Japanese Online and Offline Hybrid Teaching-Learning Analysis Model Built on Big Data Technology

Jingshu He(✉)

Foreign Languages School, Sichuan University Jinjiang College, Meishan 620860, Sichuan, China
249392550@qq.com

Abstract. Modern society has embraced the use of big data technology in all spheres of existence. Modern education technology has advanced and developed as a result of the big data era's emergence. The use of information technology for learning analytics has significantly changed teaching and learning in the age of big data in education, offering a fresh viewpoint on the development of new educational innovations. The use of this technology provides precise analysis of large-scale learners' behavior, cognition, psychology, and emotion. This paper intends to build a learning analysis model for blended university Japanese language teaching based on big data technology from data sources, data analysis, and mining, and application of data visualization techniques to help Japanese language education and students improve teaching and learning and improve teaching quality by analyzing learning behaviors. In contemporary times, the application of big data technology has penetrated into all walks of life. The advent of the era of big data has also brought new changes and promoted the development of Japanese language teaching and learning.

Keywords: big data technology · learning analysis model · Japanese language teaching · Blended Learning

1 Introduction

The United Nations published a white paper titled "Big Data for Development: Challenges and Opportunities" in 2012, stating that "the era of big data has come, and its emergence will have a profound impact on all aspects of society." Human society has entered the era of big data, thanks to the rapid advancement of artificial intelligence technology. Big data application research in numerous industries has recently become a popular topic. The marriage of big data and education, according to Viktor Mayer-Schonberger, will go beyond prior "innovations of little power" to truly upset the old education model, leading the way in teaching, learning, research, service, management, and education. The "Thirteenth Five-Year Plan for Education Informatization" was announced by the Ministry of Education in 2016, highlighting "the vital role of big data in education management and learning space." The Ministry of Education released the

V. Balakrishnan et al. (Eds.): ICMEIM 2022, AHSSEH 6, pp. 327–334, 2022.
https://doi.org/10.2991/978-94-6463-044-2_43

"Education Informatization 2.0 Action Plan" in 2018, stating that the plan aims to "improve the top-level design of education management informatization, comprehensively improve the ability to use big data to ensure education management, decision-making, and public services…… Deepen the application of big data in education." And the newest development in educational information technology is the use of big data technologies for learning analytics research [1]. Currently, there is no learning analytics system built for blended university Japanese language teaching, Therefore, as a Japanese language educator, it is in line with the times to use big data technology to build out a Japanese language learning analysis model to optimize teaching quality.

2 Definition and Characteristics of Big Data

In May 2011, McKinsey & Company published a report titled "Big Data: The Next Frontier for Innovation, Competition, and Productivity" [2], claiming that data has permeated every industry and business today and has become a significant factor in production. The mining and utilization of huge amounts of data are ushering a new wave of productivity growth and consumer surplus, according to the report. Data has pervaded every industry and business today and has become a significant factor in production, according to the paper. In this perspective, "big data" serves as a predictor of innovation, competition, and productivity. Since then, the explosive growth of big data has drew unprecedented global attention. Big Data is defined as "data with characteristics of large volume, diverse structure, and strong timeliness; new computer architecture and new technologies such as algorithms are required to handle Big Data; the application of Big Data emphasizes the application of new concepts to assist in decision-making and analysis" in the White Paper on Big Data (2014) published by the Institute of Telecommunication Research of the Ministry of Industry and Information Technology of China [3].IBM utilizes three terms to define big data characteristics in general: size, speed, and type. The 4vs, or data size (volume), data processing speed (velocity), data type (variety), and accuracy, are typically used to summarize the main properties of Big Data (veracity).

3 Definition and Characteristics of Blended Learning

The Sloan Consortium (Online Learning Consortium, OLC) was the first to propose the concept of blended instruction, defining it as a combination of face-to-face instruction and online instruction. At this stage, scholars regarded blended when-teaching as a transitional stage between pure face-to-face teaching and pure online teaching, a simple combination of both based on information technology. The White Paper on Educational Technology in the United States published in 2000 served as the first publication to discuss blended learning. Following 2007, the Sloan Consortium revised the idea and clarified the proportion of online to face-to-face education in blended learning, proposing that teaching content with online instruction accounting for 30%–79% of the content can be referred to as blended learning.

There is no uniform definition of blended learning. According to He Kekang, blended learning is a new type of learning that combines the advantages of traditional classroom teaching and digital learning [4]. According to Qin Nan, blended learning primarily

combines traditional classroom instruction with online learning, individualized learning with in-person cooperation, a variety of tools, etc. [5].

