Original Research Paper Computer Science & Technology
Volume : 6 ¦ Issue : 3 ¦ Mar 2021 ¦ e-ISSN : 2456-5040

RESEARCH ON RECOMMENDATION TECHNOLOGY BASED ON KNOWLEDGE GRAPH

Feng Guo[1], Baoshan Sun[1]

[1]School of Computer Science and Technology, TianGong University, Tianjin, China.

ABSTRACT

The recommendation system is of great significance for screening effective information and improving the efficiency of information acquisition. It is widely used in many network scenarios to deal with information overload caused by massive information data, so as to improve user experience. In view of the strong practicability of the recommender system, researchers have conducted extensive research on its methods and applications. In recent years, many works have found that the rich information contained in the knowledge graph can effectively solve a series of key problems in the recommendation system, such as data sparseness, cold start, and recommendation diversity. Therefore, this article first introduces the main personalized recommendation technology and the related content of the knowledge graph, and then uses the knowledge graph to design a model of the fusion of the recommendation system and the knowledge graph. Finally, it is tested according to the data set, and the experiment proves that this model improves the task performance of the recommendation system.

KEYWORDS: Knowledge graph; recommendation system; collaborative filtering; alternate training.

1. INTRODUCTION:

With the promotion and popularization of social informatization, the rapid development of Internet technology has caused the explosive growth of information content, and it is difficult for users to select some of the information that is really useful to them under the situation of information overload. Therefore, how to effectively filter information for users is a difficult problem in the era of big data. The main problem of recommendation system research is how to find the content that each user is interested in from these huge information groups, and push these content to users.

The important challenges faced by recommender systems are mainly data sparsity and cold start. Data sparseness refers to the fact that compared with a large number of users and items, only a small number of items have been evaluated by users. It is difficult to obtain similar users or similar items based on this, which makes traditional recommendation methods invalid. The cold start problem refers to the fact that the system does not know the historical behavior of newly added users and cannot recommend items to them. Similarly, newly added items cannot be specifically recommended to users because they have not been evaluated or purchased by users.

Therefore, consider using the knowledge graph to improve the recommendation system. The auxiliary information of the knowledge graph can enrich the description of users and items, enhance the mining ability of the recommendation algorithm, and effectively solve the problem of sparsity and cold start, and improve the accuracy and diversity of the recommendation results. And interpretability. Therefore, how to effectively integrate various auxiliary data into the recommendation algorithm according to the characteristics of the specific recommendation scene has become a hot and difficult point in the research field of recommendation systems.

In order to enhance the recommendation effect, this paper designs a new model based on the basic idea of fusing the recommendation system and the knowledge graph, then inputs the data set to train the model, and finally conducts comparison experiments to prove that this model does improve the recommendation performance based on the evaluation indicators.

2. PERSONALIZED RECOMMENDATION TECHNOLOGY:

The recommendation algorithm is the core part of the personalized recommendation system, and its algorithm determines the speed of data processing and whether the recommendation result of the entire system can reach the best.

2.1 Similarity measure:

In the recommendation algorithm based on collaborative filtering, the method of object similarity measurement needs to be used to calculate user similarity and item similarity.

2.1.1 Pearson Correlation Coefficient:

Pearson's correlation coefficient uses covariance to measure similarity. Covariance is an indicator that reflects the degree of correlation between two random variables. If one variable becomes larger as the other variable becomes larger, then the covariance of the two variables is positive, and vice versa. The formula for covariance is as follows:

$$cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] \qquad (2-1)$$

Covariance can describe the degree of correlation between two random variables to a certain extent. When the covariance is greater than zero, it means the two are positively correlated, and when the covariance is less than zero, it means the two are negatively correlated. However, the value of covariance cannot be a good measure of the degree of correlation between two random variables. Therefore, we introduce the Pearson correlation coefficient and divide by the standard deviation of the two random variables on the basis of the covariance to better measure the degree of correlation between the two random variables. The formula for Pearson's correlation coefficient is as follows:

$$sim_{pearson}(X,Y) = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \qquad (2-2)$$

Where X and Y represent two sample points. When the Pearson similarity is greater than zero, it means that the random variables X and Y are positively correlated, and when the Pearson similarity is less than zero, it means that X and Y are negatively correlated.

