Studies on the evaluation of college classroom teaching quality based on SVM multiclass classification algorithm

Xinghui Wu, Yuping Zhou and Haihua Xing*

College of Information Science & Technology, Hainan Normal University, Haikou, Hainan Province, 571158 R.R.China

*Corresponding author’s e-mail: szfwxh@qq.com

Abstract: In order to improve the evaluation accuracy and efficiency of the quality of classroom teaching in colleges and universities, combined with the teaching process and evaluation system of colleges and universities, a multi-class classification algorithm is proposed, namely, the Distance binary tree support vector machine based on Euclidean distance (DBT-SVM) algorithm, and the algorithm is used to predict the quality evaluation of classroom teaching in colleges and universities. The algorithm uses the Euclidean distance of the center of the two nearest sample classes to influence the classification, so that the first isolated classes can be separated first. Experimentally, the algorithm can improve the efficiency and accuracy of classification and solve the problem of multi-class classification.

1. Introduction

The quality of classroom teaching in colleges and universities is an important part of teaching management, and how to evaluate the teaching of teachers is a subject that everyone pays attention to[1].

The evaluation of the quality of classroom teaching in colleges and universities directly affects the evaluation of teachers' teaching effectiveness and students' learning effect, so it is necessary to establish a rapid, accurate, effective and objective evaluation system model of teaching quality [2]. The evaluation of teaching quality in colleges and universities is a multi-level, multi-objective optimization problem, the evaluation content is many, there are many factors affecting the evaluation model, so in the establishment of the evaluation model, we need to consider the various components of the evaluation model carefully, but also to their combination of comprehensive assessment. At present, the evaluation of teaching quality in colleges and universities mainly includes: multi-linear regression, gray correlation method, hierarchical analysis method, fuzzy clustering analysis method, neural network method and support vector machine [2-4]. Neural network is a kind of artificial intelligence technology with strong classification ability, which can improve the accuracy of teaching quality, but because of its limitations in the application process, it is easy to fall into local extreme points, and the evaluation results and actual results will not match, and the efficiency will be low. Support vector machine is a small sample, nonlinear approximation ability of machine learning method, but in the evaluation of the quality of classroom teaching in colleges and universities, the accuracy of its evaluation has a great relationship with the selected parameters. In order to obtain the evaluation results with high accuracy, it is necessary to optimize the parameters of the support vector machine evaluation model, and finding the parameters of the optimization support vector machine is the key to improve the evaluation of the model.
In order to improve the accuracy of classroom teaching quality evaluation in colleges and universities combined with the evaluation method of classroom teaching quality in colleges and universities, a multi-class classification method based on support vector machine is proposed. This method uses a binary tree support vector machine method based on Eucalyptus distance. Experiments show that this method can improve the accuracy of teaching quality evaluation in colleges and universities[5, 6].

2. BT-SVM

2.1. BT-SVM multiclass classification

The binary tree support vector machine multiclass classification algorithm (BT-SVM) is to divide all classes into two categories, and then subclasses into two subclasses, so that the loop, until all nodes contain only a separate category, and then at each non-leaf node to train a two-value SVM classifier, when the structure of the binary tree is close to the normal binary tree can achieve the best training speed and classification accuracy.

In the multi-classification problem, the BT-SVM classification algorithm can not only effectively overcome the inseparable problem, but also greatly reduce the number of classifiers, because for the k-class problem, BT-SVM only needs to construct k-1 two-value classifier, but this method also has two more serious problems: First, the binary tree structure is uncertain, for the same multi-class classification problem, there can be many different two-fork tree structure, and different binary tree structure will get different classification models, Figure 1 shows two binary tree structures commonly used in four types of problems. Second, this method may lead to the phenomenon of "error accumulation" [7-9], that is, if the classification at the upper node once misclassified, the error will be passed on, the subsequent nodes will lose the meaning of classification, so the BT-SVM upper node sub-classifier has a greater impact on the performance of the entire classifier. In order to obtain the best classification effect, it is necessary to construct a more reasonable binary tree structure according to the actual situation, that is, to construct a better binary classification tree [10, 11].

