Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Learning Across Wells: A Robust Framework for Lithology Classification and Prediction

Yasir Bashir

dryasir.bashir@live.com

Istanbul Technical University

Zoonash Arshad

University of Houston

Hülya Kurt

Istanbul Technical University

Caner İmren

Istanbul Technical University

Muhsan Ehsan

Bahria University

Numair Ahmed Siddiqui

Universiti Teknologi PETRONAS

Research Article

Keywords: machine learning, Random Forest, K-Nearest Neighbours, lithology classification, well-log analysis, blind-well testing

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11229597/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Learning Across Wells: A Robust Framework for Lithology Classification and Prediction

Yasir Bashir*[, a], Zoonash Arshad [b], Hülya Kurt [a], Caner İmren [a], Muhsan Ehsan[c], Numair Ahmed Siddiqui[d]

[a] *Department of Geophysical Engineering, Faculty of Mines, Istanbul Technical University, Istanbul, Turkey*

[b] *Department of Earth and Atmospheric Sciences, University of Houston, Houston, Texas, USA*

[c] *Department of Earth and Environmental Sciences, Bahria University, Islamabad, Pakistan*

[d] *Geoscience Department, Universiti Teknologi PETRONAS, Seri Iskandar, Perak, Malaysia*

* Corresponding Author: Yasir Bashir, dryasir.bashir@live.com, ybashir@itu.edu.tr, 0000-0002-4614-796X

Keywords:

machine learning;

Random Forest;

K-Nearest Neighbours;

lithology classification;

well-log analysis;

blind-well testing;

ABSTRACT

Automated lithology classification from wireline logs plays a crucial role in enhancing subsurface characterization, reservoir delineation, and completion design efficiency. This paper develops two widely used supervised machine-learning classifiers, Random Forest (RF) and K-Nearest Neighbors (KNN), for predicting multi-well lithology. The evaluation uses four standard wireline logs: gamma ray (GR), bulk density (RHOB), corrected sonic transit time (DT_{CORR}), and effective porosity (PHIE), collected from four North Sea wells (F02-01, F03-02, F03-04, and F06-01). Four lithology classes Shale, Sandstone, Tight Sand, and Carbonate were used to conduct the evaluation of the models, which were conducted without core-calibrated ground truth or fixed petrophysical cut-offs. The assessment involved blind-well testing and Leave-One-Well-Out (LOWO) cross-validation. Hyperparameters were improved through the

application of stratified five-fold cross-validation. In the blind F06-01 well, Random Forest (RF) demonstrated an accuracy of 100.0% (weighted F1 = 1.000), in contrast to KNN's accuracy of 98.45%. Additionally, LOWO validation produced mean accuracies of 99.58% for RF and 97.10% for KNN. GR, RHOB, and PHIE were the primary features of significance, while DT_{CORR} had a negligible contribution. The models reached a consensus on 98.4% of blind-well samples, with RF accurately addressing all discrepancies. Even without feature scaling, Random Forest's full, feature-randomized ensemble structure surpasses the distance-based KNN classifier in inter-well log variability robustness. We address the limitations of rule-based labeling without core data, review machine learning research for lithology and facies classification, and propose resistivity logs, ensemble boosting, and core-calibrated labels for future research.

1. Introduction

Subsurface lithology identification is essential in groundwater hydrogeology and petroleum geoscience. Rock formation lithology affects mechanical properties, fluid storage capacity, permeability, and porosity, which affect reservoir characterisation, completion design, geomechanically modelling, and field development decisions. Lithology interpretation has traditionally been done by skilled petrophysicists using wireline logs, cross-plots, deterministic cut-offs, and qualitative pattern recognition after extensive fieldwork (Rider, 1996). This manual method is the gold standard when all the data and in-depth mineralogical analysis are available, but it's inefficient for large multi-well datasets, interpreter variability is a problem, and it's hard to maintain consistency across borehole environments like a large field or basin.

Over the past decade, supervised machine learning has become a potent technique for automating wireline log lithology and facies categorization. Support Vector Machines, K-Nearest Neighbours, Gradient Boosting, Random Forest, and Artificial Neural Networks have been tested with mixed results. K-Nearest Neighbours and Random Forest (RF) are supervised learning algorithms (Breiman, 2001; Chen & Guestrin, 2016; Cover & Hart, 1967; Hall, 2016). RF ensemble bagging combines multiple decision trees trained on bootstrapped data and randomly selected feature sets. KNN, an instance-based, non-parametric technique, assigns a new sample's class based on the feature space's majority class (Jamshidi Gohari et al., 2023; Pang et al., 2025). Geoscience applications benefit from both algorithms' interpretable outputs (feature importance for RF and local neighbourhood structure for KNN), minimal hyperparameter tuning, and the ability to represent non-linear log response-lithology correlations.

Despite their prominence, geoscience literature mostly compares RF and KNN in single-well or randomly split train/test investigations (Jiang et al., 2024). Direct and controlled comparisons under genuine multi-well situations, where a model trained on particular wells must generalize to unknown wells, are rare (H. Zhang et al., 2024; Z. Zhang et al., 2025). Random train/test splits inside a well tend to exaggerate model performance because close depth samples are highly autocorrelated. This lets the model memorize local log characteristics instead of lithology-based correlations across boreholes(Li et al., 2025; Sun et al., 2026a; Zhao et al., 2024). This paper examines how well the model predicts lithology in an unfamiliar well, not how well it aligns with logs. This question matters more in industry.

Machine learning for wireline log interpretation has grown due to the widespread adoption of data-driven methods in petroleum and groundwater geosciences during the past 15 years (Sun et al., 2026b; H. Zhang et al., 2024;

Zhao et al., 2024). Early Artificial Neural Network studies focused on facies and lithology categorization (Jiang et al., 2024; Mousavi & Hosseini-Nasab, 2024; Mukherjee et al., 2024; Pang et al., 2025). This work found that non-linear correlations between log responses and lithological classes outperformed linear discriminant or cross-plot-based cut-off approaches in sections with diverse geological features(Bestagini et al., 2017). However, ANN-based approaches were often criticized for being too dependent on training hyperparameters and network architecture and for making "black-box" conclusions. Researchers investigated numerous algorithm families to find a better balance between prediction strength and ease of comprehension.

Research showing SVMs' resistance to limited training data and capacity to construct non-linear decision boundaries using the kernel method without feature engineering has made them attractive lithofacies classification tools. In common wireline-log facies classification situations, SVM is competitive but not always better than tree-based ensemble techniques (Tschannen et al., 2017). Instead of being universal, these algorithms' rankings depend on the dataset and attributes used.

Breiman refined bagging and random subspace ensemble methods using Random Forest in the early 2000s (Breiman, 2001). It has since been frequently employed in petroleum geoscience. These include seismic facies categorization, crack identification, permeability and porosity estimation, and lithofacies prediction (Tewari & Dwivedi, 2019). The literature emphasizes its robustness to noisy or outlier-contaminated log measurements, insensitivity to input feature scale, and use of mean decrease in impurity to intrinsically prioritize features, which aids geological interpretation of model behavior. Multiple studies on multi-well lithology have demonstrated that Random Forest can classify with 85%–99% accuracy. These findings depend on the number of lithology classes, feature set, and validation method (within a well or across independently held-out wells) (Alaudah et al., 2019). Despite being more rigorous, the second validation approach always has worse accuracy than random, single-well train/test splits. In naive random splitting, spatial and depth autocorrelation may improve performance, according to the research.

In addition to its lengthy history as a general-purpose pattern recognition method, K-Nearest Neighbours is used in wireline log facies classification to compare more advanced algorithms. Multiple studies show that feature scaling and distance metric choices affect KNN performance. Its accuracy may also decrease when applied to wells with considerably different log properties from the training wells(Chawshin et al., 2021; Dunham et al., 2020) . Changes in tool calibration, mud weight on density logging, borehole circumstances, and depositional settings may induce this (Ao et al., 2018). This sensitivity makes Random Forest, Gradient Boosting, and SVM better than KNN

in multiple wells or cross-basin validation. KNN performs similarly or better in basic, single-well validation cases (Bressan et al., 2020). This research uses a controlled dataset of four North Sea wells to directly and statistically examine this tendency. It also explores feature relevance and sample-level model agreement, unlike most existing literature.

Deep learning architectures, gradient boosting frameworks (e.g., XGBoost and LightGBM), and multi-curve log sequences have enabled recent advances in automated lithology and facies prediction. Though it requires more training data, is more computationally intensive, and is less interpretable than Random Forest, some studies show it improves accuracy slightly. More complex designs are not examined in this research. A controlled comparison of two baseline classifiers with an emphasis on interpretability is our focus.

This research examines the central question of which classifier, Random Forest or K-Nearest Neighbours (Breiman, 2001; Mukherjee et al., 2024), yields more accurate and consistent lithology predictions when trained on wireline logs from multiple wells and evaluated on log data from a well that was entirely withheld from training. A secondary question pertains to interpretability: which wireline curves are most influential in lithology discrimination, and do the two algorithms concur on the relative significance of those curves? Responding to these inquiries holds significant practical importance for subsurface teams as they determine the appropriate algorithm to implement in production lithology-prediction pipelines. This is especially relevant in scenarios where core data is limited or absent, and the labels utilized for training are based on petrophysical cut-off rules instead of actual core descriptions.

