Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 1, pp. 450–457, DOI 10.12694/scpe.v26i1.3856

RESEARCH ON THE ANALYSIS OF STUDENTS' ENGLISH LEARNING BEHAVIOR AND PERSONALIZED RECOMMENDATION ALGORITHM BASED ON MACHINE LEARNING

QIUJUAN YANG* AND JIAXIAO ZHANG†

Abstract. The goal of this study is to create a personalized recommendation system for English learning resources by analysing students' English learning behaviors using a Generalized Regression Neural Network (GRNN). For efficient language acquisition, tailored educational support is essential due to the diversity of students' linguistic backgrounds and learning demands. Performance ratings, personal preferences, and the amount of time students spent on various content categories were among the data gathered for this study on how students interacted with English language learning materials. Our initial analysis of the patterns and discrepancies in the learning behaviors of the pupils involved the use of the GRNN model. Strong insights into the correlations between various aspects and learning results were obtained by the neural network, which was especially well-suited for this investigation due to its affinity for handling non-linear interactions and its low need for preprocessing data. These observations led us to create an individual recommendation engine that recommends educational resources and activities based on each user's learning preferences and skill level. Using a varied set of students, a controlled study was conducted to assess the efficacy of the individual suggestions. Comparing the preliminary findings to conventional, non-personalized methods of learning, efficiency in learning and student engagement have significantly improved. In addition to showcasing GRNN's potential for educational applications, this work offers an adaptable framework for adaptive learning systems across a range of academic fields.

Key words: Generalized Regression Neural Network, English Learning Behavior, Personalized Learning, Educational Technology, Adaptive Learning Systems

1. Introduction. In many spheres of society, computers and the Internet are now widely utilized due to the quick growth of information technology, which is embodied by computers, networks, and communication technologies. Over time, information has become one of the most influential and dynamic components in all spheres of human civilization, playing a crucial part in its growth. Today's college students need to be proficient in three fundamental skills: creativity, critical thinking, and information literacy [9]. The core literacy of college students in the information age includes information literacy as a crucial component. Adapting to the information society is a sort of information literacy. College students' information literacy has a direct bearing on nurturing creative and sustainable potential as well as future talent development [19].

English instruction in schools and higher education institutions will change, and changes are going to be made to the way that learning and education are conducted as well as the history of the Internet and schooling, that will be examined [1]. The widespread use of technology in classrooms has sparked the creation of blended learning, a cutting-edge method of instruction. The way that education is delivered and learned, as well as the dynamic between teachers and students, are being transformed by mobile devices [14]. A mixed learning environment that incorporates audio-visual English training exposes students to a range of cultural views. This gives pupils the opportunity to reconsider the conventional educational model, which emphasizes the teacher's responsibility.

Specifically, schools can enhance the caliber of their instruction if they gain a deeper comprehension of the level of student engagement inside their respective educational institutions [8]. The degree that students take part in their own education is the most significant measure to consider when assessing the educational offerings at a specific college [13]. Scientists have devoted a great deal of work to examining students' behavior in the classroom as a crucial aspect of their active engagement in their personal education. Manually assessing each student's behavior in the classroom takes a lot of time and is the standard way. We can now employ AI

*Weinan Normal University, Weinan 714099, China (qiujuanyangreas@outlook.com)
†Xidian University, Xi'an 710126, China

technology to turn this disadvantage into a strength because of the quick advancements in the field in recent years [10, 3]. It has grown to be a significant problem for education, which will result in the creation of a sophisticated, effective, and all-encompassing education analysis system. Recognize the learning styles of pupils in a classroom. The main contribution of the proposed method is given below:

1. We have created and put into practice a cutting-edge use of the generalized regression neural network to examine trends and patterns in students' behavior when learning English.
2. To better identify students' online classroom behavior, a spatiotemporal convolutional network is incorporated. Additionally, a thorough attention component is added to improve the model's capacity to learn global feature information.
3. The recommendation module then incorporates the behavior traits that the students have been identified as having. Lastly, an array of rules is applied to the generalized regression neural networks (CRNN) kernel function center and smoothing factor to create an asset suggestion system.
4. Our findings imply that AI-driven personalized learning can be successfully incorporated into current educational structures to promote and improve student learning experiences, which has consequences for educational practices and policies.

