Design of Multimedia Learning Resource Recommendation System Based on Recurrent Neural Network

Zijin Xiao(✉), Ying Li, and Hai Zhou

College of Science, Engineering, Agriculture and Medicine, Gansu Radio and Television University, Lanzhou 730030, China
fddsf222@aliyun.com

Abstract. The existing learning resource recommendation system has the defect that the average absolute error of the recommendation result is large due to the limitation of its own adaptation range. For this reason, this research designed a multimedia learning resource recommendation system based on recurrent neural network. Introduce the recurrent neural network to design the architecture of the multimedia learning resource recommendation system, and design the system functional modules based on this, including the learner's demand retrieval representation module, learner preference representation module, recurrent neural network training module, database module and system management module. The simulation experiment results show that compared with the existing system, under the Gowalla data set, the average absolute error coefficient of the recommended results of this paper is reduced by 0.356; under the Yelp data set, the average absolute error coefficient of the recommended results of this paper is reduced 0.404. The above results fully show that the recommendation effect of this system is better.

Keywords: Recurrent neural network · Multimedia learning resources · Resource recommendation · Demand retrieval · Learner preference

1 Introduction

The booming development of the Internet not only brings a variety of fast and convenient to people's life, study, work, etc., but also makes people's life and the way of communication have changed greatly, and people are more and more inseparable from the Internet in every aspect of daily life. In this process, multimedia learning resources also show exponential growth [1]. Such a large number of learning resource data has brought great troubles to both learners and Internet learning resource services. For learners, in the face of such a large amount of multimedia learning resource data, it becomes very difficult to retrieve items that are relevant or interesting to them.

Massive and sufficient multimedia learning resources are the advantages of online learning. However, the excessive accumulation of online learning resources also brings

W. Fu et al. (Eds.): ICMTEL 2021, LNICST 388, pp. 243–254, 2021.
https://doi.org/10.1007/978-3-030-82565-2_20

difficulties in screening high-quality resources, which in disguise causes the plight of shortage of high-quality resources [2]. Therefore, the contradiction between the redundancy of online learning resources and the personalized learning needs of online learners makes the waste and utilization of multimedia learning resources increasingly prominent. It has become an important topic in the field of learning system research to select suitable learning resources for learners. In this context, recommendation system as a tool to provide personalized learning services, due to its high efficiency of resource retrieval, has attracted the attention of researchers, and is widely used in the field of education.

The existing recommendation system of learning resources has the defect of large average absolute error due to the limitation of its adaptive scope. Recursive neural network is an artificial neural network with tree hierarchical structure and network nodes recurse input information according to their connection order, which is one of the branches of deep learning algorithms [3]. Recursive neural networks, which have variable topology and weight sharing, are widely used in machine learning tasks containing structural relations and have been widely used in natural language processing. The aim of this study is to reduce the average absolute error of the recommendation of the design system and provide more accurate and effective recommendation service of multimedia learning resources for learners through the application of recursive neural network.

Therefore, this paper designs a new multimedia learning resource recommendation system based on recursive neural network. This system introduces the recursive neural network into the multimedia learning resource recommendation system architecture, and designs the learner demand retrieval representation module, the learner preference representation module, the recursive neural network training module, the database module and the system management module. The simulation experiment proves that the system has the advantage of small mean absolute error.

2 Design of Multimedia Learning Resource Recommendation System

In order to meet the needs of today's learners, a recurrent neural network is introduced to design a multimedia learning resource recommendation system. Its architecture is shown in Fig. 1.

Based on the system architecture shown in Fig. 1, the system functional modules are designed. The specific design process is as follows.

