

Application of machine learning for predicting potential hydrocarbon zones in wells – X field, Northern Red River basin

DO Thi Thuy Linh^{1*}, DOAN Ngoc San², LUU Khanh Linh³, TRAN Thi Oanh¹

¹ Faculty of Petroleum, PetroVietnam University, Ho Chi Minh City, Vietnam

² Hi-Technology of Applied Geology & Geophysics Company, 606/38/8, Quarter 4, Highway 13, Hiep Binh Phuoc Ward, Thu Duc city, Ho Chi Minh City, Vietnam

³ Aeron Solution Company, Ho Chi Minh City, Vietnam

* Corresponding email: linhdtt@pvu.edu.vn

Abstract: *The explosion of machine learning (ML) models that are capable of analyzing big data shed new light on oil and gas industry. This paper investigates the application of machine learning algorithms to predict zones with high potential for hydrocarbon reservoirs based on well data. By using three well data of Well-A, Well-B and Well-C in the north of Red River Basin and using ML algorithms including decision tree (DT), random forest (RF), extreme Gradient Boosting (XGB), artificial neural network (ANN), Self-Organizing Map (SOM), the attribute sets of hydrocarbon reservoir are determined by using unsupervised classification. After that, supervised classification is used to predict the potential hydrocarbon zones based on that attribute sets. Mentioned algorithms are compared with accuracy, and the results show that SOM displays very satisfactory results and predicts additional potential hydrocarbon zones. The paper also compares the performance of various models and proposes an optimization approach that integrates geophysical parameters and data preprocessing techniques to improve prediction reliability.*

Keywords: *AI, Machine learning - ML, Hydrocarbon reservoir, Well logging, Classification.*

1. Introduction

The exploration and production of oil and gas play a crucial role in the global economy, providing energy for many sectors such as manufacturing, transportation, and industry. However, exploration is becoming increasingly complex and expensive due to the depletion of traditional reservoirs and the challenges of extracting resources from hard-to-reach deposits. With the growing complexity of oil and gas reservoirs, along with the increasing volume of data from well logs and other geological attributes, machine learning (ML) is being widely adopted in oil and gas exploration and forecasting (Adibifard et al., 2014; Dong et al., 2021; Fredrick & Pius, 2021; Amir et al., 2022; Liu & Liu, 2022). The application of ML not only enhances prediction accuracy but also optimizes costs and time, reduces risks, and improves decision-making capabilities (Raeesi et al., 2012; Ren et al., 2019; Doan et al., 2019; Ta et al., 2023). These benefits are driving the growing adoption of ML throughout the entire process of oil and gas extraction and forecasting.

This study utilizes data from three exploration wells named well A, well B, and well C located in X field, northern Red River Basin. Gas has been discovered in all these wells. The study aims to predict hydrocarbon reservoirs from well data through AI tools by building an integrated model of well log data with interpreted data, and analysis of rock samples, hydrocarbon samples, well testing, and lithology. However, gaps in the raw log data (resistivity, density, gamma ray, etc.) and interpretation results (facies, lithology, porosity, permeability, etc.) - caused by measurement errors or recording failures - make accurate interpretation challenging. To address this, this data will be adjusted, and missing segments will be predicted to create a more complete dataset before proceeding with model training to achieve more accurate analysis and prediction results. Furthermore, the well log data is enriched by supplementing statistical attributes to generate valuable supporting information, increasing the reliability of AI predictions (Doan et al., 2021). Various ML algorithms are employed to forecast the potential hydrocarbon zones of these wells.

Case study and available data

The study area, situated in the northern part of the Red River Basin (Figure 1), is characterized by complex geological structures and sedimentary environments (Nguyen Manh Huyen & Ho Dac Hoai, 2007; Bui, 2015). It is considered as high potential area for oil and gas, with multiple discoveries reported in various reservoirs, including fractured carbonate basement from the pre-Cenozoic and clastic sedimentary formations from the Oligocene and Miocene (Bui, 2015; Dau et al., 2014; Nguyen Manh Huyen & Ho Dac Hoai, 2007; Nguyen Hiep et al., 2007). As oil and gas reserves decline and major reservoirs become

depleted, the search for potential reservoirs has become a key focus in exploration efforts, particularly in regions with confirmed petroleum systems and prior discoveries. This study, therefore, aims to support the prediction of potential hydrocarbon reservoirs in wells drilled within Area X, located in the northern Red River Basin.

