

A SURVEY ON COMPREHENSIVE TRENDS IN RECOMMENDATION SYSTEMS & APPLICATIONS

Ssvr Kumar Addagarla
School of Computer Science and Engineering, Vellore Institute of
Technology (VIT)
Vellore - 632014, Tamil Nadu, India
ssvrkumar@gmail.com

Anthoniraj Amalanathan
School of Computer Science and Engineering, Vellore Institute of
Technology (VIT)
Vellore - 632014, Tamil Nadu, India
aanthoniraj@vit.ac.in

ABSTRACT

Recommendation System (RS) gains considerable popularity from the past decade in E-Commerce and other allied fields. This paper investigates the various traditional Recommendation System like Content-based (CB), Collaboration Filtering-based (CF), Demographic-based, Knowledge-based and discussed current trends in recommendation system like location-aware, context-aware, and Deep Learning techniques. Various improvements and limitations in Recommendation systems have been listed out with evolution metrics for analyzing the accuracy of the algorithms. This paper well-elaborated for the past, present and future scope of the Recommendation System which would be useful for researchers to get familiarity with this domain.

Keywords: Recommendation System, Content-Based, Collaboration-Based, Knowledge-Based, Demographic, Deep Learning, Recommendation System Applications.

1. INTRODUCTION

In day to day life, we often provide recommendations to others people based on their interests that diverse to various categories. After evolving the web applications earlier 1990s and increasing in the usage of the computers almost by every person in their regular life, the need of recommendation system has been evolved to handle a huge amount of data which are being gathered especially through internet by various websites. Basically, Recommendation systems (RS) provides the personalized and group-based suggestions of various items (e.g., E-commerce products, movies, news articles, holiday spots, insurance policies, websites etc). Nowadays, RS has become an integral part of the majority of e-commerce websites¹. Thus RS such as Netflix, Amazon, Flipkart, Redbus, Uber, LinkedIn, Facebook and other similar websites are interested to identify user preferences based on their activities & information provided on their websites and in turn they try to provide the necessary information and item suggestions (predictions) by using RS algorithm as per their interest. For example, Professional social networking website LinkedIn recommendation engine provide suggestions that people we might know to grow our network on the site. As more than 250 million active users² out of 500 million users currently registered on the LinkedIn's website and not every user is recommended whereas the recommendation model is able to narrow it down and only a few suggestions are made to the users based on the various parameters like working organization, location, working profession, other users collaborations etc that actually people we might know³. In Figure 1, shows classification of RS.

Traditional Recommendation systems can be classified into content-based, collaborative-based, utility-based, demographic-based and knowledge-based systems⁴. Apart from that, there are several new recommendation systems like hybrid RS, trust based social network RS and deep learning based RS⁵. Recommendation systems generally consider the various factors such as users interest, user profile, item features, user to user correlation, item to item correlation, user-item correlation, geographical location etc⁶.

Several research is done by the various researchers in the field of recommendation systems in past decade, in this paper, various types of recommendations system which is illustrated in section 3. In the later section 4 various research works have been discussed within RS. In section 5, the limitations are also stated of RS and finally in section 6 concluded with the possible research scope in the recommendation system.

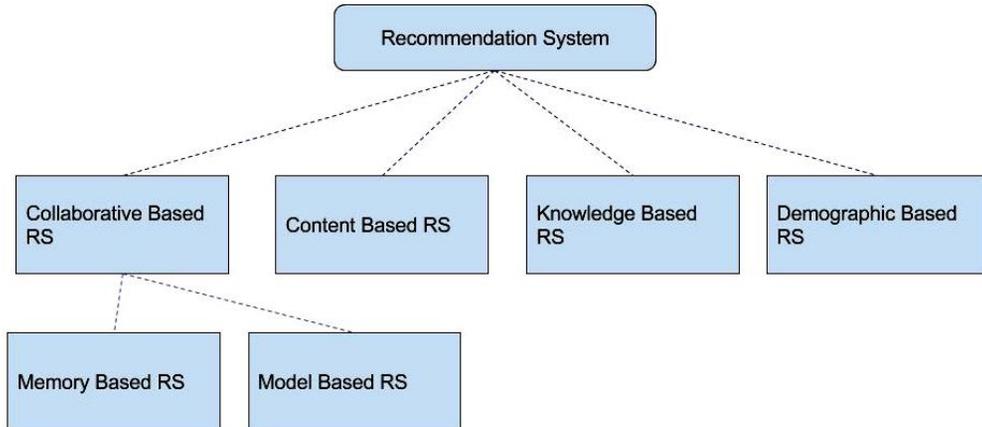


Figure 1. Recommendation system classification

2. THE SIGNIFICANCE OF REVIEW

The number of users in online is growing rapidly especially in developing country like India in the past few years⁷. People tend to depend on various online applications ranging from routine needs to generalize needs like household product purchases, news feeds, music, movies reviews, YouTube videos, digital marketing, web conferences, Facebook, Twitter, and LinkedIn. There is tremendous growth in the usage of the various online applications in the past decade⁸ where the increase in a number of users and their usage in web applications, it is necessary to generate the huge amount of data, based on their searches, interests in viewing items, visiting the web pages, tags, likes or dislikes in social network platforms, comments, reviews, and feedbacks. People have a tendency to confuse for choosing their desires from the wide range of applications, products, newsfeeds, and web links when the data is huge to search. In order to provide better solutions to the users, the need for a recommendation system is required to explore the user experience on the internet with a wide variety of applications. Figure 2 depicts some of the statistics of users' growth in India in terms of buying through online from 2014 to 2020 (estimated)⁷.

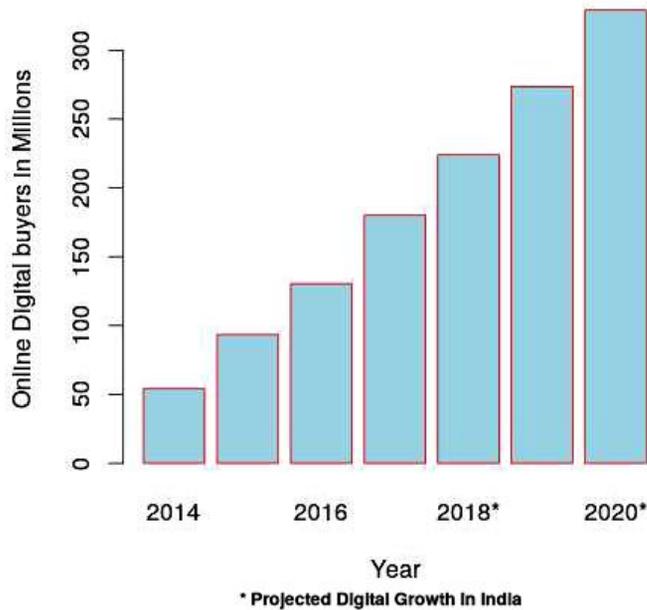


Figure 2. Digital buying growth in India

3. TYPES OF RECOMMENDATION SYSTEMS

Over the past two decades, several types of recommendation systems have been defined based on the advancements and requirements in online and offline user needs. Various researchers have discussed content based RS^{4,9-12}, collaborative RS^{9,13-15}, demographic based RS^{4,16}, utility-based RS⁴, knowledge-based RS^{4,13,17}, hybrid RS^{4,5}, intelligent fuzzy based RS¹⁸, social recommender system⁵, deep learning recommender system¹⁹⁻²¹. In this section, some of the important recommendation systems have been described to analyze and understand the RS domain. In Table 1, below we have mentioned various comparison types of classical recommendation systems.

