International Journal of Research in Engineering, Science and Management
Volume-3, Issue-4, April-2020
www.ijresm.com | ISSN (Online): 2581-5792

A Comprehensive Review of Approaches and Challenges of a Recommendation System

Lakshmi Narke[1], Azra Nasreen[2]
[1]*Student, Department of Computer Science and Engineering, Rashtreeya Vidyalaya College of Engineering, Bangalore, India*
[2]*Assistant Professor, Department of Computer Science and Engineering, Rashtreeya Vidyalaya College of Engineering, Bangalore, India*

***Abstract*: The digital era with its never ending explosion of content has created a paradigm of too many choices. A user who is new to the platform or is in search of an entity is lost in this ocean. Is there a base where she/he can start off with? Yes we can resort to static pre curated lists of popular/ most preferred items. This system's inherent flaw is that it considers all users to probably have the same behaviour and interests but on the contrary human interest spans differently across different domains and demographics. Given a fixed landing screen real estate and the wavering human mind, we need to curate user specific lists to take maximum advantage of this span of attention. Users can also be introduced to new content based on their persona. A Recommendation System assists users to find enthralling content in a large digital mine. In a box it can be broken down into a simple system that matches users with items they may be interested in. This kicks off the horses to find the best and efficient way to do it. So how can we model a system that can predict user preferences with the limited knowledge of the user and content. This paper explores various techniques used for implementation of recommendation systems such collaborative filtering, content based filtering, hybrid models of recommendation and demographic based filtering technique. The paper dives into the pros and cons of each method. The problems faced at scale. What are the attributes of the entities we look at and the ease of adding new items. The study compares other eminent research publications to comprehensively evaluate these problems. We tie all this together by looking at a hybrid approach and evaluate how multiple techniques work in tandem to create a balanced and practical model in the field of recommendation systems.**

***Keywords*: Collaborative filtering, Content-based filtering, Demographic based Filtering, Hybrid Recommendations, Recommender system.**

1. Introduction

Now-a-days, enormous volume of data is generated from numerous sources through Internet, IoT platforms, social media, smartphones etc. The generated data is a set of objects and their attributes. Objects might include record, item, observation or instance while attributes include characteristics or features [1]. Any meaningful insights cannot be drawn with only data; a process is required to derive relevant information from this data to give meaningful insights. Most of the time this is achieved using data mining.

One way of carrying out this data mining is through recommendation systems. These are the techniques providing suggestions for items to the user for instance recommending electronic related items if a user is trying to purchase a mobile on an ecommerce website. These suggestions helps the user to make an effective decision regarding the enormous choices the Web world has to offer, such as what products to purchase, which book to read, which movie to watch.

RSs are mainly directed towards individuals who usually lack competence to assess the massive number of alternatives the Internet has to offer leaving the users perplexed [1]. Recommendations can be personalised or non-personalised. In personalised recommendations different user segments receive different suggestions based on the proclivity and constraints of the user segment whereas non-personalised suggestions do not take features of any specific user segment into account.

For Example, consider a movie recommender system, the popular website Netflix.com employs a recommender system to personalise the home page for each of it's customers. Such recommendations fall under personalised recommendations. While recommending top 10 movies come under non-personalised recommendations which are easier to predict but do not attract much research. Personalised recommender systems are mostly addressed by the researchers to analyse profuse volume of data to provide the users with personalised content and services.

2. Types of recommender systems

Recommendation system is a concept that is established in the domain where *items (books, products, movies)* to be suggested to *users (customers, web page visitors, online article readers). A Recommender system gains information about* the user. This information is gained using different methods of data mining on the raw data captured by different sources like Internet/Social Network, Media, databases [2].

Recommendation process can be divided into 3 phases with a feedback. Refer to figure 1 for diagrammatic representation of the three phases.

