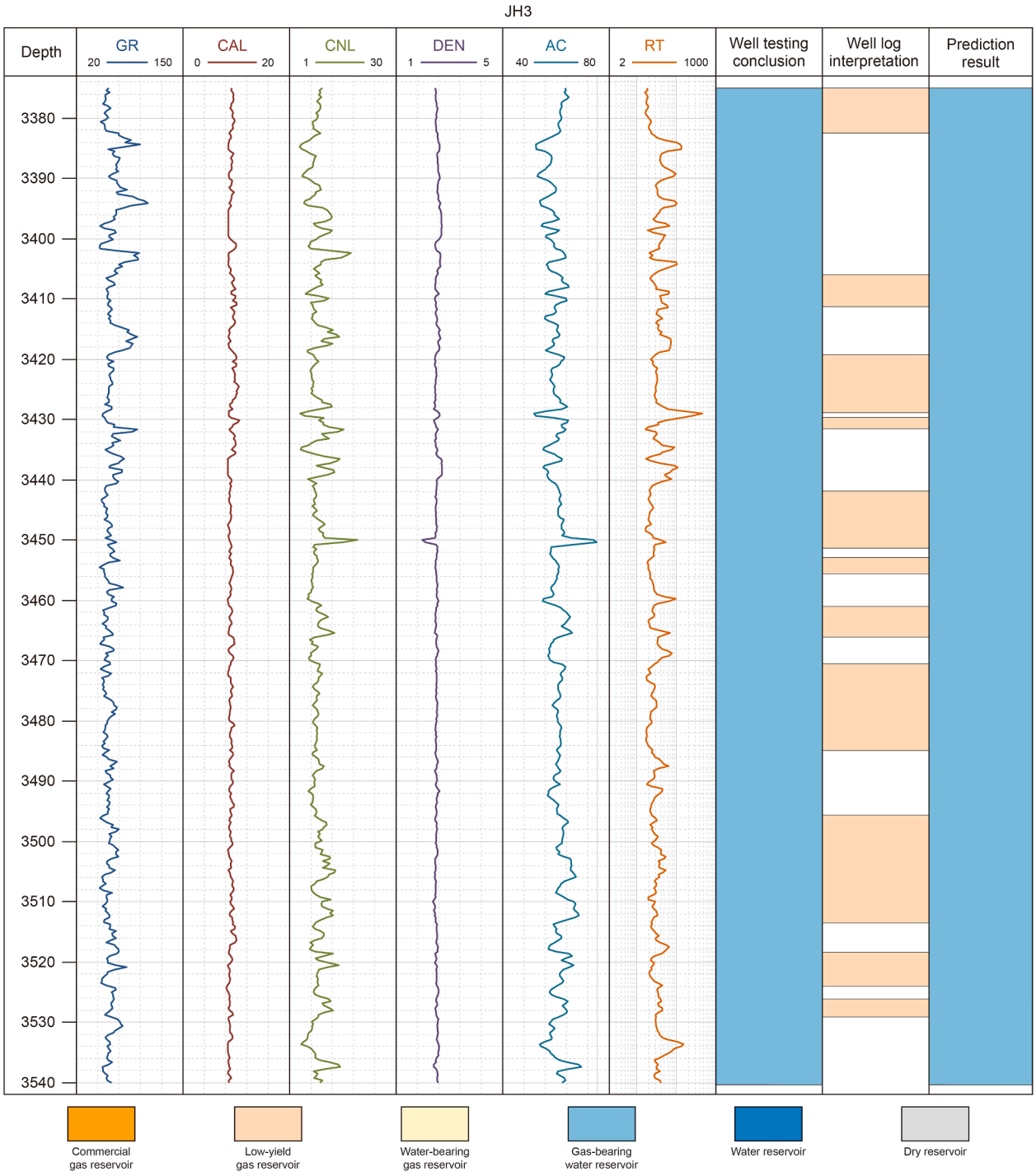

Fig. 23. Comparison of well testing, logging interpretation and ML prediction results of JH3.

5. Discussion

5.1. High-precision interpretable ML framework

This study's core contribution is the proposal of a novel predictive framework.

Unlike traditional methods that manually classify seismic data as gas or no gas based on logging curves, this framework directly links gas-bearing properties to logging data through well testing results. By eliminating intermediate data or labels, it preserves all gas-bearing information. The high vertical resolution of logging data allows the model to achieve decimeter-level accuracy. This ML framework is advanced and interpretable, and offers a cost-effective method for gas-bearing prediction, providing both theoretical and practical value in addressing global gas prediction challenges in tight sandstone reservoirs.

Having established the overall advantage of the framework, the following sections will sequentially validate its key properties.

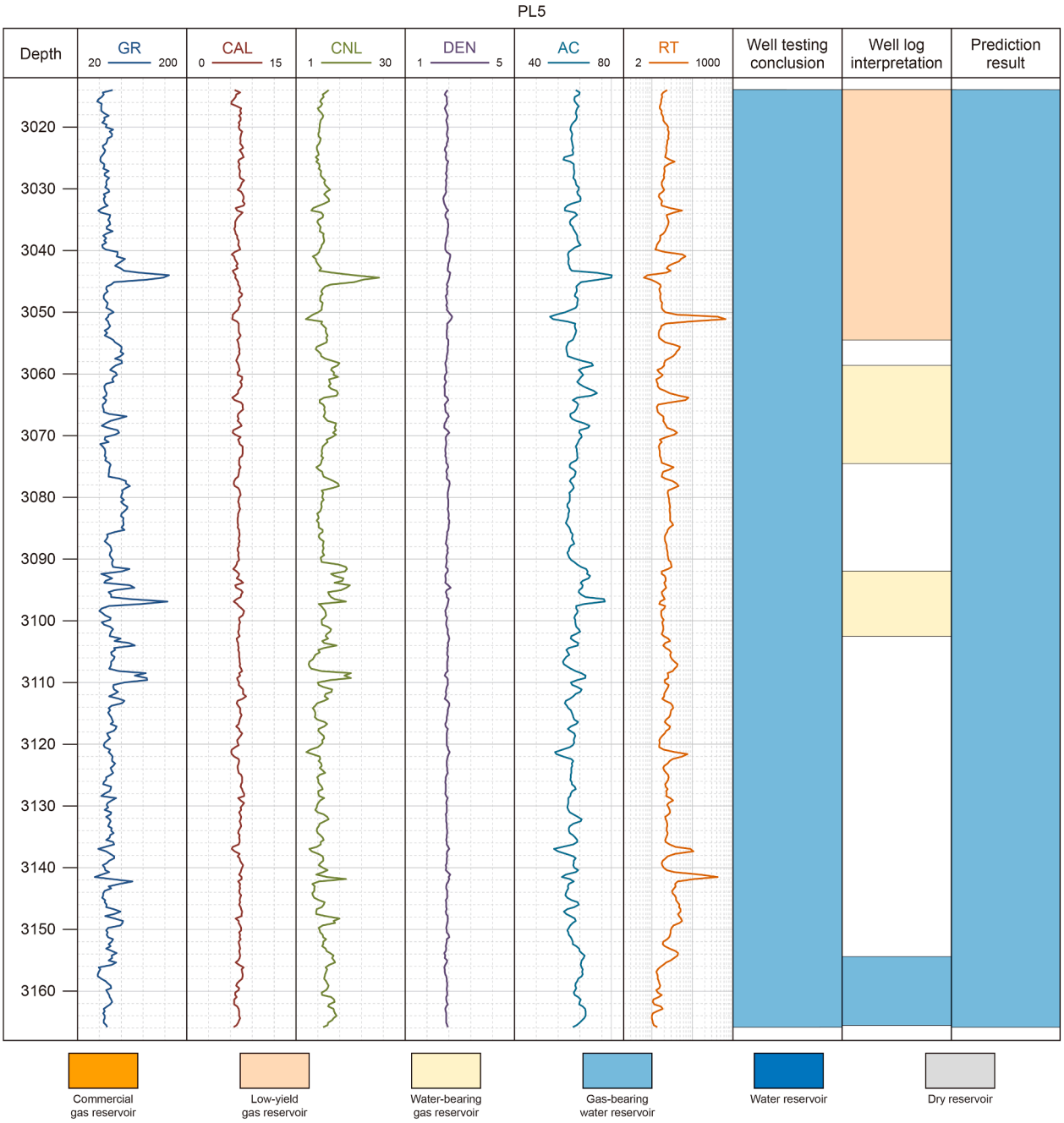

Fig. 24. Comparison of well testing, logging interpretation and ML prediction results of PL5.

