RESEARCH ON LANDSLIDE PREDICTION MODEL BASED ON SUPPORT VECTOR MODEL

Xiaowen Zhao [a,*], Min Ji [a], Xianguo Cui [a]

[a] Geomatics College, Shandong University of Science and Technology, 579 Qianwangang Road, Economic & Technical Development Zone, Qingdao, China,266510,wen1987zi@163.com, jimin@sdust.edu.cn, cxg824@163.com

KEY WORDS: mine landslide, SVM, prediction model, GIS, LIBSVM, cross validation, grid search

ABSTRACT:

The Landslide, which is caused by mining activities, has become an important factor which constrains the sustainable development of mining area. Thus it becomes very important to predict the landslide in order to reduce and even to avoid the loss in hazards. The paper is to address the landslide prediction problem in the environment of GIS by establishing the landslide prediction model based on SVM (support vector machine). Through differentiating the stability, it achieves the prediction of the landslide hazard. In the process of modelling, the impact factors of the landslide are analyzed with the spatial analysis function of GIS. Since the model parameters are determined by cross validation and grid search, and the sample data are trained by LIBSVM, traditional support vector machine will be optimized, and its stability and accuracy will be greatly increased. This gives a strong support to the avoidance and reduction of the hazard in mining area.

1. INTRODUCTION

The increasing demand for coal of the industrial society has led to more and more serious coal mining. The Mining-induced landslide hazard has seriously influenced the sustainable development of mining areas. So it is important for us to select a suitable method to predict mine slope stability. However, there are many influencing factors for landslide and the effects of the same factors in various areas are different. The mathematical relationship between the factors which impact landslide and the landslide stability prediction is hard to obtain. Therefore, it is a comparatively accurate method to get a statistical analysis model with the historical data. SVM can get solution by solving a convex quadratic programming question. The solution is global optimal solution, and its ratio is high. The Prediction with SVM will use structural risk minimization principle instead of the empirical risk minimization principle, maximize the generalization ability of learning machine, make sure that the independent test set which was gotten from a limited sample of training set remains a small error, and get a non-linear mathematical relationship with a higher dimension at the same time. In this paper, we identified the complex relationship between influencing factors and stability prediction of landslide by training samples with SVM and predicted results of unknown data with the relationship. It was proved that mine landslide prediction based on SVM got satisfactory result, and it was a prediction model with a high accuracy and stability.

2. SVM (SUPPORT VECTOR MACHINE)

SVM was a new general-purpose machine learning method based on statistical learning theory, and it was built under the theory framework and general approach of machine learning with limited samples. Its basic thought was to transform the input space to a high-dimension one by using the non-linear transformation defined by inner product function and to find a non-linear relationship between the input variables and the output ones in the high-dimension. SVM had a better generalization than neural network which used empirical risk minimization principle. (Hsu, 2009)

2.1 The generalized optimal separating hyper plane

SVM developed from the optimal separating hyper plane in a linear condition. To make it clear, we started from the two-dimension situations.

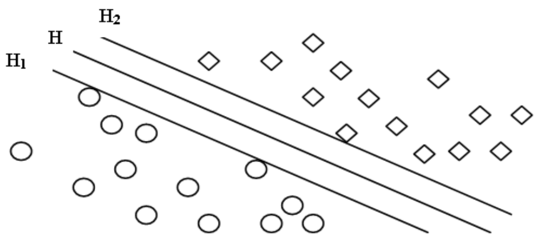

Figure 1. Separating hyper plane

As Figure 1 showed, set circular and diamond graphics as two kinds of samples, the straight line H as category line, H_1 and H_2 as straight lines which lines the samples nearest to category line were on and paralleled to the category line, the distance between the two lines called class interval.(Li X.Z.,2009; Zhao H.B., 2008). The optimal separating line asked for correct separation of the two kinds of samples and the largest class interval. The general equation could be formula (1):

$$x \cdot \omega + b = 0 \tag{1}$$

x - the samples data
w, b – parameter for the line

* Corresponding author. E-mail address: wen1987zi@163.com;