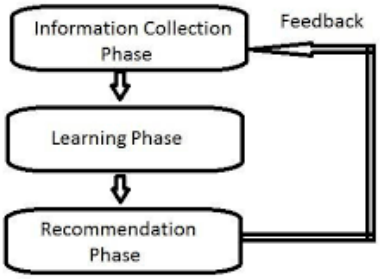

Fig. 1. General conceptual flow of a recommender system

1. *Information Collection Phase:* This phase captures data about the users from the different sources mentioned above. Data is mined to discover appropriate and relevant information.
2. *Learning Phase:* This phase employs the filtering techniques and algorithms on the information gathered implicitly or explicitly to derive meaningful insights from the information which is required in the next phase.
3. *Recommendation Phase:* This is the final phase where on the basis of data pre-processed in the first phase and trained in the subsequent phase, recommendations are made to the users in the shape of personalised suggestions, content or services.

A feedback is given after the recommendation phase based on user's actions like ratings, comments, visited links.

A Recommender software gives a prediction based on one of the following techniques.

A. Collaborative filtering

Collaborative Filtering is a method used by Recommendation systems where predictions are made about preference of a user by exploiting similarities amid various users or products for predicting missing scores and thus make satisfactory recommendations. It implies that collaborative filtering models can suggest an item to user X based on the preferences of similar user's Y and Z.

User Y: Prefers item2, item6, item9

User Z: Prefers item1, item6, item9

If User X prefers item6 then the Collaborative based recommendation system predicts that User X might be interested in item9 as well.

This technique are of two types:

- Memory-based collaborative filtering technique
- Model-based collaborative filtering technique

This technique is advantageous as there is no requirement of pre-requisite knowledge in that domain to learn the embeddings. It acts as a great system to recommend new items to the user. But this technique encounters the issue of cold start which is the scenario that arises due to addition of a new item.

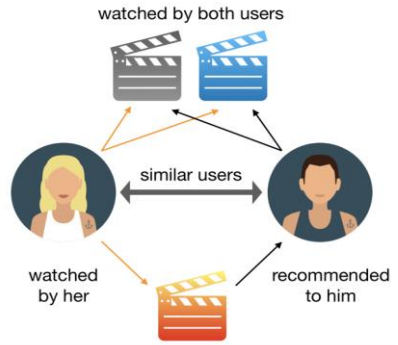

Fig. 2. Collaborative based filtering technique

Figure 2 illustrates collaborative based approach with an example of Movie recommendation.

B. Content based filtering

Content-based filtering technique is one of the common methods in building recommendation systems. This technique is even referred as cognitive filtering, that suggests items based on the comparisons of item description and a profile of user's interest. The content of every item is mapped to a collection of descriptors and represented for the items or related terms, mainly the terms that occur in a document. This technique analyses a collection of descriptors fed to it as inputs for a specific item scored by the user earlier. It then builds a model which recommends a user. These recommendations are generated by capturing users' predilection [3].

For instance, if a user views an internet page with the words "Outfits", "Accessories" and related items the model employing content filtering technique will suggest pages related to fashion and apparel. The advantage of this technique is that it scales very well to a large number of users but the main drawback is a lot of domain knowledge is required. The model turns out to be as good as the amount of knowledge we engineer into it.

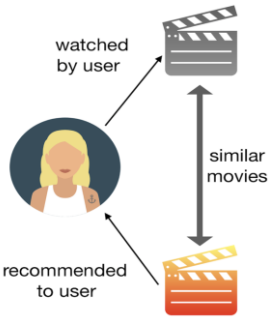

Fig. 3. Content based filtering technique

Figure 3 illustrates content based approach with an example of Movie recommendation.

C. Demographic based filtering

This recommendation methodology uses user profile data such as gender, age, education, demographic region, interests and their views on the scoring of items and identifies common

users with similar scores of items and preferences that segregates users by age group and living area [4] The conceptual ground of this system is an assumption that varied suggestions needs to be generated to account for varied demographic niches.

D. Hybrid recommendations

Hybrid Recommender systems are amalgamation of Collaborative based recommendation, content based recommendation and other techniques. This technique is more preferred in todays' consumer world due to its improved quality and performance by blending in different techniques. For instance, Collaborative filtering doesn't suffer from ramp-up problem, hence can be combined with other methodologies to nullify the same issue.

