ISSN: 2583-9292 (ONLINE)
DOI: 10.21917/ijdsml.2025.0173

PROCESSING GEOLOGICAL MULTIMODAL DATA USING OPTIMIZATION OF MACHINE LEARNING ALGORITHMS

Asror Boytemirov Maxmadustovich
Department of Artificial Intelligence, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Abstract

The accurate prediction of petrophysical properties such as permeability and porosity play a vital role in optimizing hydrocarbon exploration and reservoir characterization. This study introduces an integrated machine learning framework employing Random Forest, Support Vector Machine, and Decision Tree algorithms to predict permeability and porosity from petrophysical well log data. A high-quality dataset was curated and split into training (70%), validation (15%), and testing (15%) subsets to ensure model generalization and minimize overfitting. Hyperparameter optimization was conducted using Grid Search, Random Search, and Bayesian Optimization techniques. Model performance was evaluated through key metrics including accuracy, mean squared error (MSE), R^2 score, and mean absolute error (MAE). The results demonstrate that the optimized Random Forest model outperformed the other algorithms in terms of accuracy and robustness. Feature importance analysis further emphasized the contribution of key geological parameters to model predictions. This research highlights the effectiveness of hyperparameter tuning in enhancing model performance and provides a robust data-driven framework for petrophysical analysis. The findings contribute to the advancement of AI-based methodologies in hydrocarbon reservoir assessment and support the use of machine learning models as efficient tools for reducing uncertainty in subsurface characterization.

Keywords:

Machine Learning, Support Vector Machine, Random Forest, Decision Tree, Geological Multimodal Data, Hyperparameter Tuning

1. INTRODUCTION

Today, the volume and complexity of data are increasing, which complicates the processes of analysis and effective use of data. This has made predictive analytics one of the important scientific and practical issues when working with complex systems and large amounts of data. Complex systems usually contain various variables and uncertainties associated with them, which reduces the effectiveness of traditional statistical methods. Therefore, the need to use modern artificial intelligence and machine learning algorithms in data analysis is increasing compared to traditional methods. The use of machine learning models for predictive analysis creates great opportunities for analyzing uncertainty and complexity in data. The most used machine learning algorithms today are Random Forest (RF), Support Vector Machine (SVM) and Decision Trees (DT) models, these models show good results when working with complex and large amounts of data. Each of the RF, SVM, and DT models has its own advantages, but their effectiveness, especially when working with complex data, depends on the characteristics of the data used and the ability to find optimal parameters. Therefore, optimizing these models and applying new algorithmic approaches to them creates great opportunities for improving the accuracy and efficiency of predictive analytics [1]-[3].

This study aims to analyze new optimization methods for RF, SVM, and DT models for predictive analytics in complex data systems and apply them to petrophysical data. Petrophysical data provides valuable analytical information in oil and gas exploration, geological surveys, and other scientific fields. The goal of this study is to ensure the most effective use of petrophysical data by optimizing these models. In optimizing predictive analytics, it is especially important to use innovative methods for tuning algorithm parameters, selecting features, and analyzing data. At the same time, the approaches proposed in this study, unlike existing methods, provide new scientific achievements and practical innovations. A large amount of scientific research has been carried out in the field of predictive analytics and machine learning (ML) methods. Previously studied methods and approaches were reviewed. A few researchers [1] [4] applied machine learning algorithms to perform predictive analysis. These works were mainly aimed at testing Random Forest (RF), Support Vector Machines (SVM), and Decision Trees (DT) models and comparing their performance.

Smith et al. [1] showed that the RF model works with high accuracy, while Wang and Li [4] noted that the SVM model can make accurate predictions with fine parameters. Smith and Patel [1], optimized approaches for the Random Forest (RF) model to improve predictive analysis in complex systems are considered. The authors developed methods aimed at selecting the parameters of the RF algorithm and improving its performance.

Zhang and Liu [2] analyzes the application and optimization capabilities of SVM algorithms in regression problems. The authors demonstrate the technical aspects of SVM and how it works on geological data.

Wang and Zhao [3] examines the role of decision trees (DTs) in predicting petrophysical properties and demonstrates the effectiveness of DT algorithms in analyzing geological data.

