The CSI Journal on
Computer Science and Engineering
Vol. 20, No. 1, 2026
Pages 46-52
doi.org/ 10.22034/jcse.2025.557349.1064
Regular Paper

Porosity Evaluation Using Artificial Neural Network, Optimized with GA and PSO

Peyman Torabideh[1], Abbas Khaksar Manshad[2], Mohammad Amin Joghatayi[3], Alireza Zangeneh Var [4]

[1]Chemical Engineering Department, Persian Gulf University of Bushehr, Iran
[2]Department of Petroleum Engineering, Abadan Faculty of Petroleum, Petroleum University of Technology (PUT), Abadan, Iran
[3]Department of Intelligent Systems Engineering and Data Science, Persian Gulf University of Bushehr, Iran
[4]Graduated Association of Sharif University (Technical Manager of Petro Pars Company)

Abstract

The precise evaluation of porosity is fundamental to reservoir characterization and volumetric assessment. While direct measurements from core analysis provide accurate results, they are economically and operationally prohibitive for continuous formation evaluation. This study presents a robust machine learning framework that employs Artificial Neural Networks (ANNs)—Multilayer Perceptron (MLP) and Radial Basis Function (RBF)—to predict porosity from conventional well logs. To maximize predictive accuracy and convergence efficiency, the models are optimized using two metaheuristic algorithms: the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Applied to a dataset from a carbonate reservoir in South-West Iran, the hybrid models (MLP-GA, MLP-PSO, RBF-GA, RBF-PSO) demonstrate a superior performance compared to their non-optimized counterparts. The optimization led to a significant increase in the correlation coefficient (R) and a substantial reduction in the Mean Square Error (MSE) for both vertical and horizontal porosity estimates. This research conclusively establishes that the synergy of ANNs with evolutionary optimizers offers a reliable, cost-effective, and rapid solution for porosity prediction, with strong potential for broader application in petrophysical property estimation.

Keywords: Artificial Neural Network, Porosity, Core Analysis, Reservoir

1. Introduction

Well logging records the physical parameters of the reservoir versus depth, qualitative and quantitative evaluation of reservoir petrophysical properties of the near wellbore reservoir using well logs is one significant phase of reservoir studies. Adequate reservoir data is essential to a proper evaluation and development plan. Porosity, as one important parameter in evaluating the economic value of a reservoir, can either be evaluated through direct measurement in costly core analysis lab work, indirect qualitative and quantitative evaluation through well logs and empirical correlations with their inherent inaccuracies, or recently developed, indirectly optimized through intelligent methods as fuzzy logic, artificial neural network [1] and the other intelligent approaches.
In fact, many problems associated with the cost and inaccuracy of the former methods have been rectified by use of intelligent processing techniques, thereby improving the effectiveness of the reservoir characterization and development plans [2].
Using intelligent approaches to predict properties of reservoirs and petroleum characteristics are an emerging approach [12], [13], [14], [15], [16].
In this paper, we present a model to predict porosity of oil and gas reservoir using neural network optimized by particle swarm optimization and genetic algorithm. Results shows that on the studied dataset, the proposed algorithm shows better performance than the previous solutions.

2. An overview to Artificial Intelligence

2.1. Artificial Neural Network (ANN)

ANN has a 50-year history behind, however, their application in scientific began two decades ago and is progressing at a high pace. ANN have been modeled from the neural system of the animals, mammals and especially human, and consists of units called neuron. ANN has, in fact, been inspired from the complicated structure of the mammals', human's, brain neural system where information is saved and processed to solve problems by a network of millions of interconnected neurons [3]. A neuron consists of three main components namely, dendrite, cell body and axon. Dendrite act as connections which receive information from other neural cells and axons are the exit gate of the cells, the input data enters the neuron as chemicals the cell body converts the information into electrical signals and passes it through dendrite to axon. When the electrical signal at the end of axon is big enough, the cell reverts and electrical discharge throughout the axon and a chemical signal is created in the synapse, the connection between the axon and the dendrite of the next cell [4].

