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 computer vision, object recognition and natural language processing, the deep learning gained much attention. The recommendation task is influenced by the deep learning trend which shows its significant effectiveness. The deep learning based recommender models provides 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 is preceded by a presentation of the main lines of the recommendation approaches and the deep learning techniques. 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 [44], [46]. 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 for users as it facilitates the task of information filtering for him especially, the user become lazier and he face usually problems in expressing his information need. Guessing our thoughts by the recommender system seem to be great for time saving purposes. The recommendation process is driven by various embedding 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 four main recommendation models which are collaborative filtering, content-based recommender system, demographic filtering 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, data sparsity and Over-specialization. We mean by the cold-start problem, the issue in which the system is unable to draw any interpretations about items or users due to a lack of information gathered. The data sparsity problem is the issue related to the sufficient information about each item or user in a large data set. The over-specialization problem occur with content-based recommendation as all predicted items will be similar to items rated in last time by the user and then the same topic will be preserved and no new items related to new topic that may interest the user will be predicted. To overcome these problems, many solutions are proposed, 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, object recognition, speech recognition, natural language processing and robotic control where it shows its capability in solving theses complex tasks. In this context, the deep learning has been adopted for enhancing the recommendation approaches by improving the user experiences and ensuring his satisfaction. This is accomplished thanks to the advances of the deep learning in catching items and users' embeddings independently to the data sources nature that might be textual, visual or contextual and predict the suitable recommendations items [2].

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 which impact the collaborative filtering approaches as well as the content based recommendation systems. 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 misplaced 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 four models used for recommendation depending on the nature of the information used as inputs. These models are content based recommender system, collaborative filtering, demographic filtering and hybrid recommendation.

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 basing on items features 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 [42], [43]. 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.

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.

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 as in the collaborative filtering.

Hybrid model: This approach joins more than one type of the recommendation approaches. The content-based approach can be used to predict similar items and the collaborative filtering and demographic filtering is used to more refine the selection according to user preferences. The hybrid model is applied to overcome the limitation of content-based approach and collaborative filtering. The use of hybrid model is possible while, according to the same task, the inputs data is very varied and then we have the flexibility to use different methods simultaneously to improve the quality of the system as a whole. Many combination techniques have been explored. In this context, different weights are given to each recommendation techniques and used for boosting one of them. Another method is to use suggestions issued from one technique as input to process the second technique.

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 approaches are: movies recommendation, news recommendation, e-commerce services recommendation, books recommendation, e-learning recommendation, songs recommendation, websites recommendation, travel destinations recommendation, applications recommendation and so on [2].

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.

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, 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 [6] 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 the 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).

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 [7]-[9]. 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 is the collaborative filtering approach which rely only on past user behavior without requiring the creation of explicit profiles.

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 are 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.

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 [10], [11]. Some systems are based only on the collaborative filtering approach like the work of [12] 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.

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 [13].

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 make them widely applied in practice [14]. 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+.

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. Author in [15] proposes an efficient statistical relational learning approach which is used for constructing a hybrid job recommendation system. Author in [16] 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

Concerning datasets used for the 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

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 output parameters which can be either a rating prediction metrics or ranking score prediction. From the technical system viewpoint, evaluating the ranking score in recommender systems is typically treated as if their main purpose is searching, using metrics from information retrieval such as recall, precision, F1 score and NDCG … .

Besides the accuracy evaluation, other novel evaluation metrics are considered for making better user satisfaction such as novelty, diversity, serendipity, coverage, 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 [17], [45], [47].

The serendipity measure how surprising the relevant recommendations are.

The coverage metric [18] 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 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 [19].

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 [20].

Table II provides the most used set of classical recommender system evaluations metrics as well as the novel ones. Each evaluation metric is associated with its mathematic formulation in the field of recommender systems.