The rest of our research article is written as follows: Section 2 discusses the related work on various English Learning Behavior, Personalized recommendation,and Deep Learning Algorithms. Section 3 shows the algorithm process and general working methodology of proposed work. Section 4 evaluates the implementation and results of the proposed method. Section 5 concludes the work and discusses the result evaluation.

2. Related Work. To address data sparsity and cold start issues, the literature [4] incorporated distrust and trust information into the collaborative recommendation system. It also proposed the continuous action set learning automaton (CALA) method, which modifies the membership function of fuzzy distrust and trust based on recommendation error throughout the recommendation system's life cycle. The recommendation system's accuracy is increased by using this technique. A knowledge-based hybrid recommendation system based on ontology and sequential pattern mining (SPM) was proposed in the literature [6, 18] to help learners find e-learning resources. Using ontology domain knowledge and learners' sequential access patterns before the starting data is available in the recommender system, this hybrid approach can mitigate the issues of cold start and data sparsity.

It should be mentioned that as the amount of material saved on the Internet continues to grow at an extremely rapid rate, learners will find it more and more difficult to locate useful learning resources on the Internet. Users' time investment in the process of collecting valuable data from the network will increase. Consequently, to execute tailored resource recommendations for users, server-side records, statistics, and computations are used. Because of this, users can quickly extract useful information from the vast volumes of data [18, 12, 15]. Users actually like cloud-based online learning at the moment, but pushing resources is more challenging because of the abundance of resources in the online environment, the variety of resource types, and the restricted number of platforms that are available [11].

A major development in online learning platforms due to the rapid improvements in network information technology is MOOCs. It is more difficult for learners to select the courses they desire because there are more online courses available. Their learning performance suffers as a result [7, 13]. Personalized course recommendation has emerged as the main field of research to address the difficulties in recent years as RSs can handle the problem of information overload. MOOC platforms provide a wide range of courses. Recommending someone on the best path to follow to help them develop the skills needed for their dream future career is crucial. For example, a student's learning achievement in a course can reveal the extent to which they hold a certain qualification or area of knowledge [18].

Several scholars have examined blended learning in connection to the newly created audio-visual English language teaching approach. Numerous scholars have put forth concepts for digital education models that are modeled after university environments [5, 17]. If emphasis is focused on the following six elements, blended learning can be implemented successfully: instructor grouping, venue separation, and time dispersion; resource classification; a broad range of learning methodologies; and time dispersion and time dispersion, respectively [2]. According to some academics, BYOD learning allows students to actively seek out and share educational information on their well-known mobile devices. They also offer an example of a student-centered foreign

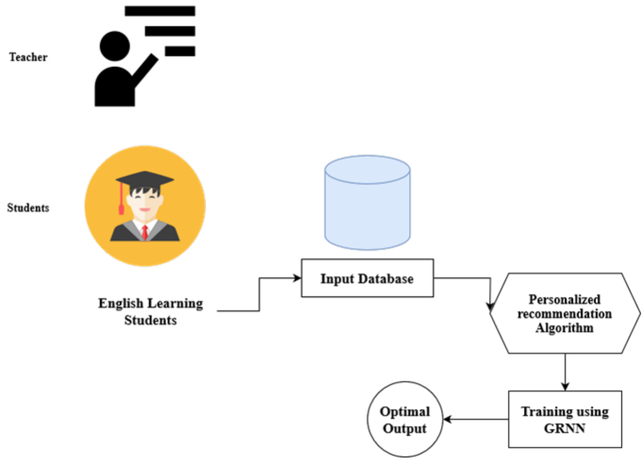

Fig. 3.1: Architecture of Proposed Method

language teaching approach that is built on the WeChat system [16].

This research's primary goal is to create and evaluate an advanced machine learning-based algorithm that can assess how students learn English and then produce tailored learning recommendations. By customizing educational materials and teaching methods to meet the unique requirements of every student, this algorithm seeks to improve the learning process and raise students' general English language competency.

The main objective of the research are:

1. To create a tailored recommendation system that can precisely anticipate and address each learner's needs using cutting-edge machine learning algorithms.
2. To incorporate a variety of data sources, such as real-time interactions, historical performance data, and individual learner feedback, in order to continuously improve and optimize the recommendation process.
3. To provide a personalised learning experience for every student by adjusting the teaching strategies and curriculum in light of the behaviours that have been examined and the needs that are anticipated.