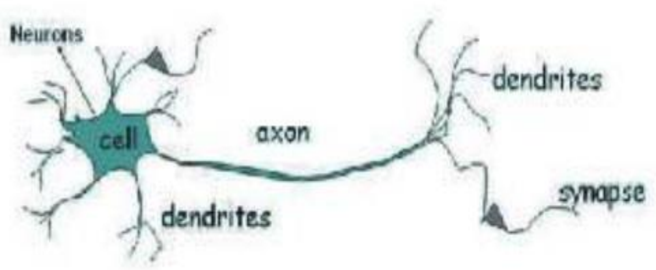

Figure 1. Schematics of a living neural cell

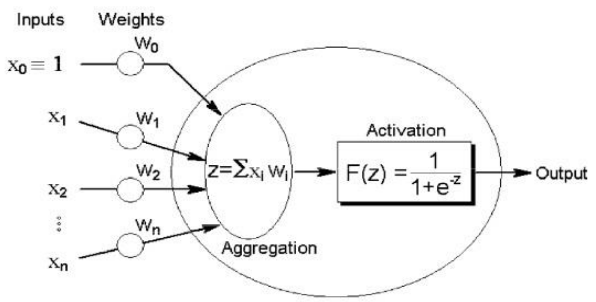

Figure 2. Illustration of the functions and data flow in ANN

ANNs are composed of multiple layers, through the first of which the information enters the network, then comes the intermediate layers (hidden layers) whose number may vary depending upon the nature of the problem, they take the role of processing the data and the connection between input and output. Here the data are analyzed, weighted, processed, coded, etc. and finally passed to the last layer, namely output layer. Hence, one may say the output is a combination of inputs, weights, bias and the hidden layer elements [1], [2].

2.2. Multilayer Perceptron Networks (MLP)

MPLs are of most frequently applied feedforward ANNs in permeability prediction and modeling in oil and gas industry. In MLP, every neuron in every layer is connected to neurons in the preceding layer, such networks are known as completely interconnected layers [5]. Here, the network is first produced with one hidden layer and in case of malfunction the number of layers will be increased, however, the number of layers has to be the smallest possible since the system may overfit if the number of layers is greater than a certain value and it will begin to preserve the data instead of analyzing them, thereby, errors high will rise. To prevent this, cross validation is used [6], [7]. Figure 3 shows the schematic of an MLP with an input layer, two hidden layers and an output layer.

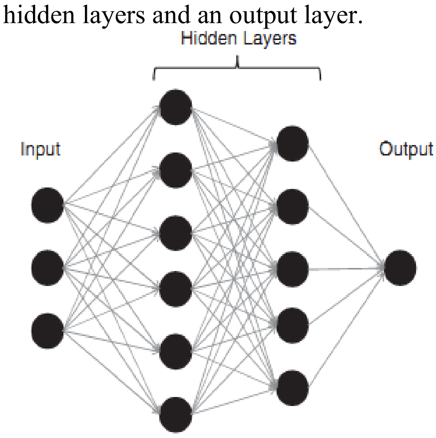

Figure 3. An MLP ANN with an input layer, two hidden layers and an output layer [5]

2.3. Radial basis function neural networks (RBF)

These networks need more neurons compared with standard feedforward networks with error backpropagation, however such networks can be learned in a shorter time than feedforward networks. RBFs serve well when there is a great deal of input available, they are of the feedforward type and their structure is like MLPs [8]. Figure 4 depicts a typical schematic of RFB networks, this network maps an N-dimensional input pattern into a Z-dimensional output pattern using the adjacent layer nodes. The inputs of each neuron are different from the others in RBFs, the inputs to the transfer function are equal to the distance vector between the weights and the inputs are multiplied by the bias.

2.4. Genetic algorithm optimization (GA)

GA, a learning method based on biological evolution, was proposed by John Holland in 1975. This method provides a large set of possible solutions [9], each of which is analyzed with a cross-over function, and then a number of the solutions create new solutions that cause evolution of solutions; accordingly, the process proceeds to the desired solution [10]. This method showed effective since the input parameters were adequately selected.

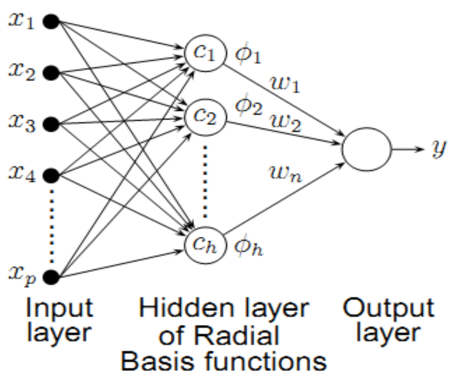

Figure 4. A schematic of an RBF ANN [8]

2.5. Particle swarm optimization (PSO)

PSO is an over exploration algorithm. It has been inspired from birds fight model in 1995 by Kennedy and Eberhart. PSO is one of the most successful algorithms in discrete and continuous optimization where every solution is modeled as a particle of magnitude and correspondence (or similarity). Since it is a method population based and there's no solution combination in it, it usually converges fast. In general, PSO operates based upon the intelligence and interaction of the particles and uses the concept of social interactions to solve optimization problems.

