RESEARCH ON IMPROVING PROCESS ASSESSMENT SYSTEM OF DISTANCE EDUCATION BASED ON DATA MINING

Yao Jun, Cheng Hua, and Ying Wei Yong
Distance Education Institute of East China
University of Science and Technology
China

Abstract

In order to promote students' interests in courses, and enhance teaching design for management, the process assessment system is applied to the student's whole learning process, such as attending lectures, assignments, discussions and examination. The parameters in assessments system are usually designated via experiences. According to the normal distribution of population, we utilize the k-means to cluster the parameters of assessment, and discuss the relationship between the parameters and the final score. It is concluded that the assessment system need to be perfected further, and more help should be offered for the distance learners.

Preface

In China, the exploration of Internet-based distance education lasts for eleven years. As a new mode of life-long education, it not only promotes the education level of the society, but also provides the comprehensive and convenient mode for the construction of a learning society. However, now there are many problems in the distance education in China. For example, in the teaching process, the subject construction does not completely match the distance education, and the learning mode does not match the learning evaluation mode. All of these make the teaching and learning quality unsatisfied for the learners and teachers. Some researchers propose that it is essential to strengthen the whole process monitoring, and it includes three aspects: monitoring input quality, running quality and output quality.

Process assessment is important for the whole process monitoring on teaching and learning quality. It is an evaluation of learning process of the students, and the instruction on the learning process and improvement on the teaching are emphasized. Black (1998) thinks that the information collected in the formative evaluation (similar with the process evaluation) is useful in adjusting the teaching and learning process. Boston (2002) points out that formative evaluation includes inspection of the teacher, discussion in the classroom, and analysis of the learning process (including homework and examination). The Assessment Reform Group, ARP UK, issued ten principles of formative evaluation.

In the practice of distance education, the educators found that it's important to construct the learning evaluation system, which is different from that in the traditional education mode. FAST (The Formative Assessment in Science Teaching) undertaken by OU (Open University of UK) has developed the online formative evaluation framework. In the Open University of Hong Kong, students must pass both formative assessment and terminative examination in some courses. The weight of the formative assessment in the final score is different in different countries and districts. In some Open Universities in India, South Korea, Pakistan, Sri Lanka, it's 30%, and in the Open University of Hong Kong, it's 50%, and in China Central Radio and TV University, it's 20%. However, it's inevitable to increase the weight of formative assessment in the distance education (Chen Nai Lin, 2006).

In China, there are 67 universities engaged in the modern distance education trial work. In a broad sense, the concept of formative assessment includes all the activities of teachers and students, and the activities can be used diagnostically to alter teaching and learning (Huang Rui-hong, 2006). However, in this passage, we mainly discuss the learning process of the students, and we call it process assessment system.

Now many methods are provided for the distant learner, such as courseware through the Internet, discussion in BBS, question and answer on Internet through audio and video, and homework in the Internet. The question is how to gain our ends by these methods, and how to reflect this to the final assessment (Zhang Jianhua, 2007; Zhu Yu-ping, 2006). We have established the evaluation system based on process assessment. In practice, the criteria system including six parameters was adopted to evaluate the learner, and the six parameters were set to different weight to calculate the final score.

This paper is organized as follows: the step of k-means algorithm is briefly introduced, the process assessment system is proposed, an illustrative example is provided from ECUST (East China University of Science and Technology), and some conclusions are drawn.

Cluster Analysis

According to the principle of "maximizing the similarity in the cluster, minimizing the similarity among the clusters," the data objects are grouped into different classes or clusters. The objects in the same cluster have higher similarity, while the objects in different clusters have lower similarity. Clustering analysis mainly includes partition-based algorithm, hierarchy-based algorithm, density-based algorithm, grid-based algorithm, and model-based algorithm. K-means belongs to

the partition-based algorithm, and its merit is simple algorithm structure and fast convergence rate, and fitness for the large-scale data analysis (He Ling, 2007).

In the iteration of k-means algorithm, the center of every cluster C_i in previous iteration was calculated as the seed of current iteration. The procedure is as follows:

1. Select k different data objects as the center of k clusters;
2. Assign every data objects to the class which represent the nearest cluster, and k clusters are formed, then the cluster criteria function is calculated;
3. Calculate the center of every cluster in previous iteration, and regard it as the center of the current iteration;
4. If these centers are the same as the centers in the previous iteration or the value of cluster criteria function is lower than the threshold value for more than twice, then terminate or carry out the next iteration. (Jiawei Han, 2007)

