Middle-East Journal of Scientific Research 13 (11): 1438-1444, 2013
ISSN 1990-9233
© IDOSI Publications, 2013
DOI: 10.5829/idosi.mejsr.2013.13.11.1349

Developing a Committee Machine Model for Predicting Reservoir Porosity from Image Analysis of Thin Sections

Amir Hatampour

Faculty of Chemical Engineering,
Dashtestan branch of Islamic Azad University, Borazjan, Iran

Abstract: The use of computational intelligence techniques has become increasingly popular in the field of geosciences. Artificial intelligence models are extremely useful for reservoir characterization, which requires high accuracy prediction for a good exploitation of the oil and gas resources. Porosity is one of the most important rock properties for fluid volume prediction and modeling of a petroleum reservoir. This study presents an improved approach utilizing the integration of intelligent systems to predict porosity. Firstly, the image analysis techniques were used for extracting petrographic parameters from thin section images. Consequently, the extracted parameters were input to two intelligent models, including a neural network (NN) and a fuzzy logic (FL) model. To improve the accuracy of the predicted porosity from NN and FL a genetic algorithm (GA) is employed, which combine the outputs of individual models. This methodology is known as committee machine with intelligent models (CMIS). The results show that CMIS performed better than NN and FL models acting alone. The proposed methodology was applied to South Pars gas field, which is the largest non-associated gas accumulation in the world.

Key words: Reservoir characterization · Image analysis · Thin section image · Committee machine with intelligent systems · Genetic algorithm

INTRODUCTION

Porosity is an important factor that identifies the space available for storage of hydrocarbon and it is a measurement of the storage capacity of a rock. An accurate estimation of porosity is functional for the geologists and petroleum engineers since the porosity can be used for developing static and dynamic models of an oil and gas accumulator. There are various approaches for determining the porosity value of a rock. The conventional studies used laboratory measurement (mercury injection), which is the most expensive and time-consuming approach. In addition, cores are not available in the whole depth of a reservoir. Developments of computer technology in software programming and hardware equipments yearned petroleum engineers to construct intelligent models for predicting reservoir characteristics based on integration of well log data and core laboratory measurements. Nowadays, the integration of intelligent systems and easily accessible reservoir characteristics, such as well log and petrographic data for prediction of essential and valuable characteristics, such as porosity and permeability, is known as a popular approach from the standpoint of the petroleum engineers. Intelligent models are able to predict porosity and permeability with acceptable accuracy and without requiring time-consuming and expensive laboratory measurements. Various researches were conducted to predict reservoir characteristics through individual intelligent systems and well log data [1-3]. Petrographic data, which can be obtained from thin section images, are also useful, cheap and easily accessible reservoir rock characteristics; however, they have received less attention for predicting reservoir characteristics compared to well log data; therefore, the authors of current paper used the petrographic data as inputs of intelligent models to compensate this deficiency in geological researches. On the other hand, hitherto, the researchers have focused on the applications of individual intelligent systems, independently. The current paper tries to integrate the outputs of two intelligent models and reaps the advantage of each model for improving the accuracy of final

Corresponding Author: Amir Hatampour, Faculty of Chemical Engineering, Dashtestan Branch of Islamic Azad University, Borazjan, Iran.

