

Title	Potentiality Evaluation for Revegetation of abandoned lands from coal mining activities based on Support Vector Machine
Author(s)	ZHUANG, Chuanli; LI, Daoliang; GUO, Xiangyun; FU, Zetian
Citation	
Issue Date	2005-11
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/3953
Rights	2005 JAIST Press
Description	The original publication is available at JAIST Press http://www.jaist.ac.jp/library/jaist-press/index.html , IFSR 2005 : Proceedings of the First World Congress of the International Federation for Systems Research : The New Roles of Systems Sciences For a Knowledge-based Society : Nov. 14-17, 2163, Kobe, Japan, Symposium 3, Session 8 : Intelligent Information Technology and Applications Computational Intelligence (2)

Potentiality Evaluation for Revegetation of abandoned lands from coal mining activities based on Support Vector Machine

ZHUANG Chuanli¹, LI Daoliang², GUO Xiangyun³ and FU Zetian^{4*}

¹ College of Economics & Management, China Agricultural University, 100083, Beijing P.R. China
zhuangchuanli@163.com

^{2,3,4} College of Engineering, China Agricultural University, 100083, Beijing P.R. China

⁴ Key Laboratory of Modern Precision Agriculture System Integration, Ministry of Education
100083, Beijing P.R. China

*corresponding author:cauzcl@cau.edu.cn

ABSTRACT

Activities of excavating coal exscind natural vegetation and deposit stone on the natural land that modified the natural land contribute to waste farmlands. Additionally, pollution of environment, losses of human life, human settlements and the infrastructure are also rising, which certainly demands urgent attention. Revegetation had been considered as a kind of cost-effective means. A reasonable potentiality for Revegetation of abandoned lands from coal mining activities is benefit for planning of Revegetation.

In this paper, 34 instances were investigated, and seven attributes such as slope angle, elevation, topographic wetness index, lineaments, geological formations, soil types, condition of traffic, correlative with abandoned lands were recorded. A potentiality evaluation method based on support vector machine was proposed and was tested on those data. The purpose of SVM 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. We randomly select 22 data instances to construct the SVM model. Testing is made by rest data.

The results show SVM the good performance of potentiality evaluation with RBF kernel. it manages to achieve 95% success on the training set and 75% success on the testing set. Experiments performed also show that the performance of this method is mostly superior to that of artificial neural networks and SVM can be employed as an efficient method for evaluating the revegetation potentiality of abandoned lands from coal mining activities.

Keywords: support vector machine, revegetation, potentiality evaluation, abandoned lands

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, pollution of environment, losses of human life, human settlements and the infrastructure are also rising, which certainly demands urgent attention.

Revegetation had been considered as a kind of cost-effective means. The first step is that the potentiality for revegetation of abandoned lands from coal mining activities should be evaluated in order to decide where should be revegetated firstly with lower cost and in shorter time. Literatures also show that some attempts had been done that the schemes of revegetation potentiality vary with the different kinds of abandoned land in several countries [3]. However, there are less systematic approaches to estimate potential of abandoned land from mining via innovative technology, especially in China.

Although several techniques are available for the potentiality evaluation of abandoned land from mining investigation[4,5], 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[6-8]. However, in the light of present knowledge, few research studies have attempted to consider the application of alternative techniques such as artificial neural networks, support vector machine in order to resolve these problems [9].

The main aim of this study is to develop an abandoned land from mining potentiality evaluation model for evaluating value using support vector machines. The brief of the support vector machine (SVM) and artificial neural network (ANN) are described in Section 2. Sections 3 discuss the application of the potentiality evaluation model for abandoned land from coal mining activity. Some conclusions are presented in Section 4.

2. THE BRIEF OF SVM AND ANN

2.1 the brief of SVM

The SVM 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). 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

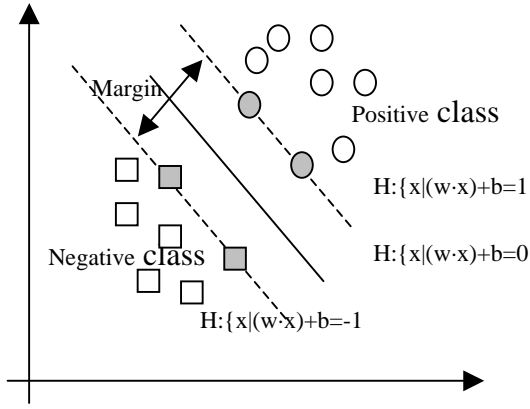


Fig 1. An example of classification of two classes by SVM.

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 $\{(x_i, y_i), i = 1, \dots, N\}$ 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:

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

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:

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

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:

$$f(x) = \text{sign}(w \cdot x + b) \quad (3)$$

For Gaussian kernels every finite training set is linearly separable in feature space [11]. Then the optimal hyperplane separating the data can be obtained as a solution to the following constrained optimisation problem [12]: find $w \in R^N$ to

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

Subject to

$$y_i(w \cdot x_i + b) \geq 1 (i = 1, \dots, N) \quad (5)$$

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 $\alpha_i \geq 0, i = 1, 2, \dots, N$; one for each of the constraints in Eq. (5), we obtain the following Lagrangian:

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

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:

$$\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0 \quad (7)$$

which translate into

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

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

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

Subject to

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

Thus, by solving the dual optimization problem, one obtains the coefficients $\alpha_i, i = 1, 2, \dots, N$

which needs to express the vector w to solve Eq. (4). This leads to the non-linear decision

function

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

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 SVM, 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:

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

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

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

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

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:

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

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 SVM.