5.2. Strong generalization and potential for broader transferability

The reliability of the framework is first demonstrated by its generalization ability.

LightGBM reached 99.76% Accuracy (WCEL = 0.0133, FL = 0.4436) after 10-fold cross-validation, while DeepFM achieved 97.79% Accuracy (WCEL = 0.1616, FL = 0.0205) among non-ensemble algorithms. Crucially, the performance of these algorithms remained stable or improved after 10-fold cross-validation compared to the initial training. These results highlight their strong generalization capabilities and potential for broader application to new data. The robust performance suggests that the framework may possess the capability for expansion to other blocks or similar geological settings, indicating promising prospects for broader application.

To understand the drivers of this robust performance, we next examine how the chosen algorithms align with the distinctive characteristics of our dataset.

5.3. Tree-based algorithms are more suitable for moderately imbalanced datasets

The superior performance of tree-based algorithms, particularly LightGBM, can be partly attributed to their inherent suitability for handling the moderately imbalanced dataset used in this study (as noted in Section 3.1). Traditional accuracy metrics are often misleading under such conditions. Therefore, we employed WCEL, which assigns higher weights to minority classes, and FL, which focuses on hard-to-classify samples, for a more nuanced evaluation.

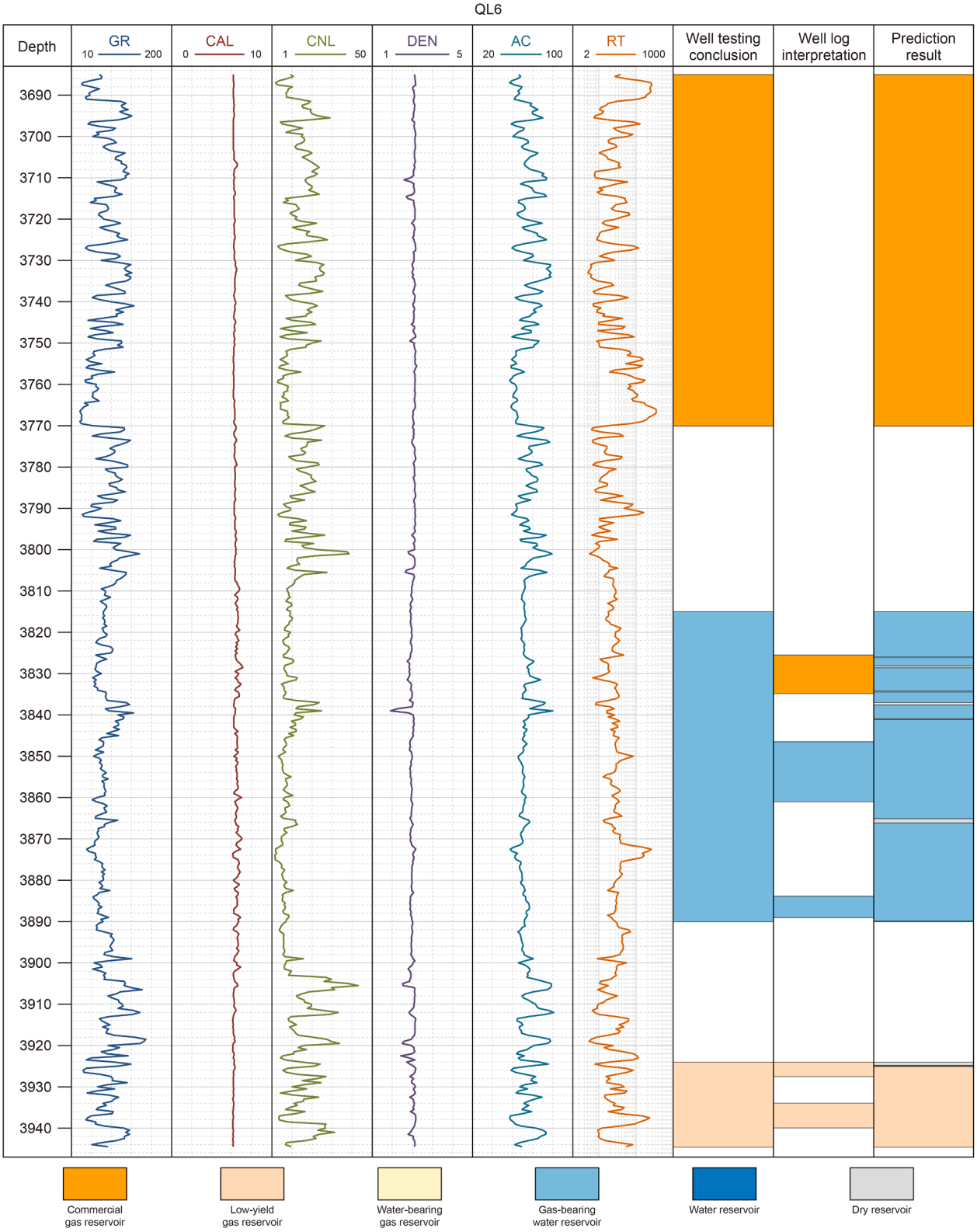

Fig. 25. Comparison of well testing, logging interpretation and ML prediction results of QL6.

The results (Tables 3 and 4) reveal a distinct pattern: tree-based ensemble methods (e.g., LightGBM, XGBoost) consistently achieve lower (better) WCEL values compared to deep learning models. In contrast, some deep learning models (e.g., DeepFM) attain lower FL values. Given that our dataset is moderately imbalanced rather than extremely so, WCEL is deemed the more appropriate metric for this context. The superior WCEL performance of tree-based algorithms indicates their effectiveness in learning a balanced

