

Analysis of Recommendation System Design Techniques

Sanjeev Dhawan¹, Kulvinder Singh² and Etika Gupta³

^{1,2}Faculty of Computer Science and Engineering, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra-136119, Haryana, India

³M.tech Computer Engineering, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra-136119, Haryana, India

ABSTRACT

The recommendation systems are generally designed to recommend data based on the current information. The recommendation systems are generally designed with the technique of machine learning. The recommendation systems have various phases like the pre-processing, feature extraction and classification. In the technique of pre-processing missing and redundant values will be removed from the input data. In the second phase of feature extraction relationship between attribute set and target set will be established for the classification. In the last phase, the technique of classification will be applied in which system will be trained on the certain attributes. In this paper, various techniques to design efficient recommendation system is reviewed in terms of certain parameters. This proposed paper has been described in two sections. Section 1 describes the introduction part; Section 2 describes the related work done.

Keywords: Classification, Feature Extraction, Machine learning, Recommendation system

Date of Submission: 18-05-2020

Date of Acceptance: 03-06-2020

I. INTRODUCTION

A system that helps individuals to make decisions of their personal interest based on the knowledge available is known as recommendation system. The preferences to the objects or services provided to the users can be predicted by these systems which are also known as the subclass of information filtering [1]. For personalizing the recommendation and handling the information overload, search is still going on since it is applied in several applications even though it has been explored on large scale in the past. Therefore, the other techniques such as model based or memory based approaches face various challenges due to such demands. For generating a better system, improvements are still needed in the recommender system. Any idea related to the object in which user is interested can be given by this sharp recommendation system.

There are various approaches designed for recommendation which are broadly categorized into collaborative filtering, content based and hybrid approaches. For using rating or content information various algorithms and approaches are applied for recommendation [2]. However, there are certain limitations faced by the collaborative filtering as well as content based approaches. A hybrid approach was then constructed by researchers with the aim of overcoming these

limitations [3]. The content information and ratings are combined to design this hybrid approach.

1.1 Issues and Challenges in Recommendation Systems

The most common challenges faced when deploying recommendation systems are enlisted below:

- i. Sparse RSs: Majority of items are not rated by the maximum number of users and due to this research there is sparse ratings matrix. The chances of identifying a set of users with alike ratings are reduced due to the data sparsity issue [4][5]. This is a major drawback which can be eliminated by providing additional information related to domain.
- ii. Cold-start problem: A problem related to new item and new user for recommendation systems is the cold-start problem. Initially, when introducing new item in the CF system with no ratings, a new item cannot be recommended. Since finding similar users is not possible, the new-user problem is difficult to be manages.
- iii. Scalability Problem: Scalability of algorithms that include large real-world datasets is the most important concern within recommendation systems. Managing the huge and dynamic datasets generated by item-users interactions is a major challenge. Therefore, to handle such problem, several

advanced large-scale assessment approaches have been designed.

iv. Privacy Issue: Collecting as much data as possible and exploiting it to the highest level is the task of recommendation systems such that the quality of personalized recommendations can be generated. However, the users might get a negative impression about the privacy since too much knowledge about them is known to the system. Therefore, to use the data in highly sensible and careful manner, various techniques need to be introduced [6].

v. Robustness of RSs: Robustness to attacks is another major challenge of recommendation systems. It is also known as its performance measure. Based on few attack models, some fake user profiles could be generated by an attacker with the aim of achieving personal profits. These affect the performance outputs of systems.

vi. Recommenders in Mobile Devices: With the increase in development of wireless networks and mobile devices, the popularity of location-based services is increasing. Thus, an important role is played here by the geographical information. When moving across various locations, the users might attain several recommendations. To use the available limited resources in effective and efficient manner, possible computational solutions are required by mobile user interfaces.