In this article, ANN has been used for optimization with two indexes, correlation coefficient (R) and mean square error (MSE):

a. Correlation coefficient is a dimensionless value ranging between 0 and 1, and is calculated through either of the following equations (1), (2):

$$R = \frac{\sum_{n=1}^{n}[(E_n-(\sum_{n=1}^{n}E_n)/n)(R_n-(\sum_{n=1}^{n}R_n)/n)]}{\sqrt{\sum_{n=1}^{n}(E_n-(\sum_{n=1}^{n}E_n)/n)^2\sum_{n=1}^{n}(R_n-(\sum_{n=1}^{n}R_n)/n)^2}} \quad (1)$$

$$R^2 = \frac{\sum_{n=1}^{n}[(E_n-(\sum_{n=1}^{n}R_n)/n)]^2}{\sum_{n=1}^{n}[(R_n-(\sum_{n=1}^{n}R_n)/n)]^2} \quad (2)$$

b. The following equation denotes MSE (3):

$$\varepsilon(n) = MSE = \frac{\sum_{n=1}^{n}(E_n - R_n)^2}{n} \quad (3)$$

Where R_n is the actual observed value, E_n is the calculated value, n is the amount of data for learning network. The weights have to be varied to decrease the error or increase the R to reach the best fitness [8].

3. Case study

As mentioned above, one of the south-western Iranian fields has been selected for this study. A great number of wells with an acceptable amount of data were available in this field whose most influential collection has been used in this study. four adequately spaced wells were selected from which vertical and horizontal cores were available, and other wells which were too distant from these wells were neglected for this study.

The ANN inputs include the following logs: SGR, CGR, DT, ILD, Density, NPHI, Caliper and Depth. The outputs include the porosities of the vertical and horizontal cores.

3.1. Data analysis an ANN building methods

The log and core data in the four selected wells had good consistency, besides the cores had good quality to be used in ANN. In this work, to the programs were written in different files for network learning analysis convenience. The method and stages of building the ANNs are as follows:

a) The first step is to characterize the effect of the input parameters. overpopulating the input data caused an oversize of the ANN and hence decelerated the learning process. Therefore, networks with different input configurations were used and those inputs of most influence on the efficiency of the results were selected.

b) Before introducing the inputs into the network, preprocessing the data in order to assure that the input parameters receive equal attention from the network contributes to improving the overall efficiency of the network. The input values were mostly large, this reduced the recognition rate of the network, therefore the data were normalized to fall in the interval between 0 and 1 using equation (4).

$$Xn = \frac{(x - MinX)}{(MaxX - MinX)} \quad (4)$$

Where X is the magnitude of the one-dimensional vector of the data, X_n is the normalized value of X.

c) Categorizing the data: Among all the data (petrophysical data, horizontal and vertical core data), 70% were used training data, 15% as validation and 15% as test data in MLP, whereas in RBF 70% of the data were used as training data and the remaining 30% were used to test/verify the network.

d) Network structure design: In order to build the network structure, the number of hidden layers as well as the number of neurons in each layer had to be determined. The number of neurons in the input and output layer were the number of input and output parameters respectively, the number of neurons in the intermediate layers was obtained by trial and error.

e) Learning: To prevent preprocessing, the learning data set were used, the number of cycles in each network was unknown before execution and as soon as the error in the learning data grew the learning algorithm stopped.

f) Validation: This series of data were introduced into the network after training (only MLP) and the produced results were compared with the actual values of horizontal and vertical porosities, the consistency of these two sets of values was used as a criterion to universalizability of the network.

g) Test/Verification: This stage is applicable only to neural networks trained by backpropagation. In this stage, the training error was monitored after each cycle, in order to stop the learning algorithm and prevent the network from saving the learned weights if the training error exceeded the required accuracy. Software packages such as Weka, MATLAB, Excel, and Neuro Solutions were used in this stage.

