International Journal of Innovation and Applied Studies
ISSN 2028-9324 Vol. 7 No. 2 Aug. 2014, pp. 546-556

http://www.ijias.issr-journals.org/

Artificial intelligence approach to reservoir fluid classification

***Abbas Mohamed Al-Khudafi*[1] *and Khaled. A Abd-El Fattah*[2]**

[1]Department of Petroleum Engineering,
Hadhramout University of Science and Technology,
Al-Mukalla, Yemen

[2]Petroleum Department,
Cairo University,
Cairo, Egypt

ABSTRACT: Fluid classification is a critical factor in decision of reservoir and production problems. Reservoir fluid can be classified into five types according to laboratory and production data as black oil, volatile oil, gas condensate, wet gas and dry gas. In this work a novel application of Neural Networks (ANN) is presented. Based on production and laboratory data neural networks model is developed for automatic classification of reservoir FLUID. More than 450 samples of five types of reservoir fluids are used to develop the neural network model. About 70 % of data are accepted for neural network training, 15 % for validation and 15 % are used as test set. The importance of different input fluid properties in classification was studied.
The different types of architectures for different groups of input data were tested to select the optimal neural network architecture by fitness criteria. The optimized neural network model was capable of classifying the reservoir fluids with high accuracy. The performance of ANNs models was determined by classification quality index and network error.
The model has been applied successfully to classification of Yemeni fluids using different range of parameters. The results show that the proposed novel ANN model can achieve high accuracy.

KEYWORDS: Classification, reservoir fluid type, artificial neural network, model, Contribution of input.

1 INTRODUCTION

Determining the exact nature of a reservoir fluid of course is a matter of considerable practical importance. In automated fluid classification, data such composition, formation volume factor, gravity, etc. are generally utilized to classify the fluid's type. Fluid type can be identified by rule of thumb [17]. However Reservoir fluid type can be confirmed only by observation in the laboratory. According to rule of thumb, three properties are must readily available to indicate the type of fluid in reservoir: the initial producing gas- oil ratio, the gravity of the stock-tank liquid, and the color of the stock tank liquid. Initial producing gas oil ratio is by far the most significant criteria of fluid type identification. However, stock-tank liquid gravity and color are helpful in confirming the fluid type indicated by the producing gas-oil ratio.

If all three indicators- initial gas-oil ratio, stock tank liquid gravity, and stock tank liquid color – do not fit within the ranges given in the rules of thumb, the rules fail and the reservoir fluid must be observed in the laboratory to determine its type.

Artificial neural networks (ANN) have been involved in many applications to solve real world problems. In petroleum engineering ANNs can be applied to solve many engineering problems such as classifications, prediction, pattern recognition, and non-linear problems where the issues are very difficult or might be impossible to solve through normal mathematical processes. Intelligent techniques are powerful tools which overcome incompleteness, imprecise and uncertainty existent in reservoir parameters. These systems can recognize the possible patterns between input and output spaces.

The objective of this work is to classify the reservoir fluid (Black oils, volatile oils, retrograde gas condensates, wet gas, dry gas) based on a laboratory and field dataset of several fluid samples. The classification was carried out using the artificial neural networks).

2 DATA SET

A set of PVT fluid reports were analyzed to acquire various types of reservoir fluids. The database for this classification study consisted in 454 fluids samples with different characteristics from different sedimentary basins. The following parameters were used as input variables in the classification model API, C7+, GOR, MWC7+, SG, TR, Pd, Pb. The model covers wide range of input parameters (Table 1).