1.2 Machine Learning

Machine learning learns machines the way of handling data in more efficient manner. Occasionally, it becomes impossible for someone for interpreting the pattern or extracting knowledge from the data. In such a situation, machine learning plays an important role. The popularity of machine learning is increasing with the availability of massive volume of datasets [7]. A number of applications make use of machine learning for extracting important knowledge. These applications include medicine, military; agriculture etc. learning from data is the main objective of machine learning. A lot of research works have been carried out for making machines independent in learning. In order to resolve this issue, various mathematicians and programmers implement different techniques.

A. Supervised Learning: The algorithms requiring outside support are known as supervised machine learning algorithms. These algorithms divide input dataset into training and test dataset. The train dataset comprises output variable [8]. The prediction or classification of this variable is vital. The learning of some sort of patterns is carried out by all algorithms from the training dataset. These patterns are applied to the test dataset which need

to be predicted or classified. Figure 2 explains the working of supervised machine learning algorithms. In this work, a discussion have made on three most popular supervised machine learning algorithms. These algorithms are:

- i. Decision Tree: DT or Decision tree performs the grouping of features by arranging them on the basis of their values. This algorithm is particularly employed to do classification. There are the nodes and branches in every tree. Every node displays features in a group selected for classification. Every branch embodies a value possible for a node [9].
- ii. Naïve Bayes: NB (Naïve Bayes) generally focuses on the text classification. It is mostly employed to do clustering and classification. The fundamental design of Naïve Bayes is based on the conditional probability. It generates trees on the basis of their probability of occurrence. Bayesian Network is the other name given to these trees.
- iii. Support Vector Machine: SVM (Support Vector Machine) is a very popular existing machine learning algorithm. It is generally employed to do classification. The working of this algorithm is based on the concept of margin computation. On the whole, it represents margins among the classes[10].

II. RELATED WORK

A new technique was designed for giving rating to merchandise by M. Vamsee Krishna Kiran, et al. (2017) [11]. This approach efficiently rated the product on the basis of its technological specification. For this purpose, this approach analyzed the multiple consumer reviews. The extraction of these reviews was carried out from various e-commerce sites. The key aim here was to assist customers in the buying of preferred merchandise. The future work would be focused on extracting particular reviews from different E-Commerce sites and blogs that especially considered electronic tools. Nishigandha Karbhari, et al. (2017) proposed a novel system for generating recommendations on the basis of score obtained by students. These models discovered optimum solutions that could remain hidden otherwise [12]. The analysis carried out on the recommendation system was implemented in college campus. This analysis provided recommendations for helping students in getting jobs in companies in the minimum possible duration. This system could be used to get results by considering the needs of companies and employees. A general idea of the recommendation systems was presented by Kunal Shah, et al. (2017). This work also described the current

generation of recommendation algorithms. Recommendation systems (RSs) are a division of information filtering system [13]. These systems had software tools and methods that provided recommendations to the customer as per their requirement. A lot of famous Ecommerce sites extensively utilized these systems for recommending news, music, case studies, books, and goods. These systems made use of online available private, implied and local information. This works tried to explain several shortcomings of recommendation systems along with their benefits. In Thailand, an ontology based recommendation model was constructed by Natedao Thotharat, *et al.* (2017) [14]. This model used similar features of Thai local goods for product recommendation. The mapping of goods was done by this approach for identifying their features. The tested outcomes depicted that the developed approach could gain customer's interests on the basis of commonly produced recommended product. Anna Gatzzioura, *et al.* (2017) designed a hybrid Case-Based Reasoning system to generate and recommend music playlists [15]. It was analyzed that the designed system could give recommendations of identical and better accuracy than some commonly employed recommendation systems. The future work would be focused on extending this system for providing more help to new recommendations. An analysis of Hadoop-based farming goods e-commerce recommendation system was presented by Jiahuan Li, *et al.* (2018) [16]. The main aim of this work was to eliminate the lack of personalized services on the available farming goods e-commerce platform. The suggested system along with resolving the issue related to the storage of large volume of data could also speedily study the behavior data of client using distributed computing. In that way, this system discovered the interests of users and provided personalized suggestion service to them. This system achieved accurate advertising of targeted customers. Simultaneously, this work provided a suggestion for the evolution of e-commerce based on farming goods. Rahul Kumar Chaurasiya, *et al.* (2018) presented a recommendation system based on reviews. The main aim of this system was to make recommendations [17]. The presented system carried out the processing of review content as the response of customer for making predictions. For this purpose, this approach used a standard collaborative filtering approach based on item. The preprocessing of unprocessed review dataset was carried out after the attainment of scores of sentiment level. The review dataset signified the comprehensive response of customer regarding the product. The recommended system outperformed

