

# Improving the classification of facies quality in tight sands by petrophysical logs

Yousef Asgari Nezhad <sup>a</sup>, Ali Moradzadeh <sup>a,\*</sup>

<sup>a</sup> School of Mining Engineering, College of Engineering, University of Tehran, Tehran, Iran

## Article History:

Received: 18 November 2020,

Revised: 01 January 2021,

Accepted: 08 January 2021.

## ABSTRACT

As conventional hydrocarbon reserves are running out, attention is now being paid to unconventional hydrocarbon resources and reserves such as tight sands and hydrocarbon shales for future energy supplies. To achieve this, the identification of tight sand facies is based on zones containing mature hydrocarbons in priority. Organic geochemical methods are the commonest methods to evaluate the quality of these reservoirs. In this study, using a deep learning approach and using petrophysical logs, a suitable classification model for facies quality is presented. Moreover, the proposed method has been compared with two common methods: multilinear regression and multilayer perceptron neural network. The results indicated that the accuracy of facies classification using these three methods is about 63%, 71%, and 84% for linear multilinear regression, perceptron multilayer neural network, and deep learning, respectively. Finally, the accuracy of the deep learning networks was optimized using two gravitational search and whale optimization algorithms. It has been shown that the accuracy of deep learning was increased from 84% to 87% and 90.5% using the gravitational search algorithms and whale algorithms, respectively.

**Keywords:** *Facies quality, Deep learning, optimization algorithm, tight sands, classification*

## Nomenclature

**CALI:** Caliper

**DT:** Interval Transit Time

**GR:** Gamma Ray

**LLD:** LathroLog Deep

**LLS:** LathroLog Short

**NPHI:** Neutron Porosity

**RHOB:** Bulk Density

**SP:** Spontaneous Potential

**TOC:** Total Organic Carbon

**T<sub>max</sub>:** Maximum of Temperature

**BQ:** Bad Quality

**GQ:** Good Quality

**MLR:** Multilinear Regression

**MLP-NN:** Multi-Layer Perceptron Neural Network

**RNN:** Recurrent Neural Network

**GSA:** Gravitational Search Algorithms

**WOA:** Whale Optimization Algorithms

**ROC:** Receiver Operating Characteristic

**LSTM:** Long Short - Term Memory

## 1. Introduction

Identification of good quality facies is an important factor in the drilling and production of tight hydrocarbon sands. Some of the important parameters for the identification of hydrocarbon facies are total organic carbon (TOC) and maturity (T<sub>max</sub> or R<sub>0</sub>) [1]. Instrumental analysis methods are the most common tools to study these parameters. But due to the high cost of these methods, the use of

petrophysical logs, which is one of the most common data available in any well, has been considered. Also using these logs is a very beneficial idea to predict organic matter and maturity [2-9]. After using simple fitting between logs and geochemical parameters, simple and multilinear regression has been used for modeling of TOC and maturity (T<sub>max</sub> or vitrinite index) [10-15].

Due complexity relation between geochemical variables and geophysical parameters (such as logs and seismic), intelligent methods were developed. Some researchers used types of neural networks such as Back Propagation neural network [16-18] and Radial Basic Function [19], and the other combined classic ( $\Delta\log R$ ) and intelligent (neural network) methods for prediction of TOC [20]. Also, support vector machines have been used by many researchers to modeling of organic geochemical parameters [21, 22]. Asgari Nezhad et al (2018) have been modeled TOC by neural network and compared it and geostatistics technique [23]. Besides, deep learning techniques are used for modeling TOC by geophysical data [24-26]. Recently Wang et al, (2019) have improved neural networks to estimate TOC, S<sub>1</sub>, and S<sub>2</sub> by a convolutional neural network (CNN) [27].

Based on the above overview, it is necessary to use an algorithm for studying the complexity of the relationship between quality facies and petrophysical logs. So in this study, it has been used recurrent neural network (RNN) as a deep learning technique (compared to multilinear regression and Multi-Layer Perceptron) and petrophysical logs to categorize the quality of tight sand facies. White Hills-1 is well case study in this manuscript. Finally, two optimization algorithms are used to increase the accuracy of the classification.

## 2. Geology

The Canning basin is one of the most potential hydrocarbon basins

\* Corresponding author. Tel: +98-2161114565, E-mail address: [a\\_moradzadeh@ut.ac.ir](mailto:a_moradzadeh@ut.ac.ir) (A. Moradzadeh).

in Western Australia (Fig. 1). One of the wells in this basin where oil and gas evidence was recorded is the White Hills-1. This evidence was mainly found in the Luluigui formations. Some source rock potential has been indicated in fine-grained clastic rocks from the Anderson, Fairfield, and parts of the Luluigui formations (Fig. 2), but vitrinite and geochemical data indicate a poor potential for liquid hydrocarbons less than 2500m [28].