2.1 The Module of Learners' Needs Retrieval and Representation

In the records generated by the learner's browsing system, the sequence of multimedia learning resources interacting with the learner is recorded. The interactive behaviors generated by these multimedia learning resources and the learner are often of different types. For example, every learner has behaviors such as clicking, browsing, collecting, downloading, etc. [4]. This research intuitively and reasonably assumes that the different interactive behaviors of learners and multimedia learning resources represent different

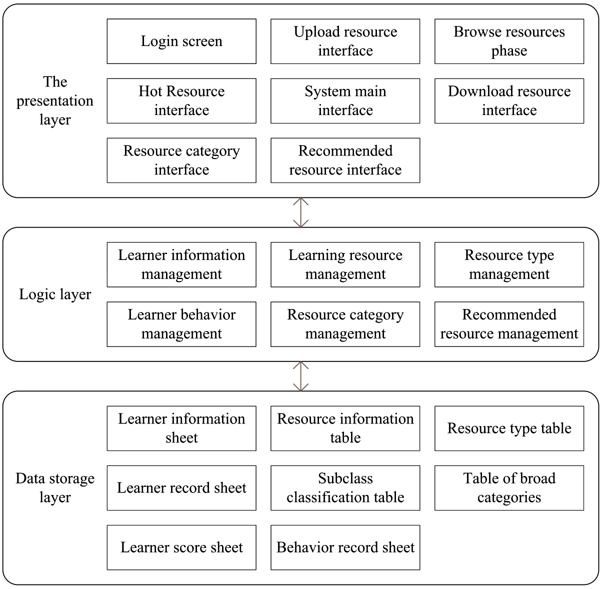

Fig. 1. Architecture of multimedia learning resource recommendation system

behavioral characteristics of learners. The purpose of the recurrent neural network algorithm is to obtain the expression of potential learner characteristics from the learner's demand retrieval. Therefore, an embedding layer of learner behavior characteristics is designed.

Specifically, the conventional process of processing discrete features by recurrent neural network is used to transform the ID of multimedia learning resources into the form of one hot coding. The dimension of one hot coding vector is the number of multimedia learning resources in the data set, which is a high-dimensional sparse vector. The embedding layer is essentially a full connection layer, which maps the one hot coding of learning resource ID from high-dimensional sparse vector to low-dimensional dense vector, which is conducive to feature extraction and abstraction of neural network. Based on this, after getting the dense embedding vectors of learning resources, the average pooling strategy is used to aggregate the learning resources of different interaction types in the history behavior records of learners, which are used as the feature vectors of different types of behaviors of learners, and these behavior feature vectors are connected to form the input layer of multi-layer perceptron.

The learner's demand retrieval program is shown in Fig. 2.

In order to better abstract the learner's demand retrieval feature into a vectorized expression, the vector representation of the potential learner's demand retrieval can be obtained from the learner's behavior feature combination vector through the fully connected hidden layer [5].

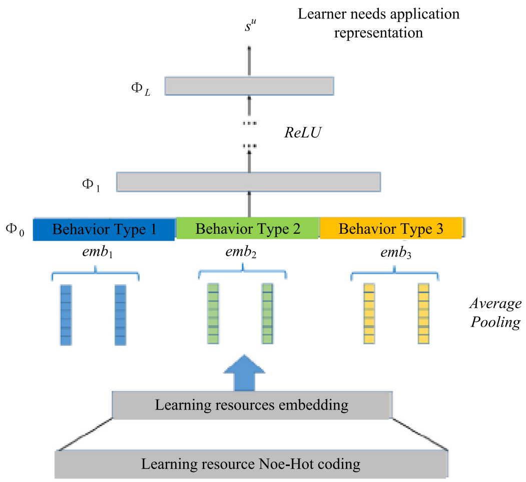

Fig. 2. Program diagram of learners' needs retrieval

As shown in Fig. 2, $emb_1, emb_2, \cdots, emb_n$ refers to embedding vectors of different behavior characteristics in learners' demand retrieval, and the input layer and hidden layer of recurrent neural network are represented as

$$\begin{cases} z_0 = \Phi_0(emb_1, \cdots, emb_n) = [emb_1; \cdots; emb_n] \\ z_1 = \Phi_0(z_0) = \sigma(W_1 z_0 + b_1) \\ \cdots \\ z_L = \Phi_L(z_{L-1}) = \sigma(W_L z_{L-1} + b_L) \end{cases} \tag{1}$$