Formula (1) could be normalized. We needed to make linear separable sample set (2) satisfy formula (3):

$$(x_i, y_i)\ (\mathrm{i}=1,\cdots,\mathrm{n}, \mathrm{x}\in \mathrm{R}^{\mathrm{d}}, \mathrm{y}\in\{+1,-1\}) \tag{2}$$

$$y_i[(\omega\cdot x_i+b)]-1\ge 0(i=1,2,\cdots,n) \tag{3}$$

x_i, y_i - the sample point data

At this time, the classification interval was $\frac{2}{\|\omega\|}$. The maximum of classification interval equalled the minimum $\|\omega\|^2$. So the hyper plane which satisfied the formula (3) and made the $\|\omega\|^2$ get the minimum was the optimal separating hyper plane. The sample points on two straight lines (H_1 and H_2) were support vectors. (Jiang Q. W., 2005)

We used Lagrange method to transform the hyper plane question to a dual one, subjected to:

$$\sum_{i=1}^{n} y_i\alpha_i = 0(\alpha_i \ge 0, i=1,2\cdots,n) \tag{4}$$

α_i - Lagrange multiplier parameter

And Looked for the maximum value with the parameter α_i from the function $Q(\alpha)$, its expression was:

$$Q(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j) \tag{5}$$

Parameters α_i were Lagrange multipliers corresponding to every parameter, so function $Q(\alpha)$ was a question which asked for the best solution in the quadratic function under inequality constraint, there was a unique solution. There were few non-zero α_i, so the samples which were corresponding to these α_i were support vectors. The optimal separating function $f(x)$ could be obtained from the questions above:

$$f(x)=sign\{(\omega\cdot x)-b\}=sign\{\sum_{i=1}^{n}\alpha_i y_i(x_i\cdot x)-b^*\} \tag{6}$$

The summation in formula(6)just contained the support vectors in fact, and b^* were classification threshold which could be obtained by any of the support vectors (vectors which satisfies the equation in formula(3)) or the mid-value of any two support vectors from the two kinds of samples.(Dong J. X., 2003)

2.2 Non-linear SVM

In non-linear condition, added a relaxation $\xi_i \ge 0$ to formula (3), then it changed to (7):

$$y_i[(\omega\cdot x_i+b)]-1+\xi_i\ge 0(i=1,2,\cdots,n) \tag{7}$$

The goal changed to be the minimum value of formula (8):

$$(\omega,\xi)=\frac{1}{2}\|\omega\|^2+C(\sum_{i=1}^{n}\xi_i) \tag{8}$$

That was equally to the generalized optimal separating hyper plane considering of the least wrongly classified sample and the maximum separating interval. C > 0 was a constant and the penalty parameter of the error term. The dual problem of optimal separating hyper plane in non-linear situation was almost the same as the linear ones. The multipliers α_i objected to:

$$0\le\alpha_i\le C(i=1,2,\cdots,n) \tag{9}$$

The method which SVM used to construct separating decision function in non-linear condition contained two steps. First step was translating training data from raw mode to high dimension space by non-linear transformation of special kernel function. Second step was looking for optimal separating hyper plane in feature space. The hyper plane was corresponding to non-linear separating surface in the raw mode. So, there was only one more mapping link in non-linear condition than the linear ones when using SVM. We supposed the non-linear mapping to be $x\to\varphi(x)$, the function $Q(\alpha)$ changed to:

$$Q(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i\cdot x_j) \tag{10}$$

$K(x_i\cdot x_j)=\varphi(x)\cdot\varphi(x)$ - (kernel parameter)

The separating decision function in non-linear SVM changed to:

$$f(x)=sign\{(\omega\cdot x)-b\}=sign\{\sum_{i=1}^{n}\alpha_i y_i K(x_i\cdot x)-b^*\} \tag{11}$$

The kernel functions $K(x_i\cdot x_j)$ in formula (11) were in accordance with the mercer condition, and in corresponding to the inner product in the transformation space. (Gallus D.)
In the choice of kernel function, there were three options:

(1)Multinomial kernel: $K(x \cdot x_i) = [(x_i \cdot x_j) + 1]^d$ (12)

(2) RBF: $K(x \cdot x_i) = \exp\{-\frac{|x_i - x_j|^2}{\sigma^2}\}$ (13)

(3) Sigmoid function: $K(x \cdot x_i) = \tanh[v(x_i \cdot x_j) + a]$ (14)

3. THE APPLICATION OF SVM IN PREDICTION FOR LANDSLIDE IN MINING AREA

The prediction of the landslide by using the SUV contained the prediction on time-series and prediction on space. We used prediction on space which depended on identification of the stability of side slope. We chose on proper influencing factors for stability and picked up these factors to construct distinguishing model. We used training of samples data in the mining area to get certain discriminant function of side slope stability and then used the function to get decision outcomes of unknown sample data and the stability outcome of the sample points.

