

Deep Learning-Based Recommendation: Current Issues and Challenges

Rim Fakhfakh

REGIM-Lab: REsearch Groups in
Intelligent Machines
University of Sfax, National
Engineering School of Sfax
(ENIS), BP 1173, Sfax, 3038,
Tunisia

Anis Ben Ammar

REGIM-Lab: REsearch Groups in
Intelligent Machines
University of Sfax, National
Engineering School of Sfax
(ENIS), BP 1173, Sfax, 3038,
Tunisia

Chokri Ben Amar

REGIM-Lab: REsearch Groups in
Intelligent Machines
University of Sfax, National
Engineering School of Sfax
(ENIS), BP 1173, Sfax, 3038,
Tunisia

Abstract—Due to the revolutionary advances of deep learning achieved in the field of image processing, speech recognition and natural language processing, the deep learning gains much attention. The recommendation task is influenced by the deep learning trend which shows its significant effectiveness and the high-quality of recommendations. The deep learning based recommender models provide a better detention of user preferences, item features and users-items interactions history. In this paper, we provide a recent literature review of researches dealing with deep learning based recommendation approaches which are preceded by a presentation of the main lines of the recommendation approaches and the deep learning techniques. We propose also classification criteria of the different deep learning integration model. Then we finish by presenting the recommendation approach adopted by the most popular video recommendation platform YouTube which is based essentially on deep learning advances.

Keywords—Recommender system; deep learning; neural network; YouTube recommendation

I. INTRODUCTION

With the rapid growth in the amount of published data in the Net characterized by an exponential evolution, the task of managing this information becomes more and more difficult. Thereby, the user has more ambiguity in finding the most relevant content for his information need. The recommendation systems serve as information filtering tools which are useful for helping users in discovering new contents, services and products they probably are interested in. The recommender system shows its importance not only for users considered as customers but also for digital companies due to its vital and essential role played in various information access systems which is helpful in boosting business and facilitating decision-making process. The recommendation process is driven by various embeddings features including item features, user preferences and user-item history interactions. Additional contextual information like temporal and spatial data and the used device can be used also for the generation of recommendation items. According to the input data of the recommender system, we can distinguish between three main recommendation models which are collaborative filtering, content-based recommender system and hybrid recommender system. Despite the widespread of the effectiveness of these

models proved through the time, we can deny that these models suffer from some limitations in dealing with the problems of cold-start and data sparsity. We mean by the cold-start problem, the issue that the system is unable to draw any interpretations about items or users due to a lack of information gathered and data sparsity problem the issue related to the little information about each item or user in a large data set. Another problem related to these models can be witnessed when balancing the recommendation effectiveness in terms of diverse evaluation metrics. To overcome these problems, many solutions are proposed in the literature including the use of deep learning for the enhancement of the recommendation models.

In the recent decades, the deep learning has witnessed a great success in many application fields such as computer vision, speech recognition and natural language processing. Due to capability of the deep learning in solving many complex tasks, this trend is adopted by the academic researches as well as by the industry who have followed the race with applying the deep learning in many applications. In this context, the deep learning has been adopted for enhancing the recommendation approaches by improving the user experiences and ensuring his satisfaction. The deep learning has shown his effectiveness in capturing the non-linearity of the user-item interconnections. It also allows the abstraction of the most complex data representation due to the intelligence driven by communicated layers in the architecture. Furthermore, the deep learning is able to catch the involved relationships within the different data sources such as textual, visual and contextual information and predict the circumstance's recommendation.

The remainder of this paper is organized as follows: In Section 2, we present the background of recommender systems and the deep learning. Section 3 is reserved to the deep learning-based recommender approaches and its classification according to the integration way of the deep learning and the dependence level on it. This classification is followed by the identification of the new challenges of the deep learning based recommendation. In Section 4, we focus on YouTube as a deep learning based recommender system and we study its architecture.

II. BACKGROUND

A. Recommender System

1) Approaches

A recommender system aims to estimate the preference of a user on a new item which he has not seen. The output of a recommender system varies according to the nature of the system, its utility and information treated as inputs. It can be either a rating prediction or ranking prediction. Rating prediction aims to predict the rating scale to an item which is not seen by the user by filling the missing entries of the user-item rating matrix. Ranking prediction aims to predict the top n items and produces a ranked list according to its similarity with user profile or items features or both of them.

There are mainly three models used for recommendation depending on the nature of the information used as inputs. These models are content based recommender system, collaborative filtering, and hybrid recommendation.

a) Content-Based recommendation

The input in this type of approaches is the content information. This approach is based on the construction of a user profile based on the features of the items that the user interacts with it by rating, clicking or any explicit or implicit means of interaction. Treated items can be texts, images or videos. This profile is used to identify new interesting items for the user which is relevant to his profile. The recommendation model is based on the comparison between items and users features. In this category of recommendation, if a user is interested on the item X and the item Y is highly similar to X so Y is predicted to be relevant to the user and recommended for him.

b) Collaborative filtering

It is an alternative to content based recommender system. As inputs, it relies only on past user behavior. The user behavior is learned from the previous interactions of the user with items presented as user-item matrix. There is no requirement for explicit user profile creation. There are two ways for the interaction with items available to the user, either in explicit way by rating items or using the implicit feedback deducted from clicks through the Net, browsing histories or user interactions in social networks. Information about the user is useful for predicting new items basing on Matrix factorization algorithms.

c) Demographic filtering

This type of recommendation classifies users under a set of demographic classes representing the demographic characteristics of users known from their age, nationality, gender, occupation and location. The major benefit of this algorithm is the no need for the user ratings history.

d) Hybrid model

This type of approach integrates any two of the above recommendation approaches to predict new items for the user.

