

Study on application of Fuzzy theory and Principal Component Analysis in teaching evaluation

Hongfu Ai^{1,a}, Ying Shi^{1,b}

¹Information teaching and Management Center of Jilin Agricultural University, Jilin Changchun, 130118

^aaixin1115@163.com, ^b7833949@qq.com

Keywords: Big data; Data mining; Teaching evaluation

Abstract. In the era of Big Data, according to the characteristics of basic computer courses, the author reasonably applied education data generated in the process of teaching. And the author made an in-depth analysis of the data collected by using the fuzzy theory and principal component. It provides a scientific basis for decision making to improve basic computer education, and improves the quality of teaching basic computer. This study reflects data can drive decision and revolutionize education too.

Computer basic education under the big data

Higher education as an important part of the social subsystem was heavily influenced by big data. Data mining education has become an important way to improve the level of teaching management and quality in developed countries [1]. The overall goal of computer basic education in university is popularization of computer culture, cultivation of professional application ability and computational thinking ability training [2]. It plays an important role in higher education, thus improving the quality of teaching basic computer is necessary.

Teaching evaluation

With the rapid development of computing education, the traditional evaluation method is difficult to reflect the actual teaching. It is necessary to dig out valuable information from vast amounts of educational data, and reasonably applied to teaching evaluation work. The purpose of assessment is to improve teaching methods, adjust teaching content and improve the quality of talent cultivation [3]. An in-depth analysis of the data can provide decision makers with reliable scientific basis for decision making; can accurately formulate reasonable teaching methods; develop a more realistic educational strategy to improve the quality of teaching.

The principal component analysis

Principal component analysis using the idea of dimension reduction turns many indicators into few comprehensive indexes [4].

The calculation steps of principal component analysis:

1. Calculate the covariance matrix: $\sum (s_{ij})_{p \times p}$ Among them:

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i) (x_{kj} - \bar{x}_j) \quad i, j = 1, 2, \dots, p \quad (1)$$

Calculate the eigenvalues of $\sum \lambda_i$ and the corresponding orthogonal eigenvectors of α_i units

Variance contribution rate (information) of the main components reflects the size of information, α_i as follows:

$$\alpha_i = \lambda_i / \sum_{k=1}^m \lambda_k \quad (2)$$

2. Select main component

To select a few principal components, F_1, F_2, \dots, F_m which are determined by the variance (information) accumulative contribution rate of $G(m)$.

$$G(m) = \sum_{i=1}^m \lambda_i / \sum_{k=1}^p \lambda_k \quad (3)$$

When the cumulative contribution rate is more than 85%, it can sufficiently reflect the original variable information. The corresponding M is the m principal component extraction.

3. The calculation of principal component scores

The formula is as follows:

$$F_i = \alpha_{1i} X_1 + \alpha_{2i} X_2 + \dots + \alpha_{pi} X_p \quad i = 1, 2, \dots, m \quad (4)$$

Fuzzy theory

The concept of the fuzzy set is presented by L. A. Zadeh, a USA cybernetics expert [5.6]. Fuzzy set took two absolute concept in ordinary set theory of flexibility, which is u either to A or does not belong to A , by the membership function instead of the absolute "belongs to" or "does not belong to the" relationship. Its membership arbitrary real values between 0 and 1

$$0 \leq \mu_A(u) \leq 1 \quad (5)$$

$\mu_A(u)$: A (fuzzy subset) membership function. It represents a collection of any element in the u membership degree of fuzzy subset of A . We can select according to the requirements of different λ (confidence level) to determine the relationship of subordination which can be expressed as :

$$0 \leq \lambda \leq 1 \quad (6)$$

When $\mu_A(u) \leq 1$ is $u \in A$, else $u \notin A$. λ is from 1 to 0, A is gradually expanding. Therefore Fuzzy subset A is a set of boundaries with vacillation. It will become larger when λ becomes smaller. The membership function in the practical application of the selection is the key.

The experimental data acquisition and analysis

Setting of teaching evaluation under the constructivist theory The questionnaire investigation of the basic teaching of computer was carrying out for grade 2013 of Jilin Agricultural University students. The question in the questionnaire is 10 assessment indicators of teaching-related. Respectively, X_1 represents computer terminology accurately, X_2 represents the interaction of teacher and students, X_3 represents teaching with practice, X_4 represents highlight the practical technology, X_5 represents clear context, X_6 represents rich knowledge, X_7 represents explaining profound theories in simple language, X_8 represents paying attention to the efficiency, X_9 represents reasonable use of teaching methods, X_{10} represents stimulating students interest.

Here is a sample of 10 data of mining analysis. Part of the data is shown in Table1.

