Advances in Modelling and Analysis A
Vol. 61, No. 2, June, 2018, pp. 64-69
Journal homepage: http://iieta.org/Journals/AMA/AMA_A

Comparative study on traditional recommender systems and deep learning based recommender systems

N.L. Anantha[1*], Bhanu P. Bathula[2]

[1] Acharya Nagarjuna University, Department of IT, VFSTR University, Vadlamudi, Guntur 522213, India
[2] Department of CSE, Thirumala Engineering College, Jonnalagadda (V), Narasaraopet-522601, Andhra Pradesh, India

Corresponding Author Email: anlakshmipathi@gmail.com

Received: April 17 2018
Accepted: June 4 2018

Keywords:
recommender systems, deep learning, item-based collaborative filtering, user-based collaborative filtering, matrix factorization

ABSTRACT

Recommender systems is a big breakthrough for the field of e-commerce. Product recommendation is challenging task to e-commerce companies. Traditional Recommender Systems provided the solutions in recommending the products. This in turn help companies to generate good revenue. Now a day Deep Learning is using in every domain. Deep Learning techniques in the field of Recommender Systems can be directly applied. Deep Learning has ample number of algorithms. These algorithms can be used to give recommendations to users to purchase products. In this paper performance of Traditional Recommender Systems and Deep Learning-based Recommender Systems are compared.

1. INTRODUCTION

Recommender System (RS) is a filtering technique which is a sub-domain of Information Retrieval. RS filters the data from www and gives context-oriented data to the user. Recommender Systems could predict the products gives those products to the users. Recommender systems gives services to both producer and consumer. Recommender systems helps in improving the revenue of e-commerce websites. Similarly, RS helps in recommending friends, music, documents and videos. Recommender systems classified into 3 categories. Content-based filtering, collaborative filtering and knowledge-based systems. These 3 categories are traditional Recommender Systems. Lot of research has been done on traditional Recommender Systems. The goal of Recommender Systems is either to perform prediction or Ranking. Since, couple of years lot of research is taking on Deep Learning based Recommender Systems. In this paper, authors are discussing mainly five Deep learning techniques. They are as follows. Multi-Layer Perceptron (MLP), AutoEncoder (AE), Restricted Boltzmann Machine (RBM), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN). Recommender systems are used in different areas. The following table1 [1] describes examples of products recommended by Recommender Systems.

Table 1. Examples of product recommendation

System	Product
Amazon.com	Books and other products
Netflix	DVDs, Streaming Video
Jester	Jokes
GroupLens	News
MovieLens	Movies
last.fm	Music
Google News	News
Google Search	Advertisements

Table 2. Traditional recommender systems

System	Product
Facebook	Friends, Advertisements
Pandora	Music
YouTube	Online videos
Tripadvisor	Travel products
IMDb	Movies

Recommender Systems [2] has the following phases. The first phase is Information Collection phase, second phase is Learning phase and the final is Prediction/ Recommending Phase.

Recommendation systems are first coined by Tapestry [3]. Recommender systems are also called as Recommendation Systems. The purpose of the Recommender systems is, it provides predicted products or predicting the ratings of the products. Recommender Systems are categorized into mainly 3 types

1) Content based filtering
2) Collaborative filtering
3) Knowledge-based Systems.

The content in Content Based filtering means descriptions. The descriptive attributes are used to give recommendations in Content based filtering. Content Based filtering works as, User given good rating to a movie. Authors don't have any other information. But movie has its type. Genre is nothing but Movie Type. Authors have following movie types. They are Action, Animation, Comedy, Fiction, Thriller, etc. Each movie falls under one of the categories or more than one category. Based on the Genre the movies are recommended to the user. The advantages of content-based recommendations are when user doesn't have enough information then this is the best approach used for recommending the products.

News Dude is a personal news system, which reads new with the help of synthesized speech. To give recommendations it uses TF-IDF model is used to identify the descriptions of the news stories and uses Cosine similarity measure to identify similar news.

