Support Vector Machine Approach and Petroleum Engineering

Seyed Reza Na'imi[1], Seyed Reza Shadizadeh[1], Mohammad Ali Riahi[2], Marzieh Mirzakhanian[3]

[1] Department of Petroleum Exploration Engineering, Abadan Facility of Petroleum University of Technology, Petroleum University of Technology, Abadan, Iran
[2]Geophysics Institute, University of Tehran, Tehran, Iran
[3]Geophysics Office, Exploration Directorate of National Iranian Oil Company, Tehran, Iran
shadizadeh@put.ac.ir

Abstract: Support Vector Machine is a supervised computer learning algorithm which is originated from Statistical Learning Theory and is used for both classification and regression tasks in wide variety of engineering problems. SVM implementations show that it gives rise to more accurate results rather than neural networks and statistical methods in most applications. Furthermore, Support Vector Machine is more convenient for situations where the populations are small and non-linear. The basic ideas behind the Support Vector Machine algorithm, however, can be explained without ever reading an equation. So in this paper, a brief description of Support Vector Machine method is first brought and after that some important implementations in petroleum engineering are discussed shortly.

[Na'imi S R, Shadizadeh S R, Riahi M A, Mirzakhanian M. **Support Vector Machine Approach and Petroleum Engineering.** *J Am Sci* 2012;8(12):309-314]. (ISSN: 1545-1003). http://www.jofamericanscience.org. 44

Keywords: Support Vector Machine, Support Vector Classification, Support Vector Regression, Kernel Function

1. Introduction

The last decade has witnessed noticeable progresses in the study and application of intelligent systems as robust tools for extracting quantitative formulation between two sets of data (inputs/outputs) that have an underlying dependency in the petroleum industry. Several studies done by open-minded researchers indicate that intelligent systems are in the vanguard of potent tools for solving complicated petroleum problem in both regression and classification approaches (Todorov et al., 1997; Cuddy, 1998; Balch et al., 1999; Trappe and Hellmich, 2000; Nikravesh and Aminzadeh, 2001; Wong and Nikravesh, 2001; Russell et al., 2003; Nikravesh and Hassibi, 2003; Saggaf and Nebrija, 2003; Russell, 2004; Soubotcheva and Stewart 2004; Zahuczki and Barany 2005; Aristimuno and Aldana, 2006; Soubotcheva and Stewart, 2006; Kadkhodaie-Ilkhchi et al., 2009).

Arelatively new promising method for learning separating functions in pattern recognition (classification) tasks or for performing functional estimation in regression problems is the Support Vector Machine (SVM) which is originated from Statistical Learning Theory (SLT) developed by Vapnik and Chervonenkis. SVMs are supervised machine learning algorithm that have been introduced in the framework of Structural Risk Minimization (SRM) and in the theory of Vapnik-Chevronenkis (VC) bound and are especially suitable for use with non-linear multiattributes. For the cases where the populations are small (i.e., only a few well-seismic attribute pairs) statistical significance may be impossible to achieve and neural network can be easily over trained and result in "over fitting" and poor predictions in validation trials, SVM has good performance on unseen data (good generalization). (Li et al., 2000; Lu et al., 2001; Choisy and Belaid, 2001; Gao et al., 2001; Kim et al., 2001; Ma et al., 2001; Van Gestel et al., 2001; Li, 2005, ; Yanzhou et al., 2010).

The basics of the SVM algorithm, however, can be explained without ever reading an equation and one need only to grasp basic concepts. Therefore, in this paper, a brief description of SVM methods and its fundamental concepts are clarified first and subsequently, some important implementations in petroleum engineering are shortly discussed.

2. Support vector Machine

Experimental data modeling is relevant to many engineering applications. For build up model of a system, the process of induction is used and it is tried to infer the response that have to be observed. Because of its observational nature, data obtained is finite and sampled; typically this sampling is non-uniform and due to the high dimensional nature of the problem the data will form only a sparse distribution in the input space (Gunn, 1998). Traditional neural network approaches are objected to difficulties with generalization, yielding models that can overfit the data. This is because of the optimization algorithms used for parameter selection and the statistical measures used to select the 'best' model (Gunn, 1998).

