The Learning Behavior Analysis of Online Vocational Education Students and Learning Resource Recommendation Based on Big Data

https://doi.org/10.3991/ijet.v17i20.34521

Yufeng Jia[1], Qi Zhao[2](✉)
[1]Ministry of Student Work and Social Sciences, Shijiazhuang University of Applied Technology, Shijiazhuang, China
[2]Department of Preschool Education, Hebei Women's Vocational College, Shijiazhuang, China
zhaoqiedu@163.com

Abstract—The Learning Behavior Analysis (LBA) of students is a major function provided by most online education platforms. Based on the results of LBA, students can get customized learning path and recommended learning resources that suit their own situations, which can help them plan their individual training programs of Online Vocational Education (OVE). At first, this paper studied the resource recommendation algorithm based on the modelling of students' online learning behavior sequences, and solved the problem of conventional learning resource recommendation algorithms of ignoring the dynamic changes in students' learning preferences. Then, this paper presented the process of LBA applicable for OVE students, and gave a diagram of the flow of using big data to perform LBA. After that, this paper developed a bidirectional encoder based on self-attention model, and described the co-occurrence characteristics and dependencies of students' online learning behavior sequences. At last, the model pre-training and fine-tuning processes were introduced in detail, and the experimental results verified the effectiveness of the proposed analysis and recommendation methods.

Keywords—big data analysis, Online Vocational Education (OVE), Learning Behavior Analysis (LBA), learning resource recommendation

1 Introduction

Internet and big data technologies enable researchers and teachers to mine the data of students' learning behaviors and combine classroom teaching with the data mining results, thereby creating new education modes and adjusting teaching strategies constantly [1–9]. When participating in OVE, students can acquire learning resources suitable for themselves, create, edit, and share their learning experience based on the acquired professional knowledge and skills, communicate with teachers and other students to solve their own learning problems or help others solve their learning problems. During these processes, students exhibit a high degree of participation, interaction, and sharing [10–17]. The LBA of students is an important function of online education platforms. Based on the results of LBA, students can get customized learning path and recommended learning resources that suit their own situations, which can help them plan their individual training programs of OVE [18–24].

The analysis given by existing studies on the learning behaviors of different-type learning groups is generally not extensive or deep enough, therefore, Li et al. [25] explored a learning decision model based on the influence of group learning behavior, they took the advantage of Q-learning to improve the conventional behavior tree model and constructed a novel model which was then applied to the research on group learning behavior in the study; the authors combined the decision-making idea with the game model, and adopted a complex network structure to figure out the evolution law of the decision-making of multi-game group learning. Zhang and Li [26] studied the influence of smart learning environment on students' learning behavior in higher vocational colleges and analyzed the related influencing factors, they also compared the specific data of these influencing factors and discussed the significance of smart learning environment to education and teaching. Alqaheri and Panda [27] researched the automatic discovery process of Inductive Vision Miner and generated an effective educational process mining model based on the Directly Follows Vision Miner algorithm to figure out and predict students' learning behavior, and the results indicated that the developed process model can effectively help experts understand students' learning behavior patterns. Zhang et al. [28] introduced a new analysis scheme to analyze the tracking data and visualize students' autonomous learning strategies in a mastery-based online learning platform, the design of the platform reduced event types and the variability in student tracking data, the introduced analysis scheme adopted three-layer clustering analysis to cope with these challenges.

Conventional learning resource recommendation algorithms generally obtain students' learning preferences based on the existing learning behavior big data, and they cannot reflect the dynamic changes in the learning preferences of students. However, in real-world OVE scenarios, the students' static preferences and recent dynamic preferences together determine their choices of learning resources. In order to solve this shortcoming of conventional learning resource recommendation algorithms, this paper aims to study a learning resource recommendation algorithm based on the modelling of students' online learning behavior sequences. The second chapter presented the process of LBA applicable for OVE students, and gave a diagram of the flow of using big data to perform LBA; the third chapter developed a bidirectional encoder based on self-attention model and described the co-occurrence characteristics and dependencies of students' online learning behavior sequences. The fourth chapter elaborated on the model pre-training and fine-tuning processes. At last, the experimental results verified the effectiveness of the proposed analysis and recommendation methods.

