IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

Review Paper: New Advances in Recommender Systems

Shubham Jain[1], Manjula R[2]

School of Computer Science and Engineering, VIT University, Vellore, India[1]

Associate Professor, School of Computer Science and Engineering, VIT University, Vellore, India[2]

Abstract: Recommender Systems have been known for the past forty plus years but with the explosion of ecommerce sites and social media. In the new era it has become a major need to personalize the recommender system for every user, collaborative filtering does this job well but only but as users and products scale up we need newer techniques. Privacy is being demanded by internet users so how can one build recommender systems with privacy embedded in it is now one of the issue discussed in this paper. A recommender system also faces attacks on it like fake users, fake ratings the methods to deal with such attacks are listed. Lastly deep learning has shown several advances in the field of computer vision how to use the power of deep learning and Convolution and Recurrent Neural Networks in Recommender systems this paper explicitly covers that.

I. INTRODUCTION

Recommender systems have been used since a very long time and with lots of data and advent of ecommerce sites notably like amazon, and Netflix its use has boomeranged. The entire business model for amazon and Netflix rely on how well the recommender system has been trained to produce most suitable recommendations. Recommender systems have advanced from initial popularity based recommender, to ones which use deep learning system. This paper gives an overview of all the recommender system techniques which are available it also talks about the challenges in the recommender systems such as cold start and increasing the privacy in recommender systems. Of late deep learning has shown great improvement in the field of computer vision with the use of GPUs. So how can the deep learning is used to make better predictions is also discussed in this paper.

II. LITERATURE SURVEY

Initial Recommender system uses techniques like popularity model; like in [1] it has shown the limits of popularity based recommender system. Popularity recommender system basically points to the most popular item on the subset of items. It does not take care of user preferences. YouTube video recommendations and Amazon recommendation show reviews based on it. However if an item is being bought frequently can also drown out the product which is not sold quite often. This is a challenge with popularity based recommender system. The other types of recommender system techniques includes Content based Recommender system and Collaborative filtering techniques. [2] Talks about Content based filtering.

Content based filtering works on the principle the “if the user likes this he will that too”. Like if user likes book A, and book A is similar to book B my recommendation to the user would be book B. This is a technique which amazon, Netflix applies a lot of times while recommending products and films. However this approach comes with a lot of challenges –Firstly it is impossible or impractical to have a huge dataset of products mapping or tagging products (in other words building a similarity index of each and every product).Secondly the recommendation given by content based filtering is not diverse and makes very little sense to the user.

For example a user purchases an umbrella online and so by content based filtering his next recommendation will be a raincoat, so this makes a very little sense because it is highly unlikely that the same user will purchased both umbrella and raincoat. However Content based filtering is widely used for movies and song recommendation sites which recommend similar types of movies and songs like Netflix and Pandora. The Second Technique in Recommendation systems would be collaborative filtering. The idea of collaborative filtering is that when to check similarity between users and predict the most suitable product for users with similar interests. For finding users with similar interests we need may use Jaccard similarity [3] but the problem with jaccard similarity is it ignores any rating values given by the user. It just tells similarity about which users have rated which movie or product. Another way to know about similar customers is to use cosine based distances or centered cosine distances. Centered cosine distances handle similarities better. These similarity indexes can further be used for ratings prediction. Collaborative filtering can happen user to user and item to item however item – item collaborative shows better results than user-user [4].

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

III. CHALLENGES WITH RECOMMENDER SYSTEM

The main issues with recommender systems can be listed as privacy, solving cold start, differentiating between false ratings or attacks on recommender system, Deep Learning for Recommender Systems and an overview about some of the tools available for recommender systems. Each of these problems is explained here in later section. Implementing privacy in Recommender system is another concern, how can recommender system prove to user that their data is not being used or given to other parties. Attacks on recommender system [5] such as shilling attacks, probe attacks, false rating strategy are also some of the attacks which are important in recommender systems. Auto-Encoders can be used for learning a particular input from a given output [18].For collaborative filtering GPUs and CPUs (with multiprogramming) can be used for fastening the output [19].

IV. COLD START IN RECOMMENDER SYSTEMS

There are two types of cold start problem – New User cold start, and New Item cold start item. It is difficult to predict the likes of new user and similarly it is difficult to give reviews and ratings of new items. However user reviews are difficult to predict in cold start since user likes and dislikes are likely to change [6].To solve this problem ,[7] proposes to collect some user data before when user registers meaning that user likes and dislikes should be recorded, and it also recommends the A LinUCB algorithm[7].For running this algorithm a matrix must be created the with rows and columns being the users and their respective interests.[8] talks about the use of filter bots for rating users and items, thereby it lists out which user ,might purchase which item by mapping the item's characteristics and user's taste.[9] talks about two types of methods – namely node recommendations and batch recommendations for cold start. In active node recommendations the what user has browsed initially in the web is known to map out his likely purchases.[9] lists out a formula to know the user interest. Batch recommendation which is defined in [9] is more like Google's page rank algorithm which is detailed in [10].