![Figure 1. Common binary classification tree structures.](image)

Because the sub-classifier of the upper node of BT-SVM has a great influence on the whole classifier, it is generally the class that is easiest to separate first, and then the class that is more difficult to separate, so that the classification error is as far away from the root node as possible, so as to obtain the optimal binary tree hierarchy. The commonly used optimal binary tree generation algorithm is the shortest distance method[12], the idea of which is to treat the Eugene distance of the two nearest sample vectors between two classes as the distance of the class, so that the classes with the Euclidean distance from the other classes are first separated, and the difference between the two classes is measured by the distance between classes. However, this method only considers the distribution of samples without considering the distance between classes, and finally obtains a partial binary tree, it is difficult to generate a complete binary tree, so that classification accuracy and classification efficiency are relatively low. To this end, this paper adopts the Distance binary tree support vector machine (DBT-SVM) multiclass classification algorithm based on Euclidean distance.
2.2 DBT-SVM Multiclass classification algorithm.

The specific algorithm is as follows:

Step 1: Put different classes into collection $E$ in ascending order by class label;

Step 2: If there are $k$ categories in the sample set, if $c_i$ and $c_j$ are the sample centers of class $i$ and class $j$, then the sample center Euclidean distance between class $i$ and class $j$ is $s_{ij}$,

$$s_{ij} = ||c_i - c_j||$$

The matrix $S$ of the $k$ class classification is:

$$S = \begin{bmatrix}
1 & 2 & s_{1,2} \\
1 & 3 & s_{1,3} \\
\vdots & \vdots & \vdots \\
k - 1 & k & s_{k(k-1)}
\end{bmatrix}$$

Where, the first column of the matrix is listed as the label of class $i$, the second is listed as the label of class $j$, and the third is listed as the relative distance of class $i, j$.

Step 3: Find the two categories $i$ and $j$ with the largest distance in set $E$ in Matrix $S$, deposit $E_1$ and $E_2$ according to size, and create an SVM classifier at the node of the binary tree, so that $E = (E_1 \cup E_2)$;

Step 4: If $E = \emptyset$, go to Step 6;

Step 5: Find the various types of minimum class center Euclidean distance $s_{im}$ and $s_{jm}$ of class $m$ (where $m \in E$) to $E_1, E_2$ in $S$, if $s_{im} \leq s_{jm}$, then add class $m$ to collection $E_1$, otherwise add to collection $E_2$, and turn to step 2.

Step 6, respectively, the collection $E_1$ and set $E_2$ as the left and right sub-tree of the binary tree, complete a class ii classification.

Step 7, assume $E=E_1$, until each sub-tree contains only one category that can no longer be divided, at which point the category is used as the leaf node of the binary tree, and the algorithm ends. Otherwise, go to step 2 and further divide the left sub-tree into 2 sub-trees.

Step 8, assuming $E=E_2$, until each sub-tree contains only one category that can no longer be divided, at which point the category ends as the leaf node of the binary tree. Otherwise, go to step 2 and further divide the right sub-tree into 2 sub-trees. Until each class becomes the leaf node of the binary tree.

3. The evaluation model of classroom teaching quality based on DBT-SVM

3.1. Determination of the evaluation index system.

After in-depth study and analysis of the actual teaching situation in domestic colleges and universities, this paper follows the principle of establishing the teaching evaluation index system, using the method of establishing the evaluation index system to develop a set of more reasonable evaluation indicators of the quality of teaching in colleges and universities, this paper from the teaching attitude, teaching content, teaching methods and teaching results to build the quality evaluation index system of classroom teaching in colleges and universities, as shown in Table 1:

| Table 1 | The classroom teaching quality indicators in colleges and universities. |
|---------|---------------------------------------------------------------------|
| Evaluation indexes | Evaluation contents | Score |
| teaching attitude | The class is serious and full of spirit. | 10 |
| | Get in/out of class on time and don't use your cell phone during the course of class; | 7 |
| | Sincerely care about students, both inside and outside the class are willing to answer questions and answers. | 8 |
| Content of courses | The emphasis is prominent, the details are appropriate, the difficulty, the depth is appropriate. | 10 |
| | | 8 |
Theory is related to practice, and the teaching content reflects the progress of the subject.  

