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.

# A Hybrid Ensemble Learning Framework for Permeability Prediction Using Integrated Petrophysical Logs, Core Measurements, and Drill Stem Test Data

**Pardis Ebrahimi**[1]
University of Tehran
**Ali Safaei**[2]
ali.safaei@ut.ac.ir
University of Tehran
**Atefeh Hassan-Zadeh**
University of Tehran
**Alireza Salamat**[3]
University of Tehran

**Article**

**Keywords:** Permeability estimation, Machine learning, Ensemble Learning, Drill Stem Test, Petrophysical well logs, SHAP analysis, Reservoir engineering

**Posted Date:** September 8th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10641204/v1

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

**Additional Declarations:** No competing interests reported.

**A Hybrid Ensemble Learning Framework for Permeability Prediction Using Integrated Petrophysical Logs, Core Measurements, and Drill Stem Test Data**

**Authors**

**Pardis Ebrahimi[1], Ali Safaei[2,*], Atefeh Hassan-Zadeh[2], Alireza Salamat[3]**

[1] M.Sc. in Petroleum Engineering, Fouman Faculty of Engineering, College of Engineering, University of Tehran, Iran

[2] Fouman Faculty of Engineering, College of Engineering, University of Tehran, Iran, P.O.Box: 43581-39115

[3] Pars Oil and Gas Company (POGC), Tehran, Iran
*(Corresponding authors: Ali Safaei, Email: [*ali.safaei@ut.ac.ir*])*

# List of abbreviations

| Acronym | Definition |
|---|---|
| **AI** | Artificial Intelligence |
| **ANN** | Artificial Neural Network |
| **BPP** | Bivariate Prediction Model |
| **CART** | Classification and Regression Trees |
| **CDF** | Compaction Derived Factor (Formation Compressibility) |
| **CGR** | Corrected Gamma Ray |
| **CV** | Cross-Validation |
| **DI_HRLT** | High-Resolution Resistivity Logging |
| **DST** | Drill Stem Test |
| **ELM** | Extreme Learning Machine |
| **GBM** | Gradient Boosting Machine |
| **GR** | Gamma Ray |
| **HDAR / HDRA** | High-Resolution Density and Acoustic Log |
| **HTEM** | High-Resolution Temperature Log |
| **ICV** | Invasion Control Value |
| **IHV** | Invasion Hydrocarbon Volume |
| **KNN** | K-Nearest Neighbors |
| **LR** | Logistic Regression |
| **MAE** | Mean Absolute Error |
| **ML** | Machine Learning |
| **MLP** | Multi-Layer Perceptron |
| **MSE** | Mean Squared Error |
| **NPHI** | Neutron Porosity |
| **PEFZ** | Corrected Photoelectric Factor |
| **PNN** | Probabilistic Neural Network |
| **POTA** | Potassium Content |
| **RF** | Random Forest |
| **RHOB** | Bulk Density |
| **RHOZ** | Corrected Bulk Density (Density Log) |
| **RLA1–RLA5** | Array Laterolog Resistivity Logs |

| | |
|---|---|
| **ResNet** | Deep Residual Neural Network |
| **RMSE** | Root Mean Square Error |
| **ROC** | Receiver Operating Characteristic |
| **RT_HRLT** | Deep True (High-Resolution Laterolog) Resistivity |
| **RXOZ** | Flushed-Zone Resistivity |
| **SGR** | Spectral Gamma Ray |
| **SHAP** | SHapley Additive exPlanations |
| **SVM** | Support Vector Machine |
| **SVR** | Support Vector Regression |
| **TENS** | Tensile Strength Log |
| **THOR** | Thorium Content |
| **TURA** | Thorium-to-Uranium Ratio Indicator |
| **UPP** | Univariate Prediction Model |
| **UPRA** | Uranium Ratio Indicator |
| **URAN** | Uranium Content |
| **XGBoost** | Extreme Gradient Boosting |


## Abstract

Permeability is one of the most fundamental petrophysical parameters governing reservoir quality assessment, fluid-flow simulation, and hydrocarbon field development planning. Although conventional determination methods, such as core analysis and well testing, offer high measurement accuracy, their operational constraints, elevated costs, and limited vertical coverage have rendered data-driven intelligent approaches indispensable in reservoir engineering practice. In this work, a comprehensive framework built upon machine learning and hybrid Ensemble Learning techniques is presented for permeability estimation, integrating petrophysical well-log data, core measurements, and Drill Stem Test (DST) records acquired from a producing well in an Iranian gas field.

To examine reservoir behavior from two complementary standpoints, the available data were organized into two principal subsets: a DST–Log dataset and a Core–Log dataset. Permeability estimation within the DST–Log subset was formulated as a classification problem, whereas the Core–Log subset was treated as a regression task. Following a preprocessing sequence comprising outlier removal, normalization, depth synchronization, and feature selection, several machine-learning algorithms—namely Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree, and hybrid Voting/Stacking ensembles—were developed and evaluated. In addition, SHapley Additive exPlanations (SHAP) analysis was employed to enhance model interpretability and quantify the contribution of individual input features.

The results indicate that tree-based algorithms and hybrid ensemble models consistently outperformed standalone learners across both classification and regression tasks. In the classification stage, the SVM, Decision Tree, and Ensemble models achieved accuracies in the range of 0.99–1.00, whereas Logistic Regression exhibited comparatively weaker performance in zones characterized by lithological overlap. In the regression stage, the Ensemble and Decision Tree models yielded the lowest prediction errors and the greatest stability in reconstructing continuous permeability profiles. Comparative analysis further demonstrated that hybrid Ensemble Learning, by integrating linear, nonlinear, local, and hierarchical learning mechanisms, substantially improved generalization capability,

reduced prediction uncertainty, and enhanced robustness against noisy, heterogeneous petrophysical data.

Feature-importance analysis revealed that resistivity logs (the RLA series, RXOZ, and RT_HRLT), gamma-ray indicators (CGR and SGR), and porosity–density parameters (NPHI and RHOB/RHOZ) exert the dominant control on permeability prediction. Collectively, these findings demonstrate that coupling dynamic well-testing data with static petrophysical measurements, within an integrated machine-learning and Ensemble Learning framework, provides a reliable and interpretable basis for permeability estimation in structurally complex, heterogeneous reservoirs. The proposed methodology holds practical relevance for reservoir characterization, well-completion design, production management, and future field-development planning.

**Keywords:** Permeability estimation; Machine learning; Ensemble Learning; Drill Stem Test; Petrophysical well logs; SHAP analysis; Reservoir engineering

## 1. Introduction

Permeability, as a fundamental petrophysical property, characterizes the ability of reservoir rock to transmit fluids and exerts a direct influence on production-capacity assessment, field-development planning, and the selection of enhanced-recovery strategies [1]. Given its central importance, accurate permeability estimation has remained a core challenge in petroleum engineering since the earliest stages of reservoir characterization.

In the early stages of the discipline, permeability was determined primarily through core analysis and laboratory experimentation. Although these methods offer high measurement accuracy, their high cost, time-intensive nature, and inability to provide complete reservoir coverage motivated the development of indirect techniques, including the interpretation of well-test data and petrophysical well logs. Among the empirical correlations proposed for this purpose, the Kozeny–Carman relationship remains one of the most widely used, relating permeability to porosity and the specific surface area of rock grains [2,3]:

$$\frac{1}{C}\cdot\frac{\emptyset^2}{S^2(1-\emptyset)^2}=k \qquad (1)$$

where $k$ denotes permeability (m²), $\varphi$ is porosity, $S$ is the specific surface area of the grains, and $C$ is an empirical constant related to pore geometry. Despite its historical significance, this relationship loses predictive accuracy in reservoirs exhibiting pronounced heterogeneity, extensive fracturing, or substantial lithological variation [4].

Under conditions where mathematical modeling proves inadequate, the application of intelligent, data-driven systems becomes essential [5]. This necessity generally stems from one of two circumstances: either the full set of governing parameters remains incompletely characterized, or the interactions among these parameters are too complex to be captured through conventional mathematical formulations. In such cases, artificial

intelligence offers a practical means of observing and reproducing system behavior without requiring an explicit physical model [6,7].

A substantial body of research has addressed permeability and related petrophysical predictions using artificial-intelligence techniques. Artificial neural networks (ANNs) have been employed for porosity estimation [8], while the integration of seismic attributes with ANN architectures has been used both for porosity prediction [9] and for the broader estimation of reservoir parameters [10]. Nooruddin proposed an ANN-based model for permeability prediction using laboratory-derived mercury-injection capillary-pressure data as model input [11]. Al-Anazi demonstrated that support vector regression (SVR) can predict both permeability and porosity of sandstone reservoirs with satisfactory accuracy when applied to diverse input datasets [12]. In a related study, Singh employed three principal well-log inputs—sonic, density, and resistivity—to estimate porosity through artificial-intelligence–based methods [13].

A generalized and computationally efficient learning paradigm known as the extreme learning machine (ELM) has also been introduced, offering considerably faster training than conventional neural-network architectures [14]. Building on this approach, Saffarzadeh and Shadizadeh applied support vector regression to permeability prediction using resistivity, neutron, density, gamma-ray, and acoustic logs [15], while Olatunji et al. developed an ELM-based framework for permeability estimation [16]. Xiao et al. subsequently extended the ELM methodology by incorporating multiple hidden-layer architectures to improve predictive performance [17].

In a further contribution, Adeniran investigated permeability determination in both homogeneous and heterogeneous reservoirs through the application of data-clustering techniques [18]. Zhang, meanwhile, evaluated permeability estimation in sandstone formations using a range of modeling strategies, including univariate (UPP) and bivariate (BPP) prediction models, support vector regression, random forest (RF), and a deep residual neural network (ResNet) [19].

## 2. Literature Review

Artificial neural networks (ANNs) are mathematically formulated systems designed to receive, process, and reproduce information. Structurally, they consist of two principal components: (i) nodes, or neurons, and (ii) weighted connections linking these neurons. A basic network architecture typically comprises three layers: an input layer that receives signals from the external environment or from preceding neurons, a hidden layer responsible for processing these inputs and extracting relevant features, and an output layer that delivers the final result.

Relative to classical deterministic or statistical techniques, ANNs offer several notable advantages. Perhaps the most significant is their independence from a predefined functional form: rather than imposing fixed rules or predetermined relationships, the network is left to establish the mapping between inputs and the desired outputs through the training process itself. In addition, whereas linear regression models tend to shrink predicted values toward the mean, ANNs preserve the underlying variability present in

the data, which generally translates into higher accuracy when modeling nonlinear relationships [20].

Training in ANN-based approaches follows a supervised paradigm and therefore requires two categories of variables: a set of independent parameters and a single dependent parameter, with the network's task being to estimate the latter from the former. In the specific case of permeability prediction, the independent variables typically consist of conventional well-log measurements—such as GR, RHOB, NPHI, and DT—together with core-derived data, while the dependent variable corresponds to laboratory-measured permeability. Because the learning process is supervised, the observed permeability values serve as target outputs used to adjust the network's weights and biases, ultimately enabling accurate permeability estimation [21].

The decision-tree algorithm was first introduced by Quinlan in 1986 [22]. A decision tree provides a tree-structured representation of the feature space, partitioning the data according to a series of splitting criteria. While several tree-based variants have since been developed, the present study focuses exclusively on the Classification and Regression Trees (CART) approach [23]. A classification tree produces categorical outputs—assigning observations to discrete classes—whereas a regression tree yields continuous numerical predictions. Structurally, a decision tree consists of internal nodes and terminal leaves; although trees of arbitrary depth and complexity can, in principle, be constructed, excessively deep trees are prone to overfitting, and maximum tree depth is therefore commonly constrained during model development.

Tree construction proceeds through a recursive partitioning process aimed at achieving an optimal tree structure. The criteria conventionally used to evaluate candidate splits differ depending on the nature of the prediction task: for classification problems, the Gini index and mutual information are typically employed, whereas for regression problems, mean squared error (MSE) and mean absolute error (MAE) are more commonly used. These metrics quantify split quality and guide the algorithm toward the most informative partition of the data at each node. One practical advantage of decision trees lies in their capacity to indicate which input variables contribute most strongly to the prediction, even though they do not provide a fully quantitative ranking of feature importance [24].

The gradient boosting algorithm was originally proposed by Friedman in 2000 [25]. Built upon the boosting principle, the method constructs an ensemble of weak learners—typically shallow decision trees—in a sequential manner, such that each successive tree is trained to correct the residual errors left by its predecessor. This iterative process can be understood as fitting a series of regression models to the pseudo-residuals generated at each stage. As with most machine-learning algorithms, gradient boosting relies on the minimization of a specified loss function, with gradient descent used to progressively reduce the error contributed by each newly added estimator. The final predictive model is obtained by combining the initial baseline estimate with all subsequent weighted estimates. In the present study, gradient boosting was implemented using the Gradient Boosting Regressor available in the Scikit-learn library, complemented by the XGBoost library, which additionally incorporates regularization into the training procedure [26].

The support vector machine (SVM) was originally introduced by Cortes and Vapnik in 1995. In its regression form—support vector regression (SVR)—the algorithm seeks a function that provides the closest approximation to the observed data while satisfying a prescribed error tolerance [27]. This function is required to satisfy two conditions: first, deviations from the observed training values $y_i$ must not exceed a prescribed tolerance ε; and second, the function should remain as flat as possible in order to avoid overfitting. A linear approximation $f(X)$ can accordingly be expressed as:

$$f(X) = wX + b \tag{2}$$

The corresponding SVM optimization problem is formulated as:

$$min_w^{\frac{1}{2}} ||w^2|| \tag{3}$$

subject to the constraints:

$$y_i - wX - b \leq \varepsilon$$
$$wX - b - y_i \leq \varepsilon$$

where the slack variables $\xi_i^*$and $\xi_i$ are introduced to accommodate deviations exceeding the tolerance ε. The solution to this optimization problem is typically obtained through its dual formulation, which involves computing the Lagrange multipliers $\alpha$ and $\alpha_i^{'}$ [28]:

$$f(x) = \sum_{i=1}^{m} \left(\alpha_i^{'} - \alpha_i\right) K(x_i, x) + b \tag{4}$$

where $K(x_i, x)$ denotes the kernel function [29]. In the present study, the standard Gaussian (radial basis) kernel was adopted:

$$K\left(x, x^{'}\right) = \exp\left(-\frac{\left||x - x^{'2}|\right|}{2\sigma^2}\right) \tag{5}$$

The support vector machine algorithm was implemented in the Python environment using the SVR () method available in the Scikit-learn library [24].

The probabilistic neural network (PNN) was first introduced by Specht in 1990 and is widely recognized as an efficient nonlinear algorithm for pattern classification. PNN represents a class of supervised neural networks primarily applied to classification and pattern-recognition problems; its theoretical foundation draws on Bayesian network theory together with Fisher's kernel discriminant analysis framework [30].