2) Application fields

Many of the modern internet services require the recommendation approaches for the perdition of new items to the user. This requirement comes from its vital role in boosting

business, facilitating the decision making and tracking the user intention without his explicit intervention. Among the application fields which depend essentially on recommendation systems we find movies recommendation, news recommendation, e-commerce services recommendation, e-learning recommendation, books recommendation, songs recommendation, applications recommendation, websites recommendation, travel destinations recommendation and so on. Each recommendation scenario has its specificity for choosing entrees attributes and the suitable approaches. In the following we present some related works dealing with different applications fields.

a) Multimedia platforms recommendation

This type of recommendation deals with content based approach as well as the collaborative filtering approach. There several commercialized Multimedia platforms recommender systems dealing with movies (IMDb, Netflix,), videos (YouTube, DailyMotion...), music (Deezer, Spotify...) or images (Flickr). Among the movie recommendation systems, we find, MovieLens, MovRec, and Netflix. MovieLens is based on collaborative filtering approach to make recommendations of movies which are not yet seen by the user based on the previous user movie ratings. MovRec [5] is based also on collaborative filtering approach to recommend movies which are judged to be most suitable to the user at that time using Matrix factorization and k-means algorithms. Netflix is an hybrid recommender system which makes recommendations by fetching users having similar profiles (collaborative filtering) as well as by predicting movies sharing similar features with movies highly interested by the user (content-based filtering).

b) News recommendation

This type or recommendation focuses more on the freshness of the news article. The two approaches content-based recommendation and collaborative filtering have been adopted for the purpose of news recommendation and mostly the two strategies are combined [6]-[8]. First, as for the content based news recommendation, a profile for each user is created and used for matching the news articles basing on article features, user profile or both for hybrid recommendation. Second, the collaborative filtering approach which rely only on past user behavior without requiring the creation of explicit profiles.

c) E-commerce services recommendation

Many of the largest E-commerce Web sites are based on recommendation techniques to help their customers to find the most valuable products among the available ones and recommend them to be purchased by the user. This technique plays a major role in increasing the sales of these E-commerce sites. The most famous E-commerce websites we find Amazon.com and eBay. Recommendation in these websites is generated based on the likelihood of the available items and the previous purchased, clicked or liked items by users.

d) E-learning recommendation

This type of recommendation is adopted for the personalization of educational content. Many systems are based on hybrid recommendation approach which takes

advantage of the rating data or the users' feedback and tags associated to the courses to recommend the suitable pedagogical resources to users [9], [11]. Some systems are based only on the collaborative filtering approach like the work of [10] adopted for the recommendation of learning materials by the consideration of the context, the students' profile and the learning materials properties. These techniques are exploited by the e-learning platforms Coursera and Moodle to satisfy the user profile and his intention.

e) Social network recommendation

The power of social networks comes from its capability in connecting users in the easiest way and recommending the suitable information to the users without his explicit intervention. Behind this power, we find a great importance related to the development of link recommendation features and handling the social graph basing on the topology of existing links and leveraging quantities such as node degree and edge density [12]. Link recommendation techniques are categorized into learning-based techniques and proximity-based techniques. The Learning-based techniques are based on training algorithms for the prediction of the association likelihood to the link. Otherwise, the proximity-based techniques do not need a construction of training data. They are characterized by the easiness of implementation which makes them widely applied in practice [13]. These techniques dealing with common neighbor are used by the major of online social networks for recommendation of new friends, new groups, new pages, or new connections. This is available through the functionality "People You May Know" and "mutual friend" in Facebook, "shared connection" and "People You May Know" in LinkedIn and "You May Know" in Google+.

f) Job recommendation

This type of recommendation is the core of the intelligent recruitment platforms which deals with the matching between job-seekers and vacancies. This platform can be useful for job-seekers as well as for employers who are looking for specific skills. This type of recommendation learn generalizations between user profile and job posting based on similarity in title, skill, location, etc. In [14], authors propose an efficient statistical relational learning approach which is used for constructing a hybrid job recommendation system. In [15], authors propose a directed weighted graph where the nodes are users, jobs and employer. The recommendation process is applied based on the similarity computing between any two profiles of objects.

3) Datasets and evaluation metrics

In the literature review, there are a representative set of existing evaluation metrics used for testing the performance of the proposed recommender systems. These evaluations measures have their standard formulations which are generally applied on a group of open recommender system public databases which are generated for the purposes. The used evaluation metrics can be classified into two different groups depending on the outputs parameters which can be either a rating prediction metrics or ranking score prediction. Table II present the most used set of classical recommender system evaluations metrics. Besides the accuracy evaluation, other novel evaluation metrics are considered for making better user

satisfaction such as novelty, serendipity, coverage, diversity, stability, reliability, privacy, trustworthiness and interpretability.

The novelty metric specifies the difference degree between recommended items and items already visited by the user. Otherwise, the diversity metric specifies the differentiation degree among recommended items [16]. The coverage metric [19] is a factor that estimates the quality of the prediction in a way that indicates the situations percentage in which at least one of the user k-neighbors rate a new item not yet rated by that user.

The reliability metric informs about how seriously we may consider the prediction value. In this way the evaluation of a recommender system will be based on a pair of values: prediction value and reliability value through which users may balance their preferences and consider them for taking their decisions. The reliability value depends on the similarity of the user neighbors who are used for making the prediction and the degree of disagreement between them on rating a predicted item. In other words, a prediction of 4.5 out of 5 is much more reliable if it has obtained by a big number of similar users than if it has obtained by only two similar users [18].

The stability quality metric quantifies the stability of the system over the time. A recommender system is stable if the provided predictions and recommendations do not change strongly over a short period of time. This metric reflects the users' trust towards the recommender system [17]. The serendipity measure how surprising the relevant recommendations are.

Table II provides the most used evaluation metrics associated with their mathematic formulation in the field of recommender systems.

As far datasets used for evaluation of new proposed approaches dealing with recommender system techniques, there are several ones which are used depending on the type of application fields and the input-output parameters. Table I presents the famous evaluation datasets according to the specific application field.