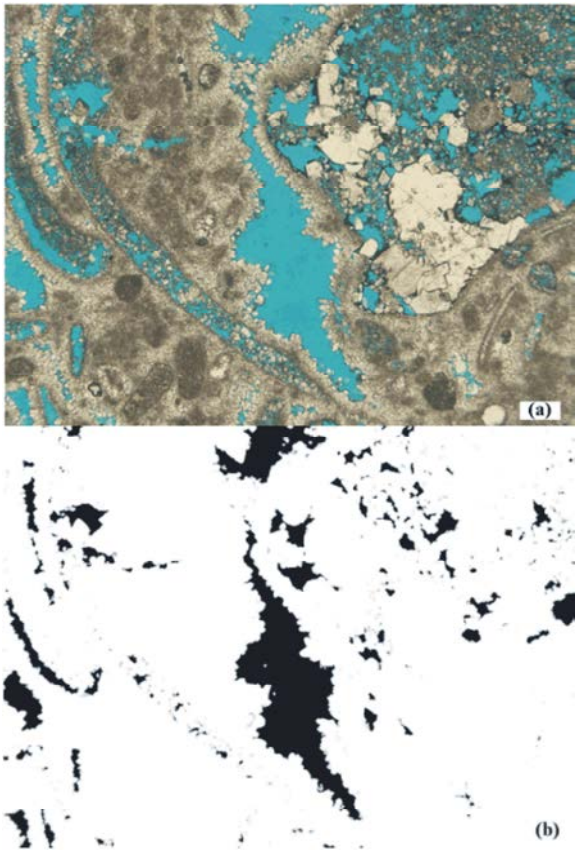

Fig. 1: Representative digital and segmented image of thin section: (a) Digital image of thin section. (b) Segmented pores image in which pores are shown with black.

predicted value of porosity. This concept can be called committee machine with intelligent systems. A committee machine has a parallel structure that produces a final output by combining the results of individual experts using an optimization technique [4]. Therefore, the model accuracy could be increased in comparison with the best individual expert. This methodology was successfully applied for predicting porosity from thin section petrographic data with an example from South Pars gas field. This study aims to cover the forfeiture of a comprehensive research about the behavior of intelligent systems for prediction of core porosity when petrographic data are used as input.

Image Analysis: The last decade is the witness of many advances in image analysis techniques, which have been employed in various branches of scientific fields because of their accurate results, rapid measurements and cost effectiveness. Petrographic image analysis is one of the most important subcategories of image analysis, which provides new ways of reservoir characterization for petroleum geologists and engineers. The process of petrographic image analysis for porosity prediction comprises four main steps, including image collection, image filtering, image segmentation and image measurement.

Image Collection: This step includes the process of image capturing from thin sections using a digital camera attached to an optical petrographic microscope. In this step, the numbers of captured images from each thin section and the microscope's field of view are two important factors. Here, ten images were captured from each thin section and the field of view at X12.5 magnification is considered large enough to analyze the pore spaces accurately. Fig. 1(a) shows a captured image from a thin section, which is used for extracting petrographic data.

Image Filtering: Images taken from the microscope were pre-processed to stabilize their light intensity and light polarity and also for removing the noises. The polishing procedures and image collection are the most important procedures that make image noises. Only some of the studied images required filtering and most of them had an acceptable quality without filtering procedure.

Image Segmentation: The segmentation involved the determination of pixels that belong to the same category. For extraction of petrographic parameters, the pixels of each image were divided to two categories named pore spaces and rock matrix. The blue pixels show the pore area and the remains indicate rock matrix. To study the pore areas easily, a segmented image was prepared from each image in which values of one (black) was allocated to pixels with the correct value for pores and zeros (white) for the other ones (Fig. 1(b)).

Image Measurement: Various features can be extracted from thin section images in this stage. These features are the inputs of intelligent systems for predicting porosity. The effective parameters on porosity that can be appeared in thin sections are investigated in the preliminary studies of this research and finally, ten petrographic parameters, including grain size, intergranular porosity, interagranular porosity, moldic porosity, microporosity, cement, limestone, dolomite, type of texture and mean geometrical

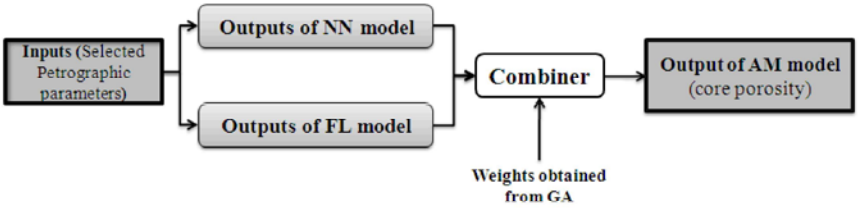

Fig. 2: Schematic diagram of CMIS designed in this research

shape coefficient of grains (γ_{mean}) were acquired for each pore area using the petrographic image analysis techniques and image processing toolbox in Matlab software.