One of the most notable applications of PNN in petroleum engineering lies in lithofacies classification, where the algorithm offers markedly faster training than conventional

backpropagation-based classifiers. Beyond this computational advantage, PNN offers several additional benefits, including:

- an inherently parallel network structure;
- the capacity to incorporate or remove training samples without requiring complete retraining of the network; and
- a theoretical guarantee of convergence toward an optimal classifier as the volume of training data increases [31].

In this study, the complete lithofacies-classification workflow based on PNN was implemented in the R statistical programming environment, using the *pnn* and *doParallel* software packages [32,33].

The gradient boosting regression model represents one of the most powerful algorithms in the machine-learning literature, originally developed by Friedman (2001, 2002) for modeling complex nonlinear dependencies. This approach can be regarded as an extension of both the AdaBoost algorithm and the gradient boosting machine (GBM) framework, both of which were formalized by Friedman [34].

GBM has since been widely applied across numerous data-driven problems for the accurate modeling and prediction of response variables. The underlying principle of the method is to iteratively fit a sequence of models such that the discrepancy between predicted and observed values is progressively minimized [35].

Within this algorithm, model learning proceeds in a direction that maximizes correlation with the negative gradient of a specified loss function. For continuous response variables, the loss function may be selected from the Gaussian or Laplace family, whereas for categorical outcomes, loss functions such as the binomial deviance or the AdaBoost loss are more appropriate. The central premise underlying the GBM loss function is that large deviations from the target value should be penalized, while comparatively small residuals may be disregarded. For continuous responses, the most commonly adopted loss function is the squared-error (L2) loss, whose derivative corresponds directly to the residual; GBM can therefore be interpreted, in essence, as a residual-fitting procedure.

The GBM workflow begins with the selection of a differentiable loss function that quantifies the discrepancy, or variance, between the observed and predicted values of the response variable y. An initial baseline model, F, is then defined—commonly taken as the simple mean of the response variable—after which a sequence of iterative updates is carried out until convergence is achieved. At each iteration, the negative gradient is computed as:

$$-g(X_i) = -\frac{\partial L\left(y_i, F_{(x_i)}\right)}{\partial F_{(x_i)}} \tag{6}$$

In the present study, the complete gradient boosting regression workflow was implemented using the *gbm* package within the R statistical programming environment,

which is widely recognized as one of the most robust and widely adopted open-source platforms for statistical computing [36].

# 3. Materials and Methods

## 3.1. Data Acquisition

### 3.1.1. Petrophysical Well Logs

Petrophysical well logs constitute the primary source of information for characterizing rock and fluid properties within the reservoir. These measurements are recorded continuously along the wellbore, and in the present study, the following comprehensive suite of logs was utilized:

- CGR (Corrected Gamma Ray): A uranium-corrected variant of the conventional gamma-ray log, obtained by removing the contribution of uranium present in the drilling mud. Relative to the standard GR log, CGR provides improved lithological discrimination and more reliable identification of clean-sand intervals.
- SGR (Spectral Gamma Ray): A spectral decomposition of the gamma-ray signal into its three principal radioactive constituents—potassium (K), thorium (Th), and uranium (U). This log is particularly valuable for determining clay provenance and for distinguishing uranium-rich from thorium-rich shale intervals.
- CDF (Compaction Derived Factor / Formation Compressibility): This log is used to evaluate the compressibility and mechanical response of the formation under in-situ reservoir stresses. CDF values provide meaningful insight into rock compressibility, the potential for compaction during production, and lithological variations associated with formation strength; such information is of particular relevance to geomechanical reservoir modeling and to the analysis of stress-dependent permeability behavior.
- ICV (Invasion Control Value): This parameter quantifies the depth and severity of drilling-fluid invasion into the formation by comparing the resistivity of the flushed zone with that of the deeper, uninvaded zone. ICV is useful for assessing cement quality, discriminating between permeable intervals, and evaluating the extent of formation damage.
- IHV (Invasion Hydrocarbon Volume): This parameter estimates the volume of hydrocarbon displaced or invaded within the flushed zone surrounding the wellbore. IHV serves as a key indicator of hydrocarbon saturation, productive-zone quality, and the extent to which drilling fluids have influenced the recorded petrophysical measurements.
- POTA (Potassium Content): Expresses the potassium content of the formation as a weight percentage and, in conjunction with THOR and URAN, supports detailed lithological interpretation and clay-source identification.
- THOR (Thorium Content): Reflects the thorium concentration within the rock matrix. Elevated THOR values generally indicate the presence of aluminosilicate shales, which typically exert an adverse influence on permeability.

- URAN (Uranium Content): Quantifies uranium concentration and is useful for identifying organic-rich shales and radioactive-enriched intervals, with additional relevance to geochemical interpretation.
- UPRA and TURA (Uranium Ratio Indicators): Computed ratios between uranium and other radioactive elements, employed to differentiate organic from non-organic shale intervals.
- RLA1–RLA5 (Array Laterolog Resistivity): A suite of resistivity measurements acquired at multiple radial depths of investigation, enabling a detailed characterization of drilling-fluid invasion and fluid saturation at progressively deeper zones within the formation.
- RXOZ (Resistivity of the Flushed Zone): Represents the resistivity of the invaded (flushed) zone; its contrast with the RLA measurements can indicate the degree of permeability and the extent of hydrocarbon saturation.
- RHOZ (Corrected Bulk Density): A corrected version of the bulk-density log from which borehole washout and environmental effects have been removed, thereby improving the accuracy of subsequent porosity calculations.
- HDRA and HDAR (High-Resolution Density and Acoustic Logs): High-resolution density and acoustic measurements employed for fine-scale analysis of heterogeneous lithologies [37].
- HTEM (High-Resolution Temperature Log): Records temperature variation along the wellbore and, when analyzed jointly with resistivity logs, can assist in identifying active fluid-flow zones.
- PEFZ (Corrected Photoelectric Factor): A refined version of the PEF log intended to improve mineralogical interpretation and to suppress tool-related noise.
- TENS (Tensile Strength Log): Represents the tensile strength of the formation and is of particular relevance to the mechanical analysis of the reservoir rock and to wellbore-stability assessment [38].

#### 3.1.2. Drill Stem Test (DST) Data

Drill Stem Test (DST) data were incorporated in this study as one of the principal reservoir-engineering datasets. DST is an in-situ dynamic testing procedure that enables direct evaluation of reservoir pressure, flow rate, effective permeability, and wellbore skin/damage index. The test is conducted by temporarily isolating the wellbore from the drilling-fluid column and inducing a series of alternating flow and shut-in periods; the resulting continuous pressure–time record provides critical information regarding the hydrodynamic characteristics of the reservoir interval under investigation.

#### 3.1.3. Core Data

Core data represent one of the most reliable sources for determining reservoir-rock properties, including porosity, permeability, fluid saturation, and textural characteristics. Core samples are recovered from selected intervals of the wellbore and subjected to a series of routine and advanced laboratory analyses. Although core data are inherently limited in vertical and areal coverage—and their acquisition entails substantial cost—the high measurement fidelity of core-derived properties renders them the primary reference for calibrating both petrophysical interpretations and machine-learning algorithms.

### 3.2. Data Cleaning and Preparation

Among the most critical steps in the preprocessing of petrophysical and reservoir-engineering datasets are the removal of outlier observations, quality control of individual measurements, and the harmonization of depth intervals across data sources. Poor data quality can lead to unstable model training and a consequent reduction in permeability-prediction accuracy [1]. In the present study, the data-cleaning procedure was carried out in three principal stages, summarized in Table 1.

**Table 1.** Data-cleaning procedure applied in this study.

| Stage | Criterion / Constraint | Operation Description |
|---|---|---|
| **Depth-interval restriction** | Depths below 2888 m and above 3060 m (second well) excluded | Removal of data outside the productive interval and unrelated to permeability |
| **DST-interval matching** | Only data falling within tested DST intervals retained | Synchronization of core and log data with DST intervals |
| **Primary data cleaning** | Removal of NaN and infinite (Inf) values | Identification and correction/removal of erroneous measurements to prevent computational errors |
| **Permeability filtering** | Permeability (Kg) ≤ 0 excluded | Removal of non-physical, non-productive data |
| **Fine permeability filtering** | Permeability < 0.001 mD excluded | Exclusion of non-productive zones with negligible permeability |
| **Data normalization** | Logarithmic transformation of Kg | Normalization of the permeability distribution |
| **Engineered-feature construction** | RES_RATIO = RLA5 / RXOZ | New feature designed to identify hydrocarbon-bearing zones based on the ratio of deep to flushed-zone resistivity |
| | DENS_NEUT_DIFF = NPHI − (RHOB / 2.71) | — |
| **Shale filtering** | SGR < 80 API | Exclusion of shale-dominated intervals to isolate clean reservoir rock |

**Stage 1 — Depth-Interval Restriction:** Core and log data lying outside the productive reservoir interval (i.e., depths shallower than 2888 m or deeper than 3060 m in the second well) were excluded, ensuring that only measurements relevant to the target reservoir zone were retained for subsequent analysis.

**Table 2.** DST intervals and derived permeability values for the second well.

| Depth Interval (m) | Interval Length (m) | Permeability (mD) |
|---|---|---|
| 2888–2900 | 12 | 41.66666667 |
| 2911–2938 | 27 | 18.51851852 |
| 2948–2963 | 15 | 33.33333333 |
| 2977–2985 | 8 | 6.25 |
| 3002–3035 | 33 | 10.60606061 |
| 3046–3060 | 14 | 14.28571429 |

**Stage 2 (DST synchronization):**
Core and log data were retained only within the six intervals of the target well for which corresponding DST measurements were available. This procedure ensured that model training was restricted to depth intervals in which dynamic reservoir data could be directly incorporated.

**Stage 3 (outlier removal):**
At this stage, outlier and non-physical observations were identified and removed on the basis of a combined petrophysical, statistical, and reservoir-engineering assessment. This step is essential to ensure data quality and to prevent the introduction of noise into subsequent modeling. The filtering procedure is detailed below:

- Permeability (Kg) values equal to zero, negative, or below 0.001 mD were excluded, as such intervals are considered non-productive from a reservoir-engineering standpoint and could otherwise mislead the model.
- NaN and ±Inf values, arising from instrumental or computational errors, were removed to retain a clean dataset for model training.
- To improve feature quality and enhance the discriminative capability of the models, several engineered features were constructed, including:
    - RES_RATIO = RLA5 / RXOZ, used as an indicator for identifying hydrocarbon-bearing zones.
    - DENS_NEUT_DIFF = NPHI − (RHOB / 2.71), particularly useful for identifying gas-bearing intervals.
    - GR/K ratio and URAN_NORM, employed to separate shale-dominated and radioactive sand intervals from clean reservoir rock.

### 3.3. Data Correlation and Integration

One of the essential steps in preparing the dataset for permeability modeling involves establishing correlations among core data, DST data, and petrophysical log measurements.

Although core data provide high measurement accuracy at the laboratory scale, they are acquired only at discrete depths. Log data, by contrast, are recorded continuously along the wellbore. To integrate these two data sources, a linear interpolation procedure was applied: the depths at which core samples were recovered were used as reference points, and the corresponding petrophysical log values at those same depths were estimated through linear interpolation. This procedure yielded a combined core–log dataset comprising laboratory-measured permeability values paired with the corresponding log-derived properties at identical depths.

DST measurements, in contrast, provide interval-scale permeability estimates and are typically reported over discrete depth ranges rather than at individual depth points. To integrate these data with the continuous log records, the six DST intervals were first defined as reference zones. A new column was then appended to the log dataset, assigning the DST-derived permeability value to every depth point within its corresponding interval. This procedure produced a DST–log dataset in which the continuous log

measurements were correlated with the interval-based permeability values obtained from DST.

### 3.4. Feature Selection

In this study, the complete set of recorded well logs for the target well was evaluated, and their relationship with permeability was assessed on the basis of reservoir-engineering principles. Logs exhibiting a meaningful relationship with permeability were retained for subsequent modeling, whereas those lacking direct relevance were excluded. The results of this evaluation for the studied well are presented below.

**Table 3.** Evaluation of petrophysical logs for feature selection.

| Log Name | Measured Property | Petroleum-Engineering Interpretation | Final Decision |
|---|---|---|---|
| **CDF** | Flow compaction | Primarily qualitative in nature; not directly related to permeability | Removed |
| **CGR** | Corrected gamma ray | Key indicator for distinguishing shale from clean-sand intervals | Retained |
| **DI_HRLT** | High-resolution resistivity | Corrected resistivity data at multiple depths, related to saturation and permeability | Retained |
| **HCAL, HDRA** | Caliper and corrected density | Quality-control tools; no direct relationship with permeability | Removed |
| **HTEM** | Borehole temperature | Indirect effect on fluid flow; not a priority for permeability modeling | Removed |
| **ICV, IHV** | Volumetric flow indicators | Directly related to fluid-flow capacity within the formation; useful for permeability prediction | Retained |
| **NPHI, PEFZ, RHOZ** | Porosity and density | Among the most important logs for permeability modeling | Retained |
| **RLA1–5, RXOZ** | Deep and shallow resistivity | Indicative of fluid saturation and permeability | Retained |
| **SGR, THOR, URAN** | Radioactivity logs | Indirect influence through shale identification | Retained |
| **TENS** | Rock tensile strength | Primarily applicable to rock mechanics; not directly related to permeability | Removed |

### 3.5. Training and Testing Data Partitioning

Following the completion of data acquisition, cleaning, correlation, and feature-selection procedures, the prepared dataset was partitioned into training and testing subsets prior to model development. This step represents a fundamental component of machine-learning model design, as it enables a rigorous and unbiased assessment of predictive performance.

Training dataset. The training subset comprised the majority of the available data (70% of the total dataset) and was used to train the machine-learning algorithms. Based on the reservoir-engineering analysis conducted during the feature-selection stage, the

petrophysical parameters exhibiting the strongest association with permeability were selected as the final model inputs.

Testing dataset. The testing subset consisted of the actual permeability values obtained from DST measurements and core data. These observations were withheld entirely from the training process and were used exclusively for model validation and for assessing generalization capability. In effect, the testing subset served as unseen data from the model's perspective, thereby providing a reliable basis for evaluating predictive accuracy.

## 3.6. Research Methodology

### 3.6.1. Case-Study Methodology

The dataset employed in this study was acquired from an Iranian gas field. This field was selected with the aim of developing an accurate and generalizable model for permeability estimation based on petrophysical logs and well-test data. A single producing well from the field, designated SPD14B-08, was examined in this investigation. The collected data were organized into two principal categories. The first category comprised the DST–log dataset, in which pressure–time records obtained from Drill Stem Tests (DST) were matched with the complete suite of petrophysical logs. The second category comprised the core–log dataset, which combined log measurements with core-analysis results; owing to its capacity to provide real, continuous permeability values, this dataset was employed in the development of regression-based models.

