Information Dynamics and Applications
https://www.acadlore.com/journals/IDA

A Data-Driven Innovation Model of Big Data Digital Learning and Its Empirical Study

Jin He, Kuo-Yi Lin*, Ya Dai

College of Business, Guilin University of Electronic Technology, 541000 Guilin, China

* Correspondence: Kuo-Yi Lin (kylink1008@hotmail.com)

Received: 09-09-2022 **Revised:** 10-12-2022 **Accepted:** 10-28-2022

Citation: J. He, K.-Y. Lin, Y. Dai, "A data-driven innovation model of big data digital learning and its empirical study," *Inf. Dyn. Appl.*, vol. 1, no. 1, pp. 35-43, 2022. https://doi.org/10.56578/ida010105.

Abstract: Digital learning is the use of telecommunication technology to deliver information for education and training. As the increased acceleration of the propagation speed of the web, a lot of data collected by automated or semi-automated way. The 4s (Volume, Velocity, Variety and Veracity) of big data increase the challenge to extract useful value via systemic framework. This study aims to construct the data model of big data digital learning. Based on the digital learning data, data-driven innovation framework was proposed to identify data form and collect data. Bayesian network was proposed to capture learning model to extract user experience of students to enhance learning efficiency. Empirical study was conducted on a university to validate the proposed approach. The results have been implemented to support the strategies to improve student learning outcomes and competitiveness.

Keywords: Big data model; Digital learning; Education; Research; Data-driven Innovation; User experience

1. Introduction

The education science develops with consumer electronics, storage devices, network communication and education management. The data storage space continues to increase and the network transmission speed accelerates, so that educational data can be collected and stored through manual processing and computer operations in an automatic or semi-automatic manner [1], and educational achievements can be accumulated continuously. Among them, digital learning refers to education and training achieved through information technology [2]. In the process of digital learning, data with characteristics like the 4 Vs (Volume, Velocity, Variety and Veracity) [3] continue to accumulate. For example, through the application of mobile terminals to assist users in learning, as well as the behavior of users searching and clicking in the process of digital learning, data are continuously collected every second and accumulate into the big data, including but not limited to the change data of key words and sequences that users search for digital learning, the teaching materials for user learning and interaction (e.g., structured and unstructured data such as text, pictures, and data), the data of doubtful authenticity recorded by user accounts and social network accounts generated by public networks or computers. With the routine courses of different semesters, digital learning data also has a periodic nature.

With the continuous collection of educational data, the existing data are often scattered and stored in different databases, the data formats are inconsistent, and the data definitions of each unit are different, resulting in difficulties in data collection and integration. If the data can be organized in a systematic way, the rules and information hidden in the data can be mined out [4]. Otherwise, the data will only become a storage burden in the database. In the past, however, there was a lack of a systematic digital learning big data data-driven innovation model as the basis for data processing. Hence, no systematic process mechanism is available for implementation and improvement, making it difficult for various units to coordinate and communicate with each other. Therefore, a systematic digital learning data definition and analysis framework is needed to organize and analyze digital learning data, find out the rules hidden in it, and provide references for educational industry, education and research units, so that teachers and administrators to assist in the subsequent strategies developed to improve learning outcomes and student competitiveness.

The goal of this study is to develop a big data-driven innovation model for digital learning. The authors developed formats, gathered data on the digital learning process, and used Bayesian networks to analyze it in order to define digital learning data and develop a framework for data-driven innovation. The study's findings support improved pattern design, enhanced learning outcomes, and a better knowledge of user experience and student learning patterns. A university was chosen as the subject of the empirical study in order to plan and gather annual data on digital learning, examine students' learning behavior, and present improvement recommendations.