LIBRA is a content-based book Recommendation System that analyzes the books gathered from the web. It uses Naïve Bayes classifier to learn user profile and predicts the books for that user.

Collaborative Filtering is most powerful technique in Recommender Systems. Its main focus is the rating given by the user on products. The big challenge for the Collaborative Filtering is sparse data and Cold start problem. It is mainly used in recommending the products based on the user's interest. Collaborative Filtering is classified in 2 categories. First is Item-based recommendation and second is User-based recommendation. Item-based recommendation identifies the similarity between the items. User-based recommendation identifies the similarity between the users based on the products purchased or rated by the users. After identified the similarity between users or items, recommends the products. In finding similar products or similar users it uses K-nearest Neighbor or Matrix factorization approaches. In another context, Collaborative Filtering is further divided into 2 ways. One is Memory Based Collaborative Filtering, and another is Model Based Collaborative filtering. In the model-based filtering, authors supply some part of the data for training. With that data model learns and applies on the test data for either recommendation or prediction. While in Memory based filtering, entire data is given to the model, then model learns and uses that knowledge for recommendation or prediction.

Group Lens is client/server-based architecture uses CF System; the system recommends Usenet news which is a high-volume discussion list service on the Internet. Ringo [7] is a user-based CF system which gives recommendations of music artists and albums. Ringo makes new user to rate list of 125 artists.

Knowledge based systems are used in, the products are not frequently being purchased. This context authors get in Cold start problem. Knowledge based systems provide recommendation with the combination of user ratings, item attributes and domain knowledge. Knowledge based systems are further divided into the following types. Constraint-based recommender systems, Case-based recommender systems.
The following diagram gives the classification of different approaches in Recommender Systems.

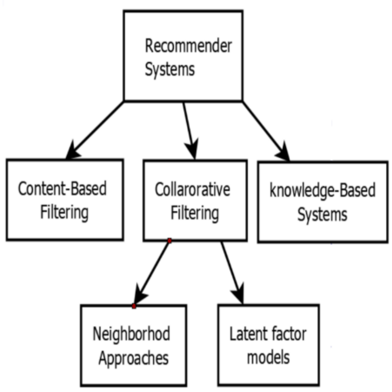

Figure 1. Recommender systems

1.1 Results analysis on traditional recommender systems

In finding the similarity between users and items in this paper Pearson correlation coefficient, Cosine Similarity and Jaccard metrics are used.

Cosine similarity between two products is calculated as.

$$CS(i,j)=\frac{\sum_{u\epsilon U_{ij}} r_{Ui} r_{Uj}}{\sqrt{\sum_{u\in Ui}\mu_{Ui}^2}\sqrt{\sum_{u\in Uj}\mu_{Uj}^2}} \quad (1)$$

where *Ui* is product i, is rated by the user group, and *Uij* is the set of users rated products and *j* are rated by user group.

Another widespread measure to compute similarity is Pearson Correlation similarity:

$$PS(i,j)=\frac{\sum_{u\epsilon U_{ij}}(r_{ui}-r_i)(r_{ui}-r_j)}{\sqrt{\sum_{u\epsilon U_{ij}}(r_{ui}-r_i)^2}\sqrt{\sum_{u\epsilon U_{ij}}(r_{uj}-r_j)^2}} \quad (2)$$

where *Ui* is product i, is rated by the user group, and *Uij* is the set of users rated products *i* and *j* are rated by user group.

Jaccard similarity is used to measure the similarity between two set of elements. The Jaccard similarity between two items is computed as

$$JS(i,j) = \frac{|U_i \cap U_j|}{U_i \cup U_j} \quad (3)$$

where *Ui* is the set of users rated an item *i* similarly *Uj* is the set of users rated an item *j*.

The following metrics are used to evaluate Traditional Recommender Systems. They are Accuracy, MAE, RMSE, precision and recall, ROC curve.
Mean Absolute Error [8], Root Mean Absolute Error [9], Precision-recall curve [10] metrics are statistical accuracy metrics.