The reliability of the results was ensured by partitioning the data into training and testing subsets and comparing the model predictions against the actual permeability values derived from DST and core measurements. This combined dataset structure enabled the simultaneous investigation of reservoir behavior from both classification and regression perspectives, thereby yielding a comprehensive and reliable characterization of permeability within the studied field.

### 3.6.2. Modeling Methodology

**Design of the machine-learning models**

In this study, the development of the machine-learning models followed two distinct pathways, since the nature of the output data differed between the two data sources (DST and core) and, accordingly, required different modeling approaches:

1. Modeling based on DST–log data
   Because the data derived from the Drill Stem Test (DST) are reported as discrete, classified permeability intervals, this problem was formulated as a classification task.
2. Modeling based on core–log data
   In this case, core-derived permeability values are continuous in nature, and the problem was accordingly formulated as a regression task.

**(a) Classification modeling using DST data**

The objective of this section was to identify permeability classes based on a comprehensive set of log-derived features. To this end, several base algorithms, together with a hybrid ensemble model, were developed.

*1. Logistic Regression*

- Selected as a simple and interpretable method, serving as a baseline model against which the performance of more complex algorithms could be compared.
- Advantages: simplicity, stability, and computational efficiency.
- Limitations: reliance on the assumption of a linear decision boundary, resulting in reduced performance on datasets exhibiting nonlinear patterns [39].

*2. K-Nearest Neighbors (KNN) Classifier*

- Well suited to identifying permeability classes located within similar depth ranges.
- Advantages: requires no complex model construction; performs well on data exhibiting local distributional structure.
- Limitations: sensitive to noise and to increasing feature dimensionality [40].

*3. Support Vector Machine (SVM) Classifier*

- Achieves strong performance in resolving nonlinear boundaries between DST-derived permeability classes.
- Advantages: capable of identifying optimal decision boundaries; well suited to high-dimensional data.
- Limitations: high computational cost and sensitivity to kernel selection [41].

*4. Decision Tree Classifier*

- The simple, tree-based structure of this model allows the direct interpretation of the influence of each log on permeability classification.
- Advantages: interpretable and transparent.
- Limitations: susceptible to overfitting on small datasets [23].

**Ensemble Classification**

To improve predictive accuracy relative to the individual models, a hybrid ensemble model was developed, combining the following four algorithms:

- Decision Tree Classifier
- Logistic Regression
- K-Nearest Neighbors Classifier
- SVM Classifier

Using techniques such as the Voting Classifier or Stacking, this model combines the outputs of multiple algorithms in order to:

- reduce the error associated with individual models, and
- improve the overall accuracy of permeability-class prediction [26].

**(b) Regression modeling using core data**

In this section, permeability is treated as a continuous numerical variable, and the problem was accordingly analyzed within a regression framework.

*1. Decision Tree Regression*

- A baseline regression model with a relatively simple structure.
- Advantages: interpretable; capable of capturing complex data patterns.
- Limitations: prone to overfitting on small datasets.

*2. KNN Regression*

- Well suited to estimating permeability at neighboring depths.
- Advantages: simple in structure and effective for densely sampled data.
- Limitations: sensitive to feature scaling and to the presence of noise [40].

*3. Support Vector Regression (SVR)*
A robust algorithm for learning nonlinear relationships between permeability and log-derived features.

- Advantages: high flexibility when an appropriate kernel is selected.
- Limitations: relatively high computational cost.

**Ensemble Regression**

To achieve the highest possible predictive accuracy, a three-algorithm hybrid ensemble model was constructed, comprising:

- K-Nearest Neighbors Regressor
- Support Vector Regression (SVR)
- Decision Tree Regressor

By combining the outputs of these individual models—through weighted averaging or stacking—this approach yielded:

- reduced RMSE and MSE values,
- improved model stability, and
- reduced prediction uncertainty [42].

3.6.3. Workflow of Classification and Ensemble Learning Models

Figure 1 presents the workflow developed for permeability classification using the DST–Log dataset. Following data acquisition and preprocessing, four standalone classifiers—Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and

Decision Tree—were implemented as base learners. Each model was trained using cross-validation combined with hyperparameter tuning, after which performance was evaluated on the testing dataset using accuracy, precision, recall, F1-score, ROC–AUC, and confusion-matrix analysis.

To enhance predictive stability and mitigate the limitations inherent to individual algorithms, a hybrid ensemble-learning framework based on Voting and Stacking techniques was subsequently developed by combining the outputs of all standalone classifiers. This ensemble model integrates linear, nonlinear, distance-based, and tree-based learning behaviors, thereby providing more robust permeability classification under heterogeneous reservoir conditions. In the final stage, the performance of the standalone and ensemble models was comparatively evaluated in order to identify the most reliable classification framework.

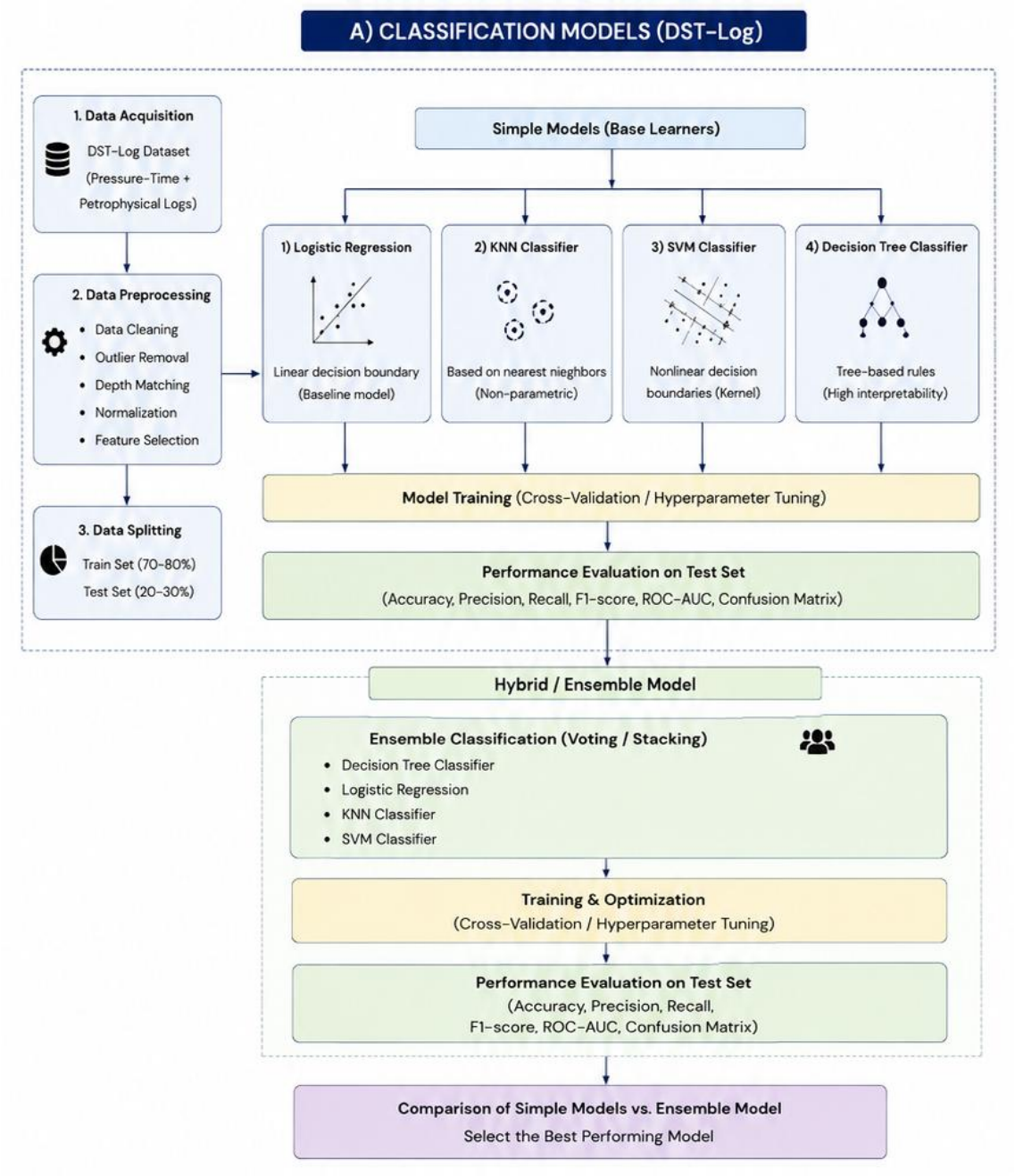



**Figure 1.** Integrated workflow of classification and ensemble-learning models developed for DST–Log permeability prediction.

### 3.6.4. Workflow of Regression and Ensemble Learning Models

Figure 2 illustrates the corresponding workflow developed for continuous permeability prediction using the Core–Log dataset. Following data preprocessing and train–test partitioning, three standalone regression models—Decision Tree Regressor, KNN Regressor, and Support Vector Regressor (SVR)—were developed and subsequently optimized through cross-validation and hyperparameter tuning.

Model performance was evaluated using RMSE, MSE, MAE, $R^2$, and adjusted $R^2$. A hybrid ensemble-regression framework, likewise based on Voting/Stacking strategies, was then implemented to integrate the predictive strengths of the individual regressors. Relative to the standalone models, this ensemble approach yielded improved stability, reduced prediction uncertainty, and superior generalization capability across heterogeneous reservoir intervals. In the final stage, a comparative analysis between the simple and hybrid models was carried out to determine the optimal permeability-prediction framework.

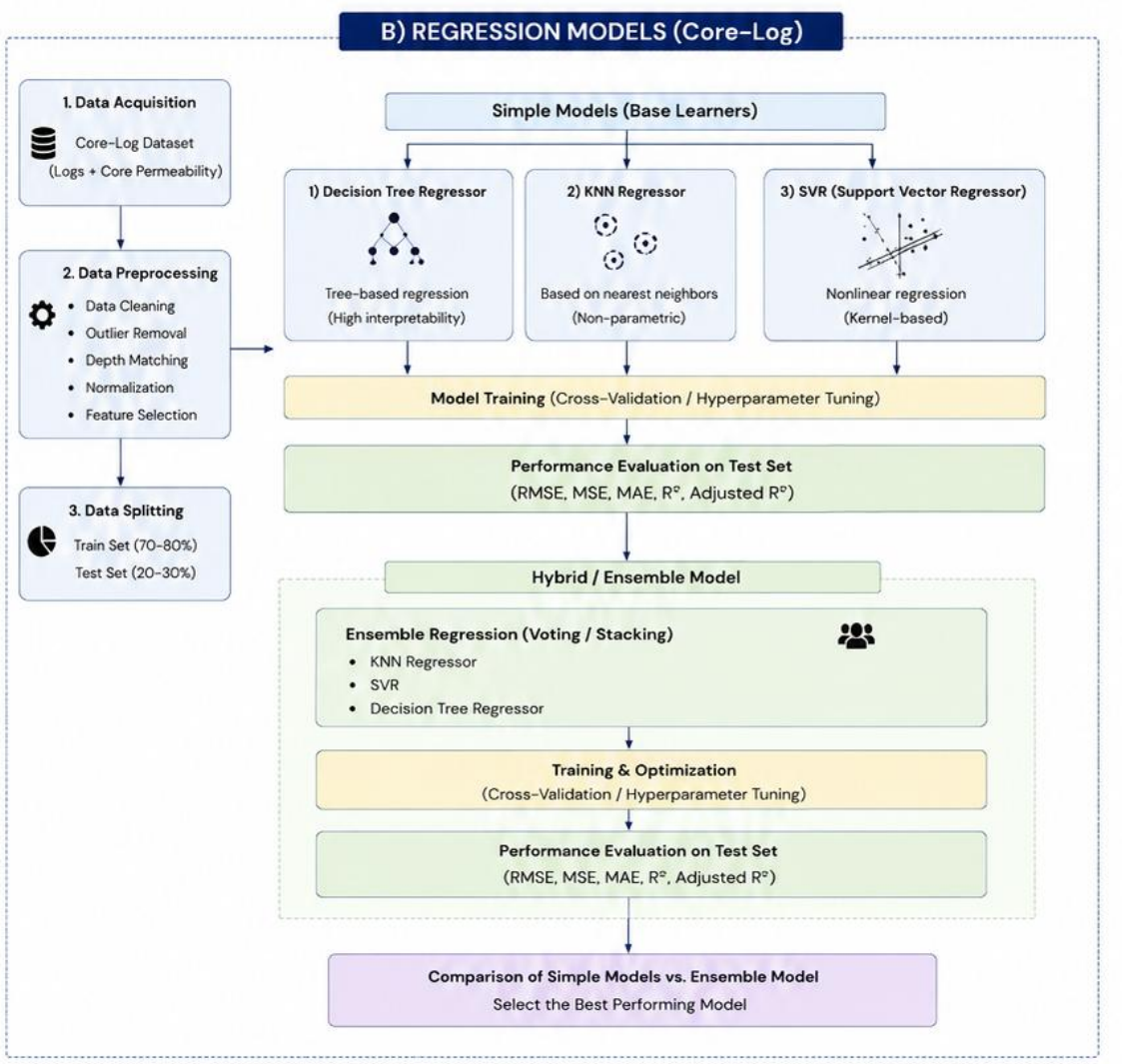



**Figure 2.** Integrated workflow of regression and ensemble-learning models developed for Core–Log permeability prediction.

### 3.7. Model Training and Validation Using Training and Testing Data

In this study, the machine-learning modeling procedure was focused on the regression-based Log–Core dataset, and a series of filtering and preprocessing steps was applied to improve the quality of the input features. These filtering procedures were designed in accordance with the geological and petrophysical characteristics of the reservoir and

differed between algorithm groups (tree-based and ensemble models versus distance-based and neighborhood algorithms).