Table 1. Classification model parameters and range

Parameter	Maximum	Minimum	Average
S.G	1.0317	0	0.0408
MWC7+	534	0	50.42
C_{7+}	86.1	0.08	15.64
Bo	2.368	0	0.2782
GOR	150036	1	6826.89
API	72	0	8.14
TR	340	100	34.14
P_b	7450	0	1172.24

3 BUILDING THE MODEL

More than 450 data points were collected from PVT reports analyzed and accepted for neural network training. The input parameters to model are C_{7+} , Bo, GOR, API, MWC7+, TR S.G, P_b, P_d. Fluid class used as an output parameter. Depending on accepted input parameters some parameters were disabled. Data randomly were partitioned. Data Partition means division each dataset onto three sets: the training set, the validation set and the test set. The Training set is a part of input dataset used for neural network training, i.e. for adjustment of network weights. The Validation set is a part of the data used to tune network topology or network parameters other than weights, for example, the number of hidden units. Validation set is used to calculate generalization loss and retain the best network (the network with the lowest error on validation set). The Test set is a part of the input data set used only to test how well the neural network will perform on new data. The test set is used after the network is trained, to test what errors will occur during future network application. This set is not used during training and thus can be considered as consisting of new data entered by the user for the neural network application.

About 312 records (70% of data to training set, 71 records (15%) to validation set and 71 records (15%) to test the model. An anomaly detection model to predict whether a data point is typical for a given distribution or not. Neural network Architecture (input layer and number of neuron, hidden layer and number of neurons in hidden layer) was selected manually. Hidden layers activation, Error function and activation function are also specified.
Data preprocessed using scaling range: (-1..1) for input parameters. Output variable was transformed to a scale between 0 and 1. The network training is accomplished by the back-propagation algorithm.

The Heuristic search method is used finding optimal neural network architectures to find the number of nodes in the input and hidden layers was done according to a maximum fitness function (minimum training error). There are two neural network classification models: Winner-takes-all and Confidence-limits [7].

In present study With the Winner-takes-all model classification performed by selecting an output unit with the biggest activation level.

The network is trained by iterations process. When desired error is achieved training stopped and the best network was tracked when best correct classification rate is get. Overtraining is identified using the Validation set. The situation when the network error increases on the validation set during several iterations while still decreasing on the Training set is identified as the starting point of overtraining. Neural network automatically tested after training completion. In the testing process, the actual vs. output are compared error values for each data point from the input dataset is calculated.

The neural network classification model consists of three layers, which comprise an input layer, a hidden layer, and an output layer (Fig.1).

4 RESULTS AND DISCUSSION

The ANNs developed was successfully trained and tested using the available data sets (training, validation, testing and all). The validity and performance of ANNs models was determined by classification quality index (CQI) and network error.

4.1 CLASSIFICATION QUALITY INDEX

CQI is as a quality index used to test neural networks' capabilities for classification. This rate is calculated by dividing the number of correctly recognized records by the total number of records.

$$\mathrm{CQI} = \frac{\mathrm{N_c}}{\mathrm{N}} \mathrm{x}100$$

Where

N_c= the number of correctly recognized records

N=total number of records

4.2 NETWORK ERROR

Another parameter used to determine performance is the network error. A value used to rate the quality of the neural network training process. The smaller the network's error is, the better the network has been trained. Minimization of the error is the main objective of neural network training.

4.3 CONTRIBUTION OF INPUT FLUID PARAMETER

In order to determine the significant parameters that affect the classification in this study, different ANNs models were developed. The different sets of input parameters were used. Table 2 present the different group of input parameters to the model to classify reservoir fluid. Initially, input for the first model used four parameters with fluid class as desired output. Then the second model only uses another group of data. For each group of input data the performance of the model is evaluated. Table 3 shows the results of CQI and training error for each group of parameters affecting the fluid classification.

We used an approach one variable is excluded in each group is evaluated in order to determine the percentage of contribution by the variable. The sensitivity for each input is estimated by the change of the output for each class with respect to the input feature perturbation.

Fig. 2 shows the error of trained neural network where one the error dependence on various group of input data can be seen. This graph allows you to analyze which ranges of the selected input tend to produce bigger or smaller network errors. CQI graph plot for various training, testing and validation dataset and for all of them is presented in Fig.3.