1.1. Study Area Wells and Data

This research used four North Sea hydrocarbon-bearing basin wells: F02-01, F03-02, F03-04, and F06-01. These wells cover depth intervals from 29-30 m to 1,700-3,150 m and include 4,400–13,800 depth samples with complete (non-null) records across the four shared wireline curves used in this study: gamma ray (GR, API units), bulk density (RHOB, g/cc), sonic transit time corrected for cycle-skipping (DT_CORR, microseconds per foot), and effective porosity. These four curves were chosen because they were constant across all four wells, a common constraint in multi-well geoscience datasets. Variations in drilling schedule or operators cause logging suite variations. Three wells (F02-01, F03-02, F03-04) were chosen as the training set, while the fourth well (F06-01) was withheld as a blind test well, unseen by either model during training and hyperparameter tuning. Leave-One-Well-Out (LOWO)

cross-validation was then used for a more rigorous evaluation. Each of the four wells was withheld as the test well while the other three were used for training. This yielded four distinct train/test folds and a more statistically reliable inter-well generalization estimate.

Random Forest was chosen because of its reputation in applied geoscience machine learning literature as a dependable, low-maintenance baseline. It automatically handles non-linear and non-monotonic log curve-lithology class relationships, is robust to outliers, and does not require feature scaling. It also provides a cost-effective feature priority ranking based on the mean impurity decrease across its constituent trees. K-Nearest Neighbours was chosen as a contrasting, conceptually simple baseline that does not assume the functional structure of the feature-class label connection. Only local similarity in an appropriately scaled feature space is used. KNN's predictions are more affected by measurement scale and systematic inter-well log characteristics (such as tool calibration, mud weight impacts on density, or borehole-specific sonic cycle-skipping) than RF, which divides feature space through axis-aligned, scale-invariant threshold splits. This study uses multi-well log data to test the theoretical expectation in practice.

2. Methodology

Wireline log data for the four study wells were loaded and examined for completeness across the four common curves of interest: GR, RHOB, DT_CORR, and PHIE. Null values, represented in the source files by the standard LAS sentinel value of -999.25, were substituted with NaN and subsequently excluded from analysis through row-wise deletion on the feature columns. Following the cleaning process, the four wells retained a total of complete depth samples ranging from approximately 4,400 to 13,800, as summarized in Table 1. Depth ranges exhibited significant variation among the wells, with the shallowest commencing at approximately 29-30 m below the reference datum and reaching a maximum logged depth of 3,150 m in well F02-01. This variation is indicative of differing well objectives, completion intervals, and logging programs throughout the study area.

Table 1. Summary of well-log datasets used in this study, after removal of incomplete depth samples.

Well	Role	Depth Range (m)	Complete Samples	Mean GR (API)
F02-01	Training	30 – 3,150	4,418	59.0
F03-02	Training	30 – 2,140	13,771	39.8
F03-04	Training	30 – 1,900	11,789	46.7
F06-01	Blind test	29 – 1,700	10,826	47.4

No core descriptions or laboratory-derived mineralogical ground truth were accessible for any of the four study wells. In this operational scenario, a consistent set of petrophysical cut-offs was uniformly applied across all wells to assign a categorical lithology label to each depth sample. This approach ensured that any classifier trained on these labels was learning a well-defined, reproducible target function rather than an arbitrary or well-specific labeling convention. Four lithology classes were established: Shale, Sandstone, Tight Sand, and Carbonate. The decision logic, applied in sequence to each depth sample, is outlined in Table 2

Table 2. Rule-based petrophysical cut-offs are used to assign lithology labels in the absence of core data.

Lithology	GR (API)	RHOB (g/cc)	PHIE (frac.)	Label Code
Shale	≥ 70	—	—	0
Sandstone	< 70	< 2.20	≥ 0.25	1
Tight Sand	< 70	2.20 – 2.60	< 0.25	2
Carbonate	< 35	≥ 2.60	—	3

The petrophysical convention classifies samples with GR ≥ 70 API as shale, as they contain radioactive clay minerals and have higher natural gamma radiation than clean reservoir lithologies. Since limestone and dolomite have a higher matrix density than siliciclastic lithologies, samples with GR below 35 API and RHOB above 2.60 g/cc were categorized as carbonate. Sandstone samples were clean, well-sorted, high-porosity reservoir-quality sand with a RHOB of less than 2.20 g/cc and a PHIE of 0.25 or higher. Samples that did not meet Sandstone requirements were classified as residual Tight Sand. This categorization includes compacted, cemented, or partially shaly sand intervals with intermediate density and lower porosity than clear Sandstone. This labeling logic was applied equally across all four wells, preventing well-specific bias in the target variable definition.

Consider that this labeling approach simplifies lithological complexity and does not account for diagenetic effects, grain-size variation within a lithology class, or transitional/interbedded facies. In Section 4.4, we revisit this constraint. Since the research aims to compare classifier performance in a controlled context, the rule-based technique provides a consistent, reproducible, and visible target variable.

2.1. Feature Set

For this model, we utilized four wireline curves: gamma ray (GR), bulk density (RHOB), sonic transit time adjusted for cycle-skipping (DT_CORR), and effective porosity (PHIE). These curves were found in all the wells in

the study and were considered to be of consistent quality. No new engineered features, such as log-derivative curves, rolling-window statistics, or cross-curve ratios, were introduced to ensure that the impact of classifier selection was distinct from the influence of feature engineering. This investigation did not include resistivity logs because, despite their widespread usage in lithology and fluid-type discrimination, they were not consistently available across all four wells.

2.2. Train/Test Partitioning Strategies

Blind-Well Holdout

The main evaluation strategy adhered to established industry practices for determining whether a lithology classifier generalizes across different borehole environments, rather than simply conforming to the specific log characteristics of wells previously encountered during training. Three wells (F02-01, F03-02, F03-04) were combined to create a training dataset consisting of 29,978 samples. The fourth well, F06-01, includes 10,826 complete samples and was entirely withheld as a blind test set, ensuring it was never used in any capacity during model training or hyperparameter tuning. This partitioning directly simulates the operational scenario of implementing a lithology model, developed using offset wells, on a newly drilled well before or in the absence of core acquisition.

Leave-One-Well-Out Cross-Validation

Although the blind-well holdout offers a singular, realistic generalization estimate, it represents merely one specific selection of the well that is excluded. To achieve a more statistically robust and independent estimate of generalization performance, a Leave-One-Well-Out (LOWO) cross-validation scheme was implemented. In this scheme, each of the four wells was sequentially designated as the test well, while the other three wells were aggregated to create the training set. This process resulted in a total of four independent train/test folds. Both Random Forest and KNN were retrained from scratch within each fold utilizing the identical hyperparameters chosen during the blind-well tuning stage (Section 2.5). Weighted accuracy and weighted F1-score were documented for each fold. The average and per-well performance across all four folds represents the most stringent evaluation of inter-well generalization demonstrated.

2.3. Model Architectures and Hyperparameter Tuning

Random Forest

Random Forest builds a collection of decision trees, where each tree is trained on a separate bootstrap sample of the training data. During the training process, each split in the trees is limited to a random subset of the features. In this case, the square-root heuristic was applied, setting max_features to 'sqrt', which means that each split only considered two out of the four available features. Final predictions are derived through a majority vote among all trees within the ensemble. This dual source of randomness — bootstrap row sampling and random feature subsetting at each split — serves to decorrelate the individual trees, thereby reducing the variance of the ensemble prediction in comparison to any single deep decision tree, while maintaining bias largely unchanged. This phenomenon exemplifies the classical bias-variance decomposition that underpins ensemble bagging methods. A minimum of two samples per leaf (min_samples_leaf = 2) was established to ensure a degree of regularization against overfitting to noise in individual leaves.

The number of trees in the forest (n_estimators) was optimized through stratified five-fold cross-validation on the combined training set, evaluating candidate values of 10, 25, 50, 100, 150, 200, and 300 trees. Stratified folds were implemented to maintain the class proportions of the imbalanced training data in each fold. Cross-validation accuracy remained largely consistent throughout the entire range examined, fluctuating by less than 0.0001 between the optimal and least effective ensemble sizes (all values falling between roughly 0.99975 and 0.99985); twenty-five estimators reached the peak cross-validation accuracy of 0.9999 (Figure 5). Twenty-five estimators were selected as the operating configuration for all subsequent Random Forest results reported in this study. This choice achieves equivalent cross-validated accuracy to significantly larger forests while incurring a fraction of the computational and memory cost. This characteristic is advantageous for production deployment on large multi-well datasets and further indicates that this four-feature, four-class classification problem is well within the modeling capacity of even a small Random Forest ensemble.

K-Nearest Neighbours

K-Nearest Neighbours classifies a query sample based on a distance-weighted vote from its K nearest neighbours in feature space. In this study, the Euclidean metric is employed, with neighbour contributions weighted by the inverse of their distance (weights = 'distance'). This approach ensures that closer neighbours have a proportionally greater influence on the predicted class compared to those that are more distant. Due to the sensitivity of Euclidean distance to the relative numerical scale of each input feature, all four wireline curves underwent standardization (zero mean,

unit variance) utilizing a StandardScaler that was fitted on the training data. This scaler was then applied to the held-out test data without refitting, thereby preventing any information leakage from the test set into the scaling transformation.

The parameter K, representing the number of neighbors, was adjusted using a stratified five-fold cross-validation method similar to that used in a Random Forest, exploring integer values of K ranging from 1 to 30. Cross-validation accuracy peaked at 0.9930 when K was set to 4. All tested K values achieved accuracy within roughly 0.0008 of each other, and the resulting curve displayed a non-monotonic, multi-modal pattern instead of a single smooth peak (Figure 6). K = 4 was chosen as the operating configuration for all subsequent KNN results reported in this study due to its highest single cross-validation score

2.4. Evaluation Metrics

Model performance was assessed through overall classification accuracy, alongside per-class precision, recall, and F1-score. The weighted average F1-score, adjusted for class support, was highlighted as the primary summary statistic to address the significant class imbalance in the dataset, where Sandstone and Tight Sand are notably more common than Shale, and Carbonate is exceedingly rare in the examined wells. Confusion matrices, displayed in both raw counts and row-normalised proportions, were implemented to illustrate the distinct patterns of class confusion demonstrated by each classifier. Feature importance was evaluated for Random Forest using the mean decrease in Gini impurity across all trees and splits. For KNN, permutation importance was employed, where each feature is shuffled independently, and the resulting decline in held-out accuracy is noted as the importance of that feature. This method is necessary for KNN since, unlike Random Forest, KNN does not have an inherent model-based understanding of feature importance

2.5. Software and Computational Environment

All analysis was conducted in Python with the help of scikit-learn for model training, hyperparameter tuning, and evaluation metrics; pandas and NumPy for data manipulation; and Matplotlib for visualization. Figures were exported at publication resolution (600 dots per inch, raster and vector PDF formats) through a standardized figure-export routine to maintain consistent typography, color palettes, and sizing across all plots included in this paper. A fixed

random seed (42) was utilized consistently for all stochastic operations, encompassing bootstrap sampling in Random Forest and cross-validation fold assignment, to guarantee the reproducibility of the reported results.