3.1 Determination of influencing factors for stability of landslide in mining area

Because of the special environment in mining area, influencing factors were more complicated than normal areas. In normal areas, the factors such as altitude, slope, aspect, vegetation coverage, litho logical character, geologic structures and so on appreciably affected the occurrence of landslide. From the survey on the study of mining area we found that the most appreciably influencing factors contained litho logical character, geologic structures, thickness of overlying strata, precipitation, precipitation intensity, slope, aspect, slope mining conditions. The data of Litho logical character, geologic structures, slope, aspect could be achieved from remote sensing image interpretation, thickness of overlying strata from field reconnaissance, precipitation and precipitation intensity from updating information on the web, slope mining conditions could be grated after field Reconnaissance.

3.2 Acquisition of sample points for prediction of landslide

In order to get enough sample data with an acceptable quality, we needed to choose suitable sample points in mining area, and get the influencing factors on each sample point. We found that it will not be enough if we only used the landslide points in mining area to be the historical data for training, so we needed to pick up certain numbers points besides centre points.

We set certain principle when picked up points besides centre points in order to make sure the samples were well distributed so that they would not affect the accuracy and stability. When the landslide area were less than 4 pixel areas, we picked up 1 point; when 4-5 pixel areas, we picked up 4 points; when 5-9 pixel areas, we picked up 5 points; according to landslide area, we may choose 1, 4, 5, 9, 16, 25 points on one landslide. We added sample points on the basis of primary points, for the primary points were certain, so we could raise the number of training samples and ensure the accuracy and stability of training model. (Ma Z. J., 2003)

3.3 Landslide database

From the analysis of database we knew that landslide database needed to contain infrastructure data layer and landslide disaster layer. The infrastructure layer involved spatial data and properties data. Non-numeric data in property layer needed to be transformed to numeric ones which were easy to use. The transformation was with certain principles. The fields in the infrastructure data list were built according to the rules in related stipulate. The graphic layer contained thematic map of each influencing factor, topographic map and so on. Landslide disaster data was the most important part in the database. There were both property and spatial data of landslide point and area. We used GIS technology to establish property database and spatial database because there were both kinds of data. The structured chart was showed in Figure.2.

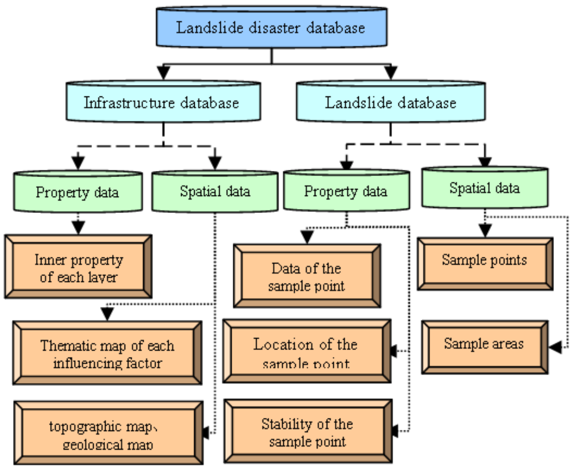

Figure 2. Database structure

3.4 Selection of parameters in landslide prediction model

We used training samples and testing samples to do the experiment. When we trained the model, fitting accuracy of testing samples was different with different input parameters in prediction model. In order to find the best fitting function, we did lots of parameter tests. The best fitting function required suitable parameters, C , σ .

We used grid search method to choose parameters and estimate the generalization of each set of parameters. Grid search method gave different numeric values to M penalty parameters for C and N kernel parameters for σ , constructed M*N combinations for different SVM models. We estimated the generalization of parameters to choose grid points of (C, σ) with the highest generalization.

The method to determine parameter based on grid search involved training of SVM and comparison of prediction accuracy. To the same grid point of (C, σ) , different training method would get discriminant function with different prediction accuracy. To n samples, picked up n-1 samples to train prediction model and got expected value of error rate. Then we used the value to determine capability of the grid point. We provided a possible interval of C (or σ) with the grid space. Then, all grid points of (C, σ) were tried to see which one gave the highest cross validation accuracy. We picked up the best parameters to train the whole training set and generate the final model.