Table 1. Comparison types of classical recommendation systems

Type	Pros	Cons
Collaborative Filtering RS	<ul style="list-style-type: none"> • Feature Selection not required. • Produces Good results. • Basic domain knowledge sufficient. 	<ul style="list-style-type: none"> • Cold Start problem for New users. • Data Sparsity Problem. • New Item Problem.
Content Based RS	<ul style="list-style-type: none"> • Not required item usage data • Avoids Popularity bias • Comparison between products are possible 	<ul style="list-style-type: none"> • Full item description required • Cold Start Problem for new user. • Lack of diversity
Knowledge based RS	<ul style="list-style-type: none"> • No Ramp-up problem • No need to compute the similarity between the users 	<ul style="list-style-type: none"> • Extensive domain-specific knowledge is required • No dynamic suggestion.
Demographic based RS	<ul style="list-style-type: none"> • Doesn't requires users ratings. • Utilizes the collaborative and content based RS approaches. 	<ul style="list-style-type: none"> • Collection of user information. • Making demographic clustering.

3.1 Content-Based RS

Content-based (CB) recommendations have its roots in Information Retrieval(IR) System community and many of the methods have adopted for recommendations from IR⁹.CB system is generally described by the associated features of an item⁴. Consider an example of a particular product consisting of various features like item name, item type, item department, and item brand which describes the associated features of a particular product. In this approach, preference has given on the item to item correlation than the user to item correlation. Several methodologies are used to learn about CB filtering like Decision Tree, Artificial Neural Networks (ANN), Vector-based representations. CB Recommended System which suggests similar items based on the user history. The Recommendation process can be performed in three stages which are content analyzer, profile learner and filtering components^{12,22}. In the content analyzer, where the type of information can be identified whether structured, semi-structured and unstructured. Especially for the unstructured information it is necessary to apply some pre-processing techniques in order to extract structured

information. Profile learner collects the data relating to users interests, likes/dislikes to construct the user profile by generalization strategy using machine learning techniques. filtering component stage recommends the items to the users based on the existing profile representations. In Figure 3, shows the architecture for the CB filtering RS.

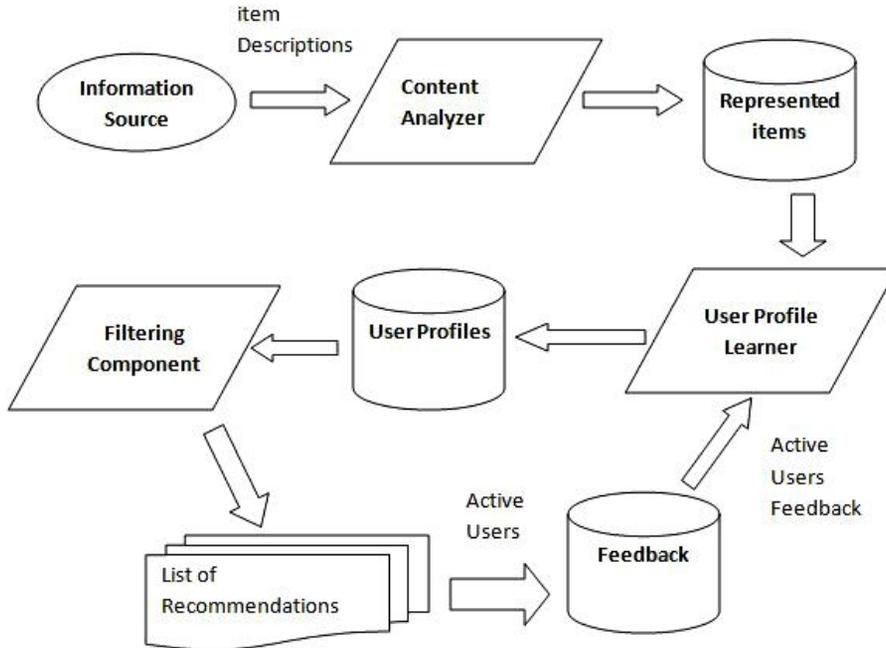


Figure 3. Architecture for the content-based information filtering recommended system

In Xia, Di Fabrizio, Vaibhav, and Datta²³ proposed a Content-based Recommender System for Personalized coupons for e-commerce websites. In the process, authors have analyzed the user's history of a shopping trip to the merchant's website, users similarity, and items to provide the right coupon recommendations to the users. They have framed the framework using random forest and gradient boosted tree classification model. For the experimentations, authors manually collected the users data from the Merchant's website and analyzed.

3.2 Collaborative Filtering RS

Earlier recommender system Tapestry developed at Xerox, they consider user actions and their opinions into the database and developed the novel search system¹⁵. In general, we see similarities between users which their likes, dislikes, interests. Based on these interests we use to suggest products or items to others to the similar group of users. Collaborative

Recommender System aggregates the recommender objects or the user ratings of the items based on the commonalities of the users⁴. Considering the following example, as shown in Figure 4, where three users U_a, U_b, U_c and items called I_1, I_2, I_3 if a user U_a like items I_1, I_2, I_3 user U_b likes items I_2, I_3 and User U_c like I_3 . In this scenario for user U_c , it may recommend the items I_1 and I_2 and for the user U_b it may recommend the items I_1 . CF uses the user to item correlation has more advantages than an item to item correlation⁵.

