Journal of Information Systems and Telecommunication
Vol.11, No.4, October-December 2023, 282-293
http://jist.acecr.org
ISSN 2322-1437 / EISSN:2345-2773

A Recommender System for Scientific Resources Based on Recurrent Neural Networks

Hadis Ahmadian Yazdi[1], Seyyed Javad Seyyed Mahdavi[2*], Maryam Kheirabadi[1]

[1].Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran
[2].Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

Received: 23 Sep 2022/ Revised: 04 Mar 2023/ Accepted: 08 Apr 2023

Abstract

Over the last few years, online training courses have had a significant increase in the number of participants. However, most web-based educational systems have drawbacks compared to traditional classrooms. On the one hand, the structure and nature of the courses directly affect the number of active participants; on the other hand, it becomes difficult for teachers to guide students in choosing the appropriate learning resource due to the abundance of online learning resources. Students also find it challenging to decide which educational resources to choose according to their condition. The resource recommender system can be used as a Guide tool for educational resource recommendations to students so that these suggestions are tailored to the preferences and needs of each student. In this paper, it was presented a resource recommender system with the help of Bi-LSTM networks. Utilizing this type of structure involves both long-term and short-term interests of the user and, due to the gradual learning property of the system, supports the learners' behavioral changes. It has more appropriate recommendations with a mean accuracy of 0.95 and a loss of 0.19 compared to a similar article.

Keywords: Deep Learning Networks; Recurrent Methods; Educational Resource; Recommender System.

1- Introduction

Nowadays, the purpose of online courses is to provide students with the opportunity to learn quickly. However, designing this course is not always the most effective method for all learners and has led to high dropout rates and low academic effectiveness. This model may be as effective as face-to-face training by enhancing artificial intelligence. The global corona epidemic has forced students to rely on technology in unprecedented ways to teach themselves. These technologies can be significantly improved with artificial intelligence and machine learning algorithms. Virtual educators can help students find the right curriculum from the available curriculum through the referral system [1,2,3].

Internet services today offer various content. This high diversity is both an opportunity and a threat. The opportunity is to present the customer's favorite content with more probability, and the customer can use it. The danger is that the customer may not find the content they need in the vast amount of content. It can be concluded that the rapid increase of the information volume, the limitation of search engines to search for information, and the increasing number of visitors to websites in recent years, are critical challenges in the recommender systems. A Student must find the required Educational Resources from the appropriate sources, which would be time-consuming and costly without information filtering and recommender systems [4].

Recommender systems can discover the users' interests and predict their preferences, and among the high volume of data, refine the items which are likely to be of interest to the user and save time by suggesting them. Of course, the only efficiency of a recommendation system (which is done by search engines) is not only searching the items in less time with less energy, but its primary purpose is to discover items. These systems, with the ability to store and analyze the user's past behaviors, also infer services and information that users have not noticed but are probably

✉ **Seyyed Javad Seyyed Mahdavi**
Mahdavi@Mashdiau.ac.ir

interested in and provide exciting results to users. Recommender systems are one of the main tools to overcome information overload and are an intelligent complement to the retrieval concepts and information filtering to analyze the users' behaviors.
In addition, students have individual differences such as educational background, study method, age, etc., which emphasizes the need to take feedback from students to guide them in the educational process better [5]. In e-learning systems, students are eager for personalized services to be automatically trained, monitored, supported, and evaluated. With such a personalized service, student loyalty increases [6].
The Recommender system can help teachers personalize the curriculum and resources for each student based on their unique skills and weaknesses. It is used to develop more complex and attractive methods of student assessment that are very time-consuming for teachers today. Artificial intelligence and machine learning have the potential to address many of the problems that have arisen in the transfer of teaching methods to online learning. Including students' resistance to changing their education, increasing curriculum planning, and addressing the loss of personal interaction between students and teachers [4].
This paper aims to design a time-based educational recommender system to suggest new resources to users based on the features that include a person's pre-clicked or downloaded educational resources and the rating that the user has given to each resource. If there is an inherent structure that the model can exploit, deep neural networks are very efficient for this issue.
Because the nature of the problem of recommending textbooks depends on the time and long-term review of student performance, the sequential structure of sessions or report clicks is very appropriate for inferential errors in conventional or recursive models. In many methods, only the user's past information is used in learning. While in the present article, having a network that looks both backward and forward can also cover changes in learner behavior and offer more up-to-date recommendations. In the following, we will review the work done on educational recommenders. In the next section, we present the proposed method, a hybrid architecture of BI-LSTM and MLP networks. In the fourth section, we review the results of implementing the proposed algorithm based on accuracy and efficiency, and finally, in the fifth section, we will present future suggestions.

2- Related Works

Session-based recommender systems are an excellent example of recommender systems. They are primarily researched, although not a new research topic [7, 8, 9]. Compared to traditional recommender systems, a session-based referral system is more suitable for learning dynamic and sequential user behaviors.
The purpose of recommender systems is to generate search results close to the user's needs and make predictions based on their priorities. In virtual learning environments, educational recommender systems have learning objects based on students' characteristics, priorities, and learning needs. A Learning Object (LO) is a unit of educational content that can assist students in their learning process [10]. A learning object is defined by the IEEE [11] as a digital or non-digital entity with educational design features that can be reused or referenced during the computer-assisted learning process.
Recently, deep learning has dramatically changed the architecture of recommenders and provided more opportunities to improve the performance of recommenders. Deep learning can capture nonlinear and meaningful user-item relationships and result in abstract data representations at higher levels. Besides, it can obtain complex relationships within data from other sources such as conceptual, textual, and visual information [12]. In the traditional technology method, the recommender system faces problems such as a large amount of data, shared filtering, poorly given information, and a cold start, which are also addressed in this study.
A new generation of algorithms is required for recommender systems due to the importance of recommender systems in online servers and the dynamics, privacy, and bulk of data and problems in these systems (such as cold start). Deep learning is offered as a solution to the problem of recommender systems.
In recent years, artificial neural networks have attracted much attention due to their increased computing power and big data storage capabilities. Many new methods in image processing, object recognition, natural language processing, and voice recognition now use deep neural networks as a primary tool [13]. The remarkable capabilities of deep learning methods encourage researchers to apply deep architectures in "recommendation". Deep learning-based recommender models can be categorized in Fig1. [14].