The implementation was as follows: we chose a range for grid points of (C,σ), for example, $\sigma = 2^{10} \sim 2^{-15}$, step-size in research was -1 and $C = 2^{-10} \sim 2^{15}$, step-size was 1, then two-dimension grids of C and σ were structured on the coordinate system. We got the accuracy of each grid point, drew contour line of the accuracy to determine the best grid point. If these points couldn't get the requested accuracy, we reduced the step-size to do ransack.

The advantage of grid research was searching 2 parameters at the time and we could finally get the best grid point to make the accuracy get the highest level. In computational process, grid points could decouple between each other in order to do parallel computation and get a high efficiency. Parallel research cut down the time for searching the best parameter value, but when the samples were large in numbers, repeating the process one by one would also cost lots of time. Then we used cross validation via to do approximate evaluation. That was to divide samples into K sets, pick up K-1 set to be training set and get a decision function, then use the function to predict the testing set. This would reduce the frequency of training and ensure the accuracy for prediction, and it was called K-fold cross-validation. (Wang X. L., Li Z. B., 2005)

3.5 Application of prediction model

The Process of prediction in mining area by using the SVM contains acquisition of training samples, selection of SVM kernel function and training software, generation of SVM prediction model and decision outcomes of unknown data. The process was showed in the Figure.3.

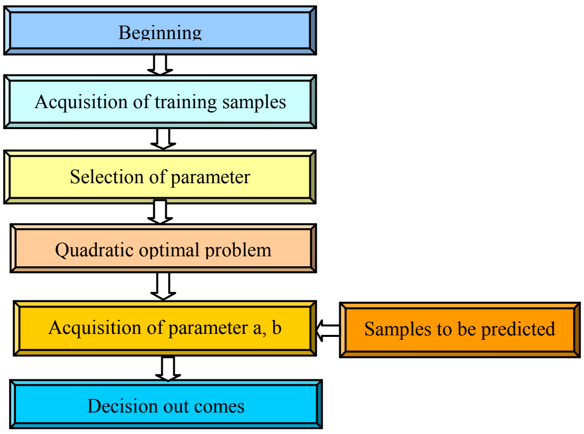

Figure 3. predicting process

3.5.1 Collection and pre-processing of materials: Before the application of SVM in landslide, we gathered materials which were to use for research. After the determination of influencing factors, we did pre-processing to the samples at first. According to assessment objective and the characteristics of the regional environment, we determined the grading standard for each factor. In order to make it easy for us to express prediction outcome in the classification diagram in GIS software on the basis of statistic analysis with the regional survey data, we found out the highest and lowest value to determine ranges for each factor data. Then we divided influencing factor value into grades according to divided principle and translated the influencing factor into quantitative value taken in accordance with actual situation.

Slope range	Raster count	Land Slide	Proportion
0-4.411999	26728	2	7.48279E-05
4.411999-9.022772	37789	12	0.000317553
9.022772-12.707438	58619	20	0.000341186
12.707438-15.93173	59643	46	0.000771256
……	…..	…..	…..

Figure 4. slope range

3.5.2 Grid unit division: Whether the division was applicable would affect reasonableness of the assessment outcomes and influence complexity of parameter acquisition in prediction process. When the unit ware was defined, we could take every evaluation unit as an independent individual. Usually, two ways were prepared for the division, regular division and irregular division. We used regular division to divide the study area into N grid units. Took each unit to be a point and extracted data from the thematic map using weighted average method.

3.5.3 Selection of kernel function: The study showed that types of kernel had little to do with the capability of prediction model, the important ones were kernel parameters and error penalty parameter. We chose RBF function as the kernel function in designing prediction model. Because there were only two parameters when using RBF function, this would keep the stability of function. (Dong H., 2007)

3.5.4 Selection of the training software: We verified from lots of literatures that using LIBSVM to train sample data would get better result than others. LIBSVM was a library for support vector machines (SVM). Its goal was to make users use SVM as a tool easily. We prepared data in special format for LIBSVM and referred to the file "*heart scale*" which is bundled in official LIBSVM source archive. The input file format is as follows.