Theory and Methodology

Why to Use the Concept of Committee Machine?: The current study describes various methods for robust prediction of porosity using soft computing (SC) techniques. Soft computing includes several expert systems, which can be integrated to produce high efficient hybrid models for reservoir rock characterization. The singular applicability of soft computing techniques is extensively proven in petroleum industries. However, this study tries to introduce a hybrid model, named committee machine (CM), to improve the accuracy of final prediction based on integrating individual experts. The experts can be empirical formula, neural network, a decision tree, or other type of algorithm [5]. The idea behind the CM is to fuse results presented from experts to produce an overall decision that is superior to any of the individual experts acting alone.

The CM described in this study comprises of two major steps. Firstly, two intelligent models, based on neural network and fuzzy model were constructed for formulating conventional petrographic data to porosity. Consequently, a genetic algorithm was used to evaluate the optimal contribution of each expert in the CM model. A schematic diagram of CM is shown in Fig. 2. The CM compensates the overestimation and/or underestimation effect of constructed models through the use of evolutionary computing rules. As a prelude, a brief introduction about employed expert systems is mentioned in following sections.

Individual Intelligent Systems: Artificial neural networks (ANNs) are computational models inspired by brain structures, mechanisms and functions. Back-propagation neural network is a supervised learning method, which means that it requires a set of training data having the desired output for any given input. The network computes the difference between the calculated output and corresponding desired output from the training data set. The error is then propagated backward through the network to adjust the connection weights and biases during a number of iterations (epochs). The training process ceases as immediate as the network outputs locate within close proximity of desired outputs [6].

The basic theory of fuzzy logic (FL) proposed by [7] as an extension of traditional Boolean logic. The main part of FL is fuzzy inference system (FIS), which formulates the inputs to an output. [8] introduced the first type of FIS and [9] initiated the second one. Both FISs, attempt to control a system by integrating a set of linguistic control rules. The distinct difference between Takagi Sugeno-FIS method and Mamdani Assilian method is that output membership functions are either linear or constant in Takagi Sugeno (TS) method. In the TS-FIS, membership functions are defined by clustering process, which divides the input-output data set into fuzzy clusters, each representing a specific part of the system behavior. Each membership function generates a set of fuzzy 'if-then' rule for formulating inputs to output. The fuzzy system makes a sum of all 'then' parts and uses a defuzzification method to give the final output.

Evolutionary Genetic Algorithm: GAs are also known as evolutionary computing approaches, which mimic their behavior from an analogy of the evolution processes in the nature. These algorithms present hopeful results when dealing with complex problems. GAs are population-based algorithms because they start with initial population of possible solutions (individuals). Consequently, a fitness evaluation unit determines a fitness value, which is a measure of the fitness of individuals. The better parent solutions are then reproduced and the next generation of solutions (children) is generated by applying two genetic operators, including crossover and mutation. The crossover operator chooses parents at random and produces children. The mutation operator randomly changes the values of elements in a string. New generations of solutions are evaluated and the whole cycle repeats until the best solution is obtained [10].

