Research on Personalized Hybrid Recommendation Algorithm Based on Deep Learning

Weiwei Guo[1,a] and Feng Liu[1,b]

[1] Heilongjiang University of Technology, JiXi,158100, China

[a]gwwguoweiwei@163.com; [b]liufeng8038@163.com

Keywords: Deep; Learning; Recommendation; Algorithm; Collaborative; Filtering

Abstract. In order to effectively improve the recommendation accuracy and recommendation quality of the recommendation system, the article firstly compares the traditional recommendation algorithm, and then combines the advantages of various recommendation algorithms to integrate the deep learning model into the recommendation system, and proposes a personalized hybrid based on deep learning. The recommendation algorithm, after deep learning mechanism, repeatedly adjusts the weight coefficient of the hybrid recommendation algorithm, and constructs an effective recommendation model. Finally, through simulation experiments, it is verified that the proposed hybrid recommendation algorithm has higher recommendation accuracy and thus effectively improves the quality of the recommended system.

Introduction

With the rapid development of cloud computing technology and big data technology, information overload has gradually become a difficult problem in network applications. In order to effectively solve the information overload problem caused by big data, scholars have introduced a recommendation system. The recommendation system can effectively solve the information overload problem. The recommendation system does not need to reconstruct the information category and information keyword. It is based on the characteristics of the user. The information, the item information that the user has browsed, and the evaluation information of the item, the user is imaged to obtain the user's point of interest, and then the item and information information of interest are recommended for the user. At present, many Internet application projects have integrated recommendation systems into network platforms, such as Jingdong, Vipshop, Taobao, and iQiyi [1].

The development of deep learning is derived from the neural network. The neural network uses the principle of pseudo-biology to imitate the human neural perception, enabling the computer to have a human-like perception system to realize artificial intelligence. The deep learning technology is based on the neural network to optimize and improve, obtain basic information from the input layer, and gradually adjust the feedback parameters through multi-layer transformation and abstract feature expression, so as to achieve image recognition, natural language processing and other functions. Research on deep learning has made some progress in the field of artificial intelligence. Deep learning first requires a nonlinear, multi-level network model that interacts with each other in the model. In the training process, deep learning needs to iterate through the established model through a large amount of data with labels, perform feature extraction, further determine the category of unlabeled data, and then perform iterative operations until the most accurate feature label is obtained [2].

In summary, the recommendation system is the main method to effectively solve the information overload. Combined with deep learning, the deep learning technology is integrated into the recommendation system, and the personalized hybrid recommendation algorithm is constructed to improve the performance and recommendation accuracy of the recommendation system， it has great research significance.

  DOI: 10.25236/icmcs.2019.008

Research Work

The recommendation system is capable of quickly helping users to filter out product information of interest. In the screening process, the user is not required to provide corresponding demand information. The screening and recommendation process is based on the user's historical behavior, obtaining the user's interest characteristics, and then screening and recommend. Commonly used recommendation algorithms are content-based recommendations, collaborative filtering recommendations, and hybrid recommendations. The content-based recommendation method is based on the content information of the user and the product, and does not require the user to evaluate the product, that is, the user behavior is not considered in the recommendation process, so the content-based recommendation method has no data sparsity and cold start. Question [3]. Collaborative filtering recommendation is to recommend the items of interest to the target users according to the historical behavior of the users. The advantage of collaborative filtering is that the historical behavior of users and products can be shared, thus avoiding inaccuracies and incompleteness in the content recommendation process. The problem, but the collaborative filtering algorithm also has drawbacks. In the recommendation process, it depends largely on the user's historical data. When the user does not score the product, or when new users add new products, it is easy to generate data sparsity and Cold boot issue. There are advantages and disadvantages in each of the two recommended algorithms, so the two can be combined to form a hybrid recommendation, and the commonly used recommendation algorithm is as follows [4].

Popularity Recommendation Algorithm. The popularity recommendation algorithm is an early proposed algorithm. In the recommendation process, the system recommends the most active products to users without any historical behavior. Here, a problem arises, how to calculate the popularity of the product. The calculation method can roughly use the number of transactions of the product, the historical behavior of the product being browsed, and the like. This kind of algorithm has certain defects. Firstly, it does not recommend according to the user's preference in the recommendation process. Subjectively, all popular products are suitable for all users. Secondly, products with high popularity are not recommended to users, and users can adopt other. Ways to obtain, low-popular products get recommended is very small [5].