Table 2. Type of input and output parameters to neural model

n	**group**	**Input parameters**	**output**
1	A	C_{7+} , Bo, GOR, API	Class
2	B	C_{7+} , Bo, GOR, API, SG	Class
3	C	C_{7+} , Bo, GOR, API, , MWC7+	Class
4	D	C_{7+} , Bo, GOR, API, SG, MWC7+	Class
5	E	C_{7+} , Bo, GOR, API, SG, MWC7+, P_b	Class
6	F	C_{7+} , Bo, GOR, API, SG, MWC7+, TR	Class
7	G	C_{7+} , Bo, GOR, API, SG, MWC7+, TR, P_b	Class

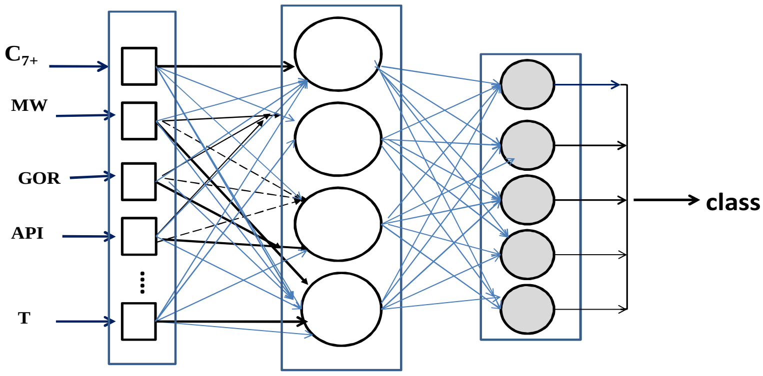

Fig. 1. ***Neural network classification model***

Table 3. CQI for various dataset and different training algorithm

group	**dataset**	**Classification quality index**					
		Training algorithm					
		Quick.	Conj.	Limit.	Quasi.	onlineBP.	BatchBP.
A	**training**	99	83.33	91.66	91.66	96.79	91.66
	validation	98.59	84.5	90.14	91.14	98.59	90.14
	testing	100	85.91	91.54	90.14	100	90.14
	all	99.12	83.92	92.07	91.18	97.57	91.18
B	**training**	99.33	84.33	93.35	93.35	99	84.38
	validation	98.57	90	94.28	95.71	97.14	90
	testing	100	88	95.71	95.71	98.57	88.57
	all	99.31	85.94	93.87	94.1	98.63	85.94
C	**training**	98.71	83.33	88.78	91.66	96.15	83.33
	validation	97.18	84.5	92.95	90.14	98.59	84.5
	testing	100	85.91	95.77	91.14	98.59	85.91
	all	98.67	83.92	90.52	90.18	96.91	83.92
D	**training**	99.03	83.33	95.91	91.34	97.11	83.65
	validation	100	84.5	91.54	90.14	98.59	84.5
	testing	100	85.91	94.36	90.14	100	85.91
	all	99.33	83.92	94.71	90.96	97.79	84.14
E	**training**	98.00	92.03	84.38	92.35	99	84.38
	validation	98.57	95.71	90	95.71	98.57	90
	testing	100	94.28	88.7	94.28	100	88.57
	all	98.41	92.97	85.94	93.19	99.09	85.94
F	**training**	99.67	83.33	91.66	91.34	96.15	83.97
	validation	100	84.5	90.14	90.14	97.18	84.5
	testing	100	85.91	90.14	90.14	97.18	85.91
	all	99.77	83.92	91.18	90.96	96.47	84.36
G	**training**	99.66	90.03	92.35	91.69	98.67	84.38
	validation	97.14	94.28	95.71	95.71	98.57	90
	testing	100	92.85	92.85	92.85	98.57	88.57
	all	99.31	91.15	92.97	92.51	98.63	85.94