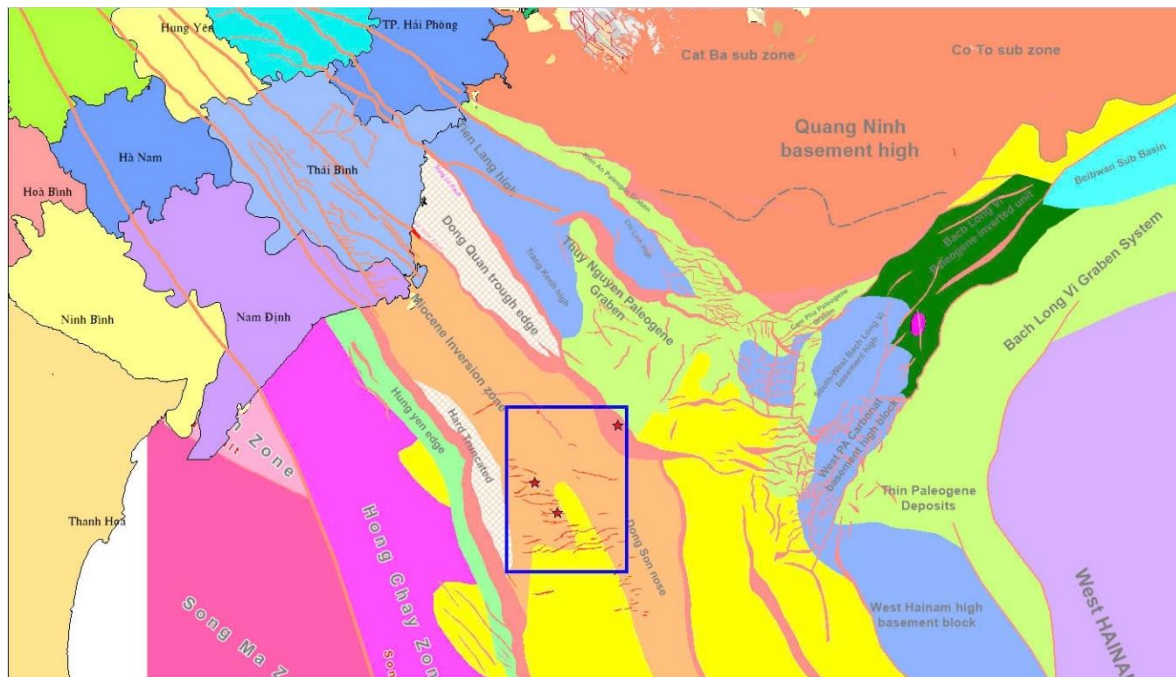


Fig. 1. Location of study area (bounded by blue line) and three wells (red stars) (Nguyen Manh Huyen & Ho Dac Hoai, 2007).

Data was collected from three exploration wells, including wireline logs such as CAL, DT, GR, NPFI, RHOB, and others, along with interpretation results like facies, lithology, porosity, and permeability and others (Table 1). Traditional interpretation methods have identified nine hydrocarbon reservoirs labeled 0 to 8 (Table 2).

Tab. 1. Data status of wells A, B, C

Data	% miss data		
	Well-A	Well-B	Well-C
TVDSS	0	0	0
CAL	0	0	22.87
DT	100	47.73	0
GR	0	0.73	0
NPFI	100	0.73	0
RHOB	0	0.73	22.87
RLA1	0.12	0	0
RLA2	0	0	22.87
RLA3	0	0	0
RLA4	0	0	22.87
RLA5	0	0	0
RT	0	0	22.87
RXOZ	0	0.32	22.87

Data	% miss data		
	Well-A	Well-B	Well-C
SP	0.44	0.82	22.87
TNPH	0	0	0
PEF	0	0	22.87
Facies	10.49	13.32	0
Perm	0	0	0
PHIE	0.16	0.64	0
PHIT	0.16	0.64	0
SW	0.16	0.64	0
VCL	0.16	0.64	0
SWT	0.38	0.64	47.87
Litho	10.51	13.31	0

TVDS - True vertical depth subsea, *CAL* - Caliper, *DT* - Interval transit time, *GR* - Gamma ray, *NPHI* - Neutron porosity index, *RHOB* - Density log, *RLA1*, *RLA2*, *RLA3*, *RLA4*, *RLA5* - Resistivity logs array, *RT* - True resistivity, *RZOZ* - Flushed zone resistivity, *SP* - Spontaneous potential log, *TNPH* - Thermal neutron porosity log, *PEF* - Photoelectric factor, *Perm* - Permeability, *PHIE* - Effective porosity, *PHIT* - Total porosity, *SW* - Water saturation, *VCL* - Volume of clay, *SWT* - Total water saturation, *Litho* - Lithology.

Tab. 2. Reservoir parameters of nine hydrocarbon reservoirs identified in 3 wells

Well	Res.	Top (mMD)	Bottom (mMD)	Top (mTVDSS)	Bottom (mTVDSS)	Fluid contact (mTVDSS)	Gross (mTVD)	Net pay (mTVD)	NTG	Av Phi (v/v)	Av Sw (v/v)	Av Vcl (v/v)	Remarks
A	0	2103.1	2119.7	2074	2090.7	GDT=2090.7 PGWC=2093 (by MDT)	16.7	12.19	0.73	0.173	0.543	0.202	DST2
	1	2456.2	2482	2427.2	2453	GDT=2453	25.8	12.5	0.484	0.128	0.618	0.249	DST1
B	2	2105.4	2140.1	2075.3	2110.1	GDT=2110.1 PGWC=2122.5 (by MDT)	34.7	24.38	0.703	0.161	0.534	0.249	DST2
	3	2705.1	2729	2675.1	2699	GDT=2699	23.9	4.9	0.204	0.157	0.613	0.344	DST1
	4	3100.1	3109.3	3070.1	3079.3	GDT=3079.3	9.16	4.27	0.466	0.113	0.642	0.279	
	5	3125.1	3140.5	3095	3110.5	GDT=3110.5	15.51	8.53	0.55	0.1	0.503	0.371	
	6	3178.9	3189.8	3148.8	3159.8	GDT=3159.8	10.9	6.32	0.585	0.122	0.592	0.277	MDT Gas sample
C	7	1737.5	1753	1705.5	1721	PGWC=2720.3 (by MDT)	15.5	9.91	0.639	0.138	0.531	0.217	DST4
	8	1942	1965.2	1910	1933.2	GDT=1933.2 PGWC= 1944 (by MDT)	23.2	18.9	0.811	0.124	0.537	0.21	DST3

Res. = reservoir; *MD* = measured depth; *TVDSS* = true vertical depth sub-sea; *GDT* = gas down to; *PGWC* = perforated gas-water contact; *NTG* = net to gross ratio; *Av Phi* (v/v) = average porosity, expressed as a volume-to-volume ratio; *Av Sw* = average water saturation; *Av Vcl* = average volume of clay; *DST* = drill stem test; *MDT* = modular dynamic tester.