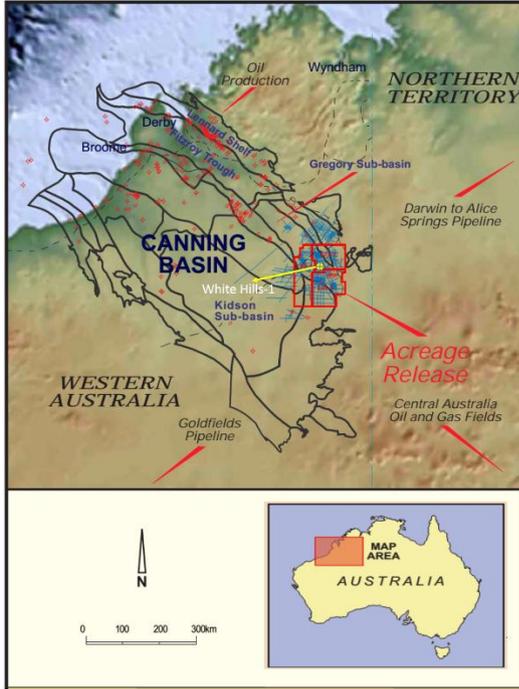


Fig. 1. Location map of the case study [28].

### 3. Methodology

#### 3.1. Multilinear Regression

To get the relationship between some variables such as petrophysical logs and the other variable (e.g, TOC and Tmax) multilinear regression (MLR) is introduced. When the sum of the squares of the differences between the actual values and the modeled values is the least, the model is appropriate [30]. That means MLR is based on least squares.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \quad (1)$$

#### 3.2. Neural Network

Neural networks are used to determine complex nonlinear relationships. These networks are made up of a set of layers (input, middle, and output) that are connected by synapses (Fig. 3).

Here a Multi-layer Perceptron (MLP) is used as an artificial neural network classifier. MLP classifier learns a function  $f(x): R^m \xrightarrow{MLP} R^o$  by training on a dataset by 'm' input and 'o' output [31, 32]. Also, due to the complexity of the relationship between input and output, the hyperbolic tangent function is used. The conjugate gradient method has also been used to optimize the network.

#### 3.3. Deep learning

Deep learning is a subset of the learning machine that enables computers to solve more complex problems. Deep learning techniques are mainly based on neural networks. This network has complexities in layers, neurons, and how they are connected. The general categorization of deep learning is as follows:

- Deep networks for supervisory learning
- Deep networks for unsupervised learning

- Deep hybrid networks

In supervised learning, the categories are defined at the beginning and each training data is assigned to specific categories, and it is said that there is an observer who provides information to the learner in addition to the training data during the training [33].

Recurrent neural network (RNN) is one of the types of neural networks that can be used with deep learning techniques. It was created in 1980, but only in the last few years, it has been widely used. Because there is a significant improvement in computing power. These types of neural networks are particularly useful for processing serial or comet data, in which each neuron or processing unit can maintain an internal state or memory to preserve information related to the previous input. RNNs are so named recurrent because the calculations in each layer affect the output of the next layer. This means that the network has a memory that can store information about the data being studied. [34, 35]. Fig. 4 shows the schematic structure of the RNN.

Like section 3.2, it uses an activation function in the form of a hyperbolic tangent in the hidden layer, and a binary code for the output layer as a quality label. The conjugate gradient method is used to optimize the network.

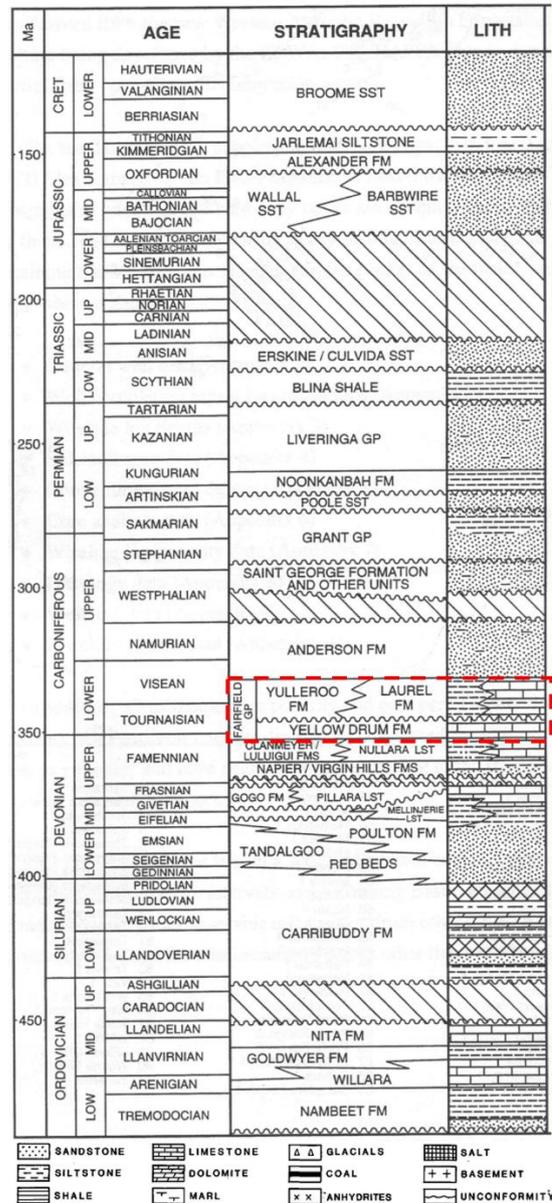


Fig. 2. Generalized stratigraphy of the Canning Basin [29].