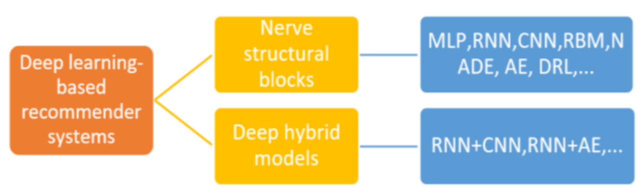


Fig. 1. General Models of Deep learning for the Recommender.

In the following, the work that has been done specifically on educational advisors will be reviewed.

2-1- Data Mining Methods

In [15], the proposed method integrates the features of online learning style, including participatory filtering (CF) features, association rules criteria, and online learning style (OLS) in the recommendation algorithm. The output of the proposed method has improved by 25% compared to the technique without students' characteristics.

In [16], attention-focused neural networks (CNN) have been used to obtain predictions of user ratings and user profiles and to recommend superior courses. Then, they integrated a participatory filter to enable real-time recommendations and reduce server workload. The model ultimately recommends courses to students. However, the proposed system may continue to suffer from the problem of recommending similar courses as MOOCs develop and the number of courses increases.

2-2- Development on Traditional Methods

Most traditional online learning systems based on refining methods depend on user's behavior towards different sources. For users who behave similarly, the results of resource recommendations are often unsatisfactory [17].

The aim is to help students make informed decisions about their learning paths using a hybrid counseling system. By combining content-based similarity and dispersion, based on structural information about module space, the detectability of long-term choices that are consistent with students' preferences and goals can be improved.

One of the advantages of this is that you can add scatter to the set of recommendations. The goal of [15] is to provide a personal reference system that leads to better recommendations in the shortest possible time. The proposed system uses user profiles to create neighborhoods and predicts weights. To overcome the problem of cold start and scattered data, student profiles are created using the learning method. Resources that are of interest to the user are suggested through calculations calculated with new features and participatory refinement method.

2-3- Machine learning-based Methods

In [18], the aim is to provide an advisory system based on reasoning theory that combines content-based, participatory, and knowledge-based recommendation methods. This method recommends training resources so that the system can generate further arguments to justify its competence.

In [19], the AROLS method is proposed. This method is an advanced recommendation integrated with a comprehensive learning style model for online students. This method considers the learning method as prior knowledge and provides recommendations. First, it creates clusters of different learning styles. Then the behavioral patterns presented by the matrix of similarity of learning resources and communication rules of each group are extracted using students' review history. Finally, it creates a set of personal recommendations based on the data mining results of the previous steps. This method presents the recommendation results more accurately while maintaining the computational advantage than the traditional participatory refining (CF) recommendation.

2-4- Artificial Intelligence-based Methods

[20] A multi-factor technology-based referral system has been developed that helps e-learning referral systems offer students the most appropriate learning resources. This work utilizes the capabilities of multi-agent technology to create a plan that combines web use and extraction algorithms such as content-based methods and collaborative refinement to find the most appropriate training resources. The performance of this combined method is better than other algorithms in it. Advances have also been made in building models for searching and retrieving learning objects stored in heterogeneous repositories.

In [21], the aggregation of two multifactorial models is introduced that can carry a specific LO corresponding to the characteristics of a student and carry the LO to the instructors to help them in creating lessons. The aim is to create an integrated multi-factor model that meets the needs of students and educators and thus improves the learning and teaching process.

2-5- Methods based on Neural Networks and Deep learning Networks

The paper [11] introduces a high-precision resource recommender model (MOOCRC) based on deep belief networks (DBNs) to increase the efficiency and enthusiasm of learners in MOOC environments. This model extracts the characteristics of learners and their curriculum content. User-lesson vector vectors are constructed as model input. Instructors' grades are processed into lessons as supervised labels. The MOOCRC model is taught without supervisor pre-training and is fine-tuned using supervisor feedback. The model's performance has been evaluated using selective data from educators obtained from the starC MOOC platform of Central China Normal University. The results show that MOOCRC has higher recommendation accuracy and faster convergence than other traditional recommending methods.

The article [22], with the aim of recommending educational resources, tries to help learners achieve better academic results. The proposed model consists of deep learning recursive layers that have been improved with the attention technique. After testing the model's performance with OULAD data, 95% accuracy was obtained, which is a better result than similar works.

3- The Proposed Method

In this paper, the purpose of the first phase is to obtain the users' database, including their interest in the study resources and the amount of use and click on these resources and related features, and then select the practical items among them. In the second phase, the recommender system is trained with acceptable accuracy using deep neural networks. The recommendation algorithm includes data extraction from OULAD information resource files, data preprocessing, building an add-on deep learning network from MLP, Bi-LSTM, initial parameterization, training, and predicting scores. Finally, it is offered resources to users using the trained network.

