ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 1 (2012) 1359–1366
Procedia Computer Science
www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2010

Optimization and data mining for fracture prediction in geosciences

Guang-ren Shi[a], Xin-She Yang[b,*]

[a]*Research Institute of Petroleum Exploration and Development, PetroChina, P. O. Box 910, 20 Xueyuan Road, Beijing 100083, P. R. China*
[b]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK*

Abstract

The application of optimization and data mining in databases in geosciences is becoming promising, though still at an early stage. We present a case study of the application of data mining and optimization in the prediction of fractures using well-logging data. We compare various approaches, including multiple regression analysis (MRA), back-propagation neural network (BPNN), and support vector machine (SVM). The modelling problem in data mining is formulated as a minimization problem, showing that we can reduce an 8-D problem to a 4-D problem by dimension reduction. The MRA, BPNN and SVM methods are used as optimization techniques for knowledge discovery in data. The calculations for both the learning samples and prediction samples show that both BPNN and SVM can have zero residuals, which suggests that these combined data-mining techniques are practical and efficient.

Keywords: multiple regression analysis; back-propagation neural network; support vector machine; optimization; dimension-reduction; knowledge discovery; well-logging

1. Introduction

Data mining has become an emerging technology to extract useful patterns and information from massive data sets. In fact, recent reviews suggested that it would become one of the most revolutionary developments over the next decades, and as one of the 10 emerging technologies that will change the world [1, 2, 3]. Over the last 20 years, data mining has seen an enormous success with a wide range of applications, following the latest development of new techniques and theoretical breakthrough [4]. Data mining is the computerized process of extracting previously unknown and important information and knowledge from large databases. Such knowledge can then be used to make crucial but informed decisions. As a result, data mining is also often referred to as the knowledge discovery in data. It has been widely used in many areas, including business, finance, marketing, image processing, pattern recognition, and scientific discovery. However, its application in geosciences for spatial data mining is still at a very early stage [5, 6, 7].

To deal with the large amount data and efficiently extract useful information from massive databases in geosciences, geoscientists can often resort to using conventional database management systems to conduct

* Guang-ren Shi. Tel.: +86-10-8359-7093; fax: +86-10-8359-7108.
E-mail address: grs@petrochina.com.cn .

doi:10.1016/j.procs.2010.04.151

applications (such as query, search and simple statistical analysis). Consequently, it is extremely challenging to obtain the knowledge structure and patterns inhered in data, which leads to a dilemma of 'rich data, but poor knowledge'. A promising solution is to apply data mining techniques in database management in geosciences, and to develop new techniques when necessary and appropriate.

The main aim of this paper is to introduce the fundamentals of data mining and its applications in geosciences. We will use a unified approach to view the data mining techniques as an optimization and data-mining framework. More specially, we will consider data mining, support vector machine, and back-propagation neural network as an integrated optimization process for knowledge discovery. We will then use these techniques to predict or discover fractures in reservoirs for well-logging interpretation as our case study. The case study will focus on the fracture prediction and comparison with field data in the Anpeng Oilfield of the Biyang Sag, Nanxiang Basin, central China. We will then evaluate the performance, feasibility and practicability of these techniques in knowledge discovery in geosciences.

2. Data Mining and Optimization

2.1. Data mining

In order to make proper knowledge discovery using data mining techniques, some appropriate pre-processing should be used to assemble the suitable target data sets and the right objective. Generally speaking, data mining consists of three main steps: data preprocessing, knowledge discovery in data, and knowledge application. Data preprocessing involves data selection, data cleaning, dandling missing data, identifying misclassifications, identifying outliers, data transformation, and min–max normalization. Knowledge discovery in data is the mainly process of choosing the right mathematical model $y = \phi(\boldsymbol{x})$ so as to find the right combination of x_i to best-fit the data by appropriate dimension-reduction algorithms. This also involves the step to select an efficient algorithm for data mining so as to discover new knowledge. This is largely a learning or training process. The next important step is knowledge application of the discovered knowledge model $y = \phi(\boldsymbol{x})$ to make new predictions.

Access to a conventional database is executed by a database management system. At the knowledge discovery stage, specific learning samples are extracted from the conventional database. Any new knowledge can be incorporated into the conventional database.

