

Design and Research of the Evaluation System of College Students' Learning in the Big Data Environment

Zhou Ruiling

Department of Computer and Information Science, Hunan Institute of Technology, Hengyang 421002, China

56779568@qq.com

Keywords: Big data; Learning behavior; Learning quality; Teaching reform

Abstract: Big data is not only a technological revolution, but also a revolution of thinking and paradigm. The essence of teaching quality in college is college students' learning quality. In the era of big data, a new paradigm of teaching quality should be set up with learners as the center and college students' learning quality as the core. Based on the data of college students' learning behavior, a support system of learning quality, with the planning and motivation of learning goal, the monitoring and early warning learning quality, the analysis and assistance of learning quality as a whole, should be set up, then circular feedback teaching reform based on big data can be implemented, so as to provide "three-dimensional" support for college students, and provide lessons and reference for fine teaching management in the era of big data.

1. Introduction

Since the beginning of the new century, the information technology revolution has been surging fiercely, constantly affecting people's value system, behavior philosophy, lifestyle, consumption habits and interpersonal relationships. In the information age, where should the reform of university teaching go? How to integrate information technology into higher education, and provide service for teaching management, and for improving teaching quality in colleges and universities?

2. The evaluation of college students' learning affect in the big data environment

Big data is not only a technological revolution, but also a revolution of thinking and paradigm. In the era of large data, the idea of "speak with data", or modern management and decision making based on data statistics and analysis has become a new normal. University teaching management also needs to transform urgently from "extensive management relying on traditional experience" to "fine management based on analyzing and interrogating data". Now promoting the comprehensive reform of higher education, reflecting big data thinking from defining teaching idea, choosing teaching mode, the process of organizing teaching to evaluating teaching effect, so as to promote the reconstruction and optimization of teaching management methods in colleges and universities. Any educational or teaching goals can be achieved only if the teaching contents are accepted by learners and transformed into learners' knowledge, skills, or conscious behaviors and autonomous cognition. "The essence of teaching quality in colleges and universities is college students' learning quality, and improving teaching quality is fundamentally to improve college students' learning quality."

Since the new century began, in order to transform successfully the great country of higher education to the powerful country of higher education, China has established the road of "strengthening the connotation and improving the quality". Thus China has implemented the "quality project", "undergraduate teaching project" and the combination of top-down and bottom-up comprehensive education reform, and effectively improved the quality of university education as well as teaching conditions, but the influence of traditional teaching paradigm and one-way thinking has not been shaken off, students as a teaching subject are ignored, students "learning" has not been paid much attention and researches a lot, and "learning" quality has gradually become the

disadvantage of improving teaching quality. In the era of big data, the teaching quality construction should be deepened continuously, “learning” instead of “teaching” should be the focus. Concerning with the evaluation of teaching effect, students’ learning process, learning effect and the quality of learning output should be laid importance on, the teaching resources and hardware construction should be diluted, and the essence of higher education, that is cultivating competent students, should be emphasized again.

3. Data analysis and reading of the evaluation system of college students learning effect in the big data environment

3.1 MapReduce computing model

In Hadoop, each MapReduce task is initialized to a Job. Each Job can be divided into two stages: the map phase and the reduce phase. The two stages are represented by two functions: the map function and the reduce function. The Map function takes an input, that is the form of <<key value>, then generates <key value> output as well, Hadoop will be responsible for putting all values that have the same intermediate key value together, and then passing them to the reduce function, the reduce takes an input, that is the form of <key, (list of value) >, and then the value collection is processed, each: educe. Or 1 output, the output of reduce is also in the form of <key, value>.