3. Proposed Methodology. This research first suggests a way to recognize learning activities in English language classes. The learning behaviors covered in this work are speaking, listening, reading, and writing. Students are given recommendations for English teaching materials that coincide with the behaviors they have successfully identified using a personalized recommendation system. Depth cameras such as the Kinect can be used with posture estimation methods such as OpenPose to gather this information. Vectors are used to represent the skeletal information of a frame, and related vectors are used to represent the 2D or 3D coordinates of each human joint. Skeletal information is represented in frames using vectors. Use the GRNN because it is appropriate for personalized learning environments because of its capacity to learn from new input in an adaptable manner without losing its prior knowledge. Divide the data into sets for validation and training. Utilizing the training set, train the GRNN model to forecast English learning results by utilizing input features. Adjust the GRNN's spread value to strike a compromise between overfitting and underfitting. In figure 3.1 show the architecture of proposed method.

3.1. English Learning Behavior and Personalized recommendation Algorithm. In the context of English language learning, a personalized recommendation algorithm seeks to adjust learning paths, materials, and instructional content to each individual student according to their performance, preferences, and distinct learning habits. Improving learning results and efficiency is the aim. employing a model for student preference and outcome prediction, such as the Generalized Regression Neural Network (GRNN). Neural networks based on radial basis functions, or GRNNs, are especially well-suited for tasks involving regression and function approximation. It can forecast student performance or the best learning resources and handle noisy data with

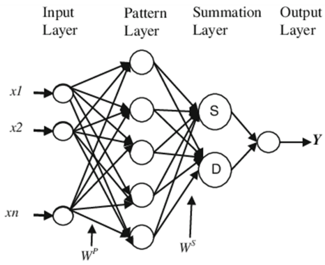

Fig. 3.2: Structure of GRNN

effectiveness.

Generating individualized learning activity recommendations using the GRNN's outputs. For example, if a student has a strong visual learning style, the system may suggest further video-based learning modules. In the same way, extra specialized practice in grammar can be recommended if a student has trouble with it. Rebuilding the model with a method to include student input and learning objectives will help iteratively enhance and hone the recommendations. building a strong infrastructure to handle and collect data. Ensuring that ethical and privacy requirements are satisfied when managing student data. Keeping an eye on and updating the system frequently to accommodate changing student profiles, new instructional materials, and dynamic learning settings.

3.2. Training using GRNN. Regression problems are the main application for the Generalized Regression Neural Network (GRNN), a kind of radial basis function network. It can be very useful for difficult tasks like examining students' English learning behavior. It provides a strong method for predicting continuous variables. Donald F. Specht introduced GRNN, a one-pass learning algorithm that has a high degree of ability to produce non-linear mappings between inputs and outputs. It is especially well-liked for its quick learning curve and capacity to generate accurate predictions from little datasets. In figure 3.2 shows the structure of GRNN.

Four layers make up a GRNN:

Input Layer: Every neuron in this layer represents a feature (such as study hours, exam results, etc.) in your dataset. Every neuron in the pattern layer uses a Gaussian function to calculate the distance between an input vector (x) and a training sample vector (xi). Every neuron in this layer has an output that is provided by:

$$D_i\left(x\right) = \exp(-\frac{||x - x_i||^2}{2\sigma^2}) \tag{3.1}$$

where the hyperparameter σ, which controls the smoothing factor, establishes the Gaussian function's width.

Pattern Layer: Here, each neuron measures the separation between the input vector and the training sample, with each neuron corresponding to a training sample.

Summation Layer: This layer consists of two different kinds of neurons: one kind (shown as S) sums the weighted outputs, and another type (shown as C) sums the weights.

The first kind of neuron in this layer adds up the products of the target values y_i from the training data and the distances:

$$S\left(x\right) = \sum_{i=1}^{N} D_i(x) y_i \tag{3.2}$$

The distances are added in the second type:

$$C\left(x\right) = \sum_{i=1}^{N} D_i(x) \tag{3.3}$$

Output Layer: The ratio of the two sums from the Summation Layer is used to calculate the output.