### 3.7.1. Group I: KNN-Based Algorithms

**Table 4.** Filtering procedures and engineered features applied for the KNN algorithm.

| Step | Filter / Feature | Description | Petroleum-Engineering Justification |
|---|---|---|---|
| **1** | `df['RES_RATIO'] = df['RLA5'] / df['RXOZ']` | Ratio of deep to flushed-zone resistivity | Indicates the presence of hydrocarbons; when RLA5 ≫ RXOZ, the flushed zone is conductive (saline water) while the deeper zone is resistive (hydrocarbon-bearing) |
| **2** | `df['PORO_RES_INDEX'] = df['NPHI'] * np.log10(df['RLA5'])` | Porosity–resistivity crossplot index | Zones exhibiting both high porosity and high resistivity generally correspond to the highest reservoir quality |
| **3** | `df['DENS_NEUT_DIFF'] = df['NPHI'] - df['RHOB']/2.71` | Neutron–density difference | — |
| **4** | `df['GR_K_RATIO'] = df['SGR'] / (df['POTA'] + 0.1)` | Gamma-to-potassium ratio | Distinguishes uranium/thorium-rich shales from clean sand intervals |
| **5** | `df['URAN_NORM'] = df['URAN'] / (df['SGR'] + 0.1)` | Normalized uranium content | Removes the non-physical gamma-ray contribution associated with radioactive sands |
| **6** | `df['LOG_PERM'] = np.log(df['PERM'])` | Logarithm of permeability | Produces a log-normal permeability distribution, resulting in a more stable model |

This advanced filtering procedure is particularly critical for neighborhood-based algorithms such as KNN, as these models are highly sensitive to data scaling, and the presence of outlier or non-physical observations can lead to overfitting or improper convergence [43].

### 3.7.2. Group II: Ensemble and Decision Tree Algorithms

**Table 5.** Filtering procedures applied for the ensemble and decision-tree algorithms.

| Step | Filter | Description | Petroleum-Engineering Justification |
|---|---|---|---|
| **1** | `df = df.replace([np.inf, -np.inf], np.nan).dropna()` | Removal of infinite and missing values | Infinite or NaN values, typically arising from instrumental error or division by zero, render decision-tree splits meaningless if retained |

| | | | |
|---|---|---|---|
| **2** | `df = df[df['Kg'] > 0]` | Removal of negative or zero permeability values | Permeability is physically meaningful only in the presence of a defined flow pathway; values ≤ 0 are non-physical |
| **3** | `df['Kg_log'] = np.log10(df['Kg'])` | Logarithmic transformation of permeability | Reduces data skewness and provides better compatibility with the wide dynamic range of Kg (from μD to kD) |

These preprocessing steps are particularly essential for tree-based and ensemble algorithms, as such models are sensitive to outlier observations, and the presence of non-physical data can lead to the formation of erroneous branches within the decision tree [41].

Overall, the data-preprocessing procedures were designed to achieve a closer correspondence between the different modeling approaches and the underlying geological and reservoir characteristics. The results demonstrated that the application of these filtering steps and engineered composite features improved the predictive capability of the models with respect to permeability estimation, while simultaneously mitigating errors arising from non-physical and noisy data.

### 3.8. Hyperparameter Tuning

One of the key stages in the development of machine-learning models is the optimal selection of hyperparameters. Hyperparameters are values defined prior to model training and exert a direct influence on the final performance and predictive accuracy of the model. Unlike weights and biases, which are updated automatically by the learning algorithm during training, hyperparameters must be selected through dedicated search and tuning procedures.

In this study, two widely used hyperparameter-optimization methods were employed:

- **GridSearchCV:** This method exhaustively evaluates all possible combinations of hyperparameters and selects the optimal configuration based on a specified performance metric, such as accuracy or prediction error [24].
- **RandomizedSearchCV:** In cases where the hyperparameter search space was excessively large, randomized search was instead employed to more efficiently identify near-optimal parameter combinations [44].

Using these two approaches, the key hyperparameters for each algorithm were identified and tuned. A detailed description of the tuning procedure for each model is provided below.

Category I: Classification Models

For data derived from Drill Stem Test (DST) measurements, which are inherently categorical in nature, several supervised classification algorithms were implemented. To

improve predictive accuracy and to prevent overfitting and underfitting, hyperparameter optimization was regarded as an essential component of the modeling workflow.

In the Logistic Regression model, the most important hyperparameter is *C* (the regularization strength), which determines the degree to which the model penalizes prediction errors. Selecting a smaller value of *C* results in a simpler model with improved generalization capability, whereas larger values increase the risk of overfitting. In addition, the choice of solver—such as *lbfgs* or *saga*—plays an important role in determining the speed and numerical stability of the model.

In the K-Nearest Neighbors model, the primary hyperparameter is *n_neighbors*, which specifies the number of neighboring samples considered during prediction. Small values result in high sensitivity to noise, whereas larger values produce smoother predictions. The choice of distance metric is likewise of considerable importance; for instance, the use of Euclidean versus Manhattan distance, selected according to the petrophysical characteristics of the data, can yield markedly different classification outcomes [45].

In the Support Vector Machine model, three key hyperparameters were tuned: *C* (the error penalty), *kernel* (the choice of kernel function, such as RBF or polynomial), and *gamma* (the radius of influence of individual training points). Proper tuning of these parameters enables the model to capture complex nonlinear relationships between petrophysical logs and permeability classification.

In the Decision Tree model, parameters such as *max_depth* (tree depth) and *criterion* (the splitting criterion, e.g., Gini or entropy) were tuned. Increasing tree depth allows the model to capture finer data detail but simultaneously increases the risk of overfitting. To address this issue, decision trees were further combined within a Random Forest framework, in which hyperparameters such as *n_estimators* (the number of trees) and *max_features* (the number of features considered at each split) play a key role in enhancing model stability and reducing variance.

Ensemble-learning techniques, including Bagging and Boosting, were also employed. In these models, the choice of *base_estimator* and the number of iterations (*n_estimators*) determine the model's capacity to correct the errors of preceding learners [46]. Among these approaches, XGBoost stands out as one of the most powerful algorithms; in addition to parameters such as *n_estimators* and *max_depth*, it further employs *learning_rate*, *subsample*, and *colsample_bytree* to maintain an appropriate balance between predictive accuracy and generalization capability.

**Table 6.** Optimal hyperparameter selection for Logistic Regression applied to permeability prediction using DST–log data.

| Hyperparameter | Search Range | Role in Model | Petroleum-Engineering Justification |
|---|---|---|---|
| **C** | [0.01, 0.1, 1, 10, 100] | L2 regularization strength | Smaller values prevent overfitting, whereas larger values increase sensitivity to training data; for noisy log data, smaller values are generally preferred |

| | | | |
|---|---|---|---|
| **solver** | ['liblinear', 'lbfgs'] | Optimization algorithm | *lbfgs* is better suited to larger datasets, whereas *liblinear* is more appropriate for smaller ones |
| **penalty** | ['l2'] | Regularization type | Stabilizes model performance on noisy petrophysical data |
| **class_weight** | ['balanced', None] | Class-weight adjustment | Compensates for the imbalance in kh intervals within the DST dataset |
| **max_iter** | [100, 200, 500] | Number of convergence iterations | Increases the likelihood of achieving a stable solution for complex reservoir data |

**Table 7.** Optimal hyperparameter selection for the KNN Classifier applied to permeability prediction using DST–log data.

| Hyperparameter | Search Range | Role in Model | Petroleum-Engineering Justification |
|---|---|---|---|
| **n_neighbors** | 1–30 | Number of neighbors | Small values result in overfitting, whereas large values result in underfitting |
| **weights** | ['uniform', 'distance'] | Weighting scheme | Distance-based weighting provides greater accuracy near fluid-contact boundaries |
| **metric** | ['euclidean', 'manhattan'] | Distance metric | Selecting an appropriate metric improves performance across features with differing scales |

**Table 8.** Optimal hyperparameter selection for the Support Vector Machine (SVM) applied to permeability prediction using DST–log data.

| Hyperparameter | Search Range | Role in Model | Petroleum-Engineering Justification |
|---|---|---|---|
| **C** | [0.1, 1, 5, 10] | Error penalty | Larger values yield a more precise model but increase the risk of overfitting |
| **kernel** | ['rbf', 'linear', 'poly'] | Kernel function type | The RBF kernel is well suited to nonlinear patterns commonly observed in well-log data |
| **gamma** | ['scale', 'auto'] | Influence radius | Determines the influence of individual training points on the decision hyperplane |
| **degree** | [2, 3] | Polynomial-kernel degree | Enables the modeling of complex nonlinear relationships |

**Table 9.** Optimal hyperparameter selection for the Decision Tree Classifier applied to permeability prediction using DST–log data.

| Hyperparameter | Search Range | Role in Model | Petroleum-Engineering Justification |
|---|---|---|---|
| **max_depth** | 1–15 | Tree depth | Controls model complexity; excessive depth leads to overfitting |

| | | | |
|---|---|---|---|
| **criterion** | ['gini', 'entropy'] | Splitting criterion | Determines the most informative feature for classification |
| **splitter** | ['best', 'random'] | Splitting strategy | The *best* strategy is generally more stable when applied to petrophysical data |
| **class_weight** | ['balanced', None] | Class balancing | Prevents bias toward the more heavily represented classes |

**Table 10.** Optimal hyperparameter selection for the ensemble classifiers (Bagging/Boosting) applied to permeability prediction using DST–log data.

| **Algorithm** | **Key Hyperparameters** | **Role** | **Petroleum-Engineering Justification** |
|---|---|---|---|
| **Logistic Regression** | C, solver | Improves the bias–variance trade-off | Low sensitivity to noise |
| **KNN** | n_neighbors | Controls neighborhood size | Improved lithofacies separation |
| **SVM** | C, kernel | Learns nonlinear decision boundaries | Well suited to complex datasets |
| **Decision Tree** | max_depth | Controls model complexity | High flexibility for reservoir data |

Category II: Regression Models

For data derived from core analysis, which comprise continuous permeability values, several regression-based models were employed. Hyperparameter tuning was of particular importance in this category, as the model output is a continuous quantity and prediction error must be minimized as far as possible.

In the Decision Tree Regressor, parameters such as *max_depth* and *min_samples_split* were tuned in order to provide the model with sufficient complexity to capture nonlinear relationships while simultaneously preventing overfitting. The splitting criterion was likewise selected based on either mean squared error or mean absolute error [47].

In the KNN Regressor, the hyperparameter *n_neighbors* governs the degree of locality in the resulting predictions. In addition, the choice of weighting scheme—uniform versus distance-based—was evaluated. The results indicated that distance-based weighting yielded better performance on the log–core dataset, since nearer samples exert a proportionally greater influence on the prediction [48].

Finally, Ensemble Regression methods, which combine multiple base learners, were employed to collectively compensate for the individual weaknesses of each model. By tuning parameters such as *n_estimators* and *learning_rate*, these methods were able to produce stable and accurate permeability predictions from the core–log dataset.

**Table 11.** Optimal hyperparameter selection for the Decision Tree Regressor applied to permeability prediction using core–log data.

| **Hyperparameter** | **Search Range** | **Role in Model** | **Petroleum-Engineering Justification** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| **max_depth** | 1–15 | Tree depth | Balances predictive accuracy and generalization capability |
| **criterion** | ['squared_error', 'friedman_mse'] | Splitting criterion | MSE is best suited to continuous permeability data |

**Table 12.** Optimal hyperparameter selection for the ensemble regressor (Voting), combining KNN, SVR, and Decision Tree, applied to core–log data.

| Algorithm | Key Hyperparameters | Role | Petroleum-Engineering Justification |
|---|---|---|---|
| **KNN** | n_neighbors | Local prediction accuracy | Captures rapid variations in permeability |
| **SVR** | C, epsilon | Error control | Adjusts sensitivity to non-normally distributed data |
| **Decision Tree** | max_depth | Complexity control | High flexibility for nonlinear reservoir data |

**Table 13.** Optimal hyperparameter selection for the KNN Regressor applied to permeability prediction using core–log data.

| Hyperparameter | Search Range | Role in Model | Petroleum-Engineering Justification |
|---|---|---|---|
| **n_neighbors** | 1–20 | Number of neighbors | Controls the bias–variance trade-off |
| **weights** | ['uniform', 'distance'] | Weighting scheme | Distance-based weighting is more physically reasonable for reservoir data |
| **metric** | ['euclidean', 'manhattan'] | Distance metric | Different metrics exhibit varying sensitivity to feature scale |

### 3.9 Research Novelty and Contributions

Several aspects distinguish the present work from earlier attempts at machine-learning-based permeability estimation. Most published studies in this area rely on a single class of algorithm applied to a single, self-contained dataset—typically either core measurements or well logs in isolation—without explicitly reconciling the static and dynamic characterizations of the reservoir. In contrast, the framework developed here couples classification of DST–log data with regression on core–log data within one coherent workflow, allowing the two complementary views of reservoir behavior (interval-based dynamic response versus continuous static properties) to inform a single, unified permeability characterization.

A second contribution concerns feature engineering. Rather than applying a generic set of input variables uniformly across all algorithms, reservoir-specific composite features—such as resistivity ratios, neutron–density difference indices, and normalized radioactivity ratios—were constructed and tailored separately to the requirements of each algorithm group (distance-based versus tree-based/ensemble methods). This algorithm-aware feature design was found to be a meaningful factor in the improved predictive performance reported in Section 4, rather than a secondary preprocessing step.

Third, model interpretability was treated as a central objective rather than an auxiliary output. SHAP analysis was applied systematically across all classification and regression models to quantify the contribution of individual petrophysical variables, enabling the predictions to be evaluated not only in terms of statistical accuracy but also against established petrophysical reasoning—for example, the expected inverse relationship between gamma-ray response and permeability, or the positive association between resistivity contrast and hydrocarbon-bearing intervals.

Finally, by directly comparing standalone algorithms against hybrid Voting/Stacking ensembles under identical preprocessing and validation conditions, this study provides quantitative evidence—rather than a general assertion—of the extent to which Ensemble Learning improves stability, reduces prediction uncertainty, and enhances generalization when petrophysical data are noisy or reservoir conditions are heterogeneous. This comparative evidence, together with the integrated DST–core–log workflow, forms the principal contribution of the present study to the reservoir-characterization literature.

# 4. Results

This section presents and discusses the results obtained from the application of the developed machine-learning models to data from a selected well within the field. Two principal datasets were used: (1) the DST–log dataset, which enabled the investigation of the dynamic behavior of the reservoir, and (2) the core–log dataset, which served as the most accurate laboratory-scale reference for permeability at the local scale.

Prior to modeling, the data were subjected to a series of preprocessing steps, including outlier removal, normalization, synchronization with DST intervals, and selection of key petrophysical features. The data were then partitioned into training and testing subsets, and cross-validation was applied in certain cases to improve model stability and reduce the risk of overfitting.

A range of classification and regression algorithms was employed for permeability prediction. In addition, SHAP analysis was used to enhance model interpretability and to quantify the influence of each input variable on the predicted permeability.

## 4.1. Well Results

4.1.1. Accuracy and Error Tables

### 1) Classification models

Tables 14 and 15 present the accuracy of the classification models. As with the previously discussed well, the SVM, Decision Tree Classifier, and Ensemble Learning models achieved perfect accuracy (1.0) on both the training and testing datasets, reflecting their capacity to capture the nonlinear relationships among key input features such as the RLA resistivity logs, GR, and NPHI. Logistic Regression, by contrast, exhibited lower accuracy and performed less effectively in resolving class boundaries.