MAE is the measure of deviation of user's actual value and predicted value. It is formulated as follows:

$$MAE= \frac{1}{N}\sum_{u,i}|Pui - r_u| \quad (4)$$

where P_{ui} is the predicted rating on item i by the user u, i is original rating and N is the total ratings on the item set. The MAE is minimum means, the prediction of ratings of the Recommender engine accurate. Also, the Root Mean Square Error (RMSE) is given by

$$RMSE= \sqrt{\frac{1}{n}\sum_{u,i}(P_{u,i} - r_{u,i})2} \quad (5)$$

The minimum RMSE means, prediction of ratings by the Recommender Engine is accurate.

Precision is the fraction of good products recommended to total recommended products and recall defined as the fraction of good products recommended those are part of the set of all useful products recommended. They are computed as

$$Precision:\frac{good\ products\ recommended}{total\ recommended\ products} \quad (6)$$

$$Recall:\frac{good\ products\ recommended}{all\ useful\ products\ recommended} \quad (7)$$

In this paper authors used RoC curve to know the performance of the algorithms. Authors have taken Jester and Movielens datasets for testing the performance of Traditional

Recommender System Algorithms. Movielens data set consists of movie name, Genre, movie id, user id, rating given by each user and other details. Rating is ranging from 1 to 5. Jester data set consists of movie name, Genre, Movie id, user id, Rating for the movie given by each user and other details. For experiment, Item-based Collaborative Filtering using Jaccard, Pearson and cosine similarity and item-based collaborative filtering using Jaccard, Pearson and cosine similarity applied on both Jester and Movielens Datasets. The following Fig 1 and Fig 2 gives the details of the performance of each algorithm on Movielens and Jester Datasets. User-based collaborative filtering with Jaccard Correlation Coefficient is performing well on Movielens dataset. Similarly, on Jester Dataset User-based collaborative filtering with Pearson is performing well. From the RoC Curve figure authors know that as the data size is increasing the model performance also increasing.

