

## **STUDY ON EARLY WARNING OF BUILDING SAFETY BASED ON SVM**

### **Li-Jun FENG<sup>1</sup>**

PhD Candidate, Tianjin University of Finance and Economics, Tianjin, China  
College of Urban and Rural Construction,  
Agriculture University of Heibei, Baoding 071001, Hebei, China

**E-mail:** flj69@eyou.com



### **Shu-Quan LI**

Tianjin University of Finance and Economics, Tianjin, China

**E-mail:** lishuquan@tjufe.edu.cn

**Abstract:** *The building is important strut industry of our national economy. But the safety accidents in building domain occur frequently every year. The safety problem of building directly influences or even restricts the development of building industry. Support Vector Machine (SVM) is a kind of new machine learning method developed on the basis of statistical learning theory. This method based on the principle of structural risk minimization can solve the problem of overfitting effectively and has good generality capability and better classification accuracy. In this paper, the author applied this method in safety management of building and carried on the study of early warning system of building safety based on SVM. At last, we carried on the data experiment to this problem and proved that the method of SVM has good generality capability.*

**Key words:** *Building safety, Early warning, generality capability, Support Vector Machine*

### **1. Introduction**

At present, the building industry has already become the important strut industry with the development of our national economy. The safety problem of building domain relates to the people life and nation property safety. Its direct aftereffect harms the people's vital interest and influences the situation of social stability. Because the building industry has these characteristics of singleness and complexity and open-air work and upper-air work and labor-intensiveness, the safety accidents in building domain occur frequently. According to the incomplete statistics, the annual number of casualties is about thousands of people and direct economic losses are over ten billion Yuan. So the safety problem in building domain has aroused the universal attention of society. In view of this situation, the government has promulgated a series of laws and regulations to strengthen the safety production of building domain. Some management personnel and technical personnel of building enterprise and

some experts have also carried through various studies and discussions to safety production of building, including having made some achievements in building safety evaluation and early warning. In allusion to the current situation of building safety management of our country, we carried on the study of early warning of building safety based on Support Vector Machine (SVM) in this paper. We hope that to develop this study can reduce casualty and property losses. In addition, it has certain model significance for the building safety management way of our country.

## 2. Support vector machine

Support Vector Machine is a kind of new machine learning method developed on the basis of statistical learning theory. This method based on the principle of structural risk minimization can solve the problem of overfitting effectively and has good generality capability and better classification accuracy. It is becoming a new research focus of machine learning field after pattern-recognition and neural network.

SVM is used to find the optimal separating hyperplane of linear classification problem first. The so-called optimal separating hyperplane not only can be used to separate the data, but also can maximize the margin. So, the problem of constructing the optimal separating hyperplane can be turned into the following optimization problem:

$$\min \phi(w) = \frac{1}{2} \|w\|^2 \quad (1)$$

$$\text{s.t. } y_i(w \cdot x_i + b) \geq 1, i=1,2,\dots,l$$

The above problem can be transformed to the following dual problem by using Lagrange optimization method:

$$\max W(\alpha) = \sum_{j=1}^l \alpha_j - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j (x_i, x_j)$$

$$\text{s.t. } \sum_{i=1}^l y_i \alpha_i = 0$$

$$\alpha_i \geq 0, i=1,2,\dots,l \quad (2)$$

where  $\alpha_i$  are the Lagrange multipliers.

Solving Equation (2) with constraints Equation determines the Lagrange multipliers, and the optimal separating hyperplane is given by,

$$f(x) = \text{sgn}\{(w^* \cdot x_i) + b^*\} = \text{sgn}\left\{\sum_i y_i \alpha_i (x_i, x) + b\right\} \quad (3)$$

where  $\text{sgn}()$  denotes the sign function.

So far the discussion has been restricted to the case where the training data is linearly separable. However, in general this will not be the case. In the case where it is misclassification, alternatively a more complex function can be used to describe the boundary. To enable the optimal separating hyperplane method to be generalized, Cortes and Vapnik (1995) introduced non-negative variables,  $\xi_i \geq 0$ . The  $\xi_i$  are a measure of the misclassification errors. The optimization problem is now posed so as to minimize the classification error as well as minimizing the bound on the VC dimension of the classifier.

