Proceedings of the 5th WSEAS International Conference on Applied Computer Science, Hangzhou, China, April 16-18, 2006 (pp522-525)

THE APPLICATION OF SUPPORT VECTOR MACHINES IN THE POTENTIALITY EVALUATION FOR REVEGETATION OF DUMP OF COAL MINE[1]

LI Daoliang[*1,2] ZHUANG Chuanli[2], FU Zetian[1, 2]
1 College of Information & Electronics Engineering,
China Agricultural University, 100083, Beijing P.R. China, 100083;
2 Key Laboratory of Modern Precision Agriculture System Integration, Ministry of Education,
100083, Beijing, P.R. CHINA.

Abstract: The problem of abandoned land from mining activities is aggravated since the coal boom. Activities of excavating coal exscind natural vegetation and deposit stone on the natural land that modified the natural land contribute to waste farmlands. This paper presents the comparableness of SVMs method to artificial neural networks in the outlier detection problem of high dimensions. Experiments performed on real dataset show that the performance of this method is mostly superior to that of artificial neural networks. The proposed method, SVMs served to exemplify that kernel-based learning algorithms can be employed as an efficient method for evaluating the revegetation potentiality of abandoned lands from coal mining activities.

Key-Words: Support Vector Machine, Potentiality Evaluation, Revegetation, dump of coal mine

1 Introduction

The problem of abandoned land from mining activities is aggravated since the coal boom[1,2]. Activities of excavating coal exscind natural vegetation and deposit stone on the natural land that modified the natural land contribute to waste farmlands. Additionally, pollutes environment, losses of human life, human settlements and the infrastructure are also rising, which certainly demands urgent attention.

It is not surprising that government agencies and local land-use planners have an interest in reducing the social and economic costs due to the abandoned land from coal mining activity. Therefore, it is desired to have a notion about the potentiality for revegetation of abandoned lands from coal mining activities, so that they can decided which areas should be revegetated firstly with lower cost and shorter time.

Several countries have attempted to discriminate areas potentiality for revegetation[3]. However, until now, systematic abandoned land from mining approaches using innovative technology to estimate potential sites are very scarce, particularly in China.

Although several techniques are available for the potentiality evaluation of abandoned land from mining investigation, a number of issues that affect the performance of these techniques have been identified, including the difficulties of handling continuous (i.e. numerical) and categorical data together. However, in the light of present knowledge, few research studies have attempted to consider the application of alternative techniques such as artificial neural networks in order to resolve these problems. The mail aim of this study is to develop an abandoned land from mining potentiality evaluation model for evaluating value using support vector machines. The objectives of the current study are to develop and validate a neural methodology applicable to the tasks of load forecasts. The brief of the support vector (SVM) is described in Section 2. Sections 3 discuss the application of potentiality evaluation for abandoned land from coal mining activity. Some conclusions are presented in Section 4.

2 The brief of Support vector machines

The SVMs can be considered to create a line or hyperplane between two sets of data for classification[4-10]. In the case of a two-dimensional situation, the action of the SVM can be explained without any loss of generality. Fig. 1 shows the classification of a series of points for two different classes of data, class A (circles) and class B (squares).

[1] This work was supported by the Sino-Greek Science & Technology cooperation program(2003DF000004) and Asia IT & C project（CN/ASIA IT&C/006(89870）

[*] corresponding author: Daoliang Li (Prof.), Tel: +86-10-62736717, Fax:+86-10-62737679, E-mail:li_daoliang@yahoo.com

The SVM attempts to place a linear boundary represented by a solid line between the two different classes, and orients it in such a way that the margin represented by dotted lines is maximized. The SVM tries to orient the boundary such that the distance between the boundary and the nearest data point in each class is maximal. The boundary is then placed in the middle of this margin between the two points. The nearest data points are used to define the margins and are known as support vectors (SVs) represented by grey circles and squares. Once the SVs are selected, the rest of the features set can be discarded, since the SVs contain all the necessary information for the classifier [11].

