Data Rate Prediction using DCCR System in Data Mining

[1]D. Saritha Reddy [2]N. Naveen

[1]Assistant Professor, Dept. of Master of Computer Applications, Narayana Engineering College, Gudur.

[2]PG Scholar, Dept. of Master of Computer Applications, Narayana Engineering College, Gudur.

Abstract- Recently, collaborative filtering combined with various kinds of deep learning models is appealing to recommender systems, which have shown a strong positive effect in an accuracy improvement. However, many studies related to deep learning model rely heavily on abundant information to improve prediction accuracy, which has stringent data requirements in addition to raw rating data. Furthermore, most of them ignore the interaction effect between users and items when building the recommendation model. To address these issues, we propose DCCR, a deep collaborative conjunctive recommender, for rating prediction tasks that are solely based on the raw ratings. A DCCR is a hybrid architecture that consists of two different kinds of neural network models (i.e., an auto encoder and a multi layered perceptron). The main function of the auto encoder is to extract the latent features from the perspectives of users and items in parallel, while the multi layered perceptron is used to represent the interaction between users and items based on fusing the user and item latent features. To further improve the performance of DCCR, an advanced activation function is proposed, which can be specified with input vectors. The extensive experiments conducted with two well-known real-world datasets and performances of the DCCR with varying settings are analysed. The results demonstrate that our DCCR model outperforms other state-of-art methods. We also discuss the performance of the DCCR with additional layers to show the extensibility of our model.

Keywords: Recommender Systems, Collaborative Filtering, Rating Prediction, Denoising Autoencoders, Multi Layered Perceptron.

1. INTRODUCTION

Recommender systems [1] are essential for the success of many online applications. Considering online shopping websites as an example, numerous goods are provided by these shopping sites, and users browse all the information about all the goods in a short time. In this context, recommender systems, as one kind of effective information filtering tool, not only can help users to obtain more valuable advice by filtering the redundant information but also gradually increase the sales volume of the websites. As a result, recommender systems have already been integrated in some large-scale websites (e.g., Amazon), which continually service thousands of people.

In the past few decades, teams of researchers have spent a considerable amount of effort on recommender design and have achieved great results. Collaborative filtering (CF) [1] is one of the great inventions in recommendation research that has been successfully used for industry applications. In contrast to traditional CF recommenders that depend on calculating the similarity between users/items with similar preferences, matrix factorization (MF) is a popular CF recommender for rating prediction tasks [11]. MF decomposes the original rating matrix R into two low-rank matrices, which represent the latent feature space of users and items. Due to their effectiveness, variants of the MF method have been proposed [2], [9]. Recently, with the success of deep neural networks, the combination with deep learning methods is a new breakthrough for recommenders.

To address these issues, this paper proposes a deep collaborative conjunctive recommender (DCCR). We focus on the rating prediction task of a recommender and format it as a regression problem. By taking advantage of deep learning and traditional CF methods, we try to extract latent features from users' explicit ratings to items without any additional information. We first present some related studies of deep learning methods and traditional methods and analyze the process of rating prediction. We then describe the feature representation with users and items and explain the model that we propose. Some techniques that we apply and the working mechanism are detailed for better features extraction with consideration of the interaction between the users and items. Particular experimental settings and processes are defined to obtain reasonable results. We test the results with the root mean squared error and the mean absolute error. To obtain the optimal results, some experiments are conducted with varying parameters settings. The experimental results are illustrated to understand the impacts of many factors. We also compare the accuracy of our proposed method with other related and recent methods.

The main contributions of this project include:

1) We present a novel recommender model that extracts deep inner features of both users and items that solely depend on the explicit ratings and extract the interaction features. We describe the details of the structure, input vector, loss function and training techniques, which are indispensable for the experiments.

2) We investigate the impacts of the parameters of the proposed model and analyze the relations of these parameters on the prediction accuracy. We also provide possible measures for improving the results from different perspectives. An improved activation function for our neural networks are proposed, which can be specified with input vectors.

3) By conducting considerable experiments on two datasets, the results show that the proposed model can achieve better accuracy for this particular rating prediction task. We also discuss the expandability of our model by analyzing the depth of neural networks. Several methods are proposed to adjust the gradient problem of the deep neural networks.

