Research on Commodity Recommendation System Based on Deep Learning

Chenyue Qian¹, Yanjuan Jin², Xiaobing Teng³

¹China Three Gorges University, Yichang, Hubei, 443000
²Zhejiang Huayun Information Technology Co., Ltd. Hangzhou, Zhejiang, 310000
³Zhejiang Huayun Power Engineering Design Consulting Co., Ltd. Hangzhou, Zhejiang, 310000

*Corresponding author: qianchenyue@ctgu.edu.cn

Abstract. The traditional matrix factorization model cannot effectively extract the features of users and items, but the feature information can be extracted well based on the deep learning model. At present, the mainstream recommendation algorithms based on deep learning only make recommendation prediction in the form of the product of neural network output or item features and user features, and cannot fully mine the relationship between users and items. Based on this, this paper proposes a recommendation algorithm based on the combination of text convolution neural network and singular value decomposition (Bias SVD) with biased terms. The text convolution neural network (Text CNN) is used to fully extract the feature information of users and items, and then the singular value decomposition method is used to make recommendations to deeply understand the document context information and further improve the accuracy of recommendation. The algorithm is widely evaluated and analyzed on two real data sets of MovieLens, and the accuracy of recommendation is obviously better than that of ConvMF algorithm and mainstream deep learning recommendation algorithm.

Keywords: Matrix factorization, singular value decomposition, deep learning, text convolutional neural network.

1. Introduction

With the advent of the cloud era, e-commerce, online news, and social media data have exploded [1]. According to IDC, the world will have 35ZB of data by 2020. These massive amounts of data have brought transformational development to human society, but at the same time they have brought about the problems of "information overload" and "long tail effect". How to obtain valuable information from these complicated data has become today's big data Difficulties to deal with. The recommendation system plays a key role as one of the methods to solve these problems, but it also faces some problems, such as the increasing sparseness of user ratings of items, and how to deal with these massive amounts of data, which all affect the quality of the recommendation system major factor.
Collaborative filtering is one of the most successful algorithms in recommendation algorithms. Especially model-based collaborative filtering algorithms have received a lot of research because of their scalability, accuracy, and ability to deal with sparse problems relatively well.

In recent years, neural networks have been widely used in natural language processing, image processing, and speech recognition, and have achieved outstanding results. Deep learning not only has the ability of self-learning, but also has powerful feature extraction and combination capabilities, so it can solve some of the problems of traditional matrix decomposition, such as sparsity, dimensional risk, calculation amount, inability to effectively extract feature information, interpretability, etc. Based on this, deep learning has been widely used in recommendation systems and has become a research focus in recent years.

This article will make improvements to this algorithm, combined with deep learning's powerful ability to learn the essential characteristics of data sets from samples, and the ability to automatically learn from multi-source heterogeneous data, using word embedding combined with feedforward neural network. The method is combined with the text convolutional neural network to extract the user and item features, and finally the neural network is used in parallel to output the predicted score items. Experiments show that the algorithm in this paper performs better than other collaborative filtering algorithms in the face of extremely sparse data sets, and is interpretable due to the use of user interaction information on items. The extensive comparative evaluation of the algorithm in this paper on the MovieLens-1M and MovieLens-10M data sets shows that it has greatly improved the RMSE and MAE indicators.

2. Method description

2.1. BiasSVD algorithm

Matrix decomposition is to infer and characterize the characteristics of users and items through the item score matrix. Because there is a high degree of consistency between the characteristics of users and items, it can be recommended. The rapid popularity of these matrix decomposition methods stems from their good scalability and relatively accurate characteristics, and they also provide great flexibility for many actual modeling situations. The matrix factorization model maps the user and item features to a joint latent factor model with dimension $\omega$. Set a vector $q_i \in \mathbb{R}^{\omega}$ associated with each item $i$, and a vector $p_u \in \mathbb{R}^{\omega}$ associated with each user $u$, and $q_i$ represents the item possessed characteristics, negative or positive, while $p_u$ represents the user’s degree of interest in these items, $r_{ui}$ represents the user’s rating of the item, and its predicted value is:

$$\hat{r}_{ui} = q_i^T p_u$$ (1)

This model is derived from SVD. When dealing with sparse scoring matrices, traditional matrix decomposition brings uncertainty, that is, large deviations will occur, so the scoring matrix often needs to be filled in the calculation. However, when encountering a large amount of data, this method is time-consuming and not easy to implement, and it is easy to overfit when dealing with fewer projects, so an improved version has appeared:

$$\min_{p,q} \sum_{(u,i) \in l} (r_{ui} - q_i^T p_u)^2 + \lambda (\| q_i \|^2 + \| p_u \|^2)$$ (2)

Among them, $l$ represents the data set, $\lambda$ is the parameter that controls the degree of regularization, and the stochastic gradient descent is used to optimize the parameters:

Errors:

$$e_{ui} = r_{ui} - q_i^T p_u$$ (3)
Optimize $q_i$ and $p_u$ according to gradient descent:

$$q_i \leftarrow q_i + \gamma (e_u p_u - \lambda q_i)$$

$$p_u \leftarrow p_u + \gamma (e_u q_i - \lambda p_u)$$

Among them, $\gamma$ is the learning rate. Because of user preference issues, the existence of these factors seriously affects the accuracy of the final results.