A relatively new promising method for learning separating functions in pattern recognition (classification) tasks or for doing functional estimation in regression problems is the SVM which is originated from Statistical Learning- learning from experimental data- Theory (SLT) developed by Vapnik and Chervonenkis. Like neural networks and decision tree learning, SVMs are supervised machine learning algorithm i.e. it is a machine learning technique for creating a function from training data. The task of the supervised learner is to predict the value of the function for any valid input point after having seen a finite number of training examples (Parrella., 2007). The formulation of SVM sustains the Structural Risk Minimization (SRM) principle, which minimizes an upper bound on the expected risk and has been shown to be superior to traditional Empirical Risk Minimization (ERM) principle which minimizes the error on the training data and used by conventional neural networks (Gunn et al., 1997, Gunn., 1998). This difference which is the destination in statistical learning causes greater ability of SVM for generalization tasks.

SVMs were first formulated for the classification problem solving, but recently they have been developed to the regression problems (Vapnik et al., 1997). The term SVM is typically used to classification description with support vector methods and support vector regression is used to describe regression with support vector methods. However, in this study the term SVM will refer to both classification and regression methods, and the terms Support Vector Classification (SVC) and Support Vector Regression (SVR) will be used for specification.

There are some principles, such as Structural Risk Minimization and VC Dimensions which are the basis of SVM but they may be not interesting and necessary to who just wants use and apply SVM Algorithms. So here, they are omitted and for detailed description, these issues are referred: Weston, 1998; Gunn, 1998; Kecman, 2001; Campbell and Ying, 2011. Furthermore, it is declared that, to understand the essence of SVM, the grasp of four basicconcepts is necessary: (i) the separating hyperplane, (ii) the maximum-margin hyperplane, (iii) the softmargin and (iv) the kernel function (Noble, 2006).

In SVM literature, a predictor variable is called an *attribute*, and a transformed attribute that is used to define the hyperplane is called a *feature*. (DTREG Tutorial, 2003)

The fundamental of the SVM is to map the original data into a higher dimensional feature space by Kernel functions. The goal of SVC modeling is to find the optimal hyperplanein feature space, that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane(Hawley, Madden, 2005). In SVR the goal is to find a regression function which has at the most ε deviation from actually obtained targets of all training data in feature space (Fu and Cheng, 2011).

2.1 Support Vector Classification (SVC)

Assume the task is to do classification for a simple 2-dimensional example and the data has a categorical target variable with two categories (e.g. Limestone and Sandstone) and there are two predictor variables with continuous values (e.g. two seismic attributes). If the data points are plotted using the value of one predictor on the X axis and the other on the Y axis it may be like the Figure 1. One category of the target variable is represented by rectangles while the other category is represented by ovals. In this idealized example, the cases are completely separated. The SVM analysis attempts to find a 1-dimensional hyperplane (i.e. a line) that separates the cases based on their target categories. There are an infinite number of possible lines; two candidate lines are shown in Figure1.

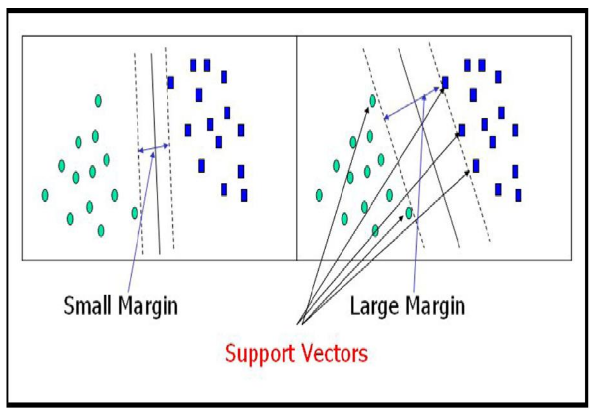

Figure 1. Data separation in SVC for two dimensional and margin illustration (DTREG Tutorial, 2003).

The dashed lines drawn parallel to the separating line, mark the distance between the dividing line and the closest vectors to the line. The distance between the dashed lines is called the margin. The vectors (points) that constrain the width of the margin are the support vectors. An SVC analysis finds the line (or, in general, hyperplane) that is oriented so that the margin between the support vectors is maximized. In the Figure 1, the line in the right panel is superior to the line in the left panel. (DTREG Tutorial, 2003)

Unfortunately, this is not commonly the case, so SVM must deal with (1) more than two predictor variables, (2) separating the points with non-linear curves, (3) handling the cases where clusters cannot be completely separated, and (4) handling classifications with more than two categories. (DTREG Tutorial, 2003)

In the cases with three predictor variable, the third value can be plotted on a third dimension so the plot ofthe points is in a 3-dimensional cube (case 1) and the points can be separated by a 2-dimensional plane.In general, the data points in N-dimensional space can be separated by a (N-1)-dimensional hyperplane.