2 The LBA of OVE students

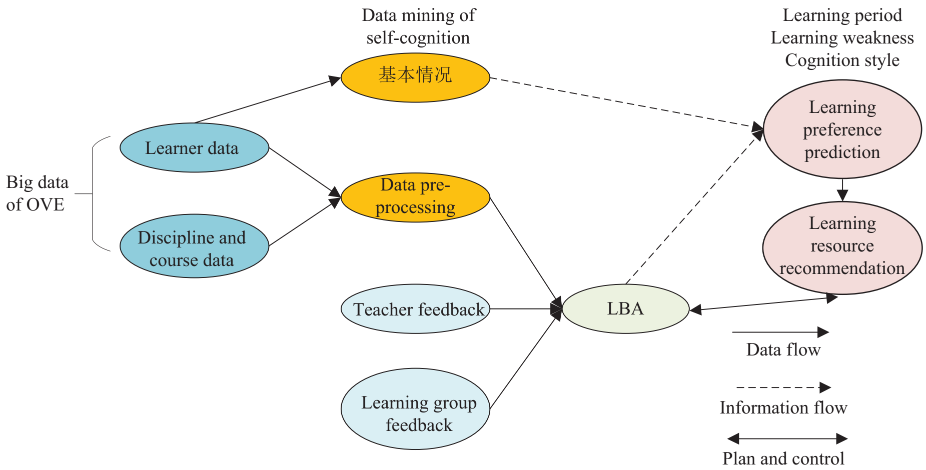


Fig. 1. The LBA process of OVE students

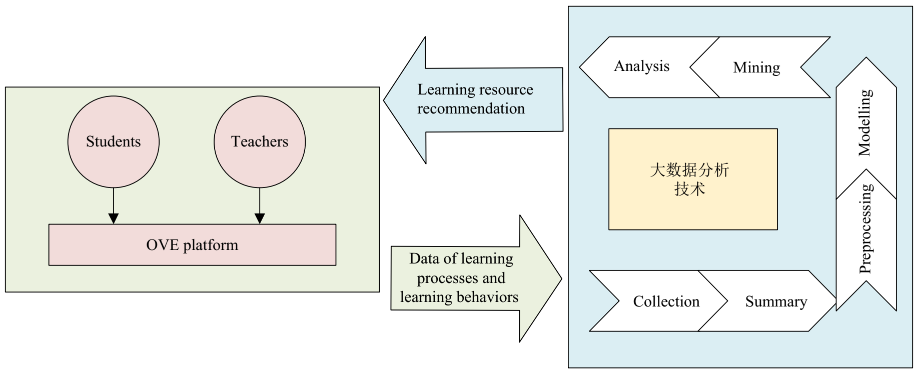


Fig. 2. Flow chart of LBA using big data

Based on the existing LBA cases of students, this paper integrated the mature LBA process to construct a novel LBA process applicable for OVE students, see Figure 1.

Figure 2 shows the flow of using big data to perform LBA. At first, the behavior data generated during the online teaching and teaching process of teachers and students are collected and summarized; then the learning behavior data of students are pre-processed, modelled, analyzed, and mined by the backstage system of the OVE platform; at last, the students' online learning behavior samples are converted into suggestions of learning resource recommendation that suit students' learning preferences. During this process, the backstage system of the OVE platform uses big data analysis to save teachers' time spent on the analysis and decision-making of students' learning resource recommendation, and ensures more objective and accurate recommendation results.

3 The LBA model of OVE students

When modelling the students' short-term behavior sequences of recent dynamic preferences, by default, the change in students' learning preferences is considered not much, but the students' online learning behavior sequences of two sequential time periods have a significant correlation. For such co-occurrence characteristics and dependencies of students' online learning behavior sequences, this paper adopted a bidirectional encoder developed based on self-attention model to model and describe them. In order to avoid large differences in the proportions of positive and negative samples of students' online learning behavior sequences, this paper employed the self-supervised method to train the constructed model.

The model consisted of three parts: input layer, sequence encoding layer, and output layer.

For a given online learning behavior sequence $r=[r_1, r_2,\ldots,r_L]$ with a length of K, this paper performed c-dimensional vector transformation on sample r_i in the i-th period in the online learning behavior sequence. Since the original online learning behavior features provided by the sample dataset of behavior sequence is a $f\times c'_t$ -dimensional feature matrix, the value of the matrix dimension c'_t is large and the sample length f in each period is not the same, so this paper converted the online learning behavior features of samples in each period into c'_t dimensional based on the average pooling process.