4 Relevant Elements of Learning Analytics

4.1 Definition and Characteristics

In a report produced in partnership with the Association for Information Technology in Higher Education, the new media consortium (NMC) first offered learning analytics (Educause). Learning analytics, according to the report, is the study and analysis of data concerning learner engagement, performance, and learning processes in order to make real-time modifications to curriculum, instruction, and assessment using loosely coupled data mobiles and analytic methodologies. Learning analytics was later characterized by other academics as "the use of intelligent data, learner-generated data, and analytic models to uncover hidden knowledge and social relationships, and to make predictions and suggestions regarding learning [6]".

The goal of learning analytics is to make learning more efficient. "The measurement, collecting, analysis, and reporting of data about learners and their circumstances to understand and optimize learning and the environment in which it occurs" is what learning analytics is all about. Some academics also note that learning analytics is the creation of tools and technologies to record and store data on learning processes and outcomes, identify patterns based on massive amounts of data, present this data in a generative and usable form, and combine the data with intelligent tools to personalize and optimize the learning environment [7].

As is clear, learning analytics uses information on students and learning processes. The goal is to improve learning and learning environments, record and analyze learning environments and learning trajectories using a variety of techniques, find learning patterns, forecast learning outcomes, offer learning strategies for students and teaching strategies for teachers, and encourage effective learning. Learning analytics that are individual differences-focused will significantly improve teaching and learning.

4.2 The Use of Data Processing Techniques in Learning Analytics

The majority of today's learning analytics technology is based on data mining approaches. In databases, it is seen as a sub-step of the knowledge discovery process. The process of learning analytics utilizes a variety of techniques. The data generated therein will undergo a data processing process. From the technical aspect, the following steps are included, including data collection, data storage, data processing, data analysis and mining, and data visualization and presentation.

As shown in Fig. 1: Data collection is the collection and integration of the required data from data sources. Data storage is the storage of data on reliable media (including cloud storage, database storage, file storage, and other methods listed in Fig. 1) that prevents data loss and provides reliable access performance. Data processing is the process of performing arithmetic and logical operations on data in order to obtain further data. The two processing architectures, batch processing and stream processing, are the

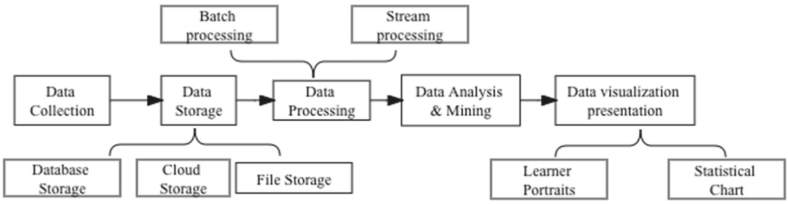

Fig. 1. Data processing flow

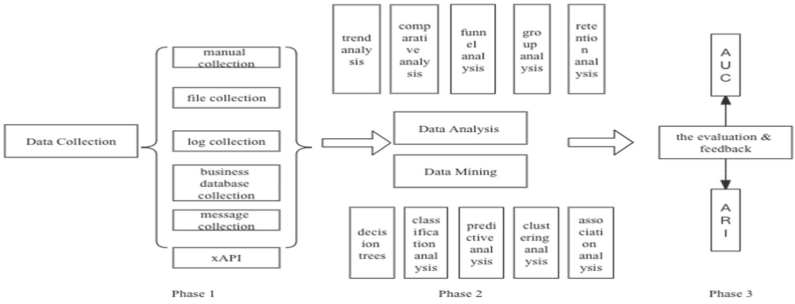

Fig. 2. The ideal learning analytics model

common technical methods used in data processing. The stored raw data is processed and calculated through data processing to generate the required data. Data analysis and mining is the process of analyzing and mining data through statistical analysis or its learning methods, extracting useful information from the data, and forming conclusions. Data visualization is the presentation of the results of data analysis in the form of learner portraits and various statistical charts (Including line graphs, pie charts, bar charts, funnel charts, radar charts, heat map dashboards, knowledge maps, etc.) to facilitate the understanding of the viewer.

In general, the architecture of the overall model for learning analytics should contain three phases: data acquisition, data analysis and mining, and evaluation and feedback. We can plan the ideal learning analytics model as Fig. 2.

Data collection is the starting point of the big data processing process, which is the process of collecting and aggregating data from various data sources to the big data storage and processing platform. The required collection techniques include manual collection, file collection, log collection, business database collection, message collection, and xAPI (A new learning technology specification released by ADL in 2013, the most central part of the TLA family of specifications. Used to track more detailed learning behaviors). Data capture can be done in terms of basic learner information, learner psychological and physiological information, and learner learning behavior.