When the two set data cannot be separated with a straight line, flat plane or an N-dimensional hyperplane and they just can be divided by a nonlinear region in input space (case 2), SVM utilizes *kernel function* to map the data into a another space which names feature space and a hyperplane can be used to do the separation there and this results in a nonlinear curves fitting to the data in original input space (Figure 2). Kernel mapping function which is very powerful concept allows SVC to perform separation even with very complex boundaries (Weston, 1998).

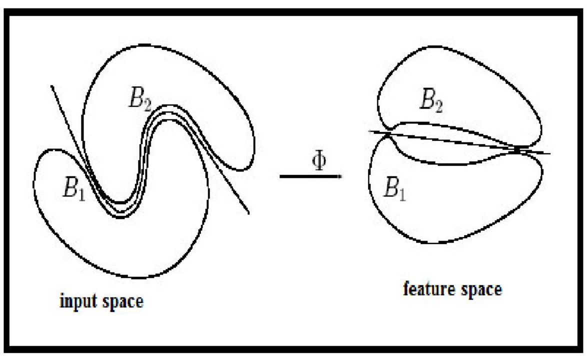

Figure 2. Mapping data from input space into feature space by Kernel function (Φ). B_1 and B_2 are two classes of data. (DTREG Tutorial, 2003)

Where the data set contains error- like many real datasets-, perfect separation may not be possible, or it may result in a model with so many feature vector dimensions which cause over fitting i.e. the model does not generalize well to other data. For these cases, SVC algorithm was customized by adding a soft margin term to be able to consider the errors in the data by allowing a few false expression profiles to be on the wrong side of the separating hyperplane without affecting the final result (case 3). For preventing the occurrence of too many misclassifications, a user specified parameter, C, which is called cost parameter is introduced that controls, roughly, how many examples are permitted to violate the separating hyperplane and how far across the hyperplane they are allowed to go. This soft margin parameter specifies a trade-off between hyperplane violations and the size of the margin and permits some misclassifications. The error for misclassified point is the distance from the point to the hyperplane multiplied by the cost factor C and technically, C is the cost of the sum of the distances of wrong-size points from the margins. Increasing the value of C increases the cost of misclassifying points and forces the creation of a more accurate model that may not generalize well (DTREG Tutorial, 2003).

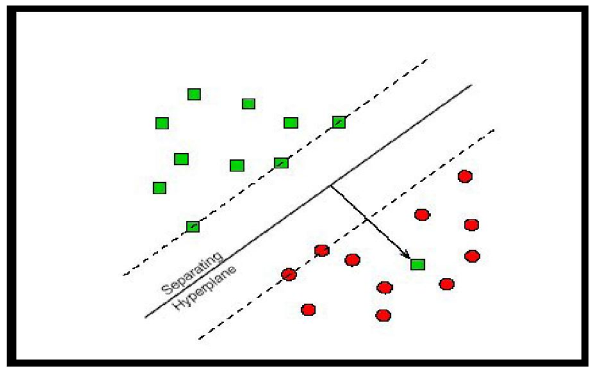

Figure 3: non separable data sets. use line separation but admit training errors. (DTREG Tutorial, 2003)

Where the target variable has more than two categories (case 4), some methods are suggested but two are the most common:(1) "one against one" where k(k-1)/2 models are constructed and k is the number of categories and, (2) "one against many" where each category is split out and all of the other categories are merged (DTREG Tutorial, 2003). The second method is more precise and computationally expensive.

2.2 The Kernel Functions

In SVM algorithm, the kernel function is a mathematical trick (i.e. $k(x, z) = \langle \phi(x), \phi(z) \rangle$) to provide solution to problem of classification or functional estimation (regression) by adding additional dimensions to the data of input space (low-dimensional) and map them into high-dimensional feature space. There are many kernel mapping functions that can be used but a few of them have been found to perform well in wide variety of applications. The process of finding proper Kernel function is trial and error task but the recommended kernel function is the Radial Basis Function (RBF). Here, 4 of the most common Kernel function are introduced:

- *Linear Kernel:* K(x, z) = $\langle x, z \rangle$
- *Polynomial Kernel:* K(x, z) = $(\langle x, z \rangle)^d$
- *RBF Kernel:* K(x, z) = $\exp(\frac{-\|x-z\|^2}{2\sigma^2})$
- *Sigmoid Kernel:* K(x, z) = $\tanh(\Upsilon * \langle x, z \rangle - \theta)$