TABLE. II. RECOMMENDER SYSTEMS EVALUATION METRICS

Evaluation metrics			Formula	Description
Standard	*Rating prediction*	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)		
	Ranking score prediction	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 Positve 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
Novel		Novelty	$$Novelty = \sum_{i\in L}\frac{log_2 P_i}{n} \; where \quad P_i = \frac{n - rank_i}{n - 1};$$	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	$$Serendipity = \frac{1}{n}\sum_{i\in n}\max(P_{user} - P_U, 0) \times rel_i$$	
		Diversity	$$diversity = \frac{a}{c}\sum_{l=1}^{c}\frac{1}{n}\sum_{j=1}^{n}i_j$$	
		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 determine 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 layer 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.
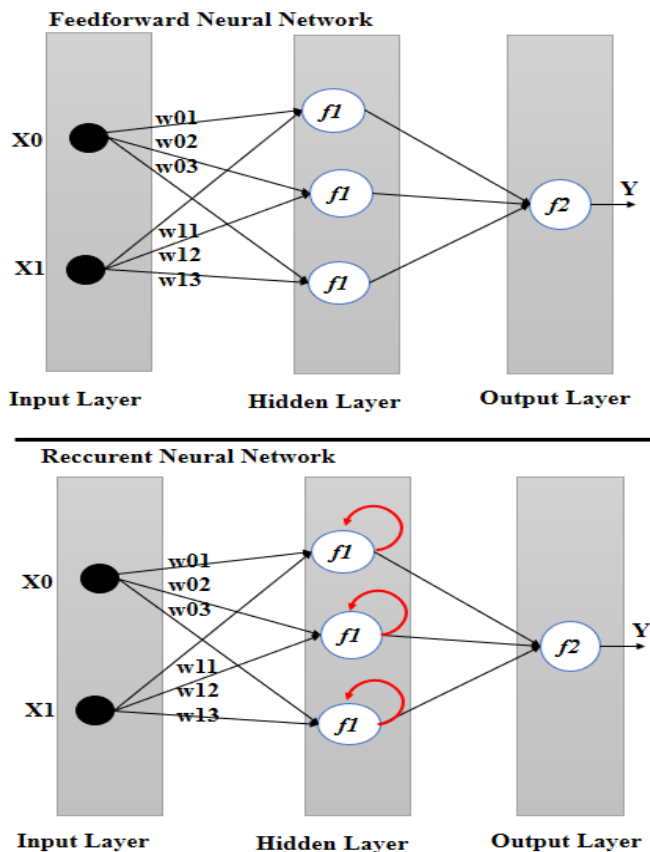

Fig. 1. Feed forward vs. recurrent neural network architecture.

There are several categories of deep learning models. Among these models we cite the following which differ in term of complexity and application fields [21]-[24]:

- Multilayer Perceptron it is a variant of Feedforward Neural Network which is the simplest deep learning approach. 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.

- Unsupervised Learning Networks: This group covers three specific architectures which are Autoencoders, Deep Belief Networks (DBNs) and Generative Adversarial Networks (GANs).

The Autoencoder is similar to 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.

The Deep Belief Network (DBNs) is composed from two types of layers Restricted Boltzmann Machines (RBMs) used for pre-training phase and feed-forward network used for finetuning phase. The specificity of the Restricted Boltzmann Machines layer is the dependence between the two used layers where no intra-communication between them is allowed hence the nomination restricted.

The Generative Adversarial Network is composed of two neural networks. The first network is used for candidates' generation termed as generator and the second for the candidates' evaluation termed as discriminator.

- Convolutional Neural Network: It is a subclass of feedforward neural network with the specificity of pooling operations and convolution layers. The convolution is a powerful concept that helps in building more robust feature space based on a signal. It is successfully applied in the computer vision and image processing domains especially object recognition and image classification.

- Recurrent Neural Network: It is a variant of feed-forward neural networks. It is characterized by its ability to send information over time-steps. using of the internal memory of the network and the loops operations which form the major difference with other feedforward neural network It is widely used in natural language processing, speech recognition and robotic control applications. LSTM (Long Short-Term Memory) is a variant of Recurrent Neural Network model.

- Recursive Neural Network: Its architecture allows the recursive network to learn varying sequences of parts of an image or words using a shared-weight matrix and a binary tree structure. The difference between a Recursive Neural Network and Recurrent Neural Network is that the recursive one can be considered as a hierarchical network where there is no time factor to the input sequence and inputs are processed hierarchically according to a tree structure. This structure is able not only to identify objects in an image, but also to degage the relation between all objects in a scene. It is useful as a scene or sentence parser.

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 different deep learning architectures used in the field of recommender system where we notice that the integration of deep learning is performed with the collaborative filtering model as well as the content-based model where different architectures can be joined in the same system [1]-[3].

