Private colleges teachers evaluation system based on support vector machine (SVM)

Xiao-YanLiu[1,a]

[1]School of Changchun University Of Science AndTechnology,Changchun130600, China.
[a]932684698 @qq.com

Keywords: kernel function of support vector machine (SVM), kernel function, the teacher evaluation , normalization.

Abstract. Private college is the important constituent of higher education system ,To improve the quality of teaching,to promote teachers' quality and enhancement the teachers' teaching ability ,it is necessary to evaluate teacher's ability.Paper is on the basis of the existing teachers' assessment, proposed based on support vector machine (SVM) method to evaluate the teachers' ability.

Introduction

Private colleges made contribution to China's popularization of higher education, made remarkable achievements.Private colleges and universities to meet the diversified demands of the people to accept higher education, for the country to train various kinds of applied talents, deepen the higher education the school system, play a positive role [1]. Private colleges can gain a sustainable development, can win a good social influence, the key is to see the quality of personnel training. To cultivate high-quality talents, it has to have good teachers. Teacher evaluation is a very important content in education evaluation [2].

Using structural risk minimization principle, SVM for cases of statistical learning of small sample problem have very good classification ability and generalization ability [3]. This article will support vector machine (SVM) theory is introduced into the private colleges of the evaluation of teachers, teaching evaluation of teachers.

The SVM learning algorithm

The principle of support vector machine (SVM).SVM from the linearly separable case evolved the optimal separating hyperplane, suppose there are n linear separable sample set x_i, and its categoryy_i,*represented as* $\{x_i,y_i\}$,i=1,2,$\cdots$,n,$x_i \in Rd$,Where n is the number of training samples, d for the dimensions of the training sample,$y_i \in \{1,+1\}$for the category label.For nonlinear can be divided into samples, through nonlinear mapping Φ transfer input variable x into a high-dimensional feature space, and then seeks the optimal classification plane in the high-dimensional space.Fig 1 is two-dimensional visual diagram of sample data.Through the solution of the problem that supports vector machine classifier become one of the common classifier.According to functional theory, a kind of kernel function K (x_i, y_j) as long as satisfy the Mercer condition, the inner product of the number of transform space, namely, K $(x_i, y_j) = \Phi (x_i) \bullet \Phi (y_i)$, without having to know the $\Phi (x)$ transform the specific form. In the original space classification of surface equation w $\Phi (x_i) + b = 0$ should satisfy the constraints

$$y_i[w\Phi(x_i)+b]-1+\xi_i \geq 0 \quad i=1, 2, \cdots n \quad (1)$$

In which w is weight vector of the classification, ξ_i is a slack variable, b is a threshold.Maximize classification interval 1/2||w|| the classification of the surface is called the optimal classification plane. So structure on the optimal classification plane problem into a type $1/2||w||^2 + c\sum_{i=1}^{n}\xi_i$ under the conditions of constraint minimum value problem,which considered in determining the optimal classification plane compromise the minimum fault samples and the largest classification interval.The constant c > 0 is punishment function and use it to control the degree of

punishment right or wrong points and sample.Using Lagrange function to convert the problem into simpler dual problem, namely in the constraint condition

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \ 0 \leq \alpha_i \leq c \quad i=1,2,\ldots n \quad （2）$$ to solve the following function under the maximum

$$LD(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad （3）$$

In which α_i is Coefficient multiplier or optimization for Lagrange. It must meet

$$\alpha_i y_i [w\Phi(x_i) + b] - 1 + \xi_i = 0 \quad i=1, \ 2, \ \ldots n \quad （4）$$

Corresponding $\alpha_i \neq 0$ called support vector.It is support vector samples decided the final classification result.For a given test sample x, support vector machine (SVM) classifier is the optimal classification function is

$$f(x) = \ \text{sign}[\sum_{SVs} a_i y_i K(x_i \cdot x) + b] \quad （5）$$

The Sign () function is symbol, Function f (x) plus or minus can determine points in that category. This is the support vector machine (SVM).

The characteristics of the support vector machine (SVM). The success of the support vector machine (SVM) to a large extent depends on the Kernel function technology,Support vector machine (SVM) model selection problem is given a kernel function, by adjusting the parameters of the nuclear and penalty factor C to improve precision of the support vector machine training, at the same time reduce the error rate.Is used in determining the error penalty parameter C data subspace, adjust learning machine believe the scope and the proportion of empirical risk, in order to make the best generalization ability learning machine . So the parameter selection of support vector machines (SVM) is a direct impact on the performance of SVM.
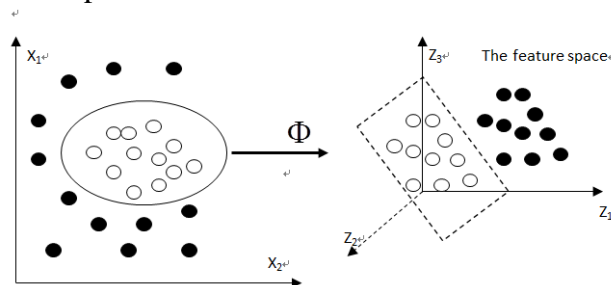

Fig.1sample space to the nonlinear mapping feature space

Run by the local university teachers' evaluation of the SVM methodData collection.

