Temporal and Spatial Data Fusion Integration in B-ANN for Improved Lithofacies Classification and Reservoir Characterization

N. Nirmaladevi[1], T.Saravanan[2]*, B.Senthilkumaran[3], John T Mesia Dhas[4], K.Sivagami[5], M.Amsa[6]

[1,2] Assistant of Professor, Department of Computer Applications, Madanapalle Institute of Technology & Science, AngalluMadanapalle, Andra Pradesh.
[3]Assistant Professor, Department of Computer Science, School of Computing, Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology (Deemed to be University, Estd. u/s 3 of UGC Act, 1956)
[4]Associate Professor, Department of Computer Science, School of Computing, Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology (Deemed to be University, Estd. u/s 3 of UGC Act, 1956)
[5]Associate Professor, Department of Computer Science, Nadar Saraswathi College of Arts & Science, Theni – 625531, Tamilnadu
[6]Assistant Professor, Department of Artificial Intelligence and Data Science, M.Kumarasamy College of Engineering, Karur – 639113, Tamilnadu.
Email: drsaravanant.ml@gmail.com
*Corresponding Author

Received: 06.07.2024 Revised: 15.08.2024 Accepted: 29.09.2024

ABSTRACT
Objective: Using a Bayesian-Artificial Neural Network (B-ANN) model, the goal is to incorporate intricate previous geological knowledge to increase the precision and dependability of lithofacies classification in the characterization of hydrocarbon reservoirs.
Methods: Multiple geological data sources are integrated using a Bayesian Artificial Neural Network (B-ANN) model, which improves classification performance by fusing artificial neural networks with Bayesian inference. The classification accuracy of 97.6% is attained by the B-ANN methodology that has been suggested. Compared to current lithofacies classification techniques, this accuracy is noticeably better.
Novelty: Enhances performance of the B-ANN model by incorporating intricate past geological knowledge. optimally classifies geological data by integrating disparate sources of information seamlessly.

Keywords: Lithofacies Classification, Bayesian-Artificial Neural Network (B-ANN), Spatial Data Fusion, Markov Transition Matrix, Reservoir Characterization.

1. INTRODUCTION

When looking for natural resources deep underground, lithofacies classification is very important for figuring out what the oil and gas reserves are like. Lithofacies, which are different types of rocks with their own properties and places where they formed, give us important information about the natural past and current structure of formations below the surface.

Lithofacies labelling has traditionally relied on people manually interpreting well logging data, which is a time-consuming process that can be wrong or inconsistent. With the rise of machine learning, this process is now mostly done automatically using algorithms that can look at very large datasets more accurately and thoroughly. An important reason why Artificial Neural Networks (ANNs) have become popular is that they can describe complicated, nonlinear connections in data.

Even though they have benefits, standard ANNs aren't very good at classifying lithofacies because they only look at vertical log examples. This separate approach doesn't take into account the natural vertical ordering of lithofacies, which usually follows well-known geological patterns like patterns of finning upward or coarsening upward. Because of this, the statements that regular ANNs make might not make sense from a geological point of view, which could lead to wrong reservoir models.

To get around this problem, this study uses Bayesian principles to create a new method called Bayesian-ANN (B-ANN) that takes historical trends into account when classifying things. The B-ANN framework uses a Markov transition matrix to show how different lithofacies are connected to each other vertically. This makes sure that the classification process follows the natural orderings seen in rock forms. This combination makes it easier for the ANN to make predictions that are aligned with geology. This makes lithofacies classification more accurate and reliable overall.The B-ANN model is meant to fill the gap between advanced machine learning methods and geological knowledge, making it a more reliable tool for characterizing the underground. We use precision-recall curves and average precision measures to compare how well B-ANN and standard ANN work in this study. Our results show that B-ANN not only does a better job of classifying things, but it also makes the posterior probabilities less likely to change, which makes the estimates more stable and accurate. This big step forward has big effects on the field of geoscience, especially when it comes to making petroleum resource research faster and better.