TABLE I. DATASETS FOR RECOMMENDER SYSTEM VALIDATION

Application fields	Datasets
Movies	MovieLens Netflix Dataset MoviePilot Dataset Jester Yahoo!
Products	Amazon eBay kaggle eCommerce Item Data
Music	Last.fm Dataset Spotify Dataset Million Song Dataset Audioscrobbler
News	YOW Dataset SmartMedia Dataset Crawling from news websites
Pedagogical content	Book-Crossing Dataset Scholarly Paper Recommendation Datasets
local businesses: dentists, hair stylists, ...	Yelp

TABLE II. RECOMMENDER SYSTEMS EVALUATION METRICS

	<i>Evaluation metrics</i>	<i>Formula</i>	<i>Description</i>	
<i>Standard</i>	<i>Rating prediction</i>	Mean Squared Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2$	n: unrated items used for the test pi: the prediction on the ith test instance ri: the corresponding rating value given by the user
		Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2}$	
		Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n p_i - r_i $	
		Normalized Mean Absolute Error (NMAE)		
	<i>Ranking score prediction</i>	Recall	$Recall = \frac{TP}{TP + FN}$	TP: True Positive: an interesting item is recommended to the user TN: True Negative: an uninteresting item is not recommended to the user FN: False Negative: an interesting item is not recommended to the user FP: False Positive an uninteresting item is recommended to the user).
		Precision	$Precision = \frac{TP}{TP + FP}$	
		F1_score	$F1_Score = \frac{2 \times Precision \times Recall}{precision + recall}$	
		Receiver Operating Characteristic (ROC)	A graphical technique that uses two metrics: TPR (True Positive Rate) and FPR (False Positive Rate) $TPR = \frac{TP}{TP + FN}$ $FPR = \frac{FP}{FP + TN}$	
		Area Under the Curve (AUC)	A graphical technique where we plot the various thresholds result in different true positive/false positive rates	
		Normalized Discounted Cumulative Gain (NDCG)	$NDCG = \frac{DCG}{IDCG}$ where $DCG = rel_1 + \sum_{i=2}^{pos} \frac{rel_i}{\log_2 i}$ $IDCG = rel_1 + \sum_{i=2}^{ h -1} \frac{rel_i}{\log_2 i}$	Rel i : the relevance of recommendation at position i Pos: the position up to which relevance is accumulated h: the set of correctly recommended items
<i>Novel</i>	Novelty	$Novelty = \sum_{i \in L} \frac{\log_2 P_i}{n}$ where $P_i = \frac{n - rank_i}{n - 1}$;	U potential users n items number N potential items L recommendation list C content vector related to each item having a length c. u: users for whom the recommender system was able to generate a recommendation lists P ₁ : old prediction P ₂ : new prediction	
	Serendipity	$Serendipity = \frac{1}{n} \sum_{i \in N} \max(P_{user} - P_{ij}, 0) \times rel_i$		
	Diversity	$diversity = \frac{a}{c} \sum_{i=1}^c \frac{1}{n} \sum_{j=1}^n i_j$		
	Coverage	$Coverage = 100 \times \frac{u}{U}$		
	Stability	$stability = \frac{1}{P_2} \sum_{i \in P_2} P_{2,i} - P_{1,i} $		

B. Deep Learning

The deep learning is a class of machine learning algorithms. It is based on the advances of the neural networks which are rebranded in the recent years as deep learning. The deep learning shows his performance in treating many application fields like speech recognition, object detection and natural language processing proved by the trust offered by the most commanding enterprise in the world such as Google, Facebook and Microsoft. The model of deep learning is represented as a cascade of nonlinear layers which form an abstraction of data. The deep learning is used for supervised and unsupervised learning tasks. In the literature, the appearance of deep learning is related to computer vision domain including object and speech recognition. The architecture of a neural network is basically composed from three layers: input layer, hidden layer and output layer.

The distinction between the different types of networks is related to the type of hidden layer and the number of hidden layers determines the depth of the neural network. The simplest artificial neural network is feedforward neural network where the information moves from the input nodes forward output nodes through the hidden nodes in the same direction and without making loops or cycles in the network. The transition between layers is controlled with an activation function which can be linear or nonlinear such as tanh, sigmoid and Rectified Linear Unit (ReLU). The activation functions manage the corresponding inputs weights w_{ij}. The architecture of neural network is illustrated in Fig. 1. The red color presents the operations applied on the simple feedforward neural network to make a recurrent neural network where neurons in the hidden layer are connected recurrently to neurons in the input layer.

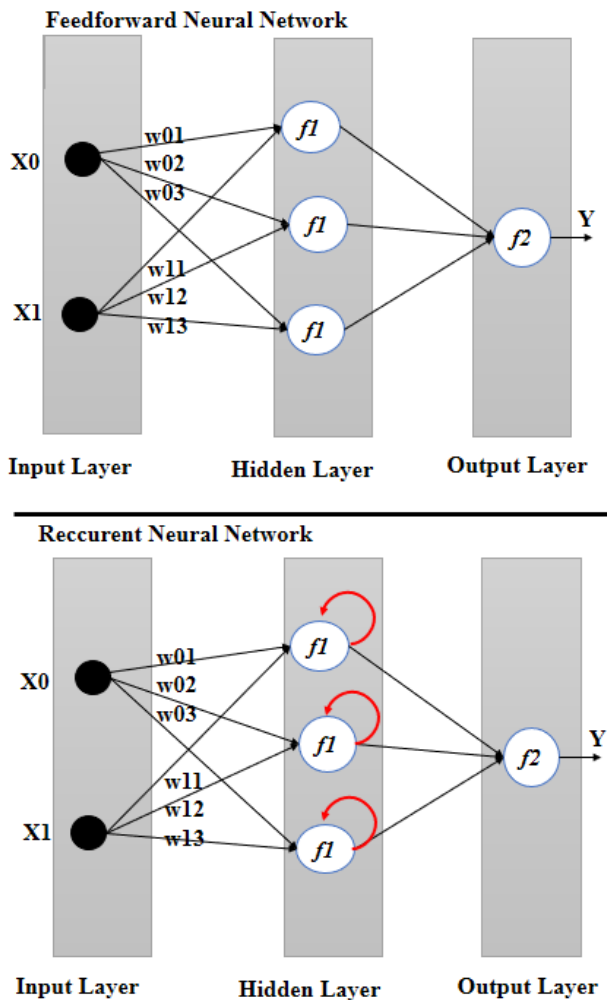


Fig. 1. Feed forward vs Recurrent neural network architecture.