To facilitate description, assuming $t_{ri}\in R^{c't}$ represents the eigenvector of the online learning behavior of period sample r_i; $T\in R^{M\times c't}$ represents the feature matrix formed by the eigenvector of all period samples in the entire dataset of behavior sequence; M represents the number of period samples. Then, based on the principal component analysis, the redundancy between all features was eliminated, and the newly obtained feature matrix is denoted as $T_g\in R^{M\times c't}$. The eigenvector of online learning behavior $t^g_{ri}\in R^{ct}$ after the dimensionality reduction of period sample r_i can be indexed from T^g, assuming $t^g_{ri}\in R^{ct}$ represents the eigenvector of period sample, then the vector t^g_{ri} after dimensionality reduction can learn an embedding matrix S_g to realize the new feature space mapping to it:

$$g_{r_i} = S_g t^g_{r_i} \tag{1}$$

This paper also introduced the ID-type features into the feature extraction of online learning behavior sequences of students, that is, to train each period sample based on the embedded eigenvector to realize the representation of time period samples without relying on any online learning behavior feature. Although it doesn't applicable for cold-start cases, in actual experimental process, data augmentation could be adopted to avoid the cold-start problem of period samples.

Besides, since the relative position information in the online learning behavior sequences cannot be represented by the self-attention layer, this paper adopted the method with better performance in the experiment to describe it, that is, to attain the relative position information in the online learning behavior sequence by each period sample's learning of the embedded vector $z_i \in R^c$. At last, the following formula gives the calculation formula of the initial vector $f^0_{ri} \in R^c$ of period sample r_i:

$$f^0_{r_i} = g_{r_i} + z_i \tag{2}$$

Through above processing, the initial vector $F^0 = [f^0_{ri},...,f^0_{rK}]$ of the online learning behavior sequence could be attained, assuming $F^0 \in R^{K\times c}$.

In the model, the sequence encoding layer is formed by the stacking of multiple self-attention layers, each self-attention layer contains two parts: the multi-head self-attention model and the position feedforward network.

The multi-head self-attention model completes the computation based on the scaled dot-product attention mechanism. For a given query denoted by W and a pair of key-values denoted by L and U, the multi-head self-attention model attains the attention score by calculating the similarity between query W and key value L, then it normalizes the calculated attention score based on the Softmax function and applies it to another key value U. Assuming: the vectors are $W \in R^{mW\times c}$, $L \in R^{mL\times c}$, $B \in R^{mU\times c}$; the numbers of input online learning behavior sequence presentations are m_W, m_L, and m_U, and they satisfy $m_L = m_U$, then Formula 3 gives the specific calculation formula:

$$Attention(W,L,U) = softmax\left(\frac{WL^T}{\sqrt{c}}\right)U \tag{3}$$

The multi-head self-attention model can perform multiple parallel operations on the scaled dot-product attention mechanism. Assuming: f represents the count of parallel operations; $Q^W_i \in R^{cf\times c}$, $Q^L_i \in R^{cf\times c}$, $Q^U_i \in R^{cf\times c}$, and $Q^T \in R^{c\times c}$ represent the parameter matrices to be learned; then the formula below gives the linear mapping of the input online learning behavior sequence representations in different spaces:

$$\begin{aligned} MH(F_r^n) &= \left[head_1;\ldots;head_f\right]Q^T \\ head_i &= Attention(F_r^n Q_i^W, F_r^n Q_i^L, F_r^n Q_i^U) \end{aligned} \tag{4}$$

The parallel operations of the multi-head self-attention model performed on the scaled dot-product attention mechanism is mainly based on linear mapping. To enable the sequence coding layer to have the nonlinear expression ability for students' online learning behavior sequences, and to realize the interaction between online learning behavior sequences of different dimensions, this paper constructed a feedforward layer to represent each period sample in the online learning behavior sequence, the feedforward layer is composed of two fully-connected layers and a *GELD* activation function layer. Assuming that the parameters to be learned shared by each period sample are $Q_1 \in R^{c\times 4c}$, $Q_2 \in R^{4c\times c}$, $y_1 \in R^{4c}$, and $y_2 \in R^{c}$, then the following formula gives the calculation formula:

$$FFN(a) = GELU(aQ_1 + y_1)Q_2 + y_2 \tag{5}$$

Assuming: $\Phi(.)$ represents the cumulative distribution function of the Gaussian distribution, then the calculation formula of the *GELU* activation function could be expressed as:

$$GELU(a) = a\Phi(a) \tag{6}$$

Formula 7 gives the calculation formula of the position feedforward network of the *n*-th layer:

$$PFFN(F_r^n) = \left[FFN(f_{r_i}^n); \ldots FFN(f_{r_K}^n) \right] \tag{7}$$

As mentioned above, more complex dependencies of online learning behavior sequences can be attained by stacking multiple layers of self-attention, but the training difficulty will increase dramatically with the increase of the number of layers of the self-attention model. For the purpose of reducing the difficulty of model training, this paper introduced the residual connection module and the layer normalization module into the multi-layer self-attention model, and attained the hidden-layer representations of the period sample sequence one by one. Assuming: *KM* represents the layer normalization model, then the representation of the *n*-th layer in the stacked self-attention of *N* layers could be calculated by the following formula:

$$\begin{aligned} F_r^n &= Trm(F_r^{n-1}), \forall n \in \left[1, \ldots, N\right] \\ Trm(F_r^{n-1}) &= KM(X_r^{n-1} + FFN(X_r^{n-1})) \\ X_r^{n-1} &= KM(F_r^{n-1} + NF(F_r^{n-1})) \end{aligned} \tag{8}$$

4 Pre-training and fine-tuning of the model

The output layer performed classification probability prediction based on the attained masked hidden-layer representation of the period samples of students' online learning behavior, through the bidirectional encoder, the hidden-layer representation of the *M*-th layer of all period samples in the online learning behavior sequence *s* could be attained. In the initial stage of model training, the input layer of the model and the sequence encoder were trained based on the masked language model which came from the processing of natural languages. Under normal circumstances, if the online learning behavior sequence $r=[r_1,...,r_K]$ is given, and it's assumed that $[\Phi]$ represents the period sample r_i of the replacement position *i*, then, the online learning behavior sequence could be converted into $r'=[r_1,...,r_{i-1},[\Phi],r_{i+1},...,r_K]$.

Assuming: the feature of [Φ] is a randomly initialized vector, which is denoted as $g[\Phi]\in R^c$, then the feature g_{ri} of period sample r_i can be replaced by $g[\Phi]$. Finally, assuming f^N_{ri} represents the representation of the N-th hidden layer at position i; f^N_{ri} was input into the fully connected layer for classification to make judgements on the masked period sample r_i. Assuming $Q_3\in R^{c\times c}$ represents the matrix of parameters to be learned; y_3 and y_t represent bias vectors; T^g represents the feature matrix; then the following formula gives the calculation formula:

$$Z_{r_i} = softmax(GELU(f_{r_i}^N Q_3 + y_3)T^{g^O} + y_t) \tag{9}$$

Assuming $Z_{ri}\in R^M$ represents the preset probability distribution, this paper built the loss function of the model based on the negative log likelihood function; assuming r_Φ represents the set of masked period samples in the online learning behavior sequence r, then the loss function could be expressed as:

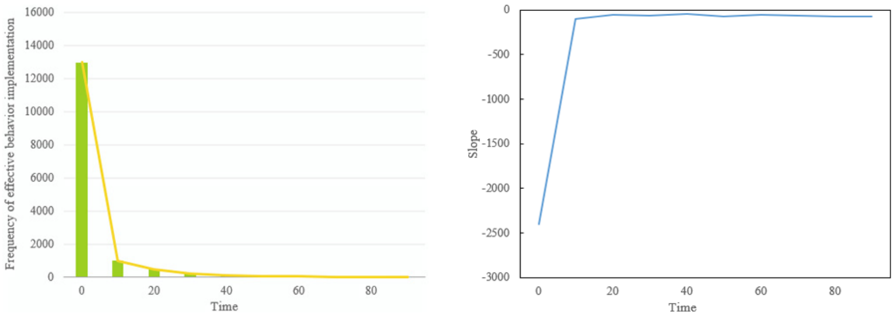
$$LOSS_Z = -\sum_{r_i\in r_{mas_l}} log(Z_{r_i}) \tag{10}$$

Through the process of model pre-training, it's known that the model training method adopted in this paper doesn't need to introduce the tags of the target period sample, it can complete the training of the input layer and the sequence encoding layer simply based on the online learning behavior sequence. This method can effectively deal with the imbalance that often occurs in the tags of target period sample; in addition, assuming each time the random masking is performed on l period samples in the online learning behavior sequence, if an online learning behavior sequence of length K is processed, then D^l_K valid training samples can be generated. Sufficient training samples can ensure that the input layer and the sequence encoding layer are fully trained so that they could better describe the co-occurrence characteristics and dependencies of online learning behavior sequences.