3.2. Methodology

3.2.1. MLP

Considering its applicability in prediction problems and its capability in generalizing the results, feedforward MLP was used for vertical and horizontal porosity prediction. The MLP includes 8 inputs in the input layer, 2 hidden layers and finally the output layer which have the results of the calculations and contains the solutions. Different hidden layers with different neurons were tried the best of which turned out to be those of greater R and smaller MSE in the two layers. Tangent sigmoid technique (Tansig) was used for introduction to the hidden layer and pure line (Purelin) was used to pass from the hidden layer to the output layer.

Nevertheless, Levenberg-Marguardt back propagation was used due to its fast convergence in small to medium size networks. A number of 212 horizontal porosity data were available whose 148, 32 and 32 numbers were applied in training, validation and test respectively. Besides, from the 190 vertical porosity data, 132, 29 and 29 numbers of them were accordingly used in training, validation and test phases. The best R and MSE values and normalized values for horizontal and vertical porosity obtained from MLP and MPL optimized by GA and PSO have been listed in Tables 1 and 2. furthermore, Tables 3 and 4 show the best R and MSE obtained with MLP optimized by GA and normalized vertical and horizontal porosities for different two-layer MLP forms.

It's worth mentioning that in single-layer and 3-layer perceptron the correlation coefficients were smaller and the MSE is greater compared with 2-layer perceptron, thus they were ignored to be presented here.

Figure 5 presents a comparison of normalized actual and predicted vertical and horizontal values at R values (highest value) and MSE values (lowest) in the best prediction mode, that is combination of MLP and GA, including the cases of total data, data used for training, validation data and test data. And Figure 6, for instance, shows the performance curve of the best normalized MLP network combined with GA in porosities of the horizontal and vertical core.

Table 1. Comparison of the best R and MSE values for horizontal core porosities, normalized by pure MLP (Marguardt-Levenberg learning) and MLP optimized by GA and PSO

Parameter	MLP	MLP optimized with GA	MLP optimized with PSO
R (Total data)	0.95715	0.99591	0.96912
MSE (Total data)	0.0039098	0.00038082	0.0038276
R (Learning data)	0.96564	0.99689	0.98268
MSE (Learning data)	0.0035528	0.00031946	0.0012266
R (Validation data)	0.8652	0.99313	0.96386
MSE (Validation data)	0.0064069	0.00075073	0.0017712
R (Testing data)	0.9245	0.99023	0.98392
MSE (Testing data)	0.0030632	0.0002949	0.0019726

Table 2. Comparison of the best R and MSE values for vertical core porosities, normalized by pure MLP (Levenberg-Marguardt learning) and MLP optimized by GA and PSO

Parameter	MLP	MLP optimized with GA	MLP optimized with PSO
R (Total data)	0.96262	0.99507	0.98813
MSE (Total data)	0.00076555	0.00021173	0.00050951
R (Learning data)	0.96968	0.99496	0.99000
MSE (Learning data)	0.00055004	0.00018985	0.00037344
R (Validation data)	0.94108	0.99043	0.96707
MSE (Validation data)	0.001369	0.00025433	0.00089356
R (Testing data)	0.96693	0.9989	0.98136
MSE (Testing data)	0.01143	0.0002687	0.00074352

Table 3. The results of applications of different two-layer normalized perceptron configurations with GA optimization for horizontal core porosity prediction in test data (15 of the total data). the best result (highest R, lowest MSE) has been bolded.

No. of Networks	No. of neurons in two-layer config.	R	MSE
1	[2 2]	0.98901	0.0012405
2	[4 3]	0.98367	0.00189204
3	**[5 4]**	**0.99023**	**0.0002949**
4	[8 6]	0.89718	0.00095807
5	[10 11]	0.90804	0.001659
6	[15 15]	0.92315	0.00091356

Table 4. The results of applications of different two-layer normalized perceptron configurations with GA optimization for vertical core porosity prediction in test data (15 of the total data). The best result (highest R, lowest MSE) has been bolded.