There are several categories of deep learning models. Among these models we cite the following which differ in term of complexity and application fields [20]-[23]:

- FF: Feedforward Neural Network: It is the simplest deep learning approach having as derivatives Single Layer Perceptron and Multilayer Perceptron MLP. MLP have multiple hidden layers which are interconnected in a feed-forward way. The mapping between the input and the output layers is driven by an arbitrary activation function.
- AE: Autoencoder: It represents an unsupervised learning model. The simplest form of it is similar to the MLP architecture with the specificity that the output layer has the same number of nodes as the input layer in order to reconstruct the inputs but in which the principle of dimensionality reduction is applied. There are various techniques derived from the autoencoder model such as sparse autoencoder, contractive autoencoder, variational autoencoder, denoising autoencoder and marginalized denoising autoencoder.

- RNN: Recurrent Neural Network: This variant of deep learning algorithms is widely used in sequence learning and natural language processing applications thanks to its capability in sequential data modelling. It is characterized by the use of the internal memory of the network and the loops operations which form the major difference with the feedforward neural network. There are several variants which derive from the RNN model such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) which are deployed to overcome problems related to vanishing gradient.
- CNN: Convolutional Neural Network: This algorithm is successfully applied in the computer vision and image processing domains. It can be defined as a subclass of feedforward neural network with the specificity of pooling operations and convolution layers. It shows its performance in capturing local and global features for item abstraction and representation.
- DSSM: Deep Structured Semantic Model: This is a specific deep learning model used for semantic representations learning and applied for measuring the semantic similarities between entities.
- BM: Boltzmann machine: This is a type of stochastic recurrent neural network where units are symmetrically connected.
- RBM: Restricted Boltzmann Machine: This is a deep neural network composed of two layers named visible and hidden layers. Its specificity comes from the dependence between the two layers. There is no intra-communication between this latter which form the source of the nomination restricted.
- NADE: Neural Autoregressive Distribution Estimation: This is considered as an unsupervised neural network which depends on the principle of feedforward neural network and the autoregressive model. It is mainly used for purposes of density and distribution estimation.
- GAN: Generative Adversarial Network: This is composed of two neural networks which are trained simultaneously by contesting with each other in a zero-sum game framework. The first network is used for candidates' generation termed as generator and the second for the candidates' evaluation termed as discriminator.

III. DEEP LEARNING FOR RECOMMENDER SYSTEM

Given the great success of the deep learning shown in many applications fields, it has recently been proposed for enhancing the recommender systems quality. In this section, we explore the used of deep learning methodology in the field of recommender system and we present two classification criteria of the deep learning based recommendation approaches according to the integration model of deep learning and its dependence level. The integration of deep learning can be performed with the classic recommendation models such as collaborative filtering and content-based approaches or in a way purely based on deep learning. The second classification criterion is the dependence level on the deep learning which

can be based on only one deep learning technique or more than one technique [1], [2].

A. Integration of Deep Learning for Recommender Systems

1) Deep learning integration with classical recommendation model

The first part of literature review is directed towards the combination of deep learning methods with the classical recommendation techniques. In several works, the integration of deep learning is combined either with collaborative filtering approaches or with content-based recommendation approaches. The most popular area of collaborative filtering is the latent factor approaches and neighborhood approaches. The deep learning reaches the two types of approaches. As for the latent factor approaches, the deep learning is applied for improving the performance of several algorithms such as factorization machine [26], matrix factorization [24], probabilistic matrix factorization [25], and K nearest neighbors’ algorithm [27]. The incorporation of deep learning with content based recommender systems can be exploited by applying the advantages of Convolutional Neural Network or Recurrent Neural Network in extracting visual or textual features of the recommended items.

2) Recommendation only-based on deep learning

This type of recommendation is based only on deep learning techniques for all the training and predicting steps of recommender system. There is no form of benefit from classical recommendation models. Many deep learning techniques are used for this purpose such as the Multilayer Perceptron network, CNN, RNN, Autoencoder, Deep Structured Semantic Model and Restricted Boltzmann Machine. Table III illustrates the most recent literature review dealing with deep learning approaches for predicting items for users.

TABLE III. LITERATURE REVIEW OF RECOMMENDATION APPROACH BASED ONLY ON DEEP LEARNING WITHOUT INTEGRATION OF THE CLASSICAL RECOMMENDATION MODELS

Recommendation based-only on deep learning	Deep learning techniques	
	Multilayer Perceptron	Recurrent Neural Network
	[29] Cheng et al, 2016	[36] Li et al, 2016
	[51] Wang et al, 2017	[52] Tan et al, 2016
	[50] Alashkar et al, 2017	[53] C.Y Wu et al, 2016
	Convolutional Neural Network	Autoencoder
	[33] Gong et al, 2016	[54] Zhuang et al, 2017
[32] Nguyen et al, 2017	[55] Y.Wu et al, 2016	
[51] Wang et al, 2017	[56] Ouyang et al, 2014	
Deep Structured Semantic Model	Restricted Boltzmann Machine	
[42] Elkahky et al, 2015	[60] Xie et al, 2016	
[57] Chen et al, 2017	[38] Liu et al, 2015	
[58] Xu et al, 2016	[59] Georgiev et al, 2013	

B. Dependence Level on the Deep Learning

1) One dependence level

This type of model is characterized by the use of only one deep learning model during the entire recommender system architecture. The selection of the deep learning technique is conditioned by the achievement of the recommendation model and its entrees parameters. Each deep learning model has its

specificity which leads its integration in recommendation process.

