

Research on Analysis Model of E-learning in Environment of Big Data

Fucheng Wan^{1,a}, Yaru Cao^{2,b,*}

¹Gansu Provincial Key Laboratory for intelligent processing of national languages, Northwest Minzu University, Lanzhou city, Gansu province, China

²Northwest Minzu University, Lanzhou city, Gansu province, China

^aWanfucheng@126.com, ^b1195825322@qq.com,

*Fucheng Wan

Keywords: Learning analysis, E-learning, Big data.

Abstract. The study gave general framework of e-learning analysis model in big data environment based on analyzing relevant factor and driving force of e-learning. Secondly, according to general framework of analysis model to construct e-learning in big data environment from left to right, from up to down, it constructed e-learning data model from multi-dimensional, multi-aspect and determined origin, method and process of data collection. Finally, it made brief analysis to the model and made feedback and guidance to learners who learning online. So that to reach purpose of promoting effective study.

1. Introduction

With development information technology, e-learning platform welcomed by learners more and more, its co-shared resources and personalized mode is benefit to learner's mind and knowledge construction. Currently, MOOC users are giant but efficiency is low, because amount of fulfillment of many courses is not too many, there are many situations of learners stopped in half study, this serious situation of user loss is not benefit to operation and development of platform. Therefore, it needs to track and analyze learners' learning behavior track in order to solve why learners stop at half study, find out which study behavior of users has large connection to successful fulfillment. Meanwhile, recommend personalized course resources based on users' learning behavior, improve his e-learning satisfaction. Based on analysis result, it can make feedback and guidance to learner's e-learning for reaching purpose of promoting effective learning.

2. Development status of e-learning analysis model

As to aspect of learning analysis technology, overseas study and development is more earlier than domestic and more improved, but as to study of e-learning in big data environment, it has got a degree of achievement in theory and practice. But there are some deficits, like lacking orientation to data acquirement of online learning of users, most of data is collected from database generally, hence it is an important fundamental job about how to extract valuable data item scientifically. While there is limitation in analysis method of e-learning analysis model, domestic experts proposed to adopt technologies of statistics and data evacuation to study online learning deeply, but its practice and application is not to depth, it hasn't find good point of penetration in big data according to practical situation. Hence, the article commits to establish analysis model of e-learning in big data environment, refers to multiple theoretical subjects, with help of education data evacuation and big data technology to construct multi-dimensional and universal C analysis model, to compensate our deficit in study of informationalized education field.

3. Construction mindset of e-learning analysis model

In big data environment, regarding behavioral science and learning analysis theory as theoretical basis, get general framework of analysis model of e-learning. Firstly, it needs to construct online learning data model, collect and deal data according to data model, secondly, design broadside procedure of model, which include making centralized analysis to behavioral data, evacuate connection between online learning study and study effect, recommend personalized course resources to its members according to feature of various groups' learning behavior. Finally, directional procedure of design model include data selection and pre-treatment of behavioral analysis, behavioral analysis method selection and application. Study object of the article constructed e-learning analysis model is subject to learners who registered in e-learning platform, study object is subject to learners' various learning behaviors in platform, specific behavior operation embodies in data model of analysis model.

4. System framework of E-learning analysis

Firstly, determine purpose of e-learning analysis, then determine behavioral science, system theory and learning analysis theory as theoretical basis, construct data model of online learning and collect and treat data, then analyze and construct model for e-learning from left to right, from up to down. Abide by construction mindset and online learning process model of online learning analysis model to divide horizontal model of analysis model into cluster analysis, recommend analysis, connection analysis. According to resolution of resolution to divide directional procedure of online learning analysis model into data treatment, method selection, analysis process and result output. General framework of learning behavior analysis model in big data environment is as figure 1.