Teaching methods are flexible and diverse, can be used appropriately with multimedia, and fit the teaching content. The explanation is easy to understand, enlightening and interacts well with the students. The homework layout is reasonable, and effective review and feedback.  

Students are highly motivated and have a good atmosphere in class. Through the teacher's teaching, students have mastered the basic knowledge and skills of the course. Through this course, students' interest in learning and learning ability have been improved.  

In this paper, the evaluation results are divided into 5 ranks, 90 to 100 points is considered as class A, 80 to 90 is as B, 70 to 80 is as C, 60 to 70 is as D, below 60 is as E.

3.2. The evaluation model of classroom teaching quality based on DBT-SVM.

In order to solve the problems existing in the traditional teaching quality evaluation model of colleges and universities, this paper puts forward a multi-class classification teaching quality evaluation model based on the binary tree support vector machine after in-depth study of the whole teaching process and teaching quality monitoring system of colleges and universities, which uses the binary tree support vector machine idea based on eucalyptus distance to combine multiple two-value classifiers to solve the multi-classification problem of teacher's teaching quality evaluation, so as to improve the efficiency and accuracy of teaching evaluation. The following workflow diagram is based on the evaluation of the quality of college classrooms based on the binary tree support vector machine, as shown in Figure 2:

![Figure 2](image_url)  

**Figure 2.** Flow chart for the evaluation of teaching quality in colleges and universities.

4. Evaluation experiments of the classroom teaching quality based on DBT-SVM.

4.1. The source and processing of the experimental dataset.

The data set of this paper is from the evaluation table of the teaching quality of teachers in a certain semester for students in a university in Hainan, and there are 180 data. By the evaluation criteria given in Table 1, students are asked to rate each item in the evaluation scale at the end of this semester's course, and to give the corresponding evaluation score according to the score range of 0 to 100. Of these, 150 are used as training sample sets and the remaining 30 as test sample sets. Each sample
contains 12 feature attributes, that is, a feature attribute corresponding to an evaluation factor in the evaluation indicator system, each of which has a value range of 0-100. Table 2 below shows sample data that was not processed by some datasets:

| No. | x1  | x2  | x3  | x4  | x5  | ... | x10 | x11 | x12 | y  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| 1   | 84  | 75  | 70  | 80  | 84  | ... | 73  | 75  | 76  | A  |
| 2   | 70  | 82  | 73  | 61  | 73  | ... | 65  | 60  | 72  | B  |
| 3   | 80  | 75  | 65  | 72  | 82  | ... | 83  | 83  | 63  | B  |
| 4   | 83  | 78  | 71  | 91  | 82  | ... | 87  | 93  | 79  | C  |
| 5   | 75  | 66  | 72  | 74  | 86  | ... | 60  | 93  | 88  | E  |
| 6   | 65  | 71  | 67  | 81  | 74  | ... | 88  | 75  | 76  | E  |
| 7   | 78  | 86  | 76  | 83  | 80  | ... | 86  | 86  | 90  | A  |
| 8   | 90  | 84  | 83  | 79  | 92  | ... | 87  | 94  | 80  | A  |
| 9   | 62  | 76  | 67  | 80  | 89  | ... | 80  | 63  | 63  | C  |
| 10  | 62  | 78  | 69  | 76  | 81  | ... | 83  | 71  | 68  | D  |
| 11  | 94  | 89  | 87  | 90  | 88  | ... | 78  | 80  | 79  | A  |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ...
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ...
| 178 | 61  | 83  | 80  | 82  | 66  | ... | 80  | 75  | 64  | E  |
| 179 | 72  | 78  | 86  | 77  | 93  | ... | 75  | 97  | 92  | A  |
| 180 | 78  | 81  | 76  | 81  | 86  | ... | 74  | 64  | 74  | B  |

In order to ensure the accuracy and effectiveness of training and test results, the data needs to be normalized to reduce its scope between 0 and 1.

4.2. Realization of the teaching evaluation model

The experiment in this paper is to use python3 implementation under the environment of anaconda3, using the scikit-learn algorithm package of machine learning, the nuclear function in the support vector machine in the experiment is using the Gauss radial base core function, the parameters are set as follows:

$$K(x_i, x_j) = \exp \left( -\frac{\gamma \|x_i - x_j\|^2}{2\gamma^2} \right)$$

Penalty parameters C and core function parameters γ are selected using grid search method, and the optimal combination of these parameters is found by 5 fold cross-validation, then determine the range of C and γ is \(2^{-5}, 2^{10}\).