2. RELATED STUDIES

Collaborative filtering (CF) has been widely used to provide users with new products and services in many industrial applications. CF provides users with products from similar users or chooses similar products from users' favorite products. The matrix factorization model is the most important method of CF and has been explored by many researchers. Among different kinds of MF models, the latent factor model (LFM) is the most popular model for rating prediction tasks. The LFM factorizes the rating matrix R into two low-latent factor matrices. However, the manual process of feature extraction consumes manpower and financial resources. Recently, deep learning methods have shown that the neural networks have the powerful ability to automatically learn the features from heterogeneous data and gain reasonable results for most tasks [4], [20]. Therefore, to achieve the goal of improving the prediction accuracy by learning the deep inner user/item features, CF combined with the neural network's methods have been proposed.

As one of the most effective deep learning methods in recommendation, autoencoders have been discussed in several papers [12]–[15]. Autoencoder is an unsupervised learning method that can automatically compress the input features to a low dimension, which has shown absolute advantages in feature extraction compared with traditional methods. Different kinds of autoencoders have been proposed for different scenarios, such as denoising autoencoders (DAE) [4], [5], marginalized autoencoders [6], [12].

To address this issue, multi layered perceptron—another neural network model—has been applied to many industry recommender systems. Multi layered perceptron combines the features of users and items, which have been extracted from neural networks to achieve better recommendation. But most of these methods are focus on the content processing, such as reviews. Normally, the reviews of users and items are employed as input data and a joint deep model are built to merge the features. Some works apply co-attention mechanisms to learn a distributed representation from user and item reviews.

In addition to the text and categorical data as side information, MLP with its derivatives can also extract features from media data, such as text content for recommendation [5]. Side information can be easily attained for a commercial company, but some information is highly sensitive. As we can see, these studies have more stringent requirements for the input data and are not appropriate for every dataset, which indicates that a substantial amount of effort should be spent on the adjustment of inputs. Besides, these related works usually focus on the interactions but not extract the feature from user and item separately which means the feature representation of user and item can be affected by each other.

Thus, we propose a new model for recommender systems that can separately represent the users and items features and merge them to make predictions more accurate. We focus on the feature representation, algorithm and model design to make better recommendations instead of using more information from specified datasets. With some techniques, this new model can precede all related studies of different dataset.

3. ARCHITECTURE OF DCCR MODEL

The proposed DCCR model. To take advantage of the deep learning model in terms of deeper inner feature extraction and fusion, the DCCR is a hybrid architecture that consists of two different kinds of neural network models (i.e., DAE and MLP). DAE extracts user and item deeper latent features for the raw ratings data in a separate way, while the main function of an MLP is to merge the user and item feature from the results of the DAE at the first layer and extract the higher features (i.e., relationships between user and item) based on the combined user and item features. The architecture of the proposed model for the rating prediction task is shown in Fig. 1.

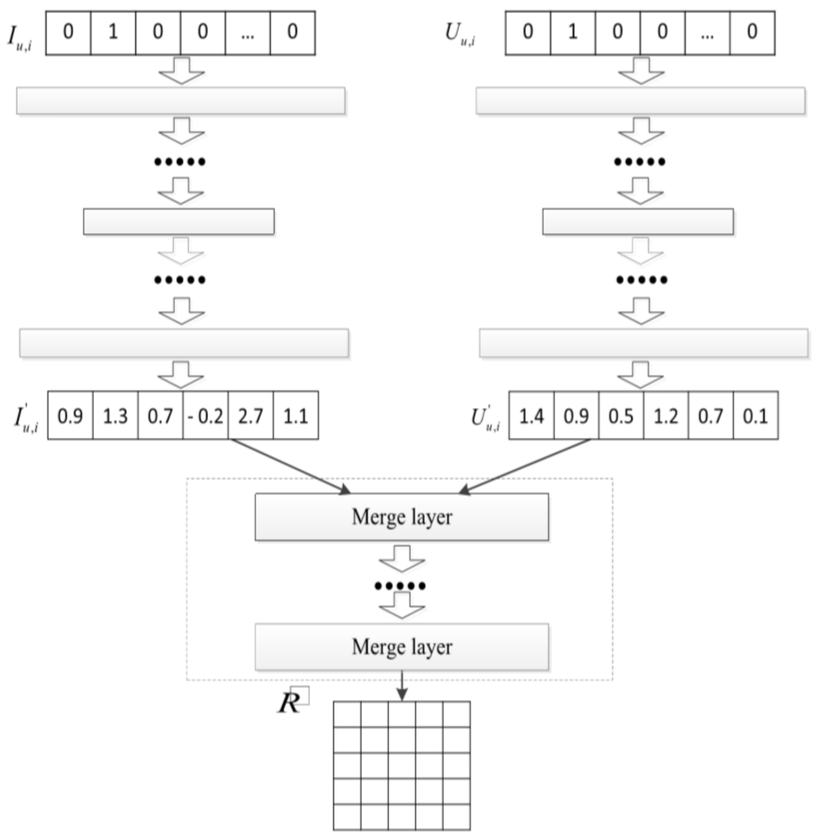

Fig 1. Rating prediction of the DCCR model

This system model consists and implements the following modules:

Admin: In this module, admin login to the system, he can view the user's information, add posts, view all posts with ratings and he collect the rating data and check the recommended posts by collaborating filtering.