### Input Layer Hidden Layer Output Layer

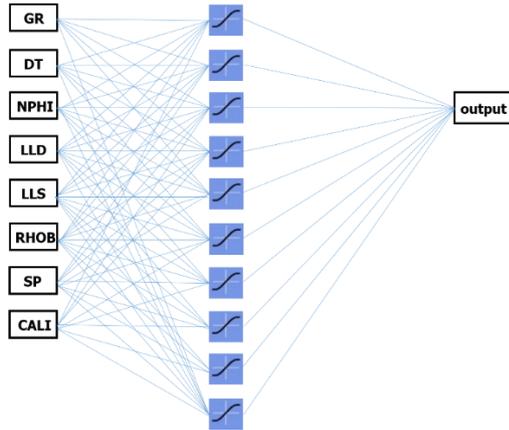


Fig. 3. MLP structure that used in the manuscript.

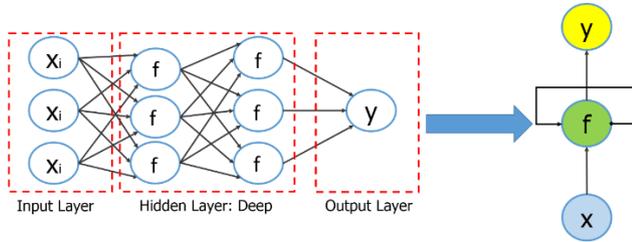


Fig. 4. RNN schematic structure.

### 3.4. Gravitational Search Algorithm of optimization (GSA)

The basis of GSA's performance as a heuristic algorithm is Newton's law of gravity and motion. [36]. In this method, each particle uses gravitational force to attract other parts. If  $X_i$  is the position of the  $i$ th agent in the search space,  $n$  is the number of agents and  $D$  is the dimension of search space in (Eq 2):

$$X_i = (X_i^1, \dots, X_i^D), i = 1, 2, \dots, n \quad (2)$$

The steps of GSA are as follows [37].

1. In the first step: generate random particle position in search space.
2. For the second step, a loop is first designed to perform the next calculation ( $t = 1$  to  $T$ ) as follow:

2.1. The objective function evaluates the fitness of each particle. The inertial mass  $M_{ii}$  is computed as follows:

$$m_i(t) = \frac{fit(X_i(t)) - worst(t)}{best(t) - worst(t)} \quad (3)$$

$$M_{ii}(t) = \frac{m_i(t)}{\sum_{j=1}^n m_j(t)} \quad (4)$$

When a force is applied, the amount of particle resistance to changing its motion states is  $M_{ii}$ . The  $best(t)$  and the  $worst(t)$  are the best and the worst fitness values of particles respectively, ( $t$  stands for  $t^{\text{th}}$  iteration) (Eq 5 and 6).

$$best(t) = \min fit_j(t), j = 1, 2, \dots, n \quad (5)$$

$$worst(t) = \max fit_j(t), j = 1, 2, \dots, n \quad (6)$$

2.2. update  $best(t)$ ,  $worst(t)$  and  $M_{ii}(t)$  for  $i = 1, \dots, n$  by means of Eqs. (3), (5) and (6) respectively.

2.3. Compute the total force in different directions based on Eq. 7:

$$F_i^d(t) = \sum_{j \in k, j \neq i}^{n} rand_j F_{ij}^d(t) \quad (7)$$

2.4. Compute the acceleration through Eq. (8) and the velocity by (9).

$$a_i^d(t) = \frac{F_i^d(t)}{M_{ii}(t)} \quad (8)$$

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \Delta t \quad (9)$$

2.5. Update the position of agents by (10).

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \Delta t \quad (10)$$

3. after the end of the loop, Supply the best agent position

### 3.5. Whale optimization algorithm (WOA)

The whale optimization algorithm (WOA) is one of the heuristic optimization procedures that was first presented by [38]. The basis of the WOA method is how to hunt the humpback whale [39]. Humpback whales create a curtain surrounded by bubbles to hunt small fish. In this algorithm, the position of the whale relative to the prey is modeled by the following equation [38]:

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot |\vec{C} \cdot (\vec{X}^*(t) - \vec{X}(t))| \quad (11)$$

where

$t$ : number of iterations

$\vec{X}$ : the situation vectors of the whale

$\vec{X}^*$ : the situation vector of the best solution

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a}$$

$\vec{C} = 2 \cdot \vec{r}$  is a coefficient vector,

$\vec{r}$  is a random vector  $\in [0,1]$ .

The summary of the whale algorithm's workflow is as follows in Fig. 5:

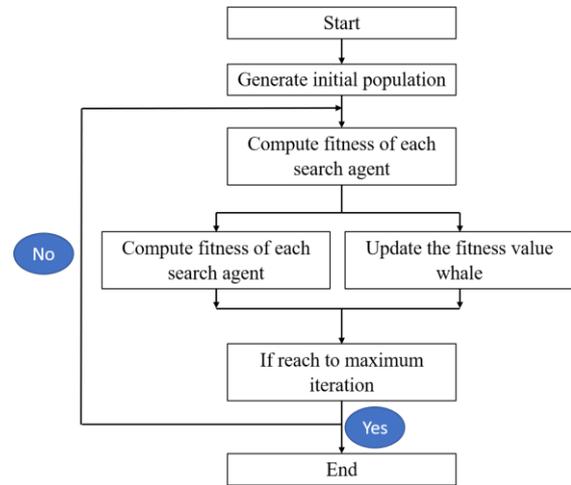


Fig. 5. Workflow of Whale optimization algorithm [38].

