Journal of Applied Data Sciences
Vol. 3, No. 2, May 2022, pp. 66-71

Research on Short Video Publishing Algorithm and Recommendation Mechanism Based on Artificial Intelligence

Lei An *

Eurasia University, Shaanxi, Xi'an, China
anlei@eurasia.edu*
* corresponding author

(Received: February 22, 2022; Revised: March 10, 2022; Accepted: April 25, 2022; Available online: May 24, 2022)

Abstract

Aiming at the problem of poor feature expression ability and model representation effect of traditional video recommendation mechanism, combined with the characteristics of traditional recommendation algorithm, this paper deeply studies the short video publishing algorithm and recommendation mechanism under artificial intelligence, and constructs a two-layer feature representation model BIFR based on attention. Firstly, the basic principle of recommendation algorithm is introduced in detail, and then the internal representation of features is studied through a multi head self attention mechanism to deeply mine the correlation between features and further improve the expressiveness of features. Then adjust the input feature crossover to learn the feature crossover more effectively. Finally, combine the two, add DNN to get the final output results, and then use the corresponding evaluation indicators to test the constructed recommendation model. The test results show that the video recommendation model constructed in this study has high accuracy, strong expressiveness and effectiveness.

Keywords: Video recommendation algorithm; Double layer feature representation; Multi head self attention

1. Introduction

With the rapid development of big data and artificial intelligence technology, short videos such as Dou yin and Kuai shou apps have gradually been integrated into People's Daily life, becoming the most popular form of entertainment in current society. As the threshold of short video creation is getting lower and lower, the number of short videos is increasing rapidly. As a result, there is too much short video information at present, and it is difficult for people to choose their favorite videos from a large number of videos. In order to improve the pertinacity of short video, the video recommendation algorithm mechanism was born. However, the traditional recommendation system has many defects such as insufficient feature expression ability and poor performance of the recommendation model in practical application. How to improve the feature extraction ability of recommendation systems and improve the video feature expression effect have become the focus of current research in the field of short video. Scholar Gao Chenfeng proposed the research of short video recommendation technology based on multi modal content analysis, and adopted the method of feature, result and mixed analysis to build an end to-end short video feature processing model, which improved the feature extraction ability and recommendation efficiency of the video recommendation system [1]; Yao Wei proposed the design of

video recommendation system based on collaborative filtering, and built a recommendation system through the traditional video recommendation algorithm, which improves the feature performance and recommendation efficiency of video to a certain extent [2]. However, the research of the above scholars only improves the feature extraction ability temporarily, and the recommendation system method is not rigorous enough, which leads to the poor recommendation effect of the final video. Based on this, combining with the empirical research of many scholars, this study constructs BIFR, a double-layer feature representation model based on attention, to improve the expression ability of video features and improve the accuracy of video recommendation through the model, and verify the expressive force and effectiveness of the recommendation model built in this study through experiments.

1.1. Basic algorithm

1.1.1. Algorithm Introduction

Traditional video recommendation algorithms mainly include collaborative filtering, content-based and hybrid recommendation algorithms. Among them, the recommendation algorithm with the highest common usage is based on collaborative filtering, which mainly includes three collaborative filtering algorithms based on users, items and models [3].

The recommendation algorithm mines users' interests and hobbies according to their historical behavior characteristics, and makes recommendations to users according to the mining situation. The calculation process of this algorithm is very simple and easy to understand. However, this algorithm also has some shortcomings, that is, most of the mining information needs to be obtained from the historical data, which is easy to appear the "cold start" problem, and the prediction accuracy will be seriously affected when the data is small. Based on this, this study explores the feature representation in the current recommendation algorithm, improves the traditional recommendation algorithm, and builds an algorithm-based model to further improve the accuracy of short video recommendation.