For most classification purposes in data mining, the data samples are divided into two categories: the learning or training samples, and the prediction samples. Each training data point contains a response (or output) y and a set of m inputs or factors $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$. Therefore, the training set can be written as $(\boldsymbol{x}_i, y_i)$ where $i = 1, 2, \ldots, n$. Mathematically, it requires that $n > m - 1$, and in practice $n >> m - 1$.

The main objective of data mining is to choose such techniques so that the overall errors (training errors and prediction errors) should be minimal. The errors, especially the training errors, can be measured using the residual sum of squares

$$\sum_{i=1}^{n} [\phi(\boldsymbol{x}_i) - y_i]^2, \tag{1}$$

which can be minimized using many techniques such as the method of least-squares. However, for classification and data mining, this is often inadequate. We have to consider model complexity and margins. Therefore, such objectives should be modified accordingly, depending on the actual techniques and formulations.

2.2. Techniques for data mining

In general, data mining can be applied to carry out classification, clustering, associate learning, and regression. Therefore, there are in general three classes of data mining techniques: dimension-reduction, classification, and regression [2, 3, 8]. However, such division is relatively arbitrary, though widely used in literature, because some techniques such as support vector machine (SVM) can be used for both classification and regression.

For dimension reduction, multiple regression analysis (MRA) is the most widely used. For classifications, common techniques include decision-tree analysis, Bayesian classification, rule-based classification, associative classification, k-nearest-neighbour classification, genetic algorithms, rough set approach, fuzzy set approach, back-propagation neural network (BPNN), and support vector machine (SVM). For the purpose of regression, popular techniques include linear regression, nonlinear regression, logistic regression, Poisson regression, regression trees, model trees, BPNN, and SVM.

In the case study to be discussed later, we will use all the three major methods: MRA, BPNN, and SVM to carry out the analyses for the same known parameters. This way, we can evaluate and then compare their performance for fracture predictions.

2.3. An optimization framework

For multiple regression analysis, the generic model can be written as

$$y = \phi(\boldsymbol{x}, \boldsymbol{\beta}), \tag{2}$$

where $\boldsymbol{\beta}$ is the vector of parameters for regression. The aim is to minimize the residual sum of the squares

$$\sum_{i=1}^{n} e_i^2 , \tag{3}$$

where $e_i = \phi(\boldsymbol{x}_i, \boldsymbol{\beta}) - y_i$ is the residual for each data point. This method is straightforward to implement and can provide insights into the selection of appropriate factors. However, the disadvantage of this method is that it often tends to over-fit the data and may bring in unnecessary factors into the equation, even though the actual geological process may not have direct dependence on these factors.

A substantial improvement is to use a support vector machine as it has been indicated by Shi [9] that SVM is superior to other methods under certain conditions. This can be understood from the optimization point of view. The mathematical model for SVM often takes the following form

$$y = \phi(\boldsymbol{x}, \boldsymbol{w}) = \boldsymbol{w} \cdot \boldsymbol{x} + b = \boldsymbol{w}^T\boldsymbol{x} + b, \tag{4}$$

where $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ is a vector of undetermined parameters. For classifications, $\boldsymbol{w}$ represents the reciprocal of the separation distances between the hyperplanes, therefore, a good classification requires to minimize $||\boldsymbol{w}||$. This can often be written as

$$\psi(\boldsymbol{w}, b) = \frac{1}{2} || \boldsymbol{w} ||^2 + \lambda \sum_{i=1}^{n} L\left[y_i \phi(\boldsymbol{x}_i, \boldsymbol{w}) \right] , \tag{5}$$

where $||\boldsymbol{w}||^2$ is l_2-norm or the Euclidean norm, and λ is a Lagrange-type parameter. $L(u)$ is the hinge-loss function. That is $L = \max(0, 1 - u)$. The first term in the above equation is to maximize the margins, while the second term means to minimize the training errors. This objective has to be subject to the constraint

$$y_i (\boldsymbol{w}^T\boldsymbol{x}_i + b) \geq 1. \tag{6}$$

Now the aim is to find the optimal solution $\boldsymbol{w}$ and b to the above optimization problem. It is possible to convert this optimization into a quadratic programming problem which can be solved efficiently. However, a kernel technique is widely used, which is often formulated in terms of kernel functions K

$$\boldsymbol{y} = \sum_{i=1}^{n} \alpha_i y_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b, \qquad \sum_{i=1}^{n} \alpha_i y_i = 0, \tag{7}$$

where $0 \leq \alpha_i \leq C$ are the Lagrange multipliers and C is a penalty constant [9, 10]. Here $K(\boldsymbol{x}_i, \boldsymbol{x})$ is the kernel function. For linear problems, a linear kernel $K(\boldsymbol{x}_i, \boldsymbol{x}) = \boldsymbol{x}_i^T\boldsymbol{x}$ can be used, while for most problems, we can use the radial basis function (RBF) as the kernel:

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = \exp(-\gamma\|\boldsymbol{x}_i - \boldsymbol{x}\|^2), \quad (8)$$

where $\gamma > 0$ is a parameter. This RBF kernel is one of the most widely used and often leads to robust results.