Project-Based Collaborative Filtering Recommendation Algorithm. The project-based collaborative filtering algorithm is widely used on the network platform, and many algorithms are derived from this. The idea of this recommendation method is to recommend to the user similar items that the user has purchased the item, that is, to recommend according to the user's interests. In the recommendation process, it is mainly divided into two processes. The first process is to establish a model through the user's historical data to obtain the similarity between the projects. The second process is the application of the model, and the target user is recommended to be most similar to its behavior. project. Then, when calculating the similarity between projects, the similarity between cosine similarity and Pearson coefficient is often used for recommendation. We use sim(a,b) to represent the similarity between user a and user b. The cosine similarity method is given in Equation 1, and the Pearson coefficient similarity method is given in Equation 2 [6].

$$\mathrm{sim}(a,b) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|_2 * \|\vec{b}\|_2} \quad (1)$$

$$\mathrm{sim(a,b)} = \frac{\sum_{\mathrm{p}}(r_{a,p} - \overline{r_a})(r_{b,p} - \overline{r_b})}{\sqrt{\sum_p (r_{a,p} - \overline{r_a})^2}\sqrt{\sum_p (r_{b,p} - \overline{r_b})^2}} \quad (2)$$

Among them, $\overline{r_b}$ is used to represent the average score of user b for all items, and $\overline{r_{\mathrm{a}}}$ is used to represent the average score of user a for all items.

In addition to using the two methods mentioned above to calculate the similarity between projects, there are many other methods, which are not described here. In the process of use, you can choose the appropriate similarity calculation according to the actual situation method [7].

Content-Based Recommendation Algorithm. The content-based recommendation algorithm comes from the field of information retrieval and is an important research content. The main idea of this recommendation algorithm is to obtain the user's target model based on the data information selected by the user history. In the process of acquiring the target model, firstly, the content information feature of the user selected item is acquired, the user's preference model is constructed through the content information feature, and then the content feature information of the item is compared and matched with the user's interest point, and finally the similarity is high. The project is recommended to the target user. For example, in the video on demand system platform, the system analyzes the hidden information such as the type of the user watching the movie, and then uses the information to discover the user's interests and hobbies, and finally recommends the movie that the user may like according to the user's preference. We can express the preference of user a for item i as the similarity between user interest and project content, which can be expressed by the following formula 3 [8].

$$\mathrm{L(a,i)} = sim(user\_L_a, c_i) \quad (3)$$

Calculating the similarity by project content and user's preference is the most important part of the recommendation process. The cosine similarity or Pearson coefficient proposed by Equation 2 or Equation 1 can be used in the calculation process [9].

The content-based recommendation algorithm method is very simple, and the recommended project information is also intuitive and easy to understand, and there is no cold start and data sparsity problem to the collaborative filtering recommendation algorithm. However, in the content-based recommendation process, there are certain defects in the leaf. First, when there are new users, because the user has not selected the project information, there is a cold start problem for the new user, so it is difficult to obtain the user's interest information, which will When the information is recommended for the user, it has a similar phenomenon with the historical information, which leads to the over-fitting of the information [10]. The content-based recommendation algorithm is subject to certain constraints in the feature extraction process of users and projects, and has limitations on the recommendation of unstructured audio, video and text information. These defects will affect the accuracy and recommendation efficiency recommended to users [11].

Recommendation Algorithm Based On The Markov Chain. Markov Chain (MC) is a stochastic process with probability of Markov property and existence in discrete exponential sets and state spaces in probability and mathematical statistics. Markov chains suitable for continuous exponential sets are called Markov processes, but are sometimes considered as subsets of Markov chains, ie continuous-time Markov chains (Continuous-Time MC, CTMC), corresponding to Discrete-Time MC (DTMC), so Markov chain is a broader concept. Markov chains can be defined by transfer matrices and transition graphs. In addition to Markov properties, Markov chains may have irreducibility, reproducibility, periodicity, and ergodicity. An irreducible and positively reproducing Markov chain is a strictly stationary Markov chain with a uniquely stable distribution. The limit distribution of the traversed Markov chain converges to its stationary distribution. Time-series data can be processed well using Markov chains. In general, it is irrelevant to use future data to predict future data given current data knowledge and information. The user's consumption behavior or purchase behavior is a time series data, so the Markov chain can be applied to the recommendation algorithm [12].