**Table 14.** Graphical accuracy results for the classification models on the training and testing datasets.

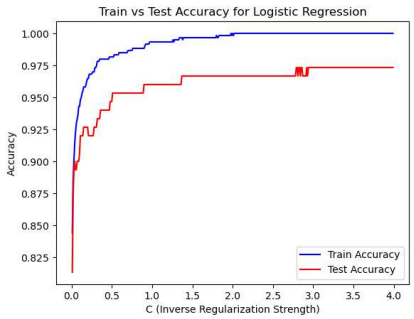
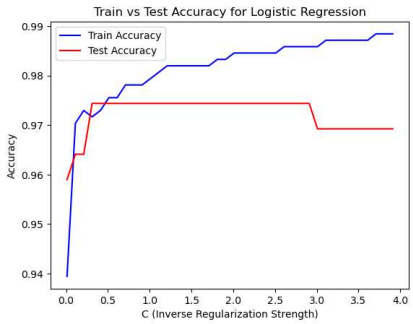
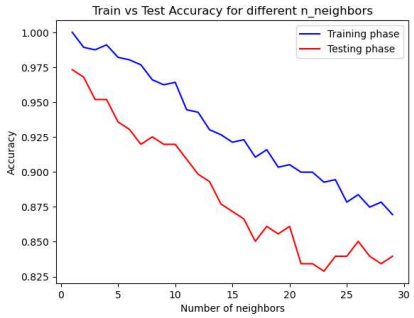
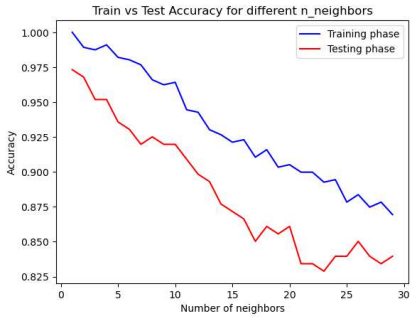
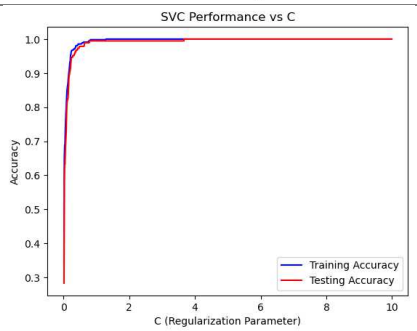
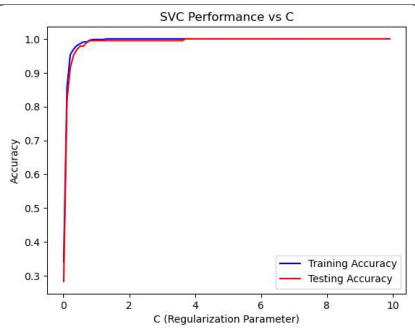
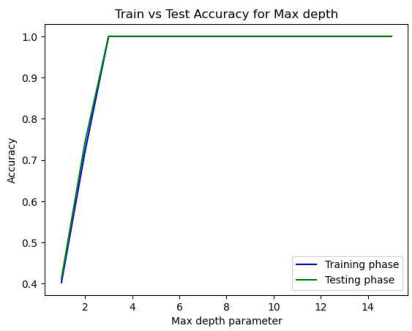
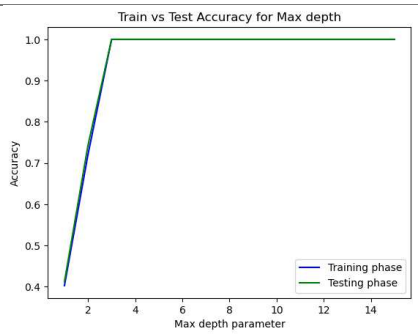
| Model | Simple | Hyperparameters |
|---|---|---|
| **Logistic Regression** |  |  |
| **KNN Classifier** |  |  |
| **SVM** |  |  |
| **Decision Tree Classifier** |  |  |

**Table 15.** Accuracy values of the classification models on the training and testing datasets.

| Model | Training Accuracy (Simple) | Testing Accuracy (Simple) | Training Accuracy (Tuned) | Testing Accuracy (Tuned) |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| **Logistic Regression** | 0.99 | 0.99 | 1.0 | 0.973 |
| **KNN Classifier** | 0.94 | 0.89 | 1.0 | 0.989 |
| **SVM** | 0.996 | 0.994 | 1.0 | 1.0 |
| **Decision Tree Classifier** | 1.0 | 1.0 | 1.0 | 1.0 |
| **Ensemble Learning** | 0.989 | 0.986 | 1.0 | 1.0 |

Table 16 presents the corresponding error metrics for the classification models in this well. The Decision Tree and SVM models achieved the lowest error values, with RMSE approaching zero, indicating superior performance, whereas Logistic Regression exhibited a somewhat higher error (approximately 0.12–0.2).

**Table 16.** Error metrics of the classification models on the training and testing datasets.

| Model | RMSE (Simple) | MSE (Simple) | MAE (Simple) | RMSE (Tuned) | MSE (Tuned) | MAE (Tuned) |
|---|---|---|---|---|---|---|
| **Logistic Regression** | 3.89 | 15.13 | 3.39 | 3.79 | 14.40 | 2.28 |
| **KNN Classifier** | 0.40 | 0.16 | 0.12 | 0.50 | 0.25 | 0.19 |
| **SVM** | 0.07 | 0.005 | 0.005 | 0.0 | 0.0 | 0.0 |
| **Decision Tree Classifier** | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **Ensemble Learning** | 0.013 | 0.013 | — | 0.0 | 0.0 | 0.0 |

Overall, the SVM, Decision Tree Classifier, and Ensemble Learning models produced the strongest results, with accuracies approaching 1.0 and near-zero prediction error. These models proved capable of effectively learning the complex relationships between key petrophysical variables (such as NPHI, RHOZ, CGR, and the resistivity logs) and permeability class. Simpler models, such as Logistic Regression, achieved acceptable performance overall but showed reduced accuracy near class boundaries.

### 2) Regression models

Tables 17 and 18 summarize the accuracy of the regression models. In this well, the Decision Tree Regressor again achieved the best overall performance. The combination of NPHI and RHOZ—representing porosity and bulk density—together with the electrical resistivity logs (RLA series), played the dominant role in permeability prediction. This finding is consistent with established petrophysical reasoning, as clean sand intervals characterized by high porosity and low gamma-ray response typically exhibit higher permeability.

**Table 17.** Graphical accuracy results for the regression models on the training and testing datasets.

| Model | Simple | Hyperparameters |
|---|---|---|

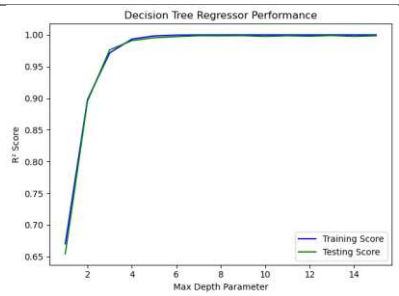
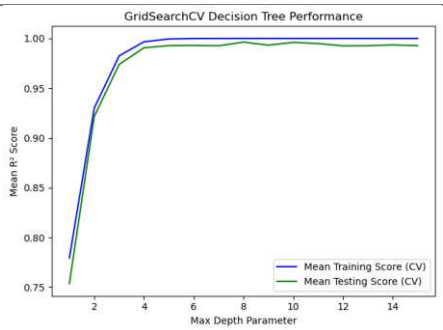
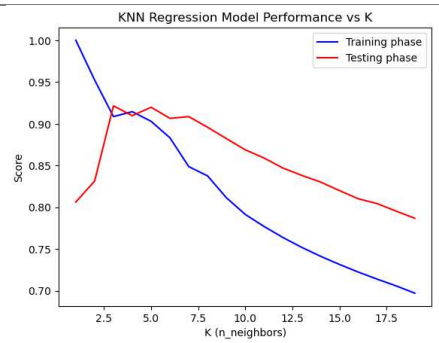
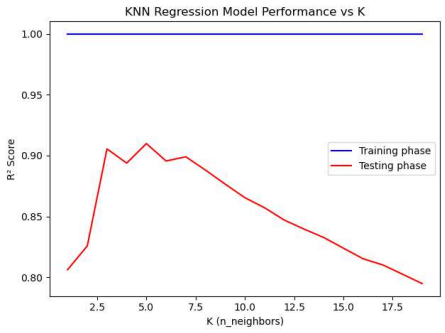
| | | |
|---|---|---|
| **Decision Tree Regressor** |  |  |
| **KNN Regressor** |  |  |

**Table 18.** Accuracy values of the regression models on the training and testing datasets.

| Model | Training Accuracy (Simple) | Testing Accuracy (Simple) | Training Accuracy (Tuned) | Testing Accuracy (Tuned) |
|---|---|---|---|---|
| **Decision Tree Regressor** | 0.897 | 0.895 | 0.999 | 0.980 |
| **Ensemble Regressor** | 0.957 | 0.955 | 0.942 | 0.913 |
| **KNN Regressor** | 0.952 | 0.831 | 0.999 | 0.899 |

Table 19 presents the corresponding error metrics. In this well, the Decision Tree Regressor exhibited the lowest error (RMSE approaching zero), whereas the KNN Regressor showed the weakest performance, with comparatively higher error values. The error associated with the MLP model was moderate, indicating that although neural networks are capable of capturing complex relationships, their performance remains strongly dependent on careful architectural tuning and effective noise suppression.

**Table 19.** Error metrics of the regression models on the training and testing datasets.

| Model | RMSE (Simple) | MSE (Simple) | MAE (Simple) | RMSE (Tuned) | MSE (Tuned) | MAE (Tuned) |
|---|---|---|---|---|---|---|
| **Decision Tree Regressor** | 0.055 | 0.003 | 0.027 | 0.159 | 0.025 | 0.064 |
| **Ensemble Regressor** | 0.239 | 0.057 | 0.192 | 0.335 | 0.112 | 0.259 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **KNN Regressor** | 0.0024 | 0.0 | 0.0008 | 0.001 | 0.0 | 0.0006 |

In the regression analysis, the ensemble-based regressor achieved the strongest overall performance, with an RMSE below 0.05 on the testing dataset, whereas the Decision Tree Regressor exhibited weaker generalization owing to overfitting on the training data.

From a geological standpoint, porosity (NPHI), bulk density (RHOZ), and the resistivity logs contributed most substantially to the reconstruction of continuous permeability.

**Overall comparison between classification and regression**

- In the classification task, the objective was to identify discrete permeability classes; tree-based models provided the best performance.
- In the regression task, the objective was to predict continuous permeability values; tree-based models again yielded the highest accuracy.
- From a geological perspective, the results collectively confirm that clean sand intervals characterized by high NPHI and low GR are reliably identified, and that electrical resistivity logs play a fundamental role in delineating hydrocarbon-saturated zones.

4.1.2. SHAP Analysis

To further examine the behavior of the trained models and to identify the petrophysical parameters exerting the greatest influence on permeability prediction, SHAP (SHapley Additive exPlanations) analysis was applied. This method quantifies the contribution of each input feature to the model output and thus provides a powerful basis for interpreting model decision-making [49].

**1) Classification models**

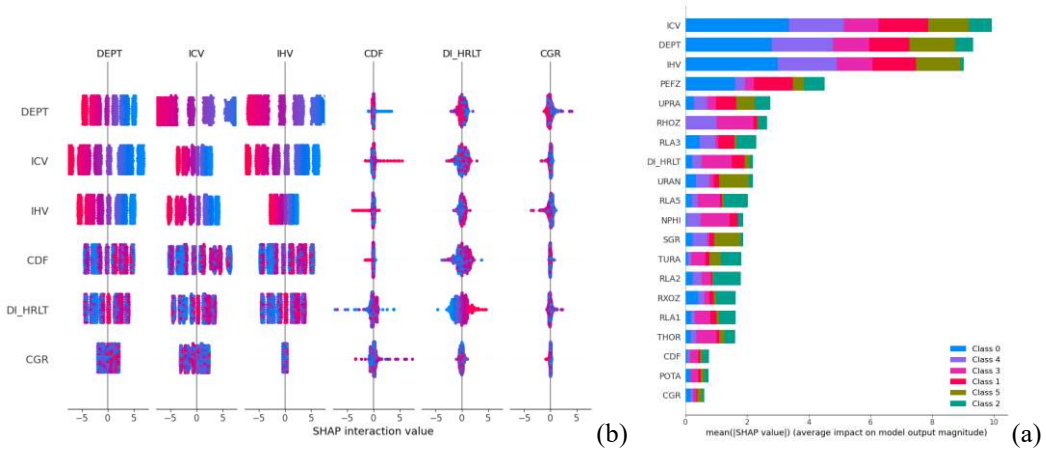



**Figure 3.** SHAP results for the Logistic Regression model: (a) feature-importance bar plot; (b) feature-interaction summary plot for permeability classification.

The summary plot (panel a) illustrates the directional influence of each feature on the predicted permeability class. The dominant contributions are associated with the gamma-ray parameters CGR and SGR, followed by neutron porosity (NPHI) and bulk density (RHOZ). High CGR/SGR values (red points with negative SHAP values) are clearly correlated with a reduced likelihood of higher-permeability classes, whereas high NPHI values (blue points on the positive side) increase this likelihood. RHOZ exhibits an inverse trend, whereby higher bulk density—corresponding to lower porosity—is associated with negative SHAP values. Despite being a linear classifier, Logistic Regression thus captures the direction of the dominant feature effects in a manner consistent with expected petrophysical behavior.

The feature-importance plot (panel b) reports the mean absolute SHAP value for each variable; CGR and SGR exhibit the highest average magnitude, followed by NPHI and RHOZ. From a reservoir-engineering standpoint, this indicates that the linear decision boundary of the Logistic Regression model is governed primarily by shale content and porosity, implying a correspondingly elevated risk of misclassification in intervals where shale and sand signatures overlap.