The back-propagation neural network is a back-propagation algorithm in which the artificial neurons are arranged in multiple layers. It uses supervised learning in which the algorithm computes the weights w_{ij} by matching the training inputs and outputs. The errors between the training data sets and the actual outputs are computed, and the objective is to minimize these errors until the neural network has really 'learnt' the training data by adjusting its weights to be optimal.

We can see from the above description that all the three methods: MRA, SVM and BPNN intend to minimize certain objectives such as errors by training the model to learn and best-fit to the known data sets. This means that we can put all these methods in an integrated optimization framework. In the rest of the paper, we will use these methods to predict fractures in reservoirs from well-logging data. We will then compare their performance and discuss the relevant results.

3. Fracture Prediction

3.1. Well-Logging Data

The understanding of fracture formation in reservoirs is crucially important because of its direct link with the oil and gas formation. The objective of this case study is to predict fractures using conventional well-logging data, which has important practical applications, especially in the case when imaging log and core sample data are sparse or limited.

Located in the southeast of the Biyang Sag in the Nanxiang Basin, the Anpeng Oilfield covers an area of about 17.5 km^2, and is close to the Tanghe-zaoyuan northwest-west striking large boundary fault in the south, and a deep sag in the east. As an inherited nose-structure plunging from northwest to southeast, this oilfield has a simple structure without any faults, where commercial oil and gas flows have been discovered [11, 12]. One of its favourable pool-forming conditions is that fracturing is found to be well-developed in the formations, deeper than 2,800 m. These fractures provide favourable oil-gas migration pathways, and also enlarge the accumulation space.

From the data of 7 well logs and 1 imaging log in Wells An1 and An2 of the Anpeng Oilfield, 33 samples were selected [13], in which 29 samples were taken as the training set and 4 samples for prediction (Table 1). We will use all MRA, BPNN and SVM, respectively, to predict the fracturing status.

Let x_1 be the acoustic log (Δt), x_2 be the compensated neutron density log (ρ), x_3 be the compensated neutron porosity log (φ_N), x_4 be the micro-spherically focused log (R_{xo}), x_5 be the deep laterolog (R_{LLD}), x_6 be the shallow laterolog (R_{LLS}), and x_7 be the absolute difference of R_{LLD} and R_{LLS} (R_{DS}). Let y be the predicted fracture as the response. The data of the 7 well logs were normalized over the interval [0, 1]. In the learning samples, y as the input data is denoted as y^*, and is determined by fracture identification of the imaging log (*IL*). In the prediction samples, y is not the input, but rather it is the predicted data, determined by the methods of BPNN and SVM (Table 1).

3.2. Data preprocessing

Using the 29 samples as the learning set (Table 1) and MRA [9, 14, 15, 16], we found that the obtained ordered (successively discovered) factors or variables are: x_5, x_2, x_4, x_7, x_3, x_6, x_1; and their corresponding mean square errors are: 0.48623, 0.31563, 0.29336, 0.27841, 0.25372, 0.25209, 0.25054, respectively. We can see that the values of the last three mean square errors are approximately the same, around 0.25, and the correlations between two independent variables (x_6 and x_1) and the fracture value y are very low. Consequently, the two independent variables x_6 and x_1 can safely be deleted in data mining. If so, the number of independent variables is reduced from 7 to 5, and the 8-D problem (x_1, x_2, x_3, x_4, x_5, x_6, x_7, y) essentially becomes a 6-D problem (x_2, x_3, x_4, x_5, x_7, y). That is the essence of data dimension-reduction.