Let $\{(x_i, y_i), i = 1, \cdots N\}$ be a training sample set S; each sample $x_i \in R^N$ belongs to a class by $y_i \in \{-1,1\}$. The goal is to define a hyperplane which divides S, such that all the points with the same label are on the same side of the hyperplane while maximizing the distance between the two classes A, B and the hyperplane. The boundary can be expressed as follows:

$$w \cdot x + b = 0, w \in R^N, b \in R \qquad (1)$$

where the vector w defines the boundary, x is the input vector of dimension N and b is a scalar threshold. At the margins, where the SVs are located, the equations for classes A and B, respectively, are as follows:

$$w \cdot x + b = 1, w \cdot x + b = -1 \qquad (2)$$

As SVs correspond to the extremities of the data for a given class, the following decision function can be used to classify any data point in either class A or B:

$$f(x) = sign(w \cdot x + b) \qquad (3)$$

For Gaussian kernels every finite training set is linearly separable in feature space [12]. Then the optimal hyperplane separating the data can be obtained as a solution to the following constrained optimisation problem [13]: find $w \in R^N$ to

$$\text{minimize } \tau(w) = \frac{1}{2}\|w\|^2 \qquad (4)$$

$$\text{subject to } y_i(w \cdot x_i + b) \geq 1 (i = 1, \cdots N) \quad (5)$$

where N is the number of training sets.

The solution of the optimisation problem Eq. (4) is called hard margin SVM classifier.Introducing Lagrange multipliers $\alpha_i \geq 0 \; i = 1,2 \cdots N$; one for each of the constraints in Eq. (5), we obtain the following Lagrangian:

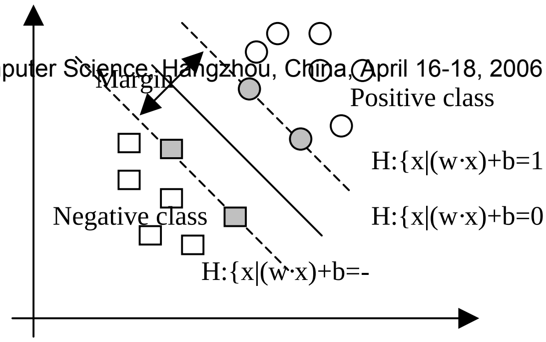

Fig 1. An example of classification of two classes by SVM.

$$L(w,b,\alpha) = \frac{1}{2}w^2 - \sum_{i=1}^{N}\alpha_i y_i (w \cdot x_i - b) + \sum_{i=1}^{N}\alpha_i \qquad (6)$$

The task is to minimise Eq. (6) with respect to *w*, *b* and to maximise it with respect to α_i. At the optimal point, we have the following saddle-point equations:

$$\frac{\partial L}{\partial w} = 0 \; \frac{\partial L}{\partial b} = 0 \qquad (7)$$

which translate into

$$w = \sum_{i=1}^{N}\alpha_i y_i x_i \, , \; \sum_{i=1}^{N}\alpha_i y_i = 0 \qquad (8)$$

From Eq. (8) it can be seen that w is contained in the subspace spanned by the x_i: By substituting Eq. (5) into Eq. (4), the dual quadratic optimisation problem can be obtained

Maximize

$$L_D(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum \alpha_i \alpha_j y_i y_j x_i \cdot x_j \qquad (9)$$

$$\text{Subject to } \alpha_i \geq 0 (i = 1,2 \cdots N), \sum_{i=1}^{N}\alpha_i y_i = 0 \quad (10)$$