2.1.2 Cosine similarity:

Cosine similarity measures the similarity of samples by calculating the cosine value of the angle between two samples in a multi-dimensional space. If the angle of the vector is smaller, the similarity of the corresponding sample is higher; The larger the angle of the vector, the lower the similarity of the corresponding sample. When the cosine value is 1, that is, when the angle of the vector is 0 degrees, the similarity of the sample reaches the maximum value, which means that the two vectors are exactly the same; When the cosine value is 0, that is, when the angle of the vector is 90 degrees, the sample similarity reaches the minimum value of 0, which means that the two vectors are completely different. The calculation formula of cosine similarity is as follows:

$$sim_{cosine}(X,Y) = cos\theta = \frac{X \cdot Y}{||X||\,||Y||} = \frac{\sum_{i=1}^{n}(x_i \times y_i)}{\sqrt{\sum_{i=1}^{n} x_i^2 \times \sum_{i=1}^{n} y_i^2}} \qquad (2-3)$$

2.2 Collaborative filtering recommendation:

The core idea of the collaborative filtering recommendation algorithm is to obtain the behavior records generated by some user groups, including comments, clicks, favorites, etc., and use these records to predict the interests of the target user group. Its algorithm system has been strongly promoted in academia and industry. For example, in industry, it is mainly used in e-commerce. When the target user group visits the website, the website can customize a personalized recommendation directory for the user on the online homepage based on the user's account data and behavior records. When users search for items, it can also help websites recommend similar items to users in real time, thereby increasing user stickiness and improving website efficiency.

2.2.1 User-based nearest neighbor recommendation:

The user-based nearest neighbor algorithm is the earliest recommendation algorithm proposed. Its core idea is: the first step is to obtain a rating data set and the current user's ID as input; For other users with similar behaviors, these other users are also called nearest neighbors; finally, for each item M that the current user has not encountered, the prediction score of M is calculated by other users. This method has an implicit condition: the interests and hobbies of users will remain similar forever, and the user's hobbies will remain the same. As shown in

Figure 2-1, to recommend items for User 1, you must first find User 2 and User 4, because their interests and hobbies are the same as User 1, that is, they all like Item 1 and also like Item 2. Also, since both user 2 and user 4 like item 4, and user 1 has not made a record of item 4, item 4 is recommended to user 1.

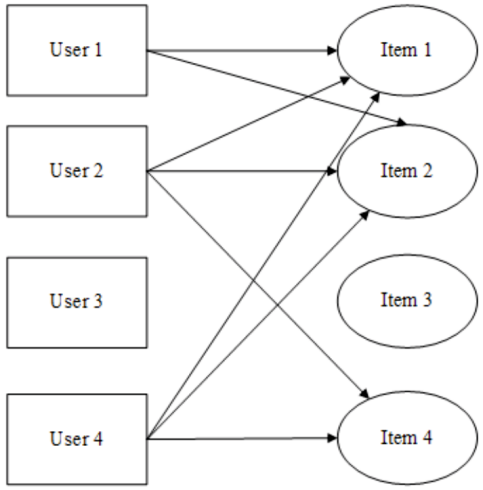

Figure 2-1: User-based nearest neighbor recommendation

2.2.2 Item-based nearest neighbor recommendation:

Although the user-based nearest neighbor algorithm was first proposed and extended to various industrial neighborhoods, it is not applicable to some large e-commerce sites with a large number of users and items, such as Taobao, Amazon, and JD. Due to the huge number of users, it is very time-consuming for the website to calculate their similarity, and it is impossible to make real-time recommendations to users. Later, in order to ensure the accuracy of the recommendation and the ability to quickly return the recommendation result, the researcher proposed an item-based recommendation technology. It is very suitable for offline calculations and can preprocess the data, that is, create an item similarity matrix privately, so although the scoring matrix is huge, it can also return to the recommendation list in time.