## 4. Results of classification

In this section, the classification has been done by three methods (MLR, MLP-NN, and RNN) by using 137 geochemical data (TOC and Tmax) and a set of petrophysical logs (GR, CALI, DT, NPFI, RHOB, LLD, LLS, and SP) for White Hills-1 borehole in western Australia.

In this study, good quality facies are defined as a unit having adequate organic matter via suitable maturity (based on [40, 41]). So, the quality of gas shale facies based on the results of drilling reports are divided into two parts of good and bad quality zones respectively (GQ and BQ) that Good quality zones have TOC > 1 and Tmax > 440. Finally, the outputs of these methods are shown in the confusion matrix. Since there are only two codes in this study, the size of this matrix will be 2 x 2. The accuracy of the proposed methods is calculated by dividing the sum of the principal diameter elements by the total amount of data.

### 4.1. Classification by MLR

The quality classification of hydrocarbon facies is performed in two parts using multilinear regression. The MLR algorithm first has to estimate Tmax, TOC values and then classify them based on the production report.

In the first part, based on the independent variables (which are petrophysical logs), the TOC and Tmax equations are obtained (based on Equation 1), which are expressed as Equations 12 and 13 for TOC and

Tmax, respectively. It should be noted that these models were established for the dataset used here.

$$TOC = 0.7 DT + 0.13 SP + 0.12 CALI - 0.26 GR + 0.73 RHOB + 0.18 NPFI + 0.97 LLD - 0.12 LLS \quad (12)$$

$$T_{MAX} = -3.62 DT + 100.03 SP + 192.37 CALI + 68.13 GR + 371.64 RHOB + 105.81 NPFI + 10.97 LLD + 14.36 LLS \quad (13)$$

Now by applying the thresholds mentioned above ( $TOC > 1$  and  $T_{max} > 440$ ), the facies studied are classified. Based on the classification matrix, the accuracy of this study is 63% (Table 1). The graphical results of this classification are presented in Figure 6. As can be seen in Table 1 and Figure 6, in this method, 39% of good quality data is incorrectly classified as poor quality and 35% of bad quality of real data is incorrectly categorized as good quality. It can be seen from Fig. 6 that the proposed model shows a significant upward shift in the model in addition to not modeling the complexities of the real data (thin and interlayers).

Table 1. Confusion matrix for the result of MLR.

		real	
		GQ	BQ
Classified	GQ	55	34
	BQ	16	32
Accuracy		% 63.5	

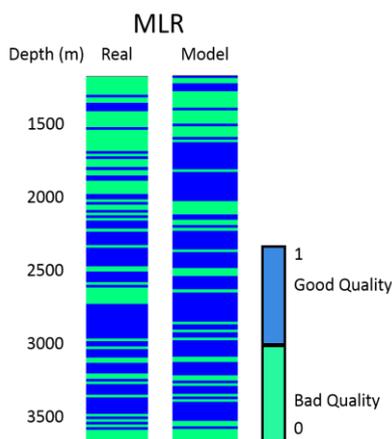


Fig. 6. The result of classification by MLR (Right column) and real data (left column).

#### 4.2. Classification by MLP-NN

To build the model, a set of 137 data (8 logs) from White Hills-1 were employed. After the input data selection, the data were normalized between 0 and 1 based on equation (14).

$$X_{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (14)$$

Now the data have been divided into three parts at a ratio of %70, %15, and %15 to training, testing, and validation, respectively. After running the network, to check the MLP network results, first, check the network performance and the receiver operating characteristic (ROC) curve. Figure 7 shows the ROC curve of the MLP neural network. Given that the graph of both classes is close to one (top left), the network performance is good. The area under the curve, which is an approximate criterion of network performance, is about 0.8, which confirms the proper performance of the algorithm. The confusion matrix is also used to check the accuracy, which in table 2 shows 71% accuracy for this algorithm.

The graphical results of the MLP-NN classifier (Fig. 8) show that the network performance is good but not very accurate at high variations (at 1500 to 2500 m). In general, this algorithm has a good ability to detect the reservoir zone (depth is more than 2700 m) but it is not properly identified interlayers.

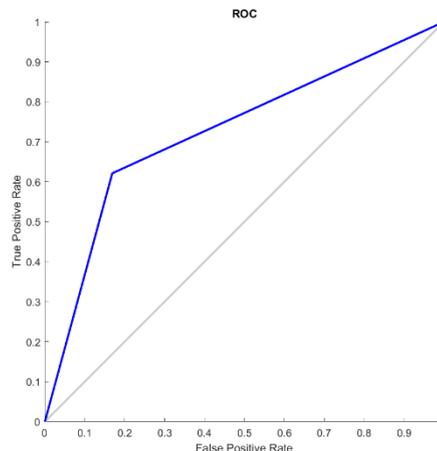


Fig. 7. ROC curve of MLP-NN for two class.