Data analysis and mining are both processes that use appropriate algorithms to process data, extract useful information and form conclusions. Data analysis mainly uses statistical analysis and business understanding methods, while data mining is more often implemented using machine learning algorithms and computer programming. Data analysis often combines multiple analysis methods to provide a comprehensive analysis from

all angles. The main analysis methods include: trend analysis (the most basic and common data analysis method), comparative analysis, funnel analysis (can be more into the analysis of the causes of problems), group analysis (requires analysts to have extensive experience in analysis), retention analysis, etc. In data mining, its mainstream algorithms include: decision trees (methods mainly include classification and regression trees and cardinality automatic interaction detection), classification analysis (common classification algorithms include rule-based, decision trees, nearest neighbor, random forests, plain Bayes, artificial neural networks, support vector machines, etc.), predictive analysis, clustering analysis (common clustering algorithms include K-means, agglomerative hierarchical clustering, DBSCAN, etc.), association analysis (common association analysis algorithms include Apriori algorithm, FP growth algorithm), etc.

Through data analysis and data mining, it is possible to specifically analyze the overall situation of learners online and offline, learners and learning contents, and teaching staff and learners in three latitudes. In terms of technical architecture, ElasticSearch and Hadoop technologies are mainly used to store and manage the corresponding data.

The purpose of the evaluation and feedback is to make the learning analysis model's effect more precise. For example, using indicators such as the ARI (Adjusted Rand Index) to analyze the benefits and drawbacks of clustering algorithms, using indicators such as the AUC (area under the curve) to judge the benefits and drawbacks of classifiers or prediction models, and explaining the models so that teaching and learning staff can avoid learning risks and make appropriate teaching responses based on the data mining results. Furthermore, the aforesaid algorithms may be used to examine students' learning state and ability, as well as gather and anticipate their learning using their various data in the network. The following is a list of JEONG and BISWAS' primary research goals [8]: (1) creating student models (which contain learners' knowledge, motivation, metacognition, and attitude) to predict learners' future learning models; (2) exploring or improving domain models (which describe what has been learned and the ideal instructional sequence); (3) studying the effects of instructional support provided by various types of learning software; (4) increasing educators' scientific understanding of learning and learners by creating computational mode (including learner models, domain models, and software instructional models).

5 Japanese Language Learning Analysis Model Based on the Background of Big Data Technology

5.1 The Importance of Learning Analysis Techniques for Japanese Language Teaching

As a method of assessing data regarding learning behaviors, learning analytics offers a road to mastering the situations that students face during the learning process. Today, big data technologies are being applied in a variety of businesses. However, while learning analytics has been adopted by specific fields in education, there is no learning analytics model for Japanese language instruction and learners. As a result, the development of a Japanese language discipline necessitates the development of a big data-based Japanese language learning analytics model.

5.2 Features of the Learning Analytics Model for Blended Online and Offline Teaching of Japanese

The process of learning a Japanese language is fraught with ambiguities and emotional aspects, such as the personality of the teacher, the personality of the student, learning motivation, and belief processes, all of which can affect Japanese language learners. As a result, in the modeling phase of Japanese language learning analysis, the uncertainties must be taken into account. Furthermore, hybrid online and offline teaching is mostly a combination of traditional classroom and online learning. The merging of individualized learning and real-time writing, resource blending, and a variety of other factors [9].

(1) Learning in traditional offline classrooms is influenced more by the teaching environment, the teacher's subjectivity, and the students' beliefs, and these data, which are mirrored in the environment and scenario, are more represented as semi-structured data. These are significant influences on Japanese language teaching and learning, and it is critical to track and collect semi-structured data on them.

(2) The process of language learning is non-linear. The non-linear character of the language learning process is manifested in the autonomy of the language content and the autonomy of the learner. In different contexts, the interactions between the elements are intricate and complex and appear in completely different ways of organization. Therefore, this is reflected in the learning of Japanese, which means that in the process of learning Japanese, when the learner's psychology is far from equilibrium and in a state of disequilibrium, behavioral variation will occur. Due to the intertwined relationship between learning contents and learners' cognitive ability, an important phenomenon arises, that is, learners' learning styles do not linearly emerge one after another, but show non-linear and sometimes counter-directional characteristics. Therefore, in the process of building an analytical model of Chinese language learning, the nonlinear characteristics of Japanese language learning should be taken into account.

6 Construction of Learning Analytics Model

The semi-structured data and non-linear properties of the Japanese learning process are the subject of the learning analysis of Japanese blended education based on big data technology. It must also reflect scientific, holistic, openness, and operability concepts. As a result, the author analyzed these elements and developed a learning analysis model for blended Japanese language education that included both online and offline components. (As seen in Fig. 3).