In the teaching evaluation, the evaluation process is generally determined according to the actual situation of school teachers evaluation index, and then according to the index to evaluate teachers' data collection.Can be seen from the data collected if each one teacher evaluation data including n index, can regard it as a n d vector X, can be expressed as $X = (X_1, X_2, \ X_3, \cdots X_n)$ So for the sample data of teaching evaluation can be expressed as:

$$X_i = (X_{i1}, X_{i2}, \ X_{i3}, \cdots X_{in}) \quad i = 1,2,3 \cdots m$$

Among them, n for throughout the evaluation system of evaluation index number, m for teachers involved in the teaching evaluation, Xij to participate in the evaluation of the ith a teacher evaluation index data of the first j a numerical value.

In this paper, according to the current widespread teachers evaluation index established evaluation index system, selection of a private colleges and universities over the years according to the teachers, students, and unfairness feedback that evaluate the most reasonable 208 samples as training sample and test data. Each set of data has 12 evaluation indexes, namely feature vector of each sample data a total of 12 properties.

Teachers' level of evaluation is divided into three levels, including Ⅰ is excellent, Ⅱ for good, Ⅲ is poor.

Normalization and the selection of kernel function. Before using the SVM classification method, the extraction of the characteristics of normalization is very important.For teacher evaluation we normalized processing using the following formula: $x_i' = \dfrac{2(\max(x_i) - x_i)}{\max(x_i) - \min(x_i)} - 1$ Among them x_i , x_i' respectively before and after transformation in value,$\max(x_i)$, $\min(x_i)$ isthe maximum and the minimum sample respectively.This system takes advantage of the evaluate the Libsvm-mat toolbox of Matlab environment.All samples according to the specific formatting LIBSVM, through the function of the sample data to zoom in, range of [1, 1], and then conduct training and classification tests.Which - 1 for poor, 0 corresponding good, 1 corresponds to excellent, thus constitute the 208 sample data.Select one of the 100 groups of data as the training sample, the other as the test sample, call the SVM classifier to training, the training sample data to get a teacher evaluation function.Applied to the evaluation function of data vector is evaluated with future teachers, to determine whether the qualified.

In the selection of kernel function, choosing RBF kernel function respectively, First select RBF kernel function, the penalty factor C experiment, the method is to use the same gamma value.In the selection of gamma value, make is equal to the vector, the reciprocal of dimension n, namely is equal to 1 / n.Application Libsmv-mat, changing of RBF kernel function is the value of penalty parameter C, view the results. The experimental results is shown in table1.

Table 1. RBF kernel function is punishment parameter C experiment

Penalty factor（C）	Kernel function parameters（γ）	Prediction accuracy
$C=2^3$	γ =0.0833	79.5586%
$C=2^5$	γ =0.0833	89.6046%
$C=2^6$	γ =0.0833	93.3568%
$C=2^7$	γ =0.0833	94.3529%
$C=2^8$	γ =0.0833	94.3529%

Training and the evaluation results of the SVM method. We are on the sample data and adopt the same penalty factor C parameters of gamma experiments. The experimental results it can be seen that prediction accuracy in gamma value is limited to 0.01-0.09 on this interval is better. When the gamma value of 0.0833 (1 / n), reached the highest accuracy of 94.3529%. Even after gamma value continues to increase, accuracy and no change again.

Conclusion

Support vector machine (SVM) method for small sample data has good classification ability and generalization ability and learning, better solve the small sample, nonlinear, high dimension and local minimum point and so on practical problems. Teacher ability evaluation can be seen as a kind of to the teachers' ability and their data of the complex nonlinear function relationship between approximation problem. Using vector machine (SVM) to evaluate the teachers' ability to eliminate the man-made factors, to teacher's evaluation is more realistic, but also save the time.

References

[1] Hu Chunsen. "private universities human resources development and management research" Ph.D. Dissertation. [D] 2008 (1), 2-4

[2] Chen Jie. On the management of human resources in colleges and universities. [J]. Journal of huaihai institute of technology (social science edition). 2005 (1), 92-94

[3]　Rocco C　M, Moreno J　A.Fastnontecarlo reliability evaluation using　support　vector machine[J].Reliability Engineering and System Safety, 2002, 76: 237-243.