User: In this module, user can register and login to the system. He searches the post and provide rating and comments on that post. And view the recommended items based their ratings. He searches friends and send request to another user.

Feature Representation

Raw data collected from industrial applications are noisy with a substantial amount of useless information and cannot be directly applied for recommender tasks. Thus, feature representation is an important step in our study. Based on the theory of collaborative filtering, we define our item and user feature representation. We do not use the extra information but only use the rating data since we try to use the least amount of information to make better recommendations in the rating prediction task.

Feature Extracting with Autoencoders

- An autoencoder is selected to extract the latent feature from raw rating data. Each branch of the first part of the DCCR is a complete denoising autoencoder with all layers and differs from existing methods that only use the output of hidden layers.
- Autoencoders usually have three layers: the input layer, the hidden layer and the output layer.
- To force the neural network to learn features from the input, the number of output layer units is equivalent to the number of input layer units, and the number of hidden layer units is less than the number of input layer units.
- We consider that the output vectors of an autoencoder can be another way to express the implicit features.

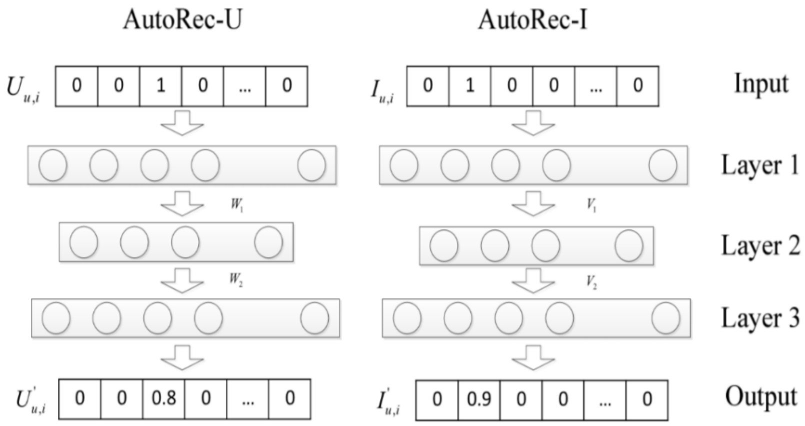

Fig 2. Rating prediction of autoencoders. AutoRec-U and AutoRec-I separately use the user features and item features as input vectors. These two models have different dimensions of input and parameter size and accuracy of predication.

4. EXPERIMENTS AND RESULTS

In this section, we present our experiments in detail. Experiments on several conditions show the key factors in our proposed model. For all the experiments, we run5-foldcross validation and take the average to report the results.

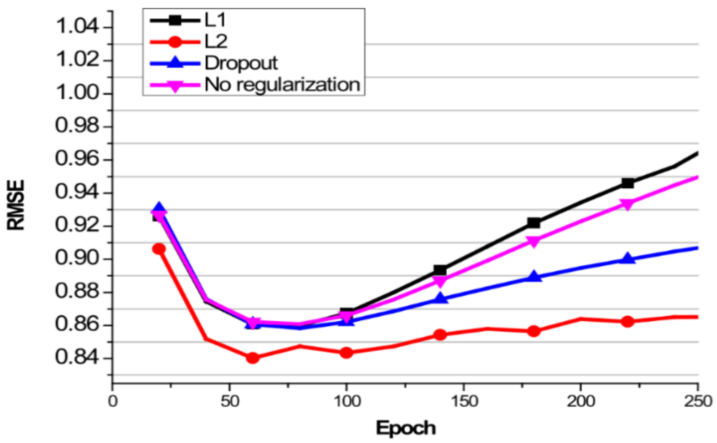

Fig 3. Impact of different

regularization term. The red line demonstrates that the L2 regularization term is better than the others.

To find the most effective regularization method, in this set of experiments, we compare the of the DCCR model with different types of regularization and non-regularization. The results are shown in Fig. 3. From the results in Fig. 3, we discover that the L2 regularization outperforms the L1 regularization. Without the regularization term, the training process of the model has a serious overfitting problem, which causes the model to have the worst generalization ability. We perform drop out regularization. The results with dropout are not better than the L2 regularization but are better than the non-regularization and L1 regularization.