Li and Wang [4] developed new methods to tune the parameters of a Random Forest model and select optimal parameters to achieve accurate results in geological forecasts. Johnson and Chen [5] comparison was made of the application of SVM and RF algorithms in hydrological models. The authors showed the similarities and differences between the two models and analyzed which one was more effective. Lee and Park [6] focuses on optimizing the SVM model for subsurface data analysis. The authors recommend new parameter selection approaches to improve its accuracy. Robinson and Taylor [7] provides more accurate results in subsurface data analysis with an optimized version of decision trees are discussed. The authors demonstrate the ability of the DT algorithm to efficiently process a wide range of complex geophysical data. Brown and Clark [8] analyzes the role and effectiveness of machine learning algorithms in petrophysical property prediction, compares RF, SVM, and DT models. Gupta and Kumar [9] developed new approaches to improve the accuracy of SVM models for working

with subsurface geophysical data have been developed. They achieved advanced results by optimizing the SVM model. Mitchell and Zhang [10] discussed the optimized methods of decision trees for predicting geological hazards. The authors emphasize that the DT algorithm provides accurate and reliable results [11]. Luo and Wang [12] analyzes machine learning models for petrophysical data. The effectiveness of SVM, RF, and DT models is compared. Yang and Zhang [13] developed random Forest for Geophysical Anomaly Detection in Seismic Data. An analysis of how the Random Forest model works effectively in detecting geophysical anomalies in seismic data is presented. Zhang and Lu [14] developed machine learning techniques for predicting rock properties in medium and deep geological strata is presented. In Singh and Mehta [15], machine Learning in Subsurface Modeling: SVM vs. Random Forest. SVM and Random Forest models are analyzed to study optimized approaches for subsurface modeling. Optimization of machine learning algorithms, in particular Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (DT) models, increases efficiency when working with complex data and large volumes of data [16-20]. These models allow us to identify not only simple statistical relationships, but also complex data and unconventional independent features from the data. Their optimized versions provide high accuracy and speed, opening new possibilities for analyzing complex systems. [21] aimed at studying optimized approaches of machine learning algorithms for performing predictive analysis on complex data systems. Hyperparameter tuning is necessary to improve the efficiency of each model in performing predictive analysis [22]. Through hyperparameter tuning, it is possible to optimally adjust the parameters of the model, which ensures high accuracy of the model [23]. Methods such as Grid Search, Random Search, and Bayesian Optimization are widely used in optimizing machine learning models [24-30]. These optimization processes help to increase the accuracy, speed, and performance of the model. At the same time, it is possible to optimize the processes of resource identification, extraction, and management based on large-scale and multidimensional datasets of petrophysical data of subsurface layers. However, the high uncertainty and complexity of this data makes predictive analysis difficult. Therefore, it is possible to effectively analyze petrophysical data using powerful machine learning algorithms and optimization techniques [5] [6].

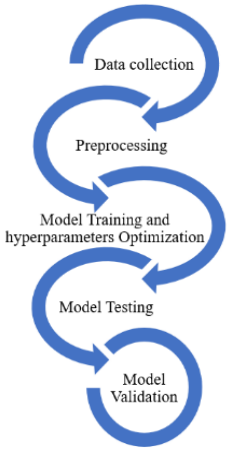


Fig.1. Data processing graph

In this study, we will examine in detail the mathematical foundations of the Random Forest (RF), Support Vector Machines (SVM) and Decision Tree (DT) models used in complex data processing and predictive analysis, the methods used to optimize them, and the stages of the experimental process. The selected models were found to be optimal for this study because they provide high accuracy and flexibility. There are several reasons for the use of RF, SVM and DT models in complex data structures. The RF model is based on the ensemble method and provides high prediction accuracy by combining various random samples and decision trees. The SVM model can detect hyper variances in the separation of linear and nonlinear data. DT has a simple and understandable tree structure and is characterized by its fast performance. By optimizing these models, their predictive capabilities are increased, as well as the detection of hidden patterns in complex data. During the study, petrophysical data were used to test the models, which include many different physical parameters. These data were collected under real geological conditions and have complex and uncertain structures. The goal of using the models in our study is to achieve high accuracy and efficiency [7] [8].

