# Deep Learning Based Recommendation: A Survey

Juntao Liu[1(✉)] and Caihua Wu[2]

[1] 709th Research Institute, China Shipbuilding Industry Corporation, Wuhan, China
prolay@163.com
[2] Section of Automatic Command, Huang Pi NCO School, Air Force Early-Warning Academy, Wuhan, China
wucaihua2009@163.com

**Abstract.** Due to the great success, deep learning gains much attentions in the research field of recommendation. In this paper, we review the deep learning based recommendation approaches and propose a classification framework, by which the deep learning based recommendation approaches are divided according to the input and output of the approaches. We also give the possible research directions in the future.

**Keywords:** Deep learning · Recommendation · Neural network

## 1 Introduction

In recent years, deep learning, as a kind of machine learning approach, is applied in many different research domains successfully, such as computer vision, speech recognition, natural language processing and so on. In these fields, compared with traditional approaches, deep learning based approaches improve the performance remarkably. Due to the great success of deep learning, some researchers try to use deep learning in recommendation systems [1–11, 13–15], and wish these newly proposed model can improve the performance of the recommendation systems, just as the deep learning models do in other research fields.

In this paper, the deep learning based recommendation approaches are reviewed under the proposed classification framework for this kind of approaches. In this framework, the deep learning based recommendation approaches are classified by the input and the output of the approaches. Introducing deep learning in the recommendation systems is a related new research direction. Several approaches are proposed, but few is used in practice. Scalability is a problem in the recommendation systems, in which there are huge number of items and users. According to this fact, we also point out the future research direction of the deep learning based recommendation.

This paper is organized as follows. First, the background of recommendation and deep learning are introduced in Sect. 2. In Sect. 3, the classification framework for the deep learning approaches are introduced, and these approaches are also reviewed. The future research directions are given in Sect. 4 followed by the conclusion in Sect. 5.


K. Kim and N. Joukov (eds.), *Information Science and Applications 2017*,
Lecture Notes in Electrical Engineering 424, DOI 10.1007/978-981-10-4154-9_52

## 2 Background

### 2.1 Recommendation

Traditionally, recommendation systems are divided into three kinds: content-based, collaborative and hybrid recommendation approaches, in the view of the recommendation approaches [17]. Recommendation systems are also classified according to the application fields [18]. In this paper, we review the recommendation systems in the view of the underlying recommendation problems. In other word, we divide the recommendation systems according to the input and output of the recommendation problems addressed in the systems.

Recommender systems usually collect the users' activities in the systems, including rating, clicking, buying, comment and so on, which are the input of the recommendation approaches. Rating is the most used input in recommendation systems, in which users are allowed to rate the item by a k-point integer. Recommendation results are generated according to the ratings. The typical approaches are PMF [19] and its extensions. Usually, rating is regarded as a kind of explicit feedback. Some approaches take implicit feedback, such as clicks, view and so on, as input. Compared with explicit feedback, implicit feedback can be collected easily by the recommendation systems. BPR [20] infers the binary preference relation of a user between items from the implicit feedback of this user, but it can handle only one kind of implicit feedbacks. Instead of learning from only one kind of feedback, some recommender systems lavage multi-kind of feedback [6, 21]. Wu et al. [6] also use the time information of the feedback in their approach.

The output of recommender systems is the recommendation result given to the users. Different recommender systems give different kinds of recommendation results. Some systems predict the ratings that the users have not issued. Matrix factorization based approaches are the most popular approaches resolving the rating prediction problem. Some recommendation systems predict the preference order of the uses among the items, which are called ranking based recommendation approaches. This kind of approaches include BPR [20], ListPMF [22] and QPMF [23]. In some cases, users may expect a combination of products, such as a jacket and the matched pant. It is raised the combination recommendation problem. [24, 25] address this problem in the field of cloth recommendation.

### 2.2 Deep Learning

Nowadays, deep learning refers to class of machine learning algorithm. Usually, the model of deep learning contains a cascade of nonlinear transformation layers. The parameters in the models are learned by end-to-end optimization.

Several kinds of deep learning models are proposed. One of the most used model is the feedforward neural network. A typical feedforward neural network with only one hidden layer is shown in Fig. 1(a). The feedforward neural networks used in deep learning usually have several hidden layers with different dimensions in order to encode the input in high level abstractions. The output $O$ in Fig. 1(a) is computed as follows,

$$h = \sigma(Wx)$$
$$O = \delta(Vh) \tag{1}$$

where $V$ and $W$ are the weight matrices for hidden layer to output layer and input layer to hidden layer. Functions $\delta$ and $\sigma$ are the nonlinear transformation functions, such as tanh and sigmoid function.

