

# Big Data-Based Behavior Analysis of Autonomous English Learning in Distance Education

<https://doi.org/10.3991/ijet.v17i13.32799>

Bing Han

Department of Public Basic Education, Hebi Polytechnic, Hebi, China  
hanbing@hbzy.edu.cn

**Abstract**—Distance English education platforms like ABC360 and Rocky English provides a new learning model, which facilitates the autonomous English learning. This study aims to clarify the relationship between learning process and learning efficiency of distance English education, and provide personalized services for distance English teaching. To this end, we carried out a big data analysis on the behavior of autonomous English learning in distance education, in a bid to reveal the learning law and behavior pattern of the learners. Section 2 provides the research strategy for the motivation and behavior of autonomous English learning in distance education, and establishes a research model for such motivation and behavior. Section 3 constructs a behavior analysis model for autonomous English learning in distance education, computes the correlation between the target autonomous learning resources (ALRs) and the series of clicked ALRs, and calculates the probability for autonomous English learners in distance education to click on new resources. The proposed network model can handle a massive amount of learning behavior data, i.e., boast a certain ability to process the big data. The three parts of the model were introduced in details. The effectiveness of the model was demonstrated through experiments.

**Keywords**—distance education, autonomous English learning, big data analysis, attention mechanism

## 1 Introduction

English learning is a learning process integrating various psychological activities, such as observation, simulation, cognition, reflection, and memorization [1-4]. From learning words, memorizing words, understanding grammar, to situational application, fragmented and personalized learning models enable students to establish conditional reflection to English sentences and contexts [5-10]. Distance English education platforms like ABC360 and Rocky English provides a new learning model, which facilitates the autonomous English learning [11-16]. Given a computer or smartphone with Internet access, students can learn English autonomously anytime, anywhere, including applied oral English like English for overseas studies and workplaces [17-24]. With students as the subjects, the autonomous English learning is immune to the interference from any people or things. Through autonomous learning, the students can

continuously improve their English proficiency, through reading, listening, speaking, research, observation, and practice. This novel learning model emphasizes the cultivation of active and conscious learning, and pays attention to students' learning preferences. To a certain extent, the learning activities of autonomous learners are stochastic and uncertain. Thus, it is difficult to analyze their learning behaviors.

During the Covid-19 epidemic, Cui and Yang [25] surveyed the self-regulation of college English learners in online learning. The survey was carried out in a presentation-assimilation-discussion (PAD) classroom in a local college of China. During the online learning, PAD classroom was implemented for a semester. After that, the learners' self-regulation ability was tested through a questionnaire survey in three dimensions: self-preparation, self-management, and self-evaluation. The teaching effect of English listening courses in Chinese colleges is generally unsatisfactory. Zhou [26] designed an autonomous learning platform based on artificial intelligence (AI), which attempts to improve the listening ability of students with the aid of AI. The design concept and operation process of AI-based learning platforms originate from the current application of AI-assisted English listening platforms. Weng [27] surveyed how college students use self-access center (SAC) to learn English autonomously, summarized the current state of their ability of autonomous English learning, and provided feasible suggestions on cultivating their autonomous English learning in the presence of computers and the Internet. Wei [28] hailed Internet-based autonomous learning as an important learning model. Internet support comes from four key areas: knowledge internalization, task, peers, and society. The four supportive areas affect the autonomous English learning in colleges very differently. English learners can utilize different areas of Internet support to enhance their autonomous learning ability. Hence, different teaching strategies should be applied to different learners. Kamsa-Ard and Danvivaht [29] carried out a survey on the online resources utilized by students to improve their oral English, the way these resources are utilized, and whether these resources affect the students' fluency of spoken English. The respondents were Grade 3 English majors in Khon Kaen University. Three sets of data were collected via pre- and post-tests, questionnaire surveys, and semi-structured interviews. The collected data were subjected to frequency analysis.

To sum up, the existing studies have not sufficiently introduced the external factors affecting the learning pattern, and the analysis approaches for learning pattern. Moreover, the correlation between learning pattern and learning efficiency has not been deeply analyzed. This study aims to clarify the relationship between learning process and learning efficiency of distance English education, and provide personalized services for distance English teaching. To this end, we carried out a big data analysis on the behavior of autonomous English learning in distance education, in a bid to reveal the learning law and behavior pattern of the learners. Section 2 provides the research strategy for the motivation and behavior of autonomous English learning in distance education, and establishes a research model for such motivation and behavior. Section 3 constructs a behavior analysis model for autonomous English learning in distance education, computes the correlation between the target autonomous learning resources (ALRs) and the series of clicked ALRs, and calculates the probability for autonomous English learners in distance education to click on new resources. The proposed network

model can handle a massive amount of learning behavior data, i.e., boast a certain ability to process the big data. The three parts of the model were introduced in details. The effectiveness of the model was demonstrated through experiments.