The center of k-means means the averaged value of the object in the cluster. Since cluster algorithm is an unsupervised algorithm, the performance evaluation function is used for the validity judgment of the cluster output. In k-means, the Euclidean distance between the object and the center or its square deviation is used to evaluate the cluster, and the definition is:

$$E = \sum_{i-1}^{k} \sum_{p \in C_i} |p - m_i|^2$$

E represents the sum of all the square deviation of the objects, and p is a point which represents the object, and m_i is the averaged value of cluster C_i (p and m_i are both multidimensional)

Process Assessment System

The fair-and-square and all-around evaluation for the learning process of the students is very important. It's not only the evidence of the scholarship and entitled certificate, but also the motivation of the student on study. Especially in the distance education, the target population is the amateurish learner. The evaluation mode in traditional university education mainly depends on the final exam score, and is obviously not fit for the students in distance education. For these students, the following factors should be taken into account: the learning process, the capability to learn and to implement in the work. So a new learning

evaluation system (or as we call it "process assessment system") must be established.

The approaches provided for the learners are also the reference criteria in the process assessment. The following factors could be the reference criteria: the log of the count and the time length of the courseware study (the granularity of navigation map in courseware for the learner is a week, and the granularity of content is a study guide in every chapter); the valid post in the study community (we have established the study community for every course, and the learner can put question here and there are tutors to answer the questions and confirm the valid post); the record of the online consultation through audio and video based on the Internet; the record provided by the learning centers of attending class and discussion and the evaluation of their activity; the score of the phase exercises and homework based on Internet.

In our teaching and learning system, we trace the learning process through the record of courseware study, online homework, examination, BBS and so on. According to the experience in practice, we adopt the criteria system with six parameters, including the count and time length of the courseware study, valid post number on BBS, regular evaluation from the teacher, scores on the final exam and scores in the online homework. The six parameters were set to different weight to calculate the final score. In need of data mining analysis, several fields in the database from different table were picked up. The meaning of the criteria is shown in Table 1.

Table 1: Parameters of Learning Process on One Course

Parameter	**Unit**	**Weight in final score**	**Remark**
the count of the courseware study	times	10%	Not less than 10 times in a semester
time length of the courseware study	Minutes		Not less than 15 minutes
Regular evaluation from the teacher	Score	10%	
valid post number on BBS	Number of the post	5%	Not less than 5 times
scores in the online homework (1)	Score	7.5%	
scores in the online homework (2)	Score	7.5%	
scores on the final exam	Score	60%	

Since the object is the cluster characteristic of learning process in one course with statistical meaning. The characteristic of learner and course should be hidden. The six parameters almost cover the learning process, including having classes, attending discussion, test and examination. The weight of the criterion parameters are set for the calculation of the final score.

However, the process assessment system is established by experience, and it is essential to analyze and prove the validity of the system, especially the parameter setting. We try to find the validity and the aspect to be improved by cluster analysis.

Experiment

If the quality of the data source to analyze is poor, it's difficult to guarantee the output of the data mining. Usually, there is noise, data loss, data inconsistency in database, so the data pretreatment is necessary.

Data Pretreatment

The steps of Data Pretreatment are follows:

1. Data clean-up. Firstly, the fragmentary data should be disposed. The absent record in six parameters, for example, the record which is zero in the count of the courseware study, should be deleted. That is to say, the object is the learners who have really taken part into the distance education. The next step is dealing with the data noise. The data noise is the data which deviate from the cluster badly or do not conform to the reality. There are two ways to get rid of the noise. The first is to set the critical value, for example, all the records of the time length in the courseware study which is higher is more than 100 hours should be deleted; the second is to delete the data which deviate from the cluster obviously by the pre-cluster method.

2. Data reduction. Data reduction is a measure of data pretreatment in data mining. The data set after reduction is much smaller than the original. However, the integrity of the data is maintained furthest, that is to say, the analysis result is similar with the one which uses original data set and has higher computation complexity. The strategy of data reduction includes the data cube collection, property subset choice, latitude reduction, value reduction, discretization, concept hierarchy and so on. In the experiment, the non-layback Simple random sampling is adopted, and finally 1127 records without characteristic of the course

and 100 records in *ethic* course are decided to taken into the final computation.

3. Standardization. This is a required step of data pretreatment for cluster analysis. Since the Euclidean distance has to be computed for the evaluation, the step to eliminate the dimensions brought by different parameters is essential. So we standardize the property data and make them fall into the region, and the standardization formula is $X_{ij} = \frac{x_{ij} - \overline{x}_j}{S_j}$.

K-means Clustering Result

Firstly**,** it is an important technique to choose the appropriate cluster count. K-means would select initial centers from K clusters, and it is badly influential to change the count of the selected clusters. The disposal strategy of k-means is to experimentalize repeatedly to decide the most appropriate K value. In the process to select the K value, we could judge the estimation of the cluster result by analyzing the center square deviation parameter.