Hybrid approaches have multi-methods [5]:

- *Weighted:* In this method, every recommended component is numerically combined, the different scores given by each system.
- *Switching:* In this approach, the system has several options of varied suggested products to the user and amidst them it selects an appropriate one based on user predilection.
- *Mixed:* This method employs a system that suggests several different items all together to the user.
- *Feature combination:* This approach involves amalgamation of several knowledge sources to develop features of a recommendation system.
- *Feature augmentation:* This technique involves computation of collection of traits of recommendation systems.
- *Cascade:* List of recommendations has a weighted priority. items with higher rating appear first, followed by the items with a lower rating score.
- *Meta-level:* It is an input technique used to generate and produce a model to the next step of the recommender system algorithm.

The combination of these various techniques improves efficiency and dampens the challenges and problems that occur by solely using either content based technique or collaborative filtering technique exclusively.

3. Related work

An improvisation on accuracy by including semantic data [6] focuses on the use of Hybrid Collaborative Filtering with Semantic Information model which uses both historic user data and semantic item data and the analysis provided an improved accuracy.

Enhancement of Movie suggestion by incorporating Support Vector Machine [7] proposes the use of a Hybrid model with Support Vector Machines as the base for Collaborative filtering and the results and optimised using genetic algorithm. This technique addresses the issue of scalability of the solution.

Use of Case based reasoning [8] is a hybrid approach that address cold start and over specialization problems. It combines both models to get higher accuracy in larger datasets.

A simple collaborative approach [9] with the standard cosine similarity rule for efficient searching and filtering mechanisms are presented.

Scalability of collaborative filtering using methods that account for interactions within a neighbourhood and that do not require imputations is presented in [10] which showcases the use of complementary models with collaborative filtering. The experimental results suggest that significantly high accuracy is achieved in large real world datasets.

A comparison of KNN and SVM based collaborative filtering framework is presented in [11]. The outcomes of the implementation showcased in this paper confronts the k-Nearest Neighbour (kNN) algorithm with Support Vector Machine(SVM) with different properties in the collaborative filtering framework. The results showcase that a model based approach like SVM is preferred for extremely sparse datasets whereas kNN is more feasible for regular datasets.

A content-based recommendation framework is proposed called FIRSt, which combines user generated content (UGC) with semantic analysis of content is proposed [12]. First's contribution here is a combined approach that allows a content-based recommender to deduce

4. Conclusion

Recommendation systems are increasingly becoming prevalent in today's consumer world. There is a strong need to identify and solve the problems that come at this scale in building these systems. This paper details the nuances of major implementations of Recommender systems like Content based filtering technique, Collaborative filtering technique, a Hybrid model and Demographic based filtering.

The challenges of each like cold start and scale on larger datasets is discussed and it has been found that the hybrid models are effective in solving these issues. The future scope of the work includes finding more complementary models that can assist these systems in achieving larger scales and better accuracy there is also dearth in comparison metrics of these models. Only using a simple accuracy metric does not show how well the model has identified some of the local semantic relationships. This can colour our vision to choose an appropriate system. Finer nuances like these need to be addressed.

References

[1] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor, "Recommender Systems Handbook," Springer 2010.
[2] P. Priyanga, A. R. Nadira Banu Kamal, "Methods of mining the data from big data and social networks based on recommender System". International Journal of Advanced Networking & Applications, vol. 8(5) 2017, pp. 55-60.
[3] Debashis Das, Laxman Sahoo, Sujoy Datta, "A survey on recommendation system," International Journal of Computer Applications, vol. 160, no. 7, pp. 6-10, February 2017.