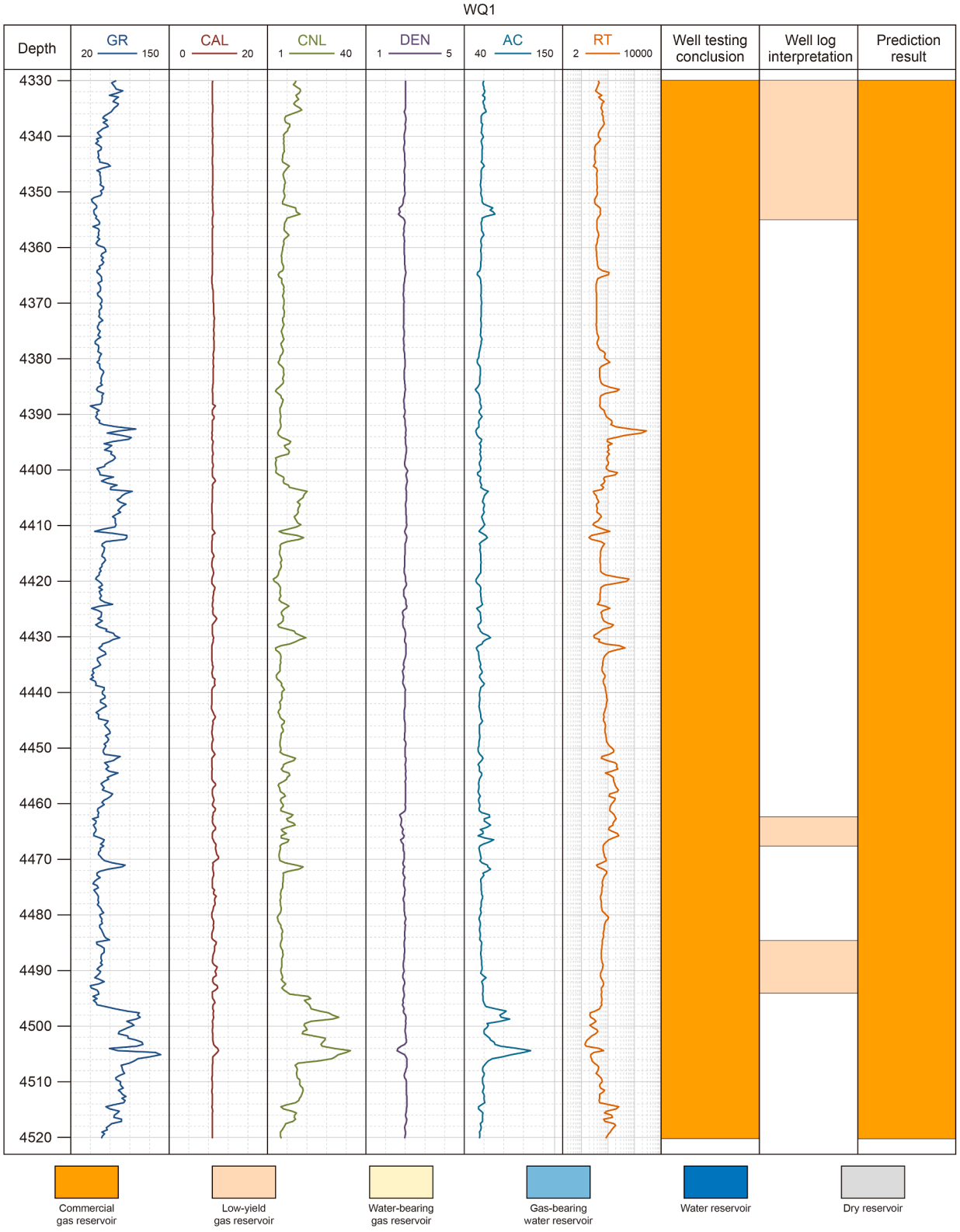

Fig. 26. Comparison of well testing, logging interpretation and ML prediction results of WQ1.

representation from all classes in a moderately imbalanced setting.

This advantage of tree-based models is rooted not only in their handling of class imbalance but also in fundamental characteristics that align well with the tabular, geoscientific nature of our data. The following interpretability analysis will shed light on how these characteristics manifest in the model's decision-making process.

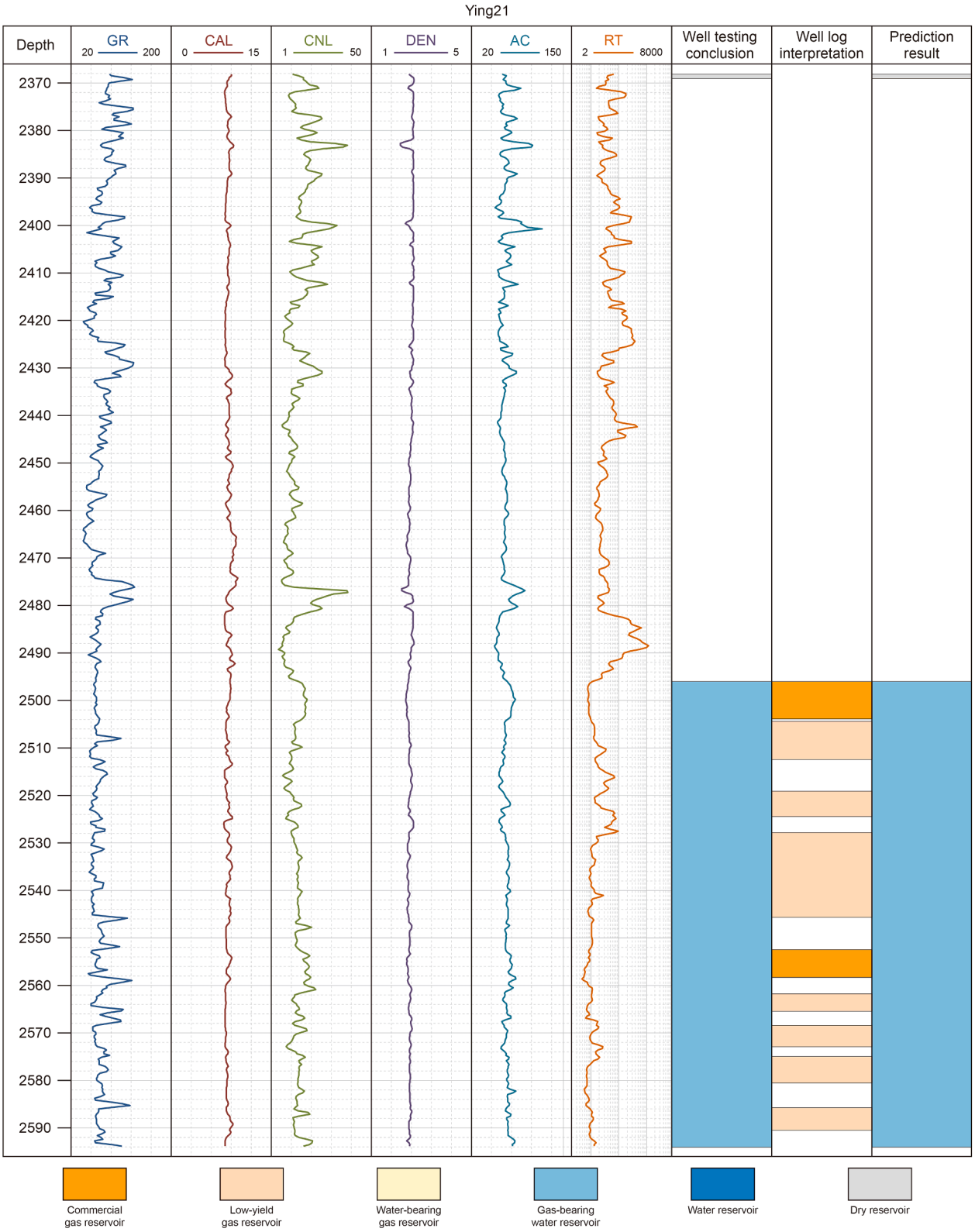

Fig. 27. Comparison of well testing, logging interpretation and ML prediction results of Ying21.