## **2 Research strategy**

When the demand of a student for distance English learning is not satisfied, he/she will generate internal stimuli for autonomous learning, which pressurize the student to compete with other students, and motivate him/her to learn autonomously. Then, the student will prepare an autonomous learning plan and determine the target behavior, aiming to satisfy the learning demand. Once the demand is satisfied, the English learning effect is improved, and the motivation for autonomous learning is further enhanced. In this case, the student will produce new demand and engage in new activities of distance learning. The entire process is shown in Figure 1.

Figure 2 displays the research model of distance English learning motivation and behavior. It can be clicked that: statistical behavior features, learning preferences, and learning motivation directly act on the students' motivation of using distance English education platform, which in turn affects the learning activities. The learning time and learning location are important factors of distance education. With the proliferation of the Internet technology and smart terminals (e.g., laptop, smartphone, and tablet computer), distance education transcends the geographical limitation, and facilitates fragmented learning. Under distance education, the students differ significantly in learning time, and their learning locations vary greatly.

To model the long-term behavior of autonomous English learning in distance education, this paper builds up a stratification model for behavior series, which illustrates the long-distance dependence amidst the behavior series on autonomous English learning in distance education. In this way, the dynamic preference variation is clearly characterized. This novel modeling algorithm divides the behavior series into multiple layers based on the time windows. Different attention models were established, and applied separately inside and between time windows, in order to characterize the preferences of autonomous English learners in distance education.

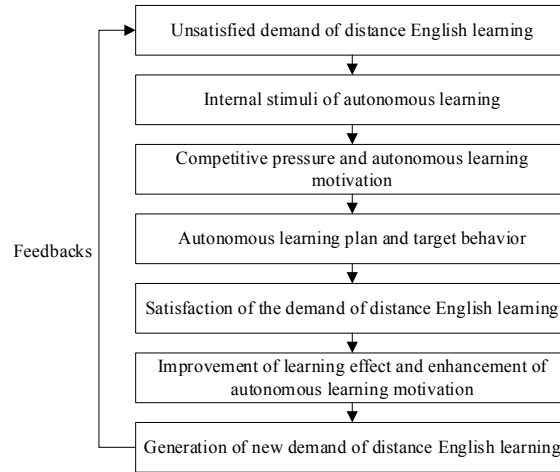


Fig. 1. Relationship between motivation and behavior of distance English learning

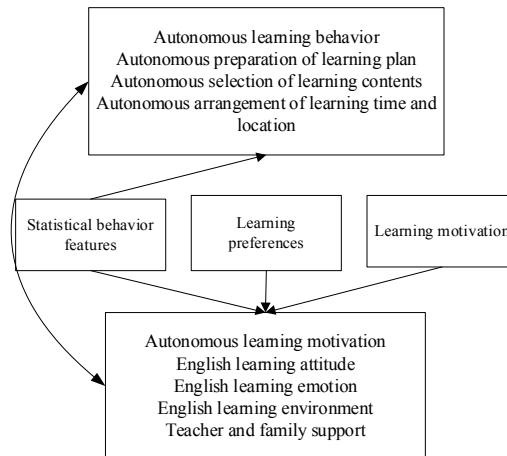


Fig. 2. Research model of distance English learning motivation and behavior

### 3 Behavior analysis model

The cold start problem is a prominent issue in ALR recommendation. This paper computes the correlation between the target ALRs and the series of clicked ALRs, before solving the probability for autonomous English learners in distance education to click on new resources. The traditional content-based filtering methods often ignore the sequential information among autonomous English learning behavior in distance education. In this study, the ALRs clicked by autonomous English learners in distance education are innovatively modeled as a series, in the order of the interaction time.

The resources of autonomous English learning take different forms, namely, audios, videos, and text files. Our network model needs to handle a massive amount of data on learning behavior, i.e., acquire the ability to process the big data. Our model consists of three parts: the feature space for ALR mapping, which is responsible for obtaining the eigenvector corresponding to each resource; the sequential stratified attention model; the prediction module of the click rate of a given new ALR. The three parts are detailed as follows:

To obtain the eigenvector corresponding to each ALR, this study embeds the ARL in an  $e$ -dimensional space. For an autonomous English learner in distance education, the series of clicked ALRs is denoted by  $\{p_1, p_2, \dots, p_m\}$ , and mapped into a eigenvector series of learning resources  $\{a_1, a_2, \dots, a_m\}$ , where  $a_i \in R^e$ . Here, Inception-v3 model is pretrained on ImageNet, and then applied to extract the eigenvector  $p_i^g$  of each autonomous learning behavior from the collected dataset of such behaviors. The high-dimensional eigenvector  $p_i^g$  is dimensionally reduced by learning an embedded matrix  $D_g$ . Let  $g_i \in R^{e_g}$  be the eigenvector of autonomous learning behavior;  $p_i^z$  be the class of ALR  $p_i$ . Then, we have:

$$g_i = D_g p_i^g \quad (1)$$

Since each ALR has and only has one class, the learning resource class can be characterized as a one-hot vector. Let  $p_i^z$  be the class of ALR  $p_i$ ;  $z_i \in R^{e_z}$  be the eigenvector of the class. Similarly, the class eigenvector can be dimensionally reduced by learning  $D_z$ :

$$z_i = D_z p_i^z \quad (2)$$

After dimensionality reduction, the eigenvector of autonomous learning behavior is stitched by the class eigenvector. Let  $[*]$  denote the stitching operation. Then, we have  $a_i[g_i; d_i]$ . As a result,  $a_i \in R^e$ , where  $e = e_g + e_z$ .