The generalized optimal separating hyperplane is determined by the vector  $w$  which that minimizes the functional,

$$\min \phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \quad (4)$$

where  $C$  is a given viable. The generalized optimal separating hyperplane is nearly the same as to linearly separable problem, just the constraints Equation turns into:

$$0 \leq \alpha_i \leq C, \quad i=1,2\dots l.$$

To non-linear problem, we can transform it to the problem of a high dimensional feature space by the use of reproducing kernels. The idea of the kernel function is to enable operations to be performed in the input space rather than the potentially high dimensional feature space. Hence the inner product does not need to be evaluated in the feature space. This provides a way of addressing the curse of dimensionality. So the optimal separating hyperplane is transformed to:

$$f(x) = \text{sgn} \left\{ \sum_i y_i \alpha_i K(x_i, x) + b \right\} \quad (5)$$

where  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$

Hence, if we select different kernel function, we can acquire different support vector machines.

### 3. Early warning of building safety based on SVM

The early warning problem can be regarded as classification problem. Considering the simplest situation, we can classify it into having warning and no having warning. Then we regard the early warning problem as classification problem of two classes by integrating the history data and the present data. Hence, we can solve it by SVM. If we hope to infer different warning degree, we may regard early warning problem as classification problem of multiclass.

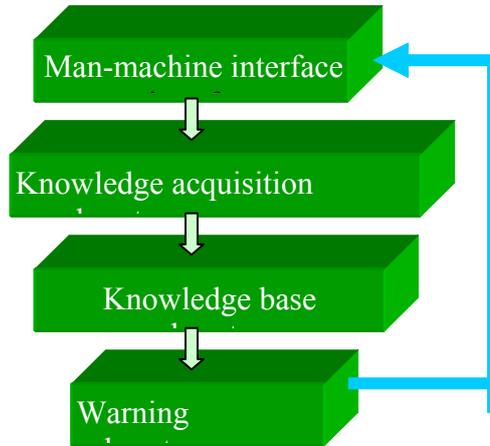
The SVM learning algorithms include SVMlight, SMO, Chunking, Decomposing and so on. We select SVMlight when constructing the early warning system. The systematic design idea is to design the effective algorithm and choose working set of including  $q$  non-0 weight and decompose the original optimal problem.

The process of early warning of building safety can be regarded as a process of pattern recognition. Comparing the new early warning sample of unknown warning degree with standard sample of known warning degree, we can determine mode category of early warning which new early warning sample should belong to.

According to the above analysis, we constructed the framework of deal with early warning problem by computer as Figure1:

In Figure 1, the man-machine interface is the program interface which users use the early warning system. The original data trained and relevant parameters are inputted through the man-machine interface and the alarm signal is also displayed through the man-machine interface. The knowledge acquisition subsystem formed by SVM classification algorithm is a core of the early warning system. The knowledge base subsystem is a carrier of the early warning knowledge, and is also a guide part of the warning subsystem. Warning subsystem is equivalent to the wrong inspection system in the pattern recognition system.

Under the drive of the new early warning data, it received the warning result by the calculation of decision-making function.



**Figure 1.** Early warning system framework

In Figure 1, the man-machine interface is the program interface which users use the early warning system. The original data trained and relevant parameters are inputted through the man-machine interface and the alarm signal is also displayed through the man-machine interface. The knowledge acquisition subsystem formed by SVM classification algorithm is a core of the early warning system. The knowledge base subsystem is a carrier of the early warning knowledge, and is also a guide part of the warning subsystem. Warning subsystem is equivalent to the wrong inspection system in the pattern recognition system. Under the drive of the new early warning data, it received the warning result by the calculation of decision-making function.

#### 4. Experiments

We validated the capability of early warning system by the data of some construction group. This data was collected from 1984 year to 2003 year, altogether 20 years.

We regarded the data from 1984 year to 1998 year as training samples. We chose 32 influence features of building safety about each training sample, as in table 1<sup>2</sup>.