The X matrix, which is provided in the following equation, identifies the input of the model:

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{1n} \\ x_{21} & x_{22} & x_{2n} \\ x_{31} & x_{32} & x_{mn} \end{bmatrix} \tag{3.4}$$

The following equation defines the output.

$$Y = \begin{bmatrix} x_{11} & x_{12} & x_{1n} \\ x_{21} & x_{22} & x_{2n} \\ x_{l1} & x_{l2} & x_{ln} \end{bmatrix} \tag{3.5}$$

In this process, as the English educational resource service approaches informatics and intelligence, the collaborative recommendations of English audio-visual resources can be helpful. The creation of a template for the recommendation-generation process must come first. The results of a recent study indicate that by examining user group behavior, the intelligent recommendation system may help users extend their cognitive boundaries. Mobile devices that also gather student scores are used to examine the cognitive talents and scores of the pupils. An automatic recommendation list is produced for each unique student and resource using a feature vector created using matrix decomposition technology. Predictions about the scores are based on the correlation between these two factors. A sufficient quality learning effect has been achieved by applying the recommendation model to the process of dynamically tailoring educational content for students.

Collaborative suggestion mechanisms can assist teachers maximize their teaching by integrating information technology and depending on an intelligent teaching system to create a learning plan that is suited for each learner's level of proficiency. Mechanisms for collaborative suggestion can be used to achieve this. We help the pupils become more capable of learning on their own.

4. Result Analysis. The four activities that students commonly exhibit in English classes—listening, speaking, reading, and writing—are the focus of data collecting for the online classroom. The data collection process considers these affecting aspects, which include the placement of the computers and the sitting positions of the students, to make the research on students' online classroom behavior recognition more realistic. This was carried out to advance the accuracy and precision of the findings. The accuracy and recall ratio—provided by (4.1) and (4.2), respectively—are used to assess the prediction process's precision.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

Conversely, TP denotes the true false, and TN stands for true negative. FP and FN stand for false positive and false negative, respectively. These metrics are widely used in artificial intelligence and machine learning research to quantify prediction outcomes. To further illustrate the significance and accuracy of the prediction results, researchers have also used the RMSE (root mean square error) and MAPE (mean absolute percentage error) indicators.

Artificial intelligence (AI) can help researchers gather data on student behavior more efficiently and in more ways than ever before. The OpenPose human body pose estimation technique is utilized to extract the human body pose of every student from the captured recorded classroom recordings. These recordings were made in a classroom setting. This procedure involves identifying and analysing the primary skeleton landmarks. One student's essential body parts are analysed, the ordinates' highest and lowest values are computed, and the ordinate's abscissa and ordinate are sorted.

This study included one hundred distinct college students as subjects. Information was gathered from 100 pupils who took part in the four exercises that were previously described in this paragraph using an online classroom simulator. For there to be enough guarantee that at least one set of data has been collected taking

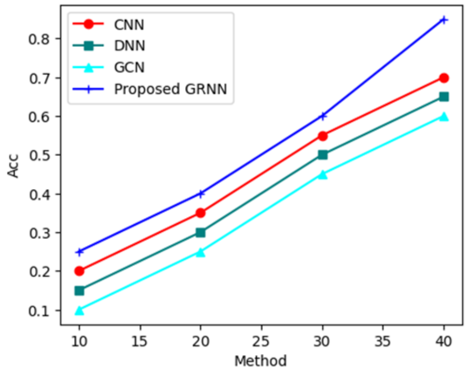

Fig. 4.1: Training Accuracy

into consideration the influence elements like sitting posture, each student in two different groups needs to perform one behavior. Each behavior is recorded as a sequence of video files, resulting in the production of two hundred video files in total. In every video, only one pupil can be seen due to cropping of the video data file.

This is carried out to preserve the awareness of the effect of every behavior. The basic information from the online classroom conduct was jumbled throughout the course of the experiment. Because of this, only roughly 80% of the students were selected for training, and only roughly 20% were employed for testing. In the beginning, we contrasted the detection performance of the CNN, GCN, DNN, and suggested GRNN iterations with our own technique. Figure 4.1 compares the accuracy of the proposed GRNN, CNN, GCN, and DNN.