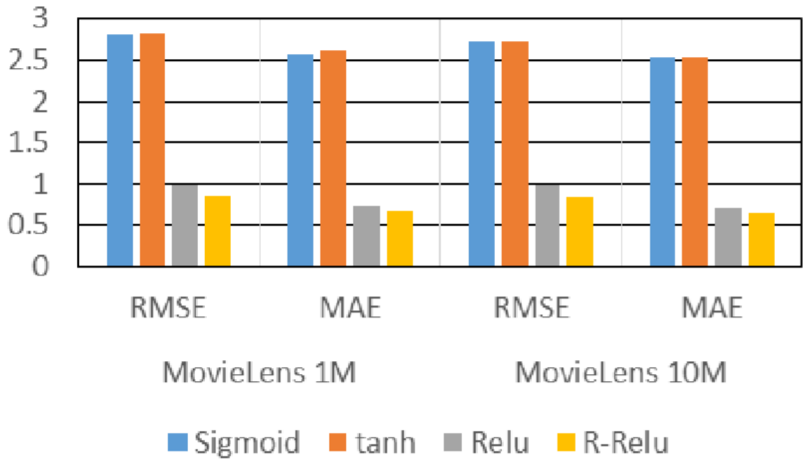

Fig 4. The performance of different activation function.

To prove the effectiveness of the R-Relu, we compare the results of some other activation functions including sigmoid, tanh and Relu, which is shown as Fig.4. The experiments are conducted without any other techniques to show the effects of different activation functions. From the results, we can see the R-Relu can outperform the other activation functions.

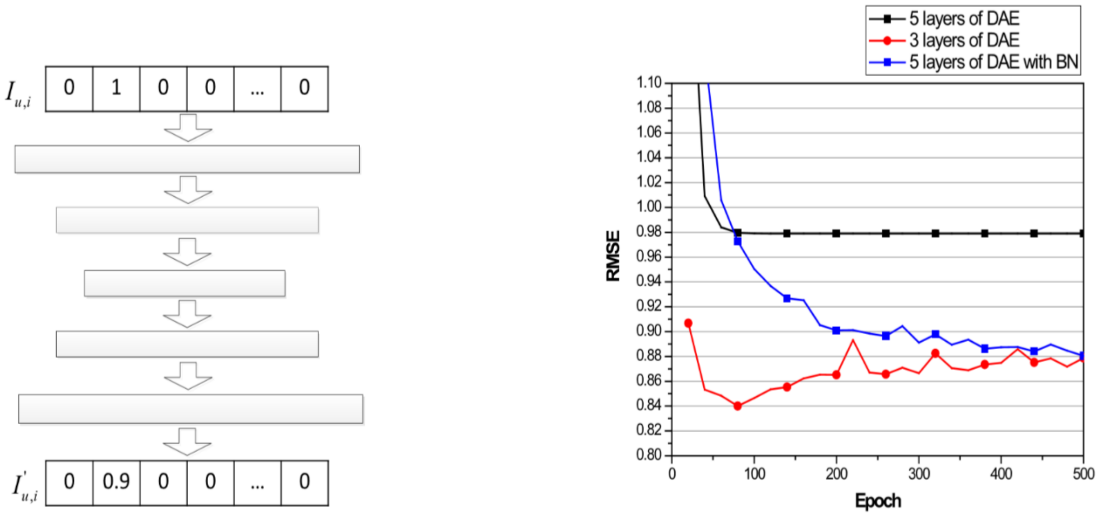

Fig 5. Multilayer autoencoders for rating prediction of recommendations (left). Results of multilayer autoencoders (right).

In our model, we design the basic structure with three hidden layers for the feature extraction part, and one layers for the fusion part. Five-layer autoencoder can be designed as Fig.5 (left) in DCCR. This situation causes network degradation, which indicates that deeper neural networks learn even less than shallow networks. Without exception, we stack the layers to predict the ratings and acquire inferior accuracy, which is shown in Fig. 5 (right).