Table 1. Parameters and calculation results for fracture prediction of Wells An1 and An2 in the Anpeng Oilfield of the Biyang Sag, Nanxiang Basin, central China

Sample type	Sample no.	Well no.	Depth/m	Sample parameters Δt	ρ	φ_N	R_{xo}	R_{LLD}	R_{LLS}	R_{DS}	IL	Predicted fracture results: BPNN 8-D program (x_1, x_2, x_3, x_4, x_5, x_6, x_7, y) by 8465 iterations	BPNN 4-D program (x_2, x_4, x_5, y) by 18059 iterations	SVM (for both 8-D and 4-D programs)
				x_1	x_2	x_3	x_4	x_5	x_6	x_7	y^*	y	y	y
Learning sample	1		3065.13	0.5557	0.2516	0.8795	0.3548	0.6857	0.6688	0.0169	1	1	1	1
	2	An1	3089.68	0.9908	0.0110	0.8999	0.6792	0.5421	0.4071	0.1350	1	1	1	1
	3		3098.21	0.4444	0.1961	0.5211	0.7160	0.7304	0.6879	0.0425	1	1	1	1
	4		3102.33	0.4028	0.3506	0.5875	0.6218	0.6127	0.5840	0.0287	1	1	1	1
	5		3173.25	0.3995	0.3853	0.0845	0.5074	0.8920	0.8410	0.0510	1	1	1	1
	6		3180.37	0.6117	0.6420	0.0993	0.6478	0.9029	0.8511	0.0518	1	1	1	1
	7		3202.00	0.6463	0.5205	0.5351	0.7744	0.2919	0.3870	0.0951	2	2	2	2
	8		3265.37	0.4154	0.9545	0.4397	0.6763	0.2906	0.5173	0.2267	2	2	2	2
	9		3269.87	0.7901	0.6601	0.1487	0.8994	0.9257	0.9325	0.0068	1	1	1	1
	10		3307.87	0.7162	0.1475	0.4481	0.9164	0.7827	0.7992	0.0165	1	1	1	1
	11		3357.37	0.5546	0.4778	0.0741	0.7725	0.9756	0.9237	0.0519	1	1	1	1
	12		3377.03	0.4909	0.3654	0.1816	0.7625	0.8520	0.8237	0.0283	1	1	1	1
	13		3416.48	0.2567	0.5843	0.2043	0.3412	0.7369	0.7454	0.0085	2	2	2	2
	14		3445.37	0.0944	0.9818	0.5124	0.7614	0.5943	0.6321	0.0378	2	2	2	2
	15		3446.12	0.5215	0.8091	0.7594	0.6924	0.7186	0.7572	0.0386	2	2	2	2
	16		3485.25	0.9443	0.2647	0.9904	0.4794	0.4189	0.4776	0.0587	2	2	2	2
	17		3575.00	0.2078	0.0000	0.0358	0.8246	0.9872	0.9800	0.0072	1	1	1	1
	18		3645.00	0.1193	0.6953	0.8879	0.7839	0.8323	0.8409	0.0086	1	1	1	1
	19		3789.37	0.0579	0.6889	0.9418	0.7261	0.8902	0.8947	0.0045	1	1	1	1
	20	An2	992.795	0.3471	0.9624	0.3848	0.6115	0.8245	0.8388	0.0143	2	2	2	2
	21		1525.37	0.5256	0.3256	0.0821	0.7450	0.9888	0.9234	0.0654	1	1	1	1
	22		1527.25	0.0753	0.5441	0.1345	0.6750	0.8468	0.9255	0.0787	1	1	1	1
	23		1867.12	0.3145	0.1325	0.0368	0.5744	0.9425	0.8547	0.0878	1	1	1	1
	24		1880.00	0.7755	0.8347	0.5546	0.4578	0.1894	0.4265	0.2371	2	2	2	2
	25		2045.87	0.4928	0.2110	0.5977	0.6892	0.7411	0.6071	0.1340	1	1	1	1
	26		2085.25	0.8678	0.0833	0.9997	0.4085	0.1973	0.4117	0.2144	2	2	2	2
	27		2112.13	0.5467	0.2961	0.8235	0.7250	0.6328	0.6825	0.0497	1	1	1	1
	28		2355.37	0.4524	0.3426	0.6005	0.7658	0.8992	0.8346	0.0646	1	1	1	1
	29		2358.00	0.6463	0.5205	0.5351	0.7744	0.2919	0.3870	0.0951	2	2	2	2
Prediction sample	30	An1	3164.00	0.5300	0.3333	0.0758	0.8939	0.9918	0.9863	0.0055	(1)	1	1	1
	31		3166.50	0.5282	0.4589	0.0459	0.7140	1.0000	1.0000	0.0000	(1)	1	1	1
	32	An2	980.485	0.2024	0.4288	0.2149	0.5581	0.8489	0.8504	0.0015	(1)	1	1	1
	33		987.018	0.0631	0.5278	0.3450	0.7403	0.7368	0.7295	0.0073	(1)	1	1	1

y^* = 1, fracture; y^* = 2, nonfracture; the numbers in parenthesis are not input data,, but only for checking the predicted results.