the conventional system based on rating in terms of recall and RMSE value. Another model for recommendation system was designed by Yograj Meena, *et al.* (2018). This model captured picture and attributes of pictures and provided recommendations to the customer [18]. A picture included various attributes that could distinguish it from other pictures. The concept of classifying picture was based on the mean and standard deviation of image parameters and computation of distance matrix with other pictures for their classification. It was analyzed that this model worked on an undersized dataset and achieved acceptable accuracy rate. Teddy Oswari, *et al.* (2018) presented music product based recommendation system. The main aim of this system was to give comfort to the customers in the selection of music products on the basis of their interests [19]. This work described the architecture of a music recommendation system for an online site. For this purpose, this work used the idea of RAD. This approach involved a recommendation system model. This work did not made discussion on the background procedure of the system. The architecture of the system was configured and modified to satisfy the requirements of clients.

Tessy Badriyah, *et al.* (2018) designed a web-based recommendation model. This model used content-based filtering method for feature selection [20]. This model provided information about feature on the basis of client's behavior. For this purpose, this model searched earlier searched marketing content. This model would help potential customers in the selection of feature product they wish to purchase. This suggested model could do this in less time. As a continuity of this work, Situated Product Recommendations were analyzed by Brandon Huynh, *et al.* (2018) in Augmented Reality. This work considered online purchase for exploring the effect of this type of recommendation system [21]. This system allowed users to analysis adapted product recommendations in the physical locations. The generated outcomes highlighted possible disparities in consumer insight of the recommended products in augmented reality. Based on the user's implicit feedback (BUIF), a personalized recommendation algorithm was designed by Bo Wang, *et al.* (2018) [22]. Not only the purchasing behavior of user but also the comparison behavior and item sequences of user were used in this model. A secondary reordering model was designed for generating user based recommendation results. The outcomes showed improvement in recommendation accuracy when applying proposed model. Ni Made Satvika Iswari, *et al.* (2019) made use of Slope One algorithm for constructing a product recommendation system on

e-commerce platform as per the requirements of customers [23]. In the presented algorithm, the input rating was provided on the basis of the domain ontology of the product. There could be more variation in the suggestions (or recommendations). These variations could increase with the increased interest of customers. The future work would be focused on implementing this technique on ecommerce platforms. Hence, efficiency of this approach could be assessed by employing more applicable data. A new product recommendation system was designed by Udom Janjarassuk, *et al.* (2019). This system was based on genetic algorithm [24]. The main aim of this system was to locate the optimum suggestion for a

mixture of merchandise to the users. The model was evaluated on the basis of user's choices and merchandise needs along with attribute ratings from the product specialists. The testing of this system was done using a review from recommendation of energy unit election for recording studio. One more attempt had been made to build a recommendation model by Phan Duy Hung, *et al.* (2019). The main aim of this work was to present a new study in data visualization, level of measure with the help of heuristics [25]. It was noticed that a user-collaborate filtering approach could make better use of data rather than item-collaborate approach.