No. of Networks	No. of neurons in two-layer config.	R	MSE
1	[3 2]	0.98456	0.0017861
2	**[4 5]**	**0.99890**	**0.0002687**
3	[5 6]	0.97271	0.0097802
4	[8 9]	0.942504	0.0025708
5	[11 10]	0.95147	0.004486
6	[12 13]	0.90806	0.005238

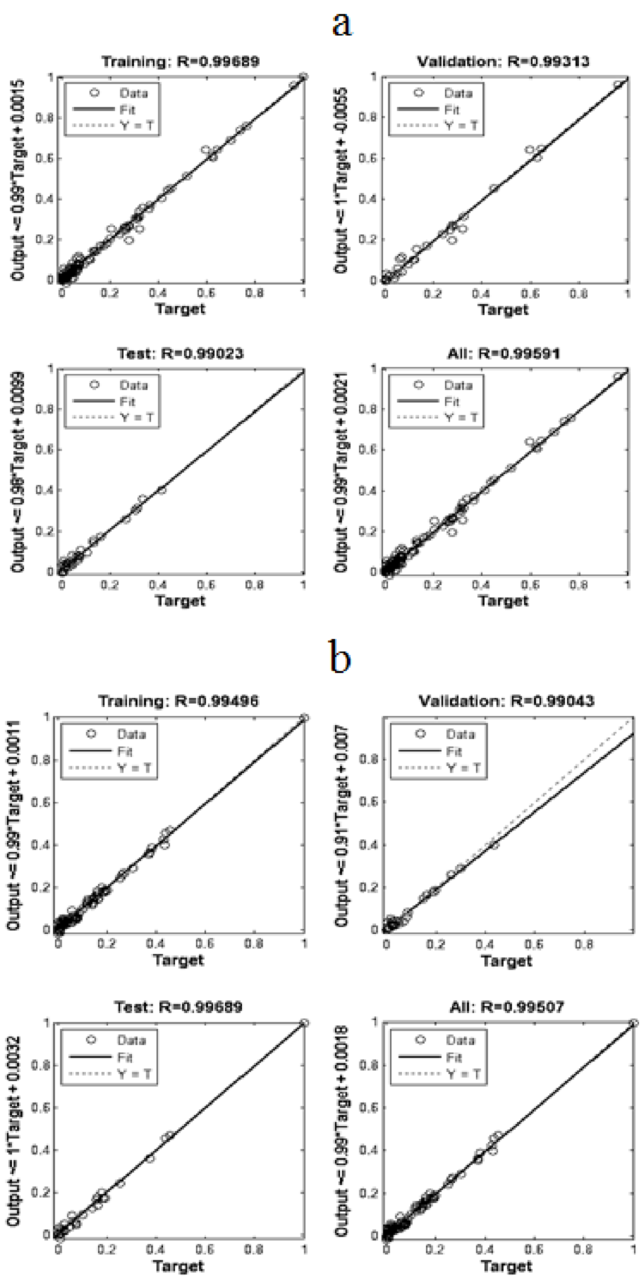

Figure 5. The best R in actual normalized porosity (horizontal axis) vs. that in normalized porosity predicted by MLP optimized by GA (vertical axis). (a) Horizontal cores, (b) Vertical cores

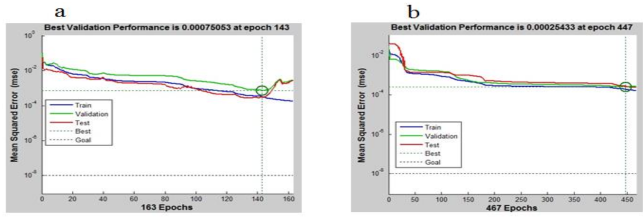

Figure 6. Performance curve, normalized MLP optimized by GA. (a) Horizontal core, (b) Vertical core

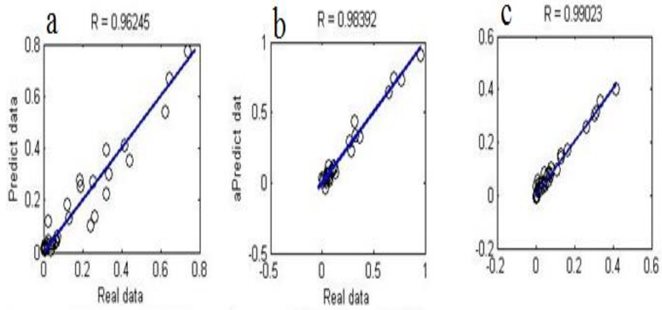

Figure 7. Comparison of correlation coefficient (R) in horizontal core. (a) MLP, (b) Optimized by PSO, (c) Optimized by GA