The advantages of data dimension-reduction are to decrease the number of factors and reduce the amount of data so as to speed up the data mining process. The relationship $y = \phi(\boldsymbol{x})$ is usually nonlinear for most subsurface studies in geosciences; however, the equation constructed by MRA is typically a linear function, so we should carry out the dimension-reduction using nonlinear algorithms such as BPNN and SVM.

According to the order of successively introduced independent variables by MRA, we can delete x_1 firstly, to reduce the 8-D problem (x_1, x_2, x_3, x_4, x_5, x_6, x_7, y) to a 7-D problem (x_2, x_3, x_4, x_5, x_6, x_7, y) if the results by BPNN and SVM remain correct. We can hen proceed to delete x_6 so as to reduce the 7-D problem (x_2, x_3, x_4, x_5, x_6, x_7, y) to a 6-D problem (x_2, x_3, x_4, x_5, x_7, y) if the results of both BPNN and SVM are still correct. We can proceed in a similar manner. In the end, the results indicate that the original 8-D problem (x_1, x_2, x_3, x_4, x_5, x_6, x_7, y) can be reduced to a 4-D problem (x_2, x_4, x_5, y).

It is worth pointing out that the learning time counts or iterations used in BPNN are 8465, 8791, 9197, 12311 and 18059 for 8-D, 7-D, 6-D, 5-D and 4-D problems, respectively. In addition, the calculated results y by both BPNN and SVM are exactly the same as y^* (Table 1) with zero residuals.

3.3. Knowledge discovery

Now we can solve the 4-D problem to predict the fracturing by BPNN and SVM. Using the 29 learning samples (Table 1) and the methods of BPNN [9, 14] and SVM [9, 10, 17, 18, 19], the relationship between the predicted y and the three well logs (x_2, x_4, x_5) has been correlated.

3.3.1. BPNN

The actual BPNN method consists of three nodes in the input layer, one output node in the output layer, and 7 nodes in the hidden layer; the value of the network learning rate for the output layer and the hidden layer is 0.6, and the learning time count is 18059. This results in an implicit relationship:

$$y = BPNN(x_2, x_4, x_5), \tag{9}$$

where *BPNN* is a nonlinear function. Equation (9) typically yields zero residuals, indicating a very good fitting.

3.3.2. SVM

Using a C-SVM binary classifier with a RBF (radial basis function) kernel function, we have $C = 512$ and $\gamma = 0.007813$. The cross-validation accuracy is 100%. The result obtained is an explicit expression:

$$y = SVM\,(x_2, x_4, x_5), \tag{10}$$

where *SVM* is a nonlinear function that can in principle be expressed as a mathematical formula. However, this actual formula is not reproduced here, as it is lengthy and only specific to this particular case study. Similarly, equation (10) also yields zero residuals, indicating a very good fitting.

Both equations (9) and (10) are new knowledge discovered by data mining techniques BPNN and SVM, respectively.

3.4. Knowledge Application for Fracture Prediction

Substituting the independent variables determined from the 29 learning samples in $BPNN(x_2, x_4, x_5)$ and $SVM(x_2, x_4, x_5)$, respectively, we can obtain the predicted fracturing, represented as a value, for both methods, and these predictions can then be used to verify the accuracy of the prediction by each method.

The predicted values for the fractures of the four prediction samples have also been summarized in Table 1. We can see that all the predicted results for y by both BPNN and SVM are exactly the same as y^* (Table 1) with zero residuals. These accurate predictions imply that the learned model equations (9) and (10) are both correct and accurate.