2.2 The brief of ANN

Artificial Neural networks are composed of simple elements operating in parallel[11-13]. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. We can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements. Commonly neural networks are adjusted, or

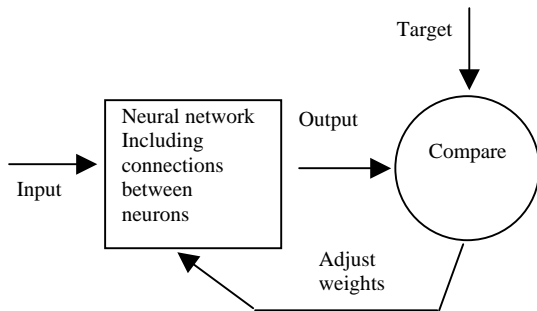


Fig.2 The flow of artificial neural network

trained, so that a particular input leads to a specific target output. Such a situation is shown below. There, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are used, in this supervised learning, to train a network. Batch training of a network proceeds by making weight and bias changes based on an entire set (batch) of input vectors. Incremental training changes the weights and biases of a network as needed after presentation of each individual input vector. Incremental training is sometimes referred to as "on line" or "adaptive" training. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision and control systems.

The following algorithm outlines the operation of the SCL network.

We assume that the number of nerve cell in input layer is R; the competed layer composes S nerve cells; the number of input model is N and input model is

$$P_k = (p_1^k, p_2^k \cdots p_R^k) \quad k = 1, 2, \dots, N$$

Step1:

Endow the linked weight weights w_{ij} of networks with random value in the region of $[0, 1]$, $(i = 1, 2, \dots, R, j = 1, 2, \dots, S)$,

Normalized the input model, and subject to

$$\sum_{i=1}^R w_{ij} = 1 \quad (j = 1 \dots S);$$

Define the original value of learning rate α ($0 < \alpha < 1$).

Define the biases b_i ($0 < b_i < 1$) $i = 1, 2, \dots, S$

Define the total learning time is T

Let $t=1$. t represents the training time

Step2:

Select any one mode $P = (p_1, p_2, \dots, p_R)$ in model (N) to input layer.

Step3

Calculate the distance between

$$W_j = (w_{1j}, w_{2j} \cdots w_{Rj}) \quad j = 1, 2, \dots, S \text{ and } P; \text{ and}$$

let

$$n_j = -\sqrt{\sum_{i=1}^R (p_i - w_{ij})^2} + b_i \quad (j = 1 \dots S)$$

Step4

Find out the least distance n_j , define the winner neuron g .

$$n_g = \max(n_i) \quad i = 1, 2, \dots, S$$

Step5

Adjust the weight linked with the winner neuron.

$$w_{ig}(t) = w_{ig}(t-1) + \alpha[p_i - w_{ig}(t-1)]$$

$$i = 1, 2, \dots, R$$

Step6:

Select another learning mode to networks input layer. Return to step3. the learning process will not end until all the learning modes have been provided to networks.

Step7:

Let $t=t+1$. Return to step2. The whole process will not end until $t=T$.

3.APPLICATION

3.1 Data description

The benchmark data set contains 34 instances. Every single instance consists of two types of Information. One is information 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.

The purpose of SVM 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 22 data instances to construct the SVM model. Testing (or ex post facto prediction) is made by rest data.

Because SVM method required the input data are quantificational, Scale method was adopted for qualitative data expressed as a quantity. The data was classified into 10 categories, labeled 0.1, 0.2...1 separately.

As a matter of convenience, we have normalized all input data. Function of membership grade was used to normalize the quantificational data, for example: maximum-minimum Function of membership grade.

$$r_1(x) = \begin{cases} 0 & 0 \leq x < q_{\min} \\ \frac{x - q_{\min}}{q_{\max} - q_{\min}} & q_{\min} \leq x < q_{\max} \\ 1 & q_{\max} \leq x \end{cases}$$

which q_{\max} and q_{\min} are the max value and min value of collected data.

3.2 Matlab toolboxes

Two software package of MATLAB were used in this work. SVM6.81 that developed by the students of Prof. Deng Naiyang in China agricultural university is a good tool for classification and regression. all parameters of SVM can be input interactively by user.



Fig3. The interface of SVM6.81

The Artificial Neural Network Toolbox we used was included in MATLAB. A number of demonstrations are included in the toolbox. Each example states a problem, shows the network used to solve the problem, and presents the final results. It is helpful for ANN model design and application.

3.3 Configuration of SVM

SVM 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, many experiments were conducted the on these parameters, and the optimal parameters are gained.

As shown in the figure 4, 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 4), the RBF kernel gives the

best results with $\gamma=1$, 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 1.