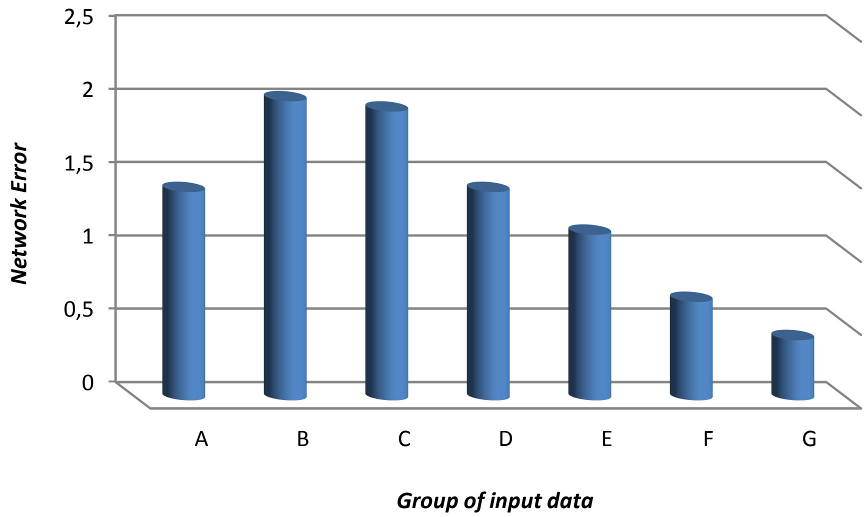

Fig.2 Neural network Error of training

Table 4. Classification quality index for various data sets depending on input group

group	training	validation	testing	all
A	99	98.59	100	99.12
B	99.33	98.57	100	99.31
C	98.71	97.18	100	98.67
D	99.03	100	100	99.33
E	99	98.57	100	99.09
F	99.67	100	100	99.77
G	99.66	97.14	100	99.31

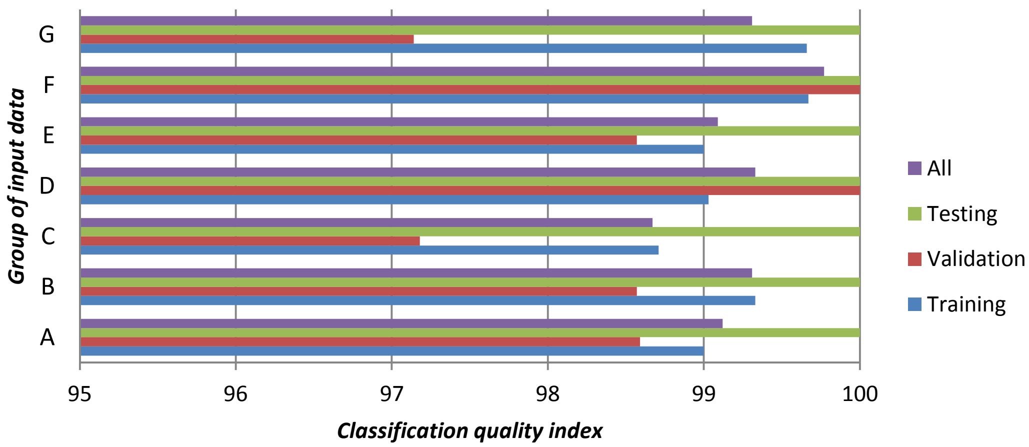

Fig.3 Classification quality index for various data set

4.4 SEARCHING THE OPTIMAL NETWORK ARCHITECTURE

In this work optimal neural network (NN) architecture is selected using automated optimization methods. Smaller test error gives the better network. This parameter is calculated as inverse mean absolute network error on the test set.

Fig. 4-10 represents relationship of quality index and iteration for training and validation data. Table 5 shows the different types of architectures for different groups of input data. Using inverse Test error as fitness criteria and heuristic search method along with the architecture search, the architecture with best fitness is selected as optimal architecture NN.

Table 5. Searching the Best Network Architecture ICQ and Network Error

	Architecture	Before				After		
		ICQ		Network Error	Architecture	ICQ		Network Error
		Training	Validation			Training	Validation	
A	**4-2-5**	98.71	100	0.0052	4-10-5	100	100	0.76
B	**5-2-5**	99.67	98.57	0.668	5-13-5	100	98.57	0.47
C	**5-2-5**	83.97	84.5	26.17	5-13-5	100	100	0.89
D	**6-3-5**	99.67	97.18	1.3	6-15-5	100	100	0.68
E	**7-8-5**	100	98.57	0.48	7-4-5	100	100	0.86
F	**7-3-5**	100	98.59	1.68	7-18-5	100	98.57	0.57
G	**8-3-5**	100	97.14	0.54	8-20-5	100	97.14	0.54