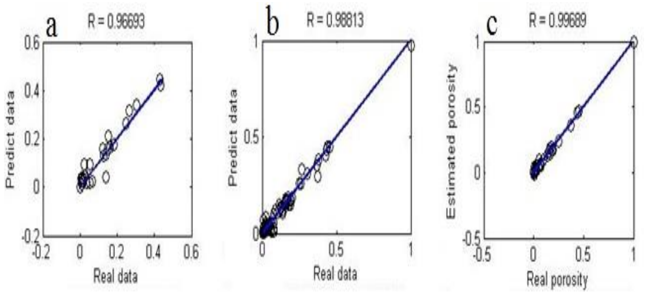

Figure 8. Comparison of correlation coefficient (R) in vertical core. (a) MLP, (b) Optimized by PSO, (c) Optimized by GA

3.2.2. The RBF method

As mentioned before, these networks operate well when there's a large amount of input data. since there's no validation stage in such networks in analogy with MLP networks, RBF uses a radial mechanism which enables it to have a better interpretation when there's large amount of input [8]. the same data as in MLP were used here, 212 horizontal core porosities of which 149 ones were used in training part and the rest were used for testing, as well as 190 vertical core porosities 133 of which were used in training stage and the rest were used for testing. Tables 5 and 6 show the best normalized R and MSE values for horizontal and vertical porosities obtained by RBF and RBF optimized by different techniques. As can be seen the combination of RBF and PSO yields the highest R and lowest MSE. Figure 9 presents a comparison of the best R values for actual porosities and those predicted in test data by normalized RBF and optimized by PSO.

Table 5. Comparison of the best R and MSE values for vertical core porosities, normalized by pure RBF (Levenberg-Marguardt learning) and RBF optimized by GA and PSO

Parameter	RBF	RBF Optimized by GA	RBF optimized by PSO
R (Total data)	0.89976	0.96715	0.98388
MSE (Total data)	0.011355	0.0039891	0.0015553
R (Learning data)	1.000	1.000	1.000
MSE (Learning data)	9.58E-17	4.078E-17	1.0104E-17
R (Testing data)	0.76411	0.92585	0.93132
MSE (Testing data)	0.037547	0.0091216	0.005152

Table 6. Comparison of the best R and MSE values for vertical core porosities, normalized by pure RBF (Levenberg-Marguardt learning) and RBF optimized by GA and PSO

Parameter	RBF	RBF Optimized by GA	RBF optimized by PSO
R (Total data)	0.91203	0.98159	0.99507
MSE (Total data)	0.0017469	0.00089621	0.00026875
R (Learning data)	1.000	1.000	1.000
MSE (Learning data)	9.996E-18	7.6805E-18	2.8039E-18
R (Testing data)	0.80169	0.9652	0.98184
MSE (Testing data)	0.0027804	0.0010366	0.00089582

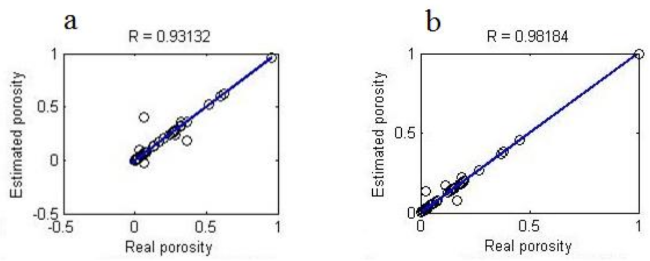

Figure 9. Comparison of R in normalized actual core porosities and those predicted by RBF optimized by PSO for test data. (a) horizontal core, (b) Vertical core

4. Conclusion

ANNs do not need complicated mathematical models, the important point is to select the adequate learning algorithm. furthermore, in ANN there's no need to know the rock lithology or the fluid filling the pores. Since the input layer neurons are functions of those parameters affecting the output, the input variables have to be selected so as to include the majority of parameters affecting the outputs, here the vertical and horizontal core porosities. ANN is a fast and effective technique to predict the porosity in comparison with time consuming and costly laboratory techniques, the nonporous spots of the rock can be distinguished by logs, results of this study to estimate the porosity in this field (horizontal and vertical cores) shows that ANN can be an effective technique to estimate reservoir parameters not only porosity, confirming the potential of ANN in reservoir studies. In general, by both artificial intelligence techniques used in this work for normalized data, ANNs, the best results were those obtained from MLP and RBF respectively combined with GA and PSO optimization. Results show that MLP and RBF, if combined with and R value optimization technique, deliver acceptable results for vertical and horizontal porosity prediction in this reservoir, however, since RBF does not have a validation stage, MLP is practically more reliable technique.