Author	Year	Description	Outcome
M. Vamsee Krishna Kiran	2017	This approach efficiently rated the product on the basis of its technological specification. For this purpose, this approach analyzed the multiple consumer reviews. The extraction of these reviews was carried out from various e-commerce sites.	The future work would be focused on extracting particular reviews from different E-Commerce sites and blogs that especially considered electronic tools
Nishigandha Karbhari, <i>etal</i>	2017	These models discovered optimum solutions that could remain hidden otherwise. The analysis carried out on the recommendation system was implemented in college campus. This analysis provided recommendations for helping students in getting jobs in companies in the minimum possible duration	This analysis provided recommendations for helping students in getting jobs in companies in the minimum possible duration. This system could be used to get results by considering the needs of companies and employees.
Kunal Shah	2017	This work also described the current generation of recommendation algorithms. Recommendation systems (RSs) are a division of information filtering system	These systems made use of online available private, implied and local information. This works tried to explain several shortcomings of recommendation systems along with their benefits
Natedao Thocharat	2017	This model used similar features of Thai local goods for product recommendation. The mapping of goods was done by this approach for identifying their features	The tested outcomes depicted that the developed approach could gain customer's interests on the basis of commonly produced recommended product
Anna Gatzoura	2017	It was analyzed that the designed system could give recommendations of identical and better accuracy than some commonly employed recommendation systems	The future work would be focused on extending this system for providing more help to new recommendations

Jiahuan Li	2018	The main aim of this work was to eliminate the lack of personalized services on the available farming goods e-commerce platform.	This system achieved accurate advertising of targeted customers. Simultaneously, this work provided a suggestion for the evolution of e-commerce based on farming goods.
Rahul Kumar Chaurasiya	2018	The presented system carried out the processing of review content as the response of customer for making predictions. For this purpose, this approach used a standard collaborative filtering approach based on item	The review dataset signified the comprehensive response of customer regarding the product. The recommended system outperformed the conventional system based on rating in terms of recall and RMSE value
Yograj Meena	2018	A picture included various attributes that could distinguish it from other pictures. The concept of classifying picture was based on the mean and standard deviation of image parameters and computation of distance matrix with other pictures for their classification	It was analyzed that this model worked on an undersized dataset and achieved acceptable accuracy rate
Teddy Oswari	2018	The main aim of this system was to give comfort to the customers in the selection of music products on the basis of their interests	This work did not made discussion on the background procedure of the system. The architecture of the system was configured and modified to satisfy the requirements of clients
Tessy Badriyah	2018	This model used content-based filtering method for feature selection. This model provided information about feature on the basis of client's behavior	This model would help potential customers in the selection of feature product they wish to purchase. This suggested model could do this in less time
Brandon Huynh	2018	This work considered online purchase for exploring the effect of this type of recommendation system. This system allowed users to analysis adapted product recommendations in the physical locations	The generated outcomes highlighted possible disparities in consumer insight of the recommended products in augmented reality
Ni Made Satvika Iswari	2019	In the presented algorithm, the input rating was provided on the basis of the domain ontology of the product. There could be more variation in the suggestions (or recommendations). These variations could increase with the increased interest of customers	There could be more variation in the suggestions (or recommendations). These variations could increase with the increased interest of customers
Udom Janjarassuk	2019	The main aim of this system was to locate the optimum suggestion for a mixture of merchandise to the users. The model was evaluated on the basis of user's choices and merchandise needs along with attribute ratings from the product specialists	The testing of this system was done using a review from recommendation of energy unit election for recording studio