2. Literature Review

2.1 Research on Educational Materials

As fertility rate gets lower and lower, educational research is getting more and more attention. Relevant researches discuss the impact of entrance examination and assessment reforms, planning of integrated information systems, and methods like video analysis, mixed research and online reading. Today's science and technology have advanced to the point where education is no longer just confined to the classroom. For instance, the Canadian federal government and non-governmental organization FuturED jointly established consumer protection-oriented service quality certification and digital learning materials mechanism - OeQLS. The ECC (E-Learing Couracware Certification) proposed by ASTD in the United States is aimed at digital learning based on the Internet (Open eQuality Learning Standard). Due to this, digital learning has evolved into a new method of learning and growth that differs from traditional learning. Numerous digital learning channels have popped up like mushrooms after rain as a result of the quick development. Although users now have countless options, there is the absence of market and product quality control procedures as well as mechanisms to safeguard consumers. The topic that has to be explored the most right now is how to create reliable indications and attain unmistakable control. The findings of the study show that there are numerous challenges in teaching assessment, learning strategies, and the development of teaching techniques, teaching evaluation, and learning objectives for digital learning.

2.2 Big Data Analysis Model and Data-Driven Innovation

According to Lin et al. [5], data-driven innovation can help with innovation decision-making by guiding innovation paths from large-scale, multi-dimensional, unstructured, and dynamic data. Data-driven innovation, which has been utilized in operational management, corporate strategy, and customer relationship management to improve user experience, has recently been recognized as one of the key strategies for businesses to increase their competitive advantages [6]. Data mining technology has become a crucial and popular tool for the data-driven innovation approach, and it has been successfully used to analyze materials in many fields to help with applications like quality improvement, productivity improvement, and engineering design [7-11].

Data mining's primary uses include association rules, classification, grouping, and prediction, among others. In order to determine the correlation between users' usage patterns and decision-making, association rules investigate the relationships between variables in the data. To categorize data samples, a certain variable is used as the classification basis. Clustering The data is separated into multiple groups to examine how comparable the data is. To estimate or predict future property values or data patterns, one uses derived models or functions. In order to comprehend students' habits and wants and further apply the findings of data analysis to drive innovation, this study studies students' usage behaviors for the needs of digital learning. On a practical level, data-driven innovation techniques have been used [12], but few research have thoroughly examined the style and implementation of educational data-driven innovation. The data-driven stages are implemented step-by-step in accordance with the systematic model, and user demands are investigated.

3. Model Construction

Digital learning and big data -driven innovation model aims to glean insights on student usage of digital learning. Problem definition, data preparation, data-driven model analysis, result interpretation, and innovation model are the steps in the enterprise analysis process. According to Figure 1, each stage's scope is engaged in various issues.