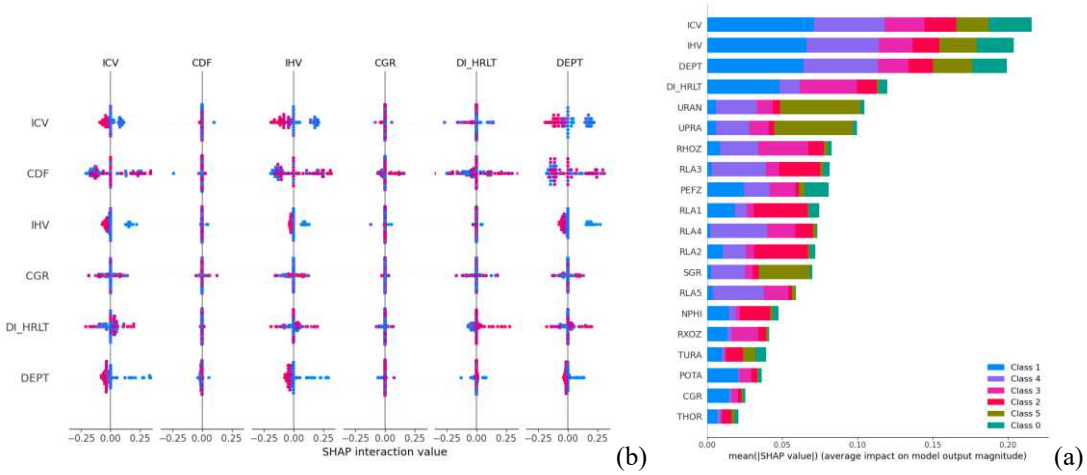



**Figure 4.** SHAP results for the K-Nearest Neighbors (KNN) model: (a) feature-importance bar plot; (b) feature-interaction summary plot for permeability classification.

For the KNN model, the summary plot indicates that NPHI and RHOZ dominate the local prediction behavior, reflecting a strong local correlation between porosity, bulk density, and permeability class among neighboring samples; the corresponding NPHI points cluster on the positive side of the plot, consistent with increasing porosity raising the probability of assignment to a higher-permeability class. RXOZ and several of the deeper resistivity logs contribute a smaller but favorable effect in the same direction. The overall scatter of points further reflects the inherent sensitivity of KNN to local data structure, with sparsely sampled or outlier-adjacent regions estimated less reliably.

Consistent with this pattern, the feature-importance plot shows NPHI and RHOZ as the two most influential variables, followed by SGR and RXOZ. Taken together, these results suggest that KNN performs well in reservoir intervals with dense, homogeneous

sampling, while its accuracy deteriorates in heterogeneous or boundary zones characterized by substantial feature overlap.

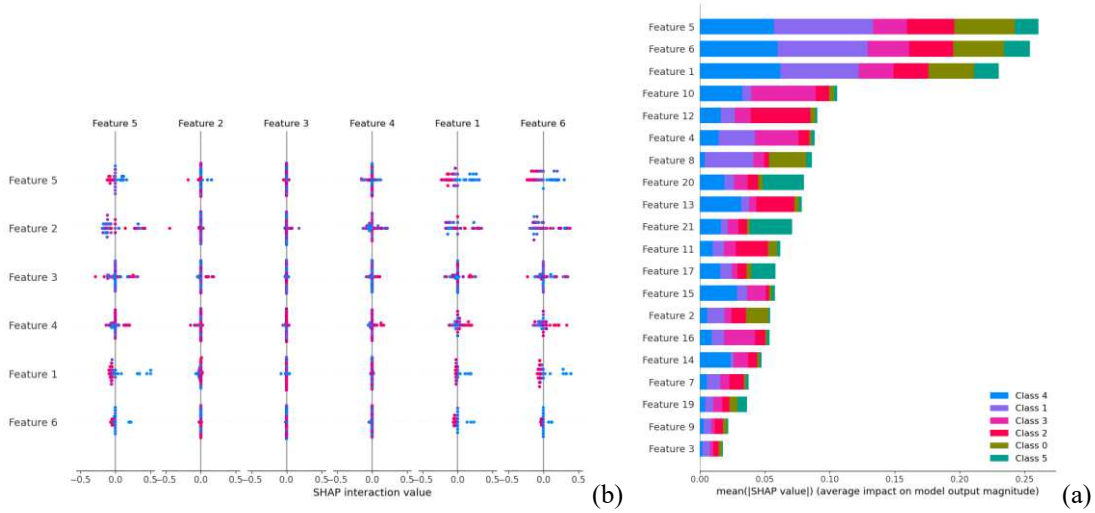



**Figure 5.** SHAP results for the Support Vector Machine (SVM) model: (a) feature-importance bar plot; (b) feature-interaction summary plot for permeability classification.

A markedly different pattern emerges for the SVM model, whose predictions are most sensitive to the deep-resistivity group—the RLA series together with RXOZ. Elevated resistivity values are generally associated with positive SHAP contributions, reflecting an increased probability of higher-permeability classes, while CGR and NPHI occupy a secondary role, with high CGR mapping to negative SHAP values and high NPHI to positive ones. The scatter of SHAP values across these features reflects the nonlinear class separation achieved through the selected kernel.

The corresponding importance plot confirms that the electrical resistivity logs (RLA series and RXOZ) represent the principal decision drivers for the SVM model, ahead of CGR and NPHI. This pattern is consistent with the capacity of SVM to exploit nonlinear interactions, effectively identifying the combination of elevated resistivity and favorable porosity as the strongest joint indicator of high permeability—an outcome that aligns closely with the theoretical strength of SVM in resolving complex, nonlinear class boundaries.

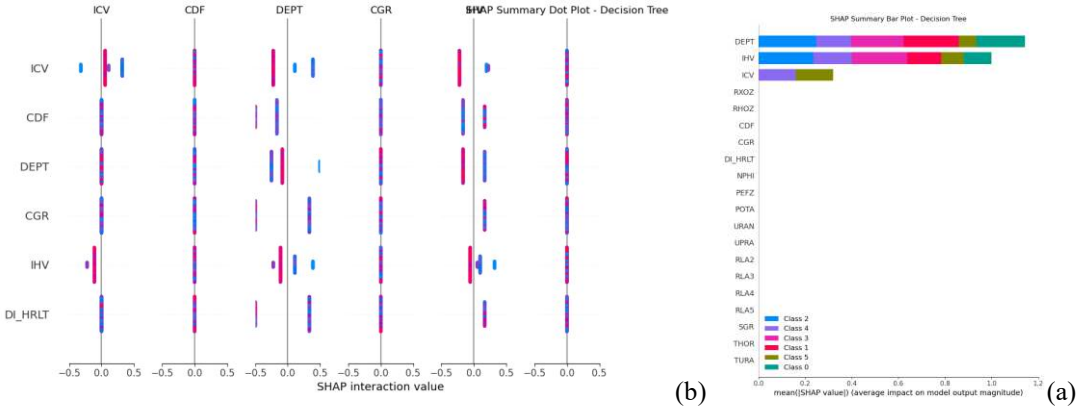



**Figure 6.** SHAP results for the Decision Tree model: (a) feature-importance bar plot; (b) feature-interaction summary plot for permeability classification.

The Decision Tree model presents yet another distinct behavior. Its summary plot identifies GR and CGR as the two variables exerting the strongest directional influence: high GR values are predominantly associated with negative SHAP contributions (a reduced probability of high permeability), whereas low GR values correspond to positive contributions. Because of its hierarchical, greedy splitting procedure, the tree tends to base its earliest partitions on one or two dominant features, resulting in a pronounced dependence on the gamma-ray response; consequently, its predictive performance may deteriorate in intervals where other variables—such as NPHI or RHOZ—carry more discriminative information than gamma ray alone.

This dependence is mirrored in the corresponding importance ranking, in which GR and CGR again occupy the top positions, indicating that the tree performs its initial split predominantly along the gamma-ray log before applying the remaining features as secondary filters. From a practical standpoint, while this splitting behavior is geophysically intuitive, it may come at the cost of discarding finer porosity- and resistivity-related information within transitional (intermediate) permeability intervals.

**Overall assessment of the classification models**

For this well, the combined evidence from model performance and SHAP interpretation points to a consistent set of conclusions:

- The tree-based algorithm exhibits the strongest overall capability, owing to a more balanced distribution of feature importance across gamma-ray, porosity, density, and resistivity variables, which allows the model to draw simultaneously on lithological and saturation-related information.
- SVM achieves robust nonlinear boundary separation, treating electrical resistivity—particularly in combination with porosity—as a key indicator of the presence of effective flow pathways.
- KNN performs well in densely sampled intervals and relies predominantly on local porosity and density characteristics (NPHI, RHOZ).

- Logistic Regression correctly identifies the direction and relative importance of the dominant features but, owing to its linear structure, remains comparatively limited in resolving overlapping class boundaries.

**2) Regression models**

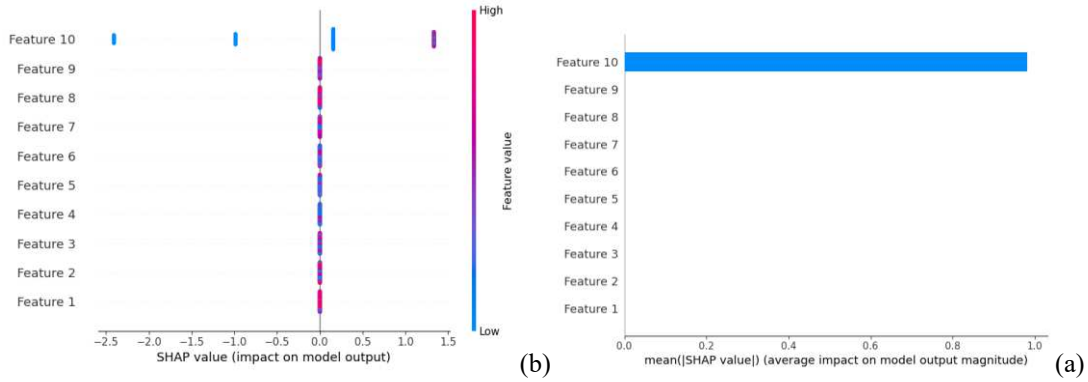



**Figure 7.** SHAP results for the Decision Tree Regressor: (a) feature-importance bar plot; (b) feature-interaction summary plot for continuous permeability prediction.

The pronounced contribution of GR and RHOB (bulk density) in this model indicates that these two variables provide the greatest discriminative power for continuous permeability, and that the Decision Tree Regressor exhibits a strong dependence on them (panel a). The corresponding importance ranking (panel b) shows that most of the observed permeability variation is associated with decreasing GR alongside increasing NPHI—a pattern consistent with the clean, high-porosity sand intervals present in this well.

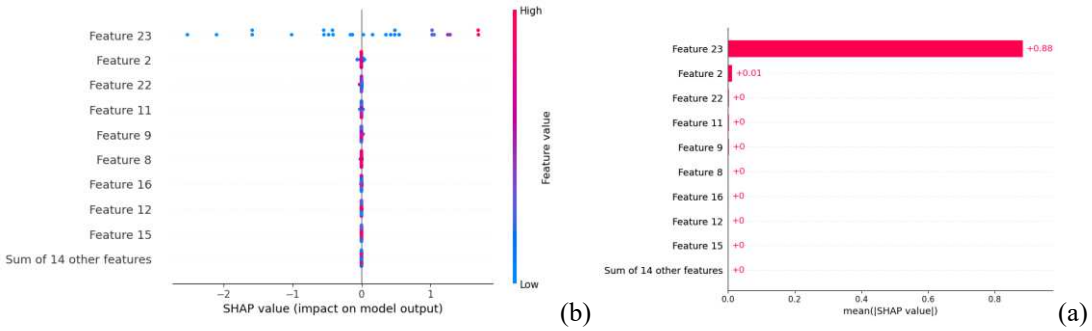



**Figure 8.** SHAP results for the Ensemble Regressor (Voting): (a) feature-importance bar plot; (b) feature-interaction summary plot for continuous permeability prediction.

For the ensemble model, the relatively balanced distribution of feature importance (panel a) suggests that combining multiple base learners reduces the influence of noise and dependence on any single dominant variable, yielding a more integrated representation of the overall feature contributions. Consistent with this, the importance plot (panel b) indicates that by aggregating several predictors, the ensemble reconstructs the underlying data trends more reliably than any individual model; NPHI and RHOB emerge as the

strongest positive contributors to permeability, while GR retains its expected negative influence.

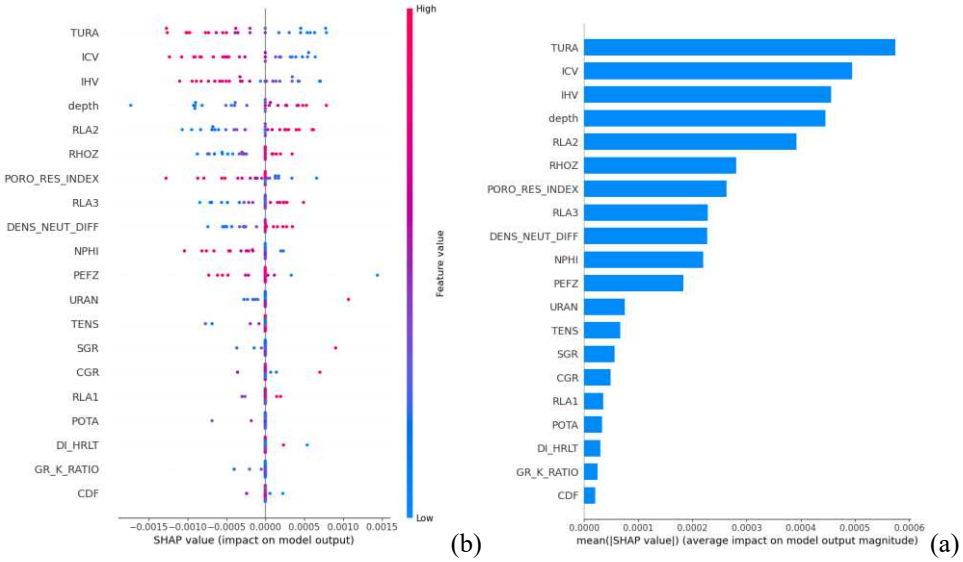



**Figure 9.** SHAP results for the KNN Regressor: (a) feature-importance bar plot; (b) feature-interaction summary plot for continuous permeability prediction.

For the KNN Regressor, RHOB ranks as the most influential variable, followed by GR and NPHI (panel a), indicating that the model's predictions are governed primarily by the local neighborhood structure of the data. The corresponding importance plot (panel b) further demonstrates a pronounced sensitivity to local data values: RHOB and NPHI dominate the reconstruction of permeability, and even small variations in these two variables produce substantial changes in the predicted output.

**Overall assessment of the regression models**

The KNN and Decision Tree models performed adequately in reconstructing relatively simple permeability trends but exhibited reduced accuracy near class boundaries and within heterogeneous intervals of the reservoir. Across all three regression models, GR, NPHI, and RHOB consistently emerged as the dominant controlling variables. This consistency with established geological reasoning confirms that clean sandstone intervals, characterized by high porosity and low shale content, play the primary role in determining reservoir quality.