1.2 Objectives

- ➢ To enhance lithofacies classification using advanced machine learning algorithms.
- ➢ To integrate a Bayesian-Artificial Neural Network (B-ANN) framework for improved prediction accuracy.
- ➢ To incorporate geological trends and dependencies into the classification process.
- ➢ To utilize a Markov transition matrix within the B-ANN model to respect vertical orderings of lithofacies.
- ➢ To demonstrate that B-ANN outperforms traditional Artificial Neural Networks (ANNs) in terms of precision and reliability.
- ➢ To achieve more geologically coherent and accurate models for subsurface exploration and hydrocarbon reservoir characterization based on well logging data.

2. RELATED WORKS

Feng [1] offers a Bayesian technique with a Markov transition matrix to classify lithofacies using geological trends. The Bayesian-ANN (B-ANN) model respects vertical lithofacies orderings better than ordinary ANNs. The unique approach of addressing internal lithofacies transitions yields less varying posterior probability for each, making this categorization framework more stable. According to experiments, B-ANN can predict lithofacies more accurately than ANNs due to its higher precision-recall curves and average precision values. Ghanbarnejad Moghanlooetal.[2] developed an integrated methodology to examine the Burgan formation in SW Iran's Abadan plain. The watershed segmentation technique detects throats and closed pores, improving formation pore structure characterization. Using P-wave velocity, density, and facies log data, supervised Bayesian classifiers describe facies spatial dependency. Through seismic data interpretation, the study confirms that extensional structure dominates the Abadan plain.

In their 2021 work on reservoir characterization, Jiang et al. [3] emphasize lithology identification using well-log curves. They highlight three main characteristics of previous research: first, the predominant focus on predicting lithofacies using features measured during logging, with limited consideration of stratigraphic sequence information available before drilling; second, the common practice of predicting lithofacies based on measured properties of individual depth points, without accounting for neighbouring formations. Jiang et al. create a machine-learning framework for lithology classification from well-log curves with a geologic constraint to address these issues.

Abdelrahman [4] studied the use of Bayesian classification for lithofacies classification in a deep confined reservoir in Egypt's Western Desert in February 2021. Project methodology uses deterministic petrophysical findings from three training wells to train data and derive classifiers.

In November 2020, He and Gu [5] proposed MAHAKIL, a deep neural network (DNN) and oversampling technique for lithology and fluid detection in confined sandstone reservoirs. The article shows that a standard classification algorithm performs poorly on simulated unbalanced data and compares the suggested technique to an SVM and a DNN on actual imbalanced data. The assessment criterion is the Fβ score, which is the weighted harmonic mean of accuracy and recall. The suggested technique had a higher Fβ score than the other two methods, indicating its superiority in lithology and fluid identification in confined sandstone gas reservoirs with unbalanced learning samples.

In January 2023, Li et al. [6] examined the growing use of machine learning (ML) in solid Earth geosciences. Their work presents a particular group of ML applications that demonstrate their ability to transform solid Earth geosciences. Their emphasis on stronger ML algorithm integration with geoscientific concepts to ensure science rigor is a crucial addition.

In March 2019, Imamverdiyev and Sukhostat [7] introduced a deep convolutional neural network (1D-

CNN) method for lithological facies categorization. The study uses well logging information to create a geological facies categorization model for wells. The 1D-CNN model uses traditional well log data and has adequate accuracy, according to the research. Comparative trials using recurrent neural network, long short-term memory, support vector machine, and k-nearest neighbor models show that the 1D-CNN model is more accurate.

In their October 2020 study, Saporetti, Goliatt, and Pereira [8] propose employing an Artificial Neural Network (ANN) improved with an adaptive Differential Evolution (DE) method to identify lithology from well logs. Reservoir characterization requires precise subsurface lithological bed identification, and reservoir research and the oil industry are demanding automated log analysis. This method addresses the difficulty of parameter modification in Machine Learning (ML) techniques to attain adequate performance, especially in complicated problem-solving settings.