For the prediction tasks of learning resource recommendation, the initialized parameter values of the input layer and the sequence encoding layer of the model directly adopted the corresponding parameter values attained in the pre-training. The student learning preference representation f_r was the clustering result of the N-th hidden layer representations $F^N_r=[f^N_{r1},\ldots,f^N_{rK}]$ of all period samples in the online learning behavior sequence, and the clustering method was the average pooling method. Assuming: t_o represents the target period sample to be predicted; t_o^g represents the online learning behavior feature of t_o indexed from feature matrix T^g; f_o represents the hidden layer representation of t_o attained from t_o^g via the multilayer cognition mechanism. At last, f_r and f_o were joint, and the joint result was input into the model interaction layer to predict the probability b^* of students accepting the learning resource. Assuming: $b\in\{0,1\}$ represents the classification tag; 0 represents the period sample that students do not accept the learning resource; 1 represents the period sample that students accept the learning resource; $\rho(.)$ represents the *sigmoid* function; then the loss function of this stage is given by the cross-entropy loss function below:

$$SQ_G = -b \cdot log\rho(\hat{b}) - (1-b)log(1-\rho(\hat{b})) \tag{11}$$

5 Experimental results and analysis



a) **b)**

Fig. 3. Trend of students' effective learning behavior implementation frequency

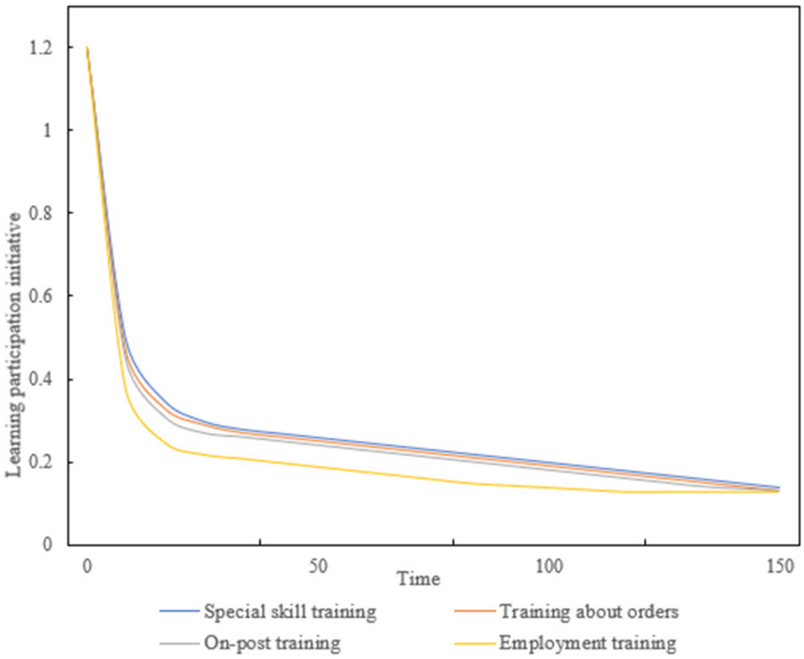


Fig. 4. Differences in students' learning participation initiative in different learning programs

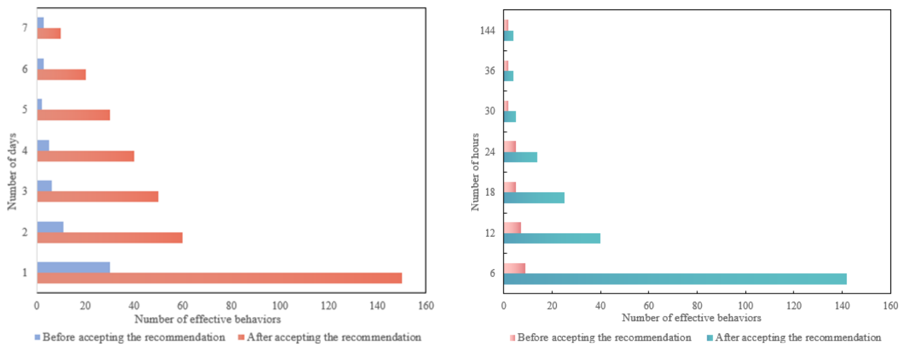
This paper counted the effective learning behaviors of students implemented during the OVE process that are conducive to their ability improvement, to observe the implementation situation of students' effective learning behaviors more intuitively, this paper plotted a bar chart, see Figure 3. According to the figure, the students' effective learning behavior implementation frequency is the highest on the day they participate in learning, then with the passing of learning end time, the implementation frequency decreases gradually until it approaches 0, which is consistent with the experimental effect shown by the slope graph.

To figure out students' situations of continuously participating in different learning programs during the OVE process, this paper compared the participation initiative of students in different learning programs. Figure 4 shows the differences in students' learning participation initiative in different learning programs. According to the figure, students' participation initiative of employment training is lower than that of special skill training, training about orders, and on-post training. This is because for different learning programs, students have different requirement orientations for the OVE. In order to further judge whether there are differences in students' participation initiative of these three types of learning programs, this paper further carried out the LogRank test and the Breslow test, and the results verified that there're significant differences in their participation initiative for different learning programs.