C. Deep Collaborative Filtering Recommendation

The Deep learning is applied for enhancing collaborative filtering based recommender system. In fact, there are two most popular area of collaborative filtering which are the latent factor approaches and the 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, matrix factorization, probabilistic matrix factorization, and K nearest neighbors' algorithm [25]-[28].

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 lead its integration in recommendation process. for example, 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. 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.

Authors in [26] have proposed a Deep matrix factorization model in which the deep learning is applied for feature learning. This model is used for CTR prediction and tested commercial data. This system takes the advantage of the Wide & Deep model proposed by Google [29] which joint trained wide linear models and deep neural networks to overcome the sparsity of user-item interactions matrix.

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 [32]-[34].

The unsupervised learning network Restricted Boltzmann Machine is used for recommender system in association with the collaborative filtering techniques 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 [35], [36].

D. Deep Content-based Recommendation

The deep learning is applied also for content-based recommender system. In this case, the main uses of the deep learning deals with the exploitation of the advances of deep learning in thanks to the Convolutional Neural Network for visual features extraction from images and Recurrent Neural Network for extracting textual features and hence improving the recommendation quality [37]-[39].

The choice the appropriate neural network is conditioned by the system requirement. 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 or selecting informative words from textual information [30]-[33].

E. Challenges and Current Issues

The collaborative filtering and the content based filtering models are commonly used for recommendation. The difference between them depends on the nature of the information used as inputs. However, we notice some drawbacks for both of types of recommendation. Indeed, data sparsity, cold start, synonymy and gray sheep problems are the most limitations of the collaborative filtering. Moreover, limited content analysis, over-specialization and new user problems present a crucial shortcoming for the content based filtering. The following table details the aforementioned problems.

The deep learning 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. To overcome limitation of the recommendation systems, the deep learning seems to be a powerful solution as it can be integrated in different steps of recommendation process. The main challenges that they remain undergoing treated by the researcher's community in terms of taking the advances of the deep learning for improving the interactivity with the user, proposing a hybrid social information and boosting the scalability.

TABLE. III.   RECOMMENDATION APPROACHES LIMITATION

Collaborative filtering problems	Content based filtering problems
Data sparsity: the collaborative filtering is not applicable if no enough ratings for both users and items are available	*Limited content analysis:* if the content does not contain enough data for items discrimination, the recommendation may will be not relevant to the user.
Cold start: this problem occurs with new users or items that they have entered the system for the first time, it is difficult to find similar ones because there is not enough information.	*Over-specialization:* we mean by the over-specialization problem the novelty' low degree of the proposed items. This problem refers to the fundamental of a content-based recommender system as it suggest usually similar content for the user and then no new content that may interest him will be suggested.
Synonymy problem: this problem can be faced in the case when several items have the same or very similar names of entries.	*New user:* when there's no enough data about the user to build his profile, the recommendation could be provided imperfectly.
Gray sheep: this problem refers to the users having special opinions which do not agree with any group of people. In this case, this user will not benefit from the collaborative filtering.	

- Interactivity Improving: The interactivity is the bedrock of the recommender system but with diverse degrees. The degree of interactivity is a double edge weapon. In fact, when increasing the interactivity degree, we have more relevant result but we risk disturbing the user. The recommender system refers mainly to implicit information which is extracted without a direct intervention of the user. This type of information is treated and modeled to form the user profile. These data can be entered via forms or other user interfaces provided for this purpose like: age, gender, job, birthday, marital status, and hobbies… or also obtained by observing user's actions like commenting, tagging, bookmarking. This alternative focused on the integration of the user in the recommendation process where he is asked to communicate with the recommender system in an interactive way. The second type of interaction is the explicit interaction which is usually deducted from the user feedback. Here, the user is asked to evaluate the system's results by mentioning if he like/dislike the results or rate them to describe his satisfaction degree. This approach presents an interactive process with user in which the initial suggested items are presented to user who will be asked to judge even a document is relevant or not. This aspect is treated by authors in [30] where they propose a deep learning-based personalized approach for tag recommendation.

- Hybrid social information: Crawling information from different websites and various social media provides hybrid social information which should be very rich and useful for improving the recommendation quality. The deep learning can be used for social network analysis [41] and for the opinion mining and sentiment analysis of users [40].

- Scalability: The deep learning techniques can be applied for boosting the scalability factor as it proves advanced performance in big data analytics. Indeed, 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.