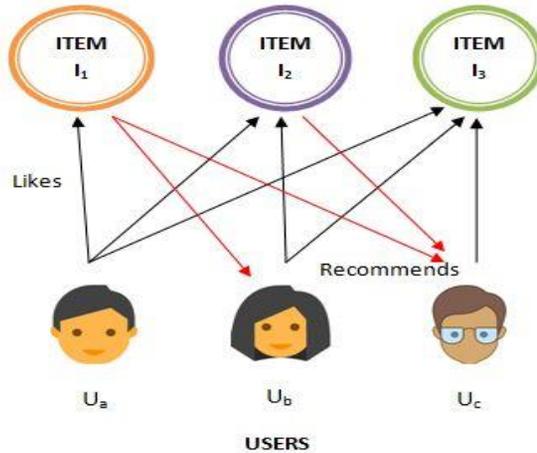


Figure 4. Collaboration filtering

Collaborative filtering is based on the similar preferences from the users who rate similarly¹⁵. There is need of sophisticated filtering techniques in order to recommend likely products to the users. For this researchers are contributed two major techniques for collaboration Filtering recommender system, which is memory based and model based⁴. Memory-based algorithms use the information for user-item correlation to recommend a product to the user which doesn't recommend before²⁴. Memory-based algorithms compute similarity measures using various approaches like Pearson correlation, cosine similarities, mean squared difference, Jaccard coefficient, Spearman correlation¹⁴, K-nearest neighbor Algorithm^{25,26}. The main challenges in the memory based techniques are data sparsity and scalability, where sparsity refers to most of the users rates only fewer items with that given model provide us the less accurate results, on the other hand, scalability refers to increasing inflow information (both the users and items) which doesn't handle the memory based techniques¹⁴. In Model-based techniques which use to predict at what extent user will like a product/item which doesn't occur previously. For this utilization of machine learning models in order to train a vector of items for a particular user, which can predict the rating of a new item that has been added using genetic

algorithms, matrix factorization technique, Bayesian classifier, fuzzy systems¹⁸, neural networks, and latent features²⁴.

Collaborative filter based RS is the most commonly used technique and in Dang and Viennet²⁷, authors proposed a social network community-based approach which uses the various e-commerce sites as Last.fm, Delicious and Epinion uses CF based RS to Predicts the similar type of tastes from their social network groups from the users relations using cosine similarity, LPTN Method and item ratings across various e-commerce websites.

3.3 Knowledge-based RS

In collaboration filtering, RS users are correlated with the item ratings. Suppose a system consisting of less number of user ratings for the items in a large amount of data, and then the system does not effectively recommend items to the users. This is commonly referred as ramp-up problem¹³. In Knowledge-based recommendation system does not have ramp-up problem where the recommender system does not depends on the user ratings and need not gather the individual user preferences because the recommender system judgments are varied to individual tastes, and knowledge base RS have the functional knowledge²⁸ that should correlate the needs of user and item⁴. In the knowledge-based RS users directly specifies the needs what they exactly want and precisely the recommender system is customizable to the specific domain¹⁷.

In Aggarwal¹⁷ discussed one of the types of knowledge-based RS is constraint-based recommender system for Online Home buying. They have taken various customer-specified attributes like marital status, city, family, Minimum required bedrooms, Maximum price etc for the data process. Based on the customer profile which is then mapped with user input values are used to recommend the desired products.

3.4 Demographic RS (DRS)

Burke⁴ elaborated demographic recommendation system relay on the various personal attributes of the users and classify them into various classes and make recommendations based on the categorizations. In DRS gathers the information of demographics of the user U and its ratings I and identifies the similar demographic users u and try to explore the user ratings of the item I to provide recommendations in demographic classes²⁸.

Zhao, Guo, He, Jiang, Wu, and Li²⁹ proposed new demographic recommender systems on microblog for product recommendations to the users. In this process they have come with a framework called METIS (merchant intelligent recommender system) consist of purchase intention detection based on the user tweets on microblogging platform, demographic

information extraction of the users from the microblogging and product demographic extraction and finally the proposed algorithms compute the similarity of the products and users are performed and which subsequently produces the recommendations to the users. For rigorous experimentation and results analysis authors has used the largest Chinese microblog platform Sina Weibo and largest e-commerce company Jingdong for product e-commerce database.

3.5 Location Aware RS

Location-based recommendation system extracts the information from mobile devices that use to understand the user location and based on the user rating history recommends various places like movie theaters, parks, malls, museums, events near user location³⁰. This motivates to further extension of location aware recommendation system (LARS). Levandoski, Sarwat, Eldawy, and Mokbel³¹ proposed the new paradigm called location aware recommendation system which methodology primarily works on three different kinds of ratings varies from spatial user ratings for nonspatial items, non-spatial user ratings for spatial items, spatial user ratings for spatial items which consider user, user location, rating, item, item location. In this LARS uses pyramid maintenance algorithm and travel maintenance algorithm which provides better recommendations than traditional RS.

Pálovics, Szalai, Pap, Frigó, Kocsis, and Benczúr³² proposed a location-aware online learning recommender system for top product recommendation using geospatial hashtags from twitter data. In their findings methods such as matrix factorization, event content based hashtag recommendation, nearest neighbor methods²⁶ are weaker in the Product recommendations. whereas in the author's framework uses the measuring popularity, learning the importance of locations, geographical hierarchy to recommend top N-product recommendations and the developed method can be applied for hotels, movies, restaurants etc.

3.6 Context-Aware RS

Regular recommendation systems rely on the attributes of the user, item, rating to predict recommendations using various approaches. To enhance the recommendations a new attribute called context³³ is introduced to better understand and recommend the items to the users under specific circumstances like when in a day, at what time³⁴, location, seasons etc³⁵.

$$\textit{Contextual RS} : (\textit{User} \times \textit{Item} \times \textit{Context}) \rightarrow \textit{Rating}$$

This contextual information³⁴ can be obtained from various approaches like explicit, implicit and inferring. Contextual user preference gathering and recommendations can be performed based on contextual pre-filtering,

contextual post-filtering, contextual modeling³⁵.

In Colombo-Mendoza, Valencia-Garcia, Rodriguez-González, Alor-Hernández, and Samper-Zapater³⁶ proposed context-aware mobile recommendation for Movie show time recommendations. In the proposed framework called RecomMetz's mainly rely on the user position, user profile and time. When a user request for the recommendations it uses the user position, profile and time builds the contextual user preferences that recommend the movies, movie theaters and their show times. Apart from the above authors have introduced the crowd as another parameter which analysis active users around for a particular movie theater and used the social networking websites especially like Foursquare as a popular location-based social network.

3.7 Topic-Based RS

There are numerous text data available in various forms on the internet. In order to get the preferred text information and articles based on the popularity, topic recommender system detects to fetch the right topics from the tag cloud to the user preference³⁷.

There is tremendous growth in micro-blog websites around the world. In the growing digital era, which is difficult to find the relevant and interesting topics in the microblog is a very tedious process. Zhang et al.³⁸ had developed a topic recommender system for the Sina Weibo a popular Chinese micro-blog website which more than 500 million register users. To overcome the users search problem in Sina Weibo, authors has proposed the topic based micro-blog recommender system based on topic based clustering. In the process authors defined four stages of development to recommend the topics namely user data processing, extraction of topic features, user clustering and microblog topic recommendations. The results shown micro-blog topic recommends to the users up to some extent but not full accurate³⁸.