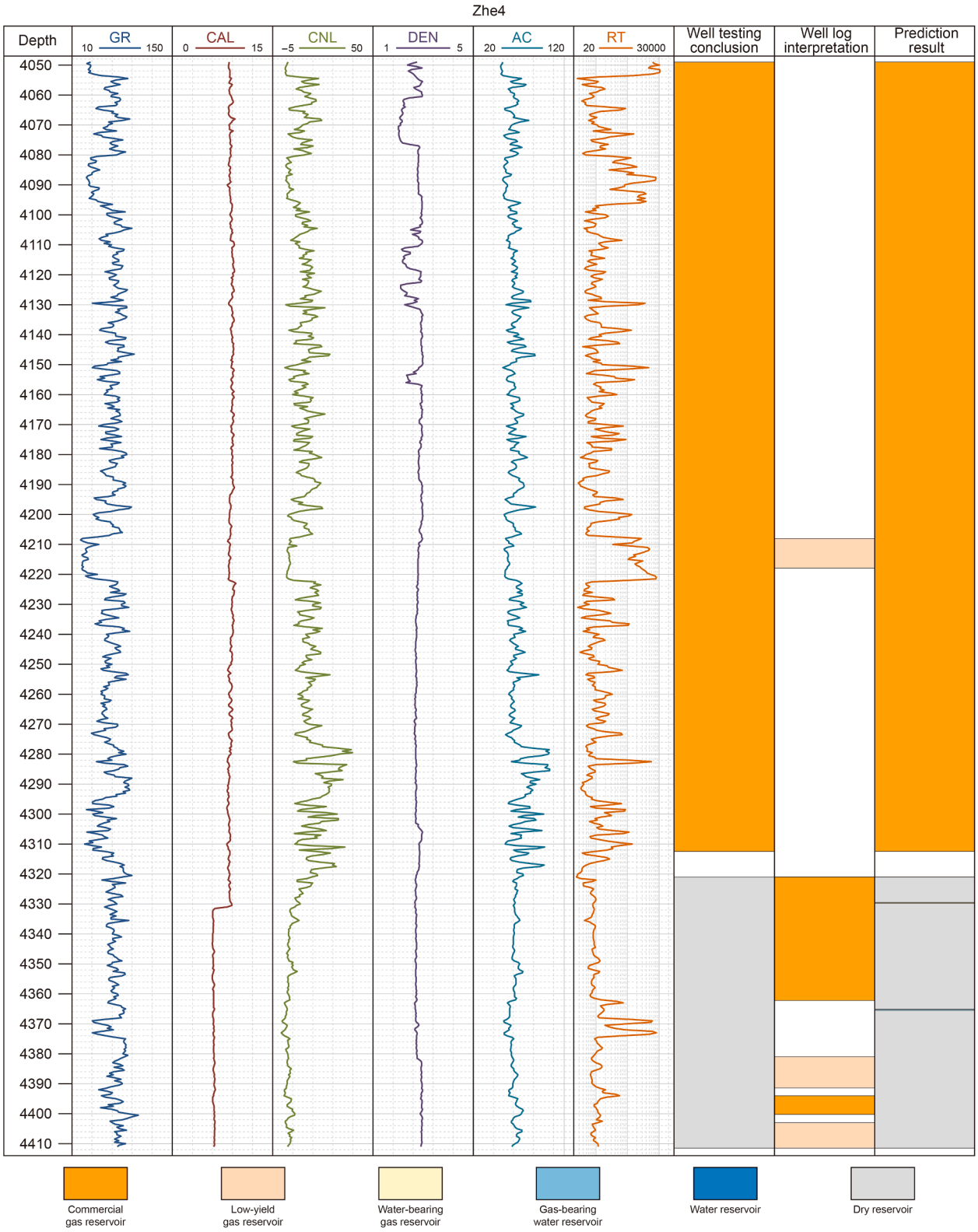

Fig. 28. Comparison of well testing, logging interpretation and ML prediction results of Zhe4.

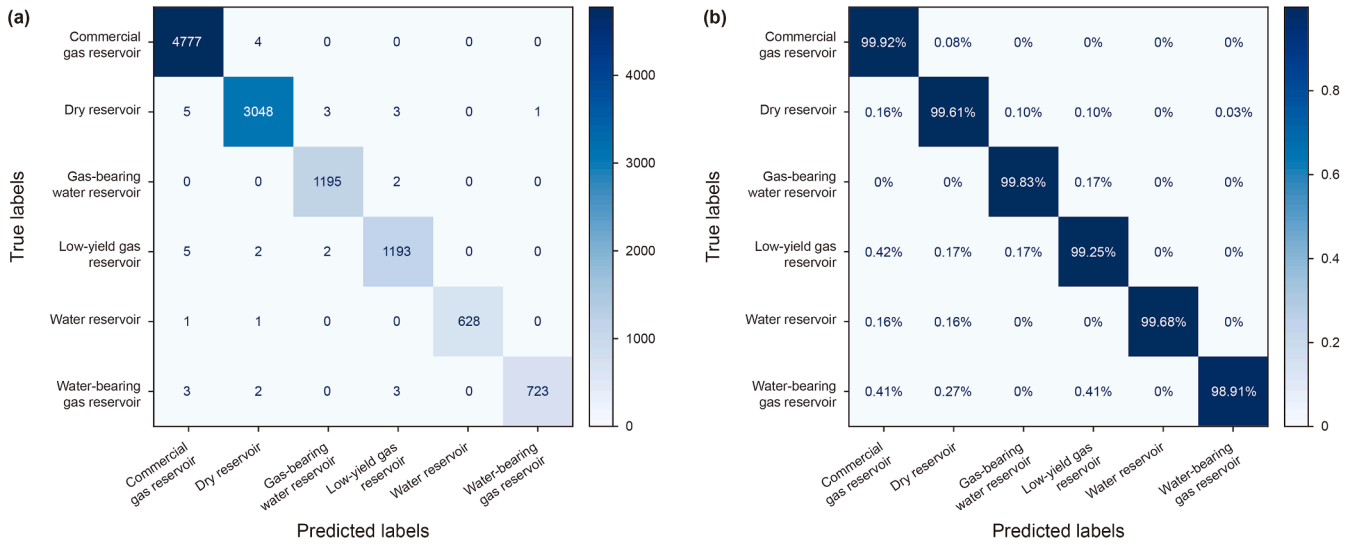

Fig. 29. Confusion matrix of LightGBM before 10-fold cross-validation: **(a)** normalized confusion matrix, **(b)** unnormalized confusion matrix.

5.4. Interpretability reports aligned with geological knowledge are crucial

The value of this framework lies not only in its high accuracy but also in the consistency of its decision-making process with geological knowledge, which enhances the credibility of the results.

By examining the Mean (|SHAP value|), we can identify which features significantly influence various model classifications and which are critical for specific classifications. In logging, the RT logging curve is crucial for reflecting reservoir fluid properties and identifying oil, gas, water, and dry zones (Ellis, 2007; Wyllie, 1963). Fig. 19 shows that RT's importance increases from gas-rich to water-bearing categories (e.g., commercial gas reservoir; low-yield gas reservoir; gas-bearing water reservoir; water-bearing gas reservoir). For significant fluid property changes (e.g., water reservoir; dry reservoir), RT's importance ranks second among logging curves. CAL, a lithology logging curve that indicates rock type and reservoir properties (Ellis, 2007; Wyllie, 1963), thus ranking first in importance among logging curves, consistent with geological knowledge. Xujiahe Formation's unique phenomenon of sand-mud interlayers further highlights the importance of lithological discrimination between sandstone and mudstone (Jin et al., 2025; Wen et al., 2024). Porosity logging curves like CNL, DEN, and AC reflect the subsurface storage space for oil and gas resources (Ellis, 2007; Wyllie, 1963); it is reasonable for them to be ranked in the middle. This SHAP-based analysis provides a detailed interpretable report about feature impacts and classifications, enhancing model interpretability and credibility. These analyses align with geological understanding and support targeted model optimization or a focus on predictive features in future data collection.

Beyond individual feature importance, the interpretability analysis further reveals a deeper mechanism influencing predictions: the complex interactions between geological features.

5.5. Interaction information among geological features is significant

The analysis of feature importance naturally leads to an investigation of the relationships between features.

Tree-based algorithms and DeepFM demonstrate superior predictive performance by capturing feature interaction information, while algorithms that fail to do so perform poorly. As shown in Section 5.4 and Figs. 20 and 21, the interaction between CAL and GR has a significant effect on model predictions. The significant interaction between CAL and GR, despite GR's lower individual importance ranking, suggests that their combined state provides unique predictive information beyond what each feature offers alone. The complex interactions between geological features, driven by the mutual coupling of geological conditions, can influence predictions positively or negatively, underscoring the need to focus on feature interactions (Cao et al., 2025). Additionally, the Depth feature, which contains significant sedimentary evolution information (Wang et al., 2024), must be carefully considered in ML model predictions.