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

YouTube can be defined as the most popular online video platform. Indeed, it reach 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 proposition of how the YouTube's recommendation system is functioning. According to [4], there are some orientation 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 [5]. 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 a 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 benefit from several features which indicate 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 improve 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. YouTube recommendation algorithm is illustrated in Fig. 2.
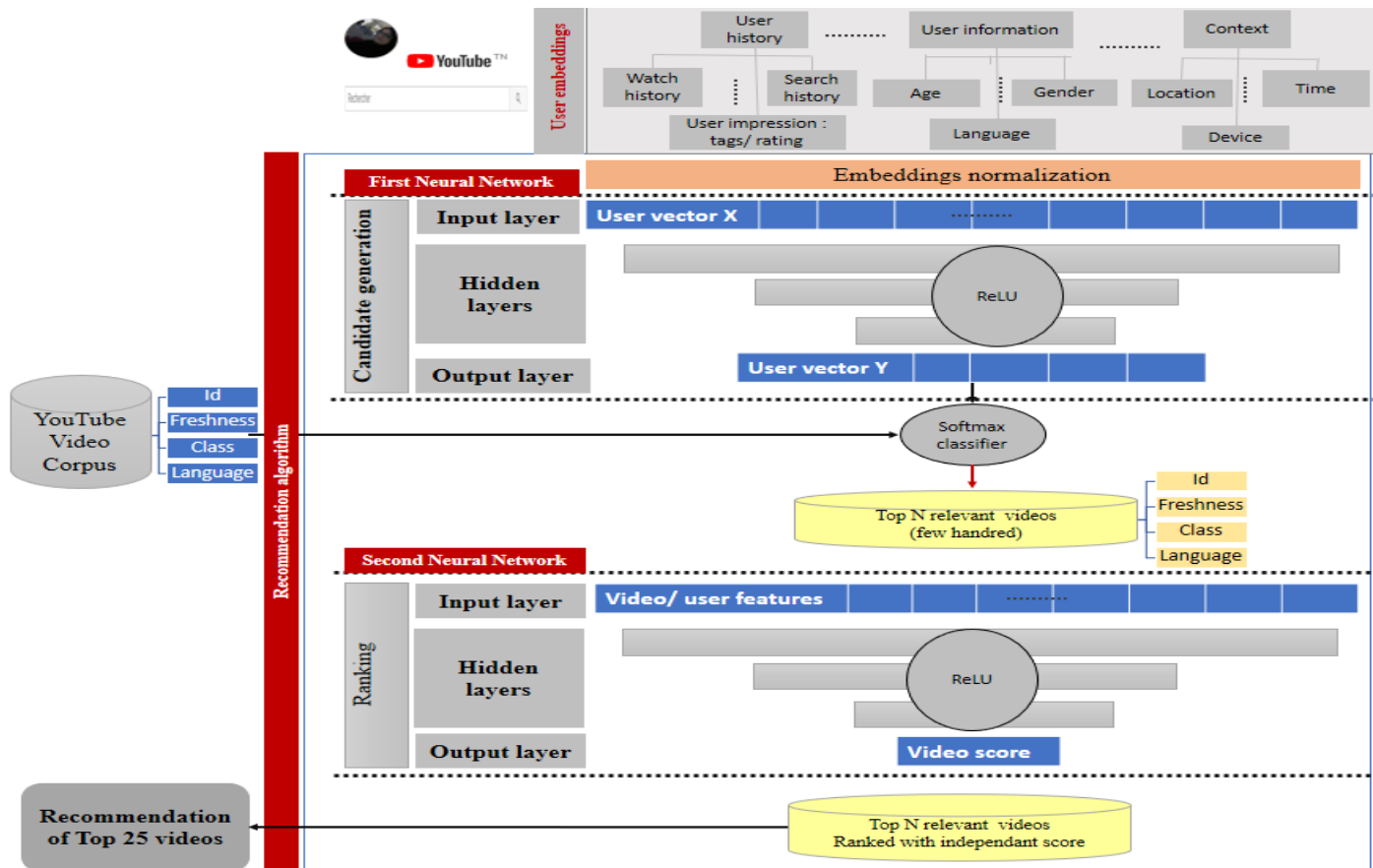

Fig. 2. Feed forward vs. recurrent neural network architecture.

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.

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.

REFERENCES

[1] Liu, Juntao & Wu, Caihua. (2017). Deep Learning Based Recommendation: A Survey. 451-458. 10.1007/978-981-10-4154-9_52.