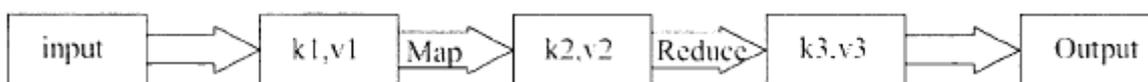


Figure 1 The basic pattern of data change in MapReduce program

3.2 Data-making of problems

First of all, we need to abstract the specific problem into a mathematical model. In this process, the most basic is to list analysis fields and analysis indicators related to the problem. This is the foundation of post analyzing and interrogating, the general analysis index and the scope of the analysis field can be delineated, and the work of data extraction can be carried out. The same time is involved among many problems, thus it is necessary to record the time of the video accurately. The completion of this task requires that the video player can accurately record the length of the playing time of the corresponding knowledge points. The watching video time is accurately measured by detecting the moving distance of the video slider. This topic designs a player that can generate a ID every time you watch and record the habits of students. Every time you move the slider a behavioral record will be generated. After watching the video, the time you spend in watching the corresponding knowledge points will be counted. On the MOOC platform, we should guide students to answer questions after class and record the correct rate of answering the questions of corresponding knowledge points. Here 1 represents the right answers, and 0, the wrong questions.

Table 1 The design of database structure

IID	The times of video fast forward	The starting time of fast forward	The ending time of fast forward	The times of video playback	The starting time of video playback	The ending time of video playback	Knowledge points	Knowledge points	Knowledge points
1	1	91	121	1	187	135	1	1	0
	2	163	201	2	164	122			
	3	127	304	3	346	272			
	4	240	285						

3.3 Model design

Before designing the algorithm model, the important basic work needs to do is to list the potential analysis fields and analysis indicators in combination with the actual case, which is the basis of analyzing and interrogating data, thus general analysis index and the scope of the analysis

field can be delineated generally, based on what the data can be extracted. The analysis indicators selected in this study include the coefficient of learning cost, the correct rate, the correct learning length and the wrong learning length.

Table 2 The analysis of learning points

	Knowledge points 1-1	Knowledge points 1-2	Knowledge points 1-3	Video 1	Knowledge points 2-1	Summary
coefficient of learning cost	1.32	1.51	1.57	1.70	1.44		1.77
Correct rate of problem setting	74%	65%	54%	64%	73%		56%
The average length of answering problem sets correctly	73	104	156	389	160		Null
The average length of answering problem sets mistakenly	45	57	97	320	187		Null

The coefficient of average learning cost is obtained by using the naive Bayes model. Here two concepts are defined first in order to express conveniently: average learning cost and the coefficient of average learning cost. Average learning cost refers to the average time students spend on learning one knowledge point (students can answer the questions of corresponding knowledge point correctly) The coefficient of average learning cost refers to the ratio between the average learning cost of the knowledge point and the time students spend on learning the knowledge point. The formula of Bias's theorem is as follows:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

X indicates the measurement and description of n number of attributes; H indicates certain assumption, for example, we assume that certain observation value X belongs to a particular category C; for classification problems, we hope to determine the P (H|X), which can be measured by the measurement and description of a given X description, then the probability of the establishment of H can be obtained, that is, the probability if the observation value belongs to the category of C can be calculated with the attribute value of X. Because P (H|X) is a posteriori probability, it is also called the posterior probability of H under the condition of X.

3.4 Excluding the records of invalid learning behavior

The main forms of invalid learning behavior include the following cases: students do something else in the learning process, leading to lengthening of learning time, and the correct rate is low; students have learned the knowledge, they directly fast forward to the end, and then answer the questions, the correct rate is very high, the learning time is very short; in order to obtain credits, students are not listening during the class and answer questions after class carelessly. These all belong to the abnormal value. When the management personnel count the learning time of students spend on answering questions correctly, a normal distribution function will be obtained, the false data will be deleted through confidence interval. Similarly, it is necessary to delete the false data when counting the learning time of students spend on answering questions mistakenly. The left and right confidence intervals of the two may be slightly different, they need to be analyzed in combination with the actual conditions.