3. Results

3.1. Raw Well-Log Character and Multi-Well Overview

Figure 1 illustrates a composite log display for well F02-01, which is the longest and deepest well in the study dataset. It features continuous traces of gamma ray, bulk density, corrected sonic transit time, and effective porosity plotted against depth. Additionally, the shale gamma-ray cut-off (70 API) is overlaid as a reference line on the gamma ray track and shaded in areas where it is exceeded. The logging suite typically captures typical vertical heterogeneity, as seen in this single-well view. Intervals of reduced porosity and elevated GR and density are consistent with intervals dominated by shale, while intervals of lower density and higher porosity that occur between these two likely represent reservoir sand.

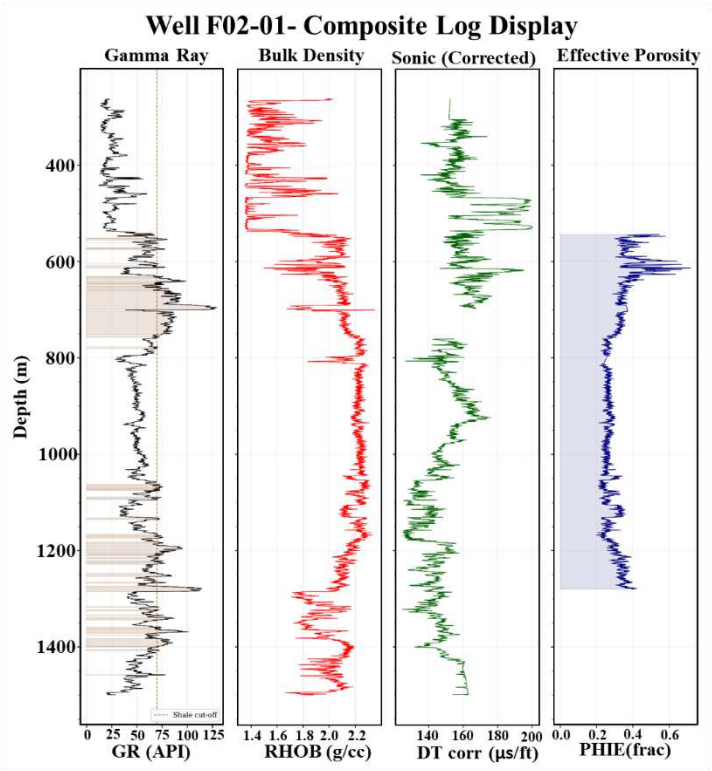


Figure 1: Composite well-log display for well F02-01 showing gamma ray (with shale cut-off reference line), bulk density, corrected sonic transit time, and effective porosity against depth.

Figure 2 presents a summary of the statistical distribution of the four wireline curves across all four study wells through the use of box plots, facilitating a direct visual comparison of the log characteristics between the wells. Although the four wells display generally comparable central tendencies for each curve, there are noticeable systematic differences. For instance, F02-01 shows a significantly higher median and a wider interquartile range in GR compared to the other three wells, aligning with its reported mean GR of 59.0 API, in contrast to the 39.8-47.4 API range observed in the remaining wells (Table 1). Inter-well shifts in log character present significant challenges for classifiers attempting to generalize from training wells to an unseen test well. This phenomenon is specifically what the blind-well and LOWO validation strategies in this study aim to investigate.

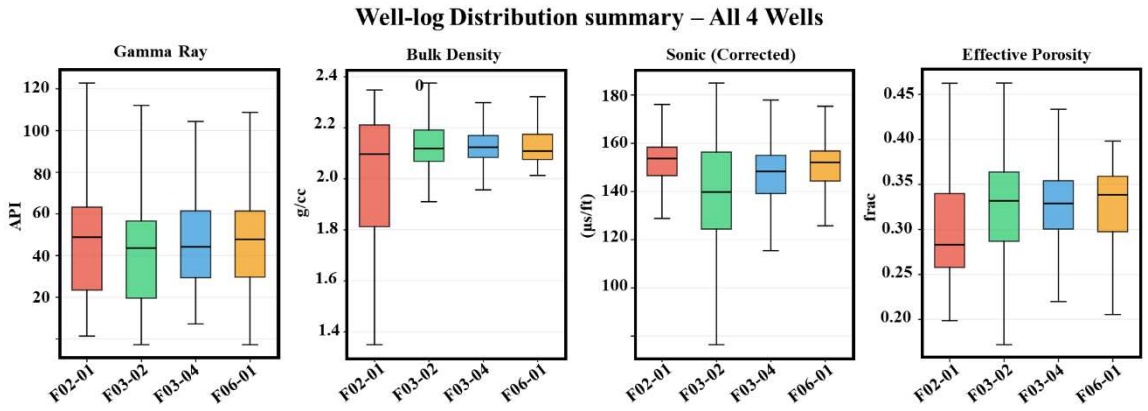


Figure 2: Box-plot summary of GR, RHOB, DT_CORR, and PHIE distributions across all four study wells, illustrating inter-well variability in log character.

Figure 3 presents a comprehensive four-by-four panel comparison of log traces, differentiated by their assigned lithology labels, for all four wells concurrently. The lithology distribution derived from the rule-based labeling scheme (Section 2.2), as summarized in Table 3, indicates that Sandstone is predominant in three of the four wells (63.3-75.3% in F03-02, F03-04, F06-01), while F02-01 exhibits a distinctly different character: Tight Sand is dominant at 63.3%, with a comparatively lower presence of clean Sandstone at 18.0%. This observation aligns with the well's distinctly elevated mean GR and the visibly different log character illustrated in Figure 2. Shale constitutes a significant minority class across all wells, ranging from 10.8% to 18.7%. Carbonate, although very rare, is indeed present in this dataset at a trace level of 0.5% in F03-02. This contrasts with the complete absence noted in some

simplified analyses of similar data, highlighting the predominantly siliciclastic, yet not exclusively siliciclastic, depositional environment of the study area.

Table 3. Lithology class distribution (%) per well, resulting from the rule-based petrophysical labelling scheme described in Section 2.2.

Well	Shale (%)	Sandstone (%)	Tight Sand (%)	Carbonate (%)
F02-01	18.7	18.0	63.3	0.0
F03-02	10.8	70.6	18.0	0.5
F03-04	14.0	75.3	10.7	0.0
F06-01	13.7	70.0	16.3	0.0

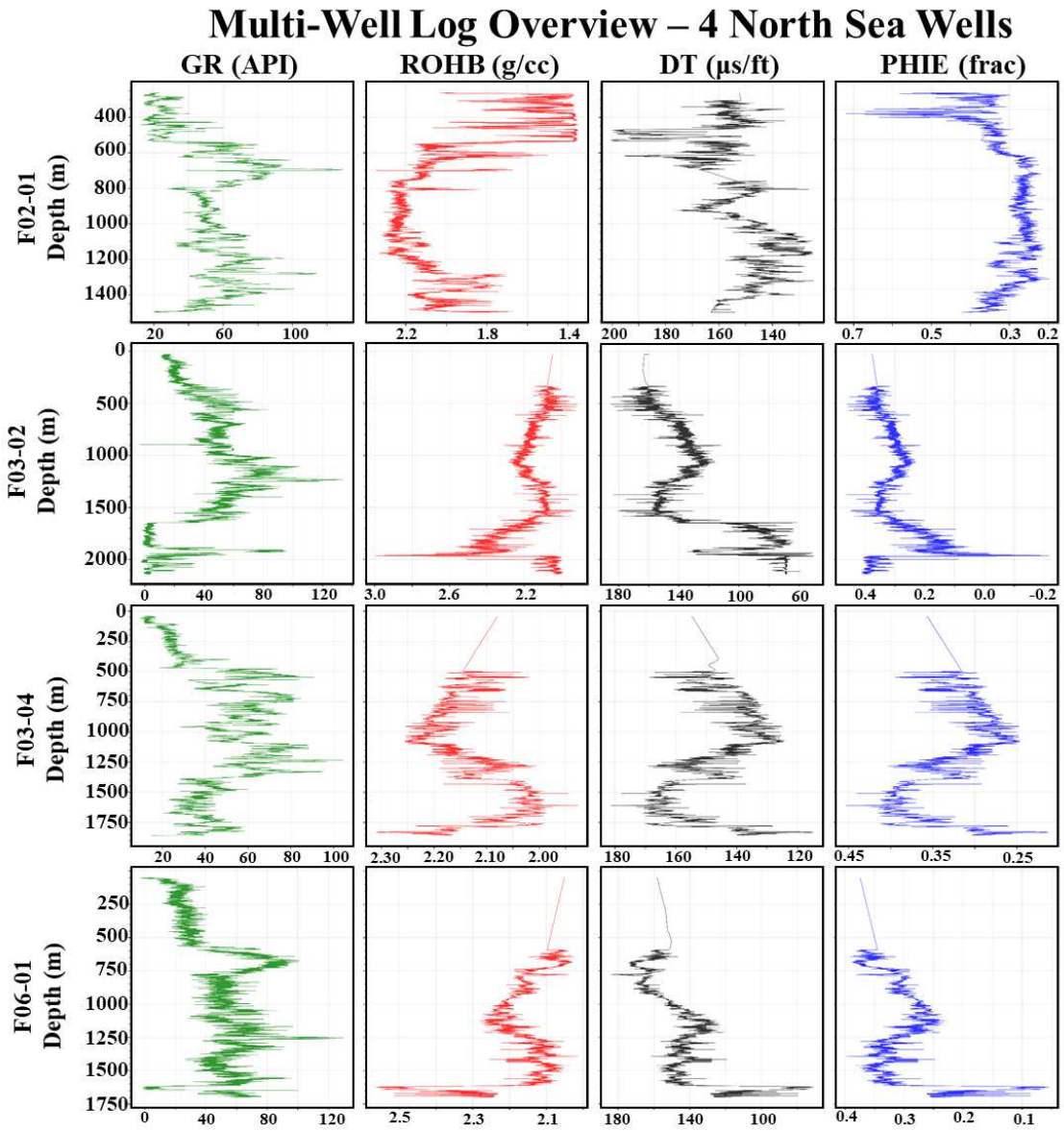


Figure 3: Multi-well overview of wireline log traces, coloured by rule-based lithology label, for all four study wells (F02-01, F03-02, F03-04, F06-01).

