CSCanada
Advances in Petroleum Exploration and Development
Vol. 10, No. 2, 2015, pp. 72-75
DOI:10.3968/7830

ISSN 1925-542X [Print]
ISSN 1925-5438 [Online]
www.cscanada.net
www.cscanada.org

Application of Support Vector Machines to a Small-Sample Prediction

MENG Minghui[a]; ZHAO Chuanfeng[b],*

[a] CNPC Great Wall Drilling Company, Beijing, China.
[b] EOR Research Institute, China University of Petroleum, Beijing, China.
*Correspondence author.

Sponsored by National Major Scientific and Technological Special Project (2011ZX05009-004).

Received 20 September 2015; accepted 13 October 2015
Published online 31 December 2015

Abstract
The support vector machines (SVMs) is one kind of novel small-sample machine learning methods based on solid theoretical background. Highly nonlinear regression and classification are their two applications. Different from conventional statistics methods, the SVMs employs the structural risk minimizing principle, which leads to high predication precision. For this method is not essentially related to probability measure and Law of Large Numbers, the final decision function is only determined by a small fraction of sample, called support vectors. Consequently, the complexity of computation only depends on the number of support vectors rather than the dimensions of the original sample space. In most occasions of oil and gas development, only small samples are available to predict the results of one measure. Introduction of SVMs into these applications can significantly improve prediction precision.

Key words: Support vector machine; Statistical learning; Prediction; Kernel function

Meng, M. H., & Zhao, C. F. (2015). Application of support vector machines to a small-sample prediction. *Advances in Petroleum Exploration and Development*, *10*(2), 72-75. Available from: URL: http://www.cscanada.net/index.php/aped/article/view/7830
DOI: http://dx.doi.org/10.3968/7830

INTRODUCTION
To predict the response of treatments in oil and gas development, the multivariate regression method was more often employed at early time and the artificial neural network later was used[1]. Both of the two methods belong to conventional statistical learning methods. Their prediction precision can be theoretically guaranteed only if the size of their sample is large enough, close to infinity. In real prediction occasions, sample sizes are always limited. Even small samples are available in some cases. No satisfactory prediction results can be obtained for these conventional statistical methods based on Law of Large Numbers[2].

A small sample statistical learning theory was put forward by Vapnik and others[3]. The use of structural risk minimization prevents SVMs from the problems that the artificial neural network method suffers. These problems include difficulty to determine the network structure, over-learning, under-learning and local minimization. Consequently, the SVMs is considered the optimal method for small-sample classification and regression. It is suitable for kinds of nonlinear fitting functions by forming the functions through the kernel function on the basis of the representative samples (support vectors) selected from the total sample[4-8].

1. PRINCIPLE OF SVMS

1.1 Linear Regression SVMs
For a given training sample set, it is a task of seeking an optimal regression hyperplane to seek an optimal linear regression function $y = f(x) = (w * x) + b$. Consider the insensitive function ϵ as the error function. When the distances from all sample points to the potential

hyperplane are all less than or equal to ϵ, this task can be further transformed to a task solving a quadratic convex programming problem. When the distances from individual sample points to the potential hyperplane are greater than ϵ, a quadratic convex programming problem seeking an optimal regression hyperplane can be established through introducing the slack variable ξ_i.

$$\min \frac{1}{2}\|w\|^2 + C\Sigma_i(\xi_i + \xi_i^*). \quad (1)$$

$$\text{s. t.} \begin{cases} y_i - (w * x_i) - b \leqslant \epsilon + \xi_i \\ (w * x_i) + b - y_i \leqslant \epsilon + \xi_i^*. \\ \xi_i, \xi_i^* \geqslant 0 \end{cases} \quad (2)$$

An optimal hyperplane linear regression function can be obtained as follows through the Lagrange mutiplier algorithm and considering the Karush-Kuhn-Tucker condition.

$$f(x) = (w * x) + b = \Sigma_{S.V.}(\alpha_i - \alpha_i^*)(x * x_i) + b. \quad (3)$$

Where, α_i, α_i^* and b are parameters to determine the optimal hyperplane, which can be obtained through the constraint condition. The support vectors (x_i, y_i) are those training samples corresponding to nonzero coefficients α_i or α_i^* ($\alpha_i\alpha_i^* = 0$ always holds true, for α_i and α_i^* can not be equal to 0 at the same time). It is indicated that the optimal hyperplane can only be determined completely by the support vectors. For both the objective function and the constraint condition are convex, it can be stated that only one global miminal solution exits for this problem according to the optimization theory.