3.8 Artificial intelligence based RS

Artificial intelligence (AI) techniques like fuzzy logic, artificial neural networks (ANN), artificial immune systems (AIS), swarm intelligence (SI), evolutionary computing (EC) are playing a vital role in recommendation systems since a decade. Fuzzy logic is used to finding out the pattern recognition and handling of the uncertainty which able to deal with missing data. ANN supports and suits for unsupervised learning and classification problems which can determine the similar types of users. AIS are useful to overcome the ambiguity issues and preference matching. Swarm intelligence algorithms are non-linear, self-organized and scalable which used to understand the collective behavior of the users to determine the

nearest neighbors and performs clustering in a dynamic environment. Evolution computing³⁹ has defined from concepts of evolutionary biology used to optimization problems and continuous optimization. There are several methodologies are available to tackle the optimization issues namely genetic algorithms, evolutionary programming and strategies. These are helpfull in recommendation systems to finding out best neighborhoods and optimizing online tasks⁴⁰.

In Katarya and Verma⁴¹ proposed a music recommendation system using particle swarm optimization which is an AI-based Technique for recommendations. Authors have proposed a hybrid music RS combining collaborative, context features along with Swarm Intelligence. For this process, they have taken the real world data set from the Last.fm a popular music listing online portal more than 19 million music listings. They have extracted the various temporal fields from dataset such as day, month and time stamp. Based on the temporal attributes, utilization of the graph based algorithms Depth-first search, Bellman-Ford algorithm, and Particle swarm optimization techniques provides the popular music recommends to the users from their preferences and behavior in different temporal features.

3.9 Natural Language Processing (NLP) Based RS

Natural Language Processing (NLP) can understand the sentences and its structure, idioms and able to understand the human generate speech. In the context, NLP has the ability to analyze the various forms of data and performs the analysis of syntax, semantics and pragmatics⁴². In Recommendation systems NLP especially used to perform the classification, clustering, automatic learn rules by analyzing a set of examples. Methods like WordNet synsets, keyword-based feature vectors, latent semantic indexing and term frequency-inverse document frequency (TF-IDF) are popular classification and clustering techniques^{43,44}.

Kamath and Kanakaraj⁴³ proposed an e-news recommendation system uses natural language processing through information extraction and domain clustering. For the process, they have developed offline and online models to process the data for e-news recommendations to the user. They have categorizes the various articles like a business, cinema, politics, science, and technology, world, sports etc. For the experimentation, they have used the popular India e-news websites The Times of India, The Hindu, The Indian Express and compared the results with existing plain keyword model and which shows 3% - 5% improvements in article recommendations.

3.10 Deep Learning (DL) Based RS

Deep learning²¹ is sub research area of Machine Learning, which gains much popularity in recent years. Traditional recommendation system alone is not sufficient to generate accurate recommendations for the new age user needs with a lot of data. It needs to a deep understanding of the users, items as well as user-item interaction. Generally three kinds of recommendation tasks rating, ranking and classification predictions applied in traditional RS. There are several deep learning techniques are widely using in various fields^{20,45-47} which solves the supervised learning and unsupervised learning tasks which provide superior results when compared to traditional RS²¹. deep learning (DL) based recommender system can be implements in various domains by understanding of various DL techniques such as multilayer perceptron (MLP), auto encoder (AE), convolutional neural networks (CNN), recurrent neural networks (RNN), deep semantic similar model (DSSM), restricted Boltzmann machine (RBM), neural autoregressive distribution estimation (NADA), generative adversarial networks (GAN), DL based natural language processing (NLP). DL is becoming popular in the last couple of years in recommendation systems to tackle various recommendation tasks like articles suggestion, music, news feeds, fashion etc. Matching the right user preferences is a big challenges nowadays from the heterogeneous data under various circumstances. Deep learning is a complex computational model consist of multiple hidden layers to process and analyze the various forms of data for right preferences to users⁴⁸.

Nguyen, Wistuba, Grabocka, Drumond, and Schmidt-Thieme⁴⁹ proposed a personalized deep learning based tag recommender system which implements through the architecture model called CNN Per MLP using the relation between the visual search image features and user. The proposed deep neural networks uses the CNN, MLP as predictor and Bayesian ranking optimization for ranking used to predicts and compute relevant tags and user images as input to produce the rankings of the relevant tags. For their implementation, they have considered the real world data sets are NUS-WIDE and Flickr-PTR for the results and comparison with the existing methods and its shows a marginal improvement from the compared methods.

4. RELATED LITERATURE SURVEY

There are several recommendation systems introduced by various researchers in mid-1990s. In Balabanović and Shoham⁹, the Fab system developed using content-based and collaborative recommendation system which is a distributed application developed as part of the library project at

Stanford University. In these experimental results proved up-to some extent of combining collaboration and the content-based system works well but not predicts the scalability issues.

Early-1999, the use of e-commerce growing steadily and e-commerce companies are attracting the users by using various recommendation systems. Schafer et al.¹⁵ has discussed various e-commerce platforms (Amazon, eBay, levis, moviefinder.com, cdnow etc) along with various recommendation strategies like non-personalized, attribute-based RS, item-item correlation people-people correlation and explores various e-commerce opportunities using recommender system. In this research, mainly focused on how best to improve the revenue of the e-commerce using recommended system has stated.

Later couple of years some good research work was done by the Robin Burke^{13,4}. In his research elaborated various RS types and noted about the importance of knowledge-based RS¹³, and hybrid RS⁴. In his survey comes with proven experimental results for model hybrid recommended a system which is EntreeC, A combination of knowledge-based RS and collaboration RS. The point of interest with the experiment of EntreeC is to prove that collaborative filtering technique has improved performance over knowledge-based RS acting alone.

As in mid-2000's the rapid growth in digital buying and the need of recommender system is much important for any electronic commerce based company. In that direction¹⁸ proposed a personalized online recommender system using fuzzy rules and techniques. Here uses both domain knowledge of the system and ephemeral of information by the user-generated with multi-attribute decision-making method to get the personalized products by the system effectively.

Vozalis and Margaritis⁵⁰ proposed new hybrid algorithms known as U-Demog and I-Demog which combines the collaborative filtering algorithms such as user based and item based algorithms with the demographic correlations between users or items and the similarity computed using the dot product of two vectors, which produced better recommendations. For this authors has taken Movie Lens dataset for experimental results⁵⁰.

Pazzani and Billsus¹¹ authors had discussed content-based recommender system for item representation on structured, semi-structured and unstructured data. This survey also focused on learning and modeling of user profiles using decision tree algorithms like ID3, rule induction algorithm like RIPPER, nearest neighbor methods, relevance feedback and Rocchio's Algorithm¹¹.

Unlike other recommender systems, the collaborative recommended

system has much scope to get better results for the adaptive web. In Schafer et al.¹⁵ briefly discussed various CF algorithms for adaptive web especially under nonprobabilistic algorithms like user based and item based nearest neighbor and its practical challenges. Then addressed the cold-start problem & possible solutions for a new item, new user, new community but not optimal. Also explores the some of the open challenges in privacy, security, and trust in recommendation system.