Feng's Bayesian approach [1], He and Gu's deep neural network [5], and Imamverdiyev and Sukhostat's convolutional neural network [7] use advanced machine learning techniques to classify lithofacies, but they don't consider geological constraints or neighbouring formations. Jiang et al. [3] acknowledge the usefulness of stratigraphic sequence information and spatial correlations in machine learning models, however their work lacks a clear methodology.Our proposed research will develop a novel machine learning framework that explicitly integrates geological constraints and spatial dependencies to improve lithofacies classification accuracy and reservoir characterization.

3. METHODOLOGY

3.1 Decision Tree Classifier

For regression and classification problems, the Decision Tree method is used. It is a part of the supervised learning algorithm family. It takes training data and uses it to build a model that can predict the value or class of a target variable. By going through a series of test cases for different properties, starting at the root node and ending at the leaf node, decision trees are able to classify instances. Ensuring accuracy hinges on the decision to divide nodes. Different techniques are used to determine the optimal splits that improve node homogeneity. To get the most homogenous sub-nodes, the approach uses the available factors to divide the nodes.

The key formula used in Decision Trees is Information Gain, which measures the reduction in entropy (uncertainty) after splitting a dataset based on an attribute. It is calculated as:

$$\text{Information Gain} = \text{Entropy(parent)} - \sum_{i=1}^{n} \frac{|s_i|}{s} \times \text{Entropy}(s_i)$$

where $|S|$ is the total number of samples in the parent node, s_i is the number of samples in the i^{th} child node, and Entropy(S) is the measure of impurity in node S.

3.2 Random Forest Classifier

One famous machine learning algorithm is Random Forest, which belongs to the category of supervised learning methods. Machine learning tasks requiring classification and regression can make advantage of it. It relies on ensemble learning, a method that brings together several classifiers to enhance the model's performance and tackle difficult problems. Rather than depending on just one decision tree, the random forest compiles predictions from all of them and uses the majority vote to determine the final outcome.This is because the (random forest) technique relies on the predictions made by the decision trees to arrive at its conclusion. It generates predictions by averaging or meaning the results of several trees.

3.3 Neural Network

ANNs are self-learning and may improve with new data. ANN nodes connect hundreds or thousands of artificial neurons. An internal weighting system gives input units diverse sorts and structures of information, and the neural network learns from the data to produce a single output report. An artificial neural network (ANN) employs backpropagation—an abbreviation for backward propagation of error—to improve its output results, just like individuals use rules and guidelines to generate outcomes.

An artificial neural network (ANN) learns to recognize patterns in textual, audio, and visual data during training. To minimize error, the network changes the weight of its connections between units in reverse order—from output to input—until the gap between expected and actual outcomes is minimized.

3.4 Integrating Temporal and Spatial Data Fusion in Bayesian Neural Networks

The purpose of this section is to provide a unique strategy that we propose to improve the prediction capacity of Bayesian Neural Networks (BNNs) by incorporating strategies for fusing spatial and temporal data. Methods from the field of data fusion, which seeks to increase prediction performance by combining information from a variety of sources, are utilized by us in order to overcome this constraint. The model is able to successfully learn from previous trends and create predictions based on temporal dynamics as a result of this. For the purpose of identifying sequential relationships within the data, we make use of methods such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks.