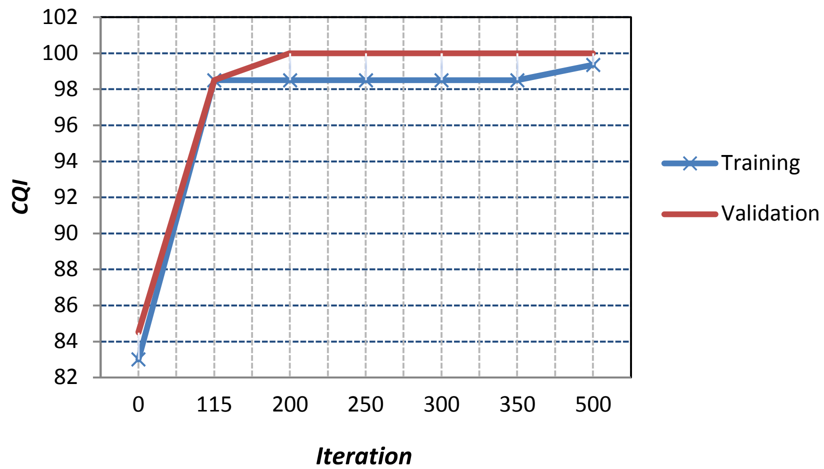

Fig.4. CQI vs iteration for training and validation datasets- A

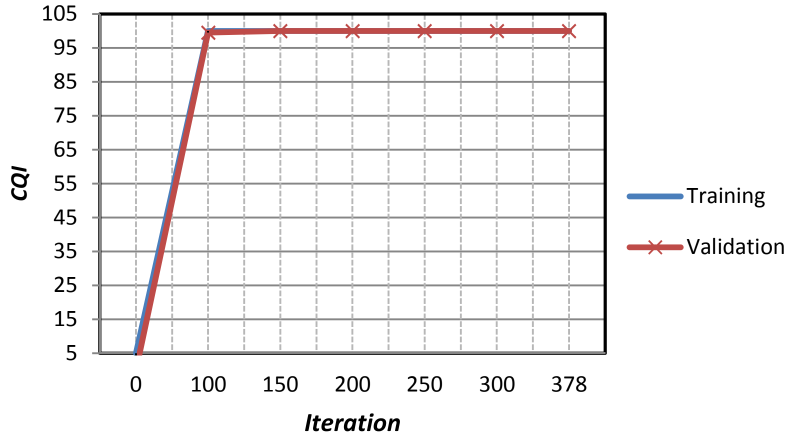

Fig. 5. CQI vs iteration for training and validation datasets- B

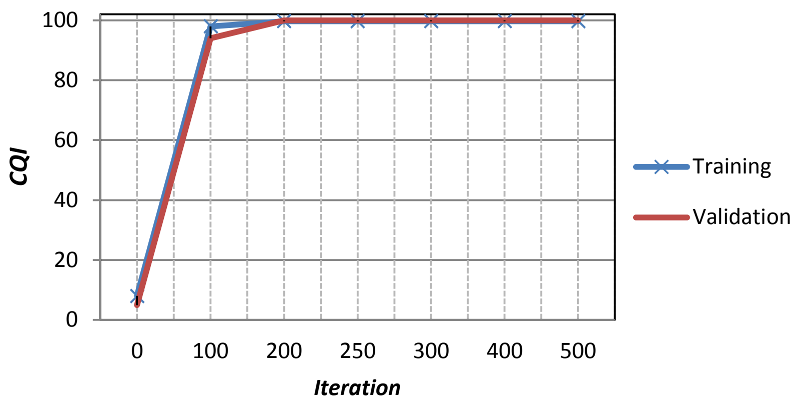

Fig. 6. CQI vs iteration for training and validation datasets- C

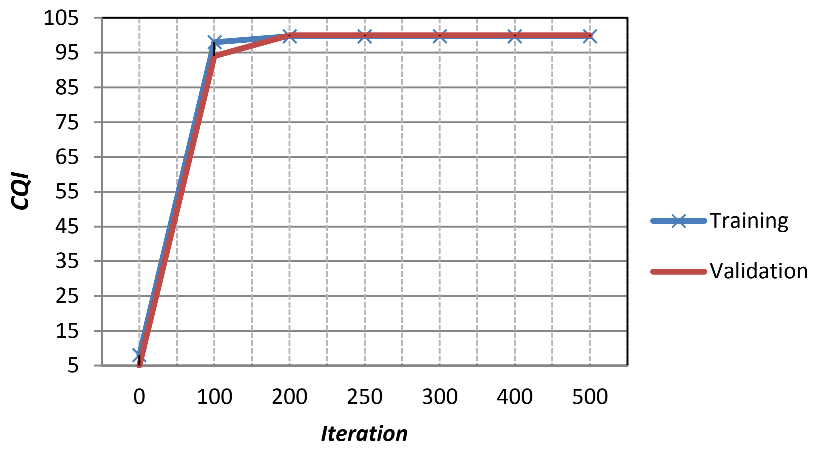

Fig.7. CQI vs iteration for training and validation datasets- D

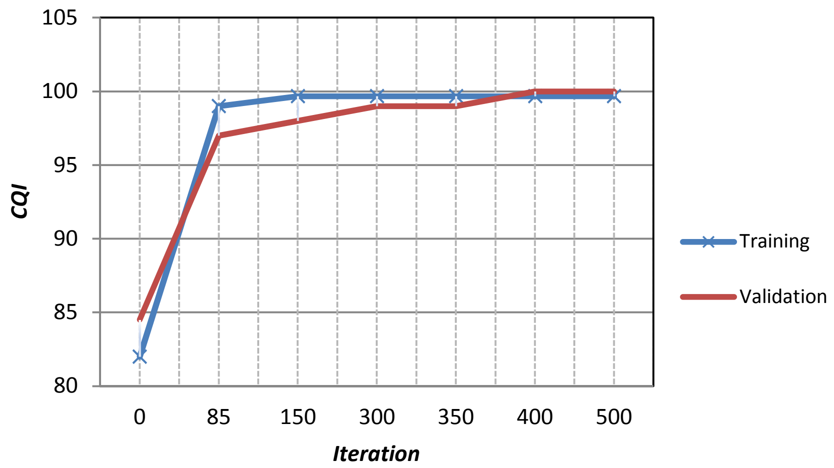

Fig.8. CQI vs iteration for training and validation datasets- E

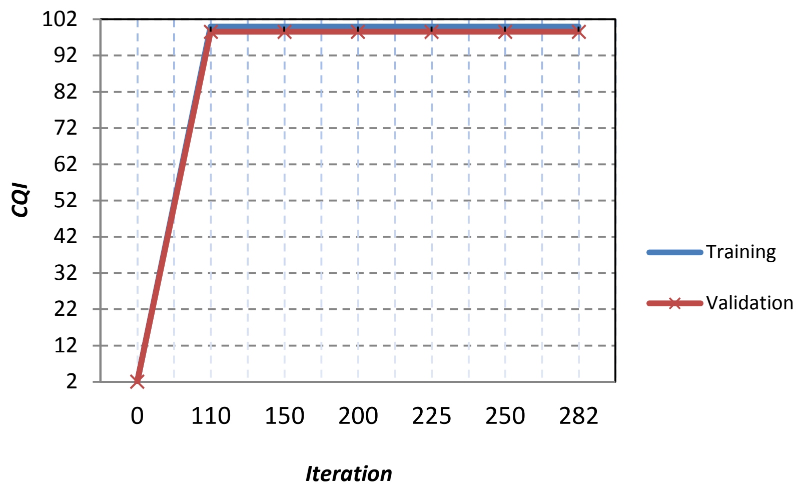