3.2. Cross-Well Petrophysical Cross-Plots

Before the training of the model, cross-well petrophysical cross-plots (Figure 4) were analyzed to ensure that the rule-based labeling scheme generated consistent and well-separated lithology clusters across all four wells, rather than artifacts specific to individual wells. The GR-RHOB and RHOB-PHIE cross-plots demonstrate a coherent clustering of Shale, Sandstone, and Tight Sand populations across all wells, maintaining the expected physical trends.

Shale samples are grouped at elevated gamma ray (GR) readings and intermediate to high bulk density (RHOB); Sandstone samples are located at low GR, low RHOB, and high porosity (PHIE); while Tight Sand is positioned in an intermediate, more dispersed area of the cross-plot space, aligning with its classification as a residual category that includes compacted or partially cemented sand. The rare Carbonate samples in F03-02 are distinctly plotted at low GR and high RHOB, aligning with their petrophysical definition (Table 2). This consistency across wells instilled confidence that a classifier trained on three wells could effectively transfer to the fourth without being hindered by a fundamentally different labeling regime in the held-out well.

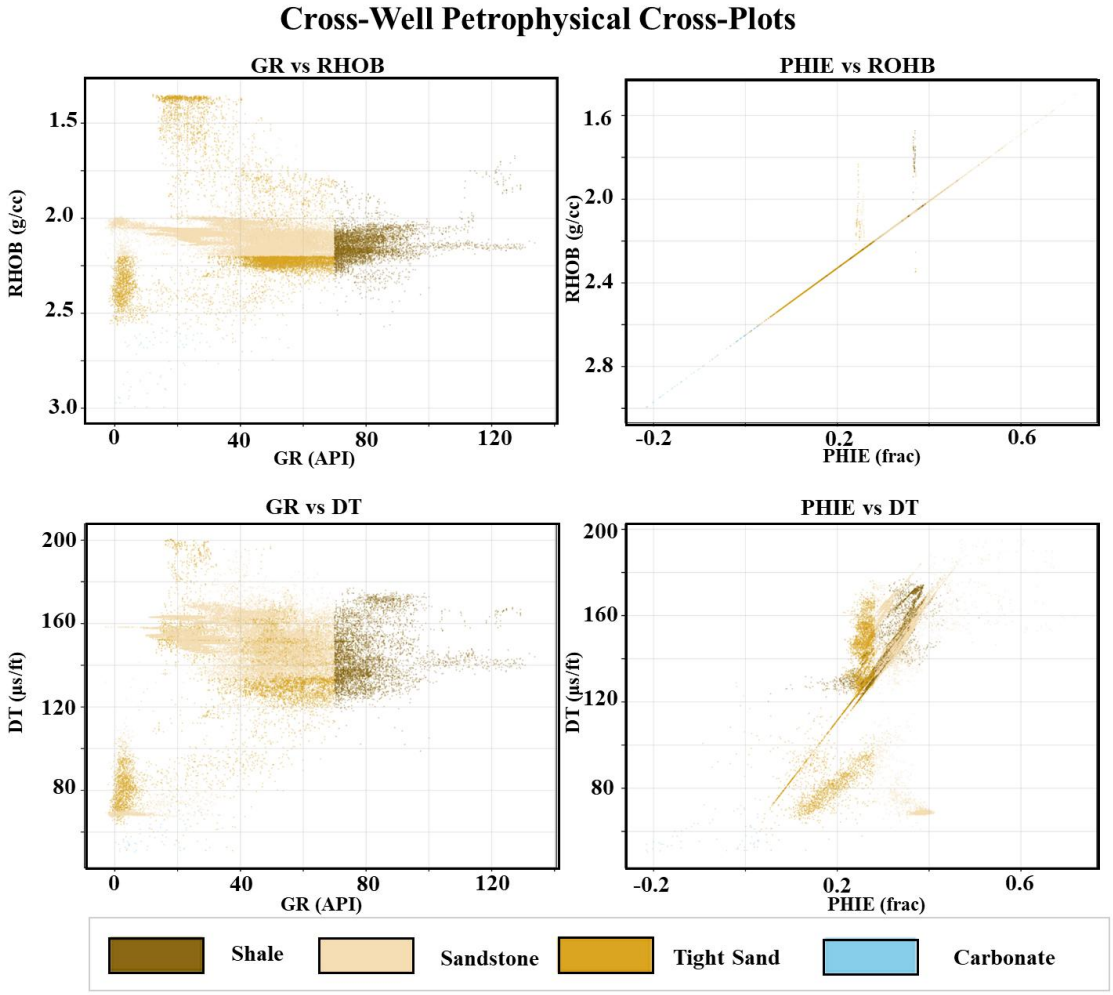


Figure 4: Cross-well petrophysical cross-plots (GR-RHOB, RHOB-PHIE, and related curve pairs) coloured by rule-based lithology label, confirming consistent class separation across all four study wells prior to model training.

3.3. Hyperparameter Tuning

Figure 5 illustrates the stratified five-fold cross-validation accuracy of the Random Forest model in relation to the number of trees in the ensemble (n_estimators). The analysis encompasses candidate values of 10, 25, 50, 100, 150, 200, and 300. Observe the significantly compressed y-axis scale in Figure 5 (values offset from 0.9998): cross-validation accuracy was nearly saturated throughout the entire range tested, with all ensemble sizes achieving accuracy within roughly 0.0001 of each other. Twenty-five trees reached the highest recorded cross-validation accuracy of 0.9999, slightly surpassing the other tested sizes, and was consequently selected as the final configuration for the Random Forest. This nearly flat tuning curve suggests that the four-feature, four-class lithology classification problem discussed here is well within the modeling capabilities of even a small ensemble. Additionally, it indicates that there is minimal additional predictive advantage to be obtained from significantly larger forests on this dataset.

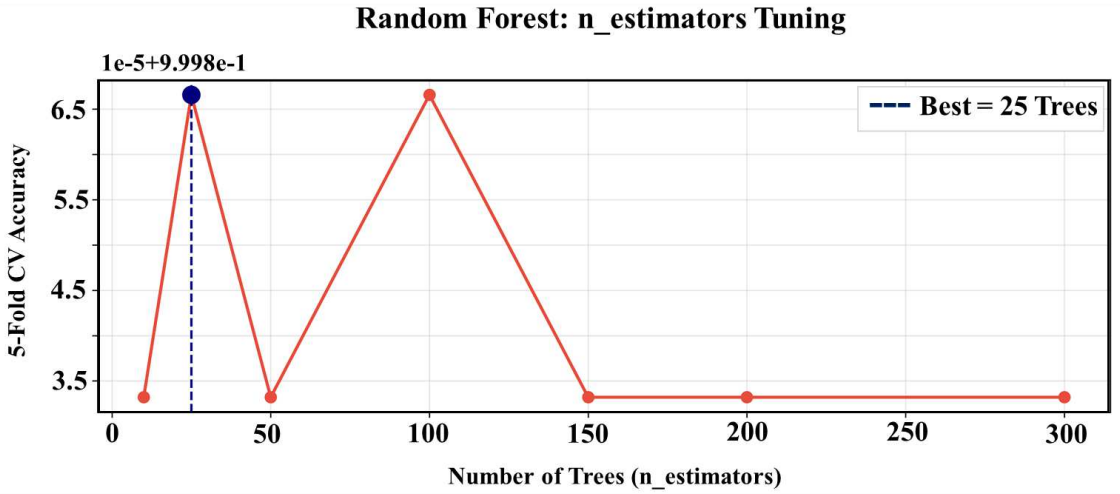


Figure 5: Stratified five-fold cross-validation accuracy of Random Forest as a function of the number of trees (n_estimators), showing rapid convergence to a near-perfect cross-validation score.

Figure 6 illustrates the cross-validation accuracy curve for KNN, depicting the relationship between the number of neighbors, K, which varies from 1 to 30. As illustrated in Figure 5, observe the compressed y-axis scale (0.9922-0.9930): all evaluated values of K attained cross-validation accuracy within roughly 0.0008 of each other. Accuracy achieved its peak value of 0.9930 at K = 4, then experienced a significant drop to a local minimum around K = 8-9. It subsequently recovered through a series of smaller secondary local peaks between K = 12 and K = 18, before gradually declining as K approached the largest tested values. This irregular, non-monotonic pattern — instead of the smooth,

single-peaked curve often idealized in textbook discussions of the bias-variance trade-off — illustrates the discrete, sample-specific characteristics of nearest-neighbour voting at this level of cross-validation accuracy. Here, successive integer values of K can each incorporate or omit a small number of borderline training samples situated near class boundaries. K = 4 was distinctly recognized as the optimal configuration and was selected as the final KNN configuration for this study.