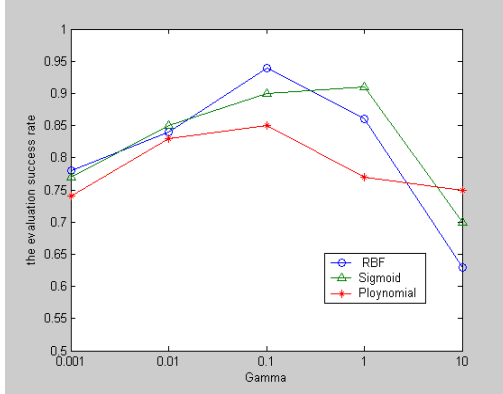


Fig 4. The evaluation success rate of different kernel

Table 1. Optimal parameters for SVM with different kernel

Kernels	γ	r	d
Linear:	-	-	-
Polynomial:	0.1	0	3
Radial basis function (RBF):	0.1		
Sigmoid:	1	1	

Here, γ , r , and d are kernel parameters. So the RBF kernels are selected with $\gamma = 1$ in our method.

3.4 Design of ANN

The ANNs have much different architecture and the most popular type of neural network is the back propagation neural network (BP). BP works well for

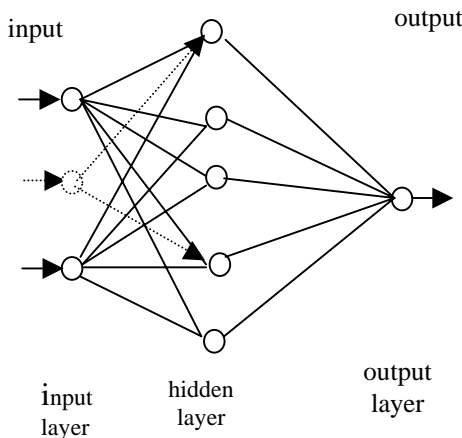


Fig 5. The ANN back-propagation model

pattern matching and for trend analysis. The BP can learn many different output patterns simultaneously with dramatic accuracy. A conventional BP uses three layers of nodes, the input layer, the hidden layer and output layer. According to above principle and the practices problem, the new models architecture in this study is as fig 5. The input pattern consists of 7 input neurons. The hidden layer consists of 18 hidden neurons. The output pattern consists of only one output neuron.

3.5 Results

We evaluated the performance of the proposed SVM method and compared it with that of Back-propagation model (an artificial neural network) in Table 2. The table shows test success rates that were achieved under different method. Examining the results, we can see that ANN manages to achieve 66% success rate on the test set and the SVM is 75%. SVM with RBF kernel gives relatively better results.

Table 2. Performance comparison between SVMs and

Kernel	Training success rate	Test success rate
ANN(BP)	100%	66%
SVM(RBF)	95%	75%

4. Conclusion

We have presented SVM that finds maximum margin hyperplanes in a high-dimensional feature space, emulating Vapnik's SVM. The objective of this paper was to show the comparableness of SVM 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 SVM 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 SVM.

5.Acknowledgements

This program is partly supported by the Sino-Greek Science & Technology cooperation program under Grant No. 2003DF000004 and EU Asia IT & C project under Grant No. CN/ASIA IT&C/006(89870).

REFERENCES

- [1]. Wang Ying, Li Daoliang: A potentiality evaluation model for revegetation of abandoned lands from coal mining activities. *Journal of China Agricultural University*, 10 (2005)88 - 92
- [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*, 64 (1998) 97-104
- [3]. Catherine Neel, Hubert Bril, Alexandra Courtin-Nomade: Factors affecting natural development of soil on 35-year-old sulphide-rich mine tailings. *Geoderma*, 111(2003) 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.: Training support vector machines: An application to face detection. In: 1997 Conference on computer vision and pattern recognition. Puerto Rico: IEEE. (1997) 130–136
- [7]. L.B. Jack, A.K. Nandi: Fault detection using support vector machines and artificial neural networks. augmented by genetic algorithms. *Mechanical Systems and Signal Processing*. 16 (2 - 3) (2002) 373 - 390
- [8]. V.N. Vapnik: An overview of statistical learning theory. *IEEE Transactions on Neural Networks* 10 (5) (1999) 988 - 999.
- [9]. K.R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf: An introduction to kernel-based learning algorithm, *IEEE Transactions on Neural Network* 12 (2) (2001) 181 - 201.
- [10]. Jie Song, Huanwen Tang: Support vector machines for classification of homo-oligomeric proteins by incorporating subsequence distributions. *Journal of Molecular Structure: THEOCHEM*. 722 (2005) 97–101
- [11]. Wen Xin, Zhou Lu. *Design for ANN Based on MATLAB*. Science press, 2001
- [12]. V.V. Gafiychuk, B.Yo. Datsko, J. Izmaylova: Analysis of Data Clusters Obtained by Self-organizing Methods. *Physica A* 341(2004)547-555
- [13]. Xu Dong, Wu Zheng: *Analysis and design for ANN based on MATLAB6.x*. Xidian University Press, 2002
- [14]. C.J.C. Burges, A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery* 1998, vol.2, pp. 955 - 974.