1.2 Nonlinear Regression SVMs

If the linear regression function can not be determined within the given sample set, the sample space needs to be mapped into a feature space supported by the eigen-function $K(x, y) = \Sigma_i \lambda_i \varphi_i(x)\ \varphi_i(y) = (\varphi_i(x) * \varphi_i(y))$ with the Mercer kernel through the nonlinear mapping function $\varphi_i(x) = (\sqrt{\lambda_1}\varphi_1(x), \sqrt{\lambda_2}\varphi_2(x), \cdots, \sqrt{\lambda_k}\varphi_k(x))$. For only the scalar product is involved in this process, it is not necessary to know the explicit expression of the mapping function. Instead, the corresponding kernel function is employed. This is the key step leaping from the linear SVMs to the nonlinear SVMs. With the use of the linear SVMs in the feature space, the regression function can be tranformed as follows.

$$f(x) = (w * \varphi(x)) + b = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(\varphi(x) * \varphi(x_i)) + b. \quad (4)$$

With the Mercer theorem taken into accout, the equation above can be simplified to the final regression function for the nonlinear SVMs.

$$f(x) = (w * \varphi(x)) + b = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) K(x * x_i) + b. \quad (5)$$

It can be seen that the nonlinear regression output for the original sample space can be obtained through the calculation with the kernel function, if only the support strength and threshold values of the support vectors are determined. With this transformation, a nonlinear SVM problem can be solved with a kernel function and the linear SVMs.

1.3 Common Kernel Functions

As stated above, one of the advantages that the SVMs holds is that only the kernel function needs to be defined in high dimension spaces instead of the explicit expressions of the feature space and the nonlinear mapping function. With this advantage, no much increase is imposed on the calculation complexity even though the space dimension increases much after transformation.

Adoption of different kernel functions will lead to different SVM algorithms. At present three kinds of kernel functions are mainly used.

(a) Multinomial Kernel Functions

$$K(x, y) = (1 + x * y)^p. \quad (6)$$

(b) Radial Basis Functions

$$K(x, y) = \exp\left[-\frac{\|x - y\|^2}{\sigma^2}\right]. \quad (7)$$

(c) Sigmoid Kernel Functions

$$K(x, y) = \tanh[v(x * y) + c]. \quad (8)$$

2. SVMS AND SMALL-SAMPLE PREDICTION

Stimulation treatments, improved or enhanced oil recovery methods are sometimes needed during the process of oil and gas field development. Prior to these treatments are employed, predictions need to be conducted on treatment response. However, the variation in oil reservoir properties makes it impossible to forecast treatment response in another oilfield with the samples from one oilfield. Even in the same oilfield, it is not suggested to forecast response in another region with the samples form one region. Predictions of this kind actually involve only small samples. Introduction of SVMs into these predictions can greatly improve the prediction precision. One example is presented as follows about an actual response prediction of conformance control treatment in water injectors. This example was used in Reference [1] with the BP artificial neural network method employed for response prediction. The obvious improvement in prediction precision can be seen with the SVMs.

2.1 Prediction Steps

When used to predict the response of conformance control treatment, the SVMs have the similar procedures to the artificial neural network method. Firstly, it is needed to select the main factors affecting the response of conformance control treatment, including the pressure index, average water injection rate and injection pressure of candidate wells prior to treatment, and the average watercut of the offset oil production well, as well as the treatment volume of the chemical agents. These five input variables, that is, the five factors, and the output variable (treatment response, that is, oil incremental) construct a SVM regression problem. It should be noted that the input and output parameters should be normalized into the range [-1, +1] in order to eliminate the effects caused by the difference in parameter dimension and unit.

The next step is to select different kernel functions and relevant parameter values to fit the training samples and predict the validation samples. The applicability of the selected model can be determined according to the validation result. If the required prediction precision can be satisfied with the selected model, the prediction model is proven suitable and reliable. Finally, this model can be used for response prediction of conformance control treatment in water injectors with similar reservoir properties.

2.2 Field Case

For the purpose of a full contrast, the SVMs and the BP artificial neural network method are seperately used to fit the whole sample set from 11 wells. The parameters in the sample set and the fitting results are shown in Table 1. And then the parameters from 8 wells are used as training samples to predict the response of the remaining three wells. The prediction results are shown in Table 2.