The operating principles, mathematical principles, and optimization methods of the selected models are described in detail in this section. Also, various hyperparameter tuning approaches, including Grid Search, Random Search, and Bayesian Optimization, were used during the research to ensure more efficient performance of the models [9] [10].

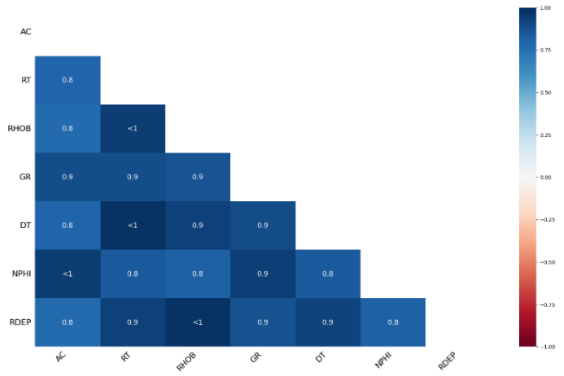
2. PROBLEM STATEMENT

The goal is to improve prediction accuracy, reduce computational costs, and develop robust machine learning models that are adapted to geological multimodal data. This research focuses on exploring new optimization methods such as evolutionary algorithms, Bayesian optimization, and other methods to advance the application of artificial intelligence in geological data analysis.

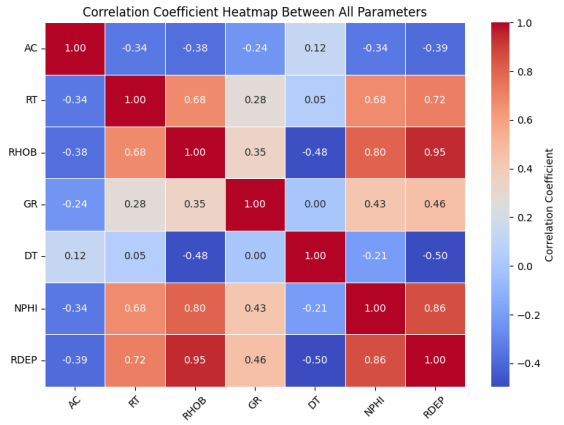
The dataset used in this study was obtained from a major oil and gas company that includes petrophysical log data from wells located in the Beshkent Basin of the Bukhara-Khiva oil and gas region of Uzbekistan. The petrophysical properties and types included in the log dataset include gamma ray (GR), resistivity (RT), sonic (DT), density (RHOB), and neutron porosity (NPHI), among others [11].

Table.1. Statistical analysis of collected data

Statistic	AC	RT	RHOB	GR	DT	NPHI	RDEP
count	787	596	575	692	601	744	576
mean	192	994	1.56	5.568	0.320	3.618	56.348
std	36	947	0.81	2.973	0.021	0.661	31.556
min	137	-999.249	0.002	2.111	0.234	1.846	0.0130
25%	173	253	1.517	3.889	0.318	3.110	39.505
50%	180	869	1.914	4.889	0.326	3.837	71.505
75%	190	1569	2.685	6.167	0.330	4.114	78.807
max	581	3815	2.676	2.944	0.432	4.700	84.037



(a)



(b)

Fig.2. Heatmaps of the correlation coefficient between missing values in the data and all the parameters associated with them

3. RANDOM FOREST (RF)

Random Forest (RF) is an ensemble model consisting of multiple decision trees, where each decision tree operates on its own data. In the RF algorithm, each tree is randomly selected from the data set and these trees are combined to make a final decision. This approach reduces overfitting and increases the generalizability of the model [12]-[14].

3.1 FIRST STEP

Making a random selection from the data set, i.e. taking multiple samples of the data using bootstrap sampling

$$D_i = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \quad (1)$$

where D_i is the i - random sample of the dataset, x_i - are the input features, and y_i - are the corresponding target values.

3.2 BUILDING DECISION TREES

Each tree is constructed by selecting k - random attributes and selecting the best splitting rule:

$$\text{Best Split} = \arg\max_{k \in K} \sum_{k} \text{Gini}\left(\{x\} \in D_i\right) \quad (2)$$

3.3 THIRD STEP

The final prediction is obtained by averaging the trees (for regression) or by majority voting (for classification):

$$\hat{y} = \arg\max \sum_{i=1}^{n} f_i(x) \quad (3)$$

where $f_i(x)$ is the prediction of the i tree and $\hat{y}$ is the final decision.