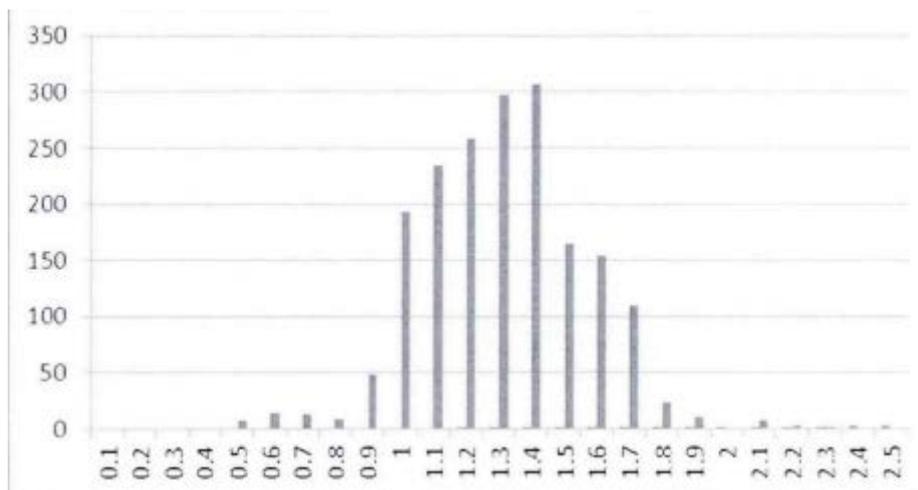


Figure 2 Summary of data analysis fields

The variance of the learning cost coefficient is calculated and the difficulty of courses is evaluated. When the expected value is obtained, the variance can be calculated. The meaning of variance in the algorithm is the average difficulty of the course. Although we have deleted the knowledge points outside the confidence interval, there is still a big difference in the degree of difficulty. The time students spend on learning some knowledge points may be two times or three times as long as the time length of knowledge point, and some may be 1.5 times as long. Comparing two courses, the expected value may be the same at last, but if the variance is large, which indicates that concerning with courses setting, some knowledge points are too simple, and some knowledge points are more difficult.

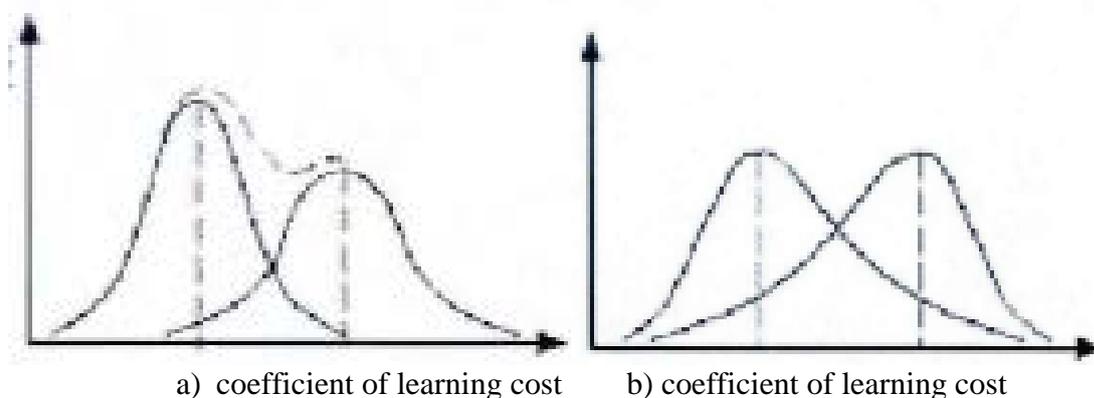


Figure 3 The expected variance of two courses

For a single course, if its coefficient of average learning cost is on the fitting curve, we think its teaching quality is high, and if it is under the fitting curve, we think its teaching quality needs to be improved.

3.5 Excluding the records of invalid learning behavior

The reasonable evaluation index is the foundation of establishing the evaluation system. According to the above evaluation system theory and design principles, combined with classroom teaching and network teaching, we establish a first level evaluation index from three dimensions of learning participation, collaborative communication and learning effect, many second level indexes are included in one first level index. The specific instructions of each index and the feedback of the system are shown in Table 3.