In formula (1), z_0 represents the output expression of the input layer, which connects different types of user behavior feature vectors $emb_1, \cdots, emb_n$ to form the output of the hidden layer; $z_1 \cdots z_L$ represents the output expression of the hidden layer 1 to L; W_i and b_i respectively Represents the weight matrix and bias of the i layer; σ represents the activation function of the fully connected layer, and its expression is:

$$\mathrm{Re}LU(x) = \max(0, x) \tag{2}$$

As the activation function of neurons, relu function is proved not to lead to supersaturation. At the same time, it supports sparse activation, which is very suitable for sparse data, so that the recurrent neural network algorithm is not over fitted. The output z_L of the last layer of relu layer is the vectorized representation s^u of learners' requirements retrieval which is obtained by recurrent neural network algorithm.

2.2 Learner Preference Representation Module

Learners' recent behavior reflects the user's recent needs or intentions. Modeling learners' short-term behavior is an important task to better understand learners' preferences. Therefore, self attention mechanism is used to model the nearest learning resource sequence in the learners' demand retrieval sequence to capture the learners' interest in short-term learning resource demand [6]. Self attention mechanism is a special case of attention mechanism. Different from the basic attention mechanism, it improves attention expression by matching a single sequence with itself.

In order to construct the temporal characteristics of the learner's demand retrieval behavior, the sequence position code is added to the learner's demand retrieval behavior sequence [7]. First, the user's demand retrieval behavior is sorted by timestamp, and then the timestamp is discretized into a time series number, and these sequences are coded using a triangular coding method. The coding formula is as follows:

$$PE_i(t) = \begin{cases} PE_{2i}(t) = \sin\left(\frac{t}{10000^{\frac{2i}{d}}}\right) \\ PE_{2i+1}(t) = \cos\left(\frac{t}{10000^{\frac{2i}{d}}}\right) \end{cases} \tag{3}$$

In formula (3), $PE_i(t)$ represents the element value of the i dimension of the $PE(t)$ vector of the time series position code; t represents the discretized time series; d represents the dimension of the position code, which is the same as the dimension of the multimedia learning resource embedding vector.

Considering the embedded expression vector sequence $X_u = \left[x_1^u, x_2^u, \cdots, x_t^u\right]$ of learners' needs retrieval sequence, the size of the observation window of attention mechanism is set to H, that is, the nearest H vectors in the learners' demand retrieval sequence are intercepted as the input of neural network, that is, the subsequence x_H^u with X_u length of H is selected as the input sequence of learners' short-term preference expression:

$$X_H^u = \left[x_{t-H+1}^u, x_{t-H+2}^u, \cdots, x_t^u\right] \tag{4}$$

In Eq. (4), $X_H^u \in R^{H \times d}$ and d denote the dimension of vector embedded vector.

Specifically, the learning resource vector in the learner demand retrieval sequence is added to the time-series position encoding vector at the corresponding time to form a new learner demand retrieval sequence, namely:

$$\tilde{X}_H^u = \left[\tilde{x}_{t-H+1}^u, \tilde{x}_{t-H+2}^u, \cdots, \tilde{x}_t^u\right] \tag{5}$$

Then, $\tilde{X}_H^u$ nonlinear mapping transformation is performed to obtain the query matrix and key value matrix of the attention mechanism. The mapping transformation formula is as follows:

$$\begin{cases} Q = \sigma\left(\tilde{X}_H^u W_Q\right) \\ K = \sigma\left(\tilde{X}_H^u W_K\right) \end{cases} \tag{6}$$

In formula (6), W_Q and W_K respectively represent the weight matrix in the query matrix and the key matrix mapping transformation; $\sigma(\cdot)$ represents the nonlinear activation function, and this study uses the ReLU function as the activation function. After that, the attention weight matrix can be obtained by calculating the attention weight:

$$M_H^u = soft\max\left(\frac{QK^T}{\sqrt{d}}\right) \tag{7}$$

In Eq. (7), $\sqrt{d}$ represents the scaling factor of inner product calculation.

The short-term preference of learners can be obtained by the above formula:

$$a^{u} = \frac{1}{H}\sum_{h=1}^{H} Att_{h} \tag{8}$$

In formula (8), it represents the attention expression matrix of the learner's search sequence.