Table 2. Confusion matrix for the result of MLP-NN.

		real	
		GQ	BQ
Classified	GQ	60	29
	BQ	11	37
Accuracy		%70.8	

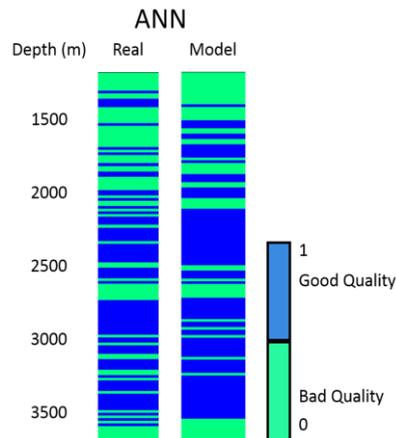


Fig. 8. Comparing the result of Classification by ANN (Right column) to real data (left column).

#### 4.3. Classification by deep learning

The facies quality is classified using a recurrent neural network as one of the deep learning techniques. The designed network has an input layer of 8 neurons (normalized petrophysical logs), two hidden layers each containing 10 neurons, and an output layer containing binary facies codes. The area under the ROC curve for the RNN is approximately 0.9, indicating the good ability of the proposed method in this study (Fig. 9). In table 3 it is explained that the accuracy of this classifier is about 84%. Also, the results of the model and the real data are graphically compared. In Fig. 10 it can be seen that the reservoir zone network is correctly identified at a depth of more than 2700 m using the RNN and is well-recognized interlayers in the reservoir.

#### 4.4. Classification by optimized deep learning by GSA

In this section, it is tried to optimize the RNN used in the previous section by the GSA. In other words, all network parameters, including input, output, hidden layers, and transfer function are the same as before, and the only change in the algorithm is the optimization of synapse weights. The algorithm works according to the process mentioned in Section 3.4, and the results are shown in Figures 10 and 11.

The deviation of the ROC curve from the 45-degree line and the area below the graph, which is greater than 0.92, indicates that the network performance has improved (Fig. 10). A noteworthy point in Figure 11 and Table 4 is that the improvement in network accuracy (3% increase) is related to the improvement of the accuracy of classification of BQ facies (at 2300 m) and GQ facies (at 2200 m), which shows that by using this optimization algorithm, most of the existing complexities were modeled by the RNN.

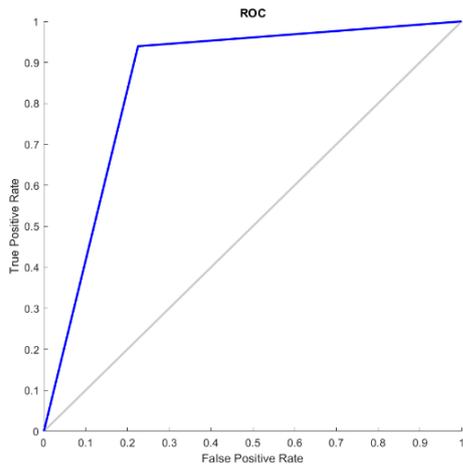


Fig. 9. ROC curve of RNN for two classes.

Table 3. Confusion matrix for the result of deep learning.

		real	
		GQ	BQ
Classified	GQ	61	12
	BQ	10	54
Accuracy		%83.9	

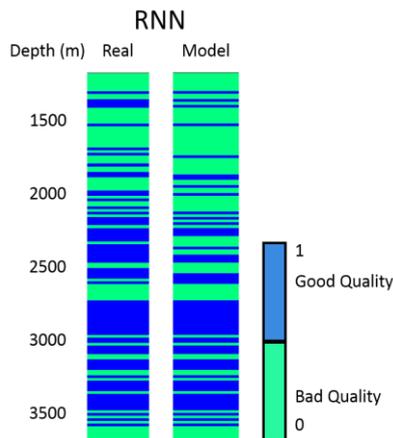


Fig. 10. The graphical result of the deep learning model (Right column) and real data (left column).

#### 4.5. Classification by optimized deep learning by WOA

In this section, as in section 3.4, the goal is to optimize the recurrent neural network, but this time by the WOA algorithm (based on the workflow mentioned in section 3.5). The ROC curve has a better trend than the GSA algorithm and the network performance has improved according to the area below the curve (Figs 12). The accuracy of the RNN has been shown in Table 5. Figure 13 also shows that the ability of the proposed model to classify tight sands facies has improved. That is, the network can model more complexities by using WOA.

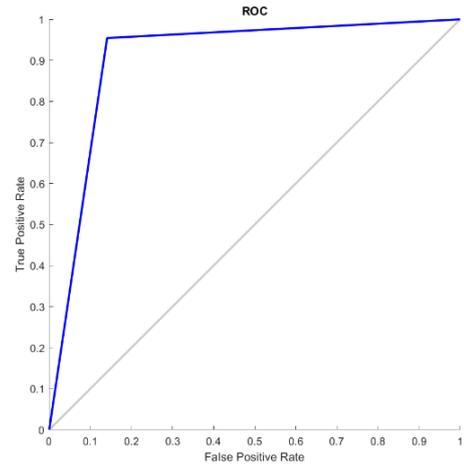


Fig. 10. ROC curve of optimized RNN by GSA for two classes.