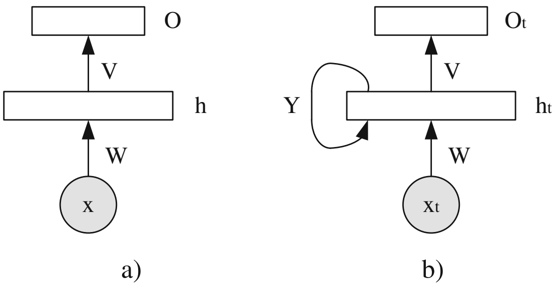


**Fig. 1.** Structure of feedforward neural network (a) and recurrent neural network (b).

Because the feedforward neural network can't analyze sequence data, such as natural language, recurrent neural network is designed. In this model, the hidden layers are connected recurrently to the input layers, as shown in Fig. 1. Structure of feedforward neural network (a) and recurrent neural network (b). The output at time t, Ot, is calculated as follows,

$$h_t = \sigma(Wx_t + Yh_{t-1})$$
$$O_t = \delta(Vh_t) \tag{2}$$

where Y is the weight matrix for previous hidden layer to current hidden layer.

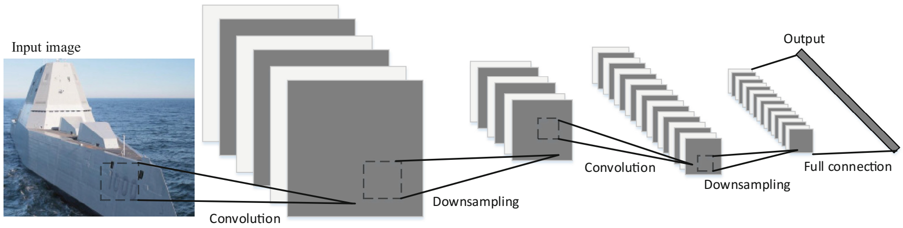


**Fig. 2.** Typical convolutional neural network for image encoding.

Convolutional neural networks are widely used for image recognition, and it contains one or more convolutional layers where the neurons are tiled in such a way that they respond to overlapping regions in the visual field. Between two neighbor convolutional layers there is always a pooling layer for subsampling. A typical convolutional neural network is shown in Fig. 2. Typical convolutional neural network for image encoding..

Because the parameters in CNN are much fewer than other deep learning model, it is easier to train. This makes them a highly attractive architecture.

# 3 Classification of Deep Learning Based Recommendation Methods

## 3.1 Classification Framework

In this paper, the deep learning based recommendation approaches are classification by the input and output of the approaches. In the aspect of input, some approaches take the content information into consideration and some approaches do not using this kind of information. In the view of output, user ratings and user activity sequence are predicted, respectively. Our classification frame work is shown in Table 1.

**Table 1.** Classification framework of deep learning based recommendation

| Output | Input | |
|---|---|---|
| | Approaches using content information | Approaches without content information |
| Rating | [1, 2, 14] | [7, 8, 10, 15] |
| Order | [3, 4, 11, 13] | [5, 6] |

## 3.2 Classification by Input

**Approaches Using Content Information.** Oord et al. [1] propose a music recommendation method based on deep neural network. In this method, the user and item latent feature vectors is learned first by weighted matrix factorization (WMF) algorithm [16]. And then the item latent feature vectors are learned by a deep convolutional network further, in which the already learned user latent feature vectors and the music signal are taken as input. The objective functions used to train the neural network are the weighted rating prediction error (WPE) just as that in WMF and the difference between the item latent feature vectors learned by WMF and those learned by the neural network. Finally, recommendation is made as in standard MF i.e. inner product between user and item latent factors. Because the item latent feature vectors are learned from item content, this method is specifically useful in Item cold start situation, where no feedback on target item is available. This method was later used at Spotify in an experiment [2].

Matrix factorization and deep belief network (DBN) are integrated in [14]. In this model, the item (music) latent feature vector is the output of DBN, which is trained previously according to the content information of the music.

Also in the music domain, Hamel et al. [3] designed two neural network models to predict music tags. The predicted tags are used for music recommendation. The input of these models is the preprocessed music feature, which is obtained by discrete Fourier transform (DFT), mel-compression and principal component analysis whitening (PCA). The author investigate the performance of several pool functions and find that combining several pool functions can improve the performance.