The RBF kernel non-linearly maps samples into a higher dimensional space, so it can handle nonlinear relationships between target and predictor attributes; a linear basis function cannot do this. Furthermore, the linear kernel is a special case of the RBF and also equal to a Polynomial Kernel of degree one and corresponds to the original input space. A sigmoid kernel behaves the same as a RBF kernel for certain parameters. The RBF function has fewer parameters to tune than a polynomial kernel, and the RBF kernel has less numerical difficulties. (Hawley and Madden, 2005)

2.3 Support Vector Regression (SVR)

In regression formulation, the objective is to estimate an unknown continuous valued function based on a finite number set of noisy samples (x_i, y_i), i= (1,2,3,…, n), where x is n-dimentional input and "*y*" unlike pattern recognition problems, where the desired outputs y_i are discrete values like Booleans, is real valued output. The SVR basic concept is to map nonlinearly the original data into a higher m-dimensional feature space and then a linear model, f(x, w) is constructed in this feature space. The solution for regression hyperplane which is linear in feature space produces a nonlinear regression hypersurface in original input space. The special property of this model function is that it has atmost ε deviation from the actually obtained targets y_i for all the training data, and at the same time is as flat as possible. In other words, the errors are ignored as long as they are less than ε, but will not accept any deviation larger than this (Figure 4). Using mathematical notation, the general type of error (loss) function introduced by Vapnik, the *linear loss function with* ε*-insensitivity zone* is (Kecman, 2001):

$$|y - f(x,w)|_\varepsilon = \begin{cases} 0 & if\ |y - f(x,w)| \le \varepsilon \\ |y - f(x,w)|_\varepsilon - \varepsilon & otherwise \end{cases}$$

Vapnik's ε -insensitivity loss function defines a ε tube (Figure 4). If the predicted value is within the tube, the loss (error or cost) is zero. For all other predicted points outside the tube, the loss is equal to the magnitude of the difference between the predicted value and the radius ε of the tube. Note that for ε= 0, Vapnik's loss function is equivalent to a least modulus function. (Kecman, 2001) as the same way like SVC, there is C constant which is chose by user.

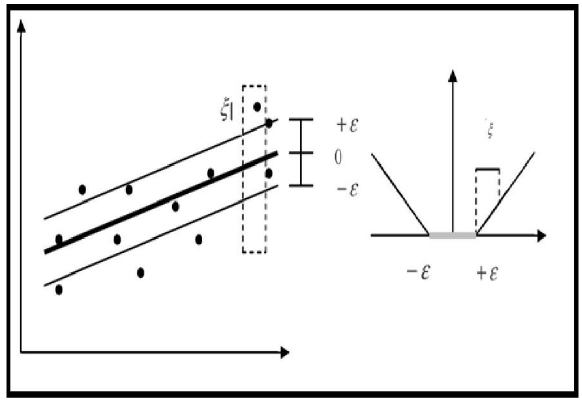

Figure 4. The soft margin lost setting for linear SVR. (Scholkopf and Smola, *2002*) ξ is positive slack variable for measurements above and below the ε tube.

3. Applications of SVM Algorithm in Petroleum Engineering

In this section, some examples of SVM approach implementations in petroleum engineering and a brief description of them are presented.

3.1 SVM and Sandstone Thickness Prediction:

Youxi and Jun (2007) utilized SVM for sandstone thickness prediction from seismic waveform in an oilfield fan. A geological model was designed for predicting different reservoir parameters and the velocity and the thickness of geobody was treated as input variables of SVM. After SVM training with only five seismic waveform traces, the final predicted velocity and theoretical velocities were extracted. The maximum error of the predicted velocity and thickness of the 101 traces was less than 4%.

3.2 SVM and Reservoir Prediction:

Li (2005) used SVM to classify 3D seismic volumes to predict oil producing intervals. He used six seismic attributes and selected well information as training data to build the SVM structure, and examined the performance of the machine with 10 remaining wells in area as test data. The results were that the ratio of correct classification as oil producing, non-commercial oil, and dry hole exceeded 85% for this application in this area and he concluded that this supervised learning approach based on a local basis function and the support vector principle has excellent generalization ability and can be used to avoid overtraining.