[4] Sarika Jain, Anjali Grover, Praveen Singh Thakur, Sourabh Kumar Choudhary, "Trends, problems and solutions of recommender System," ICCCA 2015, 2015. pp.955-958.
[5] Yagnesh G. Patel, Vishal P. Patel, "A survey on various techniques of recommendation system in web mining", International Journal of Engineering Development and Research, vol. 3(4), 2015, pp. 696-700, 2015.
[6] Bushra Alhijawi, Nadim Obeid, Arafat Awajan, Sara Tedmori, "Improving collaborative filtering recommender systems using semantic information ", ICICS, 2018.
[7] Shreya Agrawal, Pooja Jain, "An improved approach for movie recommendation system," I-SMAC, 2017.
[8] Anna Gatzioura, Miquel Sànchez-Marre, "A case based Recommendation approach for market basket data," IEEE Journal on Intelligent Systems, vol. 30, no. 1, 2015
[9] M. Viswa Murali, T. G. Vishnu, Nancy Victor, "A collaborative filtering based recommender system for suggesting new trends in any domain of research", ICACCS, 2019.
[10] Bell, R., Koren, Y., Volinsky, C.: Modeling relationships at multiple scales to improve the accuracy of large recommender systems. In: KDD 07: Proc. of the 13th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, pp. 95104. ACM, New York, NY, USA, 2007.
[11] Grcar, M., Fortuna, B., Mladenic, D., Grobelnik, M, "k-NN versus SVM in the collaborative filtering framework," Data Science and Classification pp. 251-260, 2006.
[12] Pasquale Lops, Marco de Gemmis, Giovanni Semeraro, Paolo Gissi, Cataldo Musto, Fedelucio Narducci, "Content-based filtering with tags: the FIRSt system", ISDA, 2009.
[13] Deshpande M., Karypis, G, "Item-based Top-N recommendation algorithms," ACM Transaction on Information Systems, vol. 22, no. 1, pp. 143-177, 2004.
[14] Anna Gatzioura and Miquel Snchez-Marr, "A case-based recommendation approach for market basket data," IEEE intelligent systems, 2015.
[15] Y. Yoshinari, W. Pedrycz, and K. Hirota. "Construction of fuzzy models through clustering techniques", Fuzzy sets and systems, vol. 54, pp. 157-165, 1993.
[16] Soanpet Sree Lakshmi and T. Adi Lakshmi, "Recommendation systems: issues and challenges", International Journal of Computer Science and Information Technologies, vol. 5(4), 2014, pp. 5771-5772.
[17] Using collaborative ltering to weave an information Tapestry D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, Communications of the ACM, vol. 35, no. 12, pp. 6170, 1992.
[18] G. Linden, B. Smith and J. York, "Amazon.com recommendations item-to-item collaborative filtering," in IEEE Internet Computing, Volume 7, Issue 1, January 2003, pp.76-80, 2003.
[19] Mladenic, D.: Text-learning and related intelligent agents: A Survey. IEEE Intelligent Systems14(4), 44–54 (1999).
[20] J. Ben Schafer, ShiladSen, Dan Frankowski, Jon Herlocker, (2007) "Collaborative filtering recommender systems," in the Adaptive Web, pp. 291–324. Springer Berlin / Heidelberg.
[21] F. Kong, X. Sun, S. ye, (2005), "A comparison of several algorithms for collaborative filtering in startup stage", in: Proceedings of the IEEE Networking, Sensing and Control, pp. 25–28.
[22] S. V. N. Vishwanathan, M. Narasimha Murty (2002), "SSVM: A simple SVM algorithm", Neural Networks, IJCNN '02. Proceedings of the International Joint Conference on, vol. 3, 2002.
[23] Joseph Konstan, George Karypis, BadrulSarwar, and John Riedl, "Item-based collaborative filtering recommendation algorithms", ACM, 2001.
[24] Hirdesh Shivhare, Anshul Gupta and Shalki Sharma, "Recommender system using fuzzy c-means clustering and genetic algorithm based weighted similarity measure", IC4-2015, 2015.
[25] Tanya Maan, Shikha Gupta, Dr. Atul Mishra," A survey on recommendation system", RIMEST, pp. 543-549, 2018.
[26] Hadeer Mahmoud, Abdelfatah Hegazy, Mohamed H. Khafagy, "An approach for big data security based on Hadoop distributed file system", ITCE, pp. 109-114, 2018.
[27] Mustansar Ali Ghazanfar and Adam Prugel-Bennett, "A scalable, accurate hybrid recommender system", KDD, 2010.
[28] L. Soanpet and A. Lakshmi, "Recommendation systems: issues and challenges", IJCSIT, vol. 5, pp. 5771-5772, 2014.
[29] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey", Elsevier Decision Support Systems, vol. 74, pp. 12-32, 2015.