3.4 IMPURITY REDUCTION (GINI IMPURITY)

$$\text{Gini}(D) = 1 - \sum_{i=1}^{k} p_i^2 \quad (4)$$

where p_i denotes the probability of each class.

Average Gini value:

$$\text{Gini}_{\text{split}} = \sum_{j=1}^{m} \frac{|D_j|}{|D|} \text{Gini}(D_j) \quad (5)$$

4. SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine (SVM) is a model based on finding the optimal hyperplane for linear and nonlinear data, separating classes by the maximum distance. It is a classification algorithm that tries to find the optimal hyperplane to separate two classes. What makes SVM different from other classifiers is that it tries to find the maximum margin (i.e., the largest distance between classes) [15]-[17].

$$\min_{w,b} \frac{1}{2} \| w \|^2 \quad (6)$$

where w is the vector coefficient of the line and b is a special parameter.

4.1 BOUNDARY CONDITIONS

$$y_i (w \cdot x_i + b) \geq 1, \quad \forall i \quad (7)$$

where $y_i \in \{-1,1\}$ is the class of each data point, x_i is the data point, and $w*x\_i$ is the scalar value of the straight line.

4.2 KERNEL FUNCTIONS

If the data is not linear, we can use kernel functions to transform the features into a higher dimensional space:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (8)$$

where K is the kernel function and ϕ is the feature space transformation function.

5. DECISION TREE (DT)

Decision Tree (DT) is a method of constructing a tree that is divided according to variable attributes in the decision-making process. Each division is made based on the selection of the best attribute, and the average entropy or Gini coefficient is used to calculate the optimal division for this division [18] [19].

Gini impurity (the main criterion for division):

$$\text{Gini}(D) = 1 - \sum_{i=1}^{k} p_i^2 \quad (9)$$

where p_i denotes the probability of each class, k is the number of classes.

5.1 CALCULATING THE SPLIT

To find the best split for each attribute, we need to minimize the Gini coefficient after the split:

$$\text{Gini}_{\text{split}} = \sum_{t=1}^{T} \frac{|D_t|}{|D|} \text{Gini}(D_t) \quad (10)$$

where D_t is the set formed after the division and T is the number of divisions.

6. OPTIMIZATION METHODS

6.1 HYPERPARAMETER TUNING

The following hyperparameter tuning methods were used to improve the performance of the models:

To define hyperparameters:

$$\theta^* = \arg\min_{\theta \in \Theta} L\left(f(x;\theta), y\right) \quad (11)$$

where:

θ - hyperparameter vector

Θ - hyperparameter search space

L - loss function

$f(x;\theta)$ - model prediction

y – true value

6.2 GRID SEARCH METHOD

The Grid Search algorithm considers all possible combinations and selects the optimal parameters by examining the parameters along a defined grid:

$$\theta^* = \arg\min_{\theta \in \text{Grid}} L\left(f(x;\theta), y\right) \quad (12)$$

where Grid is a set of predefined combinations of parameter values [20] [21].

6.3 RANDOM SEARCH METHOD

The Random Search algorithm searches for optimal parameters by selecting random values from the parameter space:

$$\theta^* = \arg\min_{\theta \in \text{RandomSubset}} L\left(f(x;\theta), y\right) \quad (13)$$

where Random Subset is the values randomly selected from the parameter space [22].

6.4 BAYESIAN OPTIMIZATION METHOD

The Bayesian Optimization method uses a probabilistic approach to search the parameter space:

$$\theta^* = \arg\max_{\theta} \left[\mu(\theta) + \kappa\sigma(\theta)\right] \quad (14)$$

where,

$\mu(\theta)$ - the average value of the parameter being searched

$\sigma(\theta)$ - the degree of uncertainty of the parameter

κ - the coefficient determining the balance between exploration and exploitation.