3.3 SVM and Facies Prediction

Wholberg et al., (2006) utilized support vector machines for the description of geologic facies from limited data by reconstructing a synthetic randomly generated porous medium consisting of two heterogeneous materials from a few data points and by comparing the performance of SVMs with that of the geostatistical approach. They pointed out the key differences between SVMs and geostatistics methods and then concluded that SVM do not need ergodicity and other statistical assumptions like geostatistics and also for very low sampling densities (e.g., 0.25%), which make the inference of statistical parameters meaningless, the geostatistical approach fails, while SVMs still do a reasonably good job in reconstructing the boundaries.

3.4 SVM and Litho-Facies Classification

Al-Anazi and Gates (2010) implemented SVM to classify litho-facies and to model permeability in heterogeneous reservoirs. The SVM was used to classify new patterns to the corresponding electrofacies and it assigned a permeability value for each data points in a multi-dimensional input log space. The proposed methodology was integrated with the extended fuzzy clustering method to extract clusters from both core and log data. Furthermore, a two-stage fuzzy ranking algorithms was used to identify and rank the most independent and significant permeability log drivers. An error analysis and comparison of the performance of the SVM with linear discriminant analysis and probabilistic neural networks for classification and back-propagation neural network and general regression neural network for permeability prediction revealed that SVM is comparable or superior to other methods for identifying lithology and permeability in a heterogeneous reservoir. By the comparison of log-based and core-based clustering they concluded that permeability prediction based on core-based clustering were slightly better than that of the log-based clustering.

Discussions

Support Vector Machine approach has been used in different varieties especially in petroleum engineering for classifying and calculating important parameters. Results indicate SVM is fast, robust and convenient to implement for prediction and solving complicated problems compared with conventional methods (statistical and experimental) that impose more difficulties especially for small sample size populations.

Building a substantial SVM model for creating a complete characterization for the petroleum reservoirs, including lithology identifications, permeability, porosity, water saturation and capillary pressure is recommended for future works. Prediction of petroleum reservoir characterizations can be done on sandstone and carbonate reservoir rocks (which are common in Iranian oil fields).

Corresponding Author:
Dr. Seyed Reza Shadizadeh
Department of Petroleum Exploration Engineering
Abadan Facility of Petroleum University of Technology
Petroleum University of Technology
Abadan, Iran
E-mail: shadizadeh@put.ac.ir

References

1. Al-Anazi, A., Gates, I.D., 2010. A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. Engineering Geology, ELSVIER Press.
2. Aristimuno. J., Aldana. M., 2006. Fuzzy logic prediction of petrophysical parameters (porosity and velocity) at La Fajadel Orinoco, Venezuela. European Geosciences Union, Geophysical Research Abstracts, 8, 03825.
3. Arnulf B. A., Graf, Silvio Borer, 2001.Normalization in Support Vector Machines. DAGM-Symposium 2001: 277-282
4. Balch. R.S., Stubbs. B.S., Weiss. W.W., Wo. S., 1999. Using artificial intelligence to correlate multiple seismic attributes to reservoir properties. In: Proceedings of the Society of Petroleum Engineers Annual Technical Conference, Houston, TX, USA, pp. 553–566.
5. Campbell. C., Ying. Y., 2011. Learning with Support Vector Machines. Morgan & Claypool Publishers.
6. Choisy. C., Belaid. A., 2001. Handwriting recognition using local methods for normalization and global methods for recognition. Proceedings of Sixth Int. Conference on Document Analysis and Recognition, pp. 23–27.
7. Cuddy, S.J., 1998. Litho-Facies and Permeability Prediction from Electrical Logs Using Fuzzy Logic. In: Eighth Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, Society of Petroleum Engineers (SPE) 65411 (49470), pp 1–12.
8. DTREG TUTORIAL, 2003. DTREG, Classification and Regression Trees And Support Vector Machines For Predictive Modeling and Forecasting, Phillip H. Sherrod.
9. Fu.Y., Cheng. Y., 2011. Application of an integrated support vector regression method in prediction of