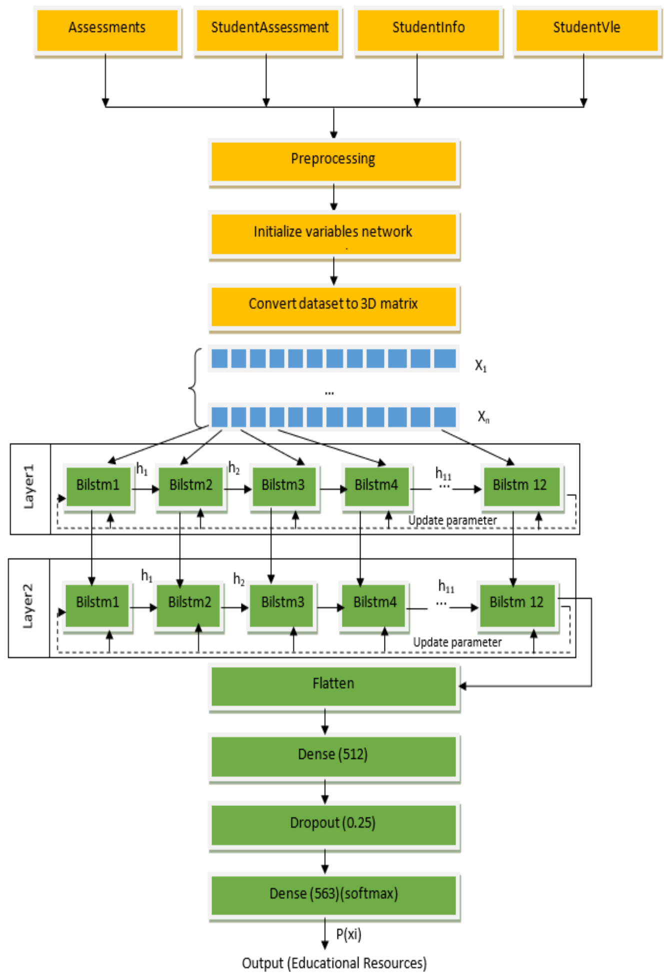
Two-way short-term memory (Bi-LSTM) is a type of recurrent neural network. This process processes data in two directions because it works with two hidden layers. It is the main point of divergence with LSTM. This method has proven promising results in natural language processing. One advantage of two-way LSTM over one-way LSTM is that two-way LSTM looks to both the past and the future to make predictions. Still, one-way LSTM only looks to the past, so two-way LSTM can be sensitive to diversity in the short-term interests of the user in both directions and thus cover learners' behavioral changes.

In the proposed architecture, as presented in Fig.2, a Bi-LSTM cell is provided in each layer for each feature in the database. The cells focus on one feature of each record, and each cell represents only one feature pattern. Consider and ultimately, the combination of these patterns will lead to better results.

Among the papers in this field, we found the works based on Bi-Lstm, which required short-term and long-term memory; or for new mechanisms that, in practice, create shortcuts and ignore several time steps. These shortcuts also allow the production error to be easily transferred to the post-diffusion phase without quickly losing. It can significantly handle the vanishing gradient problems.

3-1- Predicting Course Resource Scores

In the model training process, the data labeled class is used as a training set for the model. Then, based on the user-class feature vector, the recommendation problem becomes the category prediction problem. In this paper, using the label of rating classes, error information is published to each layer from top to bottom with fine-tuning of the parameters to the observer. After training the model to achieve a certain error, the test set can be used to test the performance of the recommender model. The data in the test set is divided into two categories: user feature -lesson vector and lesson evaluation. Each user-lesson feature vector corresponds to a category level, and each level corresponds to a point. All lessons that correspond to a user are sorted according to the expected score, and then



lesson recommendations are generated

Fig.2 The Structure of Proposed Methods.

3-2- Implementation

The selected data as input to the database implementation are divided into three sections: students, teacher, and course and includes information about 22 courses offered, 32593 students, their evaluation results, and their mutual reports with the virtual learning environment (VLE), which is provided by summaries of students' daily clicks on various "resources" (10,655,280 entries).

- **Database[1]**

Students generate various behavioral data by learning in an online learning environment. This behavioral data is collected and stored through data collection methods (OULAD). The reference database provides data sources for this platform. This curriculum database can extract content features that reflect students' interest in reference. Students' feature vectors are constructed by combining students' characteristics and lesson features, and then hybrid behavioral characteristics and user-lesson feature vectors are generated (Jacob Kozilk et al., 2007).
The frequency of each student's recorded activities is presented in Fig3. The database is anonymous using the ARX [PK15] data encryption tool. The data was reviewed for error detection and verified and published by Open Data Institute1. The frequency of each student's recorded activities is presented in Fig3. The database is anonymously using the ARX [PK15] data encryption tool. The data were reviewed for error detection and verified and published by Open Data Institute1.
The main table contains the student files attached to the courses (A student can have more than a one registered course). Each course has several assessments, which are related to students and include the history of student assessment results. There are three types of assessment: Teacher Assessment (TMA), Computer Assessment (CMA), and Final Exam (Exam).

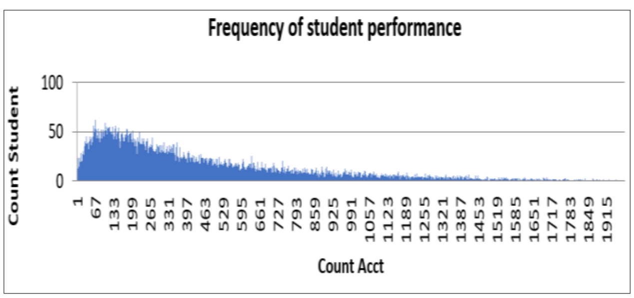


Fig.3 Frequency of Student Performance.

- **Data Preprocessing**

As presented in Fig4. In the pre-processing stage, the input first includes four sections: resources provided, students' characteristics, courses held, and student performance and assessment history in each course. These four sections are combined, and further analysis, categorization, feature mapping, clean blank or incorrect data, and feature normalization are performed on them.
The normalization of features is done in the range (0, 1).
The data must be placed at an equal distance so that the data that contains a larger range of numbers does not have a more significant effect on the algorithm than the others. This prevents network weights from fluctuating too much, and the amount of network loss fluctuates slightly when modeling data. In other words, the higher the convergence speed, the better and smoother the network model [23].