Methodology

Machine learning algorithms play a crucial role in data analysis and predictive modeling, each offering unique strengths and applications. This study employs various machine learning algorithms, including DT, RF, XGB, ANN, and SOM to complete well logging data and predict potential hydrocarbon reservoirs. Decision tree (DT) follows a structured sequence of choices, visually represented as a tree diagram. Random forest (RF) enhances DT by combining multiple trees to improve accuracy and reduce overfitting. Extreme gradient boosting (XGB) is an optimized gradient boosting algorithm that iteratively refines predictions, making it highly efficient for structured data. Artificial neural networks (ANN) simulate the human brain with interconnected layers of neurons, excelling in complex pattern recognition. Meanwhile, self-organizing maps (SOM), an unsupervised learning method, transforms high-dimensional data in a lower-dimensional space while preserving topological relationships, making it particularly useful for clustering and data visualization.

The well data includes logging measurements (resistivity, density, sonic, gamma ray, etc.), rock samples, hydrocarbon samples, well tests, lithology, etc. These data can be in numerical or image form, which are processed and analyzed to provide detailed insights into rock properties, sedimentary environments, reservoir potential, sealing capacity, and more. The well database is critical for evaluating hydrocarbon potential (Liu & Liu, 2022; Ren et al., 2019; Adibifard et al., 2014; Raeesi et al., 2012). However, log data is often incomplete due to various technical reasons (Doan et al., 2019). To address these issues, preprocessing is essential, through the following steps: removing invalid data; reducing noise in the measurement data by selecting appropriate algorithms and sliding window sizes; enhancing supplemental information layers with statistical parameters (standard deviation, slope, first derivative, curvature, difference) to improve AI predictions; selecting relevant feature combinations to predict missing data accurately. After preprocessing, to apply AI, two datasets are prepared: a training set and a testing set. Machine learning is a system designed to learn, so it requires training and testing to determine whether it has learned as intended by the user. Training data is the information used to train an algorithm, consisting of input data and corresponding expected outputs. The test data is the information used to evaluate the performance of the algorithm, consisting only of input data (Smith, 2019).

Next, the machine learning models random forest, RF, decision tree, DT, and extreme gradient boost, XGB are applied to predict missing segments in the raw logs and the lithology interpretation data of well B. This process evaluates and compares the performance of these models, allowing for the selection of the most optimal one. The chosen model is then used to predict missing segments in both the raw logs (e.g., GR, DT, NPHI, RHOB) and interpretation data (e.g., VCL, SW, PERM, Litho) across all three wells. To further validate the reliability of AI predictions, a segment of the GR log was intentionally removed. The remaining logs from the same well were then used to predict the deleted segment. This approach helps assess the model's accuracy by comparing the predicted segment with the original data. Additionally, the correlation between the predicted data and the original data, as well as their correlation with other original data are also analyzed to ensure the reasonableness and consistency of the predictions.

In wells A, B, and C, a total of nine hydrocarbon reservoirs have been identified (Table 2). The data for these nine reservoirs have been processed and predicted (when necessary) as outlined earlier and will serve as input (Figure 2) for various ML algorithms to predict potential locations of hydrocarbon reservoirs across three wells. The input data for ML includes both raw logs (GR, NPHI, RHOB, DT, PEF, SP, RLA1-5, TNPH), interpreted data (Perm, PHIT, PHIE, SW, VCL), and enriched data (standard deviation, difference, first derivative, slope, curvature). Next, using the input database consisting of attributes from the nine known reservoirs, different ML models, including RF, XGB, ANN, and SOM, will be applied to predict potential locations of hydrocarbon reservoirs across the three wells.

Reserver	Well	TVDSS	X	Y	Interval_Ve	Average_Ve	AI	AI_avg	AI_std	AI_dif	AI_slop
7	C	-1704.86	713181	2208976	3277	2292.063	8382.637	8365.672	103.8115	0.11537	314.303
7	C	-1705.01	713181	2208976	362	2291.004	8450.708	8403.801	70.63963	0.0877	178.301
7	C	-1705.16	713181	2208976	362	2291.076	8473.202	8418.77	44.36047	0.02897	21.4518
7	C	-1705.31	713181	2208976	362	2291.149	8419.351	8411.012	54.49407	-0.06935	-52.3437
7	C	-1705.47	713181	2208976	362	2291.222	8367.951	8403.755	50.12725	0.06618	49.19
7	C	-1705.62	713181	2208976	362	2291.295	8343.846	8426.115	94.31226	-0.03103	394.275
7	C	-1705.77	713181	2208976	362	2291.368	8414.425	8523.923	234.6756	0.09084	992.871
7	C	-1705.92	713181	2208976	362	2291.441	8585.004	8728.606	429.8097	0.21951	1704.26
7	C	-1706.07	713181	2208976	362	2291.513	8908.39	9045.514	618.1066	0.41607	2344.54

Fig. 2. Training sample

The table in Figure 2 illustrates a part of the input parameters for machine learning, including the reservoir number (according to Table 2) and well name, true vertical depth subsea, measurement point coordinates, interval velocity (Interval_Ve), average velocity (Average_Ve), AI - acoustic impedance, and enriched data for acoustic impedance (average – AI_avg, standard deviation – AI_std, difference – AI_dif, slope – AI_slop).