financial returns. I.J. Information Engineering and Electronic Business, 2011, 3, 37-43.
10. Gao. D., Zhou. J., Xin. L., 2001. SVM-based detection of moving vehicles for automatic traffic monitoring. IEEE Intelligent Transportation System 745–749.
11. Gunn. S. R., Brown. M., Bossley. K.M., 1997. Network performance assessment for neuro fuzzy data modeling. Intelligent Data Analysis, volume 1208 of Lecture Notes in Computer Science, pages 313–323.
12. Gunn. Steve R., 1998. Support Vector Machines for Classification and Regression. Technical Report Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science.
13. Hawley. T., Madden. M., 2005, The Genetic Kernel Support Vector Machine:Description and Evaluation., Kluwer Academic Publishers. Page 1-16.
14. Kadkhodaie-Ilkhchi. A.a, Rezaee. M.R., Rahimpour-Bonab. H., Chehrazi. A., 2009. Petrophysical data prediction from seismic attributes using committee fuzzy inference system. ElSEVIER Press.
15. Kecman. V., 2001, Learning and Soft Computing, Support Vector Machine, Neural Network and Fuzzy Logic Models, The MIT Press.
16. Kim. K., Jung. K., Park. S., Kim. H.J., 2001. Support vector machine based text detection in digital video. Pattern Recognition 34, 527–529.
17. Li. J., 2005. Multiattributes pattern recognition for reservoir prediction.2005 CSEG National Convention.
18. Li. Z.,Weida. Z., Licheng. J., 2000. Radar target recognition based on support vector machine. Proceedings of 5th Int. Conference on Signal processing, vol. 3, pp. 1453–1456.
19. Lu. J., Plataniotis. K., Ventesanopoulos. A., 2001. Face recognition using feature optimization and v-support vector machine. IEEE Neural Networks for Signal Processing XI 373–382.
20. Ma. C., Randolph. M., Drish J., 2001. A support vector machines-based rejection technique for speech recognition. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing 1, 381–384.
21. Nikravesh. M., Aminzadeh. F., 2001. Mining and fusion of petroleum data with fuzzy logic and neural network agents. Journal of Petroleum Science and Engineering 29, 221–238.
22. Nikravesh. M., Hassibi. M., 2003. Intelligent reservoir characterization (IRESC). In: Proceedings of IEEE International Conference on Industrial Informatics 2003. Banff, Alberta, Issue, 21–24, pp. 369–373.
23. Noble. William S., 2006, What is a support vector machine?, NATURE BIOTECHNOLOGY VOLUME 24 NUMBER 12 DECEMBER 2006.
24. Parrella. F., 2007. Online Support Vector Regression. A Thesis presented for the degree of Information Science, Department of Information Science, University of Genoa, Italy.
25. Russell. B.H., Lines. L.R., Hampson. D.P., 2003. Application of the radial basis function neural network to the prediction of log properties from seismic data. Exploration Geophysics 34, 15–23.
26. Saggaf. M.M., Nebrija. E.d.L., 2003. A fuzzy approach for the estimation of facies from wireline logs. AAPG Bulletin 87, 1223–1240.
27. Scholkopf. B. and Smola. A.J. 2002. Learning with Kernels. MIT Press.
28. Soubotcheva. N., Stewart. R., 2004, Predicting porosity logs from seismic attributes using geostatistics. CREWES Research Report — Volume 16 (2004).
29. Trappe. H., Hellmich. C., 2000. Using neural networks to predict porosity thickness from 3D seismic. First Break 18 (9), 377–384.
30. Todorov. T., Hampson. D., Russell. B., 1997. Sonic log predictions using seismic attributes. CREWES Research Report, Volume 9, pages 1-39.
31. Van Gestel. T., Suykens. J., Baestaens. D., Lambrechts. A., Lanckriet. G., Vandaele. B., De Moor. B., Vandewalle. J., 2001. Financial time series prediction using least squares support vector machines within the evidence framework. IEEE Transactions. On Neural Networks 12 (4), 809–821.
32. Vapnik. V., Golowich. S., Smola. A., 1997. Support vector method for function approximation, regression estimation, and signal processing. Advances in Neural Information Processing Systems 9, pages 281– 287, Cambridge, MA, MIT Press.
33. Weston. J., 1998. Support Vector Machine (and Statistical Learning Theory) Tutorial. NEC Labs America, Princeton.
34. Wholberg. B., Tartakovsky. D., Guadagini. A., 2006, Subsurface characterization with Support Vector Machine, IEE Transactions on geosciences and remote sensing.
35. Wong. P.M., Nikravesh. M., 2001. Introduction: field applications of intelligent computing techniques. Journal of Petroleum Geology 24 (4), 381–387.
36. Yanzhou. Z., Yeling. L. Baoying. X., 2010. Application of SVM in prediction of reservoir thickness. Progress In Exploration Geophysics Journal.
37. Youxi. Y., Jun. W., 2007, SVM method for predicting the thickness of sandstone, Applied Geophysics.
38. Zahuczki. P., Barany. A.,2005. Porosity prediction from seismic attribute maps in a Hungarian limestone natural gas reservoir. MOL Hungarian Oil and Gas Plc. Exploration and Production Division.

10/26/2012