### 2.2. Combination of deep learning and BiasSVD algorithm

This article mainly aims at improving the formula (1) of the BiasSVD algorithm. Instead of using formula (1) to represent the prediction scoring item, it uses the deep learning neural network to output the prediction scoring item. The deep learning neural network part has 5 layers, which are divided into embedding layer, convolutional layer, pooling layer, fully connected layer, layer 1 and layer 2. Enter the encoded user and item information data in the embedding layer, such as user age, user occupation, user ID, user gender, movie name, movie year, movie classification, user evaluation and other information. One-hot encoding is not used here. Because it will cause the word vector dimension to be very large, and there is no connection between upper and lower words. This paper uses the embedding layer of Keras neural network, which maps the words represented by a dense vector to a continuous vector space. Keras's embedding layer updates the weights based on the information of the tags, so as to achieve the purpose of higher supervised learning, and it will also learn the relationship between words. Although Keras is not as efficient as Word2vec, it can be used as a part of deep learning to facilitate the implementation of the algorithm in this paper. The convolutional layer, the pooling layer, and the first fully connected layer are mainly used to process some text information of users and items, such as movie names, user reviews, etc. These text information are emotional and have a great influence on users’ choices. Need to use text convolutional neural network for deep extraction, in order to play a better role in text classification. Especially the activation function in the first layer network adopts ReLU. The ReLU function has the advantages of overcoming the disappearance of the gradient and speeding up the training speed. $U$ represents the user information output through the fully connected layer 2 and $I$ represent the item information output through the fully connected layer 2. Then the user and item matrix output by the fully connected layer 3 is $\hat{r}_{ui}$, and the mathematical expression is:

$$\hat{r}_{ui} = \text{conlayer}(U, I)$$

Finally, optimized by Adam algorithm:

$$n \sum_{(u, i) \in l} (r_{ui} - \mu - b_i - b_u - \hat{r}_{ui})^2 + \lambda (\|b_u\|^2 + \|b_i\|^2)$$

After getting the movie feature matrix and user feature matrix, use Top-N to recommend.

### 3. Experiment analysis

#### 3.1. Data analysis

This experiment uses the MovieLens dataset. The MovieLens dataset contains ratings data of multiple users on multiple movies, as well as movie metadata information and user attribute information. This data set is often used as a recommendation system, a test data set of machines learning algorithms, especially in the field of recommendation systems. Many recommendation algorithm literatures are based on this data set. When performing statistical analysis on the entire MovieLens data set, it is found that user ratings show an increasing trend with age, which indicates that the older the age, the lower the rating standard for movies on this data set. In order to make the algorithm more accurate in the sentiment analysis of users, the hidden feature of age is added when extracting features.
In the statistical analysis between movie ages and user ratings, it is found that user ratings show a decreasing trend with the increase of movie ages, which indicates that a large number of users on the entire MovieLens data set may prefer older movies. These are all factors that affect user ratings, so this implicit feature is also added when evaluating the algorithm.

3.2. Experimental results and analysis

The experimental results of this algorithm are compared with those of seven commonly used algorithms.

As can be seen from Table 1, the RMSE of this algorithm on the Movielens-1m dataset improves by about 19 over the most advanced LSTM. Over 75%, MAE about 17. More than 82%. In the table, the calculation method of lifting ratio is: (LSTM algorithm value - the algorithm value in this paper)/LSTM algorithm value.

| Arithmetic. | RMSE   | MAE   |
|-------------|--------|-------|
| SVD         | 0.8730 | 0.686 |
| SVD++       | 0.8620 | 0.673 |
| CDL         | 0.8879 | 0.691 |
| MLP         | 0.8773 | 0.687 |
| ConvMF      | 0.8549 | 0.676 |
| NeuMF       | 0.8631 | 0.674 |
| LSTM        | 0.8480 | 0.668 |
| Algorithm in this paper. | 0.6805 | 0.5489 |

The minimum performance improvement ratio of this algorithm is /%. 19.75 17.82

In Table 2, this paper has processed the tags of interactive information between users and items. As a result, compared with LSTM algorithm, RMSE of the algorithm in this paper has been improved by 35.6% and MAE of 47.4%, it can be concluded that the interaction information between users and items plays a very important role in recommending to users.

| Arithmetic. | RMSE   | MAE   |
|-------------|--------|-------|
| SVD         | 0.8135 | 0.631 |
| SVD++       | 0.8108 | 0.629 |
| CDL         | 0.8186 | 0.637 |
| MLP         | 0.8143 | 0.632 |
| ConvMF      | 0.7930 | 0.606 |
| NeuMF       | 0.8107 | 0.674 |
| LSTM        | 0.7882 | 0.601 |
| Algorithm in this paper. | 0.5073 | 0.316 |

The minimum performance improvement ratio of this algorithm is /%. 35.6 47.4

4. Conclusions

Traditional matrix decomposition algorithm for simply the inner product of matrix with users, the item cannot be effectively extract the user and item characteristics, convolution matrix decomposition ConvMF LSTM successively with both short-term and long-term memory neural network for its improvement, but the understanding of users and items document context is not enough, lead to poor rating accuracy, this paper put forward the term embedded combination of convolution neural network feedforward neural network and text to extract the characteristics of the methods to improve the BiasSVD algorithm. Extensive experimental comparison on two Real data sets of MovieLens shows that
the algorithm rating prediction in this paper is more accurate, which verifies that users' interaction information on items can better reflect users' preferences. Therefore, the extraction of user and item information based on deep learning is still worth further study.

References
[1] Tao Xue-jiao, HU Xiao-feng, LIU Yang. Big Data Research Review [J]. Journal of System Simulation, 2013, 25(S1): 142-146.
[2] YU S, YANG M, QU Q, et al. Contextual-boosted deep neural collaborative filtering model for interpretable recommendation [J]. Expert Systems with Applications, 2019, 136: 365-375.
[3] KOREN Y. Collaborative filtering with temporal dynamics [J]. Communications of the ACM, 2010, 53(4): 89-97.
[4] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems [J]. Computer, 2009, 42(8): 30-37.