Table 1. Regression results of student learning behaviors

Indicator	Learning Resource Acceptance		Interactive Learning			Individual Learning			Model Test Result
	Number of Video-Type Resources	Number of Test Question-Type Resources	Number of Joined Groups	Number of Group Members	Active Time of Groups	Frequency of Behavior	Initial Participation Degree	Participation Stability	
B	−0.036	−0.051	−0.395	0.015	−0.162	0.058	0.069	−1.629	
SE	0.021	0.025	0.061	0.058	0.047	0.012	0.015	0.061	*X*2=6124.84
WolD	55.174	362.912	418.295	93.624	92.581	418.325	31.259	451.728	*Df*=7
Sig.	0.014	0.025	0.016	0.036	0.095	0.027	0.015	0.028	*Sig*=0.025
Exp(*B*)	0.918	0.936	0.614	1.314	1.518	0.815	1.629	0.274	

In order to consider multiple factors that affect students' continuous participation in OVE training programs at the same time, this paper built a Cox regression model, namely the proportional hazards model, to study the significance of the influence of eight factors (independent variables) on students' continuous participation time (duration) of OVE training programs, and determine the degree of the influence of the eight factors on the risk of too-low frequency of students' effective learning behavior implementation. The eight independent variables are: the number of video-type resources, the number of test question-type resources, the number of joined groups, the number of group members, the active time of groups, the frequency of behavior, the initial participation degree, and the participation stability. The eight independent variables were selected and taken as the covariates of the Cox regression model, then the behavior frequency of samples and the observed values of corresponding independent variables were input into the SPSS software. Table 1 lists the fitting results based on the constructed model. According to the data in the table, 7 of the 8 independent variables have a significant impact on students' participation duration of OVE training programs, and these variables are the number of video-type resources, the number of test question-type resources, the number of joined groups, the active time of groups, the frequency of behavior, the initial participation degree, and the participation stability.



a) b)

Fig. 5. Distribution of students' effective learning behaviors

Next, from a global perspective, this paper analyzed the distribution of each category of online learning behaviors of students in different time periods, and calculated the distribution of effective learning behaviors of students in the behavior sequence sample dataset before and after they accept the learning resource recommendation, the results are shown in Figure 5, as can be seen from the figure, the effective learning behaviors implemented by students before accepting the recommendation of learning resources are not necessarily implemented after accepting the recommendation, and this indicates that there are differences in the behavior patterns of students before and after accepting the recommendation of learning resources.

Table 2. Statistics of students' online learning behavior sequences before and after accepting the recommendation of learning resources

		Average Sequence Length	Average Number of Categories	Average Number of Repetitions
Before accepting the recommendation of learning resources	Day	15.29	8.51	2.38
	Hour	18.54	6.19	
After accepting the recommendation of learning resources	Day	12.47	8.63	2.51
	Hour	14.29	7.48	

Based on Table 2, it's known that there are significant inter-domain differences in the online learning behavior sequences of students before and after accepting the learning resource recommendation. The behavior patterns of students before accepting the recommendation are very different from those after accepting the recommendation, however, there're also some common preferences in the online learning behavior sequences of students before and after accepting the recommendation. This paper has paid more attention to the time sequence information in the online learning behavior sequences of students, because the online learning behavior sequences of students after receiving the recommendation can better describe their short-term dynamic learning preferences. The above experimental results are the motive of this paper to conduct research on the resource recommendation based on the modelling of the online learning behavior sequences of students.

6 Conclusion

This paper studied the resource recommendation algorithm based on the modelling of the online learning behavior sequences of students, and solved the problem of conventional learning resource recommendation algorithms of ignoring the dynamic changes in students' learning preferences. At first, this paper presented the process of LBA applicable for OVE students, and gave a diagram of the flow of using big data to perform LBA. Then, the paper developed a bidirectional encoder based on self-attention model, described the co-occurrence characteristics and dependencies of students' online learning behavior sequences, and introduced the model pre-training and fine-tuning processes in detail. Combining with experiment, this paper counted the effective learning behaviors of students implemented during the OVE process that are conducive to their ability improvement, and plotted the correspond bar chart. After that, to figure out students' situations of continuously participating in different learning programs during the OVE process, LogRank test and the Breslow test were performed and the results verified that there're significant differences in their participation initiative for different learning programs. At last, from a global perspective, this paper analyzed the distribution of each category of online learning behaviors of students in different time periods, and the experimental results indicate that there are differences in the behavior patterns of students before and after accepting the recommendation of learning resources.