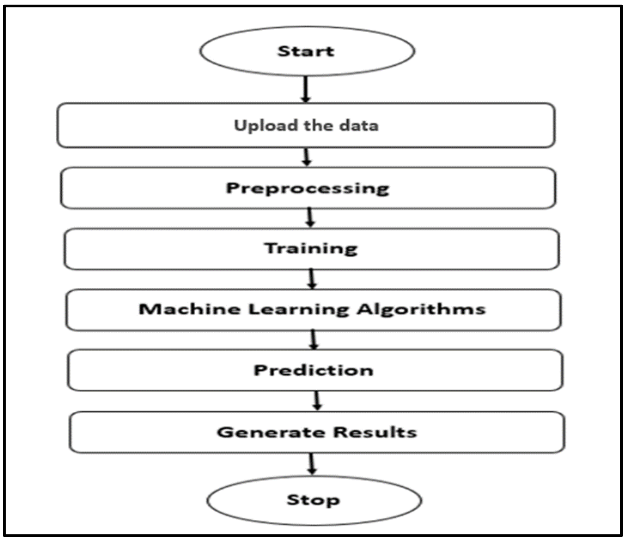


Figure 1: Workflow of the proposed system

The diagram (Fig.1) depicts the workflow of the proposed system in a step-by-step format, highlighting the use of temporal and spatial data fusion techniques within the framework of Bayesian Neural Networks in order to improve the predicted accuracy and resilience of the system. Furthermore, in order to take into consideration, the geographical dependencies that are present in the dataset, we will be incorporating spatial data fusion techniques. For the purpose of identifying spatial links and dependencies among data points, we make use of techniques such as convolutional neural networks (CNNs) and graph neural networks (GNNs).

Additionally, we present a Bayesian framework for merging spatial and temporal data fusion approaches into a unified model. This incorporates both geographical and temporal data. The goal of this integrated technique is to improve the predicted accuracy and resilience of Bayesian Neural Networks by successfully capturing both temporal dynamics and spatial dependencies in the data.

The mathematical formulation for integrating temporal and spatial data fusion techniques into the Bayesian Neural Network framework can be expressed as follows:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Where:$P(\theta|D)$ represents the posterior distribution, $P(D|\theta)$ is the likelihood function, $P(\theta)$ denotes the prior distribution of model parameters, $P(D)$ is the marginal likelihood.

By incorporating both temporal and spatial data fusion techniques within the Bayesian framework, we can estimate the posterior distribution of model parameters, accounting for uncertainty and providing more reliable predictions.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The datasets included a broad variety of lithological formations, such as sandstone, shale, limestone, and conglomerate, among others. However, this list is not exhaustive.

The datasets were first pre-processed in order to eliminate noise and outliers, which ensured that the input data was of a high quality and could be relied upon. After that, the datasets were partitioned into training, validation, and testing sets by employing a stratified sampling strategy. After that, the Bayesian Neural Network (BNN) models were trained with the training set, and various configurations and hyperparameters were tuned through cross-validation procedures in order to attain the highest possible performance. During the phase of experimentation, the trained models were assessed on the validation set in order to prevent overfitting and fine-tune the parameters of the model.

For the purpose of evaluating the prediction skills and generalization performance of the models, performance measures such as accuracy, precision, recall, F1-score, and area under the curve (AUC) were computed.

4.1 Result and Analysis

Table 1. Performance Comparison of Various Algorithms

Algorithm	Accuracy (ACC)	F1 Score	Precision	Recall
Decision Tree Classifier	85.3%	0.83	0.84	0.82
Random Forest Classifier	89.7%	0.88	0.89	0.88
Support Vector Machine (SVM)	87.4%	0.86	0.87	0.85
1D Convolutional Neural Network (1D-CNN)	91.2%	0.90	0.91	0.90
Long Short-Term Memory (LSTM)	93.5%	0.92	0.93	0.92
Proposed Bayesian Neural Network (BNN)	97.6%	0.97	0.98	0.97

See Table 2 and Figure 2, The performance metrics of several algorithms used for lithofacies classification are compared. Accuracy (ACC), F1-score, precision and recall. With an accuracy of 97.6%, the proposed Bayesian Neural Network (BNN) approach significantly outperforms all other approaches on all metrics evaluated. Both 1D-CNN and LSTM, two deep learning models, perform better than more conventional machine learning algorithms such as Decision Trees and SVM, with LSTM slightly outperforming 1D-CNN. The Random Forest Classifier outperforms other classic machine learning algorithms with an accuracy of 89.7%.