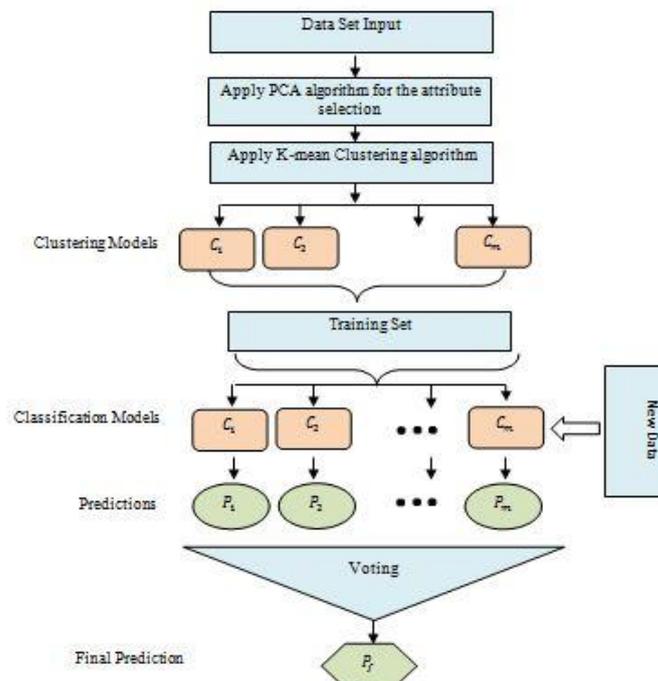


Figure 1: Voting Classification[3]

III. CONCLUSION

It is concluded that various techniques are developed to design efficient recommendation systems. The efficient recommendation systems are based machine learning algorithms. The recommendation systems have three phases which are data pre-processing, feature extraction and classification. It is analyzed that existing recommendation systems give low accuracy. In future, voting classification method will be designed which give high accuracy. The proposed technique can improved using ensemble classification algorithm

REFERENCES

- [1]. Alexandrin Popescu and Lyle H. Ungar, David M. Pennock and Steve Lawrence, "Probabilistic Models for Unified Collaborative and Content-Based Recommendation in Sparse-Data Environments", POPESCUL ET AL, volume 4, issue 12, pp- 16-22, 2001.
- [2]. Greg Linden, Brent Smith, and Jeremy York, "Amazon.com Recommendations Item-to-Item Collaborative Filtering", IEEE Computer Society, volume 28, issue 5, pp- 1634-1639, 2003.
- [3]. Jun Wang, Arjen P. de Vries, Marcel J.T. Reinders, "Unifying Userbased and Itembased Collaborative Filtering Approaches by Similarity Fusion", ACM Publications, volume 13, issue 74, pp- 532-538, 2006.
- [4]. Sergio Cleger-Tamayo, Juan M. Fern´andez-Luna, and Juan F. Huete, "A New Criteria for Selecting Neighborhood in Memory-Based Recommender Systems", Springer-Verlag Berlin Heidelberg, volume 16, issue 9, pp- 258-267, 2011.
- [5]. Pasquale Lops, Marco de Gemmis and Giovanni Semeraro, "Content-based Recommender Systems: State of the Art and Trends", Springer, volume 32, issue 8, pp- 409-416, 2011.
- [6]. Bin Xiao, E. Aimeur, J.M. Fernandez, "PCFinder: an intelligent product recommendation agent for e-commerce", IEEE International Conference on E-Commerce, CEC, volume 52, issue 69, pp. 334-345, 2003.
- [7]. Liu Guo-rong, Zhang Xi-zheng, "Collaborative Filtering Based Recommendation System for Product Bundling", International Conference on Management Science and Engineering, volume 15, issue 19, pp. 337-346, 2006.
- [8]. Yangming Zhang, Jiayin Qi, Huaying Shu, Jiantong Cao, "Personalized product recommendation based on customer value hierarchy", IEEE International Conference on Systems, Man and Cybernetics, volume 8, issue 44, pp. 1007-1015, 2007.
- [9]. Silvana Aciar, Debbie Zhang, Simeon Simoff, John Debenham, "Informed Recommender: Basing Recommendations on Consumer