4. Conclusions

We have formulated an integrated framework which views the three major techniques (MRA, BPNN, SVM) as an optimization problem for knowledge discovery in data mining. We then used these three methods to predict fractures in reservoirs successfully, based on the well-logging data in the Anpeng Oilfield of the Biyang Sag, Nanxiang Basin, central China. From our simulations and predictions, we can draw the following conclusions:

- Knowledge discovery in data mining such as fracture prediction can be formulated as a minimization problem, and can be solved by either optimization or data mining techniques efficiently;
- Data dimension-reduction can efficiently be carried out by a good combination of MRA, BPNN and SVM. In our case study, the original 8-D problem can effectively be reduced to a much smaller 4-D problem, which can significantly reduce the amount of data and subsequently increase the speed and efficiency of data mining;
- Knowledge discovery and knowledge application by both BPNN and SVM are effective and practical. Even in strongly nonlinear cases, SVM is superior to BPNN [9].

However, in the present case study, both BPNN and SVM have zero residuals, which can be attributed to the proper selection of well logs, the good quality of well-logging, and probably no strong nonlinearity.

Therefore, a good combination of data-mining techniques can provide efficiently data mining for knowledge discovery as well for prediction. The application of optimization and data mining in geosciences databases can be very promising in dealing with large data sets and modelling complex systems. Further studies will be focused on the systematic comparison of various new techniques in data mining and their applications in geosciences.

References

1. D. Hand, H. Mannila and P. Smyth, Principles of Data Mining, MIT Press., Cambridge, MA, USA, 2001.
2. D. T. Larose, Discovering Knowledge in Data, John Wiley & Sons, Inc., New York, USA, 2005.
3. D. T. Larose, Data Mining Methods and Models, John Wiley & Sons, Inc., New York, USA, 2006.
4. H. Hirsh, Data mining research: Current status and future opportunities, Statistical Analysis and Data Mining, 1(2), 2008, 104–7.
5. P. M. Wong, A novel technique for modeling fracture intensity: A case study from the Pinedale anticline in Wyoming, AAPG Bulletin, 87 (11), 2003, 1717–27.
6. F. Aminzadeh, Applications of AI and soft computing for challenging problems in the oil industry, Journal of Petroleum Science and Engineering, 47(1-2), 2005, 5–14.
7. S. D. Mohaghegh, A new methodology for the identification of best practices in the oil and gas industry, using intelligent systems, Journal of Petroleum Science and Engineering, 49(3-4), 2005, 239–60.
8. J. W. Han and M. Kamber, Data Mining: Concepts and Techniques (2nd ed.), Morgan Kaufmann, San Francisco, USA, 2006.
9. G. R. Shi, The use of support vector machine for oil and gas identification in low-porosity and low-permeability reservoirs, *Int. J. Mathematical Modelling and Numerical Optimisation*, 1(1/2), 2009, 75–87.
10. V. N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, USA, 1995.
11. H. H. Ming, Z. K. Jin, Q. Z. Li, Y. M. Qu and X. Chen, Sedimentary facies of deep sequences of Anpeng Oilfield in Biyang Depression and its control over oil-gas distribution, Journal of Earth Sciences and Environment, 27 (2), 2005, 48–51.
12. G. P.Wang, Q. J. Guo, Q. Zhao, Y. Z. Gong and L. Zhao, Effective fracturing cut-off values of extremely-low permeability reservoir—take the Anpeng Oilfield of Biyang Depression as an example, Journal of Southwest Petroleum University, 28(4), 2006, 40–3.
13. H. L. Shen and S. Y. Gao, Research on fracture identification based on BP neural network, Fault-Block Oil and Gas Field, 14(2), 2007, 60–2.
14. G. R. Shi, New Computer Application Technologies in Earth Sciences, Petroleum Industry Press, Beijing, China,1999.
15. G. R. Shi, X. X. Zhou, G. Y. Zhang, X. F. Shi and H. H. Li, The use of artificial neural network analysis and multiple regression for trap quality evaluation: A case study of the Northern Kuqa Depression of Tarim Basin in western China, Marine and Petroleum Geology, 21(3), 2004, 411–20.
16. X. S. Yang, Introduction to Computational Mathematics, World Scientific Publishing, Singapore, 2008.
17. N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Cambridge, UK, 2000.

18. B. Schölkopf, A. J. Smola, R. C. Williamson and P. L. Bartlett, New support vector algorithms, Neural Computation, 12(5), 1207–45, 2000.
19. C. C. Chang and C. J. Lin, LIBSVM: A Library for Support Vector Machines, Version 2.89 (2009), available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/.