- The multilayer perceptron model shows its performance in modelling nonlinear interactions between user’s preferences and items features which is useful for enhancing the recommendation quality. The multilayer perceptron model is integrated with the collaborative filtering to born the neural collaborative filtering techniques. This latter is adapted for several applications including traditional matrix factorization [28], neural social collaborative ranking [31], cross-domain content-boosted collaborative filtering neural network [30]. The multilayer perceptron model is used also for solving regression and classification problems in order of enhancing the diversity as well as the accuracy of recommendation. This architecture is approved by Google play App recommendation [29]
- The convolutional neural network is used in recommender systems due its performance in capturing local and global features coming from visual and textual data sources. This model is useful for solving the problem of classification and tag recommendation basing on visual features extraction from images patches [32] or selecting informative words from textual information [33].
- The recurrent neural network shows his performance in allowing the recommender system to manage the variation of rating data and content information in respect with time factor. It injects user’s short-term preferences, context information and click histories into input layers which will be processed to predict the likelihood items. Several works are based on LSTM algorithm for item recommendation taking into account user’s past session actions [34]-[36].
- The Deep Structured Semantic Model is used in information retrieval and recommendation field due to its ability in performing a semantic matching between users and items. It allows the entities projection into a common low-dimensional space which facilitates the similarity computing thanks to cosine function [37].
- The Restricted Boltzmann Machine model is used for recommender system in association with the collaborative filtering techniques. The core idea is to incorporate user features (known from the implicit feedback or the rating statement) or items features into the Restricted Boltzmann Machine model to predict the most relevant items to the user [38], [39].

2) Hybrid dependence level

This category of deep learning based recommender systems refers to more than one deep learning technique to form hybrid model. The combination of different deep learning techniques may improve the recommendation quality tanks to the complementary aspect between the different network architecture. Many possible combinations of deep learning techniques have been exploited in the field of recommender systems. Table IV illustrates several research studies where

authors are based on the combination of deep learning techniques in the field of recommendation.

TABLE IV. DEEP LEARNING COMBINATION TECHNIQUES IN THE FIELD OF RECOMMENDATION

	CNN	RNN	AE	MLP	DSSM
[41] Q. Zhang et al, 2017	x	x			
[48] P. Li et al, 2017		x	x		
[43] F. Zhang et al, 2016	x			x	
[44] Rawat et al, 2016	x		x		
[45] Lei et al, 2016	x		x		
[47] Wang et al, 2016		x		x	
[40] Lee et al, 2016	x	x			
[49] Song et al, 2016		x			x
[42] Elkahky et al, 2015	x	x			
[46] Ge et al, 2010	x				x

C. Challenges and Current Issues

Through the studied literature review, we noticed that the current deep learning techniques have a great effect in building a solid foundation of the recommender system. Each deep learning model has its specificity in treating the data. For instance, the CNNs can be considered as the ideal solution for the task of content-based feature learning which is useful for solving the cold-start problem. The RNNs are very powerful for several tasks like multi-modal learning, trends forecast and sequence based recommendation. The DSSM is an imposing solution for the semantic matching between users and items. The different types of deep learning techniques boost the recommendation performance by making a significant progress in the traditional recommender system which opens a new perspective dealing with deep recommender system. Thereby, the deep learning based recommendation as still in evolution and there are new challenges which the emerging research trends are focusing on. In the following, we identify some emerging research trends:

- Dealing with the novel evaluation metrics: Researchers have discovered that working on improving the recall/precision accuracy or reducing the prediction error is not suffusing for trusting a recommender system in the long term. There are new evaluation metrics which should be taken into account such as novelty, diversity, coverage, serendipity, privacy, trustworthiness, interpretability ... this evaluations metrics not only deal with accuracy performance of the historical modelling of the user, but also his holistic experience is mattering.
- Scalability: The scalability problem deals with the no-stop increasing in data volumes of items and users and their embeddings. Thereby, the scalability is critical factor for choosing the recommendation model where the time factor has also a principal consideration. A recommender system should satisfy the two factors in the same time. In this context, the deep learning proves its power in big data analytics.
- Hybrid dependence on deep learning: as shown in the previous section, several works depends on more than one type of neural networks where different deep learning techniques are combined together. This

technique proves its effectiveness in modelling heterogeneous data like users, items and contexts. But we have mention that not all possible combinations of different deep learning models are treated in the literature and several deep techniques couples are not tried yet such as the couple “Autoencoder and DSSM”. The combination of more than two deep learning model should be well studied and not with random depending on the practical requirement of the model.

IV. CASE STUDY: YOUTUBE AS A DEEP LEARNING-BASED RECOMMENDER SYSTEM

YouTube can be defined as the most popular online video platform. Indeed, it reaches recently (September 2017) more than 1.5 billion monthly active users. The main service allowed by this platform is the automatic suggestion of a list of related videos to a user in response to the video currently viewed and by taking into account the collected practices of the user including the history of viewed videos. YouTube interface shows by default the first 25 related videos for any watched video. In the literature review, there are some propositions of how the YouTube’s recommendation system is functioning. According to [3], there are some orientations that the video recommendation approach used by YouTube is the collaborative filtering where the principal inputs of the algorithm are patterns of shared viewership. The recommendation is predicted by exploring a video graph representation where two videos are estimated to be related if there are many users that watch the video B after the video A. Another vision in the literature which represent the minority are oriented towards a syntactical approach based on matching keywords within the title, description and tags and predict the most related videos. All these propositions are accurate in a part as the YouTube recommendation approach does not restricted to just one type of input data. The exact YouTube’s recommendation system functioning is avowed in the paper of the YouTube proprietor researchers [4]. In this paper, it is stated that the recommender system used by YouTube is driven by the Google Brain project developed by Google researchers and engineers for the purposes of conducting artificial intelligence and deep learning technologies. This project is recently open-sourced as TensorFlow (<https://www.tensorflow.org/>). This library allows the exploitation of different deep neural network architectures. The recommender system used by YouTube is composed from two neural networks.

The First Neural Network: The process treated by this network is termed as candidate generation. It is constructed for learning user and item embeddings. It takes as input, information about the user collected from his watch history. These embeddings are fed into a feedforward neural network constructed from several fully embedded layers. The use of deep learning is integrated with the traditional recommendation approach, the collaborative filtering, with the processing the Matrix Factorization algorithm. This architecture allows the generation of a distribution of hundreds of videos predicted to be relevant to the user from the YouTube corpus composed from millions of videos. The architecture of the proposed neural network allows the easily addition of new features to the model. The additional embeddings cover the search history,

demographic information (age, gender, location) used device, time, context, video class, video freshness, etc. The fed embeddings with all other model parameters are learned together using the normal gradient descent back-propagation updates algorithm. The concatenation of features is processed in the first layer which is followed by several layers connected and controlled with the activation function Rectified Linear Units (ReLU).