Gong, Ye, and Tan¹⁴ discussed the importance of combining the CF methods which are memory based and model-based to overcome the scalability and sparsity problem and produce the better recommendations compares to the traditional Strategies. As in this process new item, new user, new community problems are not fully addressed¹⁴.

Su and Khoshgoftaar¹ had surveyed various CF techniques including memory based, model-based and hybrid CF. This survey found useful in the CF algorithm such as MDP (Markov decision process) for prediction problem and latent semantic index (LSI) is a statistical method to uncover user communities and model interest profiles along with traditional CF techniques such as clustering and regression. It's important and necessary to perform computations of various prediction accuracy metrics in order to achieve accurate recommendations such as predictive accuracy, classification accuracy, rank accuracy and normalized distance based performance¹. In Su and Khoshgoftaar¹ studies, authors elaborate on metrics such as mean absolute error, root mean squared error and receiver operative characteristics (ROC) sensitivity metrics and finally, authors highlights the needs to overcomes the problems like data sparsity, shilling attack and noisy data in future recommendation systems. Other classification metrics such as precision, recall, F1-measures discussed in Bobadilla, Ortega, Hernando, and Gutiérrez²⁴.

Cremonesi et al.⁵³ presented a cross-domain recommendation system (CD-RS) which is based on the knowledge of one domain (source) should be recommended items on other domain (destination) using artificial dataset by considering various CF based algorithms⁵³. Fernández-Tobias et al.⁵⁴ provides CD-RS for ranking and diversity for Cold-start Problem using Facebook & DBpedia dataset. Whereas ranking accuracy for cold-start problem has improved and diversity showed different results which varies from the domain.

As the CB filtering techniques have an important role to recommend items to users for some extent, In¹² authors presented various CB internal information filtering methods such as content analyzer, profile learner and filtering component technique, also it is necessary to understand how an item is relevant& useful to the user by providing explicit evolutions like

relevance feedback techniques. There are several drawbacks mentioned in the CB approach like limited content analysis, serendipity & New user problem and the Authors also suggested solutions to limited content analysis through semantic analysis methods, authors provide various views of other researcher's possibilities to overcome serendipity and new item problem but not provided any optimal solution¹².

New user or new item problem⁴ generally referred as a cold-start problem¹⁵ common in CB and CF recommendation system which was addressed in Safoury and Salah⁵¹ by providing an evolution framework using user demographic attributes by considering the Movie Lens data set of 100K. and accuracy measure used for the framework is mean absolute error, root mean square error, precision, recall and F-score, as this approach solves the cold-start problem up to some extent but not completely. In Al-Shamri⁵² summarized and experimented using unified profiling and isolated profiling using Movie Lens dataset for improvement in user profiling approaches for demographic recommender system.

Zhang et al.³⁸ surveyed various deep learning techniques like multilayer perceptron (MLP), autoencoder (AE), convolutional neural networks (CNN), recurrent neural networks (RNN), deep semantic similar model (DSSM), restricted Boltzmann machine (RBM), neural autoregressive distribution estimation (NADA), generative adversarial networks (GAN). Apart from that author describes various approaches like single DL techniques, composite models and integrate DL with the traditional RS. Among these AE, RNN, MLP is much widely using DL techniques in the fields like image processing, video recommendations, speech reorganization etc²¹.

Wei et al.⁴⁵ had proposed a deep learning based solution for the cold-start problem for items using a combination of CF based Singular Value Decomposition (SVD++) and deep learning based Stacked by multiple denoising autoencoders (SDAE) by classifying cold-start items into classified to complete cold-start (CCS) and incomplete cold-start (ICS). This approach uses Netflix dataset and recommendation results shown improvement in CCS items but not in ICS items⁴⁵.

5. RS IN E-COMMERCE

The e-commerce industry is growing rapidly and attracting millions of users day by day and which is an unavoidable market sector nowadays. Many of the RS methodologies have already adopted for various portals like Amazon, Flipkart, eBay, Snapdeal for product recommendations⁵⁵.

Here in Figure 5, we presented global e-commerce retails statistics and

future trends for top 10 E-commerce markets from 2015 to 2018⁵⁶.

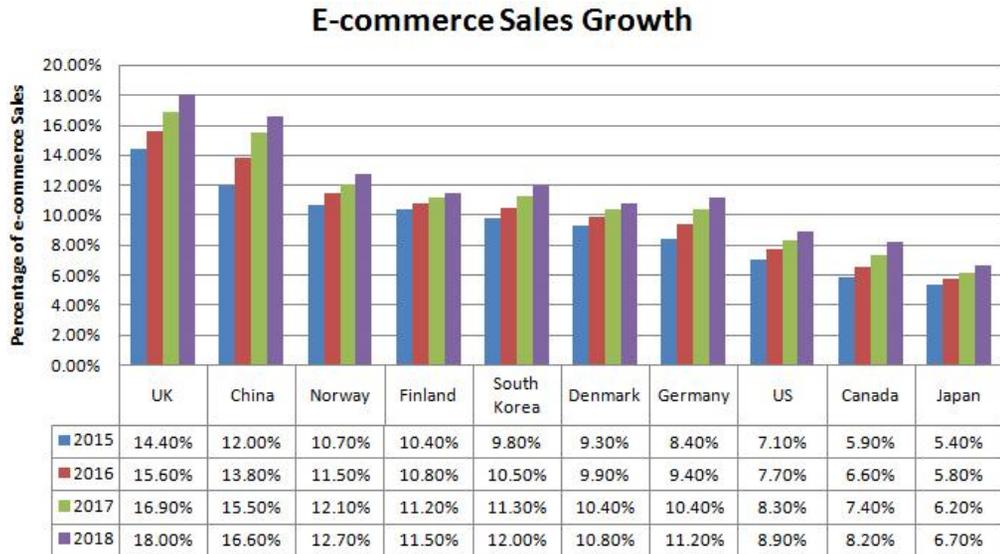


Figure 5. Top 10 countries with e-commerce sales growth

In Figure 6, presented the percentage of frequent shoppers in the global retail e-commerce industry for the year 2016, which we categorized into age groups as frequent shoppers, occasional shoppers, and infrequent shoppers⁵⁶.

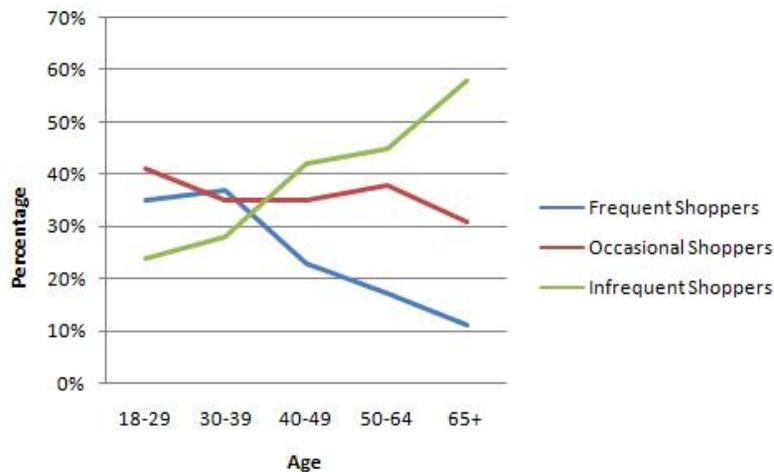


Figure 6. Percentage of frequent shoppers

Based on the above statistics, there is a tremendous need of the intelligent Recommendation System which will scatter the needs of millions of people around the globe, especially for the future e-commerce industry.