In the recommendation process, different product information can be regarded as each state in the Markov chain, and the transfer probability of the product is not defined by the system, but by the frequency of transfer of the products in the deep learning training set. Combined with the principle of the Markov chain, the current product is only related to its neighboring previous product. Therefore, the recommended method of the Markov chain only needs to recommend the user's last information behavior to the user. Several products with the highest probability of selection in the information behavior are recommended to the user. Therefore, the most basic time series data can be processed by the Markov chain [13].

Personalized Hybrid Recommendation Algorithm

Construct a deep learning model to further improve the user's accuracy in project score prediction. In the deep learning training process, it can be divided into two steps. The first step is to perform initialization training, and the parameters for initializing training are performed in the first step; the second step is to perform parameter adjustment, using reverse propagation mode, In convolutional neural networks, there are gradient descent and gradient diffusion problems at the bottom of the network using reverse propagation. Therefore, the model can be gradually converged and the optimal solution can be obtained. The training parameters initialized in the first step may be used for reference. The optimal solution, so gradient adjustment can be greatly reduced by adjusting the parameters in the second step. Through the established deep learning model, combined with the hybrid recommendation algorithm, personalized recommendation is achieved.

Deep Learning Model Construction. The deep learning model is mainly based on the human cognitive process. The usual cognitive process is advanced concept learning, and then through simple concepts and knowledge to understand complex, abstract content and knowledge, by learning more from the data. Hierarchical representation and abstraction to solve supervisory and unsupervised learning tasks. In deep learning training, the Restricted Boltzmann Machine (RBM) is usually applied to training deep learning and has achieved good results. The structure of the restricted Boltzmann machine is shown in Figure 1.

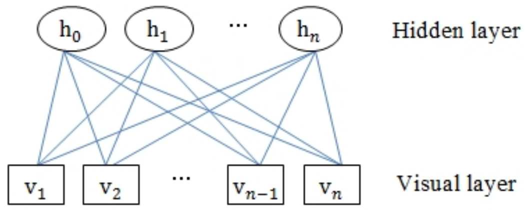

Figure 1 Restricted Boltzmann machine structure

The restricted Boltzmann machine is a two-layer undirected graph model. In the structure diagram, where v_i $(0 \leq i \leq n)$ represents a display node in the display layer, v_i $(0 \leq i \leq n)$represents a hidden node in the hidden layer, a hidden layer and a display layer The connecting lines can be represented by a matrix, which represents the weight relationship between the nodes. The nodes between the same hierarchy are independent of each other. In the usual case, the display layer represents the original input data, and the hidden layer represents the data formed by the deep learning training, indicating the hidden features existing in the original data.

Perceptron (LP) is a feedforward neural network with multiple hidden layers between the input layer and the output layer. It is a supervised learning model. It is essentially a simple feedforward neural network and a single layer neural network. As long as the input sample data is linearly separable, the perceptron algorithm can achieve convergence by multiple iterations. During the training process, the perceptron model is formed by the training of the sample data, and the test samples are classified and tested by the model.

Hybrid Recommendation Algorithm. This kind of algorithm combines a variety of different recommendation calculations according to certain conditions to make recommendations, so as to obtain more accurate recommendation results and improve the recommendation quality. In the design of the algorithm, this paper often uses content-based recommendation for new users, because new users do not have any historical data behavior, there is no way to score the project, so it is easy to generate cold start problem; for old users, collaborative filtering is adopted. The recommendation algorithm; on the whole, the ratio is based on different weight information, so as to obtain accurate recommendation results. Using the weighted hybrid recommendation strategy, by calculating the weighted sum of two or more recommendation results, they are combined according to a certain weight. Therefore, we can assume that there are n different recommendation algorithms, and the recommendation function corresponding to the recommendation function can be expressed as Re_k.

The weight corresponding to it can be expressed by β_k, so the weighted mixed recommendation formula 4 can be obtained, which is expressed as follows:

$$\mathrm{Re}_{weig}(u,i) = \sum_{k=1}^{n} \beta_k \times Re_k(u,i) \quad (4)$$

In Equation 4, all recommended item scores should be controlled within a certain range, and also satisfy $\sum_{\mathrm{k}=1}^{n} \beta_k = 1$, that is, the sum of all the weight parameters is 1 .

Simulation

The use of weighted hybrid recommendation is a relatively straightforward method. It is a more general recommendation strategy. It can combine the prediction capabilities of many recommended techniques in a weighted manner. The following is explained by simulation experiments. There are two recommendation algorithms that recommend one of the four items to the user a, as shown in Tables 1 and 2.