4.3. Results and Discussion

In order to evaluate the advantages and disadvantages of the results obtained by this model, DBT-SVM and SVM are classified and tested with SVM and BT-SVM algorithms respectively, and the accuracy results of the three algorithms and the time consumption in the experiment are obtained. The results show that the DBT-SVM algorithm is higher than the SVM and BT-SVM algorithms in predicting accuracy. In terms of algorithm time consumption, the SVM algorithm takes a little longer, but the time difference between the BT-SVM algorithm and the DBT-SVM algorithm in the classification process is very small. The DBT-SVM algorithm avoids the phenomenon caused by the accumulation of errors in the structure in the partial binary tree in the BT-SVM algorithm, and also avoids the indestructible area in the SVM algorithm, thus reducing the accuracy of classification. The DBT-SVM algorithm combines the concept of class distance in clustering with the concept of approximately complete binary trees, which allows the first isolated classes to be dissombered at the upper nodes, thus improving the accuracy of division. The results of the experiment are shown in Table 3.
### Table 3 Evaluation accuracy and time required.

| Algorithm | Accuracy (%) | Training time (ms) | Testing time (ms) |
|-----------|--------------|--------------------|-------------------|
| SVM       | 94.45        | 26.54              | 8.43              |
| BT-SVM    | 95.53        | 22.43              | 6.54              |
| DBT-SVM   | 96.46        | 21.56              | 5.83              |

5. **Conclusions**

Classroom teaching occupies a very important position in the process of teaching, and scientific and reasonable evaluation of the quality of classroom teaching is the guarantee of the development of education and teaching. To evaluate the quality of classroom teaching, not only the teacher's teaching ability, methods, attitude and teaching content, but also through the students' learning effect and evaluation of teachers' teaching effect. In order to improve the evaluation accuracy and efficiency of the quality of classroom teaching in colleges and universities, this paper puts forward a method for evaluating the effectiveness of classroom teaching in colleges and universities based on the DBT-SVM multi-category classification algorithm. Based on eucalyptus distance binary tree support vector machine multi-class classification algorithm, the calculation algorithm designs an evaluation model that conforms to the quality of classroom education and teaching in modern colleges and universities, simulates the data and analyzes the experimental results. The results show that the teaching quality evaluation model established by the improved algorithm in this paper has high accuracy and time efficiency for teacher classification, and it can be verified that the model is feasible.

### Acknowledgment

Thanks are given to the Education Department and Science & Technology Department of Hainan Province for funding the studies of this paper (QJY20181071, Hnjg2020-31, 2019RC182).

### References

[1] Hu CS, Zhao XQ, Liu ZY 2001 *J. Gannan Medical College* (in Chinese) 1 76.
[2] WU EY, Lv J 2016 *J. Chongqing Normal University (Natural Science)* (in Chinese) 33(3) 103.
[3] Li YL, Su YD 2014 *Computer Technology and Development* (in Chinese) 7 181.
[4] Li YL. 2014 The application of BT-SVM in teaching quality evaluation (in Chinese). *Thesis for master degree, Guangxi University.*
[5] Cheong S 2004 *Neural Information Processing-Letters and Reviews* 2(3) 47.
[6] Fan BC, Wang JY, Bo YM 2010 *Computer Engineering and Design* (in Chinese) 31(12) 2823.
[7] Hsu CW, Lin CJ 2002 *IEEE Transactions on Neural Networks* 13(2) 415.
[8] Zhao L 2014 *Computer Applications and Software* (in Chinese) 31(12) 233.
[9] Xia SY, Pan H, Jin LZ 2009 *Computer Engineering and Applications* (in Chinese) 45(17) 167.
[10] Song ZQ, Chen Y 2015 *J. Naval Aeronautical and Astronautical University* (in Chinese) 30(5) 442.
[11] Chen LJ 2005 *Vocational Education Research* (in Chinese) 7 150.
[12] Zou SG, Su RN 2010 *Computer Simulation* (in Chinese) 27(11) 314.