In summary, evidence from algorithm comparison to mechanistic analysis collectively points to one conclusion: tree-based models possess inherent advantages for the tabular data involved in this study.

5.6. Tree-based algorithms outperform neural networks for tabular data in oil and gas geology

Collectively, the findings from algorithm comparison, performance on imbalanced data, and interpretability analysis converge on an important, more general conclusion: tree-based models exhibit inherent advantages for the tabular data prevalent in petroleum geology.

The differences between tree-based algorithms and neural networks on tabular data are well-documented, with tree-based algorithms often outperforming neural networks even without considering their superior computational efficiency. Grinsztajn et al. (2022) demonstrated that, compared to image, speech, and audio datasets, the objective functions of tabular data—characterized by heterogeneous features, small sample sizes, and extreme values—are often non-smooth. Tree-based algorithms, which can learn piecewise constant functions, better handle these non-smooth objective functions, while neural networks tend to produce overly smooth solutions. Tabular datasets frequently include many non-informative features beyond redundancy, and neural networks lack robustness to such features. Additionally, while tabular data lacks rotational invariance, neural

networks are rotationally invariant (favoring features with less information), whereas tree-based models process each feature independently and thus lack rotational invariance (Ng, 2004; Grinsztajn et al., 2022). Comparisons of six traditional ML algorithms, seven deep learning algorithms, and six ensemble learning algorithms reveal that tree-based algorithms excel in both fitting and generalization, outperforming traditional and deep learning algorithms in original predictions and 10-fold cross-validation. Given their computational performance and the prevalence of tabular data in oil and gas geology, tree-based algorithms are recommended for modeling and prediction.

5.7. Limitations and generalizability of the interpretable framework

The interpretability results, particularly the SHAP-derived feature importance rankings, are intrinsically tied to the geological context and data distribution of the training data from the specific study area. The dominance of certain features (e.g., CAL, Depth) directly reflects the unique characteristics of the Xujiahe Formation in this region. Therefore, applying the trained model and its explanations directly to areas with divergent geological conditions would likely yield suboptimal predictions and non-transferable interpretations, as feature importance may vary considerably.

The primary contribution of this work, however, lies in the universally applicable methodological framework itself—not the specific model for this area. The key generalizable finding is the demonstration that this workflow enables the creation of transparent, trustworthy models whose decisions can be geologically validated. For new areas, the framework provides a template for building locally credible models by retraining on local data, underscoring the necessity of context-specific calibration in ML applications for geoscience.

6. Conclusions

(1) The core novelty of this study is the development of a novel, interpretable ML framework that, for the first time, enables semi-quantitative prediction of gas-bearing properties in tight sandstone reservoirs at decimeter-scale accuracy. This framework demonstrates high accuracy within the studied formation and presents a viable approach for future basin-scale applications.
(2) The framework's robustness is highlighted by its performance on imbalanced datasets. A key methodological insight is that WCEL and FL evaluation metrics are crucial for imbalanced multi-class problems. LightGBM performs best with moderately imbalanced datasets of this study.
(3) A distinctive feature of the framework is its integrated interpretability. Based on cooperative game theory (SHAP), it provides reports that clearly identify both generally important features and those specific to each class, improving trust and guiding future data collection.
(4) The framework's effectiveness stems from its ability to capture complex geological relationships. Notably, tree-based algorithms and DeepFM show superior performance by effectively learning feature interactions, which are significant for geological predictions.
(5) For tabular geoscientific data, tree-based algorithms are preferred due to their inherent advantages over neural networks, including adaptability to non-smooth functions and robustness to non-informative features, as evidenced by their superior accuracy and generalization in this study.

CRediT authorship contribution statement

Liu Cao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Fu-Jie Jiang:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition. **Zhang-Xing Chen:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition. **Li-Na Huo:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **Run-Hai Feng:** Writing – review & editing, Supervision, Methodology. **Di Chen:** Writing – review & editing, Supervision, Funding acquisition. **Meng-Yang Wang:** Writing – review & editing, Visualization. **Jian Li:** Writing – review & editing. **Yang Gao:** Writing – review & editing. **Ben-Jie-Ming Liu:** Writing – review & editing. **Yong Ma:** Project administration. **Xiao-Juan Wang:** Resources. **Zhi-Min Jin:** Resources. **Ao-Bo Zhang:** Resources.

Data availability

We confirm that all the data supporting the findings of this study are available within the study. The authors do not have permission to share data.

Declaration of competing interest

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

Acknowledgements

This work was supported by the National Science and Technology Major Project of China (No. 2025ZD1400400), National Natural Science Foundation of China (NSFC) (Nos. 42302142 and 42372147) and Collaborative Project between the PetroChina Southwest Oil & Gas Field Company and China University of Petroleum (Beijing) (HX20231362). We appreciate the Research Institute of Exploration and Development, PetroChina Southwest Oil & Gas Field Company for providing the data. We are also grateful to the reviewers for the helpful comments to improve our paper.

Appendix. Mathematical descriptions of algorithms

A1 DT

Table A1
The pseudocode of DT (Quinlan, 1986; Chou, 1991; Salzberg, 1994).

Algorithm decision tree
Input: Training dataset: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, $x_n, y_n \in R$
  if termination criteria met
    return base hypothesis $g_t(x)$
  else
    learn branching citeria $b(x)$
    split D to C parts $D_C = \{(x_n, y_n)$: $b(x_n) = c\}$
    build sub-tree $G_c \leftarrow$ Decision Tree (D_c)
  return $G(x) = \sum_{c=1}^{C} [b(x) = c] G_C(x)$

A2 KNN

Table A2
The pseudocode of KNN (Cover and Hart, 1967).

Algorithm K-Nearest Neighbors
Input: Training dataset: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, $x_n, y_n \in R$
$x_n \in R$ were variables, $y_n \in R = \{c_1, c_2, \ldots, c_k\}$ were values or classes of objective variables. $i = 1, 2, \ldots, N$.
Output: $y = \underset{cj}{\operatorname{argmax}} \sum_{x_i \in N_k(x)} I(y_i - c_i), i = 1, 2, 3 \ldots K$
I is the indicator function
$I(y_i - c_i) = \begin{cases} 1, & y_i = c_i \\ 0, & \text{otherwise} \end{cases}$

A3 AdaBoost

Table A3
The pseudocode of AdaBoost (Freund and Schapire, 1997).

Algorithm AdaBoost
Input: Training dataset: $D = \{(x_i, y_i)\}_{i=1}^{n}$, and $y_i \in \mathbb{C}$, $\mathbb{C} = \{c_1, \ldots, c_m\}$;
T: Number of iterations; I: Weak learner
Output: Boosted classifier:
$$H(x) = \underset{y \in \mathbb{C}}{\operatorname{argmax}} \sum_{t=1}^{T} \ln \frac{1}{\beta_t} [h_t(x) = y],$$
where h_t, β_t are the induced classifiers (with $h_t(x) \in \mathbb{C}$) and their assigned weights, respectively
$D_1(i) \leftarrow 1/N$ for $i = 1, \ldots, N$
for $t = 1$ to T **do**
$h_t \leftarrow I(S, D_T)$
$\varepsilon_t \leftarrow \sum_{i=1}^{N} D_t(i)[h_t(x_i) \neq y_i]$
if $\varepsilon_t > 0.5$ **then**
$T \leftarrow t - 1$
return
end if
$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$
$D_{t+1}(i) = D_t(i) \cdot \beta_t^{1-[h_t(x_i) \neq y_i]}$ for $i = 1, \ldots, N$
Normalize D_{t+1} to be a proper distribution
end for

A4 GBDT

Table A4
The pseudocode of GBDT (Friedman, 2001).