The item-based nearest neighbor recommendation utilizes the similarity between items. The core idea is to calculate the similarity between items, and then recommend items that are similar to the user's favorite items to the user, as shown in Figure 2-2.

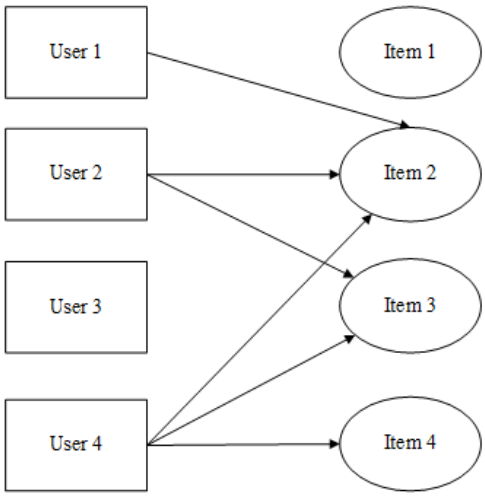

Figure 2-2: Item-based nearest neighbor recommendation

2.3 Content-based recommendation:

The collaborative filtering algorithm mainly uses the user's behavior records on the item, and ignores the user's own information, such as the user's age, gender, occupation, region, etc., as well as the item's own information, such as the type of item, manufacturer, etc.The content-based recommendation algorithm first collects and annotates the feature information, thereby constructing the portrait of the user and the item, and then realizes the personalized recommendation through the feature matching algorithm of the user portrait and the item portrait.
The recommendation algorithm based on collaborative filtering is particularly affected by the user-item matrix when calculating the similarity of users or items, or when training a recommendation model. When the user-item matrix is too sparse, the calculated user similarity and item similarity will be inaccurate, and the trained recommendation model will be inaccurate. The content-based recommendation algorithm can effectively solve the problem of data sparseness.

The content-based recommendation algorithm first digs out the relevance of the item's metadata content, and then uses the user's previous preference records to recommend relevant items to the user. The working principle of content-based recommendation algorithm is shown in Figure 2-3.

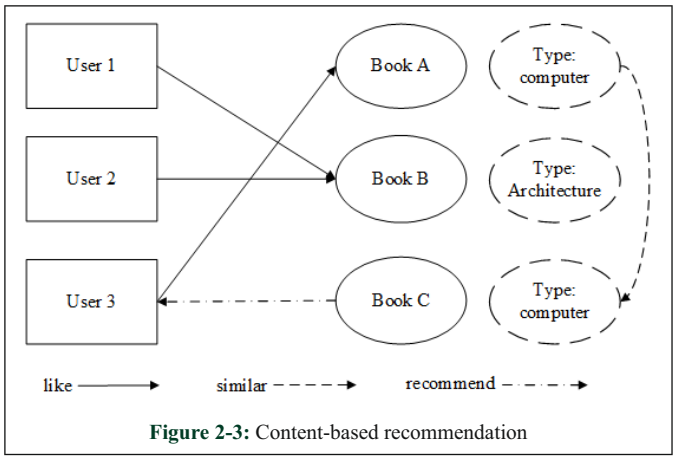

Figure 2-3: Content-based recommendation

Figure 2-3 shows a library system.We first model the metadata of the book. When modeling, we can extract the features of the item from multiple angles. Here, for simplicity, only the type features of the book are extracted.In this example, Book A and Book C are both computer-based, and Book B is architectural. Therefore, we think that Book A and Book C are similar. According to the user's preference record, it is found that user C likes book A, so we recommend similar book C to this user.

2.4 Hybrid recommendation method:

The commonly used recommendation algorithms were introduced above, but their use timing and requirements are different.After obtaining the user's historical interest, although each recommendation method can accurately recommend the user, their application environments in the engineering application field are different. Each algorithm has its advantages and disadvantages. When we obtain detailed information about users and items, as well as contextual information, often a recommendation algorithm does not use all the data, so researchers have proposed a hybrid recommendation system, which includes A variety of algorithms can not only take advantage of the advantages of one algorithm, but also make up for the shortcomings of another algorithm.