Based on the physical meaning of the clusters in reality, we could also set the initial K value purposefully. Here, we select three kinds of students: excellent, common, inferior, and this is based on the characteristic of normal distribution of population. So we set k = 3 as the base of initial cluster count. In the experiment, we have tried to set K = 3 and K = 5, and found that the two choices could classify the learning condition of the students commendably. Every cluster was classified further, and shown in Table 2, ordered by the estimation of whole behavior.

Table 2: Comparison of Different Clusters Based on 1127 Records without Characteristic of the Course

estimation of whole behavior	K=3		K=5		
	Record count	proportion	Record count	proportion	Proportion after Consolidation
Excellent	163	14.5 %	52	4.7%	11.4%
			75	6.7%	
Common	658	58.4 %	473	41.9%	72%
			340	30.1%	
Inferior	306	27.1 %	187	16.6%	16.6%

So in general, the learner can be classified into three groups:

- Spend much time on study, excellent (about 10%–20%)
- Study according to the basic demand, common (about 60%–70%)
- Spend little time on study, inferior (about 10%–20%)

Meanwhile, if the count of the cluster is not enough, the result set would be too rough, whereas we could find more details for the behavior of the learner, and provide more suggestions for the teaching setting and construction of the assessment system. So the following analysis would set K = 5 in clustering.

To know the learning status of the students better, except the requisite fields for process assessment system, we also found that the logon count in the learning system and BBS also reflects the status of the learning process, so it was regarded as the property of the cluster. We have make k-means clustering for the 1127 records without characteristic of the course and the 100 records in *ethic* course (see Table 3 and Table 4 below).

Table 3: Clustering Result Based on 1127 Records without Characteristic of the Course

Serial number of Cluster	Averaged Count for courseware study	Time length	Count of post in BBS	Score in exam	Score in homework 1	Score in homework 2	Logon Count in the learning system	Logon Count in BBS	Count of records
0	14.2	281.17	10.57	77.92	89.79	79.75	418.27	20.32	75
1	10.77	231.65	5.09	74.68	93.07	86.33	244.32	4.51	473
2	10.92	215.88	5.18	63.25	59.37	44.44	180.11	3.36	187
3	12.16	285.35	5.48	82.99	78.57	63.64	252.79	4.35	340
4	31.71	1150.39	5.98	78.38	88.08	75.63	410.13	7.58	52

Table4: Clustering Result Based on 100 Records in *Ethic* Course

Serial number of Cluster	Averaged Count for courseware study	Time length	Count of post in BBS	Score in exam	Score in homework 1	Score in homework 2	Logon Count in the learning system	Logon Count in BBS	Count of records
0	10.58	284.63	5.24	77.08	72.37	58.03	229.21	3.5	38
1	15.45	253.75	5.55	76	90.5	79	370.25	4	20
2	11.67	321.33	12.33	73	90	81.67	590.33	27.33	3
3	9.08	183.58	4.92	65.25	48.75	40.42	158.25	1.92	12
4	9.33	132.56	5.59	73.11	93.7	87.41	233.56	3.29	27

Discussion

From the k-means clustering analysis result, some conclusion can be drawn.

The Count and Time Length of Courseware Study

From Table 3, we can get the average time length of all the clusters and their score in examination. It is shown in Table 5.

Table 5: Count and Averaged Time Length of Courseware Study Based on 1127 Records without Characteristic of the Course

Serial number of Cluster	Averaged count for courseware study	Time length	Time Length/count	Averaged score in examination	Records count
0	14.2	281.17	19.8	77.92	75
1	10.77	231.65	21.51	74.68	473
2	10.92	215.88	19.77	63.25	187
3	12.16	285.35	23.47	82.99	340
4	31.71	1150.39	36.28	78.38	52

In our learning process assessment system, the learners were required to study the courseware continuously for more than 15 minutes, otherwise, the study would be judged invalid, and they are also demanded that the count of courseware study must be more than 10 times for every course. From Table 2, we know that cluster 4 is an excellent in the whole behavior. The averaged time length on courseware study is more than 36 minutes, and is much higher than threshold value. The count of courseware study is also much higher than the threshold value. It's obvious that the goal of these students is knowledge acquisition. However, they are minority. In the other 4 cluster, the time length of courseware study is almost 20 minutes, and the count is about 10 times and just more than the threshold value. So the goal of most students is to pass the process assessment. In Table 4, a similar conclusion can be drawn.

In the relationship between the courseware study and examination, the more time on courseware study doesn't mean that better score in examination, but at least it means that their score are not bad, while less study time means worse achievement. Usually, this conclusion conforms to our general knowledge. However, it also reminds us that we should take into account the efficiency in the teaching design. In the design of assessment system, we would consider the time length/count as a new factor to improve the study time.