The Second Neural Network: is used for ranking the few hundred videos issued from the first neural network of candidate generation in order to make recommendations to the user. Compared to the problem of candidate generation, the ranking is much simpler as the number of treated videos is smaller and there is sufficient information about the user and the items. This deep neural network is based on the logistic regression to assign a score to each video depending on the expected watch time. The recommendation process benefits from several features and indicates the past user behavior with items. The used features are: time since last watch, previous

impressions of the user, user language, video language, impression video ID, watched video IDs ... These features require usually a normalization process to be ready to feed the input layer of the neural network. Experiments have shown that increasing the width of hidden layers as well as their depth improves the network results. The best configuration of the hidden layers was a 1024-wide ReLU followed by a 512-wide ReLU followed by a 256-wide ReLU which present non-linear interactions between the different features.

The algorithm of YouTube recommendation is illustrated in Fig. 2. The effectiveness of the YouTube recommendation algorithm is approved by several offline metrics such as precision, recall and ranking loss. However, YouTube Community rely also on A/B testing model via live experiments in order to ensure an iterative improvement of the system by capturing the subtle changes in watch time or in click-through rate or any other feature that measure the user engagement.

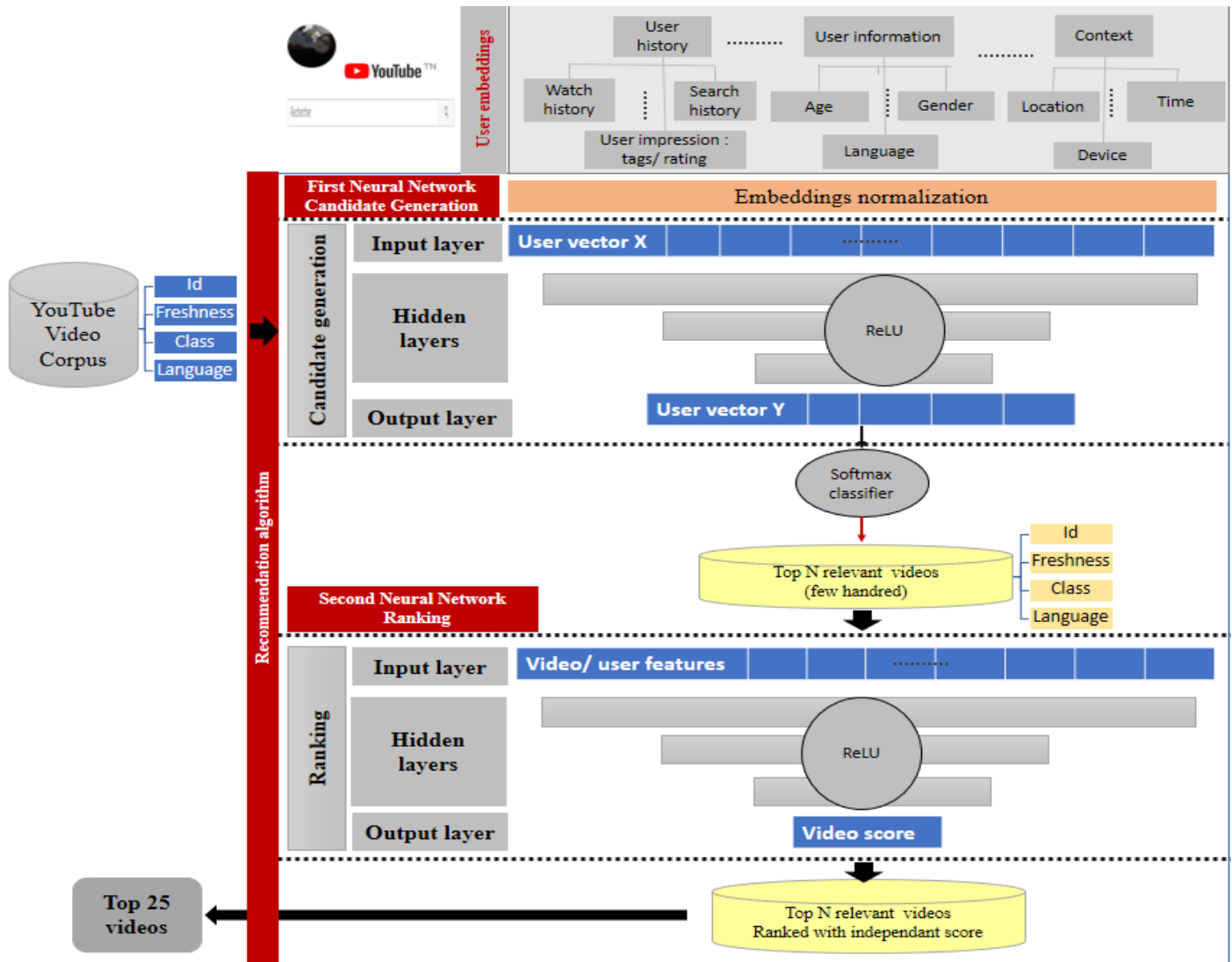


Fig. 2. Feed forward vs Recurrent neural network architecture.

V. CONCLUSION

In this paper, we are interested in the deep learning based recommender systems which benefits from the deep model for enhancing the management of users, items, contexts and user-items interconnections to guarantee the user satisfaction. We have presented the background of recommender systems as well as the deep learning architecture which is chained with the illustration of the deep learning-based recommender approaches. We have classified these approaches according two criteria: to the integration way of the deep learning and the dependence level on it. Moreover, we have identified the new challenges and the future directions dealing with the deep learning based recommendation. Finally, we have handled a real case study that implement a deep learning based recommender system which is YouTube as it can be considered as famous video recommender system driven by Google Brain team. YouTube integrate the deep neural network to learn everything about viewers' habits and preferences

ACKNOWLEDGMENT

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR61 IES48.