Table 4. Confusion matrix for the result of optimized deep learning by GSA.

		real	
		GQ	BQ
Classified	GQ	62	9
	BQ	9	57
Accuracy		%86.9	

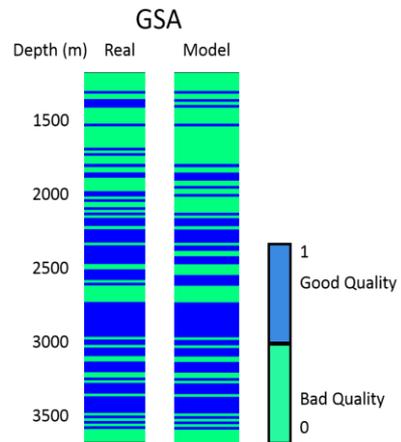


Fig. 11. The graphical result of optimized deep learning by GSA (Right column) and real data (left column).

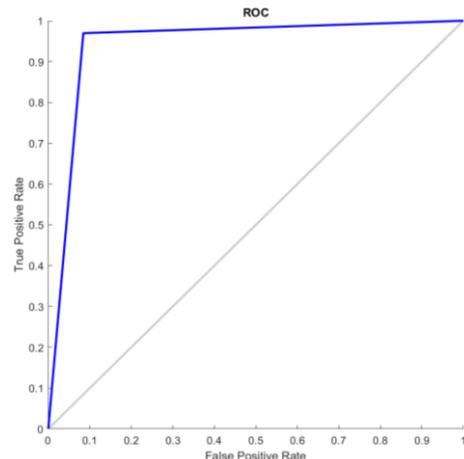


Fig. 12. ROC curve of optimized RNN by WOA for two classes.

Table 5. Confusion matrix for the result of optimized deep learning by WOA.

		real	
		GQ	BQ
Classified	GQ	65	7
	BQ	6	59
Accuracy		%90.5	

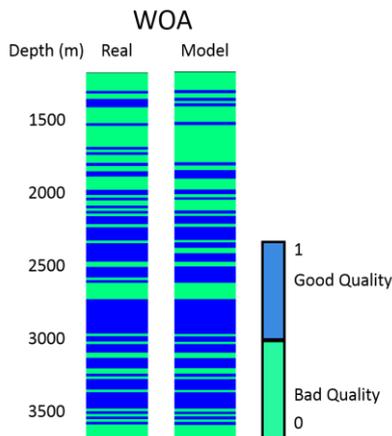


Fig. 13. The graphical result of optimized deep learning by WOA (Right column) and real data (left column).

## 5. Discussion

After classifying the quality of hydrocarbon-tight sand facies by three types of methods, the results are compared. According to Figs 5, 7, and 9, it is clear that the accuracy of the recurrent neural network more than other methods. In general, the reason for this difference inaccuracy is the accurate detection of facies boundaries using deep learning.

The MLR method is incapable of detecting these boundaries and has not provided a suitable model in the facies boundary. In the reservoir zone (more than 2700 m depth) due to the continuity of more good quality facies, the accuracy of detecting good quality facies has increased. In other words, this method is highly sensitive to facies changes. This means that if the quality changes at different depths, the accuracy of this method is reduced.

In MLP classification, accuracy is greater than MLR. In MLP classification, accuracy is greater than. Separation in the reservoir zone is almost acceptable. But with this classifier, it is not possible to accurately model severe changes. In this method, some of the thin interlayers (for example at depths of 1700m and 2200 m) are ignored and not properly modeled.

In the RNN method, the complexities of the wells are appropriately modeled. The thin layer within the reservoir zone (at 3100 m depth) has been correctly identified. Also, in the non-reservoir zone (depth less than 2700 m), the RNN classifier has been able to detect all facies, but some defects are found at depths of 1300 m and 2300 m.

To solve this problem, we tried to use two optimization algorithms to improve the performance of deep learning. Both algorithms increased network accuracy (3% for GSA and 6% for WOA). So, by using these two algorithms, the existing complexities were modeled, but an error is still seen at 1700 - 1800 meters. In these zones, the quality changes are not high, so this error is not due to the ability of the method to detect high-quality facies changes. More data (such as seismic data, PEF logs, etc.) or long-term memory methods such as Long Short - Term Memory (LSTM) or another optimization algorithm may be needed.

## 6. Conclusion

Classification of facies quality is one of the important prospects of hydrocarbon exploration. This important target is mainly achieved by

geochemical methods. For these studies, obtaining geochemical parameters such as organic carbon content and their maturity is a priority. But because of the cost of these methods, researchers are looking for indirect methods (using petrophysical and geophysical logs) to determine these parameters. In this study, the quality of hydrocarbon facies was classified using petrophysical logs and three methods: multilinear regression, multilayer perceptron neural network, and recurrent neural network. Finally, two optimization algorithms were used to improve RNN results. The results showed that with the deep learning method, severe changes in facies quality can be modeled more accurately (84%) than the other two methods (63.5% for MLR and 71% for MLP). Also optimized RNN increases model accuracy to 87% and 90% by using GSA and WOA respectively. The difference between the accuracy of optimized deep learning and the other methods shows the proper performance of optimized RNN in classifying the quality of hydrocarbon facies.