Results

By applying the RF, XGB, and ANN models, the predicted reservoirs and their respective prediction and identification rates are shown in Figures 3 - 8. In Figures 3, 5 and 7, the y-axis represents 9 known reservoirs, labeled from 0 to 8, and the x-axis corresponds to the sequential order of data sample points along the logs. The blue dots indicate reservoirs that have been earlier identified, while the orange dots represent reservoirs predicted by the machine learning model. In Figures 4, 6, and 8, the left y-axis represents the percentage accuracy of forecasting and identifying hydrocarbon reservoirs using ML. The right y-axis indicates the percentage of samples from each identified reservoir relative to the total samples of all identified reservoirs. The x-axis corresponds to the nine known reservoirs, labeled from 0 to 8. The blue bars represent the percentage of samples from each reservoir relative to the total samples of identified reservoirs. The orange line shows the forecasting accuracy of reservoirs using ML, while the red line represents the identification accuracy of reservoirs using ML.

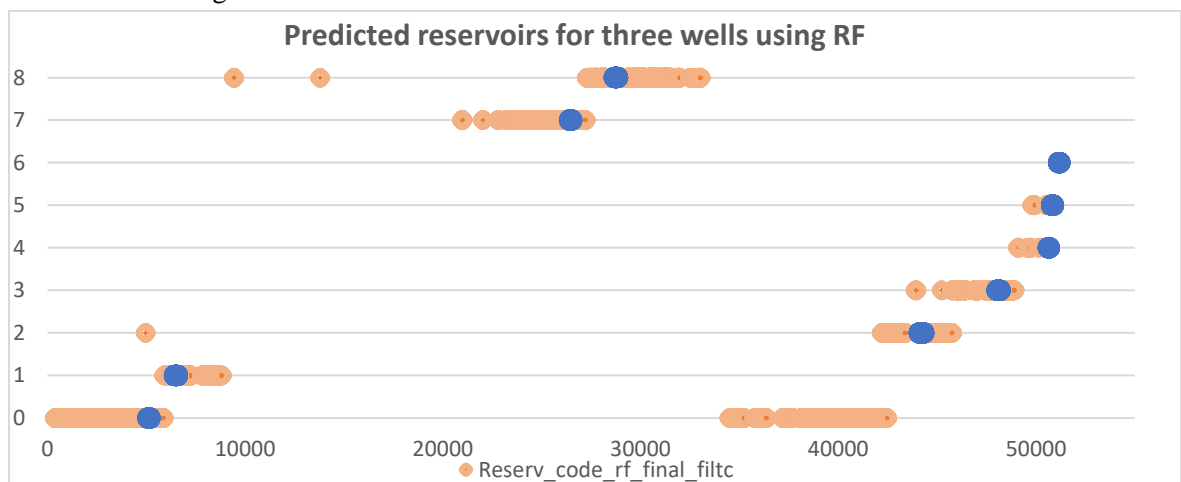


Fig. 3. Predicted reservoirs for three wells using RF

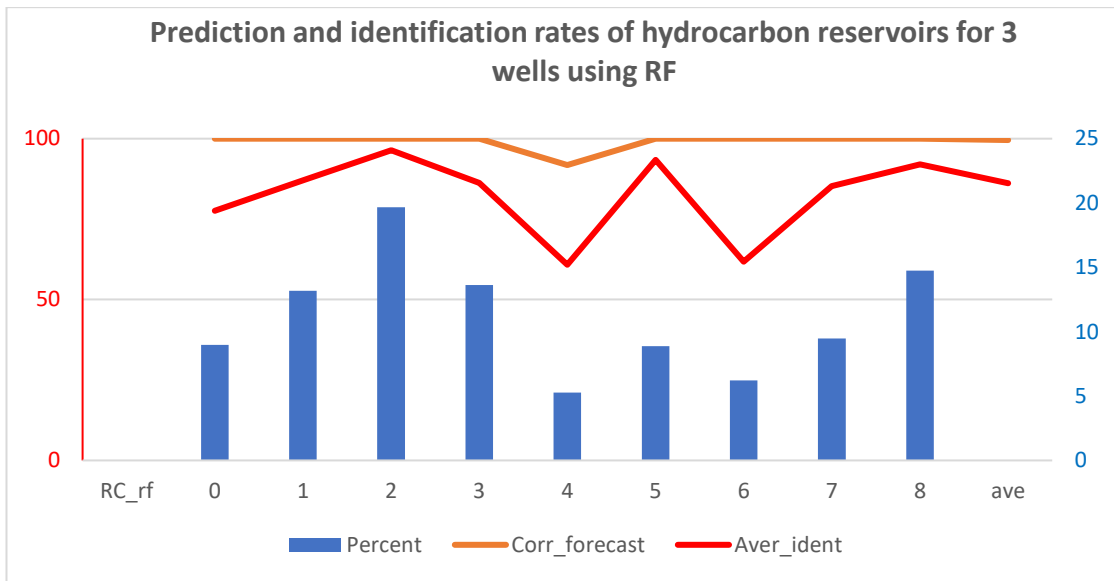


Fig. 4. Prediction and identification rates (RC_rf) of hydrocarbon reservoirs for 3 wells using RF