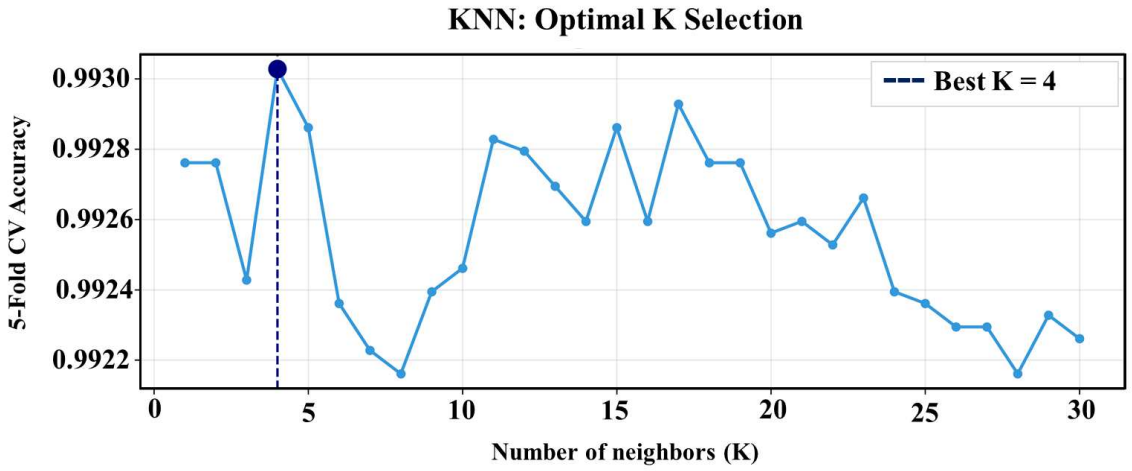


Figure 6: Stratified five-fold cross-validation accuracy of K-Nearest Neighbours as a function of the number of neighbours (K), showing a peak at K = 4.

3.4. Blind-Well Test Performance

Hyperparameters were set with twenty-five estimators for Random Forest and K = 4 for KNN. Both models were trained on the combined three-well training set, consisting of 29,978 samples, and evaluated on the completely withheld blind test well, F06-01, which included 10,826 samples. Random Forest attained an overall accuracy of 100.0% on the blind test, demonstrating perfect precision, recall, and F1-score (1.00) across all three lithology classes identified in F06-01 (Shale, Sandstone, Tight Sand; Carbonate was absent in this specific well under the established cut-offs). KNN achieved an overall accuracy of 98.45%, with a weighted F1-score of 0.984. The per-class F1-scores for KNN were 0.963 for Shale, 0.992 for Sandstone, and 0.971 for Tight Sand. This indicates that the majority of KNN's residual error was concentrated in the Shale class, particularly in recall (0.941). A significant fraction of true Shale samples were misclassified by KNN as a different lithology, likely reflecting gradational, lower-GR shale

intervals situated near the K-class decision boundary in standardized feature space. These results are encapsulated in the publication-ready results table, which is reproduced as Figure 7, and also presented in Table 4 below.

Table 4. Per-class and overall classification performance on the blind test well (F06-01) for Random Forest and KNN. Support denotes the number of true samples of each class in the test set.

Model	Lithology	Precision	Recall	F1-score	Support
Random Forest	Shale	1.000	1.000	1.000	1,480
Random Forest	Sandstone	1.000	1.000	1.000	7,578
Random Forest	Tight Sand	1.000	1.000	1.000	1,768
Random Forest	**Overall (weighted)**	**1.000**	**1.000**	**1.000**	**10,826**
KNN	Shale	0.987	0.941	0.963	1,480
KNN	Sandstone	0.990	0.994	0.992	7,578
KNN	Tight Sand	0.961	0.981	0.971	1,768
KNN	**Overall (weighted)**	**0.984**	**0.984**	**0.984**	**10,826**

Blind Well Test (F06-01) — Model Performance Summary

Model	Lithology	Precision	Recall	F1-score	Support	Accuracy
Random Forest	Shale	1.0	1.0	1.0	1480	1.0
Random Forest	Sandstone	1.0	1.0	1.0	7578	1.0
Random Forest	Tight Sand	1.0	1.0	1.0	1768	1.0
KNN	Shale	0.987	0.941	0.963	1480	0.984
KNN	Sandstone	0.99	0.994	0.992	7578	0.984
KNN	Tight Sand	0.961	0.981	0.971	1768	0.984
Random Forest	**Overall**	**—**	**—**	**1.0**	**10826**	**1.0**
KNN	**Overall**	**—**	**—**	**0.984**	**10826**	**0.984**

Figure 7: Results table summarising per-class and overall (weighted) precision, recall, F1-score, and accuracy for Random Forest and KNN on the blind test well F06-01.

3.5. Confusion Matrix Analysis

Figure 8 presents confusion matrices for both classifiers on the blind test well, represented in terms of raw sample counts and row-normalised proportions. The confusion matrix for the Random Forest model is entirely diagonal, exhibiting no off-diagonal entries across the three classes in the test well, thereby validating the 100.0% accuracy stated in Section 3.4. The KNN confusion matrix indicates a reasonable level of confusion: out of 1,480 true Shale samples, 49 (3%) were incorrectly classified as Sandstone and 39 (3%) as Tight Sand, resulting in the relatively lower Shale recall (0.941) mentioned earlier. In the comparison between Sandstone and Tight Sand, the misclassification was relatively balanced: 32 genuine Sandstone samples were incorrectly identified as Tight Sand, while 29 authentic Tight Sand samples were misclassified as Sandstone. Additionally, there were minor confusions involving Shale and Tight Sand, with 14 and 5 samples misclassified, respectively. This pattern suggests that the primary source of error for KNN on this well was the differentiation between lower-GR Shale and the neighboring Sandstone and Tight Sand populations, rather than the Sandstone-Tight Sand boundary, which might be intuitively assumed to be the most challenging to distinguish due to their common depositional origin.

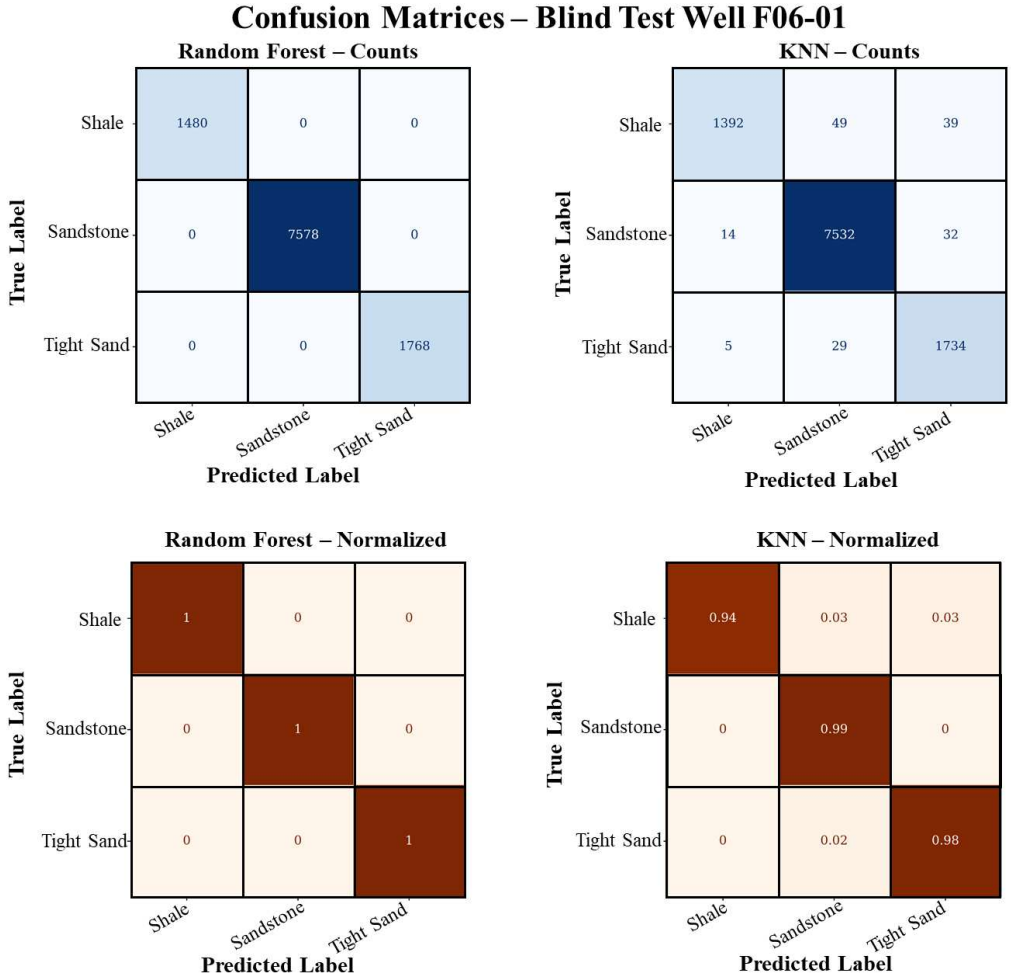


Figure 8: Confusion matrices for Random Forest (left) and KNN (right) on the blind test well F06-01, shown as raw counts (top row) and row-normalised proportions (bottom row).