Count of Logging on BBS and Post

In a similar way, the count of logging on BBS, post and the score in examination were shown in Table 6.

Table 6: Count of Logging on BBS, Post and the Score in Examination Based on 1127 Records without Characteristic of Courses

Serial number of Cluster	Averaged Count for BBS post	Averaged Count for BBS logon	**post/logon**	Averaged Score in examination	Records Count
0	10.57	20.32	**0.52**	77.92	75
1	5.09	4.51	**1.13**	74.68	473
2	5.18	3.36	**1.54**	63.25	187
3	5.48	4.35	**1.26**	82.99	340
4	5.98	7.58	**0.79**	78.38	52

In the process assessment system, the learners are required to have 5 valid posts. From Table 6, we know that all the learners in the clusters have met the requirement. However, the learners in cluster 0 frequently log on BBS and have many posts. Normally, these learners are used to discuss in BBS, and good in study. From Table 3, we know that the averaged score are similar in cluster 4 and cluster 0, and the difference is that the count of the post of the learners in cluster 0 is more than that in cluster 4, while the time length of courseware study of the learners in cluster 4 is more than that in cluster 0. And the count of logging on the learning system in the two clusters are similar and the most among all the learners. So we could consider that these learners are hard working. So they could be consolidated to one cluster.

In cluster 1 and cluster 3, most learners just log on BBS and post once, and just meet the basic requirement. In cluster 2, some learners just log on once, but open several topics, however, their examination scores are the most inferior. In Table 4, there are examples that learners in cluster 3 only have averaged 1.92 times of logging on BBS, while the averaged count of post are 4.92. It is obvious that the learners try to open all the required topics once and just get the process score. We can conclude that most learners who spend little time in BBS are also inferior to examination.

So, the count of logging on BBS is an important criterion, and we should add this to the process assessment system and guide the learners to take an active part into the discussion in BBS.

Score of Examination and Homework

From Table 2, we know that the number of learners in clusters 1 and 3 are more than that in other clusters. Except score of examination and homework, the other behavior is similar and just averaged, so their behavior reflects the characteristic of most learners. And they met all the rigid requirements such as count and time length of courseware study, post in BBS, and win all the process score.

As for the score of examination and homework, we draw them form table 3 into Table 7.

Table 7: Score of Examination and Homework Based on 1127 Records without Characteristic of the Course

Serial number of Cluster	examination	Homework 1	Homework 2	Count of record
0	77.92	89.79	79.75	75
1	74.68	93.07	86.33	473
2	63.25	59.37	44.44	187
3	82.99	78.57	63.64	340
4	78.38	88.08	75.63	52

The score of cluster 0 and cluster 4 are similar and high, while the score in cluster 2 are low in all the three fields. So we could conclude that the score of homework is the reflection of examination. In the two clusters which possess most learners, their scores of homework and examination are similar. Though there is fluctuation, according to our experience, it is because some learners reviewed the courses more while they are not satisfied with their homework and vice versa.

We found that many learners would review the courses more when they are not satisfied with their score of homework. So in the teaching design, we should provide these learners more chances to review.

Conclusion

Nowadays, the distance education in China has entered a new phase, and we have attached much importance to quality assurance. However, the traditional assessment mode is still pervasive and doesn't match the evaluation to learners' knowledge management in distance education. Based on the teaching and learning system, the process assessment system acquires the data from courseware study to examination of the distance learners, and evaluates the study of learners through

the comprehensive computation. However, the choice of the parameters and the setting of the weight are based on the experience.

This passage proposes that we could adopt the cluster analysis from data mining methodology. Through the result of clustering, we concluded that the parameters of process assessment system should be improved further. For example, the count of logging on BBS and learning system should be added into the factors. In the teaching design, the majority should be guided to more courseware study and discussion in BBS, and tutors should provide more services in BBS and the question and answer system through audio and video.

References

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Boston, C. (2002). *The concept of formative assessment*. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.

Chen, N. L. (2006). Outline of theory of distance education teaching: Start point and construction of framework. *Research of Modern Distance Education*, *3,* 5–10.

Huang, R. (2006). Quality control of formative assessment in open education. *Journal of TV University (Philosophy & Social Sciences), 3,*122–124.

Jiawei, H., & Kamber, M. (2007). *Data mining*. China: Mechanism Publishing Company.

Ling, H., Wu, L., & Cai, Y. (2007). Survey of clustering algorithms in data mining. *Application Research of Computers, 24*(1)10–13.

Zhu, Y. (2006). Thoughts on formative assessment in college distance education. *Journal of Hubei Radio & Television University, 23*(5), 10–11.

Zhang, J., & Jiang, G. (2007).Analysis on the effectiveness and timeliness of web-based formative assessment. *Open Education Research, 13*(2), 105–108.