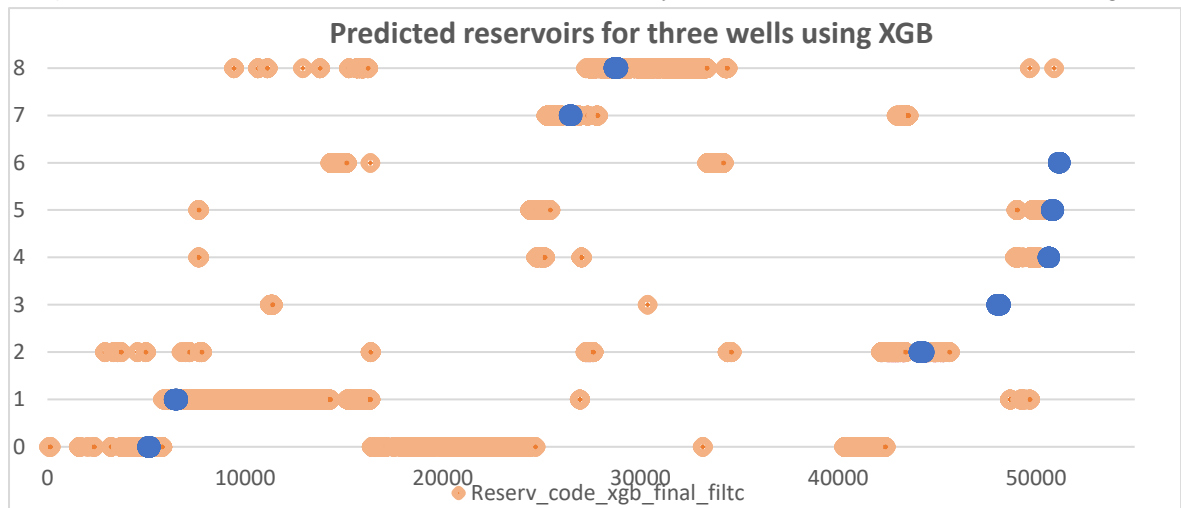


Fig. 5. Predicted reservoirs for three wells using XGB

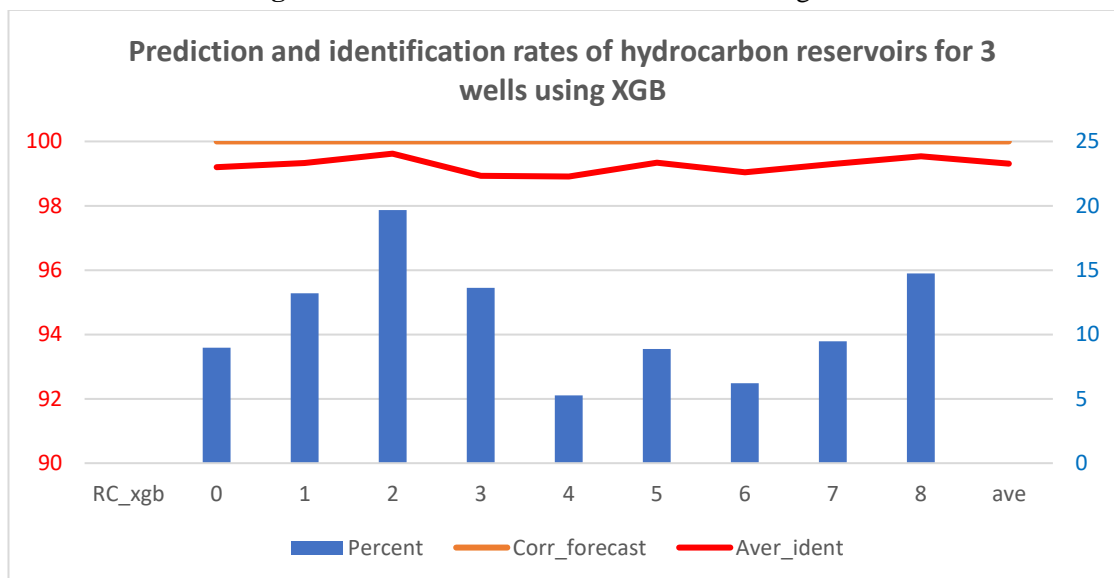


Fig. 6. Prediction and identification rates (RC_xgb) of hydrocarbon reservoirs for 3 wells using XGB