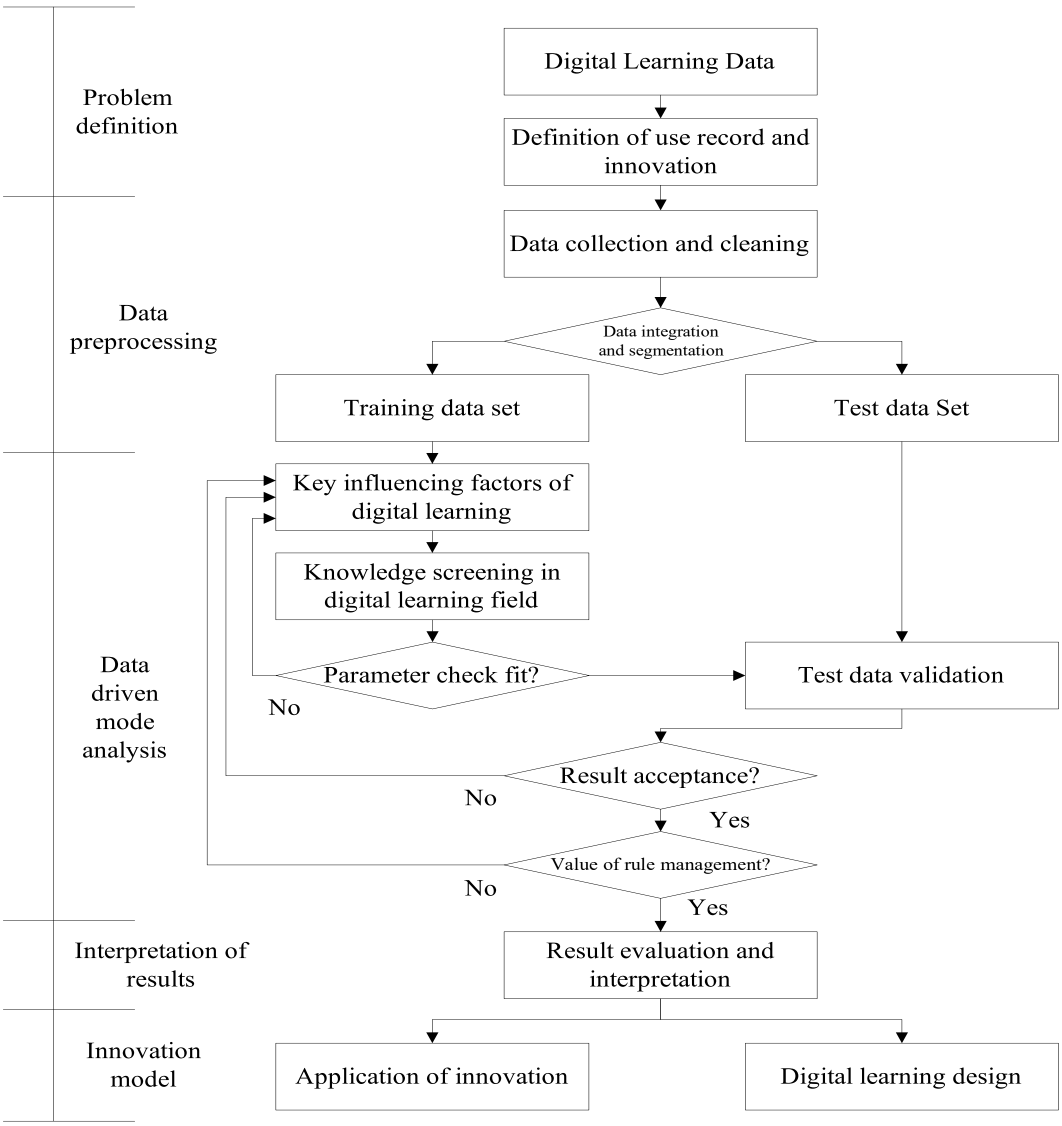

Figure 1. Definition of digital learning data and data-driven innovation framework

3.1 Problem Definition

On the network platform, digital learning data is a crucial medium for communication between students and teachers. Through the network platform, teachers give students the video resources they need to learn so they can download, search for, and debate pertinent information on course topics and gain useful skills and knowledge. However, there are an increasing number of esoteric course proper nouns as a result of the specialization and diversification of courses. The digital learning platform must be outfitted with modern technology to give students access to knowledge inquiry and access features that will enhance information extraction and decision-making. As a result, service specialization is a challenge for digital learning. In order to provide digital learning services that meet the needs of students, it is necessary to analyze the usage behavior required by students and courses.

3.2 Data Preprocessing

This study examines the content and presentation of the contents as well as how the indicators that students gathered during the discussion board's digital learning process are organized and developed. The binary record file (Centralized Binary Logging), which is used by students on the network and records the website URL, time, clicks (Clickthroughs), requests (Requests), and other details, is one of them. From this data, users' demands and behavior are deduced. Table 1 shows the data collation and description.

Table 1. Indicators of digitial learning

Indicator	Name	Description
view	Learn video materials	Learn from announcements, videos shared by teachers.
add discussion	Add discussion	Add discussion items in the discussion area.
add post	Add comment	Comment on the announcements and textbooks shared by teachers to express the students' questions and experiences.
launch	Initiate a notification	Initiate an announcement to notify all members
pre-view	Preview	Read course announcements, videos before teaching progress
recent	View recent announcements	View the latest announcement of the course
report	Upload report	Upload announcements, video topics related data and reports shared with teachers
review	Review	Read course announcements, videos after teaching progress
submit	Submit assignment	Submit the assignment requested by the teacher
update post	Update comment	Update a published comment
update submission	Update assignment	Update a submitted assignment
view assignment form	View assignment submission format	View the teacher's instructions and examples of the assignment submission format
view chapter	View chapter index	Consult the table of contents for course chapters
view discussion	View discussion	View course-related discussions
view forum	View discussion board	View discussion board
view submission	View submitted assignments	View the number of people who have uploaded the assignment
view submission grading table	View submission grading table	Check the assignment grades after the grades are announced
view summary	View summary	View summaries of teacher and classmate views on course content