4.1.3. Cross-Validation

**Classification section**

At the outset of this study, the permeability data were partitioned into six discrete classes, as follows:

- Class 1 (moderate): permeability ≈ 41.67 mD
- Class 2 (very low): permeability ≈ 18.52 mD
- Class 3 (low–moderate): permeability ≈ 33.33 mD
- Class 4 (high): permeability ≈ 62.5 mD
- Class 5 (very high): permeability ≈ 106.06 mD
- Class 6 (low): permeability ≈ 21.43 mD

These classes represent reservoir flow quality, and the models were required to correctly resolve the boundaries between them. Cross-validation was employed in this study to mitigate the effects of noise, outlier observations, and class imbalance, thereby providing a more realistic assessment of model performance. In effect, this procedure enables an evaluation of model stability while guarding against overfitting.

Given the intrinsic sensitivity of permeability to petrophysical parameters such as electrical resistivity (the RLA series and RXOZ), shale indicators (GR, CGR, SGR), and porosity (NPHI), the use of cross-validation allowed the performance of each algorithm to be assessed comprehensively across all permeability classes. This, in turn, provides petroleum engineers and geoscientists with a clearer understanding of the generalization capability of these models when applied to analogous reservoirs.

**Regression section**

In the regression component of this study, the primary objective was to predict continuous permeability values at fine resolution, without resorting to class discretization. Unlike the classification models—which partition permeability into a fixed number of categories (very low, low, moderate, high)—regression models allow a more detailed characterization of permeability variation along the wellbore. This capability is of particular importance in reservoirs exhibiting pronounced small-scale permeability heterogeneity, and it can play a decisive role in optimizing well-completion design and production planning.

Several models were applied for permeability estimation. In addition to reproducing the overall trend of permeability variation, these models were required to capture localized peaks and declines arising from abrupt changes in rock and fluid properties.

One of the principal challenges in this section stemmed from the continuous, unbounded nature of the permeability values, in contrast to the discrete class boundaries used in the classification task. Accordingly, performance metrics such as RMSE and MAE were adopted to evaluate model accuracy. This regression-based analysis provides deeper insight into reservoir behavior, enabling petroleum engineers to identify genuine, continuous variations in reservoir quality—free from the constraints of discrete permeability classes—for use in reservoir modeling and development decision-making.

**1) Classification models**

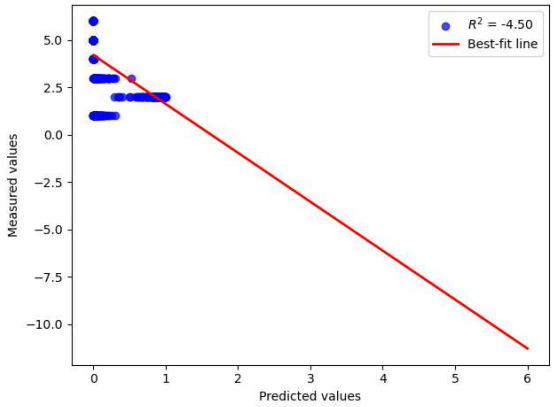



**Figure 10.** Comparison of actual and predicted permeability classes using the Logistic Regression model.

This model successfully distinguished the very-high-permeability class (106.06 mD) from the very-low-permeability class (41.67 mD) with reasonable accuracy; however, notable overlap was observed among the intermediate classes (21.43 and 33.33 mD). This limitation arises from the linear nature of the Logistic Regression decision boundary, which constrains its ability to capture the complex nonlinear relationships between petrophysical features and permeability.

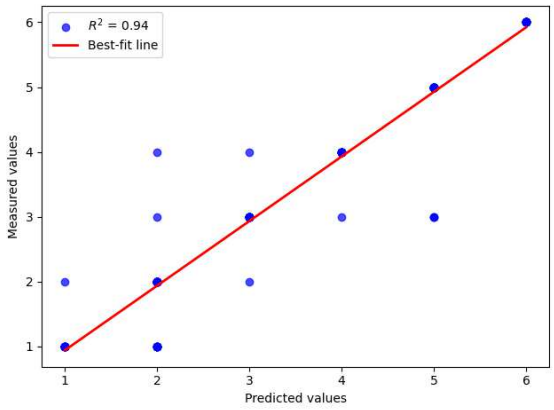



**Figure 11.** Comparison of actual and predicted permeability classes using the KNN Classifier.

The KNN Classifier performed comparatively well in identifying the intermediate classes (21.43 and 33.33 mD) but exhibited weaker performance for classes with limited sample representation (e.g., 18.52 mD). This behavior can be attributed to the low sample density within certain classes and to the sensitivity of the KNN algorithm to the number of neighboring samples. From a reservoir-engineering perspective, this weakness suggests that lower-permeability intervals may not be adequately represented, potentially resulting in an overly optimistic assessment of non-productive-zone quality.

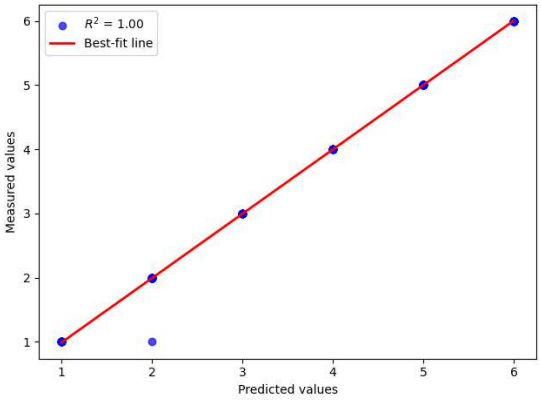



**Figure 12.** Comparison of actual and predicted permeability classes using the Support Vector Machine model.

The SVM model, employing an RBF kernel, successfully captured the complex boundaries separating permeability classes. The low- and high-permeability classes were identified with minimal error, and only limited misclassification occurred near the 33.33–41.67 mD boundary. This performance reflects the strong capability of SVM in extracting the nonlinear relationships governing petrophysical–permeability behavior.

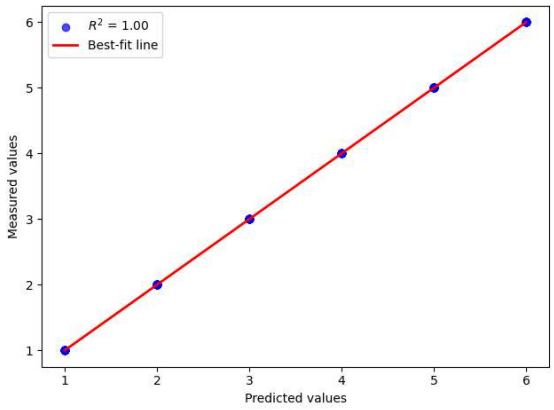



**Figure 13.** Comparison of actual and predicted permeability classes using the Decision Tree model.

The Decision Tree model achieved acceptable accuracy at the extreme ends of the permeability distribution (very low and very high) but produced oversimplified decision boundaries—and consequently greater error—within the intermediate classes (33.33 and 41.67 mD). This behavior reflects the inherent tendency of decision trees toward discrete, non-continuous partitioning, a characteristic that limits their effectiveness when applied to heterogeneous reservoir data.

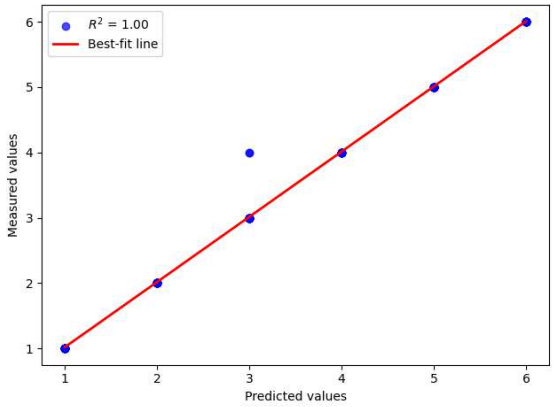



**Figure 14.** Comparison of actual and predicted permeability classes using the Ensemble (Voting) model.

The Voting ensemble produced more balanced results overall. By combining several base classifiers, the model achieved higher overall accuracy and identified the intermediate classes more reliably than any individual algorithm. This outcome is also consistent with the underlying geology, as permeability within this well is frequently governed by the combined influence of several petrophysical variables—resistivity, gamma ray, and porosity—rather than any single dominant feature.

### 2) Regression models

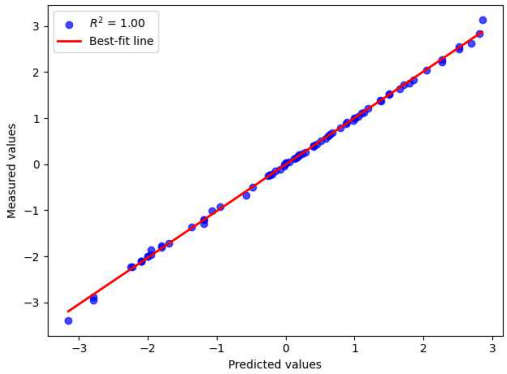



**Figure 15.** Comparison of actual and predicted permeability values using the Decision Tree Regressor.

The Decision Tree Regressor captured the overall trend of permeability variation; however, at many depth points the predicted values exhibited a stepwise, staircase-like pattern. This behavior stems from the partition-based structure of decision trees, which inherently limits their capacity to reconstruct continuous permeability variation. Given that this well exhibits both gradual and abrupt permeability changes, this limitation resulted in oversimplification within certain intervals.

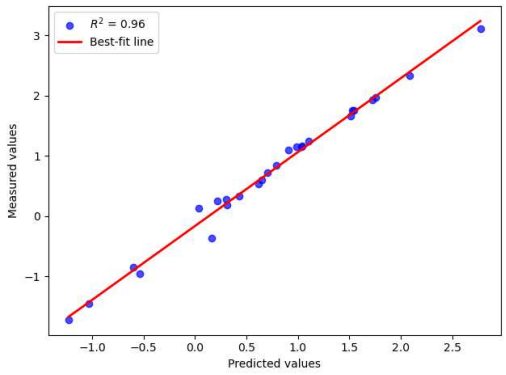



**Figure 16.** Comparison of actual and predicted permeability values using the Ensemble Regressor (Voting).

By integrating the outputs of several base learners—including KNN, Decision Tree, and SVR—the ensemble regressor produced comparatively balanced predictions. Although its accuracy did not match that of the standalone Decision Tree Regressor, the errors associated with the individual base models were substantially reduced. From a petrophysical standpoint, this model preserved the overall permeability trend while avoiding extreme, outlier-driven predictions.

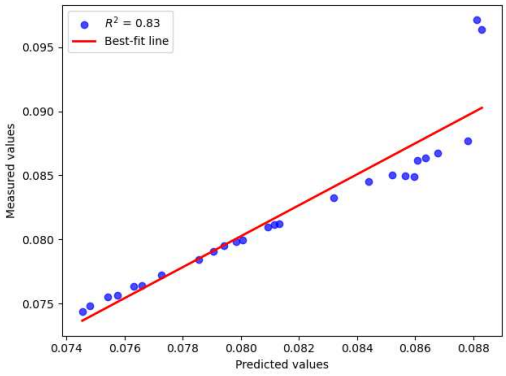



**Figure 17.** Comparison of actual and predicted permeability values using the KNN Regressor.

The KNN Regressor reconstructed permeability values reasonably well within densely sampled intervals but exhibited pronounced fluctuations in more sparsely sampled zones—a behavior attributable to the strong sensitivity of KNN to the distance and density of neighboring observations. From a reservoir-engineering perspective, in intervals where lithology changes abruptly, the KNN model's reliance on the nearest available samples produced less physically realistic permeability estimates than those obtained from the tree-based model.

## 5. Conclusions

This study analyzed petrophysical data from a producing well within an Iranian gas field with the objective of predicting reservoir permeability through machine-learning

methods. The available data were organized into two categories—DST–log data, used for classification modeling, and core–log data, used for regression modeling—and, following petrophysical filtering and feature-selection procedures, a range of machine-learning algorithms was developed and evaluated on each.

The results indicated that, despite variation in lithology and reservoir conditions, a broadly consistent pattern emerged with respect to both feature importance and relative model performance. Gamma-ray logs (SGR, CGR) and the electrical resistivity logs (RLA1–RLA5, RXOZ, RT_HRLT) consistently exerted the dominant influence on permeability prediction. This finding is consistent with established petrophysical reasoning: elevated gamma-ray response, indicative of shale content, is generally associated with reduced permeability, whereas high resistivity typically reflects hydrocarbon presence and favorable reservoir quality.

A central objective of this work was to determine whether combining multiple base learners within an Ensemble Learning framework yields a genuine improvement over any individual algorithm, rather than assuming this to be the case a priori. The comparative results, summarized across the classification and regression tasks, support a nuanced rather than unconditional answer to this question.

In the classification task, the Ensemble model (combining Logistic Regression, KNN, SVM, and Decision Tree through Voting/Stacking) achieved accuracy values statistically comparable to those of the best-performing standalone algorithms—specifically, SVM and the Decision Tree Classifier, both of which independently reached near-perfect accuracy (approaching 1.0) on the testing data. The Ensemble model did not exceed these two algorithms in raw accuracy; rather, its principal advantage lay in reducing the weaknesses associated with the weaker base learners, most notably Logistic Regression, whose linear decision boundary limited its performance in lithologically overlapping intervals. In this sense, ensembling functioned less as a means of pushing peak accuracy beyond what the strongest individual classifier could achieve, and more as a mechanism for producing balanced, class-wide reliability—an outcome of particular value in the intermediate permeability classes, where individual algorithms displayed their greatest disagreement.

In the regression task, the comparison yields a more qualified conclusion. The standalone Decision Tree Regressor achieved the lowest prediction error (RMSE approaching zero) on several performance tables, at times outperforming the Ensemble Regressor in raw accuracy; this outcome reflects the tree's capacity to fit the training data closely, though at an increased risk of overfitting and reduced generalization to unseen intervals. The Ensemble Regressor, by contrast, exhibited somewhat higher error on the reported metrics but produced smoother, more physically plausible permeability profiles, avoided the pronounced stepwise artifacts characteristic of the standalone Decision Tree, and moderated the erratic local fluctuations observed in the standalone KNN Regressor. Consequently, whether the hybrid model constitutes an improvement depends on the criterion applied: if the objective is to minimize point-wise prediction error on the available data, standalone tree-based models remained competitive or superior; if the objective is robustness, stability, and resistance to overfitting across heterogeneous reservoir intervals, the Ensemble framework provided the more dependable outcome.

Taken together, these findings indicate that Ensemble Learning does not uniformly outperform every standalone algorithm on every metric; rather, its principal contribution lies in improved generalization capability, reduced sensitivity to noise, and more consistent performance across permeability ranges—particularly in intervals where individual base learners disagree or where petrophysical responses are ambiguous. Feature-importance analysis further confirmed that, although the relative contribution of individual variables varied with lithology and reservoir conditions, gamma-ray and resistivity logs remained the most consistently influential parameters across all algorithms, while SHAP analysis additionally identified NPHI and RHOB as important secondary contributors, particularly within clean sandstone intervals exhibiting favorable pore connectivity.