1.1.2. Algorithm Improvement

Commonly used recommendation models include logistic regression model and neural network recommendation model [4]. The input feature vector of the traditional video recommendation algorithm mechanism is usually high dimensional and sparse, and the internal relationship and cross relationship between each feature are crisscrossed. This problem needs to be solved and the feature expressiveness can be enhanced to improve the recommendation effect. Therefore, based on the traditional recommendation system and the actual situation of video recommendation, this study combines feature interior and feature cross to construct BIFR, a double-layer feature representation model based on attention. Firstly, the internal correlation representation of features is learned by multi-head self attention mechanisms, and the external cross representation of features is learned by attention networks. Then combine the two, so as to improve the feature extraction ability, enhance the feature expression, and further improve the short video recommendation effect.

2. Model Construction Based on Algorithm

The overall structure of the model constructed in this study is mainly divided into five levels, namely input layer, embedding layer, MHSA layer, AC layer and output layer [5]. See Figure 1.

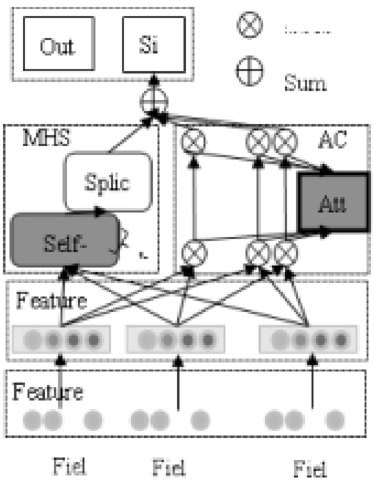

Figure. 1. Overall framework of the recommendation model based on attention-based double layer feature representation

Since BIFR model belongs to low-order feature representation, this study adopts the method of matching Deep part to learn higher-order feature information [6]. Firstly, after splicing MHSA layer output result and layer output result $S_{MHSA} = [S_1, S_2, ... S_n]$ and layer output result $P_{AC} = [P_{1,1}, P_{1,2}, ... P_{n,n}]$ bilayer model, more efficient features are learned through multi-layer perceptron. Finally, the learned output results are transformed into prediction results by sigmoid function, as shown in Equations (1) and (2).

$$a^{(0)} = concat(S_{MHSA}, P_{AC}) = [C_1, C_2, ..., C_n] \quad (1)$$

$$\widehat{y} = \sigma(W_0 + \sum_{I=1}^{n} W_i, X_i + DNN(a^{(0)}) \quad (2)$$

In the above equation, σ(.) belongs to the function ^{sigmoid}; represents the original input characteristics; w_i represents the weight of the ^{i} th input feature; $DNN(a^{(0)})$ represents the depth result, $a^{(0)}$ represents the input vector, and $a^{(0)}$ is formed by the output result set of MHSA layer and AC layer [7].

3. Simulation Test

In order to verify whether the recommendation model constructed in this study has a good recommendation effect, the data sets in this experiment mainly include two kinds, namely, MovieLens 1M dataset and 1CME data set.

3.1. Experimental Environment

In order to obtain a more realistic and accurate experimental result, python 3 and TensorFlow1.14.0 software platform were selected for the experiment, and cross entropy was set as the minimum purpose function of model construction [8].

3.2. Parameter Settings

Parameter Settings of the model constructed in this experiment are shown in Table 1:

Table. 1. Model parameter configuration

The parameter name	The parameter value	
	MovieLens-IM data set	ICME data set
Enter the feature dimension size	32 d	
Self-attention feature dimension size	16 d	
Self-attention head quantity	2	
ull connection layer number of layers	2	
Number of nodes per layer	128	
Activation function	ReLU activation function	
Improved algorithm	Adam- optimization algorithm	
Learning rate	0.002	0.0004
Batch size	2048	8192

3.3. Evaluation Index

Because the model type of this experiment is a binary classification model. Therefore, AUC and Logloss are used as the evaluation indexes of the recommended system model. AUC can avoid the interference of different data and focus on the sorting of positive and negative sample data [9]. LogLoss focuses on the error between the predicted results and the real results.