References

[1]. Bean, M., Jutten, C.,"Geophysics 65,1032–1047", Neural networks in geophysical appications, 2000.

[2]. Mohaghegh,S.,"SPE 29220 ", Artificial Neural Networks as a Valuable Tool for Petroleum Engineers, 1994.

[3]. Groupe, D.,"World ScientificPublishing Co., Singapore", Principle of Artificial Neural Networks, second ed,2007.

[4]. Doyle, A.M.,Schwaber,F.J.,Show, J.S.,"chem. Eng. 385 – 371", A dynamic neuralnetwork approach to non-linear processing modeling, 1997.

[5]. Rezaee, M. R., Nikjoo, M., Movahhed, B., Sabeti, N., "Journal of Geological Society of Iran, Vol. 1, pp. 21-27", Prediction of effective porosity and water saturation from wire line logs using artificial networks technique, 2006.

[6]. Mohaghegh, S.," SPE 58046", Virtual-intelligence applications in petroleum engineering, part1- Artificial neural networks, Distinguished Author Series, 2000.

[7]. Rezaee, M. R., Kadkhodaie Ilkhchi, A. & Alizadeh, P. M.",Journal of Geophysics and Engineering 12 - 26", intelligent approaches for the synthesis of Geophysics and Engineering, 2008.

[8]. Poggio, T., Girosi, F.,"Science 247,978–982", Regularization algorithms for learning that are equivalent to multilayer networks, 1990.

[9]. Han, J., Kamber, M."Data mining: concepts and techniques. Morgan Kaufmann (2006)".

[10]. J.H. Holland.," University of Michigan Press, Ann Arbor, Michigan re-issued by MIT Press (1992)", Adaptation in Natural and Artificial Systems, 1975

[11]. Kennedy, J., Eberhart, R.," IEEE International Conference on, vol. 4, pp. 1942-1948 (2002).", Particle swarm optimization. In: Neural Networks Proceedings, 1995.

[12]. Rostami,H., Khaksar Manshad, A.," Application of evolutionary Gaussian processes regression by particle swarm optimization for prediction of dew point pressure in gas condensate reservoirs", Neural Computing and Applications, doi={10.1007/s00521-012-1275-9}, 2012.

[13] Rostami, H., Shahkarami, A., Azin, R.,"The Prediction of the Density of Undersaturated Crude Oil Using Multilayer Feed-Forward Back-Propagation Perceptron", Petroleum Science and Technology, 30:1, 89-99(2011).

[14] Rostami, H., Azin, R., Dianat, R.," Prediction of Undersaturated Crude Oil Density Using Gaussian Process

Regression", Petroleum Science and Technology, 31:4, 418-427 (2013).
[15] Al-Bulushi, N. I., King, P. R., Blunt, M. J., Kraaijveld, M., "Artificial neural networks workflow and its application in the petroleum industry", Neural Computing and Applications, 21:3, 409-421(2012).
[16] Wu, X.J, Jiang, G.C., Wang, X. J., Fang, N., Zhao,L., Ma,Y.M, Luo, S.J., "Prediction of reservoir sensitivity using RBF neural network with trainable radial basis function", Neural Computing and Applications, 22:947–953 (2013).

Peyman Torabideh Chemical Engineering Department, Persian Gulf University of Bushehr, Iran

Email: p.torabideh@yahoo.com

Abbas Khaksar Manshad Department of Petroleum Engineering, Abadan Faculty of Petroleum, Petroleum University of Technology (PUT), Abadan, Iran
Email: Khaksar@put.ac.ir

Mohammad Amin Joghatayi received his BSc degree in electrical engineering from Khatam-al-Anbia University of Technology, Behbahan, Iran, in 2018. He is currently studying for an MSc degree in AI and robotics at Persian Gulf University, Bushehr, Iran, starting in 2024.
Email: m.a.joghatayi@mehr.pgu.ac.ir

Alireza Zangeneh Var Graduated Association of Sharif University (Technical Manager of Petro Pars Company).

Paper Handling Data:

Submitted: -
Received in revised form: -
Accepted: -
Corresponding author: Abbas Khaksar Manshad
Department of Petroleum Engineering, Abadan Faculty of Petroleum, Petroleum University of Technology (PUT), Abadan, Iran