Table 1: Performance of the neural network for predicting porosity in test data using several sets of inputs. In this table, G.S., O.P., I.P., Tex., Micro., γ_{mean}, Dol. and Lime, illustrate the abbreviation of grain size, optical porosity, interparticle porosity, type of texture, microporosity, mean geometrical shape coefficient of grains, dolomite percentage and limestone percentage, respectively

Selection of appropriate inputs for porosity prediction using NN	
Inputs	MSE ($\times 10^{-2}$)
G.S.	3.23
G.S. + O.P.	3.02
G.S. + O.P. + Tex.	2.88
G.S.+ O.P. + Tex. + Micro	1.71
G.S. + O.P. + Tex. + γ_{mean}	1.49
G.S. + O.P. + Tex. + γ_{mean} + Dol.	2.04
G.S. + O.P. + Tex. + γ_{mean} + I.P.	2.71
G.S. + O.P. + Tex. + γ_{mean} + Dol. + Cement	2.27
G.S. + O.P. + Tex. + γ_{mean} + I.P. + Cement	2.73
G.S. + O.P. + Tex. + γ_{mean} + Dol. + Cement + I.P.	4.68
G.S. + O.P. + Tex. + γ_{mean} + Micro + I.P. + Cement	4.14
G.S. + O.P. + Tex. + γ_{mean} + Micro + I.P. + Cement + Lime.	4.03

Input Selection for Intelligent Models: Thin section images of two wells of the South Pars gas field were analyzed for construction and evaluation of intelligent models. In this stage, 115 thin sections were used from which 86 samples were considered to train the experts and 29 samples to test the constructed models. As mentioned before, ten petrographic parameters were extracted for each thin section. Using ten petrographic parameters as inputs of the intelligent model increase the noise of data and reduce the performance of the model. Consequently, a neural network (NN) model has been employed to select the best inputs. Several combinations of petrographic data were introduced into NN models and the performance of the model was evaluated at each stage. By comparing the performances of constructed models, the most appropriate combination of petrographic parameters was designated. The results of NN models for input selection are shown in Table 1. In this table, G.S., O.P., I.P., Tex., Micro., γ_{mean}, Dol. and Lime, illustrate the abbreviation of grain size, optical porosity, interparticle porosity, type of texture, microporosity, mean geometrical shape coefficient of grains, dolomite percentage and limestone percentage, respectively. It is worth mentioning that the inputs and outputs of neural network and fuzzy models were normalized to the interval of [-1 1] and [0 1], respectively. This task usually enhances the performance of intelligent systems for better prediction. The MSE values of Table 1 were computed based on normalized scale.

Case Study

Prediction of Porosity Using NN: The Levenberg-Marquardt training algorithm was used for updating weights and bias values of a simple three-layered error back propagation algorithm. The hidden layers included 5 neurons. The input and output layer consisted of 4 and 1 neuron, respectively. The transfer function between the layer one and layer two was TANSIG, while PURELIN was employed for the last layer. The performance of the model was equal to 5.42E-03 for training data. After training the network, the test data were introduced to the model and the porosity was calculated. The measured mean square error (MSE) in the test well was equal to 1.49×10^{-2}, which correspond to the R^2 value of 0.816. Fig. 3 shows the correlation coefficients between the NN predicted porosity and measured one.

Prediction of Porosity Using FL: In this study, a TS-FIS was employed to predict porosity. A subtractive clustering method was used for classifying the input and output datasets. The main parameter of the subtractive clustering affecting the performance of a fuzzy model is determination of the optimal clustering radius. A small clustering radius will yield more and smaller cluster and generates more rules. In contrast, large clustering radius yields a few large clusters in data and generates few rules [11]. Searching for the optimal clustering radius was done by performing clustering process several times and gradually increasing the clustering radius from 0 to 1 (with 0.1 intervals) (Table 2).

Thus, 10 fuzzy models with different number of if-then rules were established. Then, the fuzzy model with the highest overall accuracy was selected as the optimal model. The results show that taking the clustering radius of 0.9 leads to the highest performance. The MSE of the fuzzy system for training and test data is equal to 6.01E-0.3 and 2.21×10^{-2}, respectively Fig. 4 illustrates the correlation coefficients between the real and FL predicted porosity values in the test data.