### Author declaration

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

### Acknowledgments

The authors want to express their sincere appreciation to the "Department of Mines and Petroleum of Government of WA" for their help in providing data and information.

## REFERENCES

- [1] Ding, J., Xiaozhi, C., Xiudi, J., Bin, W., & Jinmiao, Z. (2015). Application of AVF inversion on shale gas reservoir TOC prediction. Paper presented at the 2015 SEG Annual Meeting.
- [2] Passey, Q., Creaney, S., Kulla, J., Moretti, F., & Stroud, J. (1990). A practical model for organic richness from porosity and resistivity logs. *AAPG Bulletin*, 74(12), 1777-1794.
- [3] Lecompte, B., & Hursan, G. (2010). Quantifying source rock maturity from logs: how to get more than TOC from Delta Log R. Paper presented at the SPE Annual Technical Conference and Exhibition.
- [4] Euzen, T., Power, M., Crombez, V., Rohais, S., Petrovic, M., & Carpentier, B. (2014). Lithofacies, Organic Carbon and Petrophysical Evaluation of the Montney and Doig Formations (Western Canada): Contribution of Quantitative Cuttings Analysis and Electrofacies Classification. Paper presented at the CSPG CSEG CWLS Joint Annual Convention, Calgary.
- [5] Yuan, X., Lin, S., Liu, Q., Yao, J., Wang, L., Guo, H., et al. (2015). Lacustrine fine-grained sedimentary features and organic-rich shale distribution pattern: A case study of Chang 7 Member of Triassic Yanchang Formation in Ordos Basin, NW China. *Petroleum Exploration and Development*, 42(1), 34-43.
- [6] Wang, P., Chen, Z., Pang, X., Hu, K., Sun, M., & Chen, X. (2016). Revised models for determining TOC in shale play: Example from Devonian Duvernay shale, Western Canada sedimentary basin. *Marine and Petroleum Geology*, 70, 304-319.
- [7] Zhao, P., Mao, Z., Huang, Z., & Zhang, C. (2016). A new method for estimating total organic carbon content from well logs. *AAPG Bulletin*, 100(8), 1311-1327.
- [8] Zhao, P., Ma, H., Rasouli, V., Liu, W., Cai, J., & Huang, Z. (2017). An improved model for estimating the TOC in shale formations. *Marine and Petroleum Geology*, 83, 174-183.
- [9] Nie, X., Wan, Y., & Bie, F. (2017). Dual-shale-content method for total organic carbon content evaluation from wireline logs in