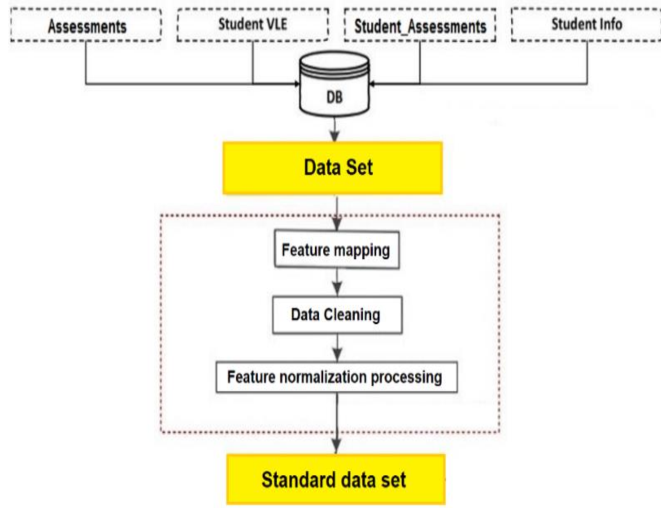


Fig.4 Data Preprocessing Steps.

Empirical evidence shows that data standardization is useful in terms of accuracy. It may be related to the descent of the gradient. It is easy to understand why normalization improves training time. Large input values saturate activation functions such as sigmoid or ReLu (negative input). This type of feedback activation function has little or no gradient in the saturated region and thus reduces the Training speed [23]. Eq1 is used to perform normalization.

$$X^* = \frac{(X - Xmin)}{(Xmax - Xmin)} \quad (1)$$

Where X min represents the lowest eigenvalue as X min {X1, X2, ..., Xn}= Xmin.Xmax represents the maximum eigenvalue as X max = max {X1, X2, ..., Xn}; X*indicates the normalized value, X represents the original data. Another step in the preprocessing step is to convert the string values in the database to numeric values.
Table1 shows an example of the values in the database that are mapped to numerical values in Table2.
After the initial steps of preprocessing, the datasheet is tagged in two following ways:

- **Maximum Click in Maximum Point Average:**

Among the set of activities registered for the standard courses for each student, the source with the most clicks can indicate the student's taste and interest) in the course that had the highest GPA (showing the practical resources studied in that course) was selected as a label. Five hundred sixty-two labels were created, mapped from 0 to 561. The frequency of count labels is presented in Fig5.

[1] [1] https://analyse.kmi.open.ac.uk/open_dataset

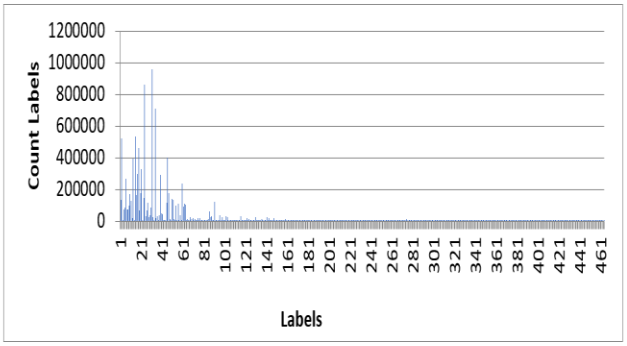


Fig. 5. Frequency of Count labels (Method 1: The Most Click in the Maximum Average Score

- **Most Recent Clicks:**

From the set of each student's activities, the source with the most clicks (indicating the student's taste and interest) were selected as the label on the last day of the student's activity, assuming that the current study subject is essential to him. One thousand five hundred ninety-four labels were created, mapped from 0 to 1593. The frequency of count labels is presented in Fig6. After preprocessing, the data is divided into a training set and a test set. Finally, the training set enters the proposed network as input.

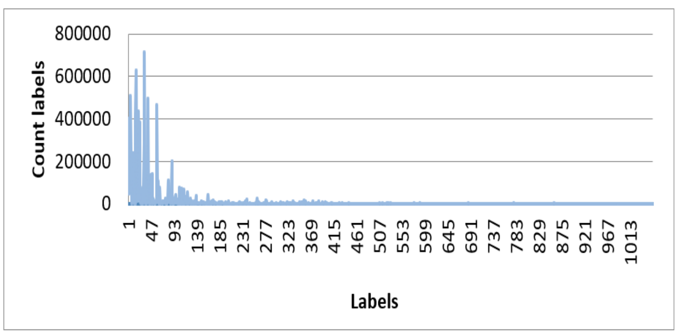


Fig. 6. Frequency of Count labels (Method 2: The Most Click)

Table 1. A Sample of Database Data before Mapping

Code module	Code presentation	Id student	gender	Age band	Sum click
AAA	2013J	11391	M	55<=	16
AAA	2013J	11391	M	55<=	44
AAA	2013J	11391	M	55<=	1
AAA	2013J	11391	M	55<=	2
AAA	2013J	11391	M	55<=	1
AAA	2013J	11391	M	55<=	2
AAA	2013J	11391	M	55<=	2
AAA	2013J	11391	M	55<=	16
AAA	2013J	11391	M	55<=	44

Highest education	final_result	score_ mean	id_site	date
HE Qualification	Pass	82	546669	-5
HE Qualification	Pass	82	546662	-5
HE Qualification	Pass	82	546652	-5
HE Qualification	Pass	82	546668	-5
HE Qualification	Pass	82	546652	-5
HE Qualification	Pass	82	546670	-7
HE Qualification	Pass	82	546671	-7
HE Qualification	Pass	82	546669	-5
HE Qualification	Pass	82	546662	-5

Table 2. The Values of Features Mapped to the Number

Code module		Code presentation		age_band		gender		highest_education		final_result	
AAA	0.1	2013B	540	0-35	0.1	F	0.1	A Level or Equivalent	0.1	Distinction	0.1
BBB	0.2	2013J	720	35-55	0.2	M	0.2	HE Qualification	0.2	Fail	0.2
CCC	0.3	2014B	180	55<=	0.3	-	-	Lower Than A Level	0.3	Pass	0.3
DDD	0.4	2014J	360	-	-	-	-	No Formal quals	0.4	Withdrawn	0.4
EEE	0.5	-	-	-	-	-	-	Post Graduate Qualification	0.5	-	-
F F F	0.6	-	-	-	-	-	-	-	-	-	-
GGG	0.7	-	-	-	-	-	-	-		-	-

- **Correlation Between Variables and Labels**

In this research, the method of the maximum average method has been used for labeling. We examined the possible correlation between the label and the existing variables that were used as input for teaching the model by correlation test. As you can see in Fig7, there is no significant correlation between the variables and their labels.