Thus, each sample vector  $x_i$  includes 32 elements. We consider the simplest situation, namely having warning and no having warning. So,  $y_i = \pm 1$ . +1 denotes having warning. -1 denotes no having warning. Then, we regard the latter 5 years data as test data. The expression method of data is to give 1 or 0 to each element of  $x_i$ . 1 denotes positive answer of influence feature. 0 denotes negative answer of influence feature. For example, if the climate is good,  $x_{i2} = 1$ , and if the climate is bad,  $x_{i2} = 0$ . The expression method of test samples is the same as training samples. After we chose training sample, we programmed in MATLAB7 to realize SVM knowledge acquisition and constructing SVM decision-making function and testing data. Test result is list in table 2.

**Table 2.** Test results and original results

Year	Test results	Original results
1999	+1	+1
2000	+1	+1
2001	+1	+1
2002	-1	-1
2003	+1	+1

According to table 2, we can know the test results are completely consistent with original results. This also indicates SVM has good generality capability and better classification accuracy.

## 5. Conclusions

The final goal of support vector machine is to solve a global quadratic optimization problem and it has generality capability. And the early warning system of building safety constructed on the basis of this theory has excellent ability of prediction and can lower the incidence of the incident of the building to a great extent. In addition, it has certain model significance for the building safety management way of our country.

## References

1. Nai-yang Deng, Ying-jie Tian, **The new method of data mining: support vector machine**, Science press, China, 2004
2. Hui Zhang, **Application of support vector machine in data mining**, Computer Engineering, vol 30(6), p. 7-8, March 2004
3. Cristianini N., Shawe-Taylor J., **Introduction to Support Vector Machines**, Cambridge, Cambridge University Press, 2000, 52~767
4. <http://support.vector.net>

<sup>1</sup> Lijun Feng graduated from Agriculture University of Hebei in 2000. Now the author is a PhD candidate of Tianjin of finance & Economics. The specialty of author is the application of computational intelligence method in building safety management. The author has published some paper, as follows:

- 1 Study On Risk Identification Method of Construction Project Based On SVM. Journal of Industrial Engineering and Engineering management. Vol. (19) Supplement, pp11-14, 2005
- 2 Study On Building Safety Evaluation Based On Training Set Of SVM Optimized By GA. Scientific Research Monthly Vol. (10), pp117-121, May 2006
- 3 Study On R&D Project Evaluation Based On LS-SVM. The Proceedings of 20th IPMA World Congress on Project Management. Vol. (2), pp728-731, October 2006
- 4 Improving the Real-Time Character of Prediction System Based On SVM Using RS Theory. Computer Technology and Development. Vol.16 (9), pp30-34, September 2006

<sup>2</sup> **Table 1.** Influence features of building safety

1. Whether the safety equipment is perfect or not.	17. The construction field is good or bad.
2. The climate is bad or good.	18. The machine device runs over load.
3. The worker's physical quality is bad or good.	19. Hoisting equipment and human overlapping work.
4. The special protective measure is bad or good.	20. The machine device is manufactured or fixed perfectly or not.
5. Whether the workers work in dangerous area or not.	21. The construction surroundings is moist or not.
6. The geological data is detailed or not.	22. The timbering system is good or bad.

7. The soil of foundation is liquefied or not.	23. The construction equipment is aging or short of servicing or not.
8. The methane give off from underground or not.	24. The materials body is big and difficult to transport or not.
9. The water on earth's surface or in pipe erode the ground or not.	25. The worker's manipulation is proper or not.
10. Construction disturbs the ground or not.	26. The worker understands the situation of construction field or not.
11. The chemical matter pollutes the soil or not.	27. The workers are short of specialty technique knowledge or not.
12. There is underground barrier or not.	28. The workers are trained or not.
13. The construction result in neighboring pipeline fracture and the dehiscence or not.	29. The inspection persons do their work well or not.
14. The power supply is steady or not.	30. The construction arrangement is good or not.
15. The ray in night is bright or not.	31. There are safety regulations or not.
16. There is dirt, noise, solarization and drench or not.	32. There are technique mistakes or not.