(Data source: the authors)

The data must initially be preprocessed to make analysis easier, and it must be regularly revised and supplemented during the research process due to the complexity of the used data, which is constrained and hampered by subjective and objective criteria like database interface and system design. Following are the steps:

(1) Basic distribution and inspection of data: The basic distribution of data is statistically analyzed, and image processing is used to thoroughly examine and comprehend the features of the data in preparation for further processing.

(2) Data cleaning: For unusual values in the data, such as deviation values, blank values, missing values, etc., carry out processing that complies with the interpretation of the actual management meaning, such as directly deleting, keeping information, or supplementing in a unique way, according to the meaning of the specific values.

(3) Conversion of data type and dimension: The type and dimension of data have a significant influence on the results of data mining. For instance, continuous data must be discretized when performing Bayesian network analysis, and the dimension must be minimized when performing principal component analysis. Some data attributes, such as time, can be split into date and particular time according to various analytical objects to improve the dimension of the data and make analysis easier.

(4) Data segmentation and integration: There is inconsistency among the data sources. In order to analyze various objects and qualities, it is required to quickly integrate and segment the data. For instance, the training data should be divided for predictive analysis and the data from various time periods should be combined when examining the general trend.

3.3 Data-driven Model

In order to meet the demand, this study verifies the data screening rules, uses the form of image maps between variables, and applies statistical reasoning models to analyze the effects of multiple uncertain events. It also uses Bayesian network analysis as a tool for classification analysis and prediction. When new data and characteristics are provided, local fine-tuning can be done without having to rebuild the entire model, effectively saving resources and allowing for speedy adjustments.

The properties of user data are rather complicated, yet the data is based on just one behavior. Numerous attributes are included in the attributes, such as the date of usage, serial number, time, address, click, request, file name, data capacity, file pricing, authorization type, file number, and file classification. The following details are therefore noted for each data in order to analyze digital learning data in this study:

(1) Date of use: The user's login date corresponds to the date they last utilized the platform.

(2) Time: The user's login time corresponds to the amount of time that has passed between their login and logout on the platform.

(3) Address: The user's network address is reflected at the point of login, which can be used to pinpoint their

location.

(4) Clicks: The quantity of times a user clicks on a certain learning resource. The learning resource is more well-liked the higher the value.

(5) Requirements: The quantity of downloads a user has made of a particular learning resource. The learning resource has more reference needs the higher its value.

(6) File number: The file that the user clicks on or requests is identified by this number.

(7) Document classification: Group the documents that users click on and request under the headings "clicks" and "requests," with the attribution method depending on the subject and sub-field of the document.

These influencing factors interact to varied degrees, which can help with the understanding and illustration of the analysis results of the digital learning indicators. By using Bayesian networks to categorize, association rules between various attributes can be discovered and supported. Significant rules between user behaviors can be discovered in historical data using confidence level (confidence level) and gain (lift) calculations and verification, which can then be used to extract and organize these rules into service design schemes as decision-making recommendations.

There can be undetected relationships between various user behavior attributes. The usage behavior is simultaneously impacted by and susceptible to the ex ante probability of other factors. By utilizing the visual representation of the image, the Bayesian network may streamline and reduce the complicated sales analysis and decision-making process into a single event judgment and option. As a result, the amount of the conditional probability can be used to evaluate the target variable's decision-making relevance, making it simple to create the user service plan.

The following three steps make up the construction of a Bayesian network:

(1) Identify the target assumptions and associated assumptions, and then express them using a set of free variables.