7 References

[1] Zeng, Y.S. (2020). Evaluation of physical education teaching quality in colleges based on the hybrid technology of data mining and hidden Markov model. International Journal of Emerging Technologies in Learning, 15(1): 4–15. https://doi.org/10.3991/ijet.v15i01.12533
[2] Neha, K., Sidiq, J., Zaman, M. (2021). Deep neural network model for identification of predictive variables and evaluation of student's academic performance. Revue d'Intelligence Artificielle, 35(5): 409–415. https://doi.org/10.18280/ria.350507
[3] Ros, S., Lázaro, J.C., Robles-Gómez, A., Caminero, A.C., Tobarra, L., Pastor, R. (2017). Analyzing content structure and Moodle milestone to classify student learning behavior in a basic desktop tools course. In Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality, 1–6. https://doi.org/10.1145/3144826.3145392
[4] Na, W. (2020). A data mining method for students' behavior understanding. International Journal of Emerging Technologies in Learning, 15(6): 18–32. https://doi.org/10.3991/ijet.v15i06.13175
[5] Man, M., Azhan, M.H.N., Hamzah, W.M.A.F.W. (2019). Conceptual model for profiling student behavior experience in e-learning. International Journal of Emerging Technologies in Learning, 14(21): 163–175. https://doi.org/10.3991/ijet.v14i21.10936
[6] Šarić-Grgić, I., Grubišić, A., Šerić, L., Robinson, T.J. (2020). Student clustering based on learning behavior data in the intelligent tutoring system. International Journal of Distance Education Technologies (IJDET), 18(2): 73–89. https://doi.org/10.4018/IJDET.2020040105

[7] Wen, J., Wei, X.C., Liu, H.P., Rong, Y.Y. (2020). Fuzzy cluster analysis on influencing factors of college student scores. Revue d'Intelligence Artificielle, 34(5): 607–616. https://doi.org/10.18280/ria.340511
[8] Gusnina, M., Wiharto, Salamah, U. (2022). Student performance prediction in Sebelas Maret University based on the random forest algorithm. Ingénierie des Systèmes d'Information, 27(3): 495–501. https://doi.org/10.18280/isi.270317
[9] Liang, K., Zhang, Y., He, Y., Zhou, Y., Tan, W., Li, X. (2017). Online behavior analysis-based student profile for intelligent e-learning. Journal of Electrical and Computer Engineering. https://doi.org/10.1155/2017/9720396
[10] Carpenter, D., Emerson, A., Mott, B.W., Saleh, A., Glazewski, K.D., Hmelo-Silver, C.E., Lester, J.C. (2020). Detecting off-task behavior from student dialogue in game-based collaborative learning. In International Conference on Artificial Intelligence in Education, 12163: 55–66. https://doi.org/10.1007/978-3-030-52237-7_5
[11] Wang, J., Zhang, Y. (2019). Clustering study of student groups based on analysis of online learning behavior. In Proceedings of the 2019 International Conference on Modern Educational Technology, 115–119. https://doi.org/10.1145/3341042.3341065
[12] Peng, W., Wang, Z., Zheng, J. (2019). A detection model for e-learning behavior problems of student based on text-mining. In International conference of artificial intelligence, medical engineering, education, 1126: 404–413. https://doi.org/10.1007/978-3-030-39162-1_37
[13] Cao, R., Sun, L. (2020). Design and practice of machine learning course based on CDIO and student behavior data. In 2020 15th International Conference on Computer Science & Education (ICCSE), 553–556. https://doi.org/10.1109/ICCSE49874.2020.9201853
[14] Yang, B., Yao, Z., Lu, H., Zhou, Y., Xu, J. (2020). In-classroom learning analytics based on student behavior, topic and teaching characteristic mining. Pattern Recognition Letters, 129: 224–231. https://doi.org/10.1016/j.patrec.2019.11.023
[15] Cantabella, M., Martínez-España, R., Ayuso, B., Yáñez, J.A., Muñoz, A. (2019). Analysis of student behavior in learning management systems through a big data framework. Future Generation Computer Systems, 90: 262–272. https://doi.org/10.1016/j.future.2018.08.003
[16] Chen, C., Meng, X. (2021). Exploring the relationship between student behavioral patterns and learning outcomes in a SPOC. International Journal of Distance Education Technologies (IJDET), 19(1): 35–49. https://doi.org/10.4018/IJDET.2021010103
[17] Tanaka, H., Utsunomiya, Y., Okuda, T. (2017). Modeling and performance evaluation of "the students' learning behavior" considering growth of a student. IEEJ Transactions on Electronics, Information and Systems, 137(3): 431–444. https://doi.org/10.1541/ieejeiss.137.431
[18] Brenner, D.G., Matlen, B.J., Timms, M.J., Gochyyev, P., Grillo-Hill, A., Luttgen, K., Varfolomeeva, M. (2017). Modeling student learning behavior patterns in an online science inquiry environment. Technology, Knowledge and Learning, 22(3): 405–425. https://doi.org/10.1007/s10758-017-9325-0
[19] Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., Lester, J. (2020). The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. Computers & Education, 147: 103781. https://doi.org/10.1016/j.compedu.2019.103781
[20] Kao, C.C., Luo, Y.J. (2020). Effects of multimedia-assisted learning on learning behaviors and student knowledge in physical education lessons: Using basketball game recording as an example. International Journal of Emerging Technologies in Learning, 15(1): 119–139. https://doi.org/10.3991/ijet.v15i01.11393
[21] Bos, N.R., Brand-Gruwel, S., Uhomoibhi, J., Costagliola, G., Zvacek, S., McLaren, B.M. (2016). Profiling student behaviour in a blended course: Closing the gap between blended teaching and blended learning. In Proceedings of the 8th International Conference on Computer Supported, 65–72. https://doi.org/10.5220/0005724300650072