- Product Reviews”, *IEEE Intelligent System*, Volume: 22, Issue: 3, pp. 108-117, 2007.
- [10]. Fuyong Yuan, Jian Liu, Chunxia Yin, Yulian Zhang, Nan Shen, “A Novel Collaborative Filtering Mechanism for Product Recommendation in P2P Networks”, *Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, volume 27, issue 33, pp- 240-246, 2007.
- [11]. M. Vamsee Krishna Kiran, R. E. Vinodhini, R. Archanaa, K. Vimalkumar, “User specific product recommendation and rating system by performing sentiment analysis on product reviews”, *4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 6, issue 14, pp. 3317-3324, 2017.
- [12]. Nishigandha Karbhari, Asmita Deshmukh, Vinayak D. Shinde, “Recommendation system using content filtering: A case study for college campus placement”, *International Conference on Energy, Communication, Data Analytics and Soft Computing*, volume 13, issue 4, pp. 1098-1105, 2017.
- [13]. Kunal Shah, Akshaykumar Salunke, Saurabh Dongare, Kisandas Antala, “Recommender systems: An overview of different approaches to recommendations”, *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, volume 15, issue 34, pp- 105-116, 2017.
- [14]. Natedao Thotharat, “Thai local product recommendation using ontological content based filtering”, *9th International Conference on Knowledge and Smart Technology (KST)*, volume 66, issue 89, pp- 423-429, 2017.
- [15]. Anna Gatzoura, Miquel Sánchez-Marrè, “A case-based reasoning framework for music playlist recommendations”, *4th International Conference on Control, Decision and Information Technologies (CoDIT)*, volume 22, issue 67, pp. 4671-4678, 2017.
- [16]. Jiahuan Li, Liqing Zhou, “Research on Recommendation System of Agricultural Products E-Commerce Platform Based on Hadoop”, *IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, volume 5, issue 9, pp. 709-718, 2018.
- [17]. Rahul Kumar Chaurasiya, Utkarsh Sahu, “Improving Performance of Product Recommendations Using User Reviews”, *3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*, volume 13, issue 4, pp. 1098-1105, 2018.
- [18]. Yograj Meena, Monika, Parveen Kumar, Ambalika Sharma, “Product Recommendation System Using Distance Measure of Product Image Features”, *Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, volume 17, issue 28, pp. 714-723, 2018.
- [19]. Teddy Oswari, Tristyanti Yunitasari, Reni Diah Kusumawati, Saurabh Mittal, “Music Recommendation System Design for Online Platform with the Concept of Unified Modeling Language (UML)”, *International Conference on Sustainable Energy, Electronics, and Computing Systems*, volume 8, issue 27, pp. 114-124, 2018.
- [20]. Tessy Badriyah, Sefryan Azvy, Wiratmoko Yuwono, Iwan Syarif, “Recommendation system for property search using content based filtering method”, *International Conference on Information and Communications Technology (ICOIACT)*, volume 11, issue 6, pp- 95-106, 2018.
- [21]. Brandon Huynh, Adam Ibrahim, Yun Suk Chang, Tobias Höllerer, John O'Donovan, “A Study of Situated Product Recommendations in Augmented Reality”, *IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, volume 16, issue 11, pp- 74-86, 2018.
- [22]. Bo Wang, Feiyue Ye and Jialu Xu, “A Personalized Recommendation Algorithm Based on the User’s Implicit Feedback in E-Commerce”, *Future Internet*, volume 10, issue 117, pp. 3307-3314, 2018.
- [23]. Ni Made Satvika Iswari, Wella Wella, Andre Rusli, “Product Recommendation for e-Commerce System based on Ontology”, *1st International Conference on Cybernetics and Intelligent System (ICORIS)*, volume 15, issue 34, pp- 105-116, 2019.
- [24]. Udom Janjarassuk, Sudatip Puengrusme, “Product recommendation based on genetic algorithm”, *5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, volume 27, issue 15, pp. 106-114, 2019.
- [25]. Phan Duy Hung, Dinh Le Huynh, “E-Commerce Recommendation System Using Mahout”, *IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, volume 13, issue 74, pp- 532-538, 2019.