5.1 Applications

Here we have listed out some of the real-world applications that make use of the various recommendation system approaches.

Achakulvisut et al.¹⁰ has developed a content-based RS to recommend the scientific publications using topic modeling techniques and tested on Society of Neuroscience 2015 conference dataset, where CB based latent semantic analysis, rocchio algorithm & nearest neighbor techniques are used to get the accurate results based on user search keywords¹⁰.

Davidson et al.⁵⁷ presented a personalized video recommendation system for popular YouTube based on the user activities like user watched videos, likes, favorite. For this author used similarity measures for the videos and ranked for top N-recommendations⁵⁷.

Wang et al.¹⁶ had presented a demographic based RS for the tourist attractions by focused on six cities: New York, Paris, London, Rome, Chicago and Berlin using TripAdvisor dataset considering museums, zoos & aquariums. for this recommendations, authors used machine learning approaches along with the demographic user data to predict the ratings and to give attraction¹⁶.

Chu and Tsai⁵⁸ had proposed a hybrid recommender system (combination of CB and CF) by considering both text and visual information from Taiwan's largest restaurant dedicated social networking platforms to predict restaurant recommendations to the users⁵⁸.

Cross-Lingual product recommender system⁵⁹ has proposed using naive CF for non-Japanese to recommend products on Japanese e-commerce website Rakuten Ichiba and collaboratively identifies the same items at Grouplens website & Japanese e-commerce site using translation pairs.

Mobile E-commerce is widely spreading nowadays and in Guo et al.⁶⁰ had proposed an enhanced apriority algorithm for fast, accurate access and retrieval of frequent itemsets for mobile e-commerce without losing customer satisfaction. Experiments conducted by taking the dataset from Taobao.com which is leading e-commerce website under Alibaba Group⁶⁰.

Zhao et al.⁶¹ proposed a demographic based product recommender system by analyzing the social networking website and shopping site. For this process authors used the Chinese microblogging site SINA WEIBO and e-commerce website JINGDONG for analyzing the tweets and online reviews for better product recommendation using user demographic attributes by matching the user profiles⁶¹.

Wu and Yan⁶² has discussed session-aware e-commerce recommendation system for anonymized users using deep neural network

framework by framing the session information embedding and list-wise ranking framework. For this author used the CIKM Cup 2016 dataset be used to test the algorithm for product recommendations⁶².

Jing et al.⁴⁷ discussed for the low-cost computations using open source deep learning framework caffe and Amazon web services used for the visual search on Pinterest using deep convolution neural networks and click-through strategies for similar item recommendations⁴⁷.

Visualization search recommendations become widely popular in the last couple of years. Shankar et al.⁴⁶ proposed a new visualization search method called VisNet based on the unified deep convolutional neural networks. The authors used the Street2Shop dataset for testing and later on applied for India's leading online e-commerce site Flipkart which proves superior visual recommendations comparing to existing methods like AlexNet, PatternNet, ColorHist⁴⁶.

6. SUMMARY

Recommendation systems have its roots in the information retrieval, cognitive science, and forecasting theories. A review on content-based, collaborative filtering, knowledge-based, demographic-based, location aware, context-aware, topic modeling, NLP, AI based RS and deep learning based RS has been presented in this paper. It is identified that most of the approaches the RS are not addressed the problems like scalability, sparsity, synonyms, gray sheep attack, and cold-start & shilling attack. Some of the hybrid RS addressed the scalability, sparsity, cold-start problems but not completely overcome and cold-start problem still an open challenge for the researchers. Recent advancements in deep learning techniques which are a part of the machine learning, providing underlying solutions for the many problems for the traditional RS that were also discussed. There are several deep learning methodologies like MLP, auto-encoders, CNN, RNN which already gained popularity in the fields like image retrieval, Image processing, speech reorganization etc. Some of other DL techniques like GAN, DSSM, MLP, RBM which needed to give more attention from the researchers and these to be integrated with traditional RS in order to develop a deep hybrid models to solve the traditional RS problems. There is a lot of scope to implement these deep hybrid models on e-commerce industry for the accurate recommendations and predictions for the millions of users in the digital world.

7. REFERENCES

- [1] X. Su, and T. M. Khoshgoftaar, A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 4, Article ID 4214252009. <http://dx.doi.org/10.1155/2009/421425>.
- [2] Omnicore, *Linkedin by the Numbers: Stats, Demographics & Fun Facts*. Retrieved on August 30, 2018, from <https://www.omnicoreagency.com/linkedin-statistics>.
- [3] sigmoidal, *Recommendation Systems - How Companies are Making Money - Sigmoidal*. Retrieved on August 30, 2018, from <https://sigmoidal.io/recommender-systems-recommendation-engine>.
- [4] R. Burke, Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4), p.331-370, 2002. <https://doi.org/10.1023/A:1021240730564>.
- [5] R. Prasad, and V. V. Kumari, A categorical review of recommender systems. *International Journal of Distributed and Parallel Systems*, 3(5), p.73-83, 2012. <http://dx.doi.org/10.5121/ijdps.2012.3507>.
- [6] G. Adomavicius, and A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), p.734-749, 2005. <http://dx.doi.org/10.1109/TKDE.2005.99>.
- [7] statista, *Digital buyers in India 2020 | Statistic*. Retrieved on January 27, 2018, from <https://www.statista.com/statistics/251631/number-of-digital-buyers-in-india>.
- [8] Inc42, *Battle Of The Indian Ecommerce Marketplaces - In Depth Comparison*. Retrieved on January 27, 2018, from <https://inc42.com/resources/ecommerce-marketplace>.
- [9] M. Balabanović, and Y. Shoham, Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3), p.66-72 ,1997. <http://dx.doi.org/10.1145/245108.245124>.
- [10] T. Achakulvisut, D. E. Acuna, T. Ruangrong, and K. Kording, Science concierge: A fast content-based recommendation system for scientific publications. *PLOS ONE*, 11(7), e0158423, 2016. <http://dx.doi.org/10.1371/journal.pone.0158423>.
- [11] M. J. Pazzani, and D. Billsus, Content-based recommendation systems. In P. Brusilovsky, A. Kobsa, W. Nejdl (eds.), *The Adaptive Web. Lecture Notes in Computer Science* (p325-341). vol 4321, Springer, Berlin, Heidelberg, 2007.
- [12] P. Lops, M. De Gemmis, and G. Semeraro, Content-based recommender systems: State of the art and trends. *Recommender Systems Handbook*, p73-105, 2011. http://dx.doi.org/10.1007/978-0-387-85820-3_3.
- [13] R. Burke, Knowledge-based recommender systems. *Encyclopedia of*