2.3 Recurrent Neural Network Training Module

Based on the learner's demand retrieval vectorized representation s^u and the learner's short-term preference representation a^u, the two are connected in vectors, and then a fully connected hidden layer is used to obtain the fusion feature expression of the learner's long-term historical behavior and short-term preference. The hidden layer The output is expressed as:

$$h = \sigma\left(W_T\left[s^u; a^u\right] + b_T\right) \tag{9}$$

In Eq. (9), W_T and b_T are the weight matrix and offset vector of the hidden layer respectively.

The fusion feature expression h is used as the input vector of the output layer, and the output layer uses softmax activation function to get the output vector expression of recurrent neural network

$$\hat{y} = soft\max(W_{\mathrm{o}}h + b_o) \tag{10}$$

In formula (10), W_{o} and b_o are the weight matrix and bias vector of the output layer, respectively.

Considering the learner's needs, the retrieval behavior is an implicit feedback behavior. In the neural network training, the learner's next interactive learning resource is used as a positive sample, and other non-interactive multimedia learning resources are used as a negative sample, so that the recurrent neural network will take the next Requirement retrieval modeling is a multi-classification task, using multi-class cross entropy as the objective function of neural network minimization:

$$Loss = -\sum_{i=1}^{|V|} y_i \log\left(\hat{y}_i\right) + (1 - y_i)\log\left(1 - \hat{y}_i\right) \tag{11}$$

In formula (11), $\hat{y}_i$ is the element value of the i dimension of the output vector $\hat{y}$ of the recurrent neural network, which represents the probability distribution vector of the learner's next learning goal. The element value $\hat{y}_i$ of each dimension of the vector represents the possible probability of learning resource v_i appearing in the next click of learners in the current sequence, which can be regarded as the recommendation basis for the next click learning resource. y_i represents whether the learner has real interaction with learning resource v_i. when the learner interacts with learning resource v_i, then

$y_i = 1$, otherwise $y_i = 0$. In the probability distribution $\hat{y}_i$ of the output vector, the Top-k recommendation of the next click multimedia learning resources can be obtained by selecting the first K maximum values.

Input the learners' demand retrieval statements into the recursive neural network algorithm, and the output result is the recommendation result of multimedia learning resources [8].

2.4 Database Module

The background of multimedia learning resource recommendation system mainly uses MySQL database for data processing. It mainly designs the learner information table, multimedia learning resource information table, multimedia learning resource type table, learner resource rating table, learning record table, large category classification table, small class classification table and learner behavior table. The specific design of these tables is introduced below.

(1) Learner information table
The learner information table stores learner-related data, such as user name, password, date of birth, gender, and major. The specific table structure is shown in Table 1.

Table 1. Learner information table

Field Chinese name	English field	Types of
User number	userID	int
Username	userName	varchar(20)
Real name	realName	varchar(20)
Password	password	varchar(20)
Gender	sex	char(2)
Date of birth	birthday	datetime
Profession	career	varchar(15)
Telephone	phone	varchar(12)
Mailbox	email	varchar(30)

(2) Information table of multimedia learning resources

The multimedia learning resource information table stores the relevant data of learning resources, such as the name of learning resources, the type of learning resources, the classification number of major categories of learning resources, the classification numbers of small categories of learning resources and the description of learning resources. The specific table structure is shown in Table 2.

Table 2. Multimedia learning resource information table

Field Chinese name	English field	Types of
Resource number	resourceID	int
Resource Name	resourceName	varchar(20)
Resource type	resourceTypeID	int
release time	releasetime	datetime
Opening Hours	openingtime	datetime
Belonging to the course	courseName	varchar(20)
Uploader number	authorID	int
description	discration	varchar(200)
Category number	subjectID	int
Subclass classification number	subsubjectID	int
Storage path	path	varchar(200)

2.5 System Management Module

The system management module is mainly divided into four parts: multimedia learning resource category management, multimedia learning resource management, learner management and recommendation management.