[22] Zhou, N., Zhang, Z.F., Li, J. (2020). Analysis on course scores of learners of online teaching platforms based on data mining. Ingénierie des Systèmes d'Information, 25(5): 609–617. https://doi.org/10.18280/isi.250508
[23] Ning, S. (2021). Analysis of college students' behavior based on machine learning and big data technology. In International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy, 97: 145–151. https://doi.org/10.1007/978-3-030-89508-2_19
[24] Zhang, C., Gao, M., Song, W., Li, C. (2022). Formative evaluation of college students' online English learning based on learning behavior analysis. International Journal of Emerging Technologies in Learning, 17(10): 240–255. https://doi.org/10.3991/ijet.v17i10.31543
[25] Li, L., Chen, D., Li, T. (2022). College students' learning decision-making based on group learning behavior. International Journal of Emerging Technologies in Learning, 17(8): 155–168. https://doi.org/10.1002/bdm.2271
[26] Zhang, X., Li, Q. (2022). Comparative analysis of students' learning behavior in smart learning environment. In International Conference on Forthcoming Networks and Sustainability in the IoT Era, 129: 336–341. https://doi.org/10.1007/978-3-030-99616-1_45
[27] AlQaheri, H., Panda, M. (2022). An education process mining framework: Unveiling meaningful information for understanding students' learning behavior and improving teaching quality. Information, 13(1): 29. https://doi.org/10.3390/info13010029
[28] Zhang, T., Taub, M., Chen, Z. (2022). A multi-level trace clustering analysis scheme for measuring students' self-regulated learning behavior in a mastery-based online learning environment. In LAK22: 12th International Learning Analytics and Knowledge Conference, 197–207. https://doi.org/10.1145/3506860.3506887

8 Authors

Yufeng Jia, female, was born in June 1987. She graduated from Hebei Normal University, and now works as lecturer at Shijiazhuang Vocational and Technical college. Her research directions include educational management, and ideological and political education. She has won 2 municipal honors and more than 10 academy honors, published 10 papers on education, ideological and political education, and student management in provincial-level journals, and hosted/participated in 10 research projects above the school level. Email: 2011110615@sjzpt.edu.cn

Qi Zhao, received a master's degree in education from Hebei Normal University, and now works as lecturer at Hebei Women's Vocational College. She mainly teaches preschool education, developmental psychology of preschool children, etc. Her course "preschool education" was recognized as the excellent online course. For four consecutive years, she won two second prizes and one first prize in Hebei Provincial Teaching Informatization Competition, and coached bronze prize winners of the Internet Plus Innovation and Entrepreneurship Competition of College Students in Hebei Province. She has authored a book named Development of Preschool Education and Its Integration with Arts, served as the associate editor of the book called Preschool Education, and co-edited three textbooks. In addition, she has published more than ten papers in provincial journals like Vocational Education Communication. Email: zhaoqiedu@163.com

Article submitted 2022-08-03. Resubmitted 2022-09-19. Final acceptance 2022-09-20. Final version published as submitted by the authors.