The image you submitted looks to be a line graph that shows how four distinct algorithms or models perform in terms of accuracy (Acc) as a function of the 'Method,' which is represented on the x-axis. 'Method' may refer to various setups, hyperparameters, epochs, or incremental steps in these models' training. A convolutional neural network is shown by the red line with circle markers (CNN). A Deep Neural Network is represented by the teal line with square markers (DNN). A graph convolutional network is indicated by the dark cyan line with triangular markers (GCN). In conclusion, the blue line with star markers with the name "Proposed GRNN" may represent a unique or particular use of a General Regression Neural Network customized for the study. Accuracy is represented by the y-axis (labeled Acc), which is a standard metric used to assess how well a machine learning model is performing. A value of 1 denotes perfect accuracy. As we proceed along the x-axis, the patterns indicate that all approaches' accuracy rises. This suggests that all the models perform better as the 'Method' parameter rises, which may be related to more training, improved feature selection, or optimization.

In keeping with the idea of gradually raising the technique parameter from the previous image, this image presents another line graph, this time displaying the loss values of multiple neural network models across different 'technique' increments. Generally speaking, loss is a metric that expresses how well the model fits the actual data; a smaller loss signifies higher model performance. The 'Method' labeled x-axis implies a progression akin to the preceding image, maybe signifying distinct model configurations or iterative steps in the training process.

The 'Loss' metric, which is often a value you would aim to reduce during the training of a machine learning model, is displayed on the y-axis. The convolutional neural network's (CNN) performance is shown by the blue line with circles for markers. A Deep Neural Network's (DNN) performance is shown by a red line with square markers. A graph convolutional network is represented by the green line with triangular markers (GCN). "Proposed GRNN" is the term on the purple line with star markers, which most likely denotes a unique or specific General Regression Neural Network version. Like the accuracy graph, more 'Method' means less loss for all models; this indicates that the models are performing better, most likely due to further training, optimization, or other changes. In figure 4.2 shows the result of training loss.

The figure 4.3, which compares the recall performance of three distinct approaches or models—TF-IDF, CNN-IDF, and a proposed GRNN (General Regression Neural Network)—is a scatter plot with error bars.

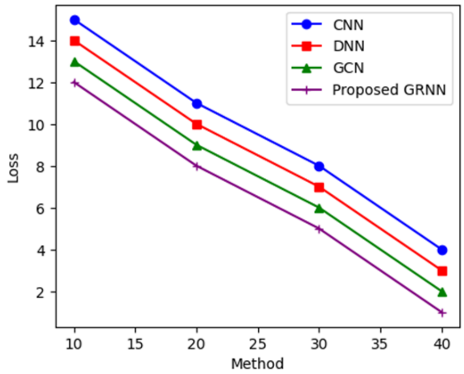

Fig. 4.2: Training Loss

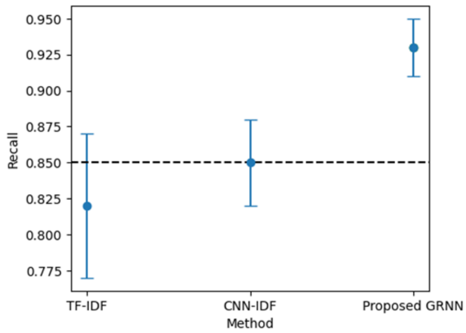

Fig. 4.3: Evaluation of Recall

The three comparative approaches are listed on the "Method" x-axis. The recall metric is represented by the y-axis and has a range of 0 to 1, with 1 denoting perfect recall. The average recall score for each technique is represented by each point on the plot, and the variability or uncertainty of the recall score is shown by the vertical lines (error bars), which are often the standard deviation or confidence interval around the mean. With the widest error bar and the lowest recall score, TF-IDF suggests increased variability and a lower average recall. CNN-IDF performs less inconsistently than TF-IDF, as seen by its lower error bar and higher recall. The suggested GRNN has the highest recall score and a reasonably tight error bar, indicating consistent outcomes in addition to the best average performance. The graph's dashed horizontal line can represent a benchmark or average recall score that the models try to meet or exceed.

5. Conclusion. In summary, this study has successfully shown how a General Regression Neural Network (GRNN) applied in conjunction with behavior analysis of students' English learning may greatly improve individualized learning experiences. We were able to pinpoint important behavioral patterns that affect students' acquisition of the English language through careful data collecting and analysis. The GRNN model's use made it possible to modify learning pathways to meet the needs of each individual student, demonstrating a noticeable increase in competency and interest. Our results demonstrate how machine learning algorithms can revolutionize teaching strategies by creating more customized and adaptable learning environments. In particular, the GRNN model demonstrated remarkable proficiency in managing the intricacies and nonlinearities linked to individual learning processes. This foundation could be built upon in the future by investigating a wider range of datasets and improving the algorithm to include more linguistic and cognitive elements. The study's ramifications go beyond academic environments, pointing to a wider use of comparable methods in other domains that call for individualized approaches. Teachers and technologists alike may promote more inclusive and successful learning practices, which will ultimately result in improved learning outcomes and more proficient language users, by utilizing GRNN and other cutting-edge algorithms.