- **Network Construction**

The Bi-Lstm library of KERAS has been used to implement the idea of this paper. The data is entered into a

two-layer Bi-Lstm architecture with 512 neurons, and finally, the output of this layer enters the MLP network.
In the model training process, the model parameters must repeatedly be adjusted to achieve better results in feature extraction. During the learning process, the 512 minibatch processing method is used to solve the problem of large data volumes. Also, the learning rate parameter 0.0001, the number of counts Epoch = 100, and the SoftMax activator function are initialized.

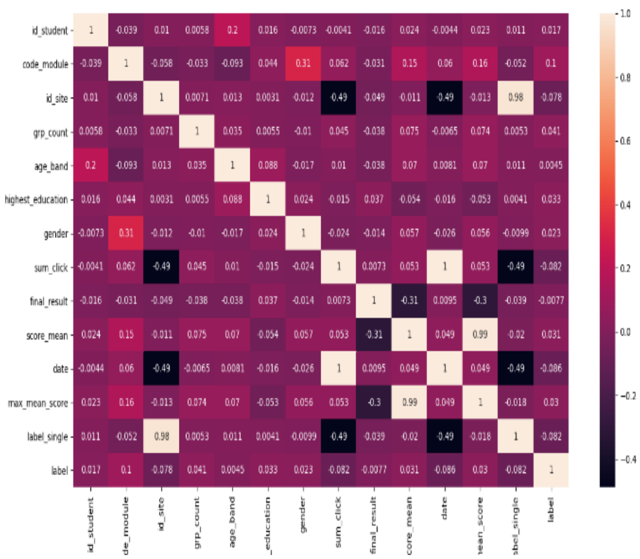


Fig7. The Result of the Correlation Test.

The number of records used in the testing process of the networks implemented in this article is Train = 8434945, test = 2108737, and validation split = 0.2. After completing Epochs, the diagrams and the results of implementations show that the accuracy and loss of the work are far better than the results of the implementation of the proposed network [24].

3-3- Methods and Tools of Data Analysis

The primary purpose of the proposed system presented in this article is to predict the best sequence of educational resources. There are many criteria for measuring different aspects of bid performance.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Indicates what percentage of experimental records are properly categorized.

$$Precision = \frac{\sum_{x\in X}|R(x)\cap H(x)|}{\sum_{x\in X}|R(x)|} \quad (3)$$

$$Recall = \frac{\sum_{x\in X}|R(x)\cap H(x)|}{\sum_{x\in X}|H(x)|} \quad (4)$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \quad (5)$$

Here x is a student from the set of all students X, R (x) represents the learning resources recommended for student x, H (x) represents the learning resources observed by learner x [25].

4- Ablation Study

4-1- Investigating the Effect of the Number of Cells in Each Layer

As observed in Table 3, the results of the implementation of three types of LSTM, GRU, and Bi-LSTM networks were implemented and examined as single-celled structures in three single-layer architectures, two layers, and three layers are not desirable. In nine architectures, one cell in each layer could not find the pattern of different features, the relationship of features to other features in combination, permutation, and different models. Increasing the number of layers has not been able to play an influential role in improving the results.
As observed in Table4, The implementation of multicellular single-layer architectures implemented and investigated in three single-layer architectures of LSTM, Bi-LSTM and GRU had more favorable results than the single-cell architectures.
In fact, by changing the layout cells in the proposed models, it is possible to use and extract features in both single and multiple forms. When entering the first feature of the model, it examines and extracts the information contained in the same feature individually. After entering the second feature of the model and extracting the information contained in this feature, it also examines and extracts the connections and information between these two features. With the introduction of the third feature, in addition to extracting the information of the same feature individually and examining the existing communications and information with the previous features in pairs, the communications are also examined on a permutation basis. As a result, the model achieves more features and is more important than single-cell networks.
Meanwhile, the GRU network with higher generalizability power than the two types of Bi-LSTM and LSTM networks with a loss of 0.2 and an accuracy of 0.91 has had a better performance.

Table 3. Results of Training and Testing of one to Three-layer Single-Cell Networks: LSTM, Bi-Lstm and GRU

LSTM	1Layer	train	loss	val_loss	accuracy	val_ acc
			2.6907	2.6617	0.2825	0.2879
		test	accuracy			
			0.28			
	2Layer	train	loss	val_loss	accuracy	val_ acc
			2.8305	2.8172	0.2607	0.2577
		test	accuracy			
			0.18			
	3Layer	train	loss	val_loss	accuracy	val_ acc
			2.8633	2.8369	0.266	0.2548
		test	accuracy			
			0.25			
GRU	1Layer	train	loss	val_loss	accuracy	val_ acc
			2.9989	2.9215	0.2389	0.2539
		test	accuracy			
			0.25			
	2Layer	train	loss	val_loss	accuracy	val_ acc
			2.9429	2.9185	0.2439	0.2562
		test	accuracy			
			0.25			
	3Layer	train	loss	val_loss	accuracy	val_ acc
			2.9811	2.9577	0.2317	0.2388
		test	accuracy			
			0.23			
Bi-Lstm	1Layer	train	loss	val_loss	accuracy	val_ acc
			1.9011	1.3059	0.4481	0.6298
		test	accuracy			
			0.62			
	2Layer	train	loss	val_loss	accuracy	val_ acc
			2.4977	2.4328	0.3133	0.3231
		test	accuracy			
			0.31			
	3Layer	train	loss	val_loss	accuracy	val_ acc
			2.5016	2.466	0.3057	0.3132
		test	Accuracy			
			0.31			