- organic shale. *Open Geosciences*, 9(1), 133-137.
- [10] Huang, Z., & Williamson, M. A. (1996). Artificial neural network modelling as an aid to source rock characterization. *Marine and Petroleum Geology*, 13(2), 277-290.
- [11] Kamali, M. R., & Mirshady, A. A. (2004). Total organic carbon content determined from well logs using  $\Delta\text{LogR}$  and Neuro-Fuzzy techniques. *Journal of Petroleum Science and Engineering*, 45(3-4), 141-148.
- [12] Kadkhodaie-Ilkhchi, A., Rahimpour-Bonab, H., & Rezaee, M. (2009). A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data: An example from Kangan and Dalan reservoirs in South Pars Gas Field, Iran. *Computers & Geosciences*, 35(3), 459-474.
- [13] Sfidari, E., Kadkhodaie-Ilkhchi, A., & Najjari, S. (2012). Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems. *Journal of Petroleum Science and Engineering*, 86, 190-205.
- [14] Alizadeh, B., Najjari, S., & Kadkhodaie-Ilkhchi, A. (2012). Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: A case study of the South Pars Gas Field, Persian Gulf, Iran. *Computers & Geosciences*, 45, 261-269.
- [15] Alizadeh, B., Maroufi, K., & Heidarifard, M. H. (2018). Estimating source rock parameters using wireline data: An example from Dezful Embayment, South West of Iran. *Journal of Petroleum Science and Engineering*, 167, 857-868 .
- [16] Khoshnoodkia, M., Mohseni, H., Rahmani, O., & Aali, J. (2010). Toc Determination of Gadvan Formation in South Pars Gas Field, Using Artificial Neural Network Technique. Paper presented at the GEO 2010.
- [17] Mahmoud, A. A. A., Elkatatny, S., Mahmoud, M., Abouelresh, M., Abdulaheem, A., & Ali, A. (2017). Determination of the total organic carbon (TOC) based on conventional well logs using artificial neural network. *International Journal of Coal Geology*, 179, 72-80 .
- [18] Zhu, L., Zhang, C., Zhang, C., Wei, Y., Zhou, X., Cheng, Y., et al. (2018). Prediction of total organic carbon content in shale reservoir based on a new integrated hybrid neural network and conventional well logging curves. *Journal of Geophysics and Engineering*, 15(3), 1050-1061.
- [19] Tan, M., Liu, Q., & Zhang, S. (2013). A dynamic adaptive radial basis function approach for total organic carbon content prediction in organic shale. *Geophysics*, 78(6), D445-D459 .
- [20] Amiri Bakhtiar, H., Telmadarreie, A., Shayesteh, M., Heidari Fard, M., Talebi, H., & Shirband, Z. (2011). Estimating total organic carbon content and source rock evaluation, applying  $\Delta\text{logR}$  and neural network methods: Ahwaz and Marun oilfields, SW of Iran. *Petroleum Science and Technology*, 29(16), 1691-1704.
- [21] Wang, P., Peng, S., & He, T. (2018). A novel approach to total organic carbon content prediction in shale gas reservoirs with well logs data, Tonghua Basin, China. *Journal of Natural Gas Science and Engineering*, 55, 1-15 .
- [22] Rui, J., Zhang, H., Zhang, D., Han, F., & Guo, Q. (2019). Total organic carbon content prediction based on a support-vector-regression machine with particle swarm optimization. *Journal of Petroleum Science and Engineering* .
- [23] Asgari Nezhad, Y., Moradzadeh, A., & Kamali, M. R. (2018). A new approach to evaluate Organic Geochemistry Parameters by geostatistical methods: A case study from western Australia. *Journal of Petroleum Science and Engineering*, 169, 813-824 .
- [24] An, P., & Cao, D. (2018). Shale content prediction based on LSTM recurrent neural network. Paper presented at the SEG 2018 Workshop: SEG Maximizing Asset Value Through Artificial Intelligence and Machine Learning, Beijing, China, 17-19 September 2018.
- [25] Zhu, L., Zhang, C., Zhang, C., Zhang, Z., Nie, X., Zhou, X., et al. (2019). Forming a new small sample deep learning model to predict total organic carbon content by combining unsupervised learning with semisupervised learning. *Applied Soft Computing*, 83, 105596.
- [26] Zhu, L., Zhang, C., Zhang, C., Zhang, Z., Zhou, X., Liu, W., et al. (2020). A new and reliable dual model-and data-driven TOC prediction concept: A TOC logging evaluation method using multiple overlapping methods integrated with semi-supervised deep learning. *Journal of Petroleum Science and Engineering*, 106944.
- [27] Wang, K., Pang, X., Zhang, H., Hu, T., Xu, T., Zheng, T., et al. (2019). Organic geochemical and petrophysical characteristics of saline lacustrine shale in the Dongpu Depression, Bohai Bay Basin, China: Implications for E3 hydrocarbon exploration. *Journal of Petroleum Science and Engineering*, 106546 .
- [28] Mobil Oil Australia, 1983. Well completion report White Hills-1, Exploration Permit 134, Canning Basin, Western Australia. Geological Survey of Western Australia, S2086A2.
- [29] Apak, S.N. and Carlsen, G.M., (1996). A Compilation and Review of Data Pertaining to the Hydrocarbon Prospectivity in the Canning Basin: Geological Survey of Western Australia, Record 1996/10
- [30] Weisberg, S. (2005). *Applied linear regression* (Vol. 528): John Wiley & Sons.
- [31] Lashin, A., & El Din, S. S. (2013). Reservoir parameters determination using artificial neural networks: Ras Fanar field, Gulf of Suez, Egypt. *Arabian Journal of Geosciences*, 6(8), 2789-2806.
- [32] Kainthola, A., Singh, P., Verma, D., Singh, R., Sarkar, K., & Singh, T. (2015). Prediction of strength parameters of Himalayan rocks: a statistical and ANFIS approach. *Geotechnical and Geological Engineering*, 33(5), 1255-1278.
- [33] Kacprzyk, J. (2008). *Studies in Computational Intelligence*, Volume 100.
- [34] Funahashi, K.-i., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous-time recurrent neural networks. *Neural networks*, 6(6), 801-806.
- [35] Hermans, M., & Schrauwen, B. (2013). Training and analyzing deep recurrent neural networks. Paper presented at the Advances in neural information processing systems.
- [36] Rashedi, E., Nezamabadi-Pour, H. and Saryazdi, S., (2009). GSA: a gravitational search algorithm. *Information sciences*, 179(13), pp.2232-2248.
- [37] Pelusi, D., Mascella, R., Tallini, L., Nayak, J., Naik, B., & Deng, Y. (2019). Improving exploration and exploitation via a Hyperbolic Gravitational Search Algorithm. *Knowledge-Based Systems*, 105404.
- [38] Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, 95, 51-67.
- [39] Watkins, W. A., & Schevill, W. E. (1979). Aerial observation of feeding behavior in four baleen whales: *Eubalaena glacialis*, *Balaenoptera borealis*, *Megaptera novaeangliae*, and

Balaenoptera physalus. *Journal of Mammalogy*, 60(1), 155-163.

[40] Tissot, B. P., & Welte, D. H. (2013). *Petroleum formation and occurrence*: Springer Science & Business Media.

[41] Suárez-Ruiz, I., Flores, D., Mendonça Filho, J. G., & Hackley, P. C. (2012). Review and update of the applications of organic petrology: Part 1, geological applications. *International Journal of Coal Geology*, 99, 54-112.