3.6. Feature Importance

Figure 9 depicts the native feature importance of Random Forest, measured by the mean decrease in Gini impurity, for all four wireline curves. This information is displayed in both a bar chart and a radar/spider chart, allowing for direct visual comparison with the corresponding KNN permutation importance profile. Random Forest analysis revealed that gamma ray (GR) is the most significant curve, exhibiting a normalized importance of 0.434. This is closely followed by bulk density (RHOB) at 0.333 and effective porosity (PHIE) at 0.221. Corrected sonic transit time (DT_CORR) had minimal impact on the ensemble's decision-making, exhibiting a normalized importance of merely 0.012. This ranking can be directly understood in relation to the rule-based labeling scheme (Table 2): GR serves as the main criterion distinguishing Shale from all other classes and plays a significant role in the Carbonate cut-off,

which likely accounts for its primary importance. Meanwhile, RHOB and PHIE are incorporated directly into the Sandstone/Tight Sand and Carbonate criteria. DT_CORR, which does not engage with the labelling logic, contributes a minimal discriminating signal. This outcome is reassuring from a model-validity perspective: the classifier has accurately learned to ignore an input variable that provides no information regarding the (rule-based) target it was designed to predict.

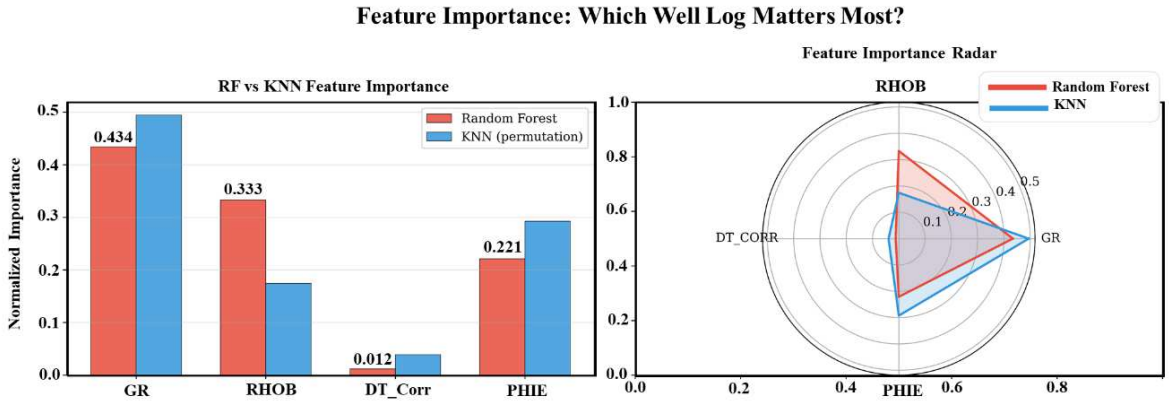


Figure 9: Feature importance comparison between Random Forest (native Gini importance) and KNN (permutation importance) across all four wireline curves, shown as a grouped bar chart (left) and radar chart (right).

3.7. Prediction Confidence and Model Agreement

With a binary agreement/disagreement track, Figure 10 analyzes the per-sample prediction confidence (highest predicted class probability) for both classifiers along the depth of the blind test well. Both models reached a consensus on the predicted lithology for 10,658 out of 10,826 blind-well samples, representing 98.4%. The disagreement was limited to 168 samples, accounting for 1.6%, and was primarily observed around the Shale boundary and, to a lesser extent, in the Sandstone/Tight Sand transition zones, as anticipated from the confusion-matrix analysis. Among the 168 samples where the two models exhibited disagreement, the predictions made by Random Forest aligned with the rule-based label in 100.0% of instances. In contrast, KNN's predictions corresponded with the rule-based label in 0.0% of the disagreement cases. This indicates that in every instance of divergence between the two classifiers, Random Forest was accurate while KNN was not. This finding offers robust, sample-level evidence (in addition to the aggregate accuracy figures presented in Section 3.4) that the classification advantage of Random Forest on this dataset is not

merely a result of class-imbalance-weighted averaging. Instead, it demonstrates a true and consistent superiority in addressing the specific ambiguous cases where the decision boundaries of the two algorithms diverge.

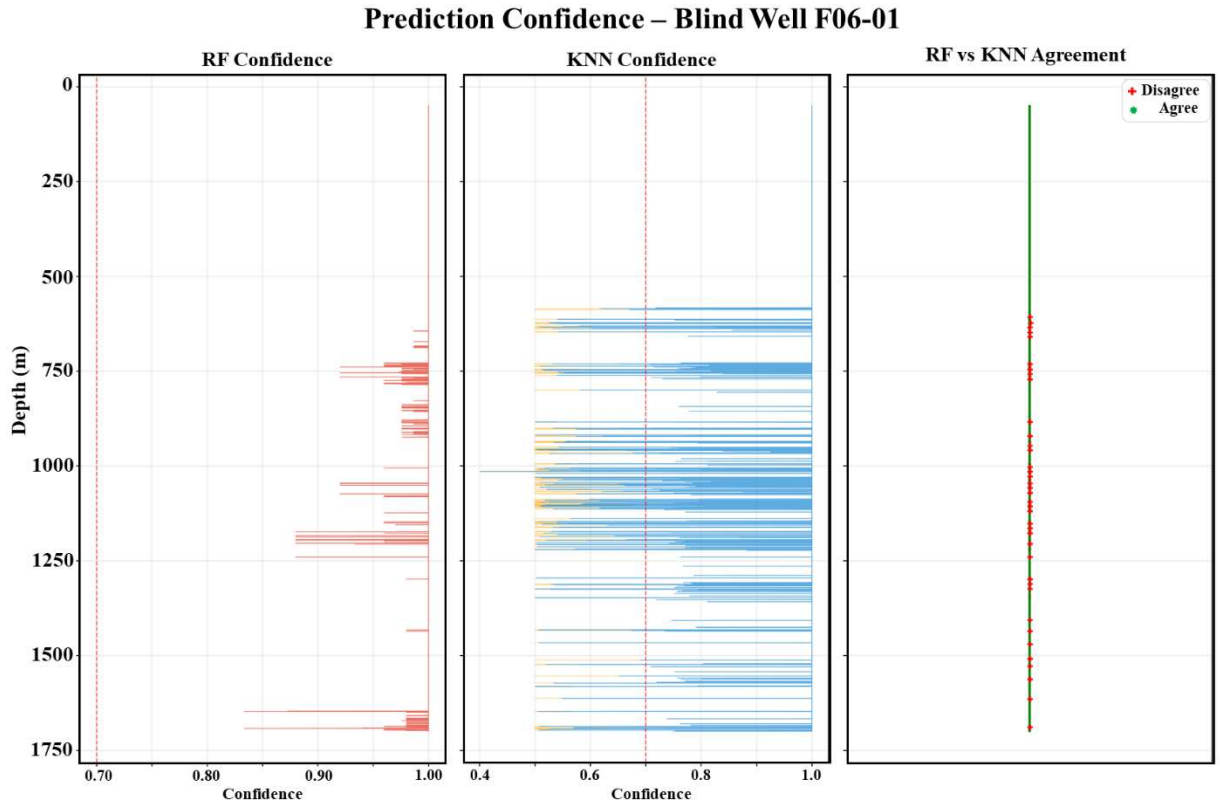


Figure 10: Prediction confidence (maximum class probability) for Random Forest (left) and KNN (middle) along the depth of the blind test well F06-01, with a binary model-agreement track (right) highlighting depth intervals of disagreement.

3.8. Full Well-Log Track Visualization

The final geoscience visualization for this analysis is shown in Figure 11, which is a seven-panel composite well-log track for the blind test well F06-01. It displays three lithology columns—the rule-based ('true') label, the Random Forest prediction, and the KNN prediction—as well as the four raw wireline curves (GR, RHOB, DT_CORR, and PHIE) against a common depth axis. Visual inspection validates the quantitative results of the confusion matrix: the Random Forest prediction track appears visually identical to the rule-based true lithology track at the shown scale, accurately reproducing each thin shale break and sand interval with remarkable fidelity. The KNN prediction track closely resembles the true track in general; however, it displays noticeable, scattered misclassifications, especially in

thin, gradational Shale intervals where individual samples are occasionally misclassified as Sandstone or Tight Sand. This observation aligns with the heightened Shale-class confusion discussed in Section 3.5. This figure demonstrates that, in addition to the overall statistics, the practical, depth-localised reliability of Random Forest predictions excels in tasks like pinpointing exact shale-cap or reservoir-to-non-reservoir contacts, where minor misclassifications could significantly impact net-pay calculations or seal-integrity evaluations.