Fig.9. CQI vs iteration for training and validation datasets- F

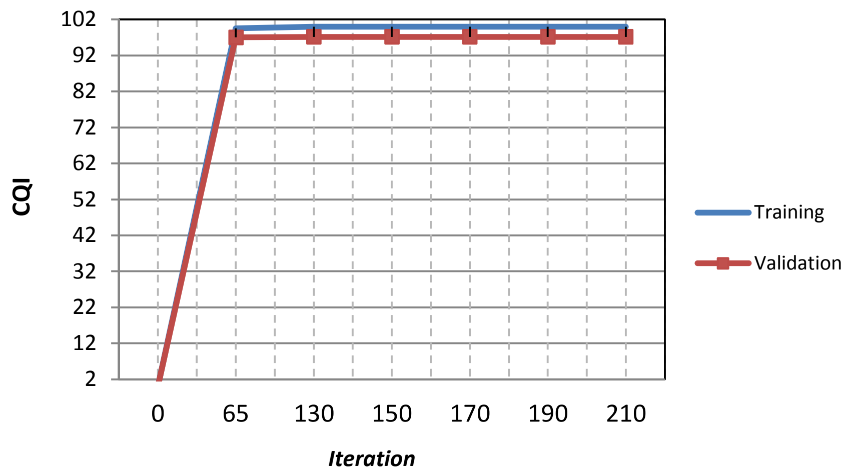

Fig. 10. CQI vs iteration for training and validation datasets- G

4.5 APPLICATION OF NEW MODEL TO YEMENI DATA

Ability of proposed model to classify the reservoir fluid of Yemeni fields was examined. The testing of novel model was based on data collecting from 16 different Yemeni fields.

The tests showed that very good results are obtained with neural networks classification model.

Fig.11 represents analysis of Classification quality index for various data set which shows excellent precision of classification