- Librart and Information Systems*, 69 (Supplement 32), p175-186, 2000.
- [14] S. Gong, H. Ye, and H. Tan, Combining memory-based and model-based collaborative filtering in recommender system. *Paper Presented at Pacific-Asia Conference on Circuits, Communications and Systems*, Chengdu, China, May 16-17, 2009. <http://dx.doi.org/10.1109/PACCS.2009.66>.
- [15] J. Ben Schafer, D. Frankowski, J. Herlocker, and S. Sen, Collaborative filtering recommender systems. In P. Brusilovsky, A. Kobsa, W. Nejdl (eds.), *The Adaptive Web. Lecture Notes in Computer Science* (p291-324). vol 4321, Springer, Berlin, Heidelberg, 2007.
- [16] Y. Wang, C. F. Chan, and G. Ngai, Applicability of demographic recommender system to tourist attractions: A case study on trip advisor. *Paper Presented at IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, United Sates, December 4-7, 2012. <http://dx.doi.org/10.1109/WI-IAT.2012.133>.
- [17] C. C. Aggarwal, Knowledge-based recommender systems. *Recommender Systems*, p167-197, 2016. http://dx.doi.org/10.1007/978-3-319-29659-3_5.
- [18] Y. Cao, and Y. Li, An intelligent fuzzy-based recommendation system for consumer electronic products. *Expert Systems with Applications*, 33(1), p230-240, 2007. <http://dx.doi.org/10.1016/j.eswa.2006.04.012>.
- [19] S. Li, J. Kawale, and Y. Fu, Deep collaborative filtering via marginalized denoising auto-encoder. In *CIKM'15, Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (p811-820). 2015. <http://dx.doi.org/10.1145/2806416.2806527>.
- [20] A. Singhal, P. Sinha, and R. Pant, Use of deep learning in modern recommendation system: A summary of recent works. *International Journal of Computer Applications*, 180(7), p17-22, 2017. <http://dx.doi.org/10.5120/ijca2017916055>.
- [21] S. Zhang, L. Yao, and A. Sun, Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 52(1), Article No. 5, 2017. <http://dx.doi.org/10.1145/3285029>.
- [22] C.C. Aggarwal, Content-based recommender systems. *Recommender Systems*, p139-166, 2016. http://dx.doi.org/10.1007/978-3-319-29659-3_4.
- [23] Y. Xia, G. Di Fabbriozio, S. Vaibhav, and A. Datta, A content-based recommender system for e-commerce offers and coupons. *Paper Presented at Proceedings of the SIGIR 2017 eCom workshop*, Tokyo, Japan , August, 2017.
- [24] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, Recommender systems survey. *Knowledge-Based Systems*, 46, p109-132, 2013. <http://dx.doi.org/10.1016/j.knosys.2013.03.012>.

- [25] C. N. Ziegler, Taxonomy-driven filtering. *Social Web Artifacts for Boosting Recommenders. Studies in Computational Intelligence*, 487, p23-45, 2013. http://dx.doi.org/10.1007/978-3-319-00527-0_3.
- [26] C. Desrosiers, and G. Karypis, A comprehensive survey of neighborhood-based recommendation methods. In *F. Ricci, L. Rokach, B. Shapira, P. Kantor (eds) Recommender Systems Handbook* (p107-144), 2011. http://dx.doi.org/10.1007/978-0-387-85820-3_4.
- [27] T. A. Dang, and E. Viennet, Collaborative filtering in social networks: A community-based approach. *2013 International Conference on Computing, Management and Telecommunications*, p128-133, 2013. <http://dx.doi.org/10.1109/ComManTel.2013.6482378>.
- [28] F. Ricci, L. Rokach, and B. Shapira, Recommender systems: introduction and challenges. In *F. Ricci, L. Rokach, B. Shapira (eds) Recommender Systems Handbook* (p1-34). Springer, Boston, MA, 2015. http://dx.doi.org/10.1007/978-1-4899-7637-6_1.
- [29] X. W. Zhao, Y. Guo, Y. He, H. Jiang, Y. Wu, and X. Li, We know what you want to buy: A demographic-based system for product recommendation on microblogs. In *Proceedings of the 20th ACM SIGKDD International conference on Knowledge discovery and data mining* (p1935-1944), 2014. <http://dx.doi.org/10.1145/2623330.2623351>.
- [30] Wikipedia, *Location-based recommendation*. Retrieved on January 28, 2018, from https://en.wikipedia.org/wiki/Location-based_recommendation#Background.
- [31] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel, Lars: A location-aware recommender system. *2012 IEEE 28th International Conference on Data Engineering* (p450-461), 2012. <http://dx.doi.org/10.1109/ICDE.2012.54>.
- [32] R. Pálóvics, P. Szalai, J. Pap, E. Frigó, L. Kocsis, and A. A. Benczúr, Location-aware online learning for top-k recommendation. *Pervasive and Mobile Computing*, 38, p490-504, 2017. <https://doi.org/10.1016/j.pmcj.2016.06.001>.
- [33] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, Recommender system application developments: A survey. *Decision Support Systems*, 74, 12-32, 2015. <http://dx.doi.org/10.1016/j.dss.2015.03.008>.
- [34] P. G. Campos, F. Díez, and I. Cantador, Time-aware recommender systems: A comprehensive survey and analysis of existing evaluation protocols. *User Modeling and User-Adapted Interaction*, 24(1-2), p67-119, 2014. <http://dx.doi.org/10.1007/s11257-012-9136-x>.
- [35] G. Adomavicius, and A. Tuzhilin, Context-aware recommender systems. In *F. Ricci, L. Rokach, B. Shapira (eds.). Recommender Systems Handbook* (p191-226). Springer, Boston, MA, 2015. http://dx.doi.org/10.1007/978-1-4899-7637-6_6.