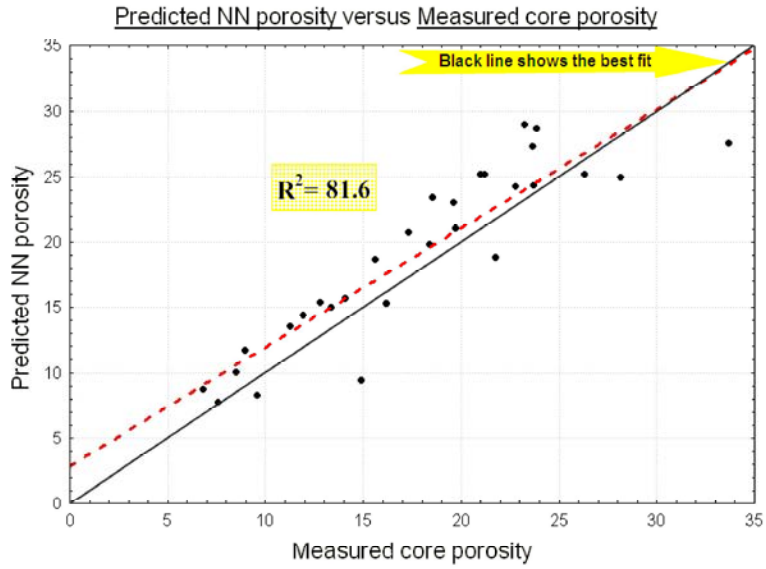

Fig. 3: Correlation coefficient between predicted porosity and measured one for test samples using neural network.

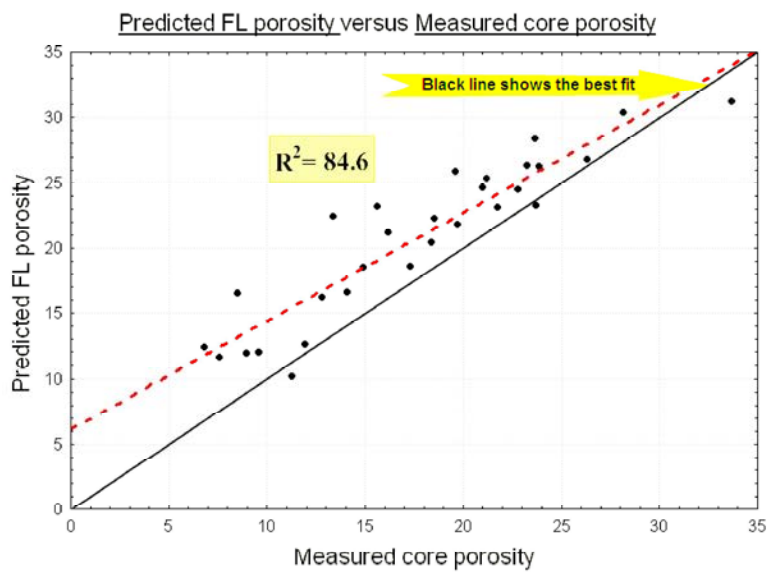

Fig. 4: Correlation coefficient between predicted porosity and measured one for test samples using fuzzy logic.

Table 2: Clustering radius, number of if-then rules and MSE for each FL model

No. FIS	Clustering radius	No. of 'if-then' rules	MSE of fuzzy model
1	0.1	81	0.0452
2	0.2	65	0.0443
3	0.3	30	0.0432
4	0.4	11	0.0386
5	0.5	5	0.0379
6	0.6	4	0.0298
7	0.7	4	0.0284
8	0.8	3	0.0267
9	0.9	2	0.0221
10	1	2	0.0234

Improving the Accuracy of Predicted Porosity Using CMIS: Finally, the main concept of this study was practically examined here. The outputs of individual intelligent models, including neural network and fuzzy logic, were integrated through the capabilities of GA for determining the optimal contribution of the individual algorithms. Fitness function that should be optimized with GA is defined as below:

$$MSE = \frac{1}{n}\sum_{1=1}^{n}(W_1 \times NN_i + W_2 \times FL_i - \phi_i)^2 \quad (8)$$