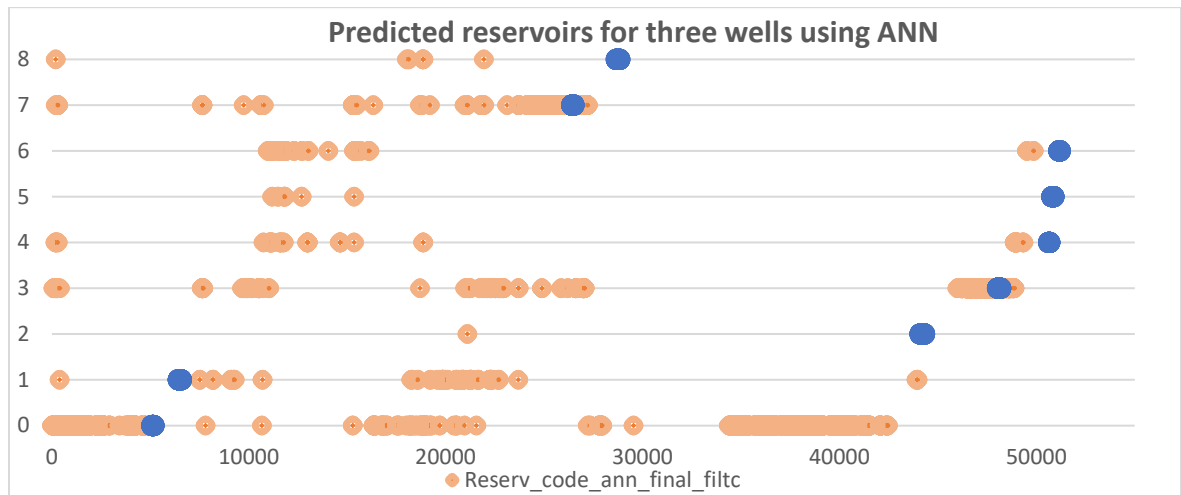


Fig. 7. Predicted reservoirs for three wells using ANN

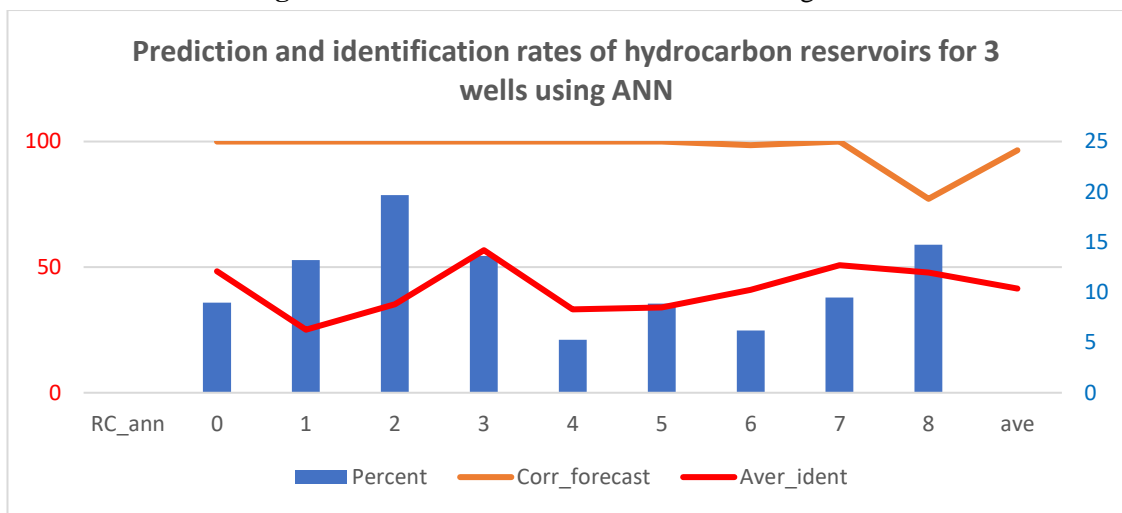


Fig. 8. Prediction and identification rates (RC_ann) of hydrocarbon reservoirs for 3 wells using ANN

The prediction results from the RF, XGB, and ANN models are not satisfactory (Figures 3, 5, 7), showing significant discrepancies in predicted thickness of several identified reservoirs. The relative thickness of the reservoir is represented by the number of measurement samples along the borehole. For example, the predicted thickness of reservoirs 0, 1, 7, and 8 using the RF model shows significant discrepancies compared to the known thicknesses (Figure 3). Similarly, the predicted thickness of the reservoirs using the XGB and ANN models are also unreasonable (Figures 5, 7). The identification rates for hydrocarbon reservoirs using the RF and ANN models are both low, with the ANN model achieving identification rates mostly below 50% (Figures 4, 8).

In contrast, using the Self-Organizing Map (SOM) model, data points with similar attributes are grouped together, helping to identify different sets of attributes (Doan et al., 2021). These sets may represent different groups of reservoirs in the study area. The research team conducted unsupervised reservoir classification using two SOM configurations: a 3x3 SOM, resulted in 9 reservoir attribute sets (SOM 0 – SOM 8) (Table 3), and a 3x5 SOM, generated 15 reservoir attribute sets (SOM 0 – SOM 14) (Table 4). In Table 3 and Table 4, count refers to the total number of data points in each cluster/group, count_% represents the percentage of data points in each group relative to the total dataset, and Rsv_0 to Rsv_8 denote the known hydrocarbon reservoirs, with the corresponding values indicating the percentage of data points within the group that belong to each specific reservoir. With the 3x5 SOM configuration, only 90.26% of the total dataset is generated in 15 reservoir attribute sets (Table 4) because some samples were rejected due to the predicted probability being lower than 75%.

Tab. 3. The results of the unsupervised reservoir classification using a 3x3 SOM.

Cluster	Count	Count_%	Rsv_0	Rsv_1	Rsv_2	Rsv_3	Rsv_4	Rsv_5	Rsv_6	Rsv_7	Rsv_8
SOM 0	171	14.74									100
SOM 1	72	6.21							100		
SOM 2	164	14.14					37.2	62.8			
SOM 3	110	9.48								100	
SOM 5	158	13.62				100					
SOM 6	228	19.66			100						
SOM 7	93	8.02		100							
SOM 8	164	14.14	63.41	36.59							
	1160	100									

Tab. 4. The results of the unsupervised reservoir classification using a 3x5 SOM.