The hybrid recommendation algorithm is composed of two or more algorithms. There are two main ways to form it: the first one is a series type. It combines multiple recommendation ideas into one algorithm, and then works together; the second is parallel. The data is input into different recommendation algorithms, various operations, and finally the recommendation results are sorted and output.

3. KNOWLEDGE GRAPH TECHNOLOGY:

Knowledge graph is a kind of knowledge base, its concept was first proposed by Google, the purpose is to improve the search quality of search engines and enhance the user's search experience.The essence of knowledge graph is a structured network that stores the relationship between knowledge entities and entities, which can help formal description and understanding of things in the real world and their relationships. With the rapid development of Internet and Internet of Things technology, the application of knowledge graphs has gradually expanded from search engines to various fields, such as recommendation systems, intelligent question answering, and text analysis.

The advantages of the knowledge graph in the recommendation system: First, it is accuracy. The knowledge graph introduces more semantic relationships, which can discover user interests at a deeper level; secondly, it is diversity. Connecting through different relationships in the knowledge graph is conducive to the recommendation results. Divergence; In addition, the knowledge graph also has the interpretability of the recommended path. As shown in Figure 3-1, the process of applying the knowledge graph recommendation system mainly has the following steps:

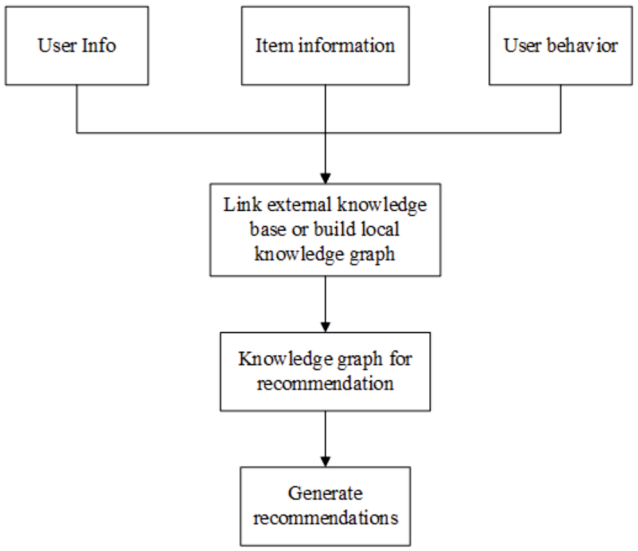

Figure 3-1: Recommendation system flow of applying knowledge graph

The first step is to collect data according to the characteristics of the recommended task. Data generally includes user information, item information, and user-item interaction information (for example, historical behaviors or ratings such as purchases, clicks, collections, etc.); the second step is to construct knowledge for recommendation based on the collected data (or link to external data) Graph; the third step is to design a recommendation model or method, and use the data collected in the first step and the knowledge graph constructed in the second step to train and verify the recommendation model or method; the fourth step is to use the trained recommendation model or method to generate Recommendation, to provide users with recommendation services.

4. RECOMMENDATION BASED ON KNOWLEDGE GRAPH:

4.1 Model framework:

In this paper, we use knowledge graphs for feature learning to obtain relation vectors and entity vectors. The recommendation system obtains the user vector and the item vector after learning. Because the entity vector and the item vector overlap, the framework of multi-task learning is adopted, and the recommendation system and knowledge map feature learning are regarded as two separate but related tasks, and alternate training and learning are performed.

According to the learning method, improve and construct a multi-task learning knowledge graph enhanced recommendation system (MKGRS) model. As shown in Figure 4-1, the model is divided into three parts. On the left is the recommendation module. The input of this part is the characteristic representation of users and items, and the estimated value of click-through rate is the output; On the right is the knowledge graph embedding module. This part uses the head node and relationship of the triple as input, and the predicted tail node as output; There is also a cross-feature sharing module in the middle, which is a link module that allows two modules to exchange information. Since the item vector and the entity vector are actually two descriptions of the same object, the cross-sharing of information between them allows both to obtain additional information from each other's modules, thereby making up for their own sparse information.