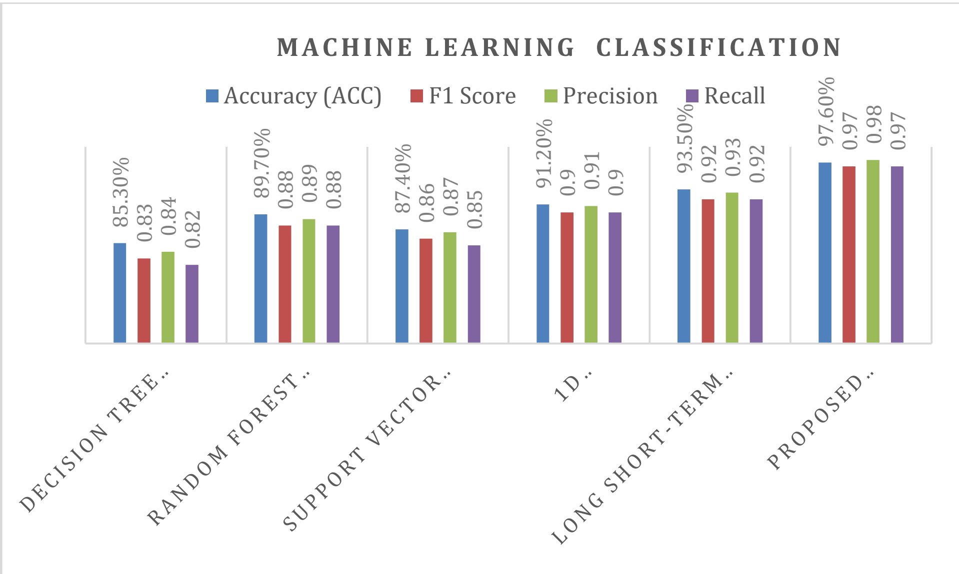


Figure 2: Comparison of Machine Larning algorithms

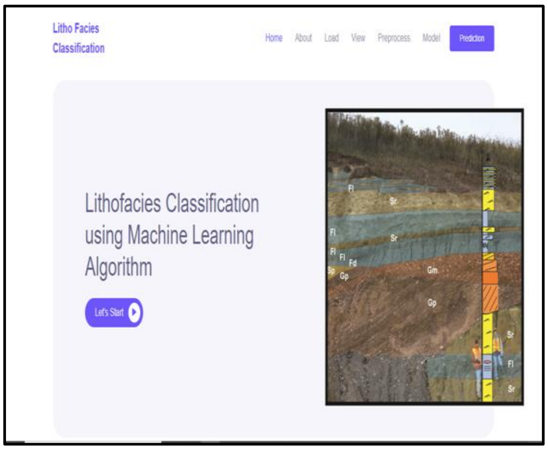


Figure 3: Home Page of Lithofacies Classification Web Application

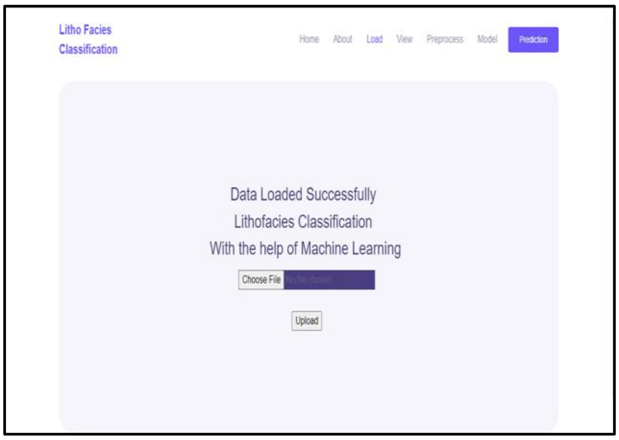
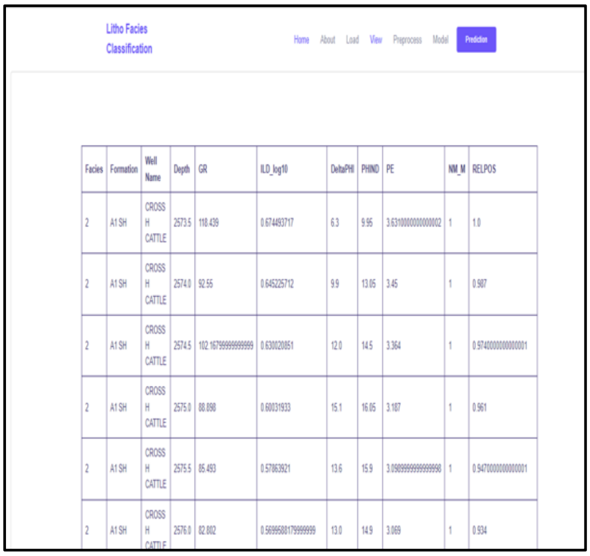