Elkahky et al. [4] propose a Deep Learning approach to map users and items to a latent space where the similarity between users and their preferred items is maximized. They learn the item feature learn from different domains. And the user feature is learned user features by a multi-view Deep Learning model. The proposed method is test for Windows Apps, news and Movie/TV recommendation.

Wang et al. [11] propose Bayesian stacked denoising autoencoder (SDAE) [12], and integrate this model with Bayesian probabilistic matrix factorization (BPMF), which is called collaborative deep learning (CDL), to address the problem of implicit feedback recommendation. In their method, the latent item feature is first generated by a previous trained SDAE model according to the item content information. And then, user latent feature vector and rating issued by this user are generated in the way similar to that in BPMF.

There are two neural network models in the recommender system of YouTube [13]. One model is used to generate hundreds of candidate items from huge number of items according to the historical behaviors of the users. The candidate items are then ranked by the other model according to the historical user behavior, context information and item feature.

**Approaches Without Content Information.** Some approaches do not use content information. In these method, only the feedback of the users, such as ratings, clicking and so on, are used to generate the recommendation results. For example, [7, 8] use the observed rating to predict the ratings of the items that the user have not accessed. In [5], implicit feedback is used to generate the ranked item orders. And [6] take multiple feedback as the input of the deep learning model. The input of [15] is the user-tag matrix.

### 3.3 Classification by Output

**Item Order Prediction Methods.** As mentioned above, the approaches using content information, such as [3, 4, 11, 13], train deep learning models and rank the items. The ranked item orders are recommended to the users.

Some approaches do not use content information. Hidasi et al. [5] propose a session-based recommendation method using recurrent neural network (RNN). In their method, the user accessed items are treated as sequences. The predicted item sequences that the users may accessed is generated by the trained RNN model in the end-to-end manner. There are embedding layer, feed forward layer and several GRU layers in the proposed model. The authors find that pair-wise loss function is better than point-wise loss function. And the model with single GRU layer is better than that with several GRU layers.

Wu et al. [6] propose a recurrent neural network based recommendation approach (RNNRec) to address the problem of time heterogeneous feedback recommendation. In this work, historical feedback activities with time stamps of the users are treated as sequences. And a recurrent neural network is trained using these feedback sequences. It is reported that the recommendation results generated RNNRec are more accurate than those generated by the traditional recommendation methods.

**Rating Prediction Methods.** Content information are used in [1, 2, 14] to predict the ratings.

There are approaches only using user feedback, such as raing and tag, to generate recommendation. Salakhutdinov et al. [10] use RBMs for collaborative filtering. RBMs can be used as a fundamental units of Deep Neural Networks. But in [10], there is only a single layer in RBM. Additionally, Edwin Chen walks through a more basic use of RBM's for collaborative filtering in this blog post: Introduction to Restricted Boltzmann Machines.

Zhang et al. [7] propose a deep learning model to predict the ratings. The input of the model is the concatenation of the embedding feature vectors of user and item. There is only one hidden layer in the model, and output of the model is the predicted rating. Zheng et al. [8] use Neural Autoregressive Distribution Estimator (NADE) [9] model to address recommendation problem. The model is modified to share parameters among ratings. And to scale to large dataset, a factorizing version model is proposed inspired by RBM [10]. In this work, authors also present a list-wise loss function. Zou et al. [15] use stacked autoencoders in tag-aware recommender systems. The user latent feature vectors are generated by the stacked autoencoders according to the user-tag matrix. Recommendation results are obtained through aggregating the user latent feature vectors and item-user rating matrix.

## 4 Future Research Directions

One of the problems of the deep learning based recommendation approaches is scalability for recommendation systems, in which there are huge number of items and users. And user feedback are collected every second. The performance of the recommendation approaches is importance in this circumstance. In other hand, training the deep learning model is time-consuming. So how to improve the scalability of the deep learning based recommendation approaches is an importance issue in the future research.

Another possible research direction is to design new kinds of deep learning model to solve special problems in recommendation. RNN and feedforward neural network are used in the existing deep learning recommendation approaches. Convolutional neural network (CNN) is seldom used in recommendation. Maybe it can get good results for some recommendation problems.