If the behavior series is very long for an autonomous English learning in distance education, the ALR eigenvector series  $\{a_1, a_2, \dots, a_m\}$  needs to be modeled by a recurrent neural network (RNN), and extended to multiple scales, using the sequential stratified attention mechanism. These moves aim to better depict the long-distance dependence of learning behavior, and reduce the computational complexity of the model.

In this paper, the behavior series of autonomous English learning in distance education is split into  $n$  time windows, each of which contains  $l$  ALRs:  $n \times l = m$ . To extract the local short-term preference features of autonomous English learning in distance education from each time window, this paper establishes an attention model based on the class level and entry level of ALRs. Next, the forward multi-head self-attention model is called to mine the correlation between time windows, such as to obtain the global preferences of autonomous learners from the 1<sup>st</sup> to the  $i$ -th time window. After that, the local preference features are combined with the global preference features, and normalized to obtain the final preference features of each time window, concerning the autonomous English learning in distance education.

The following explains how to realize the attention model based on the class level and entry level of ALRs. Let  $p_{ij}$  be the  $j$ -th ALR in the  $i$ -th time window ( $i=1, 2, \dots, n$ ,

$j=1,2,\dots,l$ ;  $Q_z \in R^{ez \times ez}$ ,  $Q_1 \in R^{ez \times ez}$ , and  $Q_2 \in R^{ez \times eg}$  be the parameter matrices to be learned;  $\varphi_1$  and  $\varphi_2$  are  $e_z$ -dimensional bias vectors;  $\varepsilon(*)$  be the activation function. Then, the class level attention score can be calculated by:

$$\beta_z(z_{ij}) = Q_z \varepsilon(Q_1 z_{ij} + Q_2 g_{ij} + \phi_1) + \phi_2 \quad (3)$$

The above score can be normalized by:

$$\tilde{\beta}_z(z_{ij}) = \frac{\exp(\beta_z(z_{ij}))}{\sum_{j=1}^l \exp(\beta_z(z_{ij}))} \quad (4)$$

Finally, the class information of the ALRs in the  $i$ -th time window are merged to represent the coarseness preference of autonomous English learners in distance education. Let  $\oplus$  be the elementwise multiplication. Then, the merge process can be defined as:

$$k_i^z = \sum_{j=1}^l \tilde{\beta}_z(z_{ij}) \oplus z_{ij} \quad (5)$$

Similarly, the entry-level attention score can be calculated by:

$$\beta_g(g_{ij}) = Q_g \varepsilon(Q_1' z_{ij} + Q_2' g_{ij} + \phi_1') + \phi_2' \quad (6)$$

Let  $Q_g \in R^{eg \times eg}$ ,  $Q_1' \in R^{eg \times ez}$ ,  $Q_2' \in R^{eg \times eg}$ ,  $\phi_1' \in R^{eg}$ , and  $\phi_2' \in R^{eg}$  be the parameters to be learned;  $k_i^z$  and  $k_i^g$  be the merged preferences of the autonomous English learners in distance education in the  $i$ -th time window ( $k_i \in R^e$ ). In the current time window, the ALR information can be merged based on the normalized attention weight:

$$k_i^g = \sum_{j=k}^l \tilde{\beta}_g(g_{ij}) \oplus g_{ij} \quad (7)$$

To capture the sequential changes in the preferences of autonomous English learners in distance education, this paper introduces the multi-head self-attention model to limit the unidirectional flow of the model information (Figure 3). Firstly, the local preference map  $\{k_i, i=1,2,\dots,n\}$  is duplicated three times. The results are denoted as  $\{w_i\}$ ,  $\{l_i\}$ , and  $\{u_i\}$ . Based on  $\{w_i\}$  and  $\{l_i\}$ , the attention score is calculated, and then subjected to weighted aggregation with  $\{u_i\}$ , producing the global preference map  $\{h_i, i=1,2,\dots,n\}$ . Specifically,  $w_i$ ,  $l_i$ , and  $u_i$  are linearly mapped  $f$  times to an  $ef$ -dimensional space, using different projection matrices ( $ef=e/f$ ). Suppose  $o=1,2,\dots,f$ ; the parameters of projection matrices are denoted as  $Q_o^w \in R^{ef \times e}$ , and  $Q_o^l \in R^{ef \times e}$ ; the bias vector is denoted as  $\phi_o \in R^{ef}$ . Then, the attention tensor can be expressed as  $F \in R^{ef \times n \times n \times f}$ . Thus, the attention score of the  $O$ -th mapping can be calculated based on  $w_i$  and  $l_i$ :