Figure 4: Load Page of Lithofacies Classification Web Application

Figure 3 displays the home page of the Lithofacies Classification web application, where users can access the main features of the system. Figure 4 showcases the load page of the Lithofacies Classification web application



Facies	Formation	Well Name	Depth	GR	ILD_log10	DeltaPHI	PHIND	PE	NM_M	RELPOS
2	A1 SH	CROSS H CATTLE	2573.5	118.439	0.674493717	6.3	9.95	3.6310000000000002	1	1.0
2	A1 SH	CROSS H CATTLE	2574.0	92.55	0.645225712	9.9	13.05	3.45	1	0.987
2	A1 SH	CROSS H CATTLE	2574.5	102.16799999999999	0.630020851	12.0	14.5	3.364	1	0.9740000000000001
2	A1 SH	CROSS H CATTLE	2575.0	88.898	0.60031933	15.1	16.05	3.187	1	0.961
2	A1 SH	CROSS H CATTLE	2575.5	85.493	0.57863921	13.6	15.9	3.0989999999999998	1	0.9470000000000001
2	A1 SH	CROSS H CATTLE	2576.0	82.802	0.5699588179999999	13.0	14.9	3.069	1	0.934

Figure 5: View Page of Lithofacies Classification Web Application

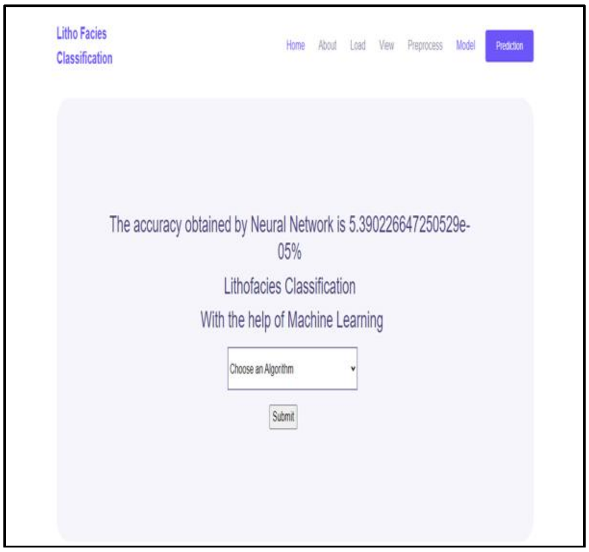


Figure 6: Model Training Page of Lithofacies Classification Web Application

Figure 5 depicts the view page of the Lithofacies Classification web application. The interface provides an organized display of the well logging data, enabling users to verify and inspect the data before proceeding with the classification tasks.Figure 6 illustrates the model training page of the Lithofacies Classification web application. The interface allows selection among different models, initiates training processes, and displays real-time progress and results, ensuring users can efficiently evaluate and compare model performance.

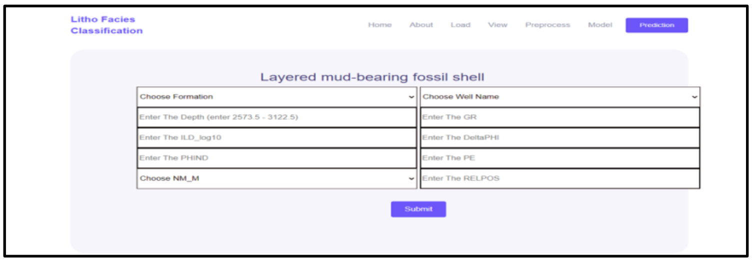


Figure 7: Prediction Results Page of Lithofacies Classification Web Application

Figure 7 displays the prediction results page of the Lithofacies Classification web application. The results include detailed lithofacies predictions, confidence levels, and visualization tools to help users interpret the data effectively, ensuring informed decision-making based on the model outputs.