## 5 Conclusion

Recently, due to the great success of deep learning, several researchers propose to use deep learning approach in recommendation systems. In this paper, the deep learning based recommendation approaches are reviewed and classified by the input and output of the approaches. It is found that the most used deep learning models are feedforward neural network and recurrent neural network. Convolutional neural network is seldom used. It is prompted that CNN based recommendation approaches is the possible research direction. Deep learning can improve the accuracy of the recommendation systems, but scalability is a critical problems for the huge number of items and users in the systems. So improving efficiency of the deep learning based recommendation approaches is the main work in this field.

**Acknowledgments.** This work is supported by National Natural Science Foundation of China (No. 61403350 and No. 61401228).

## References

1. van den Oord, A., Dieleman, S., Schrauwen, B.: Deep content-based music recommendation. In: NIPS (2013)
2. Recommending music on Spotify with deep learning (2014). http://benanne.github.io/2014/08/05/spotify-cnns.html
3. Hamel, P., Lemieux, S., Bengio, Y., Eck, D.: Temporal pooling and multiscale learning for automatic annotation and ranking of music audio. In: Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), pp. 729–734 (2011)
4. Elkahky, A., Song, Y., He, X.: A multi-view deep learning approach for cross domain user modeling in recommendation systems. In: Proceedings of the 24th International Conference on World Wide Web (WWW 2015), pp. 278–288 (2015)
5. Hidasi, B., Karatzoglou, A., Baltrunas, L., Tikk, D.: Session-based recommendations with recurrent neural networks. In: ICLR 2016 (2016)
6. Caihua, W., Wang, J., Liu, J., Liu, W.: Recurrent neural network based recommendation for time heterogeneous feedback. Knowl.-Based Syst. **109**, 90–103 (2016)
7. Junlin, Z., Heng, C., Tongwen, H., Huiping, X.: A Distributional Representation Model For Collaborative Filtering (2015). https://arxiv.org/abs/1502.04163
8. Zheng, Y., Tang, B., Ding, W., Zhou, H.: A neural autoregressive approach to collaborative filtering. In: Proceedings of the 33rd International Conference on Machine Learning (2016)
9. Larochelle, H., Murray, I.: The neural autoregressive distribution estimator. In: International Conference on Artificial Intelligence and Statistics, pp. 29–37 (2011)
10. Salakhutdinov, R., Mnih, A., Hinton, G.: Restricted boltzmann machines for collaborative filtering. In: Proceedings of the 24th International Conference on Machine Learning, pp. 791–798. ACM (2007)
11. Wang, H., Wang, N., Yeung, D.-Y.: Collaborative deep learning for recommender systems. In: KDD 2015 (2015)
12. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A.: Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. JMLR **11**, 3371–3408 (2010)
13. Covington, P., Adams, J., Sargin, E.: Deep neural networks for YouTube recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems, pp. 191–198 (2016)
14. Wang, X., Wang, Y.: Improving content-based and hybrid music recommendation using deep learning. In: Proceeding of ACM MM 2014 (2014)
15. Zuo, Y., Zeng, J., Gong, M., Jiao, L.: Tag-aware recommender systems based on deep neural networks. Neurocomputing **204**, 51–60 (2016)
16. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (2008)
17. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)

18. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Expert Syst. Appl. **39**(11), 10059–10072 (2012)
19. Salakhutdinov, R., Mnih, A.: Probabilistic matrix factorization. In: Proceedings of Advances in Neural Information Processing Systems, NIPS 2008, pp. 1257–1264 (2008)
20. Gantner, Z., Drumond, L., Freudenthaler, C., Schmidt-Thieme, L.: Bayesian personalized ranking for non-uniformly sampled items. In: Proceedings of Knowledge Discovery and Data Mining (KDD) Cup and Workshop 2011 (2011)
21. Pan, W., Zhong, H., Xu, C., Ming, Z.: Adaptive Bayesian personalized ranking for heterogeneous implicit feedbacks. Knowl.-Based Syst. **73**, 173–180 (2015)
22. Liu, J., Wu, C., Xiong, Y., Liu, W.: List-wise probabilistic matrix factorization for recommendation. Inf. Sci. **278**, 434–447 (2014)
23. Liu, W., Wu, C., Feng, B., Liu, J.: Conditional preference in recommender systems. Expert Syst. Appl. **42**(2), 774–788 (2015)
24. McAuley, J., Targett, C., Shi, Q., Hengel, A.: Image-based recommendations on styles and substitutes. In: SIGIR 2015 (2015)
25. Jagadeesh, V., Piramuthu, R., Bhardwaj, A., Di, W., Sundaresan, N.: Large scale visual recommendations from street fashion images. In: KDD 2014 (2014)