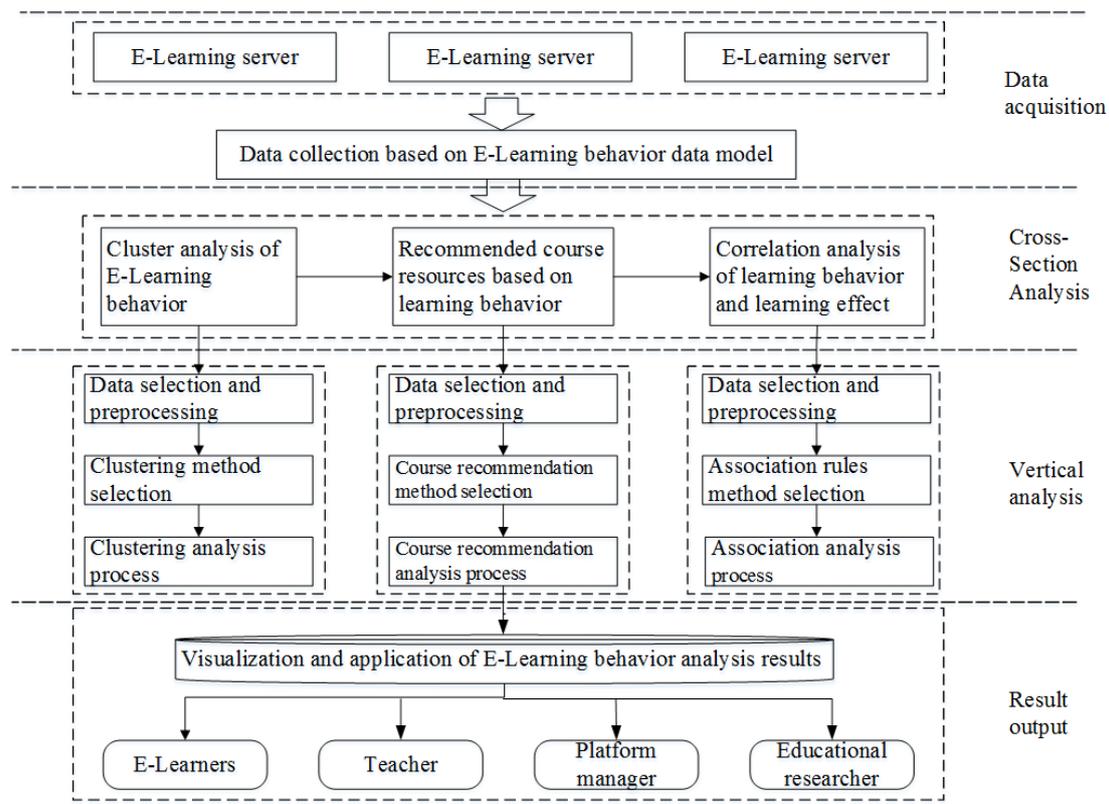


Fig. 1. General framework of online learning analysis model

4.1 Cluster analysis of e-learning behavior

Cluster analysis clusters data into different categories or clusters, it based on similarity, possibility of similarity in the same item is large, difference possibility among different items is large. The article classifies students who learning MOOC into several groups, then analyze and conclude student's

connection between learning effect and learning behavior. The writer uses nearest neighbor clustering algorithm to cluster, calculation formula of similarity of sample x_i and x_j is specific cluster process is as following:

$$d_{ij} = \|x_i - x_j\| = \left(\sum_{k=1}^n x_{ik} - x_{jk} \right)^{1/2} \quad (1)$$

① Based on selecting any sample as the first cluster of x_1 and cluster center of $C_1=x_1$, set similarity threshold as d_1 .

② Take a sample x_2 , calculate distance of d_{21} between x_2 and x_1 , if d_{21} is less than or equal to d_1 , x_2+C_1 cluster, or x_2 as new cluster center of C_2 , $C_2=x_2$, set similarity threshold as d_2 .

③ Deposit that there are k clusters from C_1 to C_k , continue to take sample x_i , and calculate similarity of cluster center of x_i and k 's cluster centers, if similarity of cluster center of x_i and cluster C_j is d_{ij} less than d_j , then x_i belongs to C_j .

④ Repeat to execute step ③, until all of data samples classified and get primary cluster set C .

⑤ As to every cluster $C_t(t=1,2,\dots,N)$, calculate distance of C_t from other clusters $d(C_t, C_j)=\|m_t-m_j\|$, among of it, m_t is average of C_t and t is not equal to j , and calculate the most nearest distance of C_t from i_t $d_t^*=\min \|C_t - C_j\|$.

⑥ Calculate average between various clusters

$$d^* = W^* 1/n \left(\sum_{t=1}^n d_t^* \right) \quad (2)$$

if d_t^* is less than d^* , then it can be seen as abnormal cluster. Then delete abnormal clusters set $C=\{C_1, C_2,\dots,C_N\}$ as cluster set of next phase.

4.2 Recommend course resources based on learning behavior

There are three kinds way of recommending technology: recommendation of based on content filter, recommendation based on synergy filter and recommend based on relevant rules. The article adopts recommending technology based on relevant rules, based on recommending technology of relevant rules is pre-make some rules, then analyze user's interest according to these rules and collected data, to weigh connection between study behavior and it. For instance, if a learner learned Java, when she clicks subject resources of "Java", the system will recommend learning resources relate to java concept, java thought, java practice. Hence these are basic knowledge of java, the next learning can be made after controlling basic knowledge, then it can recommend learning resources for her interests or demand.