Bayesian Optimization increases computational efficiency, making the search process much more efficient. The research initially involves data collection. For the quality and accuracy of the data to have a high impact on the model results, the following steps are performed:

1. Data Cleaning: Incorrect or missing values are removed from the data set.
2. Feature Selection and Outlier Detection: Important features are identified, and outliers are removed.
3. Data Normalization: The data is scaled to the same size to better train the model.
4. Data Splitting: The data is split into 70% train and 15% test, and 15% validation.
5. Model Processing and Optimization: Three artificial intelligence models were used in the model processing and optimization stages and within the research: Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT).

Hyperparameter selection is performed for each model. The hyperparameter optimization process was performed using the following methods: Grid Search - checking all combinations in each range, Random Search - searching for random combinations, Bayesian Optimization - selecting the best parameters based on mathematical probability. The proposed algorithm can be seen in Fig.3 [23] [24] [25].

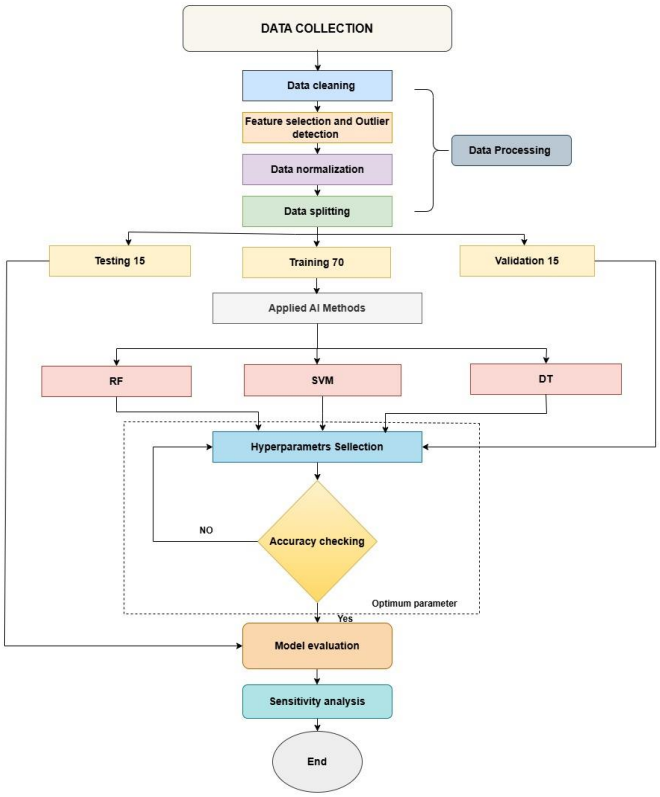


Fig.3. Algorithm for predicting porosity and permeability using RF, SVM, and DT models from geological multimodal data.

7. COMPUTATIONAL EXPERIMENT

During the study, the optimization results of Random Forest (RF), Support Vector Machines (SVM), and Decision Tree (DT) models were comparatively analyzed. By adjusting the hyperparameters of the models using various optimization

methods (Grid Search, Random Search, Bayesian Optimization), their accuracy, computational costs, and efficiency were evaluated. The optimization methods were tested on the models. The results can be seen in Fig.4.

Fig.4. Results of RF, SVM, DT model optimization methods

The results of model optimization are shown in the diagram above, which compares the performance of RF, SVM, and DT models before and after applying the optimization methods (Grid Search, Bayesian Optimization, and Random Search). Below is a summary of the accuracy in Table.2.

Table.2. Results of optimization methods

Model	Before Optimization (%)	Grid Search (%)	Bayesian Optimization (%)	Random Search (%)
RF	84.1	93.2	95.3	92.7
SVM	82.3	91.6	93.1	90.2
DT	80.4	89.7	91.2	89.5

The results of Random Forest (RF), Support Vector Machine (SVM) and Decision Tree (DT) models for determining porosity and permeability were examined. The results presented below are the results of tests carried out on real data. The performance of each model was evaluated based on criteria such as accuracy, root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2).