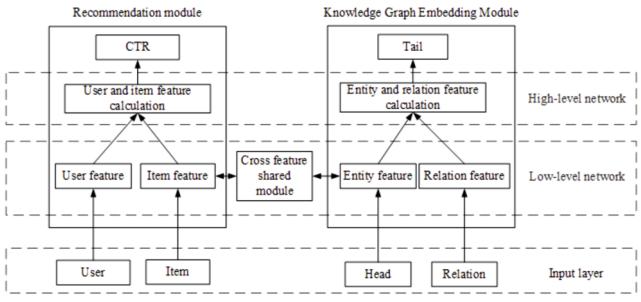

Figure 4-1: Schematic diagram of MKGRS model

4.2 Loss function:

The recommendation module, the knowledge graph embedding module, and the cross feature sharing module together form the MKGRS model. The loss function of the model is as follows, and the model is optimized by the loss function, see formula (4-1).

$$L = L_{RS} + L_{KG} + L_R$$

$$= \sum J(y'_{uv}, y_{uv}) - \lambda_1 \left(\sum score(h,r,t) - \sum score(h',r,t') \right) + \lambda_2 ||W||_2^2 \quad (4-1)$$

Where J is the cross-entropy function, the first term is the recommended module loss, the second term is the knowledge embedding module loss, and the third term is a regular term to prevent over fitting. λ_1 and λ_2 are balance parameters.

4.3 Model training:

The past gradient descent algorithm is divided into two types: the first is batch gradient descent, which is to calculate the loss function after traversing all the data, and then calculate the gradient of the parameters through the function to obtain a new gradient; the second is stochastic gradient descent, which is Calculate the loss function for each data, and then calculate the gradient to obtain new parameters. The former method is expensive because it traverses all the data samples in a loop, and can only be learned offline; the second method has a fast calculation speed because the calculated samples are small, but it may form a local optimal solution or the convergence effect is not good. , But very close to the optimal solution.

So this article uses a small batch gradient descent method to update the parameters. As a compromise between the above two gradient descent methods, it has the advantages of both. The data is divided into multiple samples, a small part of the samples are selected in each cycle, and the parameters are updated, so that the amount of calculation is small, and for a set of samples The data together determine the direction of gradient descent and increase accuracy.

5. EXPERIMENT AND RESULT ANALYSIS:

5.1 Data set:

This article uses the MovieLens-1M data set to verify the performance of the MKGRS model. The MovieLens-1M data set is a movie recommendation data set. It obtains movie ratings, movie metadata, and user characteristic information (such as gender, age, region, and industry sector) from various movie websites. The MovieLens-1M data set has about 6×10^3 users' rating details for about 4×10^3 movies. The rating of the MovieLens-1M data set is from 1 to 5 points. The data set is composed of a user information table, a movie information table, and a rating table. There are three tables in total.

5.2 Evaluation Index:

This article uses two indicators to evaluate model performance:

(1) Click-through rate forecast and evaluation. Apply the trained model to the test set, and use AUC and accuracy to evaluate the click-through rate effect. The larger the AUC value, the better the classification effect. AUC is the area of the lower part of the curve ROC on the coordinate axis. The ROC curve is drawn from the true rate and false positive rate on the coordinate axis.

(2) Select the accuracy rate and recall rate of the top K items after sorting to evaluate the recommendation performance.

5.3 Experimental design:

In order to verify the performance of the MKGRS model, this article will compare with the following models in the experiment:

(1) *CKE model:* Collaborative Knowledge Base Embedding is a model that adds text or graph data in the knowledge base to the recommendation system to improve recommendation accuracy. In this experiment, the embedded dimension of users and items is 64, and the embedded dimension of entities is 32.