Algorithm GBDT
Input: Training dataset: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, $x_n, y_n \in R$
Initialize $f_{k0}(x) = 0, k = 1, 2, \ldots, K$.
for $m = 1$ **to** M:
Set
$$P_k(x) = \frac{e^{f_k(x)}}{\sum_{\ell=1}^{K} e^{f_k(x)}}, k = 1, 2, \ldots, K.$$
for $k = 1$ **to** K:
Compute $r_{ikm} = y_{ik} - p_k(x_i), i = 1, 2, \ldots, N$.
Fit a regression tree to the targets $r_{ikm}, i = 1, 2, \ldots, N$,
Giving terminal regions $R_{jkm}, j = 1, 2, \ldots, J_m$.
Compute
$$\gamma_{ikm} = \frac{K-1}{K} \frac{\sum_{x_i \in R_{jkm}} r_{ikm}}{\sum_{x_i \in R_{jkm}} |r_{ikm}|(1 - |r_{ikm}|)}, j = 1, 2, \ldots, J_m.$$
Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} r_{jm} I\left(x \in R_{jm}\right)$.
Output: $\hat{f}(x) = f_M(x)$

A5 RF

Table A5
The pseudocode of RF (Breiman, 2001a).

Algorithm Random Forests
Input: Training dataset: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, $x_n, y_n \in R$
Base learning algorithm ♀
Number of trainings T.
Algorithm:
1. For t = 1, 2, …, T, **do**
2. $h_t = L(D, D_{bs})$
3. End for
Output: $H(x) = \underset{y \in Y}{\operatorname{argmax}} \sum_{t=1}^{T} \prod(h_t(x) = y)$

A6 XGBoost

Table A6
The pseudocode of XGBoost (Chen and Guestrin, 2016).

Algorithm XGBoost (Exact Greedy Algorithm for Split Finding)
Input: Training dataset: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, $x_n, y_n \in R$
I, instance set of current nodes
d, feature dimension
$gain \leftarrow 0$
$G \leftarrow \sum_{i \in I} g_i, H \leftarrow \sum_{i \in I} h_i$
for $k = 1$ **to** d **do**
$G_L \leftarrow 0, H_L \leftarrow 0$
for j in *sorted* (I, by x_{jk}) **do**
$G_L \leftarrow G_L + g_j, H_L \leftarrow H_L + h_j$
$G_R \leftarrow G\text{-}G_L, H_R \leftarrow H\text{-}H_L$
score ← max (score, $\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda}$)
end
end
Output: Split with max score

A7 LightGBM

Table A7
The pseudocode of LightGBM (Ke et al., 2017).

Algorithm lightGBM (Gradient-based one-side sampling)
Input: I: training data, d: iterations
Input: a: sampling ratio of large gradient data
Input: b: sampling ratio of small gradient data
Input: loss: loss function, L: weak learner
models ← { }, fact ← $\frac{1-a}{b}$ (I)
topN ← $a \times$ len(I), randN ← $b \times$ len(I)
for $i = 1$ **to** d **do**
preds ← models.predict(I)
g ← loss (I, preds), w ← {1, 1, … }
sorted ← GetSortedIndices (abs(g))
topSet ← sorted [1:topN]
randSet ← RandomPick (sorted [topN:len(I)], randN)
usedSet ← topSet + randSet
w [randSet] × = fact ▷ Assign weight fact to the small gradient data.
newModel ← L (I [usedSet], − g [usedSet], w [usedSet])
models.append (newModel)

A8 CatBoost

Table A8
The pseudocode of CatBoost (Prokhorenkova et al., 2018).

Algorithm CatBoost
Input: Training dataset: $D = \{(x_i, y_i)\}_{i=1}^{n}$, I, a, L, s, $Mode$
$\sigma_r \leftarrow$ random permutation of $[1, n]$ for r = 0 … s;
$M_0\ (i) \leftarrow 0$ for $i = 1 \ldots n$;
if $Mode = Plain$ **then**
  $M_r\ (i) \leftarrow 0$ for $r = 1 \ldots s$, i: $\sigma_r\ (i) \le 2^{j+1}$;
if Mode = Ordered then
  for $j \leftarrow 1$ **to** $\lceil \log_2 n \rceil$ **do**
    $M_{r,j}\ (i) \leftarrow 0$ for $r = 1 \ldots s$, $i = 1.2^{j+1}$;
for $t \leftarrow 1$ **to** I **do**
  T_t, $\{M_r\}_{r=1}^{s} \leftarrow$ BuildTree ($\{M_r\}_{r=1}^{s}$, $\{(x_i, y_i)\}_{i=1}^{n}$, a, L, $\{\sigma_i\}_{i=1}^{s}$, Mode);
  $\text{leaf}_0\ (i) \leftarrow$ GetLeaf (x_i, T_t, σ_0) for $i = 1 \ldots n$);
  $\text{grad}_0 \leftarrow$ CalcGradient (L, M_0, y)
  foreach leaf j in T_t **do**
    $b_j^t \leftarrow$ -avg ($\text{grad}_0\ (i)$ for i: $\text{leaf}_0(i) = j$);
    $M_0(i) \leftarrow M_0(i) + a\, b^t_{\text{leaf}_0\ (i)}$ for $i = 1 \ldots n$;
return $F(x) = \sum_{t=1}^{I} \sum_j a b_j^t 1_{\{\text{GetLeaf}(x, T_t, \text{ApplyMode}) = j\}}$

A9 SHAP

Definition 1. SHAP interprets the Shapley value as an additive feature attribution method, explaining the model's predictive output as a linear function of binary variables:

$$\begin{cases} g(\boldsymbol{z}') = \phi_0 + \sum_{i=1}^{M} \phi_i \boldsymbol{z}_i', \\ \boldsymbol{z}' \in \{0,1\}^M, \phi_i \in R. \end{cases} \tag{A1}$$

Here,
f: the original prediction;
$\boldsymbol{x}$: the input features;
g: the explanation model;
M: the number of simplified input features;
ϕ_i: the effects of each feature (i.e., Shapely value).

The additive feature attribution method is the only solution that satisfies the following three desirable conditions:

Property 1. Local accuracy:

$$f(\boldsymbol{x}) = g(\boldsymbol{x}') = \phi_0 + \sum_{i=1}^{M} \phi_i \boldsymbol{z}_i', \tag{A2}$$

when approximating the original model f for specific input $\boldsymbol{x}$, local accuracy required that the explanation model must at least match the output of f for the simplified input $\boldsymbol{x}'$.

Property 2. Missingness:

$$\boldsymbol{x}_i' = 0 \Rightarrow \phi_i \neq 0, \tag{A3}$$

if the simplified input represents the presence or absence of features, the requirement of absence implies that features missing from the input should have no impact on the outcome.

Property 3. Consistency: Let $f_x(\boldsymbol{z}') = f(h_x(\boldsymbol{z}'))$ and $\boldsymbol{z}' \backslash i$ denote setting $\boldsymbol{z}_i' = 0$. For any two models f and f', *if*

$$f_x'(\boldsymbol{z}') - f_x'(\boldsymbol{z}' \backslash i) \ge f_x(\boldsymbol{z}') - f_x(\boldsymbol{z}' \backslash i), \tag{A4}$$

for all inputs $\boldsymbol{z}' \in \{0,1\}^M$, then $\phi_i(f', \boldsymbol{x}) \ge \phi_i(f, \boldsymbol{x})$.

Consistency requires that if the model changes, the contribution of the simplified input should increase or remain the same regardless of other inputs.