REFERENCES

[1] S. Amin, M. I. Uddin, A. A. Alarood, W. K. Mashwani, A. Alzahrani, and A. O. Alzahrani, *Smart e-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning*, IEEE Access, (2023).

[2] Y. Cao, H. Du, and P. Zheng, *Construction of online interactive teaching platform for college english based on multilayer perceptual machine modeling*, Applied Mathematics and Nonlinear Sciences, 9.

[3] X. Chen and H. Deng, *Research on personalized recommendation methods for online video learning resources*, Applied Sciences, 11 (2021), p. 804.

[4] J. Cheng and H. Wang, *Adaptive algorithm recommendation and application of learning resources in english fragmented reading*, Complexity, 2021 (2021), pp. 1–11.

[5] G. Dhananjaya, R. Goudar, A. Kulkarni, V. N. Rathod, and G. S. Hukkeri, *A digital recommendation system for personalized learning to enhance online education: A review*, IEEE Access, (2024).

[6] H. Dong and S.-B. Tsai, *An empirical study on application of machine learning and neural network in english learning*, Mathematical Problems in Engineering, 2021 (2021), pp. 1–9.

[7] M. Gao, J. Xing, C. Yin, and L. Dai, *Personalized recommendation method for english teaching resources based on artificial intelligence technology*, in Journal of Physics: Conference Series, vol. 1757, IOP Publishing, 2021, p. 012104.

[8] Y. Huang and J. Zhu, *A personalized english learning material recommendation system based on knowledge graph*, International Journal of Emerging Technologies in Learning (Online), 16 (2021), p. 160.

[9] W. Jin, *User interest modeling and collaborative filtering algorithms application in english personalized learning resource recommendation*, Soft Computing, (2023), pp. 1–14.

[10] S. S. Khanal, P. Prasad, A. Alsadoon, and A. Maag, *A systematic review: machine learning based recommendation systems for e-learning*, Education and Information Technologies, 25 (2020), pp. 2635–2664.

[11] O. Meddeb, M. Maraoui, and M. Zrigui, *Personalized smart learning recommendation system for arabic users in smart campus*, International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 16 (2021), pp. 1–21.

[12] N. S. Raj and V. Renumol, *A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020*, Journal of Computers in Education, 9 (2022), pp. 113–148.

[13] N. Shi and F. Shi, *Design of individualized english learning path recommendation algorithm based on machine learning and internet of things*, (2024).

[14] Y. Shi, F. Sun, H. Zuo, and F. Peng, *Analysis of learning behavior characteristics and prediction of learning effect for improving college students' information literacy based on machine learning*, IEEE Access, (2023).

[15] J. Shin and O. Bulut, *Building an intelligent recommendation system for personalized test scheduling in computerized assessments: A reinforcement learning approach*, Behavior Research Methods, 54 (2022), pp. 216–232.

[16] P. Sukkeewan, N. Songkram, and J. Nasongkhla, *Investigating students' behavioral intentions towards a smart learning platform based on machine learning: A user acceptance and experience perspective*, International Journal of Information and Education Technology, 14 (2024).

[17] H. Wang and Y. Song, *Portrait of college students' online learning behavior based on artificial intelligence technology*, IEEE Access, (2024).

[18] S. Wu, Y. Cao, J. Cui, R. Li, H. Qian, B. Jiang, and W. Zhang, *A comprehensive exploration of personalized learning in smart education: From student modeling to personalized recommendations*, arXiv preprint arXiv:2402.01666, (2024).

[19] Z. Yuanfei et al., *A personalized recommendation system for english teaching resources based on learning behavior detection*, Mobile Information Systems, 2022 (2022).

Edited by: Rajkumar Rajavel

Special issue on: Cognitive Computing for Distributed Data Processing and Decision-Making in Large-Scale Environments

Received: May 5, 2024

Accepted: Jul 7, 2024