(2) *DKN model:* Deep Knowledge-Aware Network is a model that combines knowledge graph entity embedding and neural network for recommendation. In this experiment, the movie name is used as the text input of the DKN network, and the entity embedding dimension is 64.

The data set is divided into training set, validation set and test set, and the ratio of the three is 6:2:2. The specific values of the data set are shown in Table 5-1.

Table 5-1: Data set values

Dataset	Users	Items	Interaction	KG triples
MovieLens-1M	4787	2152	574587	14745

5.4 Experimental results:

Run the MKGRS model, CKE model and DKN model on the data set to obtain the respective AUC and ACC of the three models. The comparison results are shown in Table 5-2.

Table 5-2: Predicted click-through rate evaluation results

Model	AUC	Accuracy
CKE	0.853（-4.2%）	0.771（-8.5%）
DKN	0.726（-18.5%）	0.635（-24.6%）
MKGRS	0.891	0.843

From the results in Table 5-2, it can be seen that the AUC of the MKGRS model is 4.2% and 18.5% higher than that of the CKE model and DKN model, respectively. The accuracy of the MKGRS model is 16.55 percent higher than that of the CKE model and DKN model. It can be seen that the classification effect of the MKGRS model is relatively good, and the prediction accuracy rate is higher than that of commonly used models.

Under the condition of recommending a list of K items, and calculating the results of each model according to the precision rate and recall rate formulas, it can be seen that the MKGRS model is superior to the other two models in terms of precision rate and recall rate, and has obvious advantages. Based on the comparison results of the above models, the MKGRS model can improve the recommendation performance.

6. CONCLUSION:

This paper uses the advantages of knowledge graph mining information to design a model that combines knowledge graph and recommendation module. This model is a generalized framework based on recommendation algorithm and can be applied to common recommendation requirements. Then, based on the data set and the experimental design process, through comparative experiments, the model framework designed in this paper is verified, and the recommendation accuracy is relatively high. After integrating other evaluation indicators, it is concluded that this model can moderately improve the recommendation performance.

REFERENCES:

I. Hernando A, Bobadilla J, Ortega F.A non-negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model[J].KnowledgeBased Systems, 2016, 97: 188-202.

II. Shu J, Shen X, Hai L, et al.A content-based recommendation algorithm for learning resources[J].Multimedia Systems, 2017(1): 1-11.

III. Chu W T,Tsai Y L.A hybrid recommendation system considering visual information for predicting favorite restaurants[J].World Wide Web, 2017, 20(6): 1313-1331.

IV. Rumelhart D E.Learning representations by back-propagating errors[J].Nature, 1986, 23.

V. Zhang S, Yao L, Sun A,et al.Deep learning-based recommender system:a survey and new perspectives[J].ACM Computing Surveys, 2019, 52(1): 5.

VI. Zhou Y, Huang C, Hu Q,et al.Personalized learning full-path recommendation model based on LSTM neural networks[J].Information Sciences, 2018, 444: 135-152.

VII. Wang Li-Cai, Meng Xiang-Wu, Zhang Yu-Jie.ContextAware recommender systems. Journal of Software, 2012, 23(1): 1-20.

VIII. Chang Liang, Zhang Wei-Tao, Gu Tian-Long,et al.Review of recommendation systems based on knowledgegraph[J].CAAI transactions on intelligent systems, 2019, 14(2): 207-216.

IX. Wang Guo-Xia, Liu He-Ping.Survey of personalized recommendation system. Computer Engineering and Applications, 2012, 48(7): 66-76.

X. Ye Y, Wang X, Yao J, et al. Bayes EMbedding (BEM):Refining Representation by Integrating Knowledge Graphs and Behavior-specific Networks[C].Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019: 679-688.

XI. Xian Y, Fu Z, Muthukrishnan S, et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation[J]. arXiv preprint arXiv:1906.05237, 2019.

XII. Wang H, Zhang F, Xie X, et al. DKN: deep knowledge-aware network for news recommendation. In: Proceedings of the 2018 World Wide Web Conference, 2018: 1835-1844.