(2) Define and build the Bayesian network structure diagram, connecting various influencing factors, using the mutual influence relationship between variables.

(3) Verify the relevant variables' conditional probability distribution.

The relationship between the target variables and the conditional probability distribution is depicted using network topology in the Bayesian network architecture diagram, which combines expert judgment and domain knowledge.

3.4 Results Interpretation

The usage rules can be obtained from the main factor mining phases of the whole data through the aforementioned modeling and testing procedure. The Chien et al. [13] technique was used as the foundation for the verification method of the study's rules.

3.5 Innovation Model

In order to obtain information that satisfies the research purpose, obtain the management significance of the analysis, and then present the practical contribution of the innovation model, it is necessary to continuously discuss and exchange opinions with experts, such as teachers, students, administrators, and system developers.

The main elements influencing how students use digital learning resources are revealed by the user insights attained through the data-driven strategy used in this study. The clicks and demands that users make on courses can be used to understand their preferences as well as to inform teachers' future course loads and hours for course learning frequency.

The advantages of digital learning must be outlined for students in order to entice them to learn and form learning habits. When combined with the advancement of general courses, it can be anticipated to produce better results, boosting students' learning motivation, performance, and competitiveness.

4. Empirical Analysis

This study uses a university as the empirical object to prepare and gather annual digital learning data in order to assess the efficacy of the big data data-driven innovation paradigm for digital learning. The index system was constructed in 2015, and data collection began in 2016. In order to examine the connection between digital learning behaviors, a total of 160,041 pieces of data, including the indications given in Table 1, were gathered. To outline innovation models, evaluate students' learning behavior and offer follow-up ideas for development.

4.1 Problem Definition

A certain Chinese province has enough resources and teaching energy for digital learning, and the collected teaching energy can help with the process's continual growth and progress. However, the province's higher education is struggling with the effects of globalization, internationalization, and the dwindling birthrate. Although resources are scarce, competition is fiercer. The goal and mission of increasing the effectiveness of educational resources is crucial. In order to provide guidance for future digital learning course design, this study investigates student use of digital learning, examines the level of student engagement in digital learning, and comprehends the determinants and key indications of high student participation.

4.2 Data Preprocessing

The data considers a single instance of a digital learning activity as a unit, and it has many different features, such as the date and time of usage, serial number, address, click, request, file name, file size, file price, authorization type, file number, and file classification. The following data preparation tasks were completed for this study's data:

(1) Remove unnecessary data: If there is a phenomenon of overreach, discuss the type of data; only one item, such as address and authorization type, is reserved for study.

(2) Change the data format to improve the features, for example, by adding the seasons and summer vacations to the date of usage.

(3) Combine and identify data, such as file names, that do not correspond to the analysis's goals.

(4) Fill in blanks, such as the absence of file classification in records of transaction history.

Following the preprocessing of the data, 106,693 records were selected from the 160,041 records that had been collected.

4.3 Data-Driven Model Analysis

In this study, prior probability is changed to posterior probability based on fresh information using Bayesian network analysis. In order to calculate the likelihood of another event, conditional probability is predicated on the occurrence of one event. The calculations are broken down into the following steps:

(1) Change the posterior probability: Continuously change the posterior probability in accordance with the sample data or evidence:

$$P(\tilde{\theta}=\theta_j \mid E)=\frac{P(\theta_j \cap E)}{P(E)}=\frac{P(E \mid \tilde{\theta}=\theta_j)\times P(\tilde{\theta}=\theta_j)}{\sum_{j=1}^{m} P(E \mid \tilde{\theta}=\theta_j)\times P(\tilde{\theta}=\theta_j)} \tag{1}$$