Table 4. Results of Training and Testing Single layer Multi-Cellular Networks of LSTM, GRU and BI-LSTM

LSTM	train	loss	val_loss	accuracy	val_ acc
		0.6685	0.6067	0.7573	0.7704
	test	accuracy			
		0.77			
Bi-Lstm	train	loss	val_loss	accuracy	val_ acc
		0.3268	0.2366	0.8757	0.9036
	test	accuracy			
		0.9			
GRU	train	loss	val_loss	accuracy	val_ acc
		0.2319	0.2021	0.898	0.9099
	test	accuracy			
		0.91			

4-2- Investigating the Effect of the Number of Layers of Network Architecture

In the architecture of deep learning networks, one issue is the relationship between the cells in each layer and themselves, which was studied in detail. The second issue will be the relationship between the different layers in the implemented architecture. This part of the architecture does an overview of the features in the database. Some previous work has suggested that multilayer Bi-Lstm in neural networks can further improve classification or regression performance [26]. In addition, some related theoretical supports have shown that a deep hierarchical model is more efficient in delivering some functions than the shallow type.

It has been implemented and trained three architectures of two-layer LSTM (Fig8, a and b), two-layer GRU (Fig8, c and d), and two-layer BI-LSTM (Fig8, e and f) to evaluate the effectiveness of the relationship between the layers. As observed in Table5, the results of two-layer architectures have been more favorable compared to the single-layer architectures.

According to Table5, the proposed architecture results, based on BI-LSTM bilayer and multi-cellular, with a loss of 0.9 and accuracy of 0.95, were much more accurate and desirable than the proposed architecture [25- 28]. As observed in Fig8, the Loss reduction speed indicates that the number of selected AIPs is appropriate, and since the accuracy and validation accuracy diagrams are almost the same, over-fitting did not occur in this experiment.

As can be seen in Table 7, the results of our proposed architecture are very acceptable and desirable.

According to Figure 8, considering the loss rate, it can be seen that the number of selected epics is appropriate. Besides, according to the accuracy diagram, it can be seen that as the accuracy and validation accuracy diagrams are almost the same, overfitting does not occur in this experiment.

Table 5. Results of Training and Testing of Dual-layer Multi-Cell Networks

LSTM 2 Layer	train	loss	val_loss	accuracy	val_ acc
		0.2207	0.1836	0.9021	0.9169
	test	accuracy			
		0.91			
GRU 2 Layer	train	loss	val_loss	accuracy	val_ acc
		0.2321	0.2063	0.9013	0.9121
	test	accuracy			
		0.92			
Bi-Lstm2 Layer	train	loss	val_loss	accuracy	val_ acc
		0.1171	0.0967	0.9529	0.9539
	test	accuracy			
		0.95			

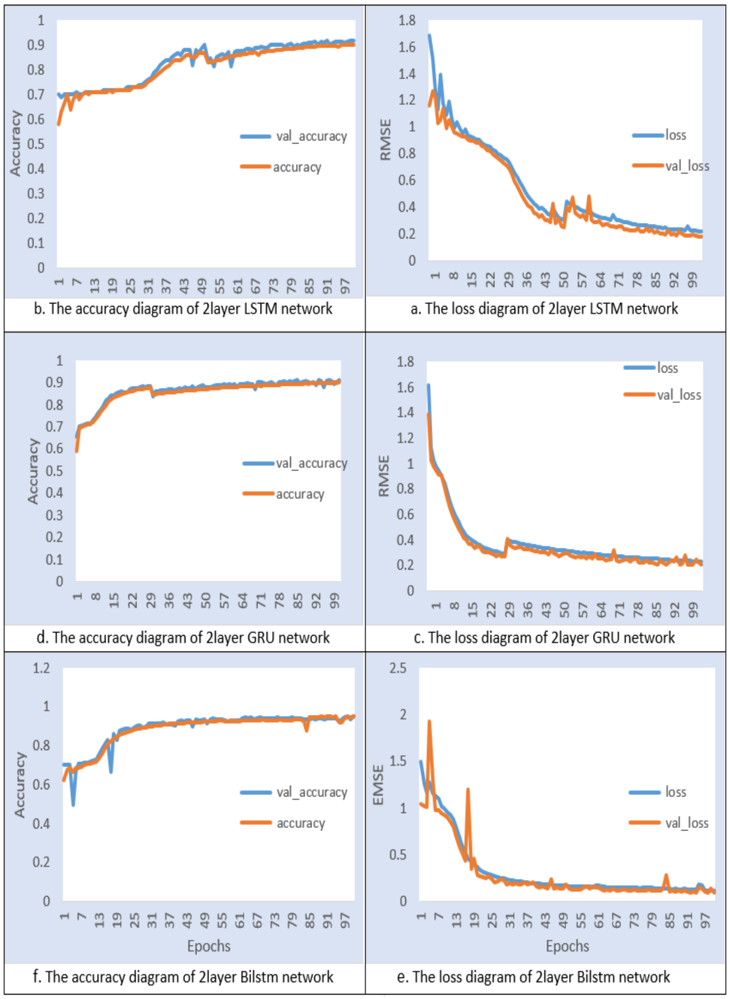


Fig. 8. Results of Training and Testing of Dual-layer Multi-Cell Networks

5- Result and Discussion

In the proposed network and the training phase, the highest Validation accuracy was 0.9529 in epoch 47, and the minimum loss was 0.0995 in epoch 45. After completing the training step, it is entered the test data was input into the network, and the final result was 0.95.

Due to the multi cellularity of the layers, the proposed network allows the study and extraction of features in both single and multiple forms. When entering the first feature of the model, it examines and extracts the information contained in the same feature individually. After entering the next features, the available communications and information with the features are checked in pairs, and the communications are also examined on a permutation basis. On the other hand, the multi-layering of multicellular structures will create a connection between the different layers in the implemented architecture. An overview of the features in the database is done by this part of the architecture.

In addition, the model consisting of Bi-LSTM layers has been more successful due to its two-way structure and the extraction of the relationship between past and future features. As a result, the model acquires more features and is more important than single-cell and single-layer networks.