[2] Zhang, S., Yao, L., & Sun, A. (2017). Deep learning based recommender system: A survey and new perspectives. arXiv preprint arXiv:1707.07435.

[3] Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-based systems, 46, 109-132.

[4] M. Airoldi, 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

[5] 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.

[6] 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

[7] 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)

[8] 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)

[9] 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)

[10] Vladoiu, 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.

[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] 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, SCI, 350, 258–278.

[13] 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).

[14] 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.

[15] 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.knosys.2017.08.017.

[16] 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.

[17] 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.

[18] Mouzhi Ge, Carla Delgado-Ba.enfeld, 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.

[19] G. Adomavicius, J. Zhang, On the stability of recommendations algorithms, in: ACM Conference on Recommender Systems, 2010, pp. 47–54.

[20] 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.

[21] Li Deng, Dong Yu, et al. 2014. Deep learning : methods and applications. Foundations and Trends® in Signal Processing 7, 3–4 (2014), 197–387.

[22] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. Deep Learning. MIT Press. h. p://www.deeplearningbook.org.

[23] 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.

[24] Lei zheng. 2016. A survey and critique of deep learning on recommender systems. Department of computer science.

[25] 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.

[26] van Baalen, M. (2016). Deep Matrix Factorization for Recommendation.

[27]] 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.

[28] Zoran, D., Lakshminarayanan, B., & Blundell, C. (2017). Learning Deep Nearest Neighbor Representations Using Differentiable Boundary Trees. arXiv preprint arXiv:1702.08833.

[29] Cheng, H. T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., ... & Anil, R. (2016, September). Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems (pp. 7-10). ACM.

[30] 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

[31] Yuyun Gong and Qi Zhang. 2016. Hashtag Recommendation Using Attention-Based Convolutional Neural Network. In IJCAI. 2782–2788.

[32] 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.

[33] [Seo et al, 2017] 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).

[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

[35] 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.

[36] 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.

[37] Kumar, V., Khattar, D., Gupta, S., Gupta, M., & Varma, V. (2017, September). Deep Neural Architecture for News Recommendation. In Working Notes of the 8th International Conference of the CLEF Initiative, Dublin, Ireland. CEUR Workshop Proceedings.

[38] Van den Oord, A., Dieleman, S., & Schrauwen, B. (2013). Deep content-based music recommendation. In Advances in neural information processing systems (pp. 2643-2651).

[39] Wang, X., & Wang, Y. (2014, November). Improving content-based and hybrid music recommendation using deep learning. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 627-636). ACM

[40] Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. Int J Adv Comput Sci Appl, 8(6), 424.

[41] Fadlullah, Z. M., Tang, F., Mao, B., Kato, N., Akashi, O., Inoue, T., & Mizutani, K. (2017). State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems. IEEE Communications Surveys & Tutorials, 19(4), 2432-2455.

[42] Fakhfakh, R., Feki, G., Ammar, A. B., & Amar, C. B. (2016, October). Personalizing information retrieval: a new model for user preferences elicitation. In Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on (pp. 002091-002096). IEEE.

[43] Fakhfakh, R., Ben Ammar, A., & Ben Amar, C. (2014, October). Fuzzy user profile modeling for information retrieval. In Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management-Volume 1 (pp. 431-436). SCITEPRESS-Science and Technology Publications, Lda.

[44] Feki, G., Fakhfakh, R., Ammar, A. B., & Amar, C. B. (2016). Query Disambiguation: User-centric Approach. Journal of Information Assurance & Security, 11(3).

[45] Feki, G., Fakhfakh, R., Bouhlel, N., Ammar, A. B., & Amar, C. B. (2016). REGIM@ 2016 Retrieving Diverse Social Images Task. In MediaEval.

[46] Feki, G., Fakhfakh, R., Ammar, A. B., & Amar, C. B. (2015, December). Knowledge Structures: Which one to use for the query disambiguation?. In Intelligent Systems Design and Applications (ISDA), 2015 15th International Conference on (pp. 499-504). IEEE.

[47] Feki, G., Ammar, A. B., & Amar, C. B. (2017, March). Towards diverse visual suggestions on Flickr. In Ninth International Conference on Machine Vision (ICMV 2016) (Vol. 10341, p. 103411Z). International Society for Optics and Photonics.