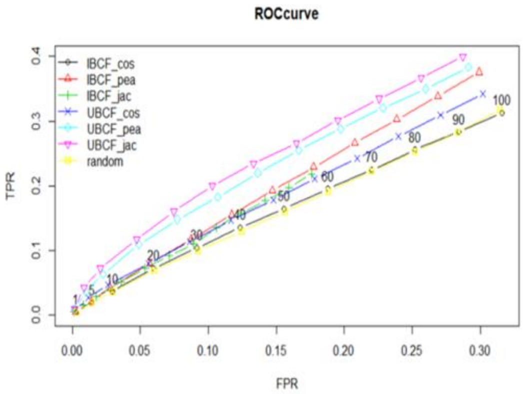

Figure 2. Roc Curve on Movielens Dataset

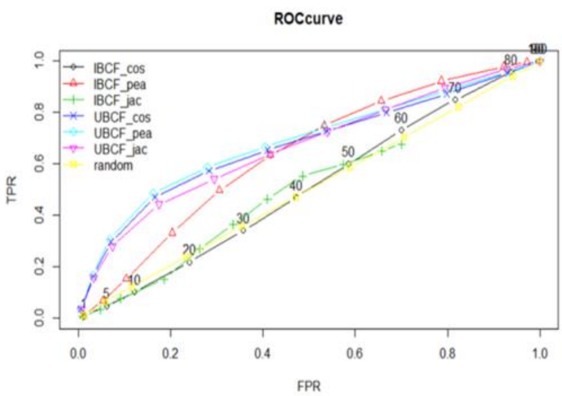

Figure 3. Roc Curve on Jester Datase

1.2 Recommender systems tools/ frameworks

RS is supported by so many frameworks, tools, package and libraries. These help us to test the performance of our algorithms. They are as follows. Easyrec is a Java-based personalized RS, offers Top-N Recommendations. PREA is a Java based personalized recommendation algorithms toolkit gives us the ability to work on collaborative filtering. LibRec is also a java-based RS offers only Collaborative filtering technique. Duine is also a java framework offers both Collaborative filtering technique and Top-N Recommendations. LensKit is a Java toolkit for recommendations offers both Collaborative filtering, Top N Recommendation. Case Recommender is a python Framework offers Collaborative Filtering Technique. Crab is a python framework offers Collaborative filtering. Recommender lab is Package Provides infrastructure for development of recommendations in R offers Collaborative filtering and Top N Recommendations. Graph lab is Machine learning platform offers Collaborative filtering, Matrix Factorization, Top N Recommendation. Scikit is a Machine Learning Framework offers Collaborative filtering, Top N Recommendations and Matrix Factorization methods. My Media Lite is a C# implementation of recommended algorithms offers Collaborative Filtering Technique.

2. DEEP LEARNING BASED TECHNIQUES

Deep learning is sub field of Machine learning. Deep learning is showing immense impact on the fields of image processing, Natural language processing, computer vision and speech recognition. Deep learning is also showing remarkable impact on the Recommender Systems.

Deep learning techniques consists of activation functions and uses Neural Networks. Each Neuron contains activation function. Deep Learning based Recommender Systems uses non-liner functions. Linear function is a function where graph is a straight line. Linear functions don't have any exponents higher than 1. A simplest form of linear function is y= mx + c, where m and c are constants. A non-linear function is a function where the graph is not a straight line. In this paper authors discuss all popular activation functions. Activation functions are used for neurons is to introduce non-linearity to the network. They are sigmoid, tanh, ReLU, softmax etc.,

Activation Functions:

Sigmoid function ranges from 0 to 1. It is also called as logistic function. It is named sigmoid because it is in s-shape. The drawbacks of sigmoid function are sigmoid saturate and kill gradients and sigmoid outputs are not zero centered.

Sigmoid function is defined as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{8}$$

Hyperbolic Tangent function (TanH) function ranges from -1 to 1. It is a trigonometric function.

$$\tanh(x) = \frac{2}{1+ e^{-2x}} - 1 \tag{9}$$

Softmax function is generalization of logistic function. The output of Softmax function is a categorical distribution. It is used in multiple classification methods. Softmax function is defined as:

$$softmax(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^{n} \exp(x_j)} \tag{10}$$

ReLU function is very much useful in feedforward neural networks. ReLU is defined as:

$$f(x) = \max\{0, x\} \tag{11}$$

Authors use the activation functions to carry the neural network. Now authors know the notations used in deep learning techniques.

In this paper authors describe functioning of deep learning techniques in Recommender systems. They are

Multilayer perceptron: is a basic model used in deep learning. MLP has a mathematical function which takes some set of inputs and maps them to output values. MLP is a feedforward neural network with multiple layers. MLP is having one input layer and one output layer. For processing data MLP is having one or more hidden layers. MLP contains perceptron. Each perceptron has one activation function. Figure 1. (a) MLP with one hidden layer.

Restricted Boltzmann Machines: Boltzmann Machine is a stochastic recurrent Neural Network consisting of binary Neurons [11]. Boltzmann Machines consist of one visible layer which takes input and another hidden layer. In Boltzmann Machine, each layer consists of set of nodes. Each node in the visible layer has connections with each node in the layer (intra-node connections) as well as each node in the visible layer has connections with hidden layers (inter-node connections).

These connections making each node dependent on each node causing inefficient sampling etc., To overcome these limitations, Paul Smolen sky proposed Restricted Boltzmann Machine [12]. RBM also has 2 layers. One is visible layer and one is hidden layer. The only change in RBM is, the nodes in a layer don't have connections with the nodes in that layer (no Intra-node connection). Node in the visible layer has connections with all the nodes in the hidden layer. If RBM uses more hidden layers they are called as deep belief networks. RBM's are using in Recommendation systems, classification, regression, clustering, anomaly detection, feature learning, dimensionality reduction. In Recommender systems, RBM is being used in collaborative filtering. Deep Boltzmann Machines, Convolutional Boltzmann Machines and etc are other variants of Boltzmann Machines. Figure 1.(b) RBM with one hidden layer.