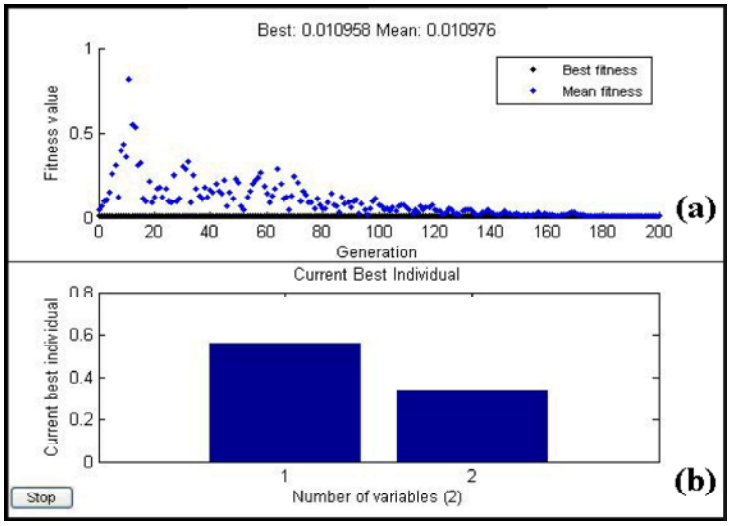

Fig. 5: Results of running GA including the best and mean fitness value for fitness function after 200 generations (a) and weight coefficients of experts (b).

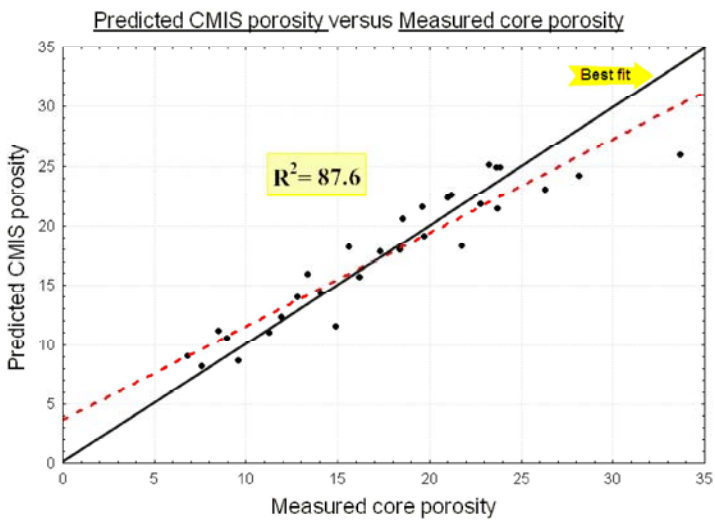

Fig. 6: Correlation coefficient between predicted porosity and measured one for test samples using CMIS.

This function shows the MSE of the CMIS prediction where W_1 and W_2 the weight factors corresponding to the NN and FL predictions, respectively. Φ_i is the target value (measured porosity from core plugs) and n is the number of training data points.

The initial population size was set to 20 which specifies the number of individuals in each generation and initial range was set to [0, 1] which specifies the range of the individuals in the initial population. The maximum number of generations that specifies the maximum number of GA iterations was set to 200. The best and mean fitness value for fitness function after 200 generations is shown in Fig 5(a). The next generation was produced using crossover and mutation operators with default functions in the GA toolbox [12]. Fig 5(b) illustrates weight coefficients of experts.