Theorem 1. Only one possible explanation model g that satisfies both definition 1, Local accuracy, Missingness and Consistency:

$$\begin{cases} \phi_i(f, x) = \sum_{\boldsymbol{z}' \subseteq \boldsymbol{x}'} \frac{|\boldsymbol{z}'|\,!\,(M - |\boldsymbol{z}'| - 1)!}{M!} [f_x(\boldsymbol{z}') - f_x(\boldsymbol{z}' \backslash i)], \\ \boldsymbol{z}' \subseteq \boldsymbol{x}'. \end{cases} \tag{A5}$$

Here,
ϕ_i: the effects of each feature (i.e., Shapely value);
$|\boldsymbol{z}'|$: the number of non-zero entries in $\boldsymbol{z}'$;
$\boldsymbol{z}' \subseteq \boldsymbol{x}'$: representing all $\boldsymbol{z}'$ vectors where the non-zero entries are a subset of the non-zero entries in $\boldsymbol{x}'$.

Theorem 1 is derived from the combinatorial results in cooperative game theory. Young (1985) demonstrated that the Shapley value is the unique solution that satisfies local accuracy, consistency, and a redundancy property.

References

Aras, S., Hanifi Van, M., 2022. An interpretable forecasting framework for energy consumption and CO_2 emissions. Appl. Energy 328, 120163. https://doi.org/10.1016/j.apenergy.2022.120163.
Bayes, F.R.S., 1958. An essay towards solving a problem in the doctrine of chances. Biometrika 45 (3–4), 296–315. https://doi.org/10.1093/biomet/45.3-4.296.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning, Information Science and Statistics. Springer, New York.
Bjorlykke, K., 2010. Petroleum Geoscience: From Sedimentary Environments to Rock Physics. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02332-3.
Breiman, L., 2001a. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.
Breiman, L., 2001b. Statistical modeling: the two cultures. Stat. Sci. 16 (3), 199–231. https://doi.org/10.1214/ss/1009213726.
Cao, L., Jiang, F., Chen, Z., et al., 2025. Data-driven interpretable machine learning for prediction of porosity and permeability of tight sandstone reservoir. Adv. Geo-Energy Res. 16 (1), 21–35. https://doi.org/10.46690/ager.2025.04.04.
Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785.
Cho, K., van Merrienboer, B., Bahdanau, D., et al., 2014. On the properties of neural machine translation: Encoder–decoder approaches. In: Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation. Association for Computational Linguistics, Doha, Qatar, pp. 103–111. https://doi.org/10.3115/v1/W14-4012.
Chou, P.A., 1991. Optimal partitioning for classification and regression trees. IEEE Trans. Pattern Anal. Mach. Intell. 13 (4), 340–354. https://doi.org/10.1109/34.88569.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. https://doi.org/10.1023/A:1022627411411.
Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inf. Theor. 13 (1), 21–27. https://doi.org/10.1109/TIT.1967.1053964.
Cox, D.R., 1958. The regression analysis of binary sequences. J. R. Stat. Soc., Ser. B Stat. Methodol. 20 (2), 215–232. https://doi.org/10.1111/j.2517-6161.1958.tb00292.x.
Dai, J., Ni, Y., Liu, Q., et al., 2021. Sichuan super gas basin in Southwest China. Petrol. Explor. Dev. 48 (6), 1251–1259. https://doi.org/10.1016/S1876-3804(21)60284-7.
Deng, J., Liu, M., Ji, Y., et al., 2022. Controlling factors of tight sandstone gas accumulation and enrichment in the slope zone of Foreland Basins: The Upper Triassic Xujiahe Formation in Western Sichuan Foreland Basin, China. J. Petrol. Sci. Eng. 214, 110474. https://doi.org/10.1016/j.petrol.2022.110474.
Dong, D.Z., Wang, Y.M., Li, D.H., et al., 2012. Global shale gas development revelation and prospect of shale gas in China. Eng. Sci. 14 (6), 69–76. https://doi.org/10.3969/j.issn.1009-1742.2012.06.010 (in Chinese).
Ellis, D.V., 2007. In: Well Logging for Earth Scientists, second ed. Springer, Dordrecht, The Netherlands.
Elman, J.L., 1990. Finding structure in time. Cogn. Sci. 14 (2), 179–211. https://doi.org/10.1207/s15516709cog1402_1.