NN model - with high agreement between actual data the corresponding neural network output.

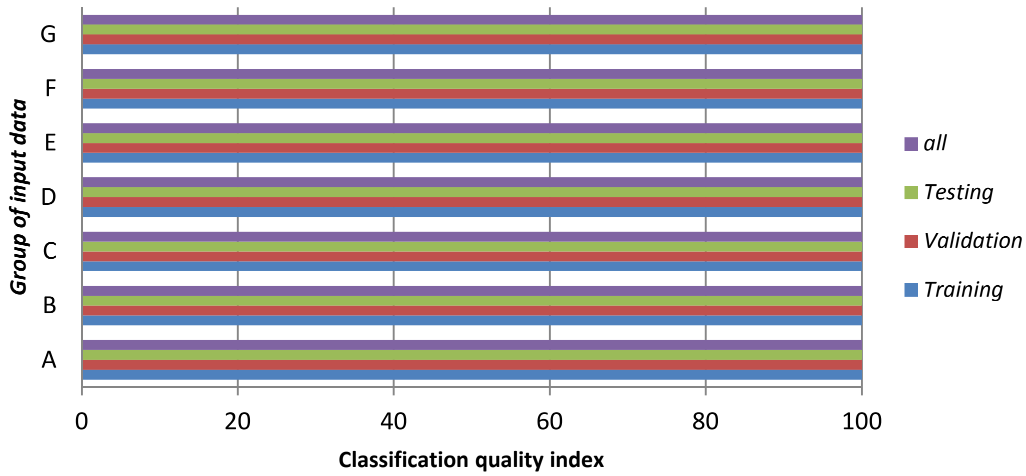

Fig. 11. Classification quality index for various data sets

5 Conclusion

The following general conclusions can be drawn from this study:

- New application for neural network is presented.
- Result show powerful ability and high accuracy of ANNs to recognize reservoir fluid.
- The significance of different input fluid properties in classification was studied.
- The novel model has been successfully applied to classify the reservoir fluid of Yemeni fields.