(2) Hypothesize conditional probability of occurrence: Considering the conditional probability of occurrence of a specific event E or evidence, in the case of event E occurring, the assumed conditional probability $P(\tilde{\theta}=\theta_j|E)$ of occurrence for θ_j can be calculated by:

$$P(\tilde{\theta}=\theta_j \mid E)=\frac{P(\theta_j \cap E)}{P(E)} \tag{2}$$

where, $P(E)$ is the probability of event E occurring; $P(\theta_j \cap E)$ is the probability of hypothesis θ_j and event E occurring at the same time. The joint probability between θ_j and event E can be expressed as:

$$P(\theta_j \cap E)=P(\theta_j \mid E)\times P(E)=P(E \mid \theta_j)\times P(\theta_j) \tag{3}$$

If $\theta_1, \theta_2, \ldots, \theta_m$ is a segment in sample space S, then event $E \subset S$, $P(\theta_j) \neq 0$, $j = 1, \cdots, m$. Then, the sum probability theorem can be expressed as:

$$P(E)=P(E \mid \theta_1)\cdot P(\theta_1)+P(E \mid \theta_2)\cdot P(\theta_2)+\cdots+P(E \mid \theta_m)\cdot P(\theta_m)=\sum_{j=1}^{m} P(E \mid \theta_j)\cdot P(\theta_j) \tag{4}$$

(3) Solve the posterior probability by:

$$P(\theta_j \mid E)=\frac{P(E \mid \theta_j)\cdot P(\theta_j)}{\sum_{j=1}^{m} P(E \mid \theta_j)\cdot P(\theta_j)} \tag{5}$$

To investigate the factors influencing students' click behaviors, Bayesian network analysis and user data were used. The results of students' click behaviors were uploaded reports, new discussions, new replies, initiating notifications, previewing, watching recent announcements, reviewing, submitting assignments, updating replies, updating assignments, viewing assignment submission formats, viewing chapter tables of contents, viewing discussions, viewing discussion areas, and viewing sub-discussions. Simply this behavior can forecast the frequency with which students would click, and it has been discovered that there are 90% possibilities of stimulating students' click behavior when there is only a high frequency of occurrence.

Then take into account the actions of viewing the summary, starting a new discussion, adding a reply, initiating a notification, viewing a preview, watching a recent announcement, uploading a report, reviewing it, submitting an assignment, updating a reply, updating an assignment, viewing an assignment submission format, and viewing the chapter directory. View discussions, discussion boards, assignments that have been submitted, and the results of those tasks. Grading tables may increase the likelihood of students clicking. The influence relationship can be depicted in the network architecture diagram using attributes and online graphics. The Bayesian network architecture design in Figure 2 should be arranged.

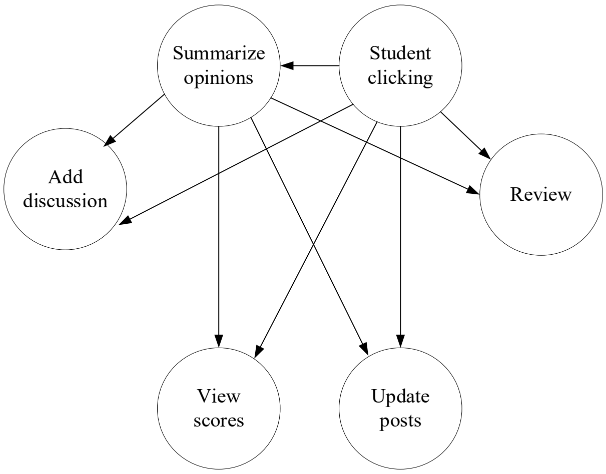

Figure 2. Structure of Bayesian network

The Bayesian network architecture makes it clear that the address, date of use, file classification, and click are all connected to the user's behavior when requesting files. By encouraging click behavior, the teaching unit can encourage students to use the online learning platform more frequently. Students who summarize ideas, add comments, check grades, update replies, and review also tend to click more frequently. To further assess the efficacy of the rules, this study sets a confidence level criterion of 80% and a gain threshold of 1.