REFERENCES

- [1] Liu, Juntao & Wu, Caihua. (2017). Deep Learning Based Recommendation: A Survey. 451-458. 10.1007/978-981-10-4154-9_52.
- [2] Shuai Zhang, Lina Yao, Alexin Sun, Nanyang Technological University Deep Learning based Recommender System: A Survey and New Perspectives ACM J. Comput. Cult. Herit., Vol. 1, No. 1, Article 35. Publication date: July 2017.
- [3] M. Airoidi, et al., Follow the algorithm: An exploratory investigation of music on YouTube, Poetics (2016), <http://dx.doi.org/10.1016/j.poetic.2016.05.001>
- [4] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 191–198.
- [5] Yiu-Kai Dennis Ng, "MovReC: a personalized movie recommendations for children based on online movie features", International Journal of Web Information Systems, <https://doi.org/10.1108/IJWIS-05-2017-0043>
- [6] Shmueli, E., Kagian, A., Koren, Y., Lempel, R.: Care to comment? Recommendations for commenting on news stories. In: Proceedings of the 21st international conference on world wide Web, pp. 429–438. ACM, New York (2012)
- [7] Abbar, S., Amer-Yahia, S., Indyk, P., Mahabadi, S.: Real-time recommendation of diverse related articles. In: Proceedings of the 22nd international conference on world wide Web, pp. 1–12. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva (2013)
- [8] Abel, F., Gao, Q., Houben, G.-J., Tao, K.: Analyzing user modeling on twitter for personalized news recommendations. In: roceedings of the 19th international conference on user modeling, adaption, and personalization, pp. 1–12. Springer, Berlin (2011)
- [9] Vladioiu, M., Constantinescu, Z., & Moise G. (2013). QORECT – a case-based framework for quality-based recommending open courseware and open educational resources In C. Bădică, N.T. Nguyen, & M. Brezovan (Eds.), ICCCI 2013, LNAI 8083, 681–690.
- [10] Lichtnow, D., et al. (2011). Recommendation of learning material through students' collaboration and user modeling in an adaptive e-Learning environment. Technology enhanced systems and tools, SCL, 350, 258–278.
- [11] Durao, F., & Dolog, P. (2010). Extending a hybrid tag-based recommender system with personalization. In Proceedings 2010 ACM Symposium on Applied Computing, SAC 2010, Sierre, Switzerland (pp. 1723–1727).
- [12] Li, Z., Fang, X., and Sheng, O. R. L. A survey of link recommendation for social networks: Methods, theoretical foundations, and future research directions. Theoretical Foundations, and Future Research Directions (October 28, 2015) (2015).
- [13] Backstrom, L., and Leskovec, J. Supervised random walks: predicting and recommending links in social networks. In Proceedings of the fourth ACM international conference on Web search and data mining (2011), ACM, pp. 635-644.
- [14] Yang, Shuo & Korayem, Mohammed & Aljadda, Khalifeh & Grainger, Trey & Natarajan, Sriraam. (2017). Combining Content-based and Collaborative Filtering for Job Recommendation System: A Cost-sensitive Statistical Relational Learning Approach. Knowledge-Based Systems. 10.1016/j.knsys.2017.08.017.
- [15] Y. Lu, S. El Helou, D. Gillet, A recommender system for job seeking and re- cruiting website, in: Proceedings of the 22Nd International Conference on World Wide Web, in: WWW '13 Companion, ACM, New York, NY, USA, 2013, pp. 963–966.
- [16] S. Vargas, P. Castells, Rank and relevance in novelty and diversity metrics for recommender systems, in: Proceedings of the 2011 ACM Conference on Recommender Systems, 2011, pp. 109–116.
- [17] G. Adomavicius, J. Zhang, On the stability of recommendations algorithms, in: ACM Conference on Recommender Systems, 2010, pp. 47–54.
- [18] A. Hernando, J. Bobadilla, F. Ortega, J. Tejedor, Incorporating reliability measurements into the predictions of a recommender systems. Information Sciences, in press, doi: 10.1016/j.ins.2013.03.018.
- [19] Mouzhi Ge, Carla Delgado-Baenfeld, and Dietmar Jannach. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In Proceedings of the fourth ACM conference on Recommender systems. ACM, 257–260.
- [20] Li Deng, Dong Yu, et al. 2014. Deep learning: methods and applications. Foundations and Trends® in Signal Processing 7, 3–4 (2014), 197–387.
- [21] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. Deep Learning. MIT Press. [h. p://www.deeplearningbook.org](http://www.deeplearningbook.org).
- [22] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. 2013. Learning deep structured semantic models for web search using clickthrough data. In Proceedings of the 22nd ACM international conference on Conference on information & knowledge management. ACM, 2333–2338.
- [23] zheng. 2016. A survey and critique of deep learning on recommender systems. Department of computer science.
- [24] Trigeorgis, G., Bousmalis, K., Zafeiriou, S., & Schuller, B. W. (2017). A deep matrix factorization method for learning attribute representations. IEEE transactions on pattern analysis and machine intelligence, 39(3), 417-429.
- [25] van Baalen, M. (2016). Deep Matrix Factorization for Recommendation.
- [26] [Guo et al, 2017] Guo, H., Tang, R., Ye, Y., Li, Z., & He, X. (2017). DeepFM: A Factorization-Machine based Neural Network for CTR Prediction. arXiv preprint arXiv:1703.04247.
- [27] Zoran, D., Lakshminarayanan, B., & Blundell, C. (2017). Learning Deep Nearest Neighbor Representations Using Differentiable Boundary Trees. arXiv preprint arXiv:1702.08833.
- [28] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Commi. ee, 173–182.
- [29] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishu Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 7–10.
- [30] [Lian et al, 2017] Jianxun Lian, Fuzheng Zhang, Xing Xie, and Guangzhong Sun. 2017. CCCFNet: 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 Commi. ee, 817–818.