Cluster	Count	Count_%	Rsv_0	Rsv_1	Rsv_2	Rsv_3	Rsv_4	Rsv_5	Rsv_6	Rsv_7	Rsv_8
SOM 0	113	9.74	0	0	0	100	0	0	0	0	0
SOM 1	45	3.88	0	0	0	100	0	0	0	0	0
SOM 2	102	8.79	0	0	0	0	0	0	0	0	100
SOM 3	49	4.22	0	0	0	0	0	0	0	0	100
SOM 4	20	1.72	0	0	0	0	0	0	0	0	100
SOM 5	61	5.26	0	0	0	0	100	0	0	0	0
SOM 6	103	8.88	0	0	0	0	0	100	0	0	0
SOM 7	72	6.21	0	0	0	0	0	0	100	0	0
SOM 8	135	11.64	0	0	100	0	0	0	0	0	0
SOM 9	40	3.45	0	0	0	0	0	0	0	100	0
SOM 10	104	8.97	100	0	0	0	0	0	0	0	0
SOM 11	78	6.72	0	100	0	0	0	0	0	0	0
SOM 12	75	6.47	0	100	0	0	0	0	0	0	0
SOM 13	93	8.02	0	0	100	0	0	0	0	0	0
SOM 14	70	6.03	0	0	0	0	0	0	0	100	0
	1047	90.26									

After the SOM model grouped the data through unsupervised classification to identify reservoir attribute sets as mentioned above, a supervised classification process was employed. It utilizes these reservoir attribute sets as training samples and scanned all three wells to predict hydrocarbon reservoirs by identifying points with attributes similar to those of the trained samples. The prediction results of hydrocarbon reservoirs using SOM are illustrated in Figure 9. The figure displays a representation of several input data such as DT, GR, RHOB, Permeability, and colored columns within the frame indicating the location and relative thickness of hydrocarbon reservoirs. The dark red columns represent known reservoirs; the light red and yellow columns correspond to reservoirs predicted by SOM (3x3) with confidence levels of 95% and 90%, respectively; while the

green and blue columns correspond to reservoirs predicted by SOM (3x5) with confidence levels of 90% and 85%, respectively. The results indicate that, with a 95% confidence level in the 3x3 SOM forecast, no new reservoirs were identified, aside from differences in reservoir thickness compared to the known ones. However, at a 90% confidence level with the 3x3 SOM forecast, four additional potential reservoir locations were predicted (Figure 9). Similarly, the 3x5 SOM forecast did not identify any new reservoirs at a 90% confidence level, but two additional potential reservoirs were predicted at an 85% confidence level.

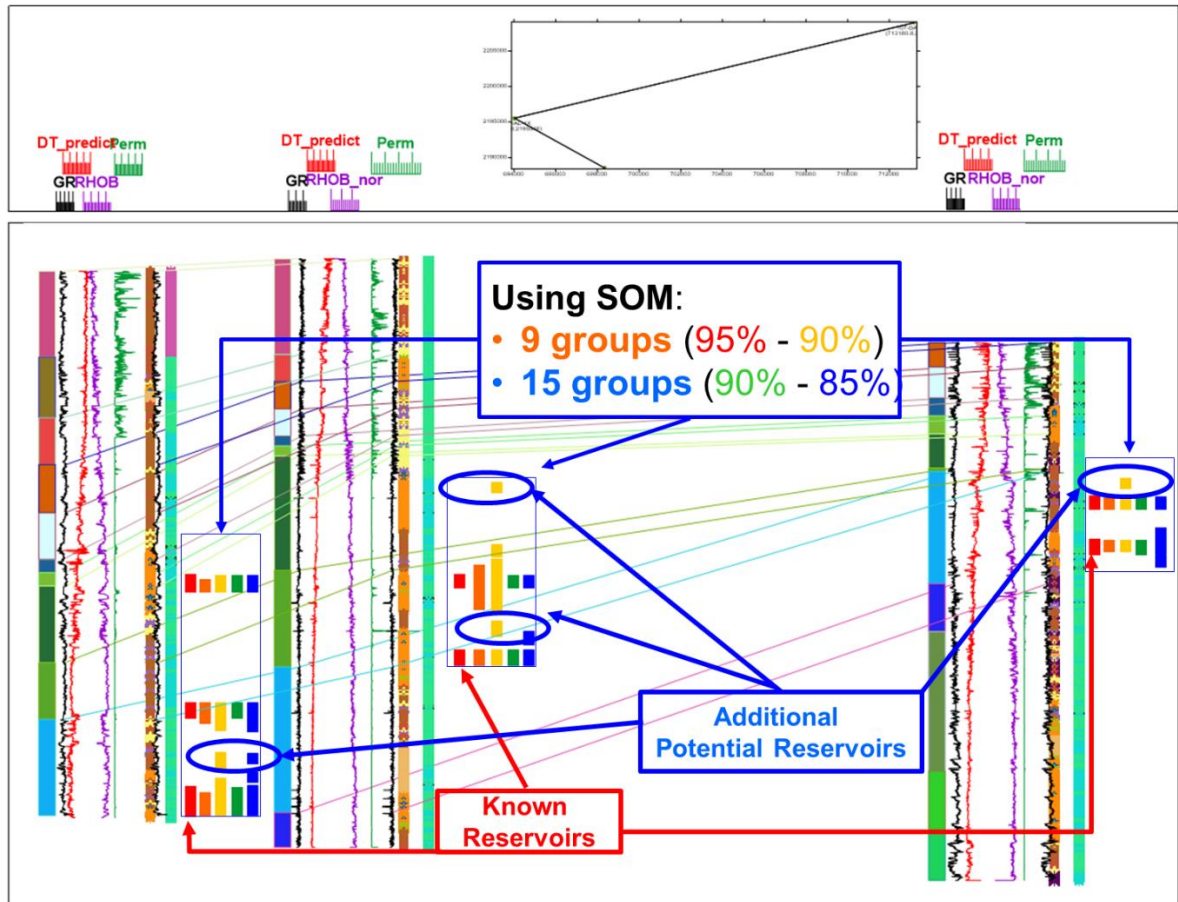


Fig. 9. Predicted hydrocarbon reservoirs for three wells using SOM; explanations in text