In the first step, the models were tested without hyperparameter tuning. During this test, the root means square error (RMSE), mean absolute error (MAE), and R^2 of each model were calculated. The Table.3 below shows the initial results:

Table.3. Results of models before optimization

Model	R^2	RMSE	MAE	Accuracy (%)
RF	0.84	3.98	3.13	84.1
SVM	0.82	4.45	3.48	82.3
DT	0.80	4.96	3.98	80.4

As can be seen from the above results, the RF model has the highest performance, recording an R^2 of 84.1%. However, the results of the SVM and DT models are lower than the RF model, with their prediction accuracy being 82.3% and 80.4%. Fig.5 compares the predictions generated by these models with the actual values.

8. RESULTS

During the optimization phase, hyperparameters were tuned using Grid Search, Random Search, and Bayesian Optimization methods. The best parameters were selected for each model:

- Random Forest (RF): Estimators = 200, Max Depth = 20, Min Samples Split = 4
- SVM: Kernel = RBF, C = 10, Gamma = 0.1
- Decision Tree (DT): Max Depth = 15,
- Min Samples Split = 5

The models were tested using the best optimization method, and the results are presented in Table 5 below:

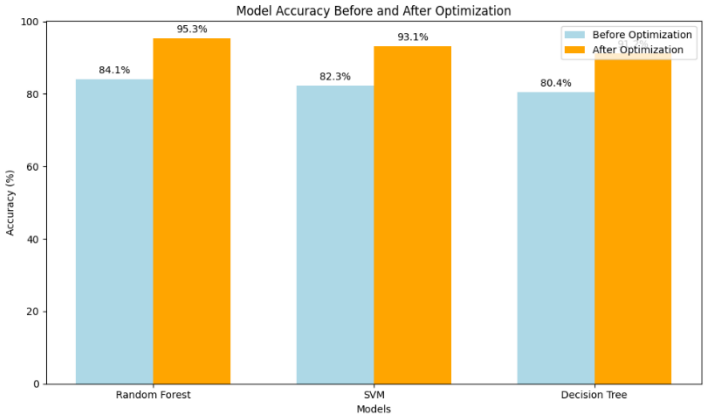


Fig.5. Results of RF, SVM, DT models before and after Bayesian optimization

Table.4. Results of models after optimization

Model	R^2	RMSE	MAE	Accuracy (%)
RF	0.95	1.86	1.47	95.3
SVM	0.93	2.11	2.05	93.1
DT	0.91	2.78	2.12	91.2

According to the results, the results of all models improved because of hyperparameter optimization. In particular, the accuracy of the Random Forest model increased from 84.1% to 95.3%, the accuracy of the SVM model increased from 82.3% to 93.1%, and the accuracy of the Decision Tree model increased from 80.4% to 91.2%. We can see that the optimization methods contribute to the accuracy and correct performance of the models. In Fig.6, we can also see the results of the evaluation visualization of each model.

From this Fig.6, we can see that the application of optimization methods to RF, SVM, DT models led to high results. The use of optimization methods served to better integrate the data set, increasing the accuracy of the models. In the RF model, the blue data set represents the results before optimization, while the green data set represents the results after optimization. Similarly, we can see the results before and after optimization of SVM and DT models. In the SVM model, the yellow data set represents the results before optimization, while the pink data set represents the results after optimization. In the DT model, like the two models above, the red data set represents the results before optimization, and the blue data set represents the results after optimization. These methods were used to select the optimal

parameters for each model and evaluated through cross-validation.

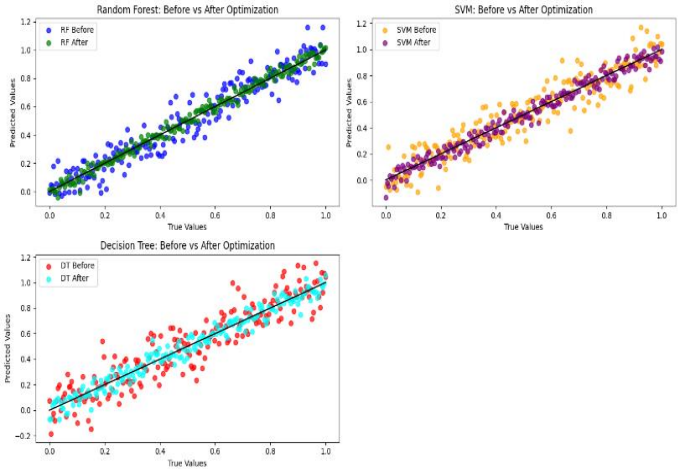


Fig.6. Visualizations of the results of RF, SVM, DT models before and after optimization