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139. https://doi.org/10.1006/jcss.1997.1504.
Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. https://doi.org/10.1214/aos/1013203451.
Gao, J., Song, Z., Gui, J., et al., 2022. Gas-bearing prediction using transfer learning and CNNs: An application to a deep tight dolomite reservoir. IEEE Geosci. Remote Sens. Lett. 19, 3001005. https://doi.org/10.1109/LGRS.2020.3035568.
Geisser, S., 1975. The predictive sample reuse method with applications. J. Am. Stat. Assoc. 70 (350), 320–328. https://doi.org/10.1080/01621459.1975.10479865.
Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on typical tabular data?. In: Advances in Neural Information Processing Systems 35. Neural Information Processing Systems Foundation. Inc. (NeurIPS), pp. 507–520. https://doi.org/10.52202/068431-0037.
Guo, H., Tang, R., Ye, Y., et al., 2017. DeepFM: A factorization-machine based neural network for CTR prediction. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Melbourne, Australia, pp. 1725–1731. https://doi.org/10.24963/ijcai.2017/239.
Hammond, A.L., 1974. Bright spot: Better seismological indicators of gas and oil. Science 185 (4150), 515–517. https://doi.org/10.1126/science.185.4150.515.
He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Las Vegas, NV, USA, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
Jia, A., Wei, Y., Guo, Z., et al., 2022. Development status and prospect of tight sandstone gas in China. Nat. Gas Ind. B 9 (5), 467–476. https://doi.org/10.1016/j.ngib.2022.10.001.
Jiang, X., 2022. Study on Hydrocarbon Detection Method for Carbonate Rocks in Deep Leikoupo Formation of Western Sichuan. Ph.D. thesis. Chengdu University of Technology, Chengdu (in Chinese).
Jiang, Z., Ran, B., Li, Z., et al., 2023. A Late Triassic depositional age for the Xujiahe Formation, Sichuan Basin: Implications for the closure of the Paleo-Tethys Ocean. Mar. Petrol. Geol. 155, 106346. https://doi.org/10.1016/j.marpetgeo.2023.106346.
Jin, Z.M., Xie, J.R., Cao, Z.L., et al., 2025. Tight sandstone diagenesis, evolution, and controls on high-graded reservoirs in slope zones of foreland basins: A case study of the fourth member of Xujiahe Formation, Tianfu gas field, Sichuan Basin. China Geol. 8 (2), 325–337. https://doi.org/10.31035/CG20230072.
Júnior, D.A.D., Batista da Cruz, L., Otávio Bandeira Diniz, J., et al., 2023. Detection of potential gas accumulations in 2D seismic images using spatio-temporal, PSO, and convolutional LSTM approaches. Expert Syst. Appl. 215, 119337. https://doi.org/10.1016/j.eswa.2022.119337.
Karpatne, A., Ebert-Uphoff, I., Ravela, S., et al., 2019. Machine learning for the geosciences: Challenges and opportunities. IEEE Trans. Knowl. Data Eng. 31 (8), 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006.
Ke, G., Meng, Q., Finley, T., et al., 2017. LightGBM: A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems 30. Annual Conference on Neural Information Processing Systems 2017, December 4–9, 2017, pp. 3146–3154. Long Beach, CA, USA.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.
LeCun, Y., Bottou, L., Bengio, Y., et al., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324. https://doi.org/10.1109/5.726791.
Lin, T.Y., Goyal, P., Girshick, R., et al., 2020. Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. 42 (2), 318–327. https://doi.org/10.1109/TPAMI.2018.2858826.
Liu, G., Gong, R., Shi, Y., et al., 2022. Construction of well logging knowledge graph and intelligent identification method of hydrocarbon-bearing formation. Petrol. Explor. Dev. 49 (3), 572–585. https://doi.org/10.1016/S1876-3804(22)60047-8.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30. 31st Annual Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, California, USA, 4–9 December 2017. Neural Information Processing Systems Foundation. Inc. (NeurIPS), pp. 4766–4775.
National Energy Administration of China, 2021. SY/T 6293–2021 specification for well test in exploration. Natl. Energy Adm. China 1–20 (in Chinese).
Ng, A.Y., 2004. Feature selection, L_1 vs. L_2 regularization, and rotational invariance. In: Proceedings of the Twenty-First International Conference on Machine Learning. ACM Press, Banff, Alberta, Canada. https://doi.org/10.1145/1015330.1015435.
Ostrander, W.J., 1984. Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence. Geophysics 49 (10), 1637–1648. https://doi.org/10.1190/1.1441571.
Prokhorenkova, L., Gusev, G., Vorobev, A., et al., 2018. CatBoost: Unbiased boosting with categorical features. In: Advances in Neural Information Processing Systems 31. Annual Conference on Neural Information Processing Systems 2018 (NeurIPS 2018), December 3–8, 2018, Montreal, Canada, pp. 6639–6649.
Qin, X.Y., Xiao, L.Z., Zhang, Y.Z., 2005. Methods of natural gas reservoir identification and evaluation of Erdos Basin. Prog. Geophys. 20 (4), 1099–1107. https://doi.org/10.3969/j.issn.1004-2903.2005.04.034 (in Chinese).
Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1 (1), 81–106. https://doi.org/10.1007/BF00116251.
Rendle, S., 2010. Factorization machines. In: 2010 IEEE International Conference on Data Mining. IEEE, Sydney, Australia, pp. 995–1000. https://doi.org/10.1109/ICDM.2010.127.
Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, San Francisco, California, USA, pp. 1135–1144. https://doi.org/10.1145/2939672.2939778.
Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1 (5), 206–215. https://doi.org/10.1038/s42256-019-0048-x.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323 (6088), 533–536. https://doi.org/10.1038/323533a0.
Salzberg, S.L., 1994. C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. Mach. Learn. 16 (3), 235–240. https://doi.org/10.1007/BF00993309.
Seydoux, L., Balestriero, R., Poli, P., et al., 2020. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. Nat. Commun. 11, 3972. https://doi.org/10.1038/s41467-020-17841-x.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
Song, Z., Li, S., He, S., et al., 2022. Gas-bearing prediction of tight sandstone reservoir using semi-supervised learning and transfer learning. IEEE Geosci. Remote Sens. Lett. 19, 3007205. https://doi.org/10.1109/LGRS.2022.3177314.
Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc., Ser. B Stat. Methodol. 36 (2), 111–133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x.
Wang, P.Y., Bing, Z.Y., 2017. Distribution characteristics and regularities of tight sandstone gas resources in the world. China's Manganese Ind. 35 (6), 23–29. https://doi.org/10.14101/j.cnki.issn.1002-4336.2017.06.006 (in Chinese).
Wang, S., Li, X.Y., Di, B., et al., 2010. Reservoir fluid substitution effects on seismic profile interpretation: A physical modeling experiment. Geophys. Res. Lett. 37 (10), L10306. https://doi.org/10.1029/2010GL043090.
Wang, W., Pang, X., Chen, Z., et al., 2020. Improved methods for determining effective sandstone reservoirs and evaluating hydrocarbon enrichment in petroliferous basins. Appl. Energy 261, 114457. https://doi.org/10.1016/j.apenergy.2019.114457.
Wang, H., Fu, T., Du, Y., et al., 2023. Scientific discovery in the age of artificial intelligence. Nature 620 (7972), 47–60. https://doi.org/10.1038/s41586-023-06221-2.
Wang, F., Chen, D., Li, M., et al., 2024. A novel method for predicting shallow hydrocarbon accumulation based on source-fault-sand (S-F-Sd) evaluation and ensemble neural network (ENN). Appl. Energy 359, 122684. https://doi.org/10.1016/j.apenergy.2024.122684.
Wen, L., Zhang, B., Jin, Z., et al., 2024. Accumulation sequence and exploration domain of continental whole petroleum system in Sichuan Basin, SW China. Petrol. Explor. Dev. 51 (5), 1151–1164. https://doi.org/10.1016/S1876-3804(25)60532-5.
Wyllie, M.R.J., 1963. The Fundamentals of Well Log Interpretation. Academic Press, New York.
Xiang, T., Cao, J., Zhao, L., et al., 2024. Gas-bearing prediction in tight sandstone reservoirs based on multinetwork integration. Interpretation 12 (2), T177–T185. https://doi.org/10.1190/INT-2023-0091.1.
Yang, Y., Wen, L., Zhou, G., et al., 2023. New fields, new types and resource potentials of hydrocarbon exploration in Sichuan Basin. Acta Pet. Sin. 44, 2045–2069. https://doi.org/10.7623/syxb202312004 (in Chinese).
Yilmaz, Ö., 2001. Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data. Society of Exploration Geophysicists, Tulsa, OK. https://doi.org/10.1190/1.9781560801580.
Young, H.P., 1985. Monotonic solutions of cooperative games. Int. J. Game Theor. 14 (2), 65–72. https://doi.org/10.1007/BF01769885.
Yuan, Z., 2017. Interpretation Methods of Tight Sandstone Reservoir with Seismic Data and Well Logs Based on Machine Learning Method and Multi-Information Fusion. Ph.D. thesis. China University of Geosciences, Beijing), Beijing (in Chinese).
Yuan, J., Luo, D., Feng, L., 2015. A review of the technical and economic evaluation techniques for shale gas development. Appl. Energy 148, 49–65. https://doi.org/10.1016/j.apenergy.2015.03.040.
Zhang, Y., 2018. Study of Reservoir Evaluation and gas-bearing Prediction of Carbonate Reservoir in the Eastern Area of Sulige Gas Field. Ordos Basin. Ph.D. thesis. China University of Geosciences, (Beijing), Beijing (in Chinese).
Zhang, G., Wang, Z., Mohaghegh, S., et al., 2021. Pattern visualization and understanding of machine learning models for permeability prediction in tight sandstone reservoirs. J. Petrol. Sci. Eng. 200, 108142. https://doi.org/10.1016/j.petrol.2020.108142.
Zhao, C., Jiang, Y., Wang, L., 2022a. Data-driven diagenetic facies classification and well-logging identification based on machine learning methods: a case study on Xujiahe tight sandstone in Sichuan Basin. J. Petrol. Sci. Eng. 217, 110798. https://doi.org/10.1016/j.petrol.2022.110798.

Zhao, X., Chen, X., Huang, Q., et al., 2022b. Logging-data-driven permeability prediction in low-permeable sandstones based on machine learning with pattern visualization: a case study in Wenchang A sag, pearl river mouth Basin. J. Petrol. Sci. Eng. 214, 110517. https://doi.org/10.1016/j.petrol.2022.110517.

Zou, C.N., Zhu, R.K., Liu, K.Y., et al., 2012. Tight gas sandstone reservoirs in China: Characteristics and recognition criteria. J. Petrol. Sci. Eng. 88–89, 82–91. https://doi.org/10.1016/j.petrol.2012.02.001.

Zou, C.N., Zhai, G.M., Zhang, G.Y., et al., 2015. Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources. Petrol. Explor. Dev. 42 (1), 14–28. https://doi.org/10.1016/S1876-3804(15)60002-7.

Zou, C.N., He, D.B., Jia, C.Y., et al., 2021. Connotation and pathway of world energy transition and its significance for carbon neutral. Acta Pet. Sin. 42 (2), 233–247. https://doi.org/10.7623/syxb202102008 (in Chinese).

Zou, C.N., Lin, M.J., Ma, F., et al., 2024. Development, challenges and strategies of natural gas industry under carbon neutral target in China. Petrol. Explor. Dev. 51 (2), 476–497. https://doi.org/10.1016/S1876-3804(24)60038-8.