Thus, by solving the dual optimization problem, one obtains the coefficients α_i, $i = 1,2 \cdots N$ which needs to express the vector w to solve Eq. (4). This leads to the non-linear decision function

$$f(x) = sign(\sum_{i=1}^{N}\alpha_i y_i (x_i \cdot x) + b) \qquad (11)$$

In cases where the linear boundary in the input spaces are not enough to separate into two classes properly, it is possible to create a hyperplane that allows linear separation in the higher dimension. In SVMs, this is achieved through the use of a transformation $\Phi(x)$ that converts the data from an N-dimensional input space to Q-dimensional feature space:

$$s = \Phi(x) \qquad (12)$$

where $x \in R^N$ and $x \in R^Q$

Substituting the transformation Eq. (12) in Eq. (3) gives the decision function as

$$f(x) = sign(\sum_{i=1}^{N}\alpha_i y_i (\Phi(x) \cdot \Phi(x_i)) + b) \qquad (13)$$

A kernel function $K(x, y) = \Phi(x) \cdot \Phi(y)$ is used to perform the transformation into higher dimensional feature space. The basic form of SVM is obtained after substituting the kernel function in Eq. (13) as follows:

$$f(x) = sign(\sum_{i=1}^{N} \alpha_i y_i K(x_i \cdot x) + b) \quad (14)$$

Any function that satisfies Mercer's theorem [13] can be used as a kernel function for computing the dot product in feature space. There are different kernel functions such as linear, polynomial, Gaussian Laplacian RBF, and Sigmoid used in SVMs.

3 APPLICATION

3.1 Data description

The benchmark data set contains 46 instances. Every single instance consists of two types of Information (see Table 1). One is information regarding target value (class labels) and attributes that contain abandoned land conditions such as Slope angle, Elevation, Topographic wetness index, Lineaments, Geological formations, Soil types, condition of traffic. The other information is about the evaluation value of abandoned land, which is target value and can be grouped into two classes:

(1) Tractable

(2) Non-tractable

Table1 1. Attributes and evaluation value of the benchmark data set

Attributes	✧ Slope angle ✧ Elevation ✧ Topographic wetness index ✧ Lineaments ✧ Geological formations ✧ Soil types ✧ Condition of traffic
Classes	✧ Tractable ✧ Non-tractable

The purpose of SVMs proposed in this paper is to construct a model that suggests target value of data instances in the testing set using only the given attributes. For training purposes, we random select 32 data instances to construct the SVM model. Testing (or ex post facto prediction) is made by rest data.

3.2 Configuration of SVMs

SVMs method has a lot of parameters to be set such as selection of kernel and the representation of the data. Moreover, it turns out to be surprisingly sensitive to specific choices of representation and kernel. In order to produce robust results, it is very important to pay attention to set these parameters. Therefore, we conducted the experiments on these parameters.

Table 2. The Configuration of different parameters

Kernels	γ	r	d
Linear: $K(x_i \cdot x_j) = x_i^T \cdot x_j$	-	-	-
Polynomial: $K(x_i \cdot x_j) = (\gamma x_i^T \cdot x_j + r)^d$	0.001,0.01, 0.1, 1, 10	0,1	2,3
Radial basis function (RBF): $K(x_i \cdot x_j) = \exp(-\gamma \Vert x_i - x_j \Vert^2)$ $\gamma > 0$	0.001,0.01, 0.1, 1, 10	-	-
Sigmoid: $K(x_i \cdot x_j) = \tan(\gamma x_i^T \cdot x_j + r)$	0.001,0.01, 0.1, 1, 10	0,1	-

Here, γ, r, and d are kernel parameters to be set for a specific problem.

3.3 Results

One result of all experiment with different kernel, as shown in the figure 2,because the experiment with linear kernel without changes of parameter it's evaluation success rate is always 0.72, so it's curve doesn't display on figure 2), the RBF kernel gives the best results with γ =1,the evaluation success rate is up to 95% on test set. We can also recognize from this validation process that it is important to choose the appropriate parameters. The optimal parameters on different kernels are summarized in Table 3.