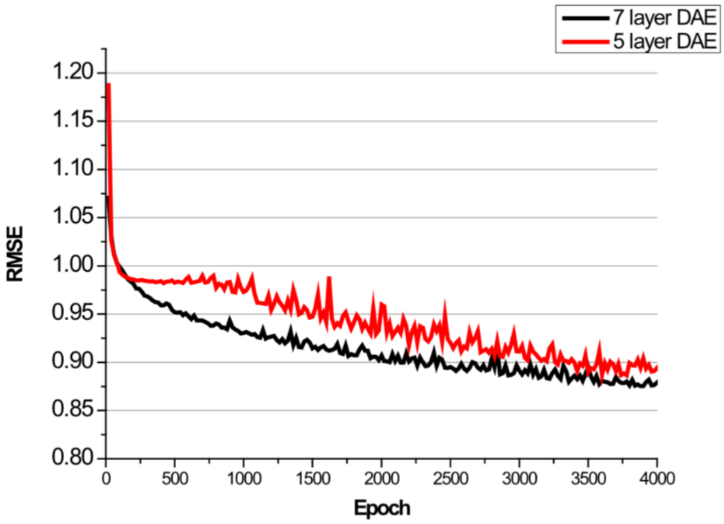

Fig 6. The comparison of RMSE of DAEs with different layers.

Fig. 6 (right) shows the results of 5 layers of DAE with BN, which is substantially better than the model without BN. We also test the results with 7 layers, which is shown in Fig. 6. Our experiments show that deep networks can assist the rating prediction when the vanished gradients have been alleviated. With a gradual increase in the networks' depth, the training process will be slower than the shallow networks, require more time and have higher requirements for hardware. This approach is a trade-off between accuracy and efficiency. We defer this issue for future research.

Our experiments show that deep networks can assist the rating prediction when the vanished gradients have been alleviated. With a gradual increase in the networks' depth, the training process will be slower than the shallow networks, require more time and have higher requirements for hardware. This approach is a trade-off between accuracy and efficiency. We defer this issue for future research.

5. CONCLUSIONS

Collaborative filtering has shown to be effective in commercial recommender systems. By combining with neural networks, CF can represent the latent features of users and items without a manual setting. However, most of related studies use a single model with a common activation function to perform a rating prediction ask without considering the traits of features and ratings. In this paper, we propose a hybrid neural network model for rating prediction that is named the deep collaborative conjunctive recommender (DCCR). This model integrates the spirits of several neural networks to separately capture the latent features from users and items and describes the interactions between these features. Solely using the explicit ratings from the data, we design this end-to-end model to improve the accuracy of rating prediction. Numerous factors affect the prediction performance. Thus, to achieve the optimal model, we evaluate the DCCR with varying factor settings by considerable contrast experiments. The results show that our DCCR model outperforms other state-of-the-art methods using two real-world datasets. We also prove that the DCCR with additional layers has a positive effect on accuracy improvement.

REFERENCES

1. M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, ''Collaborative filtering recommender systems,'' Found. Trends Hum.-Comput. Interact., vol. 4, no. 2, pp. 81–173, 2011.
2. R. Kumar, B. K. Verma, and S. S. Rastogi, ''Social popularity based SVD++ recommender system,'' Int. J. Comput. Appl., vol. 87, no. 14, pp. 33–37, 2014.
3. A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''Imagenet classification with deep convolutional neural networks,'' in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.
4. Y. Jhamb, T. Ebesu, and Y. Fang, ''Attentive contextual denoising autoencoder for recommendation,'' in Proc. ACM SIGIR Int. Conf. Theory Inf. Retr. New York, NY, USA: ACM, 2018, pp. 27–34
5. X. Cai, J. Han, and L. Yang, ''Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation,'' in Proc. 32nd AAAI Conf. Artif. Intell., 2018, pp. 1–8.
6. Y.Peng, S.Wang, and B.-L.Lu, ''Marginalized denoising auto encoder via graph regularization for domain adaptation,'' in Proc. Int. Conf. NeuralInf. Process. Berlin, Germany: Springer, 2013, pp. 156–163.
7. P.Baldi and P.J.Sadowski, ''Understanding dropout,'' in Proc. Adv. Neural Inf. Process. Syst., 2013, pp. 1–9.

Author's Profile:

D. Saritha Reddy has received her M.Tech degree in Computer Science from Acharya Nagarjuna University, Guntur in 2010 and pursuing Ph.D with area Computer Networks. At present she is working as Assistant Professor in Narayana Engineering College, Gudur, Andhra Pradesh, India.

N. Naveen has Received his B.Sc Degree in Computer Science from Vidyalaya Degree College, Gudur affiliated to Vikrama Simhapuri University, Nellore in 2016 and pursuing PG Degree in Master of Computer Applications (M.C.A) from Narayana Engineering College, Gudur affiliated to JNTU Anantapur, Andhra Pradesh, India.

