

The Adaptive Teaching in the Setting of Big Data

Li He

Department of Information Science and Technology,
Tianjin University of Finance and Economics,
Tianjin 300222, China
renke21@vip.sina.com

Abstract—Education is one of important application fields of big data, and with the technology developing of big data, education data in scale is becoming larger and larger, and the breadth and depth of data are expanding. Through techniques of learning analytics and educational data mining, educators can get information of students' learning performances, such as learning attitudes, interests and effects, predict students' learning behaviors in future, find the possible problems hidden in the process of learning, and then try to provide the most suitable teaching contents and activities for the adaptive teaching. This paper discussed the influences of big data on education and teaching, expounded the applications of learning analytics and educational data mining in the adaptive teaching, and lastly presented a functional structure of a typical adaptive teaching system.

Keywords—big data; adaptive teaching system ; learning analytics; educational data mining

I. INTRODUCTION

Big data is a popular term used to describe the exponential growth and availability of data. The McKinsey Global Institute defines big data from the perspective of volume as datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyzing [1], and Wikipedia defines big data based on the McKinsey's definition as a collection of datasets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications [2]. Gartner defines big data as high volume, high velocity and high variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision-making [3]. In general, big data has the following three Vs characteristics:

1) *Volume*: Volume is about the number of big data mentions in the press and social media. There are many factors, such as transaction-based data stores, increasing amounts of sensor data, contribute to the increasing of data volume.

2) *Variety*: Data can be collected in all types of formats today, besides structured data and numeric data in traditional databases, there are more and more unstructured data in education area, such as text documents, video, audio, stock ticker data and financial transactions.

3) *Velocity*: Velocity is frequently equated to real-time analytics, which implies data stream in at unprecedented speed and must be dealt in time.

With the appearing and developing of online-teaching and massive open online courses, education is considered to

be one of the important application fields of big data. Compared with the data in other fields, education data is hierarchical, and it can be divided into seven levels according to its sources: keystroke level, answer level, session level, teacher level, student level, classroom level and school level. Education data in different levels reflect the students' learning behaviors, preference and effects from different aspects. Analyzing education data in different level we can obtain decision-making rules for improving teaching.

II. IMPACTS OF BIG DATA ON THE ADAPTIVE TEACHING

Under the support of big data, it becomes possible to discover some rules or regular patterns related to students' learning. In the process of teaching, we can get various data, such as the number and rate of correct answers, academic performance of students and teacher-student interaction frequency etc. With these data, educational institutions and teachers can provide individual educational resources and teaching activities to support students' successful learning according to the prediction models based on data. Consequently, in the age of big data, educational data is an important basis for adaptive teaching decision-making.

A. Promoting of the Teaching Decision-making Based on Data

Teaching decision-making directly determines the arrangement of teaching contents and teaching activities, it has a remarkable influence on the allocation of education resources and teaching qualities. In the era of big data, various data collected in the teaching process reflects the students' learning performances more realistically and comprehensively, by analyzing of these data, we can find students' learning preferences and teaching activities in different knowledge units; mine associations between knowledge units based on students' learning behaviors and, the similar characteristics of different students expressed in the teaching process; obtain the usage of different types of teaching resources and the effects of teaching activities. All these above can be used to support the decision-making for the adaptive teaching. Therefore, compared with the traditional decision-making based on experiences and subjective inferences, decision-making based on big data analytics is more objective, scientific, and it can reflect better the results of teaching activities.

Data analytics has been applied to public education in USA, and become an important driven force of teaching reformation. In 2012, the education Ministry of America federal government invested 2 billions of dollars in the big data plan of the public education, and in the field of USA

education data, in addition to leading IBM, many young companies engaged in analysis of education data, such as Civitas Learning, Dream Box Learning and Knewton, have successfully developed and released their versions of the intelligent learning systems with analysis and prediction of big data.

B. Driving the Learning Analysis on the Basis of Multivariate Data

The traditional evaluations of learning analysis are mostly implemented on the structured data stored in relational database, their analyzing objects are mostly examination results and usual grades. Under the setting of big data, it is possible to collect different types of data in the teaching process with relevant software and data collecting equipments, these data may be structured or unstructured, and unstructured data can reflect the students' learning behaviors and attitudes in finer granularity. Analyzing of these data, the information about students' answering background, steps, using frequency and total time in each knowledge unit can be obtained; educators can understand students' learning states better and find abnormal actions in time, then they can adjust the teaching arrangements according to students' learning performances and, provide each student with more adaptive teaching contents and activities.

III. EDUCATION DATA ANALYSIS TECHNOLOGIES

Learning analytics and educational data mining are two typical data analysis techniques in intelligent teaching and practices, these two kinds of techniques usually overlap in many research fields. In the adaptive teaching system, learning analytics focuses on the applications of known models, and pays more attention to human judgments and the individualized teaching; educational data mining is the application of data mining in the field of education, it pays more attention to mine the implicit, previously unknown and potentially valuable knowledge from educational datasets, it is a step of knowledge discovery, and greater focuses on automated teaching adaption[5].