$$F_{i,j}^{(o)} = Q_o \varepsilon(Q_o^w w_i + Q_o^l l_j + \phi_o) \quad (8)$$

To capture the long-distance dependence of the behavior series, the above attention tensors are superimposed with a directed mask N. Then, the element can be defined as:

$$N_{i,j}^{(o)} = \begin{cases} 0, & \text{if } i > j \\ -\infty, & \text{otherwise} \end{cases} \quad (9)$$

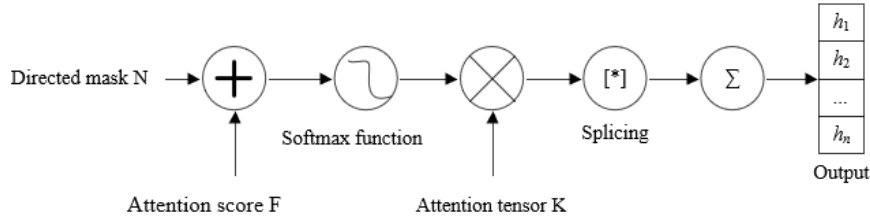


Fig. 3. Multi-head self-attention model

The directed mask N, attention score F, and attention tensor K are illustrated in Figure 4. During the softmax normalization of attention scores, if j is greater than i, then the distance learning location i contributes zero to the distance learning location j. Hence, our attention model only focuses on the scenario of j<i. The corresponding normalized attention weight applies to the  $K \in R^{ef \times n \times f}$ , where each element is an  $e_f$ -dimensional vector. Suppose  $o=1, \dots, f$ , and  $Q_o^r$  is the projection matrix parameter. Then, the element can be defined as:

$$K_{i,j}^{(o)} = Q_o^r \tau_j \quad (10)$$

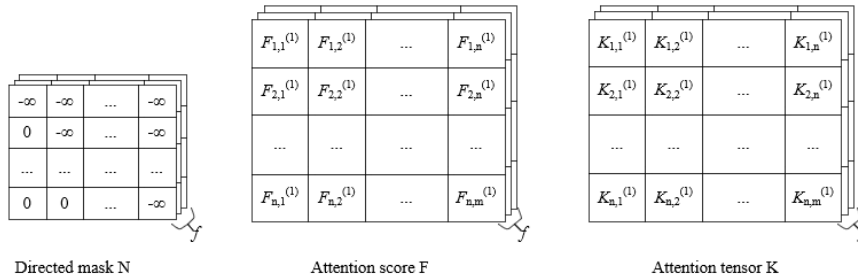


Fig. 4. Illustrations of N, F, and K

Figure 5 shows the acquisition flow of attention tensor K. Since  $e=e_f \times f$ , tensor  $K \in R^{ef \times n \times n \times f}$ , after being applied the attention weight, is stitched into  $K \in R^{n \times n \times e}$  along the dimension O. Then, superposition is performed along the dimension i, yielding  $\{h_1, h_2, \dots, h_n\}$ , with  $h_i \in R^e$ . Let  $\{v_1, \dots, v_n\}$  denote the final preference series of autonomous English learners in distance education,  $v_i \in R^e$ ; GU denote the normalization process. Then, we have:

$$v_i = GU(k_i + h_i) \quad (11)$$

The next is to predict the probability of an autonomous English learner in distance education to click on a new ALR, based on the preference series  $\{v_1, v_2, \dots, v_n\}$ . Here, a new ALR is denoted by an  $e$ -dimensional vector  $a$ . For a given  $a$ , the preference series is aggregated by the attention model. Let  $Q_v \in R^{e \times e}$ ,  $Q_3 \in R^{e \times e}$ ,  $Q_4 \in R^{e \times e}$ , and  $\phi_3 \in R^e$  be the parameters to be learned;  $\varepsilon(\cdot)$  be the activation function. Then, the aggregation function can be given as:

$$\beta_v(v_i) = Q_v \varepsilon(Q_3 v_i + Q_4 a + \phi_3) \tag{12}$$

The attentions score obtained by formula (12) is normalized into  $\beta^*_v(v_i)$ . Then, weighted aggregation is performed between  $\beta^*_v(v_i)$  and  $v_i$ , yielding  $v^e = \sum_{i=1}^n \beta^*_v(v_i) \oplus v_i$ . According to the behavior series of autonomous English learning in distance education, the long short-term dynamic preferences are represented by a vector  $v^e$ . To further capture the static preferences, this paper stitches the preference vectors  $v^e$  and  $v^r$  with the ALR vector  $a$ , and then predicts the ALR click rate, using a multilayer perceptron (MLP). Let  $[\cdot]$  denote the stitching operation ( $v^r \in R^e$ ,  $v^e \in R^e$ ), and NUG denote the MLP. Then, we have:

$$\hat{\phi} = NUG([\![v^e; v^r; a]\!]) \tag{13}$$

The training error of the network can be expressed as:

$$E(v, a) = -b \log \hat{b} - (1-b) \log(1-\hat{b}) \tag{14}$$

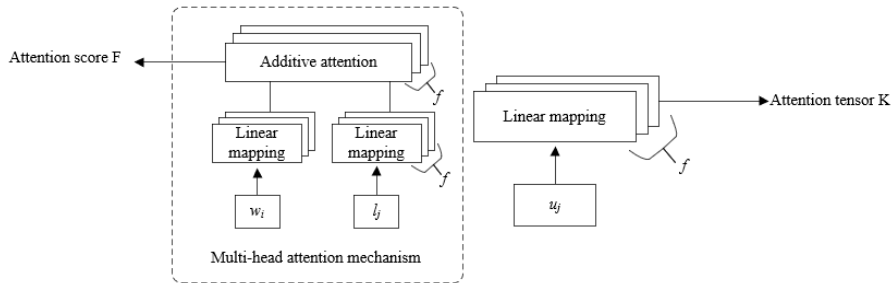


Fig. 5. Acquisition of attention tensor

## 4 Experiments and results analysis

The authors firstly tested the model performance at different time window sizes. The results in Table 1 show that: when the time window was smaller than 10, our model performed similarly on AUC, GAUC, and DNCG. When the time window was of the size 6, our model reached the best performance. With the gradual increase of the time window, the proposed model performed slightly worse on the three metrics. Overall, the preferences of autonomous English learners in distance education are close to each other, when the time window is relatively small. In this case, the proposed class-level



and entry-level attention model can easily emulate the autonomous learning behavior in each time window. If the time window is large, the autonomous learning preferences change significantly in each time window, adding to the difficulty of modeling. In this case, it is difficult for the model to achieve an ideal performance, or control the time cost of training and reasoning. To sum up, the time window size of 6 is the most suitable for our model.

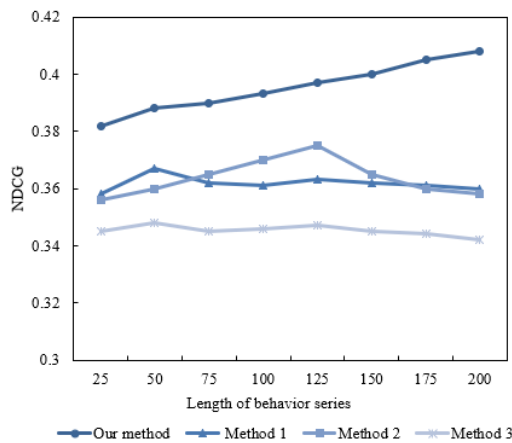
**Table 1.** Model performance at different time window sizes

Time window size	2	4	6	8	10
<i>AUC</i>	0.7152	0.7532	0.7786	0.7648	0.7695
<i>GAUC</i>	0.6258	0.6642	0.6859	0.6728	0.6725
<i>NDCG</i>	0.3625	0.3958	0.4142	0.3852	0.3862
Time window size	18	26	34	40	48
<i>AUC</i>	0.7562	0.7495	0.7341	0.7228	0.7115
<i>GAUC</i>	0.6681	0.6695	0.6536	0.6452	0.6401
<i>NDCG</i>	0.3747	0.3785	0.3625	0.3648	0.3547

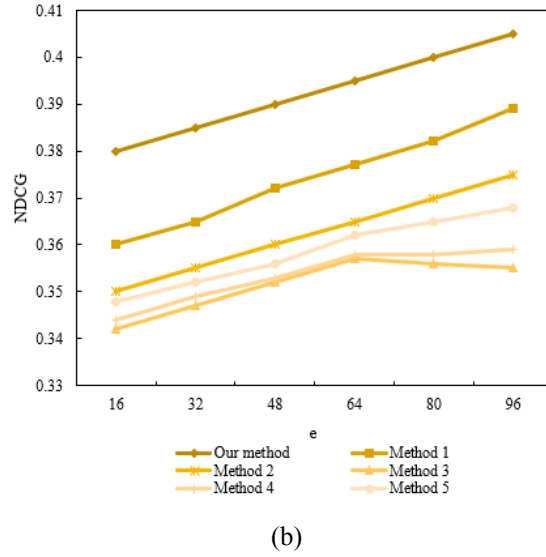
Note: *AUC*, *GAUC*, and *NDCG* are short for area under the curve, generalized area under the curve, and normalized discounted cumulative gain, respectively.