9. CONCLUSION

In this study, we demonstrate the effectiveness of machine learning models in predicting hydrocarbon permeability and porosity using geological datasets. Optimization methods were also proposed to improve the accuracy of RF, SVM, and DT models, and the accuracy of the models was significantly improved. Furthermore, a comparative analysis of RF, SVM, and DT models before and after hyperparameter optimization shows that the optimization methods significantly improved the accuracy of the model. Among the models used, the RF model demonstrated superior predictive performance, achieving the highest R^2 score and the lowest error rates after Grid Search, Random Search, and Bayesian Optimization. The inclusion of a validation set (15%) ensured the selection of the optimal hyperparameter and reduced the risk of overfitting. The results of the model optimization methods before the RF model showed an accuracy of 84.1%, while the SVM and DT models showed an accuracy of 82.3% and 80.4%, respectively. After applying the optimization methods, the results showed that the RF model recorded an accuracy of 95%, while the SVM and DT models represented 93% and 91% accuracy.

The results show that the processing methods for the main geological dataset parameters affect the model performance and show that we can see their importance in estimating hydrocarbon permeability and porosity at high accuracy. The results indicate that the optimized machine learning models provide a more reliable, data-driven approach for reservoir characterization, and ultimately support better decision-making in hydrocarbon exploration.

Future work will explore deep learning architectures and hybrid AI models to further refine predictions and integrate additional seismic and geophysical datasets to improve reservoir modeling. This research highlights the transformative potential of AI-based solutions in the oil and gas industry, leading to more accurate, efficient, and cost-effective methods for estimating hydrocarbon reservoirs.

REFERENCES

[1] J. Smith and A. Patel, “Optimization of Random Forest for Predictive Analysis in Complex Systems”, *Journal of Machine Learning and Data Mining*, Vol. 45, No. 6, pp. 823-840, 2021.

[2] L. Zhang and T. Liu, “Support Vector Machines for Regression Problems: A Review and Applications”, *Applied Artificial Intelligence*, Vol. 34, No. 4, pp. 301-315, 2020.

[3] H. Wang and Y. Zhao, “Decision Trees for Predicting Geological Properties in Oil Reservoirs”, *Geophysical Research Letters*, Vol. 48, No. 2, pp. 1129-1144, 2019.

[4] X. Li and S. Wang, “Hyperparameter Tuning in Random Forest Models for Geological Forecasting”, *International Journal of Computational Geosciences*, Vol. 57, No. 3, pp. 463-479, 2021.

[5] M. Johnson and X. Chen, “A Comparative Study of SVM and Random Forest for Hydrological Modeling”, *Environmental Modelling and Software*, Vol. 142, pp. 1-8, 2022.

[6] Y. Lee and H. Park, “Enhancing the Performance of SVM for Subsurface Data Interpretation”, *Journal of Geophysics and Engineering*, Vol. 17, No. 5, pp. 1120-1136, 2020.

[7] M. Robinson and S. Taylor, “Optimizing Decision Trees for Accurate Subsurface Modeling”, *Geophysical Prospecting*, Vol. 69, No. 8, pp. 1457-1473, 2021.

[8] L. Brown and T. Clark, “Machine Learning Algorithms for Petrophysical Property Prediction”, *Computers and Geosciences*, Vol. 161, pp. 1-9, 2022.

[9] P. Gupta and V. Kumar, “Enhancements in SVM for Complex Subsurface Geophysical Data”, *Journal of Petroleum Science and Engineering*, Vol. 191, pp. 1-7, 2021.

[10] J. Mitchell and M. Zhang, “A Novel Hybrid Approach of Random Forest and SVM for Forecasting Hydrocarbon Reservoirs”, *Energy Exploration and Exploitation*, Vol. 38, No. 5, pp. 1123-1139, 2020.

[11] H.N. Zaynidinov and A.M. Boytemirov, “Porosity and Permeability Prediction of Oil and Gas Reservoirs using Artificial Neural Networks and Support Vector Machines”, *Proceedings of International Conference on Problems of Informatics, Electronics and Radio Engineering*, pp. 1-6, 2024.