Table 1
The Whole Sample Set and the Fitting Results

Well names	Pressure index	Water injection pressure	Water injection rate	Watercut	Treatment volume	Acutal dimensionless oil incremental	Regression fitting results with SVMs	Regression fitting results with BP
	MPa	MPa	m^3/d	%	t			
1-11N13	31	4	160.3	59	6	-0.9199	-0.9193	1.2
1-16N17	2	2	134.3	96.4	30	-0.0837	-0.0835	-0.0089
1-18-19	27	4	105.7	90.6	5	-0.0428	-0.0423	0.2784
1-2-13	31	4.33	212.7	96.1	5	0.2935	0.2938	0.1858
1-3N17	0.4	1	148	96.4	36	-0.0041	-0.0035	0.4738
1-4-817	12	2	206.7	91.5	28	1.1903	1.1893	0.8154
1-5N16	53	7	256	97.3	30	-0.0483	-0.0494	-0.0637
D9-24	28	5.67	210.7	96.3	50	0.5723	1.0293	0.219
1-16-21	34	5	205.3	94.5	20	0.128	0.1273	0.1737
1-17-16	26	4	143.3	96.5	30	0.0833	0.0817	-0.0663
1-19-172	24	5	152.7	93	30	0.1185	0.1196	0.1352

Table 2
Prediction Results With SVMs and BP

Well names	Acutal dimensionless oil incremental	SVMs			BP
		Multinomial kernel function	Sigmoid kernel function	Radial basis function	
1-11N13	-0.0199	-0.3389	-0.4632	-0.9201	1.2000
1-16N17	-0.0837	0.3364	-0.0013	-0.00827	-0.00841
1-18-19	-0.0428	-0.1437	-0.4987	-0.0444	0.9979
1-2-13	0.2935	0.2494	0.5125	0.2953	0.2865
1-3N17	-0.0041	0.4873	0.4526	-0.0039	0.2622
1-4-817	1.1903	0.6093	0.7328	0.7053	0.9182
1-5N16	-0.0483	0.5327	0.4083	-0.0485	-0.0687
D9-24	0.5723	1.2007	0.2265	0.4823	0.2328
1-16-21*	0.1280	0.3575	0.2566	0.2217	0.2586
1-17-16*	0.0833	0.3503	-0.1672	0.0643	-0.0278
1-19-172*	0.1185	0.2491	-0.1705	0.1563	0.2563

Note. * indicates a well to be predicted.

It is indicated from Table 1 that the fitting results with SVMs are much tightly close to the actual oil incremental and much higher in fitting precision than the BP artificial neural network method except Well D9-24. In particular, if the actual oil incremental is a negative value (such as Wells 1-11N13, 1-16N17, 1-5N16, 1-18-19, 1-3N17), the fitting precision of SVMs is far higher than that of the BP artificial neural network method.

It can be seen from Table 2 that the prediction precision of SVMs, higher than that of the BP artificial neural network method, can satisfy the requirement of engineering application when the radial basis function is adopted. However, if an inapplicable kernel function is used, the prediction precision may be lower than that of the BP artificial neural network method.

CONCLUSION

(a) The SVMs can overcome the inherent disadvantages of the artificial neural network method. The field case indicates that the SVMs is higher in prediction precision than the BP artificial neural network.

(b) In order to improve the prediction precision of the SVMs, a suitable kernel function and relevant parameters should be employed.

(c) The SVMs has a promising prospect when used for the response prediction of improved or enhanced oil recovery techniques in oil and gas field development.

REFERENCES

[1] Feng, Q. H., Chen, Y. W., & Guo, J. L. (2003). Studies on response prediction of conformance control treatment in water injectors. *Petroleum Drilling and Production Technology*, *25*(6), 63-65.

[2] Chen, Y. Y., Yu, X. D., & Gao, X. H. (2004). An new method to solve nonlinear classification and regression problems (I). *Journal of Applied Meteorology*, *15*(3), 345-353.

[3] Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York, USA: Springer Verlag.

[4] Zhai, L., Li, Z. M., & Zhang, X. (2007). An optimization method for conformance control treatment in polymer injection based on SVMs. *Oil and Gas Reservoir Geology and Recovery*, *14*(4), 88-91.

[5] Li, D., Wang, H. L., & Du, Z. X. (2006). Regression prediction with SVMs for urban civil water consumption. *Journal of Tianjin University* (*Social Science Edition*), *8*(1), 64-66.

[6] Zhang, F., Zhang, X., & Zhang, L. (2007). Using SVMs to predict reservoir productivity. *Journal of Southwest Petroleum University*, *29*(3), 24-27.

[7] Le, Y. X., & Yuan, Q. S. (2005). Application of SVMs to oil and gas reservoir parameter prediction. *Natural Gas Industry*, *25*(12), 45-47.

[8] Mei, J. X., Duan, S., & Pan, J. B. (2002). Application of SVMs to small sample prediction. *Journal of Wuhan University* (*Science Edition*), *48*(6), 732-735.