Nomenclature

TR = temperature, °F
SG = specific gravity of heptanes plus
GOR = gas oil ratio, scf/STB
P_b = bubble point pressure, psia
P_d= dew point pressure, psia
MWC7+= molecular weight of heptanes plus
API = oil gravity in American petroleum institute degree
C_{7+} = mole fraction of heptanes plus
Bo = formation volume factor
CQI= classification quality index
Quick-Quick propagation
Conj-Conjugate Gradient Descent
Quasi-Quasi-Newton
Limit-Limited Memory Quasi-Newton
OnlineBP-Incremental back propagation
BatchBP-Batch back propagation,

Reference

[1] T. Ahmed, Hydrocarbon phase behavior. Gulf Publishing, pp. 244-86, 1989.
[2] J.K. Ali, “Neural networks: a new tool for the petroleum industry,” in: European Petroleum Computer Conference, 1994.
[3] B. Baesens, R. Setiono, C. Mues, J. Vanthienen, “Using neural network rule extraction and decision tables for credit-risk evaluation,” Management Science n°. 49, pp. 312–329, 2003.

[4] J. Benediktsson, P.H. Swain, O.K. ERSoy, “Neural network approaches versus statistical methods in classification of multisource remote sensing data,” IEEE Transactions on geoscience and remote sensing n°. 28, pp. 540–552. 1990.

[5] H. Bischof, W. Schneider, A.J. Pinz, “Multispectral classification of Landsat-images using neural networks. Geoscience and Remote Sensing,” IEEE Transactions on 30, pp.482–490, 1992.

[6] H. C. Chang, Kopaska-Merkel, D.C., Chen, H.-C., Durrans, S, “Lithofacies identification using multiple adaptive resonance theory neural networks and group decision expert system,” Computers & Geosciences n°. 26, 591–601. 2000.

[7] CMOS winner-take-all circuits, “A tutorial article on various types of winner-take-all transistor-level circuits fabricated in the CMOS technology.

[8] A. Danesh, PVT and phase behaviour of petroleum reservoir fluids. Elsevier,1998.

[9] R.O. Duda, Hart, P.E., Stork, D.G, “Pattern classification (2nd edition), Wiley, ISBN 0-471-05669-32001.

[10] Hansen, L.K., Salamon, P, “Neural network ensembles. Pattern Analysis and Machine Intelligence,” IEEE Transactions on n° 12, pp. 993–1001, 1990.

[11] Hong, Y, “Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network-Cloud Classification System, 2003.

[12] J. Wei Khan, J.S, Ringner, M., L.H Saal, M. Ladanyi, F. Westermann, F. Berthold,., M. Schwab,., C.R. Antonescu,., C. Peterson , “Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks,” Nature medicine n° 7, pp.673–679, 2001.

[13] Lawrence, Jeanette Introduction to Neural Networks,” California Scientific Software Press. ISBN 1-883157-00-5,1994.

[14] R.P. Lippmann, “Pattern classification using neural networks. Communications Magazine, IEEE n°. 27, pp. 47–50. 1989.

[15] C.G. Looney, Pattern recognition using neural networks, theory and algorithms for engineers and scientists,” Oxford University Press, Inc,1997.

[16] F.J. Lucia, “Rock-fabric/petrophysical classification of carbonate pore space for reservoir characterization. AAPG Bulletin-American Association of Petroleum Geologists n° 79, pp. 1275–1300,1995.

[17] W.D. McCain, The properties of petroleum fluids, PennWell Books, 1990.

[18] P. Meldahl, R. Heggl and, B. Bril, de Groot, P, “Identifying faults and gas chimneys using multiattributes and neural networks,” The Leading Edge n° 20, pp.474–482, 2001.

[19] Mi, Y. Ishii, M. Tsoukalas L.H, “Vertical two-phase flow identification using advanced instrumentation and neural networks,”Nuclear Engineering and Design 184, pp. 409–420, 1998.

[20] K.S. Pedersen, Christensen, P.L., Azeem, S.J, Phase behavior of petroleum reservoir fluids. CRC Press2006.

[21] Ripley, Brian D, “Pattern Recognition and Neural Networks,”Cambridge,1996.

[22] B.P. West, S.R. May, J.E. Eastwood, C. Rossen, “Interactive seismic facies classification using textural attributes and neural networks, ”The Leading Edge n° 21, pp.1042–1049, 2002.