Further, we investigated the influence of behavior series length and embedding dimension  $e$  on model performance. The test results are reported in Figures 6-8. The two eigenvectors of ALRs were set to half of the embedding dimension  $e$ .

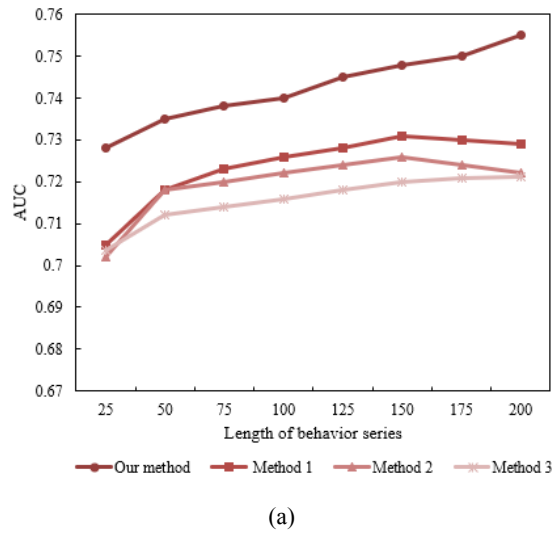
According to the subgraph (a) of Figures 6-8, the proposed model outshined the contrastive models significantly in all three metrics, regardless of the length of the behavior series. The three contrastive models are self-attentive sequential recommendation (SASRec), deep interest evolution network (DIEN), and Caser model. SASRec, DIEN, and Caser are denoted as Models 1-3, respectively.

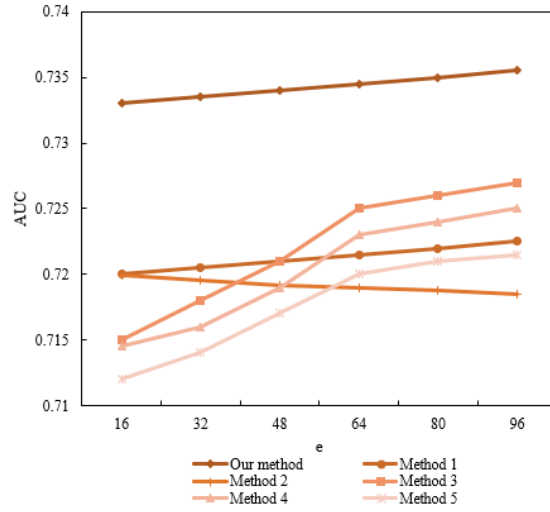


(a)



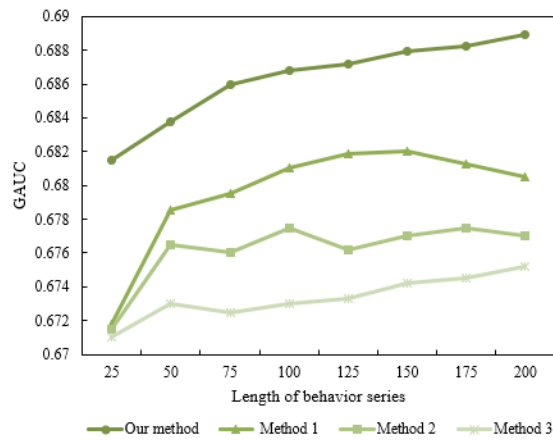
**Fig. 6.** Comparison of NDCG performance





(b)

Fig. 7. Comparison of AUC performance



(a)

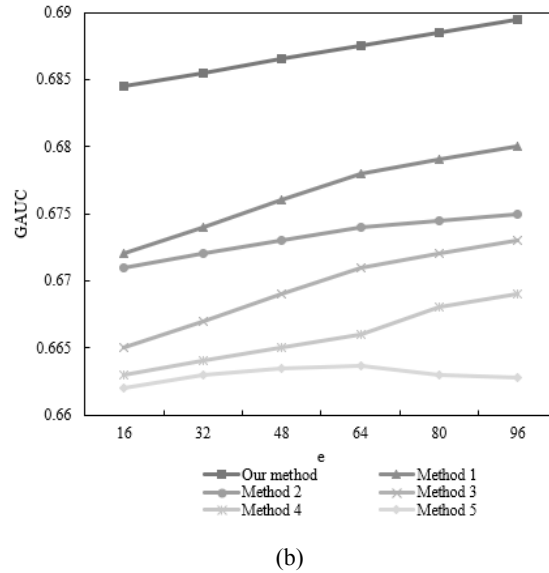


Fig. 8. Comparison of GAUC performance

Among the three metrics, GAUC is obtained by weighting the AUC of each learner. This metric can effectively rank the personalized preferences of autonomous English learners in distance education. That is why our strategy based on the behavior series of autonomous learning outperforms the other models in the ranking of global preferences of learning behavior.