Table 3 The index table of multi-dimensional evaluation

First level index	Second level index	Feedback of the system	Instructions	weight
Learning participation U1	Learning times U11	Ranking/total number	Automatic record by the system	5%
	Class performance U12		Importing data after teachers' signing and counting	3%
	Class attendance U13		Importing data after teachers' signing and counting	7%
collaborative communication U2	Resources sharing U21		The number of uploaded and downloaded resources, which is recorded automatically by the system	3%
	Online discussion U22		The times of participating discussion and post, which is recorded automatically by the system	2%
Learning effect U3	The grade of Offline homework U31		Importing data after teachers' signing and counting	15%
	The grade of Offline homework U32	Average, ranking / total	Teachers' online grading and comment	16%
	Peer evaluation U33	Average, ranking / total	Online grading and comment on peers' works	4%
	Online test scores U34	Average, ranking / total	Automatic record by the system	5%
	Final exam scores U35		Importing data after teachers' signing and counting	40%

4. Analysis of the evaluation system of college students' learning effect in the big data environment

The core impetus of the development of big data comes from human's desire to measure, record and analyze the world. In colleges and universities, this motivation comes from the understanding of the students' learning, the monitoring and accurate evaluation of learning quality. In the era of big data, the data-based information of students' whole learning and living, data processing, interrogating and utilization provide support for teaching management and teaching evaluation. For example, we can judge students' devotion in learning, learning difficulties and learning habits by analyzing the data of students' online learning behavior. Through GPS system data in campus, we can know students location, whether they are in classroom, library, playground, venue or dormitory, so as to judge their time and energy they spend on learning, the second class, sports and other activities. Students' studying can be judged by the data of attendance, accuracy rate of homework, examination results, and so on. In conclusion, we can make all information of students' learning and living in campus data-based, then interrogate and sort the data and set up students' learning model, predict students' learning situation, and provide personalized guidance for students' learning teachers' teaching. In the era of big data, learner-centered method means laying more stress on training students' ability of knowledge acquisition and integration, and the ability of self-evaluation, and further developing the ability of lifelong learning. The exam-oriented education experience in primary and middle school in China deprives of college students' deficiency of self-management, self-restraint and self-learning ability after they enter university.

In the teaching evaluation, teachers are costumed to evaluate students' learning quality with test scores, they pay attention to assessing their learning outcomes, and ignore the learning process and developmental evaluation, which are reflected in teachers' unidirectional evaluation of students, while students' self-evaluation and mutual-assessment is ignored, which cannot comprehensively reflect students' learning. Based on the support of big data, we set up an evaluation system of students' learning quality consisting of three indicators: "learning process evaluation, learning effectiveness evaluation and practical innovation ability". Conclusive evaluation is transformed to developmental and procedural evaluation, and single exam is transformed to variable evaluation includes individual, others and examination. The influence of non-intellectual elements on learning quality, from learning cognition, attitude, emotion, volition, to methods, should be emphasized; students' learning quantity should be counted with credits, their learning effect should be evaluated

with grade points, and their learning quality should be assessed with grade point average; pay attention to fostering students' ability of practice and innovation.

References

- [1] Huang Xiaoyan. Discussion on the reform of statistical practice teaching under the big data environment [J]. China management informatization, 2015,18 (02): 248-249.
- [2] Chang Peiyan. The influence of big data environment on college students' value orientation and entrepreneurship education[J]. Electronic test, 2015 (14): 134-136.
- [3] Wang Yu, Ni Yanjun. Research on college students' mobile learning situation and factors of foreign language in the big data environment[J]. Chinese Journal (foreign language teaching), 2015 (10): 159-160.
- [4] Liu Shixin, Tan Xiyu, Liu Min. Research on college students' information attainments of English learning in the big data environment[J]. Contemporary educational practice and teaching research, 2017 (02): 241.
- [5] Li Wei. The optimizing the process of quality education of embedded information in the big data environment[J]. Intelligence information work, 2016 (03): 100-106.