V. ATTACKS ON RECOMMENDER SYSTEM

The attacks on recommender systems are done to make it give inaccurate results. In many ways these can be done by automated preprogrammed bots or users who would want the recommender system to give wrong predictions. It can be done by simple giving a wrong product a high rating or low rating based on circumstances. Some recommender systems use collaborative filtering methods so a fake profile is created which has a greater similarity index margin and the ratings of this profile (fake profile) is taken more often. Similarly fake profiles can be injected to effectively prop up one product over another. [11] Tells about two metrics namely Rating Deviation from Mean Agreement (RDMA) and Degree of similarity with Top Neighbors as a way to check between malicious and genuine users. [12] Talks about shilling attacks in great detail. Shilling attacks happens in collaborative filtering and it is quite impossible to remove them however they can be detected. [5] gives a formula on how a profile can be successfully be injected by the user.

VI. IMPLEMENTING PRIVACY IN RECOMMENDER SYSTEM

[13] Tells about the need for implementing privacy in recommender systems and gives three major points for measuring trust in recommender systems – Exposure, Bias and Sabotage. Privacy is quite essential for some set of Users but recommender systems need as much as possible for the customer to give better recommendations. Some of the simple privacy techniques which can be applied are obfuscating the data while giving to central server, keeping personal information locally in the device and application of cryptography in recommender system.[15] talks about the privacy-preserving collaborative filtering (PPCF) [15] which implements privacy and not affecting the results in collaborative filtering. PPCF but the database given to the recommender system is an obfuscated database, and the recommendations here are about Euclidian similarity [15].

VII. DEEP LEARNING IN RECOMMENDER SYSTEMS

Deep Learning has had tremendous success with Computer Vision, Natural Language Processing (NLP) so it needs to be seen whether deep learning can be applied to Recommender system with gives higher improvement. Some of the deep learning algorithms which can be integrated to Recommender Systems are : Convolution Neural Networks(CNN),Restricted Boltzmann Machines(RBM) ,Auto-Encoders ,Recurrent Neural Networks(RNNs),Multi-Layer Perceptron(MLP),Generative Adversarial Networks(GAN)[16].Brief introduction about the following terminologies: Multi-Layer Perceptron: It is a feed forward neural network with many hidden layers. Convolution Neural Networks (CNN): Widely used in Computer Vision tasks. Contains large number of neural network layers called as hidden layers, an input layer and output layer. Restricted Boltzman Machine (RBM): It contains just two layers hidden and visible layer. They are used in feature learning, classification. RBM can be used for

 DOI 10.17148/IJARCCE.2018.7375

collaborative filtering, with a slight increase in performance, it can handle tabular data such as user ratings (User rating varies between 1-5) [21]. Recurrent Neural Networks (RNN): are basically connections from one node to another. Used in Natural Language Processing. Deep Learning can be implemented in Recommender system in following two ways as described in [16]: Deep Learning can be stacked up with an Existing recommender system algorithm or it can be implemented along; both these approaches are done with the above given technologies. Deep Learning can be implemented in recommender system either with just one algorithm or as collection of various deep learning algorithms [16]. Deep Composite Models[16] can be used to club 2 or 3 ML algorithms to give recommendations[16].One of the open issue in implementing Deep Learning in Recommender system is Learning from other sources More User information can be collected from other sources like websites he has visited or people he has followed.[16]

VIII. CASE STUDY: NETFLIX AND YOUTUBE RECOMMENDER SYSTEM

This section illustrates about some of the technologies being used in Netflix and YouTube. Netflix [19] gives personalized recommendation of movies to its users. Personalized Video Ranker: PVR [19] by Netflix gives the Users lists of movies of the same genre which he enjoys the most (the higher the rankings of the user given to a particular type of movie). Since PVR produces a large number of movies for the user another algorithm Top N Video Ranker [19] narrows them by reducing the movies being shown by using some metrics like time, how it has been rated by other users. Some of the other technologies which Netflix uses are Latest Trends, Similar Videos etc. Lastly a page generation algorithm [19] uses the outputs of all the above given recommendations to give the users a single page. Now coming to YouTube recommendations. YouTube Recommendations use deep learning or deep neural networks [20]. YouTube Recommendations are quite different from other recommendations available due to the scale (about 3 billion videos being uploaded in YouTube) and Freshness every day lots of videos being uploaded in YouTube and also the quality of video being shown is not known to the machine referred as Noise in the Paper [20]. The YouTube system consists of two parts candidate generation and Ranking [20]. Candidate generation gives a list of all the videos which are generated so this can be in millions and Ranking part determines the order, this is done by having appropriate feature engineering.

IX. CONCLUSION

In this paper we have discussed several methodologies and issues which come up in implementing a recommender system. Implementing deep learning with recommender system is a very new field and has scope for further improvement to give predictions for a system which has large number of diverse items to be recommended and perhaps it can also be able to solve some of the issues like attacks on recommender system.