Conclusion

In this research, we employed various ML models and calculated their performance to compare their effectiveness. The results indicate that the SOM model provides better reservoir prediction results compared to other ML models like RF, XGB, and ANN. The SOM model successfully identified all nine reservoirs that were traditionally detected. However, the thickness of these reservoirs predicted by the two methods does not match 100% exactly (Figure 9). SOM also predicted additional potential reservoirs, which need to be validated and evaluated by experts. Validated potential hydrocarbon zone of these wells can increase the prospect of oil accumulation in different compartments of the X oil field.

Acknowledgments

This work is funded by PetroVietnam University under Project “Application of machine learning for predicting potential hydrocarbon zones in wells – X field, northern Red River Basin”. The authors thank PetroVietnam University and the anonymous reviewers for making the paper more perfect.

Literature – References

1. Adibifard M., Tabatabaei-Nejad S.A.R., Khodapanah E., 2014. Artificial Neural Network (ANN) to estimate reservoir parameters in Naturally Fractured Reservoirs using well test data. *Journal of Petroleum Science and Engineering*, 122, 585-594.
2. Amir G., Mohamed H., Jebrael G., Hesam G., Hossein B., James B., Karrar A. A., 2022. Application of machine learning techniques for identifying productive zones in unconventional reservoir. *International Journal of Intelligent Networks* 3 (2022) 87–101.
3. Bui T. L., 2015. Evaluation of oil and gas potential of blocks A and B Song Hong basin and suggested exploration plan. *Science and Technology Development Journal*, 18(4), 12-31. <https://doi.org/https://doi.org/10.32508/stdj.v18i4.925>
4. Dau N. T., Huong, H. T., Nga, L. H., Thang, P. V., Nytoft, H., Petersen, H. I., Nielsen, L. H., Bojesen-Koefoed, J., & Abatzis, I., 2014. Characterisation and distribution of hydrocarbon in Song Hong basin. *Petrovietnam Journal*, 10, 45-51. Retrieved from <https://pvj.com.vn/index.php/TCDK/article/view/601>
5. Doan N. S., et al., 2019. Application of Artificial Intelligence System integrating petroleum geological data to assess hydrocarbon prospects. National research project, Ministry of Science and Technology. No. KC-4.0-01/19-25.
6. Doan N. S., Phan D. N., Nguyen T. C., Do T. T. L., Nguyen Q. K., 2021. Application of GIS - AI for detecting geological structure and perspective zoning of mineral in VanYen area. 4th Asia Pacific Meeting on Near Surface Geoscience & Engineering.
7. Dong Y., Y. Zhang, F. Liu, X. Cheng, 2021. Reservoir production prediction model based on a stacked LSTM network and transfer learning. *ACS Omega* 6 (50):34700–11. doi:10.1021/acsomega.1c05132.
8. Fredrick O. O., Pius A. E., 2021. Machine learning approach to hydrocarbon zone prediction from seismic attributes over “GEM” field, Niger-Delta, Nigeria. *International Journal of Advanced Geosciences*, 9 (2) (2021) 88-98.
9. Liu J.-J., Liu J.-C., 2022. Integrating deep learning and logging data analytics for lithofacies classification and 3D modeling of tight sandstone reservoirs. *Geoscience Frontiers* 13.1 (1):101311. doi:10.1016/j.gsf.2021.101311.
10. Nguyễn Hiệp, Trần Văn Trị, Vũ Văn Minh, 2007, Hoạt động điều tra địa chất, thăm dò và khai thác dầu khí. Địa chất và tài nguyên dầu khí Việt Nam (Nguyễn Hiệp chủ biên). Science and Technics Publishing House, 15-40 (in Vietnamese).
11. Nguyễn Mạnh Huyền and Hồ Đắc Hoài, 2007. Bể trầm tích Sông Hồng và tài nguyên dầu khí. Trong Địa chất và tài nguyên dầu khí Việt Nam (Nguyễn Hiệp chủ biên). Science and Technics Publishing House, 187-240 (in Vietnamese).
12. Raeesi M., A. Moradzadeh, F.D. Ardejani, M. Rahimi, 2012. Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks *J. Petrol. Sci. Eng.*, 82, 151-165.
13. Ren X., J. Hou, S. Song, Y. Liu, D. Chen, X. Wang, L. Dou, 2019. Lithology identification using well logs: a method by integrating artificial neural networks and sedimentary patterns *J. Petrol. Sci. Eng.*, 182.
14. Smith D., 2019. What is AI Training Data?. *Lion Bridge*. <https://lionbridge.ai/articles/what-is-ai-training-data/>
15. Ta V. C., Hoang T. L., Doan N. S., Nguyen V. T., Nguyen D. N., Pham T. T. T. & Nguyen D. N., 2023. Tabnet efficiency for facies classification and learning feature embedding from well log data. *Petroleum Science and Technology*, DOI: 10.1080/10916466.2023.2223623.