(a) Category management of multimedia learning resources. The category management of multimedia learning resources is mainly to classify learning resources into major and sub-categories. Pedagogy, philosophy, literature, economics, science, law, medicine, history, engineering and management are regarded as major categories. Divide the major categories into several sub-categories. For example, take computer, electrical, economic management, architecture, finance, English, diplomacy and psychology as subcategories. The learner can select the corresponding category for quick query according to the category of the learning resource. According to the different types of learning resources, the learning resources are divided into course resources, case resources, e-books, exam certification, material resources and other resources. Learners can choose the type of learning resources needed according to their own needs;

(b) Multimedia learning resource management. The functions of multimedia learning resource management include uploading new learning resources, filling in the name, resource type, large category classification category, small category classification category of the learning resource to be uploaded and the introduction of the resource [9, 10]. After the successful upload of learning resources, the administrator also needs to check whether the resources meet the requirements. If the learning resource information is not correct, you can modify and delete the learning resource;

(c) Learner management. The functions realized by learner management include learner registration, login, downloading learning resources, online browsing learning resources, collecting learning resources, scoring and commenting learning resources. The information registered by the learner includes user name, password, age, major and other information. At the same time, the system has recommendations of popular resources and the latest resources to facilitate learners' learning of learning resources;

(d) Recommendation management. The function of recommendation management is to recommend the interesting learning resources for different learners. The recursive neural network algorithm is applied to the multimedia learning resource recommendation system, which provides the multimedia learning resource recommendation service for learners.

Through the design of the above system architecture and functional modules, the operation of the multimedia learning resource recommendation system is realized, and learning resources more in line with their needs are provided for learners.

3 System Application Performance Test

In order to verify the application performance of the multimedia learning resource recommendation system based on recurrent neural network, MATLAB software is used to design simulation experiments. The specific experimental process is as follows.

3.1 Experimental Data Set Preparation

Select Gowalla dataset and Yelp dataset as experimental data. Each check-in record in the data set contains a timestamp, a user ID, the user's social relationship and ResourceID. The statistical overview of the two data sets is shown in Table 3.

Table 3. Statistical overview of data sets

Data set	Gowalla	Yelp
User	43074	30887
Resource	46234	18995
Check-ins	1720082	860888
Density	0.0500%	0.1399%

In order to filter noise data, for Gowalla dataset, users with less than 20 signatures and multimedia learning resources with less than 20 visits are deleted. For yelp dataset, users with less than 10 signatures and multimedia learning resources accessed less than 10 times are deleted.

3.2 Experimental Environment Construction

Each node of the recurrent neural network can have data input. For the nodes of the i level, the calculation method of the system state is as follows:

$$h^{(i)} = f\left(U^T h_c^{(i)} + W^T X + b\right) \tag{12}$$

In formula (12), $h^{(i)}$ and $h_c^{(i)}$ are the system states of the node and all its parent nodes. When there are multiple parent nodes, the system state can be combined into a matrix. If the node has no input, no calculation is performed. It is an excitation function or encapsulated feedforward neural network, the latter corresponds to gating algorithms and some deep algorithms.

The experimental recurrent neural network environment is shown in Fig. 3.

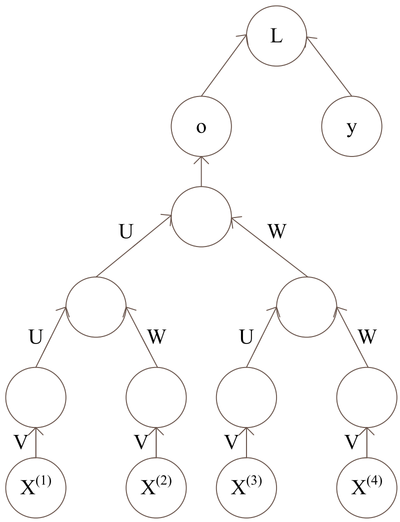

Fig. 3. Schematic diagram of recurrent neural network environment

In Fig. 3, U, W and V are weight coefficients, X is input data, o is output information, y is error, and L is candidate information.