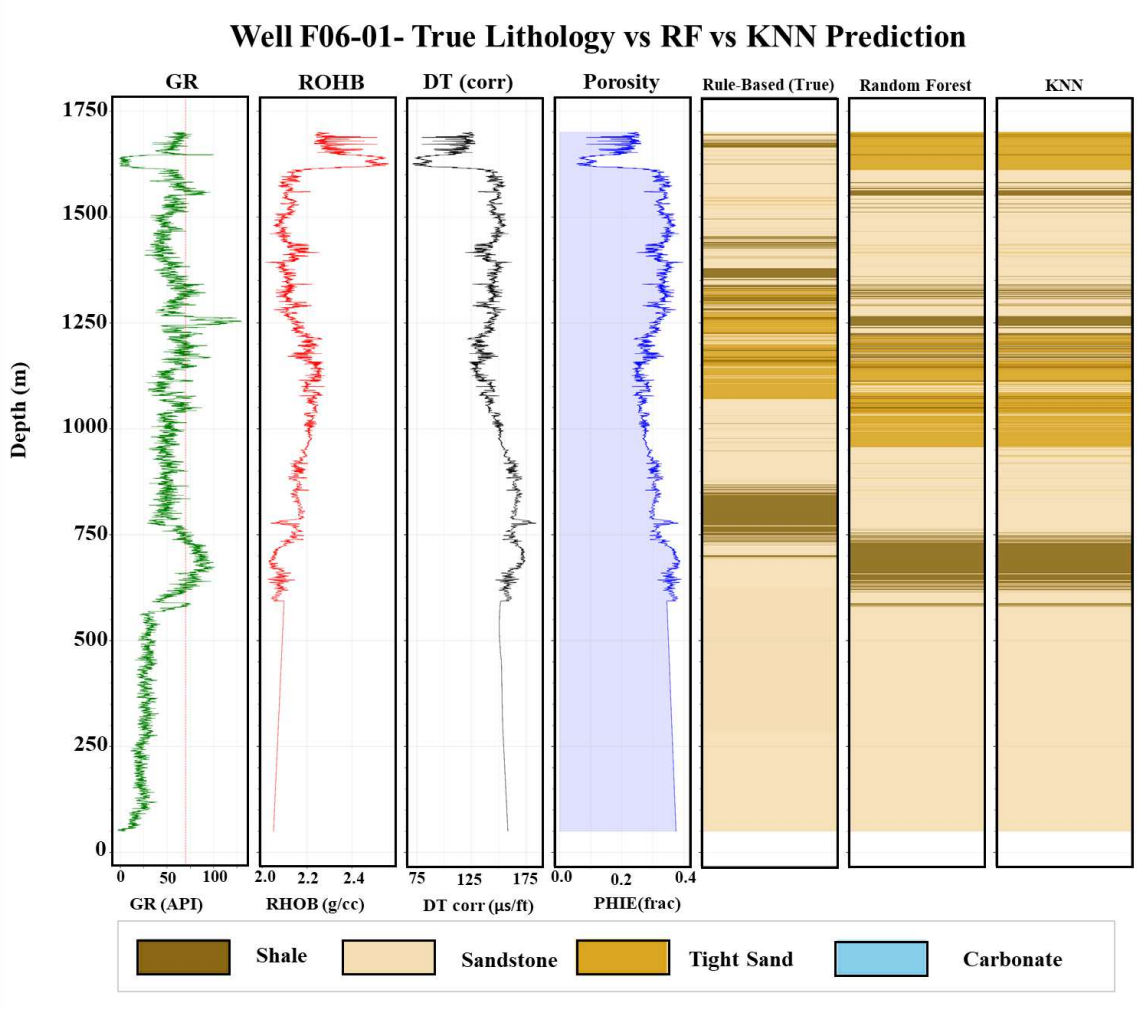


Figure 11: Full seven-panel well-log track for the blind test well F06-01: GR, RHOB, DT_CORR, and PHIE curves alongside the rule-based ('true') lithology column, Random Forest prediction column, and KNN prediction column, all plotted against a common depth axis.

3.9. Leave-One-Well-Out Cross-Validation

The result from the blind well mentioned above is based on a specific selection of the held-out well. To evaluate the generalizability of the Random Forest advantage across various selections of held-out wells, the Leave-One-Well-Out (LOWO) cross-validation method specified in Section 2.4.2 was utilized, systematically withholding each of the four wells in succession. Table 5 provides a summary of the results. Random Forest attained a mean LOWO accuracy of 99.58% (with a range of 98.82-100.00% across the four folds) and consistently outperformed KNN across all four held-out wells. KNN achieved a mean LOWO accuracy of 97.10% (range: 94.55-98.99%). Its weakest performance occurred when F02-01 was held out, resulting in an accuracy of 94.55%. This well's distinctive log character, characterized by an elevated mean GR and a Tight-Sand-dominated lithology distribution, differs significantly from the other three wells used to fit the StandardScaler and the KNN neighbourhood structure during that fold (Tables 1 and 3). This single result is particularly instructive: it exemplifies the inter-well distributional shift discussed theoretically in Section 4.1, which disproportionately degrades KNN in comparison to Random Forest. F02-01 serves as a direct, empirical illustration of that mechanism within this dataset. The accuracy of Random Forest on the held-out set (98.82%) demonstrated a significantly greater resilience to the same distributional shift.

Table 5. Leave-One-Well-Out (LOWO) cross-validation results: weighted accuracy and weighted F1-score for Random Forest and KNN, for each of the four wells held out in turn, plus the overall mean across all four folds.

Held-Out Well	RF Accuracy	RF F1	KNN Accuracy	KNN F1	Winner
F02-01	0.9882	0.9883	0.9455	0.9457	Random Forest
F03-02	0.9950	0.9926	0.9641	0.9626	Random Forest
F03-04	1.0000	1.0000	0.9899	0.9899	Random Forest
F06-01	1.0000	1.0000	0.9845	0.9844	Random Forest
Mean	**0.9958**	**0.9952**	**0.9710**	**0.9707**	**Random Forest (4/4 wells)**

Here is the final comparison dashboard (Figure 12). It combines the per-well LOWO accuracy bar chart, per-well LOWO weighted F1-score bar chart, aggregate mean-performance comparison, and a qualitative model-properties reference table. The table summarizes the practical trade-offs between the two algorithms, including their requirements for feature scaling, training and prediction speed, robustness to noise, interpretability, and typical multi-well transfer performance. It also discusses the importance of native versus permutation-based features, transparency, and robustness to noise. The consistency of the Random Forest advantage across all four LOWO folds , particularly in the

F02-01 fold, which experienced the most significant inter-well distributional shift — along with its lack of requirement for feature scaling and its inherent feature-importance output- reinforces its selection as the preferred classifier for this multi-well lithology prediction task.

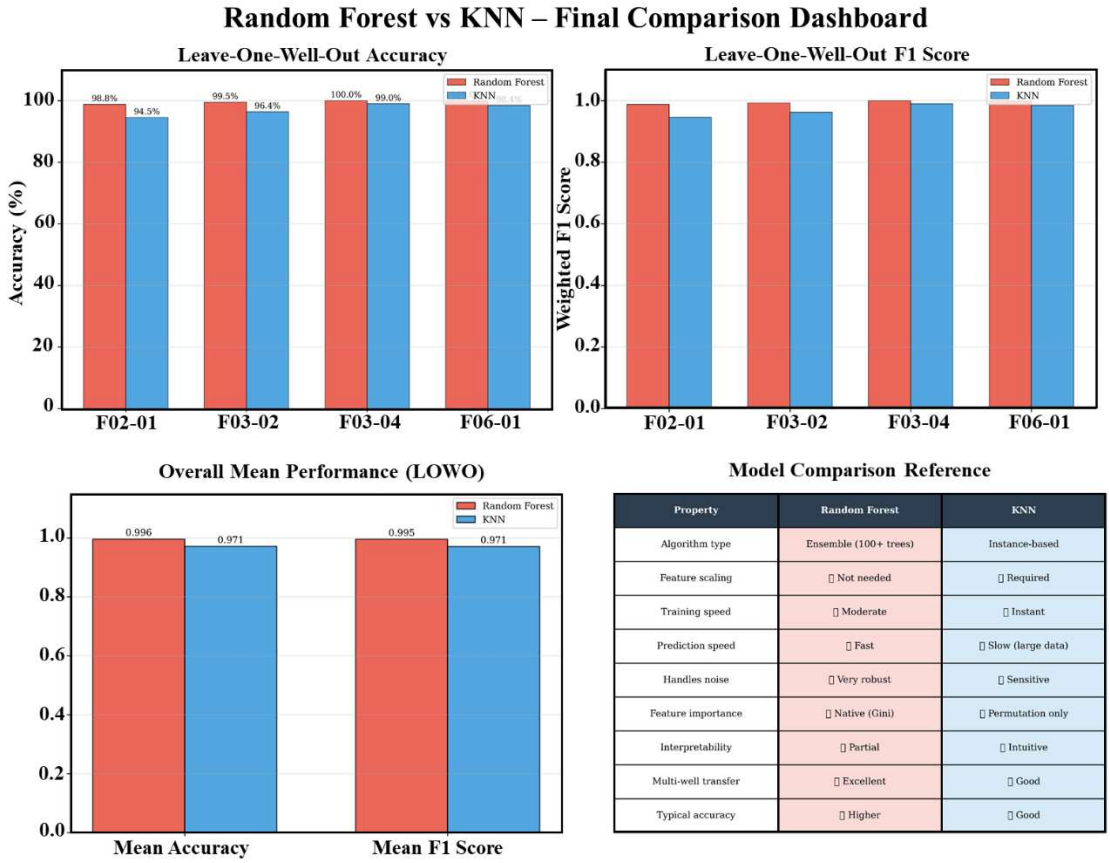


Figure 12. Final comparison dashboard: per-well LOWO accuracy (top left), per-well LOWO weighted F1-score (top right), aggregate mean-performance comparison (bottom left), and qualitative model-properties reference table (bottom right).

4. Discussion

The single blind-well test and all four folds of the Leave-One-Well-Out cross-validation showed a wide and statistically significant performance disparity between Random Forest and KNN, consistent with several theoretical and practical properties of both algorithms. Random Forest's recursive, axis-aligned partitioning of feature space remains unchanged when features are monotonically rescaled. Thus, Random Forest decision boundaries are unaffected by systematic inter-well differences in the absolute level of a single log curve, such as F02-01's higher

mean GR than the other three wells (see Table 1). This is valid if lithology and log response correlation follows a consistent ordering and threshold structure across all wells. KNN classifies using Euclidean distance and a consistent feature space. This work correctly trained the scaler exclusively on the training wells to relocate the test well into the feature space relative to the training neighborhood structure if the training data does not contain any regular inter-well offset. The nearest-neighbor vote is directly influenced by this change.

The Random Forest bagging approach averages trees based using bootstrapped subsets of the training data. This strategy frequently produces decision boundaries that are robust to each training set's noise. KNN, a one-pass, local technique, does not employ the ensemble structure's implicit regularization for training data quirks. Thirdly, as each Random Forest tree randomly subsets two of the four features at each split, the ensemble finds several partially redundant decision routes that produce the same classification result. This strategy may improve resilience if a single feature (like GR) is less reliable or more variable in an unknown well than in the training wells.