5-1- Investigating the Effect of Unconventional Data on Model Accuracy

As shown in the analysis of the existing database data and the frequency charts of student activities, Figure 3 and the frequency of classes, Figures 5 and 6, we are faced with an unbalanced data set. In this case, in addition to accuracy, it is better to use call parameters and F1-score to evaluate the model. To make sure that the model does not intelligently categorize all the data presented in an iterative class in the training process to achieve high accuracy.

We have used the content of the database in 3 different sections to teach and test the model.

At the end of the test phase, the average value for all three f1-score reminder parameters is 0.95. We examined the classification accuracy of each group separately and found that the data of all classes were well categorized. For example, Table 6 presents the results of ten groups of groups with the lowest frequency of repetition and ten groups of groups with the highest frequency of repetition.

As you see, in all three sections, 30-70, 20-80, and 10-90 as a group with a support value of one, The result of most recall and f1-score is one. On the other hand, the value one for recall and f1 points is too low in the groups with the most members. This shows that our model, in addition to the data volume challenge, has also responded well to the unconventional data challenge.

5-2- Comparison of the Performance of the Proposed Model with other Models

We have compared the result of the proposed model in line 1 of Table 7 with other methods presented in previous studies or implemented by ourselves. As can be seen, the results are more favorable for different evaluation parameters of the proposed model than other implemented methods. All evaluations were performed on OULAD shared data.

The proposed method [16] has been implemented and has been trained, tested, and evaluated with OULAD data. As can be seen in Table 7, it performed worse than our proposed model in terms of both error and accuracy criteria.

In [25], the three criteria Recall, Prec, and F1 for the 3 methods itemCF, Clustering + itemCF, and AROLS are examined and show that their proposed algorithm

Table 6. Values of Mapping Properties are Given in Numbers

Train test split(test-size=0.1)					Train test split(test-size=0.2)					Train test split(test-size=0.3)					
support	**f1-score**	**recall**	**precision**	**label**	**support**	**f1-score**	**recall**	**precision**	**label**	**support**	**f1-score**	**recall**	**precision**	**label**	
1	0.4	1	0.25	530	1	0.4	1	0.25	209	1	1	1	1	326	10 groups with the lowest frequency
1	1	1	1	498	1	1	1	1	228	1	1	1	1	418	
1	1	1	1	453	1	1	1	1	547	1	1	1	1	463	
2	0.02	0.5	0.01	477	1	0.4	1	0.25	228	1	0.4	1	0.25	391	
2	1	1	1	305	1	1	1	1	90	2	1	1	1	446	
2	1	1	1	460	2	1	1	1	184	2	0.4	1	0.25	442	
3	1	1	1	213	3	1	1	1	133	2	0.01	0.5	0.01	171	
3	1	1	1	503	3	1	1	1	391	3	1	1	1	408	
3	0.02	0.5	0.01	411	3	1	1	1	326	3	1	1	1	547	
3	1	1	1	103	4	1	1	1	557	3	1	1	1	550	
82212	0.96	0.96	0.96	8	82212	0.96	0.96	0.97	1	86317	0.95	0.95	0.96	16	10 groups with the highest frequency
84858	0.99	0.99	0.99	34	84858	0.98	0.98	0.98	57	94754	0.98	0.98	0.99	20	
99036	0.99	1	0.98	5	99036	0.98	0.99	0.98	60	112033	0.98	0.99	0.98	12	
101120	0.98	0.99	0.98	28	101120	0.96	0.96	0.97	5	112418	0.99	1	0.99	46	
111173	0.96	0.95	0.98	56	111173	0.97	0.97	0.98	11	131070	0.97	0.97	0.98	17	
137559	0.98	0.99	0.98	13	137559	0.98	0.99	0.98	18	148366	0.98	0.99	0.98	1	
165625	0.97	0.98	0.97	33	165625	0.98	0.98	0.99	16	154820	0.98	0.98	0.99	14	
198928	0.98	0.98	0.99	49	198928	0.98	0.98	0.99	19	202925	0.98	0.98	0.98	34	
262277	0.97	0.97	0.97	56	262277	0.96	0.96	0.97	26	249276	0.97	0.96	0.97	23	
265671	0.98	0.98	0.98	12	265671	0.95	0.95	0.96	41	275379	0.98	0.97	0.98	30	

(AROLS) has a better Prec compared to the other two cases. At the same time, F1 and Recall remain relatively constant during the n top recommendation.

The report in [26] shows that the performance of AROLS is much better than traditional participatory filtering. In particular, the User-AROLS call and accuracy has more than tripled, and the UserCF call and accuracy are much lower than ItemCF, Probably because UserCF focuses more on the interest of learners who are more like a particular learner. ItemCF's recommendation, on the other hand, is more personal because it largely suggests similar ones based on the learner's interest. As you can see in the first row of results, our proposed model performed better than all 7 methods reviewed in these two papers.

In [27] the results show that OLS characters can make the recommendation algorithm more accurate and robust, but as you can see in the results of row one, our proposed model performed better than both methods studied.

Table 7. Comparison of the Proposed Method with the Results Obtained from other Implementations Performed by us and Studies [16,25- 27]

Model	**Accuracy**	**Recall**	**Prec.**	**F1**	**ref**
Our proposed model	0.9529				-
Naive Bayes(nb)	0.1981	0.4725	0.2937	0.1853	
Logistic Regression(Lr)	0	0	0	0	
Latent Dirichlet Allocation(lda)	0.5415	0.0772	0.4314	0.4579	
DBN	0.2912	-	-	-	**16**
AROLS	-	0.022	0.28	0.04	**25**
itemCF	-	0.018	0.18	0.027	
Clustering + itemCF	-	0.024	0.24	0.041	
ItemCF	-	0.026	0.1334	0.0435	**26**
Item-AROLS	-	0.0406	0.1880	0.0668	
User-AROLS	-	0.0018	0.0046	0.0026	
UserCF	-	0.0005	0.0011	0.0007	
CF with ARM	-	0.6874	0.076	0.1374	**27**
Proposed article method	-	0.8647	0.1033	0.1842	

6- Conclusion

Recommendation of Educational Resources is an essential and challenging task, especially in a course with the rapid development of the Internet, which consists of a massive variety of educational resources. The challenge is due to the massive amount of educational information in almost all academic fields and the inevitable neglect of personal needs for specific knowledge. Therefore, research on timely learning of learners' behaviors and then personal guidance of their learning process becomes more necessary. This study has analyzed the online learning behaviors to improve personal recommendations in Educational Resources. It is necessary to use different sources and design a centralized framework to combine them and thus provide superior recommendations.