4.3 Connection analysis of learning behavior and learning effect

Connection analysis means find frequent model, connection and relevance among project clusters in connection data, transaction data or other information carrier, easily speaking, connection analysis finds connection between different commodities or projects through transaction data. The article uses upgrade algorithm of Apriori algorithm FP-growth, in algorithm of FP-growth, firstly, scan matter database for twice to find out frequent cluster of everything, and stored in FP-tree according to order of decreasing support. If it continues to find frequent model, it doesn't need to scan matter database repeatedly, it needs to find in former established FP-tree, and utilize recurrence adjustment method of FP-growth to get frequent mode directly, hence, it don't need to generate candidate model in the whole checking process. The merit of this algorithm not only compensated deficit of Apriori algorithm, but improved similarity rate largely.

5. Conclusion

Through study, it finds that utilize massive users' behavioral data in online learning platform can realize all-round analysis to learners, it is very valuable and meaningful, including learner's learning method and interest for teachers and platform managers to make refined and professional education management and personallized course recommendation. Teacher utilize cluster analysis result of online learning behavior can control user's learning status, analyze difficulty and blind point, adjust structure and video content rationally, meanwhile, teacher can make proper adjustment from students' feedback, expand teaching mindset, change phenomenon of unchanged teaching content in traditional teaching, at present, performance of learning platform is complex, but practical effect has large difference, utilize data to analyze and provide guidance for system manager to improve and optimize learning platform performance. From revolution of teaching model, study learner's behavior, recommend personalized study case brought new breakthrough for experts of study revolutionary teaching model, use teaching interaction mode supported by computer to comply development tread of higher education in informationized era.

Due to limited time, the study has some issues need to be solved. Firstly, it needs to improve the model further, especially, refine analysis method and enrich technical toll, and consider collection of offline data and analyze issues. Secondly, it needs to develop special analysis tool or system to decrease mistake of manual statistics, decrease teaching personnel's burden, realize automatic and intelligent learning analysis. Finally, it needs to expand trail range, to choose more learners from professional and different grades to testify effectiveness of online analysis model for expanding its adoption range.

Acknowledgement

This research was financially supported by the Gansu science education programs of "thirteenth five-year plan" (GS[2016]GHB1218).

References

- [1] SARA IBN EI A, et al, Massive open online courses: A new dawn for higher education? *International Journal on Computer Science and Engineering*, vol. 5, pp. 323-232, 2013.
- [2] Strang K, How student behavior and reflective learning impact grade in online business courses, *Journal of Applied Research in Higher Education*, vol. 8, pp. 390-410, 2016.
- [3] Anthony G. Picciano, The Evolution of Big Data and Learning Analytics in American Higher Education, *Journal of Asynchronous Learning Networks*, vol. 16, pp. 9-20, 2012.
- [4] Sam Madden, .From databases to big data, *IEEE Internet Computing*, vol. 16, pp. 4-6, 2012.
- [5] Raymond Y K L, Chunping L, and Stephen S Y L. Social Analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis, *Decision Support Systems*, vol. 65, pp. 80-94, 2014.
- [6] Zhiting Zhu, and Demei Shen, New paradigm of educational technology research based on big data, *E-education Research*, vol. 2013, pp. 5-13, 2013.
- [7] Ke Zhu, and Qingtang Liu, Development and application of learning platform based on "learning analysis" technology, *China Educational Technology*, vol. 2013, pp. 127-132, 2013.
- [8] Manli Li, and Zhenzhong Huang, Education demonstration of big data on MOOCs platform, *Science China*, vol. 60, pp. 570-580, 2015.
- [9] Zhuoxuan Jiang, Yan Zhang, and Xiaoming Li, Analysis and prediction of learning behavior based on MOOC data, *Journal of Computer Research and Development*, vol. 52, pp. 614-628, 2015.

- [10] Yang Zong, Hongtao Sun, Xiangguo Zhang, Qinhua Zheng, and Li Chen, Logistic regression analysis of learning behavior and learning effect of MOOCs, *Distance Education in China*, vol. 36, pp. 14-22, 2016.
- [11] Yiling Hu, Xiaoqing Gu, and Chun Zhao, E-learning behavior analysis modeling and mining, *Open Education Research*, vol. 20, pp. 102-110, 2014.
- [12] Shunping Wei, Characteristics of e-learning behavior and its influencing factors, *Open Education Research*, vol. 18, pp. 81-90, 2012.
- [13] Yang C, Hsien T C, Regional differences of online learning behavior patterns, *Electronic Library*, vol. 31, pp. 167-187, 2013.