As stated in Section 3.7, Random Forest was correct in 100% of the 168 blind test disagreements between the two models. This sample-level conclusion rules out the notion that class-imbalance-weighted averaging conceals compensating faults, giving Random Forest an accuracy advantage. However, Random Forest regularly beats all other classifiers, especially when borders are ambiguous and hard to resolve.

Along with its improved accuracy on this dataset, Random Forest's intrinsic and cost-effective feature importance output, generated without model evaluations beyond the training period, is a major practical benefit for geoscience applications. The discriminating curves' order—gamma ray (GR) at 0.434 (normalized Random Forest importance), bulk density (RHOB) at 0.333, effective porosity (PHIE) at 0.221, and corrected sonic transit time (DT_CORR) at 0.012—explains the results in geological terms. GR is both part of the Carbonate cut-off and the main criterion for classifying the Shale class in this rule-based labeling technique. RHOB and PHIE affect carbonate and sandstone criteria. Thus, these three curves may affect classifier performance. Two structurally independent algorithms agreeing on which input variables are most and least significant gives geoscientists more confidence in that ranking than a single model. KNN's independently computed permutation importance (Figure 9) was near to the top and bottom of the ranking, with GR being the most important curve (normalized importance 0.50) and DT CORR the least important (0.04). It's a useful crosscheck. Both algorithms ranked RHOB and PHIE in the middle differently. Random Forest favored RHOB over PHIE, whereas KNN favored PHIE due to permutation significance. This partial disagreement indicates that both curves include substantial, overlapping discriminating information for this labeling method, which

is useful. However, their accurate second-order ranking is less stable and more affected by the importance approach utilized, such as eliminating contaminants during tree construction or employing permutation-based accuracy deterioration.

The rule-based labeling approach utilized in this study partially explains the importance ranking. Table 2 indicates that this strategy labels GR, RHOB, and PHIE as explicit cut-off variables but not DT_CORR. Since the aim variable may have been an alternate labeling scheme or true core-calibrated lithology, DT_CORR's relative priority ranking may have been changed. Petrophysically, porosity and matrix lithology affect sonic transit time, although this metric was not utilized to label.

5. Conclusions

The paper presents a comprehensive and versatile, machine-learning framework designed for the automated classification of lithology from wireline logs, based on multi-well datasets. In contrast to traditional single-well workflows, the proposed framework highlights the ability to predict inter-well performance by applying blind-well testing and Leave-One-Well-Out (LOWO) validation. This approach offers a practical evaluation of how well the model can be applied to wells that have not been previously encountered. By incorporating four broadly obtained wireline logs (GR, RHOB, DT_CORR, and PHIE), the framework effectively categorized four primary lithologies (Shale, Sandstone, Tight Sand, and Carbonate) while ensuring high predictive accuracy across diverse geological conditions. The results indicate that ensemble learning provides consistent and dependable performance despite inter-well variability, underscoring its applicability for extensive subsurface characterization. Feature importance analysis revealed that gamma ray serves as the primary predictor of lithology, whereas corrected sonic transit time has a minimal contribution. This insight offers valuable direction for future log acquisition and feature selection strategies. The proposed workflow establishes a scalable and computationally efficient approach for automated lithology prediction, which can be easily extended to additional wells, larger basins, and more complex geological settings. Future work will concentrate on the integration of resistivity logs, core-calibrated lithological labels, advanced ensemble and deep learning models, along with uncertainty quantification, to improve the reliability and practical applicability of machine-learning-driven subsurface characterization.

6. Acknowledgments

The authors wish to express their gratitude to the Department of Geophysical Engineering at Istanbul Technical University for the provision of facilities utilized in this research. This work was supported by the Performance-Based Research Project (PEDAP) by BAP Istanbul Technical University.

Authorship contribution statement

Yasir Bashir: Concept and design of study, development of concept, acquisition of data, drafting the manuscript.

Zoonash Arshad: Concept and design of study, Writing, Concept of study, interpretation of data, review & editing

Hülya Kurt: Design of study, analysis, and/or interpretation of data, drafting the manuscript.

Caner İmren: Concept of study, interpretation of data, drafting and revising the manuscript.

Muhsan Ehsan: Writing: Concept of study, ML Framework, interpretation of data, review & editing

Compliance with Ethical Standards:

Conflict of Interest:

The authors declare that they have no conflict of interest.

Editorial Board Conflict of Interest:

None of the authors is an Editor-in-Chief, Guest Editor, Associate Editor, or Editorial Board Member of the journal. Therefore, no editorial conflict of interest applies.

Research Involving Human Participants and/or Animals:

This research did not involve human participants, human subjects, or animals. Therefore, ethical approval and informed consent were not required.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

The well-log data used in this study are subject to data access and confidentiality restrictions and therefore cannot be made publicly available. However, the data may be made available by the corresponding author upon reasonable request, subject to the applicable data-sharing restrictions, permissions, and conditions governing the use of the dataset.

References

Alaudah, Y., Michałowicz, P., Alfarraj, M., & AlRegib, G. (2019). A machine-learning benchmark for facies classification. *Interpretation*, *7*(3), SE175–SE187.

Ao, Y., Li, H., Zhu, L., Ali, S., & Yang, Z. (2018). Logging lithology discrimination in the prototype similarity space with random forest. *IEEE Geoscience and Remote Sensing Letters*, *16*(5), 687–691.

Bestagini, P., Lipari, V., & Tubaro, S. (2017). A machine learning approach to facies classification using well logs. *SEG International Exposition and Annual Meeting*, SEG-2017.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Bressan, T. S., de Souza, M. K., Girelli, T. J., & Junior, F. C. (2020). Evaluation of machine learning methods for lithology classification using geophysical data. *Computers & Geosciences*, *139*, 104475.

Chawshin, K., Berg, C. F., Varagnolo, D., & Lopez, O. (2021). Lithology classification of whole core CT scans using convolutional neural networks. *SN Applied Sciences*, *3*(6), 668.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27.

Dunham, M. W., Malcolm, A., & Welford, J. K. (2020). Improved well log classification using semisupervised Gaussian mixture models and a new hyper-parameter selection strategy. *Computers & Geosciences*, *140*, 104501.

Hall, B. (2016). Facies classification using machine learning. *The Leading Edge*, *35*(10), 906–909.

Jamshidi Gohari, M. S., Emami Niri, M., Sadeghnejad, S., & Ghiasi-Freez, J. (2023). An ensemble-based machine learning solution for imbalanced multiclass dataset during lithology log generation. *Scientific Reports*, *13*(1), 21622.

Jiang, L., Zhang, Z. M., Wang, Q. W., Feng, Z. B., Zhang, B. C., & Ren, T. F. (2024). Comparative study on lithology classification of oil logging data based on different machine learning models. *Geophysical and Geochemical Exploration*, *48*(02), 489–497.

Li, Z., Zhang, Z., Jiang, L., Liu, Z., Zhou, H., Sun, J., Lee, K., & Ning, F. (2025). WirMAE: Learning well-logging interval representations via masked autoencoders for gas hydrate reservoir characterization. *IEEE Transactions on Geoscience and Remote Sensing*, *63*, 1–16.

Mousavi, S. H. R., & Hosseini-Nasab, S. M. (2024). A novel approach to classify lithology of reservoir formations using GrowNet and Deep-Insight with physic-based feature augmentation. *Energy Science & Engineering*, *12*(10), 4453–4477.

Mukherjee, B., Kar, S., & Sain, K. (2024). Machine Learning Assisted State-of-the-Art-of Petrographic Classification From Geophysical Logs. *Pure & Applied Geophysics*, *181*(9), 2839.

Pang, Q., Chen, C., Sun, Y., & Pang, S. (2025). STNet: Advancing Lithology Identification with a Spatiotemporal Deep Learning Framework for Well Logging Data: STNet: Advancing Lithology Identification: Q. Pang, C. Chen, Y. Sun and S. Pang. *Natural Resources Research*, *34*(1), 327.

Rider, M. (1996). *The geological interpretation of well logs; 2*.

Sun, Y., Pang, S., Qiao, S., & Zhang, Y. (2026a). Meta-Learning and Transformer Model Through Geophysical Well Logging Data Processed via Weighted Wavelet Thresholding Denoising Algorithm for Lithology Identification. *IEEE Transactions on Geoscience and Remote Sensing*.

Sun, Y., Pang, S., Qiao, S., & Zhang, Y. (2026b). Meta-Learning and Transformer Model Through Geophysical Well Logging Data Processed via Weighted Wavelet Thresholding Denoising Algorithm for Lithology Identification. *IEEE Transactions on Geoscience and Remote Sensing*.

584 Tewari, S., & Dwivedi, U. D. (2019). Ensemble-based big data analytics of lithofacies for automatic development of
585 petroleum reservoirs. *Computers & Industrial Engineering*, *128*, 937–947.
586 Tschannen, V., Delescluse, M., Rodriguez, M., & Keuper, J. (2017). Facies classification from well logs using an
587 inception convolutional network. *ArXiv Preprint ArXiv:1706.00613*.
588 Zhang, H., Wu, W., & Song, X. (2024). Well logs reconstruction based on deep learning technology. *IEEE*
589 *Geoscience and Remote Sensing Letters*, *21*, 1–5.
590 Zhang, Z., Zhang, X., Zhang, C., & Feng, H. (2025). Category-wise sequence correlation filter for lithology
591 identification with well logs. *IEEE Transactions on Geoscience and Remote Sensing*, *63*, 1–14.
592 Zhao, N., Li, N., Xiao, Z., Wang, X., & Qin, C. (2024). Physics-driven deep learning pixel-based inversion of
593 logging while drilling in anisotropic formation. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–
594 12.
595

596