Auto Encoders: An autoencoder neural network is an unsupervised learning algorithm that applies backpropagation, setting the target values to be equal to the inputs. Autoencoder Neural Network is like Multilayer Perceptron. It consists of 3 layers, encoder and decoder. First layer is input layer, second layer is a single layer, or more than one hidden layer and third layer is output layer. Encoders role is simplifying the data representation and decoder decodes back the original data. Algorithms takes this encodes data learns more than the normal data. Denoising autoencoders, contractive autoencoders and sparse autoencoders are other variants on autoencoders. Autoencoders are mainly useful in the research areas like information retrieval and dimensionality reduction.

Convolutional Neural Network: It is a class of deep, feed-forward ANN which is a variation of MLP designed to minimize preprocessing. CNN is made up with the following layers. They are convolution layer, pooling layer and fully connected layer. Convolution layer is not fully connected layer. it takes an image as input generates a feature map or activation map. Feature map contains the information about the image. Pooling layer is also called as downsampling. Pooling layer uses either max pooing or average pooling to perform downsampling. Fully connected layer means every node in the network has connections with each node in the next layer. Convolution Neural Network takes an image as input and gives the probabilities of the objects available in the given image.

Recurrent Neural Network: MLP and other Neural Network architectures map input vector to an output vector only. But Recurrent Neural Networks maintains the information about the history of previous inputs to each output. Before producing an output in RNN [13], the network maintains the previous inputs as persistent in the memory to produce output. The best applications of RNN is Natural Language processing.