REFERENCES

[1] Marco Bressan, Stefano Leucci, Alessandro Panconesi, Prabhakar Raghavan, Erisa Terolli in The Limits of Popularity-Based commendations, and the Role of Social Ties in arXiv:1607.04263v1 [cs.SI] 14 Jul 2016
[2] Jieun Son, SeougBum Kim in Content-based filtering for recommendation systems using multiattribute networks in http://dx.doi.org/10.1016/j.eswa.2017.08.008
[3] Suphakit Niwattanakul, Jatsada Singthongchai, Ekkachai Naenudorn, Supachanun Wanapu in Using of Jaccard Coefficient for Keywords Similarity in Proceedings of the International MultiConference of Engineers and Computer Scientists 2013 Vol I, IMECS 2013, March 13 - 15, 2013, Hong Kong
[4] Abinash Pujahari, Vineet Padmanabhan in Group Recommender Systems: Combining user-user and item-item Collaborative filtering techniques in 2015 14th International Conference on Information Technology 978-1-5090-0487-4/15, 2015 IEEE DOI 10.1109/ICIT.2015.36
[5] Michael P. O'Mahony, Neil J. Hurley, Guenol ´ e C.M. Silvestre in Recommender Systems: Attack Types and Strategies in 2005, American Association for Artificial Intelligence (www.aaai.org), AAAI-05 / 334
[6] L.H. Son, Dealing with the new user cold-start problem in recommender systems: A comparative review, Information Systems (2014), http://dx.doi.org/10.1016/j.is.2014.10.001i
[7] Hai Thanh Nguyen, J´er´emie Mary ,Philippe Preux in Cold-start Problems in Recommendation Systems via Contextual-bandit Algorithms in arXiv:1405.7544v1 [cs.IR] 29 May 2014
[8] Seung-Taek Park1, David Pennock2, Omid Madani1, Nathan Good3, Dennis DeCoste1 in Naıve Filterbots for Robust Cold-Start Recommendations in 2006 ACM 1-59593-339-5/06/0008
[9] Ossama H. Embarak in A Method For Solving The Cold Start Problem In Recommendation Systems in 2011 International Conference on Innovations in Information Technology 978-1-4577-0314-0/11
[10] The Page Rank Citation Ranking: Bringing Order to the Web in http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf
[11] Wei Zhou, Junhao Wen, Yun Sing Koh, Shafiq Alam and Gillian Dobbie in Attack Detection in Recommender Systems Based on Target Item Analysis in 2014 International Joint Conference on Neural Networks (IJCNN) July 6-11, 2014 978-1-4799-1484-5/14
[12] Ihsan Gunes · Cihan Kaleli · Alper Bilge ·Huseyin Polat in Shilling attacks against recommender systems: a comprehensive survey in Artif Intell Rev DOI 10.1007/s10462-012-9364-9
[13] Shyong K Tony Lam, Dan Frankowski, and John Riedl in Do You Trust Your Recommendations? An Exploration Of Security and Privacy Issues in Recommender Systems
[14] D. Li et al, An algorithm for efficient privacy-preserving item-based collaborative filtering Future Generation Computer Systems(2014),http://dx.doi.org/10.2016/j.future.2014.11.003
[15] F. Casino et al., A k-anonymous approach to privacy preserving collaborative filtering, J. Comput. Syst. Sci. (2015), http://dx.doi.org/10.1016/j.jcss.2014.12.013

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

[16] Shuai Zhang, Lina Yao, and Aixin Sun. 2017. Deep Learning based Recommender System: A Survey and New Perspectives. ACM J. Comput. Cult. Herit. 1,1, Article35 (July 2017),35pages. DOI: arXiv:1707.07435v5
[17] Jian Wei, Jianhua He, Kai Chen, Yi Zhou, Zuoyin Tang, Collaborative Filtering and Deep Learning Based Recommendation System For Cold Start Items, Expert Systems With Applications (2016), doi: 10.1016/j.eswa.2016.09.040
[18] Efthalia Karydi and Konstantinos Margaritis. 2016. Parallel and distributed collaborative filtering: A survey. ACM Comput. Surv. 49, 2, Article 37 (August 2016), 41 pages. DOI: http://dx.doi.org/10.1145/2951952
[19] Carlos A. Gomez-Uribe and Neil Hunt. 2015. The Netflix recommender system: Algorithms, business value, and innovation. ACM Trans. Manage. Inf. Syst. 6, 4, Article 13 (December 2015), 19 pages. DOI: http://dx.doi.org/10.1145/2843948
[20] Paul Convington ,Jay Adams and Emre Sargin in Deep Neural Networks for YouTube Recommendations in RecSys '16 September 15-19, 2016, Boston , MA, USA, ACM ISBN 978-1-4503-4035-9/16/09. DOI: http://dx.doi.org/10.1145/2959100.2959190
[21] Ruslan Salakhutdinov, Andriy Mnih, and Geo.rey Hinton. 2007. Restricted Boltzmann machines for collaborative filtering. In Proceedings of the 24th international conference on Machine learning. ACM, 791–798.