Table1. The survey data

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	8	6	9	7	9	7	7	8	8	9
2	8	9	8	6	6	6	7	3	5	0
3	9	7	9	8	9	8	8	6	8	8
4	2	0	2	0	0	2	1	9	8	4
5	9	8	7	8	8	9	8	7	8	9
6	9	8	9	8	2	6	4	9	0	0
7	9	8	8	9	9	8	7	8	7	8
8	2	2	8	2	7	0	8	9	2	2
9	9	7	8	8	8	7	8	8	8	9
0	0	8	9	9	8	8	2	1	0	0
1	9	8	8	8	9	8	9	7	8	8
2	8	4	6	6	0	2	9	8	8	9
3	8	8	8	9	9	9	8	8	7	8
4	2	9	2	2	0	6	0	9	4	2
5	8	7	8	8	8	7	8	7	8	9
6	6	8	8	8	6	0	6	0	9	0
7	9	8	9	9	7	8	9	7	9	9
8	6	6	2	0	6	8	2	8	2	2
9	1	8	7	9	9	8	8	9	8	8
0	8	9	0	4	6	6	6	4	9	8

Selecting 10 sets of data samples, x_{ij} represents the i -th student in the j -th index scoring results (Highest score is 100 points), $X = (x_{ij}) (10 \times 10)$ thus a raw data matrix of 10×10 is obtained.

The results of PCA The principal component analysis was used for the data matrix of table. According to the cumulative contribution rate [7], it selects the fourth principal component F_4 , because its cumulative contribution rate has reached 92.38%, and fully meets the requirement. The main components of the original evaluation expressed as follows:

$$F_4 = -0.229X_1 - 0.006X_2 - 0.289X_3 + 0.088X_4 + 0.362X_5 + 0.550X_6 + 0.310X_7 \quad (7)$$

Fuzzy principal component analysis results Depending on the application, the actual indicators of different dimension, so the influence of dimension should be eliminate before the principal component analysis. Here the application of fuzzy theory was used to deal with the original data, eliminate the data dimension, and make data between 0 and 1. Principal component analysis was used for processed data to choose fourth principal component F_4 (cumulative contribution rate has reached 95.24%) according to corresponding contribution rate. The main components of the original evaluation expressed as follows:

$$F_4 = -0.126X_1 - 0.238X_2 + 0.245X_3 + 0.124X_4 + 0.062X_5 + 0.325X_6 - 0.007X_7 - 0.220X_8 - 0.324X_9 + 0.228X_{10} \quad (8)$$

Comparison and analysis of the experimental results By original PCA, we get 5 indexes to evaluate the computer teaching which play a major role in 10 indicators, X_4 represents highlight the practical technology, X_5 represents clear context, X_6 represents rich knowledge, X_7 represents explaining profound theories in simple language, X_8 represents paying attention to the efficiency. From fuzzy principal component analysis, the 5 indicators can evaluate effectively which is X_3 represents teaching while practice, X_4 represents highlight the practical technology, X_5 represents clear context, X_6 represents rich knowledge, X_{10} represents stimulating students interest

Two methods have reduced the teaching evaluation index. However, this research carries out on computer teaching, according to its own characteristics, especially pays attention to the actual operation ability and creativity. So to be more scientific, reasonable and in line with the objective reality, we put X_3 which is teaching while practice and X_{10} which is stimulating students interest into the main evaluation index. So the principal component analysis that based on fuzzy theory is more effective to dig out the index system of evaluation of computer teaching.

Conclusion and Prospect

Although some achievements have been obtained, but the application of data mining technology starts late in education and teaching, there are many studies to carry out. This study applies the method of principal component analysis and fuzzy principal component analysis method, sets the evaluation index of computer teaching under the constructivism theory, and carries on in-depth excavation, gets the rational analysis result. parallel computing ability and large storage capacity of cloud computing bring convenience for educational data mining , also make the traditional data mining based on data of education becomes more difficult and complex. In the future, with the rapid increase of education data, parallel data mining of cloud computing in education will be more and more important.

References

- [1] Bu Yang. Application of association rules in computer teaching [J]. Jia Mu Si College Journal, 2012.5.158-159
- [2] Siemens G. Learning and Knowledge Analytics-Knewton-the future of education? [EB/OL]. [2011-04-14].<http://www.learninganalytics.net/?p=126>
- [3] Zhao Ling-li. Construction of Higher Education Quality Which is based on Cloud Computing and Big Data [J]. Fu Dan Education Forum 2013.vol.11, No.6.
- [4] Zhang Xinhua. Data Warehouse and Data Mining Technology [M]. Beijing: Qing Hua University press, 2011.6.
- [5] Lin Heping, Liu Dinghui, Bao Naiyuan. Grey Relation Analysis and its Application [J].Journal of Jilin University, Information Science Edition.2007, 25(1):84-90.
- [6] Zhan Sheng, Xu Yuanxin, Shi Yongquan, Prediction of vehicle trajectory based on fuzzy colored Petri net [J].Computer Engineering and Applications, 2014, 50(3):227-231.
- [7] Du Shiyu, Qi Jiayin. Research on the Evaluation of Micro blog Topic Influence Index Based on PCA Methods [J].Journal Of Intelligence.2014, 5, 33(5):129-135.