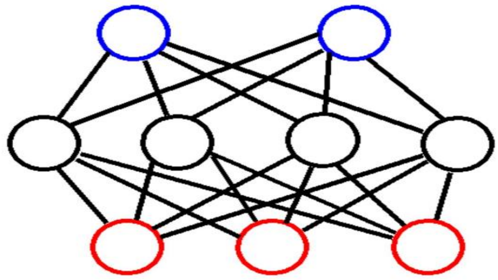

Figure 4. Multilayer perceptron

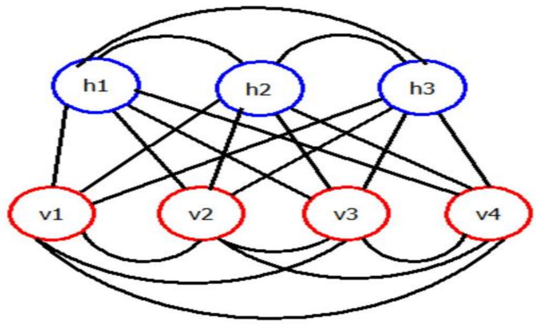

Figure 5. Bolzmann machine

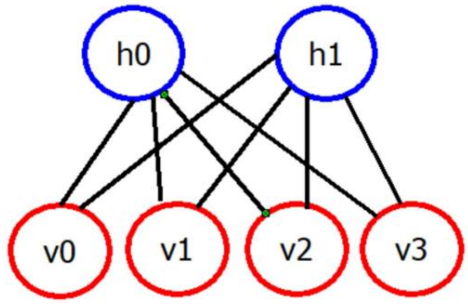

Figure 6. Restricted Botlzmann machine

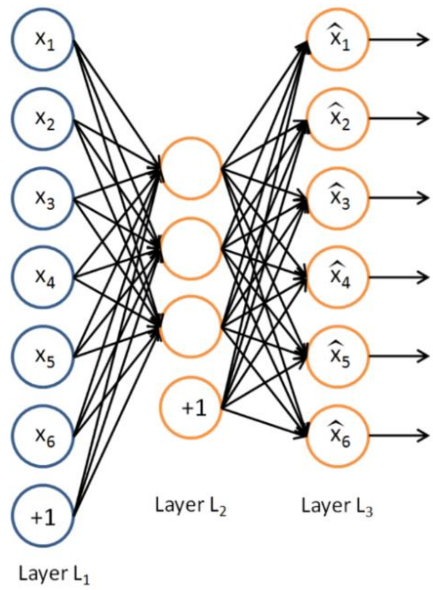

Figure 7. Autoencoder

3. DEEP LEARNING BASED RECOMMENDER SYSTEMS

3.1 MLP based recommender system

Neural collaborative filtering [14] is a MLP based technique which uses matrix Factorization approach and gives user-based and item-based recommendation. Cross-domain Content-boosted Collaborative Filtering Neural Network [15] is MLP based technique offers user-based and item-based recommendations. Deep Factorization Machine, Deep FM [16] is a MLP based technique uses Factorization Machines to provide user-based and item-based recommendations.

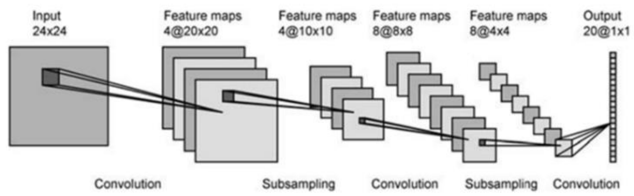

Figure 8. Recurrent neural network

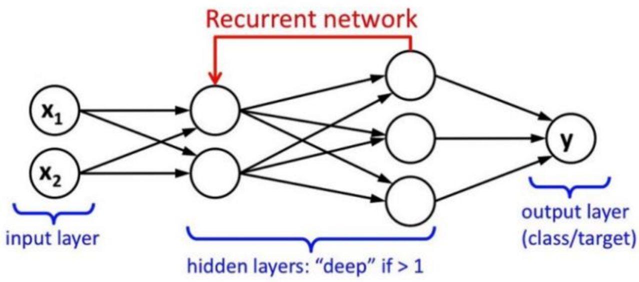

Figure 9. Convolution neural network

3.2 Restricted Boltzmann machine based recommender system

Restricted Boltzmann Machine Collaborative Filtering [RBM-CF] [17] is RBM uses Collaborative Filtering provides user-based recommendations. Hybrid RBM-CF [18] incorporates item features and offers both user-based recommendations and item-based recommendations.

3.3 AutoEncoder based recommender system

AutoRec [19] is an AutoEncoder based technique which offers user-based and item-based recommendation. Authors have separate implementation for item-based recommendation using I-AutoRec and for user-based recommendation using U-AutoRec. Collaborative Filtering Neural network (CFN) [20-21] is also an AutoEncoder based technique offers itembased and userbased recommendations using I-CFN and U-CFN.

3.4 CNN based recommender system

Deep Cooperative Neural Network [DeepCoNN] [12] is Convolutional Neural Network uses factorization Machines provide users rating predictions. ConvMF [23] is a combined model of Convolutional Neural Network and Probabilistic Matrix Factorization technique offers item-based recommendations.

3.5 RNN based recommender system

Recurrent Recommender Network (RRN) [9] is RNN uses preferences of user changes over time and temporal evolution of items seasonality and predicts the ratings.

3.6 Experimental results on deep learning based recommender systems

Authors have successfully completed experimentation on Movielens and Jester datasets using Autoencoder. Autoencoder internally uses MLP. In this experimentation, different Activation functions are used. The table 2 gives loss values of different activation functions used in Autoencoders. Among those activation functions, Autoencoder with Relu activation functions is giving better results.

Table 3. Results using autoencoder

S.No	*Dataset*	*Autoencoder*		
		Sigmoid	*Relu*	*TanH*
1	Movielens	0.18	0.16	0.19
2	Jester	0.17	0.15	1.19

4. CONCLUSION & FUTURE WORK

In this paper, authors have successfully completed experiments on both traditional and Deep learning-based recommender system Algorithms. In future, authors want to propose an algorithm on Convolutional Neural Network which gives better results than before.

REFERENCES

[1] Aggarwal CC. (2016). Recommender systems. Springer International Publishing, 2016.

[2] Billsus D, Pazzani MJ. (2000). User modeling for adaptive news access. User Model User-adapted Interact 10(2–3): 147–80.