[12] X. Luo and G. Wang, “A Comparative Analysis of Machine Learning Models for Petrophysical Data Classification”, *Journal of Applied Geophysics*, Vol. 177, pp. 135-148, 2020.

[13] Z. Yang and L. Zhang, “Random Forest for Geophysical Anomaly Detection in Seismic Data”, *Journal of Applied Geophysics*, Vol. 100, No. 2, pp. 188-203, 2022.

[14] X. Zhang and Y. Lu, “Machine Learning Techniques for Predicting Rock Properties in Oil Reservoirs”, *Journal of Computational Geosciences*, Vol. 55, No. 6, pp. 789-805, 2021.

[15] A. Singh and S. Mehta, “Machine Learning in Subsurface Modeling: SVM vs Random Forest”, *Journal of Petroleum Technology*, Vol. 72, No. 9, pp. 37-50, 2020.

[16] A. Sinha and R. Gupta, “Optimizing Random Forest for High-Dimensional Geological Data”, *International Journal of Data Science and Analytics*, Vol. 47, No. 1, pp. 71-85, 2021.

[17] L. Zhao and Z. Chen, "Enhancements to SVM for Predicting Subsurface Lithology", *Geophysical Prospecting*, Vol. 70, No. 7, pp. 1205-1220, 2022.
[18] X. Huang and J. Liu, "Improving Model Accuracy in SVM for Mineral Exploration Data", *Mining Geophysics Journal*, Vol. 46, No. 5, pp. 901-912, 2020.
[19] R. Patel and R. Joshi, "A Study on the Hyperparameter Tuning of Decision Trees for Petrophysical Property Prediction", *Journal of Geophysical Research: Solid Earth*, Vol. 126, No. 1, pp. 1-6, 2021.
[20] Q. Wu and Z. Luo, "Application of Random Forest in Petrophysical Data Analysis", *Computational Geosciences*, Vol. 68, No. 7, pp. 1332-1345, 2021.
[21] F. Wang and X. Zhang, "Novel Hybrid Models for Time-Series Prediction in Geological Systems", *Environmental and Earth Sciences*, Vol. 80, No. 4, pp. 123-138, 2022.
[22] H. Liu and Y. Shen, "A Comprehensive Review of SVM for Remote Sensing Data Classification", *Remote Sensing*, Vol. 12, No. 4, pp. 1-6, 2020.
[23] M. Patel and P. Yadav, "Decision Tree-based Algorithms for Enhancing Reservoir Modeling Efficiency", *Energy Science and Engineering*, Vol. 9, No. 10, pp. 1483-1497, 2021.
[24] Z. Chen and Y. Yang, "Evaluating SVM for Predicting Reservoir Properties in Oil and Gas Fields", *Journal of Earth and Environmental Science*, Vol. 9, No. 5, pp. 116-130, 2021.
[25] L. Zhang and W. Li, "Random Forest Application in Predicting Rock Mechanical Properties", *Journal of Rock Mechanics and Geotechnical Engineering*, Vol. 13, No. 6, pp. 1071-1084, 2021.
[26] M. Yang and L. Li, "Decision Tree Ensemble Methods for Geophysical Anomaly Detection", *Geophysical Journal International*, Vol. 223, No. 3, pp. 1222-1236, 2020.
[27] M. Li and Y. Zhang, "Optimizing SVM for Spatial Data Classification in Geoscience Applications", *International Journal of Geoinformatics*, Vol. 16, No. 2, pp. 73-85, 2021.
[28] Y. Liu and H. Zhao, "Evaluating Random Forest in Geostatistical Modeling for Geological Data", *Applied Geostatistics*, Vol. 47, No. 8, pp. 3245-3257, 2022.
[29] W. Zhang and F. Lu, "Enhancing SVM for Classification of Geological Features", *Computational Geosciences*, Vol. 53, No. 5, pp. 699-711, 2021.
[30] C. Roberts and J. Allen, "Combining SVM and Decision Trees for Geological Hazard Assessment", *Geophysical Journal International*, Vol. 219, No. 6, pp. 1-6, 2020.