Common evaluation methods include accuracy, recall rate, F_1 index and AUC. Among the evaluation indexes, confusion is easy to occur, among which, accuracy will affect the recommendation effect, recall rate, , true rate TPR, false positive rate FPR all have certain limitations. After F1 calculating the true rate TPR and false positive rate FPR, the ROC curve was obtained. TPR represents the ordinate and FPR represents the abscissa. TPR increases with the decrease of FPR, and the two are inversely proportional. In other words, the greater the fluctuation of the ROC curve, the better the model recommendation effect. Therefore, the experiment chooses AUC as the evaluation standard. The larger THE AUC, the better the model effect and the stronger the classification ability.

LogLoss is an evaluation index widely used by the dictation model, which can accurately calculate the average deviation of sample data and reduce the probability that the prediction standard becomes a category [10]. LogLoss is the logarithmic loss, which can also be called the logarithmic likelihood loss. The expression is Formula (3) :

$$-\ (y\, log(p)\ \ (1\ y)log(1\ p) \qquad (3)$$

In the above formula, P represents the probability of a positive sample, and 1-P represents the probability of a negative sample.
In the dichotomous model, it represents the existence of two results. If the input sample y {0,1) size is N, then the logarithmic loss of the model represents the average of the logarithmic loss of each sample, as shown in Equation (4) :

$$LogLoss\ =\ -\ \frac{1}{N}\sum_{1}^{N}-\ (y\, log(p)\ \ (1\ y)log(1\ p) \qquad (4)$$

LogLoss can calculate the average error of the prediction model, so as to better evaluate the performance of the model. As the model type of this experiment is a dichotomous model, AUC and LogLoss are taken as evaluation indexes.

3.4. Results and Analysis

3.4.1. Comparative Analysis of Recommended Models

Experiments were conducted on the two data sets respectively, and the performance of the traditional recommendation model and the model proposed in this study was analyzed and compared, which were divided into low-order and high-order. The comparison results are shown in the following table.

Table 2. Comparison table of model performance evaluation

	Model	MovieLens-1M data set		ICME data set		Low and high order mean difference	
		AUC	Log Loss	AUC	Log Loss	AUC	Log Loss
low order	LR	0.6984	0.3957	0.7215	0.5958	-	-
	FM	0.7958	0.4012	0.6958	0.6015	-	-
	AEM	0.7819	0.3954	0.7187	0.6024	-	-

	BIFR	0.7955	0.3899	0.6895	0.6123	-	
	DNN	0.8012	0.3911	0.6854	0.6011	-	-
Higher order	Wide & Deep (LR)	0.7781	0.4015	0.6992	0.5897	+0.0524	-0.0410
	Deep(FM)	0.7995	0.3958	0.6758	0.5902	+0.0101	-0.0080
	DeepBIFR (Ours)	0.7857	0.3846	0.6983	0.5974	+0.0090	-0.0069

It can be seen from table 2 that compared with other low-order models, BIFR model has better performance on the two data sets, and its AUC index is significantly improved. After adding DNN to output the results, the prediction ability of each model is improved. From the mean difference data, it can be seen that the DeepBIFR model obtained by adding DNN has superior model performance, strong expressiveness and better video recommendation effect compared with other low-order models.

3.4.2. Impact Analysis of Model Parameters

After comparing the performance of each model, this experiment mainly focuses on the impact of the number of full connection layers on the model after adding DNN, and obtains the comparison results of the number of full connection layers in the AUC index and log loss index of DeepBIFR model, as shown in Figure 2:

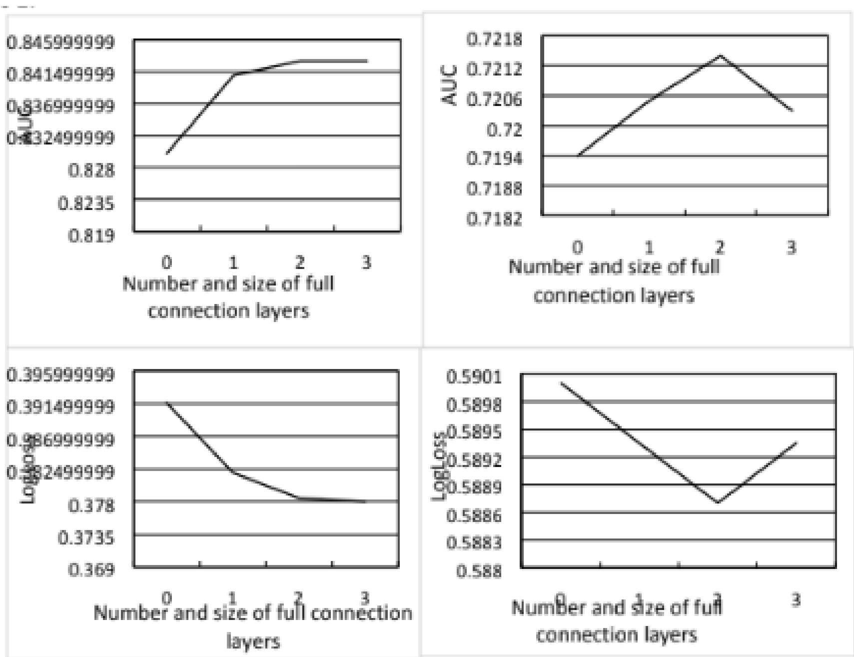

Figure. 2. Comparison between AUC index and log loss index of DeepBIFR model

As can be seen from Figure 2, when the number of full connection layers is 0, it indicates that the BIFR model is a shallow model. After adding one layer, the movieLens-1m data set and LCME data set are significantly improved. It shows that the feature combination of full connection is helpful to improve the prediction effect of the model. When the number of layers is added, the performance of the model is continuously improved. The main reason is that the multi-layer full connection takes into account the high-order feature combination and enhances the generalization ability of the model.

4. Conclusion

In conclusion, the BIFR model constructed in this study is feasible and effective. The internal feature representation combined with cross feature representation can improve the feature expression ability and the model accuracy. At the same time, after adding DNN neural network to BIFR model, the performance of the traditional recommendation model and the model proposed in this study are analyzed and compared. Compared with other low-order models, BIFR model has better performance on the two data sets, and its AUC index is significantly improved. After adding DNN to output the results, the prediction ability of each model is improved. In terms of mean difference data, compared with other low-order models, the DeepBIFR model obtained by adding DNN has superior model performance and better video recommendation effect. The test results show that the video recommendation model constructed in this study has high accuracy, strong expressiveness and effectiveness.

References

[1] Gao Chenfeng (2019). Research on short video recommendation technology for multimodal content analysis. Research on communication power, no.3 (14),pp: 277-278 + 280.

[2] Yao Wei (2021). Design of video recommendation system based on collaborative filtering. Network security technology and application, no.(05),pp: 54-55.

[3] Feng Yong, Zou Baoxu, Xu Hongyan (2021). Short video recommendation model based on the combination of video content features and bullet screen text. JOURNAL OF LIAONING UNIVERSITY (NATURAL SCIENCE EDITION), no.48 (02),pp: 108-115.

[4] Gao Chenfeng (2019). Research on short video personalized recommendation system Dou Yin for new media based on deep neural network. Satellite TV and broadband multimedia, no.(05),pp: 16-20.

[5] Feng Zhehui, Zhang Xiao (2019) ChuanYu. From the media background, the development of short video is taken as an example. stage, no. (05),pp: 97-100.

[6] Zhao Chenwei, Liu Tao, Du Hai Hong (2019). Research on video recommendation mode of short video platform under Dou Yin algorithm area. publishing wide-angle, no.(18),pp: 76-78.

[7] Cheng Xiaona, Sun Zhifeng (2020). Lfm-xgb-lr fusion recommendation algorithm in implicit feedback scenario . Computer engineering and application, no.56 (05), pp: 85-92.

[8] Zhao Nan, Pi Wenchao, Xu Changqiao (2020). A video recommendation algorithm for multidimensional feature analysis and filtering. Computer science, no.47 (04), pp: 103-107.

[9] Ding Herong (2021). The mechanism and problem of intelligent algorithm in Dou Yin short video. new media research, no.7 (10),pp: 10-12.

[10] Yan Hui, He Zixuan (2021). Research on user thinking of short video app in artificial intelligence environment. Media Forum, no.4 (13),pp: 163-164.