A. Learning Analytics

According to the 1st International Conference on Learning Analytics and Knowledge, "learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs [4]." Compared with academic analytics, learning analytics emphasizes on the learning process, its object is to enable educators and schools to tailor educational resources and actions to the needs of each student. So, learning analytics of the adaptive teaching focuses on not only the students' performance, but also the assessments of teaching resources, educators and institutions. For students, learning analytics is usually used to analyze students' learning contents, trails, activities and social

interactions so as to get all kinds of performance and features of students in the learning process; for teaching resources, in order to provide adaptive teaching resources and contents, learning analytics can be used to analyze the using frequency and effects of courses, teaching activities and other teaching resources, find the semantic relations between courses and knowledge; for educators, learning analytics is customarily used to track the situation of teaching resource provision and using, count the interactions between students and teachers; for institutions, learning analytics can be used to evaluate students' learning results, teachers' teaching effects, and the availability of teaching settings.

Many learning analysis methods, such as social network analysis, sentiment analysis, influence analytics, discourse analysis, learning prediction, sense-making models, can be used in learning analytics to discover the relationships between the teaching contents and activities; track students' learning behaviors; help institutions improve administrative decision-making and organizational resource allocation, increase organizational productivity and effectiveness; help students insight into their own learning habits and get personalized recommendations.

B. Educational Data Mining

Educational Data Mining (EDM) has emerged as a research area in recent years all over the world from different and related research areas such as: online teaching, e-learning and the adaptive teaching system [5]. EDM is an application of data mining, it can be defined as the process involved in extracting interesting, interpretable, useful and novel information from educational data, its goals is to discover some valuable information which has an important influence on educational research and teaching. Nowadays, a variety of teaching and learning software are used in educational areas and, they produce a lot of data related to students' learning processes, these data is available for EDM.

The tasks of EDM are decided by the goals of teaching management. EDM involves different groups of users, and different groups look at educational information from their own requirements and goals. In the adaptive teaching system, teachers and students are the two most important users, and they have different requirements for EDM. For teachers, their fundamental goals are to improve teaching. So, they should be able to understand the students' learning attitudes, interests and the results of teaching activities, obtain students' feedbacks about teaching effectiveness in time, develop and provide more suitable contents and activities according to students' individual requirements; for students, individualized teaching guiding and recommendations of teaching resources and activities are most important, they always expect to get the most valuable teaching resources and the most suitable teaching activities. In general, different target users have different task requirements for EDM in the adaptive teaching. This paper summarizes the main tasks of EDM as the following five points according to the requirements of educators and students in the adaptive teaching system:

1) *Grouping of students* : Because the differences of students' learning performances, in order to maximize the teaching effects, it is necessary to cluster students according to their learning performances, and students in the same cluster have similar learning behaviors, interests and preferences, each student group has different learning preferences and behavior characteristics.

2) *Analyzing of knowledge associations* : There are logical relationships between different knowledge and concepts, these relationships will influence the learning sequences of knowledge, affect the teaching results of teaching resources and activities. From this aspect, it is important to analyzing models of knowledge associations for the adaptive teaching system.

3) *Tracking students' learning behaviors* : In order to understand students' learning states, it is necessary to track students' learning behaviors and, collect different types of data in the learning process, only with these data, the adaptive teaching system can understand students' learning performances and, find problems hidden in students' learning behaviors timely. Data with students' learning behaviors in the learning process usually includes: what knowledge and skills have been learning? What the learning sequence of these knowledge and skills? How many times each group of students has participated in a special discussion? How much time each group of students has spent in knowledge and skills and what are the learning results? Are there any students' abnormal behaviors in the learning process?

4) *Evaluating students' learning performance* : Learning performance of a student is a comprehensive evaluation factor, it mainly consists of learning attitudes, interests, effectiveness and results. With the evaluation models of learning performance, we can obtain all the performance information of student groups represented in the process of teaching, and these information is the scientific basis for the decision-making in the adaptive teaching system.

5) *Personalized recommendations* : Individualized teaching is a important feature of the adaptive teaching system, and personalized recommendations of courses, learning resources and activities are important parts of the individualized teaching.