[Label1] [index1]: [value1] [index2]: [value2]...
[Label2] [index1]: [value1] [index2]: [value2]...
(http://www.csie.ntu.edu.tw/~cjlin/libsvm/)

3.5.5 Result analysis: From above tasks, we got training samples and testing samples, the factors charts of study field. Then we used training samples to do the test. We inputted different parameter in training process and did comparison on the fitting degree by testing samples. We constructed prediction models with the best parameters obtained by the above comparison, then used the model to train the testing samples and picked up support vectors in the samples set to get the discriminant function. We used LIBSVM to train samples, and then got the prediction model to estimate the unknown grid units divided according to the study field. The outcome was expressed in GIS software by showing in the map. In order to verify the method, we used another method to do the test on the accuracy with the help of the data that we've got. The similar ways contained logistic regression model, stability factor prediction and so on. We put the landslide layer in to do a test and found that SVM prediction model got a higher accuracy. (Wang H. W., 2007)

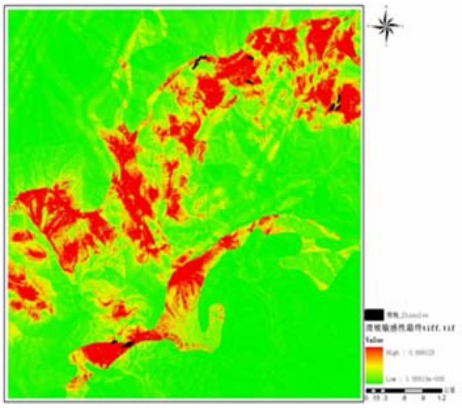

Figure 5._result chart

Classification Table[a]

			Predicted		
			flag		Percentage Correct
Observed			.00	1.00	
Step 1	flag	.00	1837	521	77.9
		1.00	393	1965	83.3
	Overall Percentage				80.6

Figure 6._accuracy list

4. CONCLUSION

The Prediction of landslide was related to several influencing factors and there were interactions in factors. It was hard to use traditional mathematical analyzing method to get a certain linear relationship between prediction outcome and influencing factors. SVM avoided these questions by getting non-linear relationship using historical data. SVM got result from solution of convex quadratic programming problem without numerous samples, and the solution was global optimal solution with a high accuracy. Small sample size will take a great advantage in mining area where sample data was hard to get. Using function which was obtained with sample data would make each possible influencing factor be considered. From principal component analysis we could determine the primary factors and it also took the interactions in factors into account. The prediction model obtained by training known landslide point data and stable point data was possessed of a high capability and accuracy. Display in GIS software was useful for expressing the outcome visually intuitive and further analysis to outcome.

The Application of landslide prediction based on SVM in mining area is always in exploration, and there are still many things which need to be researched further. For example, the initial selection of kernel parameter and penal parameter need to be directed much better in order to save the time on finding the best parameters.

REFERENCES

Dong H., 2007, Prediction of landslide displacement based on Takens theory and SVM, China Journal of Highway and Transport, PP,13-18

Dong J. X., Krzyzak A., 2003, A Fast Parallel Optimization for Training Support Vector Machine, Canada

Gallus D., Abecker A., Classification of Landslide Susceptibility in the Development of Early Warning Systems

Hsu C.W., Chang C.C., 2009, Department of Computer Science, A Practical Guide to Support Vector Classification, Taiwan

Jiang Q. W., 2005, ArcGIS-based assessment of regional landslide hazard, China

Li X. Z., Zhao Z. Y., 2009, Basis and application prospects of Support Vector Machine, SCIENCE&TECHNOLOGY INFORMATION, China, PP 431,461

Ma Z. J., 2003, Prediction of landslide based on Support Vector Machine theory, Journal of Zhejiang University (Science Edition), China, PP, 592-596

Wang H. W., 2007 No.4, Predictive modelling on multivariate linear regression, Journal of Beijing University of Aeronautics and Astronautics, China

Wang W., Zhang N., 2008, General approach of solving Convex constrained quadratic programming problem, Journal of Hainan Normal University（Natural Science）, China

Wang X. L., Li Z. B., 2005, Grid-based search of the Support Vector Machine Kernel Parameters, PERIODICAL OF OCEAN UNIVERSITY OF CHINA, China

Wu J., 2006, *Localized Support Vector Machines For Classification*

Zhao H. B., 2008, *Analysis of Support Vector Machine in Rock and Soil Mechanics and Engineering*, PP, 19-31

LIBSVM -- A Library for Support Vector Machines，Chang Chih-Chung and Lin Chih-Jen,

http://www.csie.ntu.edu.tw/~cjlin/libsvm/

APPENDIX

Supported By Chinese National Programs for High Technology Research and Develop-ment(2009AA12Z147);
Supported By the Key Laboratory of Surveying and Mapping Technology on Island and Reed, State Bureau of Surveying and Mapping(2009B14);
Supported By Key Laboratory of Advanced Engineering Surveying of SBSM(TJES0805);