4.4 Results Interpretation

Table 2. Summary list of key indicators of digital learning

Rule	Desription	Supporting documents	Rule confidene	Rule gain	Rule acceptance
1	If there is a high-frequency view summary, it will prompt students to click	59355	0.949	1.58	**Yes**
2	If there is a high-frequency view summary and a high-frequency add discussion, it will prompt students to click	55336	0.982	1.63	**Yes**
3	If there is a low-frequency view summary and a high-frequency add discussion, it will prompt students to click	4069	0.421	0.70	No
4	If there is a high-frequency view summary, and a low-frequency add discussion, and a high-frequency view submission grading table, it will prompt students to click	2369	0.688	1.14	**Yes**
5	If there is a low-frequency view summary, and a high-frequency add discussion, and a high-frequency update post, it will prompt students to click	1698	0.667	1.11	**Yes**
6	If there is a low-frequency view summary, a low-frequency add discussion, and a high-frequency review, it will prompt students to click	1197	0.280	0.467	No

Table 3. Summary list of rules

Rule	View summary	Add discussion	View assignment grade table	Update post	Student clicking decision
1	High-frequency				Click
2	High-frequency	High-frequency			Click
4	High-frequency	High-frequency	High-frequency		Click
5	High-frequency	High-frequency		High-frequency	Click

Therefore, this study further conducts data verification and confirmation on user click behavior and digital learning indicators. The results are shown in Table 2 and Table 3. High frequency indicates that the indicator exceeds 5 times a day.

Rules 1, 2, 4, and 5 are used in subsequent innovation models, which are organized as Table 3 according to the analysis results. Of the six rules given, rules 3 and 6 do not fulfill the criteria of this study for the confidence level threshold of 80% and the gain threshold of 1.

4.5 Innovation Model

Through their digital learning, students can sort a table by clicking on the rules, which show that the most important criteria are to view the summary, add comments, examine the assignment grade table, and edit the post. Among these, the occurrence of view summary will influence students' click decisions and the incentives for them to participate more. To further investigate the cause, consider that the teacher's announcement or the students' understanding of the material, which is frequently the course's core, is summarized in the view. Students have great incentives to refer to it and learn because it is simple to understand. Therefore, it is advised that courses using digital learning regularly conduct course view summaries and key discussions. In order to increase the incentives for students to participate in language learning, teachers should be encouraged to publish the assignment grade table and encourage students to participate in discussions based on the high-frequency view summary.

This study genuinely develops and gathers yearly data on digital learning, analyzes students' learning practices, and offers recommendations for improvement in the future. The empirical subjects have absorbed and put into practice the findings. The advantages of this study can be discussed in terms of its foundation for digital learning innovation, crucial elements, launch strategy, and anticipated performance. User evaluation is added to the digital learning innovation basis to complete and personalize system service innovation in addition to the evaluation and advocacy by the producers of digital learning platforms. Avoiding only the designer's imagination and estimation, starting directly from the perspective of the students and users at the school, understanding the user's needs, beginning from the user's needs and expectations, and minimizing unnecessary service and behavior waste are some of the important considerations.

5. Conclusions

The challenge and effects of the diminishing birthrate, internationalization, and globalization are being felt in the target province's higher education system. Although resources are scarce, competition is fiercer. It is a crucial goal and responsibility to figure out how to enhance educational outcomes with constrained educational resources. Building a data-driven innovation model for big data in digital learning is the aim of this study. It develops formats, gathers data on the digital learning process, analyzes it using Bayesian networks, and proposes a definition of digital learning data and a framework for data-driven innovation based on digital learning data. This study leads to better pattern design, user experience, and learning outcomes by assisting in the understanding of student learning patterns, and uses a university as an illustration to step by step explain and validate the framework it proposes.

Based on data from users of digital learning, this study employs a methodical approach to comprehend user habits and suggests acceptable service solutions to increase the rate at which learning is utilized. Through the Bayesian network analysis method, the relationship between the influencing factors and click behavior has been examined. The results show that the click behavior is clearly influenced by the view summary, add discussion, view assignment, grade grading table, update reply, etc. These factors also help with the subsequent improvement of the digital learning platform.

The discussion of user subjective satisfaction is absent from this study, which is restricted to the examination of data related to digital learning. In the future, text mining [14] and information systems [15] can be coupled with information and students' mental models to extract user experience and enhance learning results. It can additionally target to comprehend the connection between student usage behavior and school management.