3.3 Selection of Experimental Evaluation Index

The mean absolute error MAE was chosen as the evaluation index. At present, almost all the researches on recommendation choose Mae as the evaluation standard. Mae determines whether the recommendation is accurate by comparing the deviation degree. Deviation degree refers to the error between the recommendation results and the actual needs of users. The lower the MAE value is, the smaller the recommendation error is, the more accurate the recommendation is; on the contrary, the higher the MAE value, the greater the recommendation error and the worse the recommendation effect.

The average absolute error formula is defined as follows:

$$MAE_i = \frac{\sum_{j=1}^{n} |x_i - y_i|}{n} \tag{13}$$

In formula (13), n represents the total number of resources required by the user; x_i represents the system's demand score for resource i; y_i represents the user's real demand score for resource i.

3.4 Analysis of Experimental Results

According to the prepared experimental data and experimental environment, the multimedia learning resource recommendation experiment is carried out, and the average absolute error data of the recommendation is shown in Table 4.

Table 4. Recommended mean absolute error data table

(1) Gowalla data set		
Number of experiments/time	Recommended average absolute error coefficient	
	Existing systems	Text system
1	0.98	0.56
2	0.98	0.50
3	0.85	0.54
4	0.78	0.51
5	0.80	0.50
Average value	0.878	0.522
(2)Yelp data set		
Number of experiments/time	Recommended average absolute error coefficient	
	Existing systems	Text system
1	0.89	0.50
2	0.84	0.49
3	0.87	0.45
4	0.94	0.50
5	0.90	0.48
Average value	0.888	0.484

According to the data in Table 4, compared with the existing system, the average absolute error of the system in this paper is reduced by 0.356 and 0.404 in the Gowalla and Yelp data sets, which fully shows that the recommendation effect of the system in this paper is better.

4 Conclusion

In this paper, a new multimedia learning resource recommendation system based on recursive neural network is designed. This system introduces the recursive neural network into the multimedia learning resource recommendation system architecture, and designs the learner demand retrieval representation module, the learner preference representation module, the recursive neural network training module, the database module and the system management module. The system greatly reduces the average absolute error of the system recommendation, can provide more accurate multimedia learning resources for learners, and also provides more effective support for improving the utilization of learning resources.

References

1. Zhu, H., Hao, J., Niu, Y., et al.: Molecular targets of Chinese herbs: a clinical study of metastatic colorectal cancer based on network pharmacology **8**(1), 7238–7238 (2018)
2. Lee, T., Lee, I.: AraGWAB: network-based boosting of genome-wide association studies in Arabidopsis Thaliana **15**(6), 2925–2925
3. Tanabe, K., Liu, J., Kato, D., et al.: LC–MS/MS-based quantitative study of the acyl group- and site-selectivity of human sirtuins to acylated nucleosomes **37**(2), 2656–2656 (2018)
4. Radner, J.M., Ferrer, M.J.S., Dominique, M.M., et al.: Practical considerations for transitioning early childhood interventions to scale: lessons from the Saving Brains portfolio. Ann. N. Y. Acad. Sci **1419**(1), 230–248 (2018)
5. Liu, S., Li, Z.J., Zhang, Y.D., et al.: Introduction of key problems in long-distance learning and training. Mob. Netw. Appl. **24**(1), 1–4 (2019)
6. Liu, S., Lu, M.Y., Li, H.S., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)
7. Liu, S., Matt, G., Marco, Z., et al. (eds.): E-Learning, E-Education, and Online Training, pp. 1–374. Springer, Cham (20s18). https://doi.org/10.1007/978-3-319-93719-9
8. Liu, Q., Dong, Z., Wang, E.: Cut based method for comparing complex networks. Sci. Rep. **8**(17), 5134s (2018)
9. Wang, J., Cao, H., Zhang, J.Z.H., et al.: Computational protein design with deep learning neural networks. Sci. Rep. **45**(19), 6349 (2018)
10. Roffman, D., Hart, G., Girardi, M., et al.: Predicting non-melanoma skin cancer via a multi-parameterized artificial neural network. Sci. Rep. **28**(13), 1701 (2018)