According to the subgraph (b) of Figures 6-8, the proposed model outperformed the contrastive models in all three metrics, regardless of the embedding dimension  $e$ . The contrastive models are SASRec, DIEN, Caser, sequential recommendation with bidirectional encoder representations from transformer (BERT4Rec), and behavior sequence transformer (BST). SASRec, DIEN, Caser, BERT4Rec, and BST are denoted as Models 1-5, respectively.

In the behavior series of autonomous English learning in distance education, the ALR feature map can characterize the autonomous learning preferences. When the  $e$  is small, the ALR feature map contains limited information, which restrains the expressivity of the behavior series model. With the growth of  $e$ , the information volume increases in the ALR feature map. Then, the behavior series model can better illustrate the dynamic preferences of students for autonomous English learning in distance education. Meanwhile, the AUC and GAUC of the contrastive models increased relatively slightly, along with the rise of  $e$ . This means, with the increase of  $e$ , our model can describe dynamic preferences for autonomous English learning in distance education better than the static preferences.

## 5 Conclusions

This paper analyzes the autonomous learning behavior in distance English education based on the big data, and reveals the relationship between learning process and learning efficiency. Specifically, we presented the research strategy for the motivation and behavior of autonomous English learning in distance education, devised a research model for such motivation and behavior, and constructed a behavior analysis model for autonomous English learning in distance education. Next, the probability for autonomous English learners in distance education to click on new resources was obtained by solving the correlation between the target ALR and the series of clicked ALRs. The proposed network model can handle a massive amount of learning behavior data, i.e., boast a certain ability to process the big data. The three parts of the model were introduced in details.

In addition, we tested the model performance at different time window sizes, and confirmed that the time window size of 6 is the most suitable for our model. Then, we investigated the influence of behavior series length and embedding dimension  $e$  on model performance. The investigation shows that our strategy based on the behavior series of autonomous learning outperforms the other models in the ranking of global preferences of learning behavior. Moreover, with the increase of  $e$ , our model can describe dynamic preferences for autonomous English learning in distance education better than the static preferences.

## 6 References

- [1] Liu, J.W. (2020). Promoting Effects of Computer Scoring on English Learning of College Students, *International Journal of Emerging Technologies in Learning*, 15(7), 98-109.
- [2] Qin, M., Du, X., Tao, J., Qiu, X. (2016). A study on the optimal English speech level for Chinese listeners in classrooms. *Applied Acoustics*, 104: 50-56. <https://doi.org/10.1016/j.apacoust.2015.10.017>
- [3] Ajisoko, P. (2020). The Use of Duolingo Apps to Improve English Vocabulary Learning, *International Journal of Emerging Technologies in Learning*, 15(7), 149-155.
- [4] Dev, A., Bansal, S.A., Agrawal, S.S. (2021). Cross linguistic acoustic phonetic study of Punjabi and Indian English. In *International Conference on Artificial Intelligence and Speech Technology*, 1546: 151-159. [https://doi.org/10.1007/978-3-030-95711-7\\_13](https://doi.org/10.1007/978-3-030-95711-7_13)
- [5] Li, C. (2022). A study on Chinese-English machine translation based on transfer learning and neural networks. *Wireless Communications and Mobile Computing*, 2022. <https://doi.org/10.1155/2022/8282164>
- [6] Liu, W., Yu, F., Lin, H. (2014). An empirical study on imbalance of English teaching online. In *Applied Mechanics and Materials*, 687: 2375-2378. <https://doi.org/10.4028/www.scientific.net/AMM.687-691.2375>
- [7] Pham, H., Huynh, N., Nguyen, H. (2022). Virtual mobility adoption: a study of factors affecting students' satisfaction toward an online English program at a young Vietnamese higher education institution: virtual mobility adoption. In *2022 13th International Conference on E-Education, E-Business, E-Management, and E-Learning (IC4E)*, 45-52. <https://doi.org/10.1145/3514262.3514282>