- [36] L.O. Colombo-Mendoza, R. Valencia-Garcia, A. Rodriguez-González, G. Alor-Hernández, and J. J. Samper-Zapater, RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes. *Expert Systems with Applications*, 42(3), p1202-1222, 2015. <http://dx.doi.org/10.1016/j.eswa.2014.09.016>.
- [37] Á. Bogárdi-Mészöly, A. Rövid, H. Ishikawa, S. Yokoyama, and Z. Vámosy, Tag and topic recommendation systems. *Acta Polytechnica Hungarica*, 10(6), p171-191, 2013.
- [38] S. Zhang, S. Zhang, N. Y. Yen, and G. Zhu, The recommendation system of micro-blog topic based on user clustering. *Mobile Networks and Applications*, 22(2), p228-239, 2017. <http://dx.doi.org/10.1007/s11036-016-0790-9>.
- [39] M. Sipper, R. S. Olson, and J. H. Moore, *Evolutionary computation: The next major transition of artificial intelligence?* Retrieved on August 30, 2018, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5534026/>.
- [40] G. Gabrani, S. Sabharwal, and V.K. Singh, Artificial intelligence based recommender systems: A survey. In *Singh M., Gupta P., Tyagi V., Sharma A., Ören T., Grosky W. (eds.). Advances in Computing and Data Sciences. Communications in Computer and Information Science* (p50-59), vol 721. Springer, Singapore. 2016. http://dx.doi.org/10.1007/978-981-10-5427-3_6.
- [41] R. Katarya, and O. P. Verma, Efficient music recommender system using context graph and particle swarm. *Multimedia Tools and Applications*, 77(2), p2673-2687, 2017. <http://dx.doi.org/10.1007/s11042-017-4447-x>.
- [42] S. Nirenburg, and M. McShane, Natural language processing. *The Oxford Handbook of Cognitive Science*, p.337, 2016. <http://dx.doi.org/10.1093/oxfordhb/9780199842193.013.13>.
- [43] S. S. Kamath, and M. Kanakaraj, Natural language processing-based e-news recommender system using information extraction and domain clustering. *International Journal of Image Mining*, 1(1), p112-125, 2015. <http://dx.doi.org/10.1504/IJIM.2015.070031>.
- [44] Q. Marchena, *How Natural Language Processing works - infographic*. marketer. Retrieved on August 30, 2018, from <https://marketeer.co/en/blog/natural-language-processing-infographic/>.
- [45] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, p29-39, 2017. <http://dx.doi.org/10.1016/j.eswa.2016.09.040>.
- [46] D. Shankar, S. Narumanchi, H. A. Ananya, P. Kompalli, and K. Chaudhury, Deep learning based large scale visual recommendation and search for e-commerce. *Computer Vision and Pattern Recognition*,

- 2017.
- [47] Y. Jing, D. Liu, D. Kislyuk, A. Zhai, J. Xu, J. Donahue, and S. Tavel, Visual Search at Pinterest. In *KDD'15, Proceedings of the 21th ACM SIGKDD International Conference Knowledge Discovery and Data Mining*, p.1889-1898, 2015. <http://dx.doi.org/10.1145/2783258.2788621>.
- [48] E. Diaz-Aviles, *A glimpse into deep learning for recommender Systems*. Libre AI. Retrieved on August 31, 2018, from <https://medium.com/@libreai/a-glimpse-into-deep-learning-for-recommender-systems-d66ae0681775>.
- [49] H. T. H. Nguyen, M. Wistuba, J. Grabocka, L. R. Drumond, and L. Schmidt-Thieme, Personalized deep learning for tag recommendation. In *J. Kim, K. Shim, L. Cao, J. G. Lee, X. Lin, Y. S. Moon (eds.) Advances in Knowledge Discovery and Data Mining. PAKDD 2017. Lecture Notes in Computer Science* (p186-197), vol 10234. Springer, Cham, 2017. http://dx.doi.org/10.1007/978-3-319-57454-7_15.
- [50] M. Vozalis, and K.G. Margaritis, Collaborative filtering enhanced by demographic correlation. In *AIAI symposium on professional practice in AI of the 18th world computer congress*,2004.
- [51] L. Safoury, and A. Salah, Exploiting user demographic attributes for solving cold-start problem in recommender system. *Lecture Notes on Software Engineering*, 1(3), p303-307, 2013. <http://dx.doi.org/10.7763/LNSE.2013.V1.66>.
- [52] M. Y. H. Al-Shamri, User profiling approaches for demographic recommender systems. *Knowledge-Based Systems*, 100, p175-187, 2016. <https://doi.org/10.1016/j.knosys.2016.03.006>.
- [53] P. Cremonesi, A. Tripodi, and R. Turrin, Cross-domain recommender systems. *IEEE 11th International Confernece on Data Mining Workshops*, p496–503, 2011. <http://dx.doi.org/10.1109/ICDMW.2011.57>.
- [54] I. Fernández-Tobias, P. Tomeo, I. Cantador, T. Di Noia, and E. Di Sciascio, Accuracy and diversity in cross-domain recommendations for cold-start users with positive-only feedback. In *RecSys'16, Proceedings of the 10th ACM Conference Recommender Systems* (p.119-122). 2016. <http://dx.doi.org/10.1145/2959100.2959175>.
- [55] GitHub, Inc., *godgetfun/RECOMMENDER-SYSTEM-FOR-E-COMMERCE-PORTAL*. Retrieved on February 3, 2018, from <https://github.com/godgetfun/RECOMMENDER-SYSTEM-FOR-E-COMMERCE-PORTAL>.
- [56] K. Saleh, *Global Online Retail Spending - Statistics and Trends*. invesp. Retrieved on February 3, 2018, from <https://www.invespcro.com/blog/global-online-retail-spending-statistics-and-trends/>.

- [57] J. Davidson, B. Liebald, J. Liu, P. Nandy, T. Van Vleet, U. Gargi, S. Gupta, Y. He, M. Lambert, B. Livingston, D. Sampath, The YouTube video recommendation system. In *RecSys'10, Proceeding of the fourth ACM Conference Recommender Systems* (p293-296). 2010. <http://dx.doi.org/10.1145/1864708.1864770>.
- [58] W.T. Chu, and Y. L. Tsai, A hybrid recommendation system considering visual information for predicting favorite restaurants. *World Wide Web*, 20(6), p1313-1331, 2017. <http://dx.doi.org/10.1007/s11280-017-0437-1>.
- [59] K. Komiya, M. Sasaki, H. Shinnou, and Y. Kotani, Cross-lingual product recommendation system using collaborative filtering. *Journal of Natural Language Processing*, 24(4), p579-596, 2017. <http://dx.doi.org/10.5715/jnlp.24.579>.
- [60] Y. Guo, M. Wang, and X. Li, Application of an improved Apriori algorithm in a mobile e-commerce recommendation system. *Industrial Management & Data Systems*, 117(2), p287-303, 2017. <http://dx.doi.org/10.1108/IMDS-03-2016-0094>.
- [61] W. X. Zhao, S. Li, Y. He, L. Wang, J.-R. Wen, and X. Li, Exploring demographic information in social media for product recommendation. *Knowledge and Information Systems*, 49(1), p61-89, 2016. <http://dx.doi.org/10.1007/s10115-015-0897-5>.
- [62] C. Wu, and M. Yan, Session-aware information embedding for e-commerce product recommendation. In *CIKM'17, Proceedings of the 2017 ACM Conference on Information and Knowledge Management* (p.2379-2382). 2017. <http://dx.doi.org/10.1145/3132847.3133163>.