Table 3. Optimal parameters for SVM with different kernel

Kernel	γ	r	d
linear			
polynomial	0. 1	0	3
RBF	0. 1		
Sigmoid	1	1	

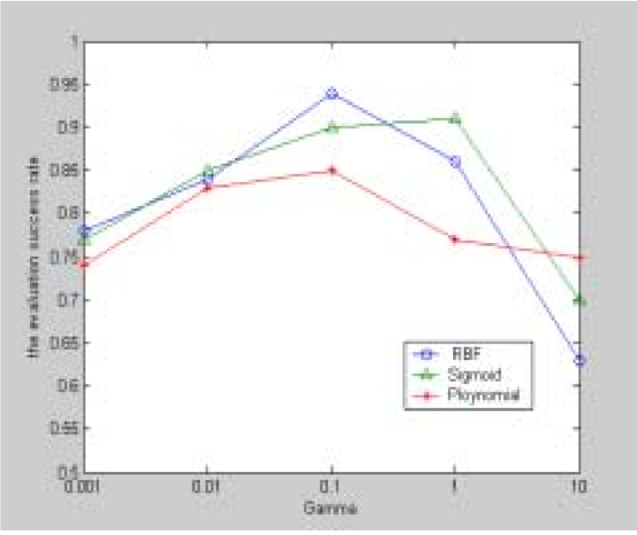

Fig 2. The evaluation success rate of different kernel

We evaluated the performance for the proposed SVMs with different types of kernel and compared it with that of Back-propagation model (an artificial neural network) in Table 4. The table shows test success rates that were achieved under each different method. As can be seen, SVMs with RBF kernel has consistently given the best performance of all five methods. Examining the results, we can see that ANNs manages to achieve 78% success on the test set. Compared to the performance of ANNs, SVMs gives relatively better results.

Table 4. Performance comparison with different kernels

Kernel	Number of classes	Test success rate
ANNs(BP)	2	78%
Linear	2	72%
Polynomial	2	82%
RBF	2	95%
Sigmoid	2	85%

4 Conclusion

We have presented SVMs that finds maximum margin hyperplanes in a high-dimensional feature space, emulating Vapnik's SVMs. The objective of this paper was to show the comparableness of SVMs method to artificial neural networks in the outlier detection problem of high dimensions.

Experiments performed on real dataset show that the performance of this method is mostly superior to that of artificial neural networks.

We can use a variety of methods to evaluate the potentiality for revegetation of abandoned land from coal mining activity. The accuracy and fastness is the barometer of reducing the cost of maintenance and operating of revegetation. The proposed method, SVMs served to exemplify that kernel-based learning algorithms can be employed as an efficient method for evaluating the revegetation potentiality of abandoned lands from coal mining activities.

Since SVMs was shown to be sensitive to the parameters and the choice of kernel, however, more reinforced consideration should be followed in how to select appropriate parameters and kernel for SVMs.

References:

[1]. Wang Ying, Li Daoliang, A potentiality evaluation model for revegetation of abandoned lands from coal mining activities, system engineer and practise, 2005.6

[2]. Z.Q. Zhang, M.H. Wong, X.P. Nie, C.Y. Lan.Effects of zinc (zinc sulfate) on rhizobia-earleaf acacia (acacia auriculaeformis) symbiotic association]. Bioresource Technology, 1998,64:97-104

[3]. Catherine Neel, Hubert Bril, Alexandra Courtin-Nomade,et al.Factors affecting natural development of soil on 35-year-old sulphide-rich mine tailings. Geoderma, 2003,111:1~20

[4]. Bo-Suk Yang, Won-Woo Hwang, Dong-Jo Kim,Andy Chit Tan Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines, Mechanical Systems and Signal Processing 19 (2005) 371–390

[5]. Hyun Joon Shina, , Dong-Hwan Eom, Sung-Shick Kim.One-class support vector machines—an application in machine fault detection and classification. Computers & Industrial Engineering 48 (2005) 395–408

[6]. Osuna, E., Freund, R., & Girosi, F. (1997). Training support vector machines: An application to face detection. In 1997 Conference on computer vision and pattern recognition (pp. 130–136). Puerto Rico: IEEE.