- [8] Luo, Y. (2022). A corpus-driven study on three elements of third personal pronoun endophora in English abstracts of Chinese and foreign theses. *Computational Intelligence and Neuroscience*. <https://doi.org/10.1155/2022/5981285>
- [9] Swanson, L.H., Bianchini, J.A., Lee, J.S. (2014). Engaging in argument and communicating information: A case study of English language learners and their science teacher in an urban high school. *Journal of Research in Science Teaching*, 51(1): 31-64. <https://doi.org/10.1002/tea.21124>
- [10] Zhang, R. (2016). Study on improving English reading ability of the electrical automation specialty students in multimedia-assisted self-learning. *International Journal of Emerging Technologies in Learning*, 11(2): 23-28. <https://doi.org/10.3991/ijet.v11i02.5256>
- [11] Guo, A. (2014). A student profile model based online English learning. *Computer Modelling and New Technologies*, 18(11): 921-926.
- [12] Jing, Z., Chen, J., Li, Z. (2014). Autonomous English language learning aided by an online teaching system. *International Journal of Information Technology and Management*, 13(1): 54-63. <https://doi.org/10.1504/IJITM.2014.059148>
- [13] Sun, M. (2015). Application of multimodal learning in online English teaching. *International Journal of Emerging Technologies in Learning*, 10(4): 54-58. <https://doi.org/10.3991/ijet.v10i4.4697>
- [14] Fang, L.X. (2015). English course online learning system design based on MVC. In 2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA), 372-375. <https://doi.org/10.1109/ISDEA.2015.99>
- [15] Sun, M. (2015). Application of multimodal learning in online English teaching. *International Journal of Emerging Technologies in Learning*, 10(4): 54-58. <https://doi.org/10.3991/ijet.v10i4.4697>
- [16] Sun, Y., Jiang, Y. (2021). Application of data mining in English online learning platform. In *Journal of Physics: Conference Series*, 1992(2): 022118. <https://doi.org/10.1088/1742-6596/1992/2/022118>
- [17] Zhou, Z.F. (2021). A resource recommendation algorithm for online English learning systems based on learning ability evaluation. *International Journal of Emerging Technologies in Learning (iJET)*, 16(9): 219-234. <https://doi.org/10.3991/ijet.v16i09.22745>
- [18] Zhou, Z. (2021). Multimedia English online learning behavior intelligent monitoring system based on face recognition. In 2021 IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI), 7-13. <https://doi.org/10.1109/IAAI54625.2021.9699901>
- [19] Tseng, F.C., Liu, P.H.E. (2021). The development of effective evaluation criteria for online English learning resources. In 2021 International Conference on Advanced Learning Technologies (ICALT), 279-281. <https://doi.org/10.1109/ICALT52272.2021.00090>
- [20] Daut, N., Abd Halim, N.D. (2021). Effects of online heutagogy learning environment towards students' creativity in English presentation. In 2021 1st Conference on Online Teaching for Mobile Education (OT4ME), 8-14. <https://doi.org/10.1109/OT4ME53559.2021.9638922>
- [21] Tan, Q. (2022). English Majors' online learning technology needs in China. In 2022 13th international conference on e-education, e-business, e-management, and e-learning (IC4E), 160-167. <https://doi.org/10.1145/3514262.3514353>
- [22] Simon Antero, E., Odanga Simon, E., Francisco Echanis, S., Castillo Dunghit, J., Laroza Silva, D., Marzan Adina, E. (2021). Logistic regression approach in analyzing the learning styles of students in science, mathematics and English online classes. In 2021 13th International Conference on Education Technology and Computers, 195-201. <https://doi.org/10.1145/3498765.3498796>

- [23] Pinandito, A., Hayashi, Y., Hirashima, T. (2021). Online collaborative kit-build concept map: Learning effect and conversation analysis in collaborative learning of English as a foreign language reading comprehension. *IEICE Transactions on Information and Systems*, 104(7): 981-991. <https://doi.org/10.1587/transinf.2020EDP7245>
- [24] Jiemsak, N. (2021). The development of online learning and teaching management in English literature using google classroom for the undergraduate students. In 2021 6th International STEM Education Conference (iSTEM-Ed), 1-4. <https://doi.org/10.1109/iSTEM-Ed52129.2021.9625143>
- [25] Cui, Z.H., Yang, Y.P. (2020). A survey on college English learners' online learning self-regulation in PAD class during pandemic. In 2020 International Conference on Modern Education and Information Management (ICMEIM), 663-666. <https://doi.org/10.1109/ICMEIM51375.2020.00149>
- [26] Zhou, J. (2020). Design of Ai-based self-learning platform for college English listening. In 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), 544-547. <https://doi.org/10.1109/MLBDBI51377.2020.00114>
- [27] Weng, Y. (2018). College students' autonomous English learning in computer and network-based self-access center. In International Conference on Frontier Computing, 542: 915-922. [https://doi.org/10.1007/978-981-13-3648-5\\_115](https://doi.org/10.1007/978-981-13-3648-5_115)
- [28] Wei, L. (2016). A research of college English self-learning based on internet supporting factors. In 2016 International Conference on Robots & Intelligent System (ICRIS), 43-46. <https://doi.org/10.1109/ICRIS.2016.9>
- [29] Kamsa-Ard, T., Danvivaht, U. (2015). Effects of using online resources by undergraduate students for self-directed learning of English speaking. Workshop Proceedings of the 23rd International Conference on Computers in Education, ICCE 2015, 193-199.

## 7 Author

**Bing Han (June 1982)**, female, master's degree, assistant professor. She is a Science and Technology Commissioner of Henan Province and a Cadre Teacher of Vocational Colleges in Henan Province. Serving as a judge of the English speaking contest in her city. She has published many articles on English teaching and presided over a number of research projects.

Article submitted 2022-03-27. Resubmitted 2022-05-26. Final acceptance 2022-05-30. Final version published as submitted by the author.