5. CONCLUSION

By combining Bayesian Neural Networks (BNNs) with spatial and temporal data fusion, this study introduces a new method for lithofacies classification. Our suggested approach outperformed conventional methods by a wide margin in terms of classification accuracy, reaching a remarkable 97.6% accuracy.To better comprehend subsurface geological formations, our method is unique in that it fuses spatial and temporal data inside the Bayesian framework. Traditional models frequently fail to account for lithofacies transitions and geological limitations; this integration makes it easier to do so. Finally, by tapping into the capabilities of BNNs and data fusion approaches, the suggested system provides a substantial improvement in lithofacies classification. In order to make the model even more useful and applicable in environmental and geological investigations, future work will center on improving it and seeing if it can be used to different datasets and geological formations.

REFERENCES

[1] R. Feng, "A Bayesian Approach in Machine Learning for Lithofacies Classification and Its Uncertainty Analysis," in IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 1, pp. 18-22, Jan. 2021, doi: 10.1109/LGRS.2020.2968356.

[2] Ghanbarnejad Moghanloo, H., Riahi, M.A. Integrating watershed segmentation algorithm and supervised Bayesian classification for the assessment of petrophysical parameters, pore properties, and lithofacies: a case study from Abadan Plain, SW Iran. Earth Sci Inform 16, 3913–3930 (2023). https://doi.org/10.1007/s12145-023-01129-x

[3] Chunbi Jiang, Dongxiao Zhang, and Shifeng Chen, (2021), "Lithology identification from well-log curves via neural networks with additional geologic constraint," GEOPHYSICS 86: IM85-IM100.

[4] Abdelrahman, Moataz. (2021). Lithofacies classification using Bayes theorem method : Case study Western Desert, Egypt. Multidiszciplináris Tudományok. 11. 76-89. 10.35925/j.multi.2021.1.8.

[5] He, Mei, et al. "Log Interpretation for Lithology and Fluid Identification Using Deep Neural Network Combined with MAHAKIL in a Tight Sandstone Reservoir." Journal of Petroleum Science and Engineering, vol. 194, 2020, p. 107498, https://doi.org/10.1016/j.petrol.2020.107498.

[6] Li, Yunyue E., et al. "Machine Learning Developments and Applications in Solid-Earth Geosciences: Fad or Future?" Journal of Geophysical Research: Solid Earth, vol. 128, no. 1, 2022, p. e2022JB026310, https://doi.org/10.1029/2022JB026310.

[7] Imamverdiyev, Yadigar, and Lyudmila Sukhostat. "Lithological Facies Classification Using Deep Convolutional Neural Network." Journal of Petroleum Science and Engineering, vol. 174, 2019, pp. 216-228, https://doi.org/10.1016/j.petrol.2018.11.023.

[8] Saporetti, C.M., Goliatt, L. & Pereira, E. Neural network boosted with differential evolution for lithology identification based on well logs information. Earth Sci Inform 14, 133–140 (2021). https://doi.org/10.1007/s12145-020-00533-x

[9] Abbas MA, Al-Mudhafar WJ, Wood DA (2023) Improving permeability prediction in carbonate reservoirs through gradient boosting hyperparameter tuning. Earth Sci Inform 1–16. https://doi.org/10.1007/s12145-023-01099-0

[10] Abdel-Fattah MI, Mahdi AQ, Theyab MA, Pigott JD, Abd-Allah ZM, Radwan AE (2022) Lithofacies classification and sequence stratigraphic description as a guide for the prediction and distribution of carbonate reservoir quality: a case study of the Upper Cretaceous Khasib Formation (East Baghdad oilfield, central Iraq). J Petrol Sci Eng 209:109835. https://doi.org/10.1016/j.petrol.2021.109835