Many data mining methods can be used in EDM, such as classification, clustering, relationship mining, discovery with models, visualization etc. Classification is a procedure in which individual items are placed into groups based on their quantitative information regarding one or more characteristics inherent in the items and based on a training set of previously labeled items. In the adaptive teaching system, classification is usually used to group student according to their learning performances, and to group teaching resources according to their characters and using conditions. The commonly used classification methods used in EDM are decision tree and Bayesian models. Clustering is a process of partitioning a set of objects, such as students or teaching

resources, into subsets, each subset is a cluster, and the objects in the same cluster are similar. Clustering methods can be used to build recommendation models for students with similar learning performances, and for clustering teaching resources with semantic and behavioral relationships. Relationship mining can be used to discover the interesting associations and correlations between teaching resources, and the associations of student learning behaviors in the adaptive teaching system. With these association rules produced in relationship mining, personalized learning contents, activities and discussions can be recommended to students. For teachers, Relationship mining can help finding students' shortcuts and anomalous behaviors, and provide the best and targeted teaching guidance according to students learning behaviors and interests. Prediction is one of the most commonly used discovery model which objective is to estimate the unknown value of a variable that describes students [5]. In the adaptive teaching system, prediction is usually used to predict students' mastering degree of knowledge and the learning behaviors and actions in the future. Some commonly prediction methods, such as regression analysis, association rules mining and sequential pattern mining, can be used in EDM.

IV. THE FUNCTIONAL STRUCTURE OF AN TYPICAL ADAPTIVE TEACHING SYSTEM

A typical adaptive teaching system should be able to track students' performances; find potential learning risks of students; adjust the corresponding teaching arrangements; provide the personalized teaching and guidance for students' successful learning. So, besides the traditional functions of student information management and teaching resources management, some basic functional components, such as learning analysis and prediction, teaching intervention and adaptive control, should be contained in a typical adaptive teaching system. The relationships and data flows between these function components of the adaptive teaching system are described in Fig. 1, it consists of seven main components as follows:

- Learning contents and activities module interacts with students to deliver personalized contents and activities for learning, collects data related to students' learning performances timely, and then stores them into the database of learning data for learning analytics.
- Teaching resource management system manages all of online courses, knowledge units, semantic relations between knowledge units and other teaching materials for supporting the content-based recommendations.
- The database of student information and learning data stores the basic information of students, such as learning level, interests, and time-stamped information about learning behaviors captured in learning process etc.
- Learning analysis and assessment system evaluates students' learning behaviors, interests and

preferences, and delivers these evaluation results to prediction system.

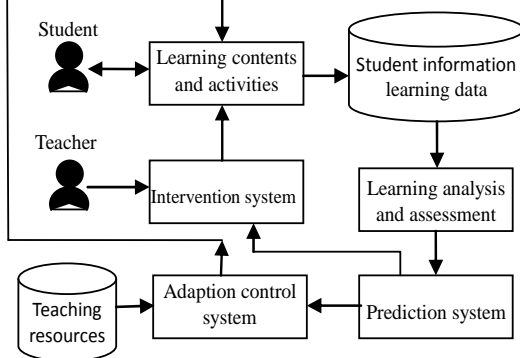


Figure 1 Structure of an individual teaching system

- Prediction system predicts students' learning behaviors and preferences in the future, and delivers the prediction results to the adaption control system and the intervention system.
- Adaption control system provides the optimal learning contents and activities to students according to the semantic relationships between knowledge and the outputs of the prediction system.
- Intervention control system allows teachers to intervene the automated teaching process according to students' current performances and their anomalous behaviors.

V. CONCLUSIONS

The rapid development of information techniques provides a broad space for the adaptive teaching. By tracking students' learning performances, the adaptive teaching system can not only predict students' learning interests and activity preferences in the future, but also discover students' learning risks timely. In order to improve the teaching quality, the adaptive teaching system should be able to tailor learning environment, provide individual contents and activities to students by the adaptive control system and the intervention system. Decision-making based on learning analytics is an important driven power for the future reformation of education and teaching, the application of EDM and learning analytics can optimize the allocation of teaching resources, improve teaching quality, and realize decision-making based on data. This paper introduces the influences of big data on the adaptive teaching, illustrates the application of EDM and learning analytics in the process of the adaptive teaching, and provides a functional structure of a typical adaptive teaching system. Next, we will focus on the reformations of teaching methods in the age of big data.

REFERENCE

- [1] Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Manyika, A. H. B. J., "Big Data: The Next Frontier for Innovation, Competition, and Productivity", McKinsey Global Institute, 2011.
- [2] Big data[EB/OL]. <http://en.wikipedia.org/wiki/>.
- [3] Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s[EB/OL]. <http://www.forbes.com/sites/gartnergroup/2013/03/27/>.

- [4] Siemens, G., & d Baker, R. S., "Learning analytics and educational data mining: towards communication and collaboration", In Proceedings of the 2nd international conference on learning analytics and knowledge, ACM, pp. 252-254, April 2012.
- [5] Romero, C., & Ventura, S, "Educational data mining: a review of the state of the art." Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40(6), 2010, pp.601-618.
- [6] Castro, F., Vellido, A., Nebot, À., & Mugica, F., "Applying data mining techniques to e-learning problems." Evolution of teaching and learning paradigms in intelligent environment. Springer Berlin Heidelberg, 2007, pp. 183-221.