Informal learning environments will also focus on teacher support. Systems must be able to participate in the teachers' tasks, especially when the continuous monitoring and assessment of student homework during the semester is needed [39].

References

[1] Konstan, J. A., |Introduction to recommender systems. In Proceedings of the 2008 ACM SIGMOD international Conference on Management of Data, Vancouver, Canada, (Jun, 2008).

[2] Resnick, P. and Varian, H. R. 1997. Recommender systems, Commun. ACM 40, 3 (Mar, 1997).

[3] Schafer, J. B., Konstan, J., and Riedi, J. Recommender systems in e-commerce. In Proceedings of the 1st ACM Conference on Electronic Commerce, Denver, Colorado, United States, (Nov, 1999).

[4] Gordan Durovic, Martina Holenko Dlab and Natasa Hoic-Bozic, Educational Recommender Systems: An Overview and Guidelines for Further Research and Development, Croatian Journal of Education Vol.20; No.2, 2018, 531-560.

[5] Paula Rodríguez et all, An educational recommender system based on argumentation theory, AI Communications 30 (2017), 19–36.

[6] learning Technology Standards Committee, IEEE Standard for Learning Object Metadata. Institute of Electrical and Electronics Engineers, New York (2002).

[7] Zhang, S.; Yao, L.; Sun, A. Deep learning based recommender system: A survey and new perspectives. arXiv 2017, arXiv:1707.07435.

[8] Ludewig,M.; Jannach, D. Evaluation of Session-based Recommendation Algorithms. arXiv 2018, arXiv:1803.09587.

[9] Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; Tikk, D. Session-based Recommendations with Recurrent Neural Networks. In Proceedings of the International Conference on Learning Representations, San Juan, Puerto Rico, 2-4 May 2016; pp. 1-10.

[10] Paula Rodriguez et all, An educational recommender system based on argumentation theory, AI Communications 30 (2017), 19-36.

[11] Zhang, H., et al., MOOCRC: A highly accurate resource recommendation model for use in MOOC environments. Mobile Networks and Applications, 2019. 24(1): p. 34-46.

[12] Shuai Zhang, et. all, \Deep Learning based Recommender System: A Survey and New Perspectives", ACM Computing Surveys, Vol. 1, No. 1,2018.

[13] Zeynep Batmaz, Ali Yurekli, Alper Bilge and Cihan Kaleli, A review on deep learning for recommender systems: challenges and remedies, Springer Nature B.V. 2018.

[14] SHUAI ZHANG,LINA YAO, AIXIN SUN and YI TAY, Nanyang Technological University Deep Learning based Recommender System: A Survey and New Perspectives, 2018, ACM Computing Surveys, Vol. 1, No. 1, Article 1.

[15] Chen, H., et al., Enhanced learning resource recommendation based on online learning style model. Tsinghua Science and Technology, 2019. 25(3): p. 348-356.

[16] Li, R., et al., Online learning style modeling for course recommendation, in Recent Developments in Intelligent Computing, Communication and Devices. 2019, Springer. p. 1035-1042.

[17] Hagemann, N., M.P. O'Mahony, and B. Smyth. Visualising module dependencies in academic recommendations. in Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion. 2019.

[18] Rodríguez, P., et al., An educational recommender system based on argumentation theory. AI Communications, 2017. 30(1): p. 19-36.

[19] Dawen Liang, Rahul G Krishnan, Mathew D Ho man, and Tony Jebara. 2018. Variational Autoencoders for Collaborative Filtering. arXiv preprint arXiv:1802.05814 (2018).

[20] Neto, J. Multi-agent web recommender system for online educational environments. in International Conference on Practical Applications of Agents and Multi-Agent Systems. 2017. Springer.

[21] Rodríguez, P., N. Duque, and S. Rodríguez, Integral Multi-agent Model Recommendation of Learning Objects, for Students and Teachers, in Management Intelligent Systems. 2013, Springer. p. 127-134.

[22] Ahmadian Yazdi, H., S.J. Seyyed Mahdavi Chabok, and M. Kheirabadi, Dynamic Educational Recommender System Based on Improved Recurrent Neural Networks Using Attention Technique. Applied Artificial Intelligence, 2021: p. 1-24.

[23] Shanker, M., M.Y. Hu, and M.S. Hung, Effect of data standardization on neural network training. Omega, 1996. 24(4): p. 385-397.

[24] Zhang, H., Huang, T., Zhihan, Lv., Liu, S., and Yang, H. MOOCRC: A Highly Accurate Resource Recommen- dation Model for Use in MOOC Environments. Springer, Mobile Networks and Applications, Springer ,2018.

[25]Charnelli, M.E., Sistemas recomendadores aplicados en Educación. 2019, Universidad Nacional de La Plata.

[26]Li, R., et al., Online learning style modeling for course recommendation, in Recent Developments in Intelligent Computing, Communication and Devices. 2019, Springer. p. 1035-1042.

[27]Yan, L., et al. Learning Resource Recommendation in E-Learning Systems Based on Online Learning Style. in International Conference on Knowledge Science, Engineering and Management. 2021. Springer.
[28] P. Resnick and H.R. Varian. Recommender systems. Communications of the ACM, 40(3):56.58, (1997).