Overall, the findings of this study confirm the practical value of machine-learning approaches for permeability prediction and demonstrate that integrating static petrophysical logs with dynamic well-test and core data can meaningfully improve reservoir characterization. The results further suggest that hybrid Ensemble Learning techniques offer the more reliable framework specifically in heterogeneous reservoirs where multiple petrophysical mechanisms jointly govern fluid-flow behavior—whereas in more homogeneous or cleanly separable intervals, standalone tree-based models may remain equally, or occasionally more, effective. These outcomes carry direct relevance for reservoir management, field-development planning, and the optimization of hydrocarbon-production strategies in Iranian reservoirs.

## 6. Discussion

The results obtained from the well investigated in this study indicate that the performance of machine-learning algorithms in permeability prediction is governed not solely by the mathematical sophistication of the models themselves, but substantially by the quality, diversity, and geological representativeness of the input petrophysical data. Consequently, the elevated accuracy observed for certain models—and the comparatively weaker performance of others—should be interpreted primarily as a reflection of the geological characteristics and data structure of the reservoir, rather than as evidence of the absolute superiority or deficiency of any single algorithm.

In intervals where electrical resistivity and gamma-ray logs exhibited pronounced variability and clear lithological contrast, both the standalone tree-based models and the Ensemble framework achieved near error-free predictions. This outcome is consistent with established reservoir-engineering principles, given that resistivity and gamma-ray responses are directly linked to fluid saturation, shale content, and lithological variation, thereby providing strong, learnable patterns. Conversely, in intervals characterized by more homogeneous or noisier petrophysical responses, the predictive capability of all models—including the Ensemble—declined, reflecting the reduced separability of the underlying permeability-related features rather than any specific algorithmic weakness.

A further noteworthy observation concerns the performance of comparatively simple algorithms. Logistic Regression, despite its restrictive linear decision boundary, remained capable of correctly identifying the very-low- and very-high-permeability classes. This finding suggests that even linear models can achieve acceptable performance when the

underlying petrophysical data are sufficiently clean and well separated; the same model, however, consistently struggled within transitional permeability zones, where nonlinear interactions among lithology, porosity, and fluid saturation became dominant—precisely the regime in which the more flexible standalone algorithms (SVM, Decision Tree) and the Ensemble framework demonstrated their comparative advantage.

From a geological and petrophysical standpoint, the integration of DST–log and core–log datasets provided a more complete representation of reservoir quality than either data source considered independently. Where core measurements were available, the models—both standalone and ensemble—established more accurate relationships between porosity, density, and permeability; in contrast, in intervals relying primarily on DST data, electrical resistivity and gamma-ray responses exerted a comparatively greater influence, owing to their stronger association with fluid-flow behavior and reservoir saturation state.

Returning to the central comparison of this study—standalone algorithms versus the hybrid Ensemble framework—the evidence indicates that the Ensemble models generally achieved superior stability and generalization performance, though not a uniform improvement in raw predictive accuracy across every model comparison. By combining the complementary learning behaviors of the base algorithms (the local sensitivity of KNN, the hierarchical partitioning of the Decision Tree, and the nonlinear boundary separation of SVM), the Ensemble framework improved the overall balance between bias and variance and reduced sensitivity to noisy or locally anomalous observations. At the same time, the results showed that this improvement was not absolute: in several localized intervals, the standalone Decision Tree or SVM models produced results equal to, or marginally better than, the corresponding Ensemble model, owing to their greater capacity to adapt directly to specific nonlinear patterns present within a given depth interval.

This qualification is important, as it indicates that the practical benefit of Ensemble Learning is conditional rather than guaranteed, and depends strongly on reservoir heterogeneity, the distribution of the underlying features, and the overall quality of the available petrophysical measurements. In reservoirs exhibiting substantial heterogeneity—where different depth intervals are governed by different combinations of controlling parameters—the Ensemble framework offers a genuine advantage by remaining robust across this variability. In more homogeneous or cleanly separable reservoir sections, however, a well-tuned standalone algorithm may achieve comparable, or occasionally superior, accuracy with lower computational cost and greater interpretability.

Overall, the findings of this study support the broader consensus emerging in the recent literature that no single machine-learning algorithm—whether standalone or hybrid—can be regarded as universally optimal across all reservoir conditions. Model selection should instead be guided by the geological characteristics, lithological complexity, and data quality specific to each reservoir. Nevertheless, on the basis of the comparative evidence obtained in this study, Ensemble Learning, together with SVM and tree-based algorithms, can be regarded as the most reliable and practically dependable approach for permeability prediction in reservoirs sharing geological characteristics similar to those of the field

investigated here—particularly where robustness and generalization across heterogeneous conditions are prioritized over marginal gains in point-wise predictive accuracy.

## Nomenclature

### Latin Letters

| Symbol | Description | Unit |
|---|---|---|
| $b$ | Bias (intercept) term in the SVR/SVM function | dimensionless |
| $C$ | Empirical constant related to pore geometry (Kozeny–Carman equation) | dimensionless |
| $C$ | Regularization / error-penalty parameter (Logistic Regression, SVM) | dimensionless |
| $f(X)$ | Predicted output function of the support vector regression model | dimensionless |
| $F$ | Initial (baseline) model estimate in gradient boosting | dimensionless |
| $g(x_i)$ | Negative gradient of the loss function at observation $i$ | dimensionless |
| $K(x_i, x)$ | Kernel function | dimensionless |
| $k$ | Permeability | $m^2$ |
| $L(y_i, F(x_i))$ | Loss function relating observed and predicted response values | dimensionless |
| $S$ | Specific surface area of rock grains | $m^{-1}$ |
| $w$ | Weight vector (coefficients) in the SVR/SVM decision function | dimensionless |
| $X$ | Input feature vector | — |
| $x_i$ | $i$-th input observation (feature vector) | — |
| $y_i / y_i$ | Observed (measured) target value for observation $i$ | mD (permeability) |

### Greek Letters

| Symbol | Description | Unit |
|---|---|---|
| $\alpha, \alpha_i'$ | Lagrange multipliers in the dual formulation of SVR | dimensionless |
| $\varepsilon$ | Error tolerance (insensitive-loss margin) in support vector regression | dimensionless |
| $\xi_i, \xi_i^*$ | Slack variables permitting deviations exceeding ε | dimensionless |
| $\sigma$ | Bandwidth (spread) parameter of the Gaussian (RBF) kernel | dimensionless |
| $\varphi$ | Porosity | fraction |

**Funding: No**

## Data Availability:

The data surveys that support the findings of this study are available from Pars Oil & Gas Company (POGC), but restrictions apply to the availability of these data, which were used under license for the current study and so are not publicly available. The data are, however, available upon request and with the permission of POGC.

## References


[1] Tiab, D., & Donaldson, E. C. (2015). *Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties* (4th ed.). Gulf Professional Publishing. https://doi.org/10.1016/C2013-0-19103-8

[2] Kozeny, J. (1927). Über kapillare Leitung des Wassers im Boden. *Sitzungsber Akad. Wiss. Wien*, 136, 271–306.

[3] Carman, P. C. (1937). Fluid flow through granular beds. *Transactions of the Institution of Chemical Engineers*, 15, 150–166.

[4] Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., & Keelan, D. K. (1993). Enhanced reservoir description: Using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. *SPE Annual Technical Conference and Exhibition*.

[5] Alimoradi, A., Hajkarimian, H., Hemati Ahooi, H,.and Salsabili, M., (2022), "Comparison between the Performance of Four Metaheuristic Algorithms in Training a Multilayer Perceptron Machine for Gold Grade Estimation," International Journal of Mining and Geo-Engineering, 56, 2, pp. 97–105.

[6] Mohaghegh, S., (2000), "Virtual-Intelligence Applications in Petroleum Engineering: Part 1 Artificial Neural Networks," Journal of Petroleum Technology, 52, 09, pp. 64–73.

[7] Mohaghegh, S. D., (2000), "Virtual Intelligence and Its Applications in Petroleum Engineering; Part 3. Fuzzy Logic, Distinguished Author Series," Journal of Petroleum Technology (JPT).

[8] Huang, Z., and Williamson, M. A,(1997) ,."Determination of Porosity and Permeability in Reservoir Intervals by Artificial Neural Network Modelling, Offshore Eastern Canada," Petroleum Geoscience, 3, 3, pp. 245–258.

[9] Hampson, D. P., Schuelke, J. S., and Quirein, J. A" ,(2001) ,.Use of Multiattribute Transforms to pp. 220–236.

[10] Russell, B. H., (2004), "The Application of Multivariate Statistics and Neural Networks to the Prediction of Reservoir Parameters Using Seismic Attributes", PhD Thesis, Department of Geology and Geophysics, Calgary, Alberta.

[11] Nooruddin, H. A., Anifowose, F., and Abdulraheem, A., (2013), "Applying Artificial Intelligence Techniques to Develop Permeability Predictive Models Using Mercury Injection CapillaryPressure Data," SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, SPE, p. SPE168109.

[12] Al-Anazi, A. F., and Gates, I. D., (2012), "Support Vector Regression to Predict Porosity and Permeability: Effect of Sample Size," Comput Geosci, 39, pp. 64–76.

[13] Singh, S., Kanli, A. I., and Sevgen, S., (2016), "A General Approach for Porosity Estimation Using Artificial Neural Network Method: A Case Study from Kansas Gas Field," Studia Geophysica et Geodaetica ,60 ,pp. 130–140.

[14] Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K" ,(2004) ,.Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), Ieee, pp. 985–990.

[15] Saffarzadeh, S., and Shadizadeh, S. R" ,(2012) ,.Reservoir Rock Permeability Prediction Using Support Vector Regression in an Iranian Oil Field", Journal of Geophysics and Engineering, 9, 3, pp. 336-334.

[16] Olatunji, S. O., Selamat, A., and Raheem, A. A. A" ,(2013) ,.Extreme Learning Machines Based Model for Predicting Permeability of Carbonate Reservoir", International Journal of Digital Content Technology and its Applications, 7, 1, p. 450.

[17] Xiao, D., Li, B., and Mao, Y., (2017), "A Multiple Hidden Layers Extreme Learning Machine Method and Its Application," Math Probl Eng, 2017.

[18] Adeniran, A. A., Adebayo, A. R., Salami, H. O,. Yahaya, M. O., and Abdulraheem, A., (2019), "A Competitive Ensemble Model for Permeability Prediction in Heterogeneous Oil and Gas Reservoirs", Applied Computing and Geosciences, 1, p. 100004.

[19] Zhang, G., Wang, Z., Mohaghegh, S., Lin, C., Sun,Y., and Pei, S., (2021), "Pattern Visualization and Understanding of Machine Learning Models for Permeability Prediction in Tight Sandstone Reservoirs," J Pet Sci Eng, 200, p. 108142.

[20] Singh, S., Kanli, A. I., and Sevgen, S., (2016), "A General Approach for Porosity Estimation Using

Artificial Neural Network Method: A Case Study from Kansas Gas Field," *Studia Geophysica et Geodaetica, 60*, pp. 130–140.

[21] Vemuri, V. R., and Rogers, R. D., (1994), "Artificial Neural Networks: Forecasting Time

[22] Quinlan JR (1986) Induction of decision trees. Machine Learning 1(1):81–106Series", Los Alamitos.

[23] Breiman L (2017) Classification and regression trees. Routledge.

[24] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E, Louppe G (2012) Scikit-learn: Machine learning in python 12

[25] Friedman J (2000) Greedy function approximation: A gradient boosting machine 29.

[26] 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, pp 785–794

[27] Cortes C, Vapnik V (1995) Support vector networks 20:273–297

[28] Hsieh CJ, Chang KW, Lin CJ, Keerthi SS, Sundararajan S (2008) A dual coordinate descent method for large-scale linear svm. In: Proceedings of the 25th international conference on Machine learning, ACM, pp 408–415

[29] Crammer K, Singer Y (2001) On the algorithmic implementation of multiclass kernel-based vector machines. Journal of machine learning research 2(Dec):265–292

Specht DF (1990) Probabilistic neural networks. Neural Network 3(1):109118 [69]

[30] Mao KZ, Tan KC, Ser W (2000) Probabilistic neural-network structure determination for pattern classification. IEEE Trans Neural Network 11:10091016

[31] El Emary IMM, Srinivasan R (2008) On the application of various probabilistic neural networks in solving different pattern classification problems. World Appl Sci J 4(6):772–780

[32] Chasset (2015) Pierre-Olivier probabilistic neural-network. https//: cran.r-project.org/web/packages/pnn/pnn.pdf

[33] Weston S, Calaway R (2015) Getting Started with doParallel and foreach. https://cran.r-project.org/web/packages/doParallel/vignettes/gettingstartedParallel.pdf

[34] Freund Y, Schapire RE (1997) A decision-theoretic generalization of online learning and an application to boosting. J Comput Syst Sci55(1):119–139.

[35] Natekin A, Knoll A (2013) Gradient boosting machines, a tutorial. Front Neurorobot 7:1–21.

[36] Ridgeway G (2017) Generalized boosted regression models. https//: cran.r-project.org/web/packages/gbm/gbm.pdf.

[37] Mavko, G., Mukerji, T., & Dvorkin, J. (2020). *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media* (3rd ed.). Cambridge University Press.

[38] Zoback, M. D. (2007). *Reservoir Geomechanics*. Cambridge University Press.

[39] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression. Wiley.

[40] Tariq, Z., et al. (2021). Feature selection for machine learning models for predicting rock properties. *Journal of Petroleum Science and Engineering*, 205, 108921.

[41] Zhang, H., Niu, W., & Liu, Y. (2020). *Application of machine learning algorithms in reservoir characterization: A review*. *Journal of Petroleum Science and Engineering*, 186, 106682. https://doi.org/10.1016/j.petrol.2019.106682

[42] Al-Fatlawi, O., Ahmed, M., & Abdulraheem, A. (2022). Machine learning-based permeability prediction using well log data: Application of artificial neural networks and ensemble methods. *Journal of Petroleum Science and Engineering*, 210, 110066.

[43] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging, 9*(4), 611–629.

[44] Bergstra, J., & Bengio, Y. (2012). *Random search for hyper-parameter optimization*. *Journal of Machine Learning Research (JMLR)*, 13(10), 281–305.

[45] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE.

[46] Dietterich, T. G. (2000). *Ensemble methods in machine learning*. In *Multiple classifier systems* (pp. 1–15). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45014-9_1

[47] Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.

[48] Altman, N. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician.

[49] Lundberg, S. M., & Lee, S. I. (2017). *A Unified Approach to Interpreting Model Predictions*. In Advances in Neural Information Processing Systems (NeurIPS 2017) (pp. 4765–4774). Curran Associates, Inc.