After running the GA, the optimized weight coefficients were applied to the test data to produce the final output:

$$K_{CMIS} = 0.579 \times \phi_{NN} + 0.378 \times \phi_{FL} \quad (9)$$

The MSE of the CMIS predicted porosity is 1.09×10^{-2} that shows the minimum MSE and maximum correlation coefficient with respect to other methods. Crossplot of Fig. 5 shows the correlation coefficient between the measured and CMIS predicted results. A comparison between the measured and CMIS predicted porosity versus depth is shown in Fig. 6.

RESULTS AND DISCUSSION

In the present paper, only two individual intelligent systems were constructed and their outputs were integrated finally. However, the CM is able to combine the results of more individual experts. The only difference should be considered in the fitness function, where the number of weights increases with respect to numbers of intelligent models. It is clear, in this situation, more accurate results will be obtained for predicting reservoir characteristics. The main results of this paper were listed as below:

- Results of our study show that integration of petrographic image analysis and intelligent models can successfully be used in the estimation of porosity from thin sections in carbonates of South Pars gas field. This methodology is able to convert the two-dimensional porosity into three-dimensional one. Using integration of the intelligent systems and petrographic data can be an effective methodology for porosity estimation specially, in old fields, where cores may not be available or well preserved. In addition, this technique can be applied on the petrographic data collected from drill cuttings.
- A NN model was develop for selecting the best combination of petrographic parameters as inputs of intelligent models. Comparisons of MSEs revealed that the set of grain size, optical porosity, texture and mean geometrical shape coefficient of grains, have lowest MSE for prediction of porosity.
- This study shows an acceptable prediction of porosity from petrographic data using intelligent systems. The MSE of the NN and FL methods for prediction of porosity in the test data are 1.49×10^{-2} and 2.21×10^{-2}, which correspond to the R^2 values of 81.6 and 84.6, respectively.
- The CMIS was carried out to obtain the optimal combination of the weights. The CMIS determines the contribution of NN and FL, using a GA. The GA derived weights for the NN and FL models are 0.579 and 0.378, respectively. MSE of the CMIS for the test data is 1.09×10^{-2}, which corresponds to the R^2 value of 87.6. It is clear that, the CMIS performs better that NN and FL models.

REFERENCES

1. Jamialahmadi, M. and F.G. Javadpour, 1999. Relationship of permeability, porosity and depth using an artificial neural network. Journal of Petroleum Science and Engineering, 26: 235-239.
2. Lim, J., 2005. Reservoir properties determination using fuzzy logic and neural network from well log data in offshore Korea. Journal of Petroleum Science and Engineering, 49: 182-192.
3. Kadkhodaie-Ilkhachi A., M.R. Rezaee and S.A. Moallemi, 2006. A fuzzy logic approach for the estimation of permeability and rock types from conventional well log data: an example from the Kangan reservoir in Iran Offshore Gas Field. Iran. Journal of Geophysics and Engineering, 3: 356-369.
4. Haykin, S., 1991.Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs, NJ, pp: 842.
5. Sharkey, A.J.C., 1996. On combining artificial neural networks. Connect. Sci., 8: 299-314.
6. Bhatt, A. and H.B. Helle, 2002. Committee neural networks for porosity and permeability prediction from well logs. Journal of Geophys. Prospect, 50: 645-660.
7. Zadeh, L.A., 1965. Fuzzy sets. Information and Control, 8: 338-353.
8. Mamdani, E.H. and S. Assilian, 1975. An experimental in linguistic synthesis with fuzzy logic control. Journal of Man-Machine Studies, 7: 1-13.
9. Saemi, M. and M. Ahmadi, 2007. Integration of genetic algorithm and a coactive neuro-fuzzy inference system for permeability prediction from well logs data. Journal of Transp Porous Med., 71: 273-288.
10. Chiu, S., 1994. Fuzzy model identification based on cluster estimation. Journal of Intelligent Fuzzy Systems, 2: 267-278.
11. Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, USA, pp: 183.
12. Matlab user's Guide. 2010a. Fuzzy logic and genetic algorithm toolboxes, Matlab CD-ROM, by the Mathworks, Inc.