- [31] Xiang Wang, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. 2017. Item Silk Road : Recommending Items from Information Domains to Social Users. Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval (2017).
- [32] Hanh T. H. Nguyen, Martin Wistuba, Josif Grabocka, Lucas Rego Drumond, and Lars Schmidt-Ieme. 2017. Personalized Deep Learning for Tag Recommendation. Springer International Publishing, Cham, 186–197. [h.ps://doi.org/10.1007/978-3-319-57454-7_15](https://doi.org/10.1007/978-3-319-57454-7_15)
- [33] Yuyun Gong and Qi Zhang. 2016. Hashtag Recommendation Using Attention-Based Convolutional Neural Network. In IJCAI. 2782–2788.
- [34] Lei Zheng, Vahid Noroozi, and Philip S. Yu. 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. [h.ps://doi.org/10.1145/3018661.3018665](https://doi.org/10.1145/3018661.3018665)
- [35] Sungyong Seo, Jing Huang, Hao Yang, and Yan Liu. 2017. Representation Learning of Users and Items for Review Rating Prediction Using Attention-based Convolutional Neural Network. In 3rd International Workshop on Machine Learning Methods for Recommender Systems (MLRec)(SDM'17).
- [36] Yang Li, Ting Liu, Jing Jiang, and Liang Zhang. 2016. Hashtag recommendation with topical attention-based LSTM. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.
- [37] Zhenghua Xu, Cheng Chen, .omas Lukasiewicz, Yishu Miao, and Xiangwu Meng. 2016. Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 1921–1924.
- [38] Xiaomeng Liu, Yuanxin Ouyang, Wenge Rong, and Zhang Xiong. 2015. Item Category Aware Conditional Restricted Boltzmann Machine Based Recommendation. In International Conference on Neural Information Processing. Springer, 609–616.
- [39] Ruslan Salakhutdinov, Andriy Mnih, and Geoffrey Hinton. 2007. Restricted Boltzmann machines for collaborative filtering. In Proceedings of the 24th international conference on Machine learning. ACM, 791–798.
- [40] Hanbit Lee, Yeonchan Ahn, Haejun Lee, Seungdo Ha, and Sang-goo Lee. 2016. Quote Recommendation in Dialogue using Deep Neural Network. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. ACM, 957–960.
- [41] Qi Zhang, Jiawen Wang, Haoran Huang, Xuanjing Huang, and Yeyun Gong. [n. d.]. Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network. In the 26th International Joint Conference on Artificial Intelligence (IJCAI 2017).
- [42] Ali Mamdouh Elkahky, Yang Song, and Xiaodong He. 2015. 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. International World Wide Web Conferences Steering Commi.ee, 278–288.
- [43] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM, 353–362.
- [44] Yogesh Singh Rawat and Mohan S Kankanhalli. 2016. ConTagNet: exploiting user context for image tag recommendation. In Proceedings of the 2016 ACM on Multimedia Conference. ACM, 1102–1106.
- [45] Chenyi Lei, Dong Liu, Weiping Li, Zheng-Jun Zha, and Houqiang Li. 2016. Comparative Deep Learning of Hybrid Representations for Image Recommendations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2545–2553.
- [46] Mouzhi Ge, Carla Delgado-Baffenfeld, and Dietmar Jannach. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In Proceedings of the fourth ACM conference on Recommender systems. ACM, 257–260.
- [47] Hao Wang, SHI Xingjian, and Dit-Yan Yeung. 2016. Collaborative recurrent autoencoder: Recommend while learning to fill in the blanks. In Advances in Neural Information Processing Systems. 415–423.
- [48] Piji Li, Zihao Wang, Zhaochun Ren, Lidong Bing, and Wai Lam. 2017. Neural Rating Regression with Abstractive Tips Generation for Recommendation. Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval (2017).
- [49] Yang Song, Ali Mamdouh Elkahky, and Xiaodong He. 2016. Multi-rate deep learning for temporal recommendation. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. ACM, 909–912.
- [50] Taleb Alashkar, Songyao Jiang, Shuyang Wang, and Yun Fu. 2017. Examples-Rules Guided Deep Neural Network for Makeup Recommendation. In AAAI. 941–947.
- [51] Xuejian Wang, Lantao Yu, Kan Ren, Guangyu Tao, Weinan Zhang, Yong Yu, and Jun Wang. 2017. Dynamic Attention Deep Model for Article Recommendation by Learning Human Editors Demonstration. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. ACM.
- [52] Yong Kiam Tan, Xinxing Xu, and Yong Liu. 2016. Improved recurrent neural networks for session-based recommendations. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 17–22.
- [53] Chao-Yuan Wu, Amr Ahmed, Alex Beutel, and Alexander J Smola. 2016. Joint Training of Ratings and Reviews with Recurrent Recommender Networks. Workshop track, ICLR 2017 (2016).
- [54] Fuzhen Zhuang, Dan Luo, Nicholas Jing Yuan, Xing Xie, and Qing He. 2017. Representation Learning with Pair-wise Constraints for Collaborative Ranking. In Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. ACM, 567–575.
- [55] Yao Wu, Christopher DuBois, Alice X Zheng, and Martin Ester. 2016. Collaborative denoising auto-encoders for top-n recommender systems. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining. ACM, 153–162.
- [56] Yuanxin Ouyang, Wenqi Liu, Wenge Rong, and Zhang Xiong. 2014. Autoencoder-based collaborative filtering. In International Conference on Neural Information Processing. Springer, 284–291.
- [57] Cheng Chen, Xiangwu Meng, Zhenghua Xu, and Thomas Lukasiewicz. 2017. Location-Aware Personalized News Recommendation with Deep Semantic Analysis. IEEE Access 5 (2017), 1624–1638.
- [58] [Zhenghua Xu, Cheng Chen, Thomas Lukasiewicz, Yishu Miao, and Xiangwu Meng. 2016. Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 1921–1924.
- [59] Kostadin Georgiev and Preslav Nakov. 2013. A non-iid framework for collaborative filtering with restricted boltzmann machines. In International Conference on Machine Learning. 1148–1156.
- [60] Weizhu Xie, Yuanxin Ouyang, Jingshuai Ouyang, Wenge Rong, and Zhang Xiong. 2016. User Occupation Aware Conditional Restricted Boltzmann Machine Based Recommendation. In 2016 IEEE International Conference on Internet of things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 454–461.