Funding

This research was supported by National Natural Science Foundation of China (No. 62173253).

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Q. N. Sun, "Evaluation model of classroom teaching quality based on improved rvm algorithm and knowledge recommendation," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 2457-2467, 2021. https://doi.org/10.3233/JIFS-189240.

[2] P. C. Sun, R. J. Tsai, G. Finger, Y. Y. Chen, and D. Yeh, "What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction," *Comput. Educ.*, vol. 50, no. 4, pp. 1183-1202, 2008. https://doi.org/10.1016/j.compedu.2006.11.007.

[3] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inform. Manage.*, vol. 35, no. 2, pp. 137-144, 2015. https://doi.org/10.1016/j.ijinfomgt.2014.10.007.

[4] R. Shadiev, T. T. Wu, and Y. M. Huang, "Enhancing learning performance, attention, and meditation using a speech-to-text recognition application: evidence from multiple data sources," *Interact. Learn. Envir.*, vol. 25, no. 2, pp. 249-261, 2017. https://doi.org/10.1080/10494820.2016.1276079.

[5] K. Y. Lin, C. F. Chien, and R. Kerh, "UNISON framework of data-driven innovation for extracting user experience of product design of wearable devices," *Comput. Ind. Eng.*, vol. 99, pp. 487-502, 2016. https://doi.org/10.1016/j.cie.2016.05.023.

[6] K. Y. Lin, A. P. I. Yu, P. C. Chu, and C. F. Chien, "User-experience-based design of experiments for new product development of consumer electronics and an empirical study," *J. Ind. Prod. Eng.*, vol. 34, no. 7, pp. 504-519, 2017. https://doi.org/10.1080/21681015.2017.1363089.

[7] C. Chien, R. Kerh, K. Lin, and A. Yu, "Data-driven innovation to capture user-experience product design: an empirical study for notebook visual aesthetics design," *Comput. Ind. Eng.*, vol. 99, 162-173, 2016. https://doi.org/10.1016/j.cie.2016.07.006.

[8] C. F. Chien and K. Y. Lin, "Manufacturing intelligence for Hsinchu Science Park semiconductor sales prediction," *J. Chin. I. Ind. Eng.*, vol. 29, no. 2, pp. 98-110, 2012. https://doi.org/10.1080/10170669.2012.660200.

[9] D. Braha, "Methods and applications," In Data Mining for Design and Manufacturing, The Netherlands: Kluwer Academic Publishers, New York, NY, USA: Springer, vol. 3, 2001

[10] A. K. Choudhary, J. A. Harding, and M. K. Tiwari, "Data mining in manufacturing: A review based on the kind of knowledge," *J. Intell. Manuf.*, vol. 20, no. 5, pp. 501-521, 2009. https://doi.org/10.1007/s10845-008-0145-x.

[11] G. S. Linoff and M. J. A. Berry, *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*, Wiley Publishing, 2011.

[12] B. Agard and A. Kusiak, "Data-mining-based methodology for the design of product families," *Int. J. Prod. Res.*, vol. 42, no. 15, pp. 2955-2969, 2004. https://doi.org/10.1080/00207540410001691929.

[13] C. F. Chien, K. Y. Lin, and A. P. I. Yu, "User-experience of tablet operating system: An experimental investigation of Windows 8, iOS 6, and Android 4.2," *Comput. Ind. Eng.*, vol. 73, pp. 75-84, 2014. https://doi.org/10.1016/j.cie.2014.04.015.

[14] K. Y. Lin, "User experience-based product design for smart production to empower industry 4.0 in the glass recycling circular economy," *Comput. Ind. Eng.*, vol. 125, pp. 729-738, 2018. https://doi.org/10.1016/j.cie.2018.06.023.

[15] K. Y. Lin, "A text mining approach to capture user experience for new product development," Int. J. Ind. Eng. Theor. Appl. Pract., vol. 25, no. 1, pp. 108-121, 2018.