Figure 3 shows a Japanese online and offline hybrid learning analysis model based on comprehensive data from both online and offline sources, as well as the non-linear properties of the Japanese language learning process and the underlying concepts for developing the model. Data collection in the traditional classroom is mainly entered manually by the teacher. This includes students' classroom performance, attendance, group discussions, answers to questions, assignments, etc. Teachers will enter these data into the system by themselves. As for the online data collection, the integration of data will rely on the online learning platform. For example, the number of tasks completed, the bibliography of videos that students watch to learn the course, the progress of chapter tests, the progress of chapter learning, the degree of learning of videos, the

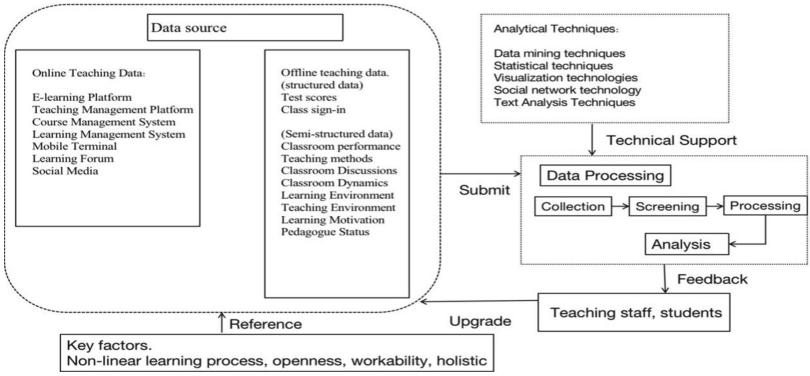

Fig. 3. Japanese online and offline hybrid teaching learning analysis model built on big data technology

number of discussions, etc. Online data collection will be more convenient for teachers compared to offline. Advanced network technology and computer-aided technology play a key role in the specific analysis process, and the data's final results are provided back to pedagogues and learners to increase teaching quality and learning status. Japanese learners' enthusiasm in learning can be sparked through scientific and individualized instructional analysis and assessment. It is hoped that the development of hybrid online and offline Japanese language instruction would be aided by this learning analytics paradigm.

7 Conclusion

In the subject of teaching and learning, learning analytics technology is being created and implemented. To improve learning efficiency and cultivate idealized learning individuals, a learning analytics model for hybrid online and offline Japanese language teaching based on big data technology is supported by computer science and network technology, as well as scientific and personalized learning analysis of teaching and learners. Learning analytics are expected to play a significant role in the future of Japanese language instruction.

References

1. Wu, Y.H., Chao, P, Ma, X.L. (2014) Developments and Challenges in Learning Analytics - A Review of the 4th International Conference on Learning Analytics and Knowledge. Education Research.12:72–80.
2. McKinsey & Company (2011) big data: The Next Frontier for Innovation, Competition, and Producitivity https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation
3. Institute of Telecommunication Research, Ministry of Industry and Information Technology, China (2014) Big Data White Paper http://www.cac.gov.cn/2014-06/18/c_1111184441.htm

4. He, K.K. (2004) The new development of educational technology theory from Blending Leaming. Electrochemical Education Research,3:1-6.
5. Ren, Y., Ma, Y.H. (2019) A Study on the "Dual-Teacher" Teaching Model of College English for Teachers and Students under the Concept of Blended Teaching. Heilongjiang Higher Education Research, 12:153-156
6. Siemens, G. (2019) what are learming analytics. https://blogs.adelaide.edu.au/myunitransform/2019/05/28/what-is-lcarning-analytics/.Larusson, J.A., White, B., (2020) Leaming analytics: from research to practice.
7. Lockyer, L., Heathcote, E., Dawson, S. (2013) Informing pedaogca action: aligning leaming analitics with learning design. American Behavioral Scientist,10: 1439-1459.
8. Jeong, H.Biswas, G.(2008)Mining student behavior or models in learning-by-teaching environments https://www.semanticscholarorg/paper/Mining-Student-Behavior-Models-in-Environments-Jeong-Biswas/b072bb4918da4b09a1495f72e5a3b98457a12512?p2df
9. Qing, N.Shandong Normal University.(2017) Research on the construction of blended teaching mode in the context of Internet+. https://ifgfy415d30141eea45d4hoxbpoouoqpup6k9fffh.res.gxlib.org.cn/kcms/detail/detail.aspx?recid=&FileName=1017240167.nh&DbName=CMFDLAST2018&DbCode=CMFD&uid=S1V0RG1qZlFIQzliQzlGOTVyVzlhaUVPOGl5dUgxSFlESzlQVWlaQis3TEVkaUpx