Table 1 Recommended Algorithm 1 Recommended Score

User name	Item 1	Item 2	Item 3	Item 4
Score	0.8	0	0.4	0.1

Table 2 Recommended Algorithm 2 Recommended Score

User name	Item 1	Item 2	Item 3	Item 4
Score	0.7	0.9	0.3	0.4

There are two different recommendation algorithms recommend different items to the user, the first recommendation algorithm recommends item 1, and the second recommendation algorithm recommends item 2 to the user. We use a balanced weighting scheme for hybrid recommendation, that is, the weight coefficient $\beta_1 = \beta_2 = 0.5$. By combining the scores of Re_1 and Re_2, the weighted scores of each item are shown in Table 3.

Table 3 Weighted scores

Item	Re_1	Re_1 sort	Re_2	Re_2 sort	Re_w	Re_w sort
Item 1	0.8	1	0.7	2	0.75	1
Item 2	0	3	0.9	1	0.45	2
Item 3	0.4	2	0.3	4	0.35	3
Item 4	0.1	4	0.4	3	0.25	4

After the weighted hybrid recommendation, item 1 has a higher recommendation order. Although item 2 has only one recommendation algorithm for recommendation, it also has a higher recommendation order after weighting. However, in the actual process, assuming that the user purchased item 1 and item 4, such purchase behavior does not conform to the recommended method of our design, so a weighting coefficient assignment method capable of reducing the target metric is needed, so that the average of the user's item rating set is obtained. The error is minimal. Therefore, the weight coefficients in the simulation experiment cannot be set equal, and the deep learning model should be trained within the specified range.

Summary

This paper studies the traditional recommendation algorithm and combines deep learning to propose a personalized hybrid recommendation algorithm based on deep learning. The simulation algorithm proves that the proposed hybrid algorithm has high recommendation after deep learning model training. Accuracy, which improves the overall recommendation quality.

Acknowledgements

This work was supported by the project of Nature Scientific Foundation of Heilongjiang Province (F2016038).

References

[1] X. Huang and S. Pan . Study on Grid Scheduling of Super-Peer Model Based on QoS[C]// Fourth International Symposium on Information Science & Engineering. IEEE Computer Society, 2012.

[2] F. Liu and W. Guo. Study on Grid Scheduling Model Based on Hierarchical Scheduling Model[J]. International Journal of Grid & Distributed Computing, 2015, 8(3): 1-10.

[3] F. Liu. Research on personalization algorithm based on collaborative filtering [J]. International Journal of u- and e- Service,Science and Technology, 2016, 9(2) :101-108.

[4] F. Liu and W. Guo. Research and Design of Task Scheduling Method Based on Grid Computing[C]// International Conference on Smart City and Systems Engineering. IEEE Computer Society, 2017:188-192.

[5] W.W. Guo. A Dynamic Resource Scheduling Strategy Based on Response and Resource Status Update Time [J]. Journal of Qufu Normal University：Natural Science Edition, 2018, 44(2): 54-58.

[6] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an Application to boosting. Journal of Computer and System Sciences, 1997, 55(1): 119-139

[7] Jiang Baining, feature selection algorithm research in machine learning,Ocean University of China, 2009

[8] L. Feng and W.W. G . Recommendation Algorithm Based on Tag Time Weighting [C]// 2018 International Conference on Smart City and Systems Engineering (ICSCSE). IEEE Computer Society, 2018.

[9] W.Q.Xiao, S.J. Yao and S.M. Wu An improved top-N collaborative filtering recommendation algorithm[J]. Journal of Computer Applications, 2018, 35(1).

[10] L.L. Yang and H.H. Yuan. Collaborative Filtering Recommendation Algorithm Based on Combination Optimization Theory[J]. Modern Electronic Technology, 2018, 41(1): 139-142.

[11] Yoav Freund, Raj Iyer, Robert E. Schapire, and Yoram Singer. An efficient boosting algorithm for combining preferences. Journal of Machine Learning Research, 2003, 4:933–969

[12] X.Q. Lü, T. Wang and X.W. Li, et al. Research on film recommendation algorithm based on content and interest drift model[J]. Journal of Computer Applications, 2018, 35(3):717-720.

[13] X.Y. Huang, Y.D. Li and L.Y. Xiong. Collaborative Filtering Recommendation Algorithm Based on Topological Potential User Clustering[J]. Computer Engineering and Design, 2018(1): 90-95.