[3] Wu CY, Ahmed A, Beutel A, Smola AJ, Jing H. (2017). Recurrent recommender networks. In Proceedings of the Tenth ACM International Conference on web Search and Data Mining. ACM 495–503.

[4] Chen LS, Hsu FH, Chen MC, Hsu YC. (2008). Developing recommender systems with the consideration of product profitability for sellers. Int J Inform Sci 178(4): 1032–48.

[5] Cotter P, Smyth BPTV. (2008). Intelligent personalized TV guides. In: Twelfth conference on innovative applications of artificial intelligence, 957–64.

[6] Donghyun K, Chanyoung P, Jinoh O, Lee SY, Yu H. (2016). Convolutional matrix factorization for document context-aware recommendation. In Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 233–240.

[7] Drosou M, Pitoura E. (2010). Search result diversification. SIGMOD Rec 39(1): 41–7.

[8] Florian S, Jeremie M. (2015). Collaborative filtering with stacked denoising auto encoders and sparse inputs. In NIPS Workshop on Machine Learning for e-Commerce.

[9] Florian S, Romaric G, J´er´emie M. (2016). Hybrid recommender system based on autoencoders. In

Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 11–16.
[10] Goldberg K, Roeder T, Gupta D, Eigentaste PC. (2015). A constant time collaborative filtering algorithm. Inform Retrieval J. 4(2): 133–51.
[11] Goldberg D, Nichols D, Oki BM, Terry D. (2014). Using collaborative filtering to weave an information tapestry. Commun. ACM 35(12): 61-70.
[12] Haykin SS. (2016) Neural networks: A comprehensive foundation.
[13] Guo HF, Tang RM, Ye YM, Li ZG, He XQ. Deep FM. (2017). A Factorization-Machine based Neural Network for CTR Prediction 2782–2788.
[14] Isinkaye FO, Folajimi YO, Ojokoh BA. (2015). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal 16(3): 261-273.
[15] Lian JX, Zhang FZ, Xie X, Sun GZ. (2017). CCCF Net: A content-boosted collaborative filtering neural network for cross domain recommender systems. in proceedings of the 26th international conference on world wide web companion. International World Wide Web Conferences Steering Committee 817–818.
[16] Konstan JA, Miller BN, Maltz D, Herlocker JL, Gordon LR, Riedl J. (2017) Applying collaborative filtering to usenet news. Commun ACM 40(3): 77–87.
[17] Zheng L, Noroozi V, Yu PS. (2017). Joint deep modeling of users and items using reviews for recommendation. In Proceedings of the Tenth ACM International Conference on web Search and Data Mining (WSDM '17). ACM, New York, NY, USA 425–434.
[18] Mooney RJ, Roy L. (2016) Content-based book recommending using learning for text categorization. In: Proceedings of the Fifth ACM Conference on Digital Libraries. ACM 195–204.
[19] Rumelhart DE, Hinton GE, Williams RJ. (2017). Learning representations by back-propagating errors. Nature 323: 6088-533.
[20] Salakhutdinov R, Mnih A, Hinton G. (2017). Restricted Boltzmann machines for collaborative filtering. In Proceedings of the 24th International Conference on Machine Learning, ACM, 791-798.
[21] Smolensky P. (2015). On the comprehension/production dilemma in child language. Linguistic Inquiry 27(4): 720-731.
[22] Sedhain S, Krishna Menon A, Sanner S, Xie LX. (2015). Autoencoders meet collaborative fillteri ng. In Proceedings of the 24th International Conference on World Wide web. ACM, 111–112.
[23] He XN, Liao LZ, Zhang HW, Nie LQ, Hu X, Chua TS. (2017). Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide web. International World Wide Web Conferences Steering Committee 173–182.
[24] Liu XM, Ouyang YX, Rong Wg, Xiong Z. (2015). Item category aware conditional restricted Boltzmann machine based recommendation. In International Conference on Neural Information Processing. Springer 609–616.