Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Engineering

Elixir Comp. Engg. 119 (2018) 51034-51041

ISSN: 2229-712X

Analysis of Product Recommendation System Using Machine Learning Algorithms

Sachin Sirohi, Sanjeev Kumar Sharma and Naveen Kumar

JPIET, Meerut, Uttar Pradesh.

ARTICLE INFO

Article history:

Received: 07 May 2018;

Received in revised form: 07 June 2018;

Accepted: 16 June 2018;

Keywords

Machine Learning,
Classification,
Euclidean Distance.

ABSTRACT

This paper presents a Product recommendation system based on metric analysis of product descriptions. The developed system ranks the catalog of products and offers corresponding items to the user's request while, at the same time, selecting the most diverse items. An algorithm for ranking is developed. Based on the request, the recommendation system finds the distance from this request to all documents from the collection of data. The request and the collection of data are sets of features. The system ranks the results in accordance with the following rules: minimizes the distance from the query to the relevant results, maximizes the distance from the query to the irrelevant results and maximizes the distance between the relevant query results. For ranking, Heterogeneous Euclidean-Overlap Metric (HEOM) of clothes catalogue items is used. HEOM metric uses different attribute distance functions to measure distances between objects in mixed scales. A dataset of clothes catalogue items is collected. The system, in addition to the basic attributes given as text descriptions of product, uses attributes based on expert description such as fashion, psychological age and attractiveness. The dataset has features of text, linear and nominal scales. The computational experiment shows the effectiveness of the proposed algorithm. The importance of features of the collection of data is defined. A software product demonstrating the recommendation system in action is developed.

© 2018 Elixir All rights reserved.

1. Introduction

As of late with springing of the web, significant measure of information is available on the web, nevertheless it is difficult to manage every one of the information and that results into information over-burden, to vanquish the issue of information over-burdening suggestion structure was introduced. The principle point of the recommendation framework is to prescribe the most appropriate things to the client. Nowadays, clients depend on recommendations or suggestions from other individuals by talked words, reports of news from news media, general studies, reference letters, travel guides and many others like this.

In this manner, recommendation system assumes a vital part in finding the best things. A recommendation system filtered the information through information investigation methods which are valuable to prescribe the reasonable things to the client. Recommendation systems work from a particular kind of data filtering system method that endeavors to suggest information things (motion pictures, TV program/appear/scene, music, books, news, pictures, website pages, logical writing and so forth.) or social components (e.g. individuals, occasions or gatherings) that are probably going to hold any importance with the user [1]. The recommendation system likewise thinks about the client profiles and looks to foresee the evaluations.

1.1 Types of Recommendation System

Moreover, the Recommendation system use opinions about the community of users and to determine content of interest using certain rules extractions. Recommendation systems are classified into 3 approaches which are collaborative, content-based or knowledge-based method to have a better recommendation. The utilization of productive and exact recommendation techniques is critical for a framework that will give great and valuable recommendation to its individual clients. This clarifies the significance of understanding the highlights and possibilities of various recommendation techniques. Figure 1.1 shows the anatomy of different recommendation filtering techniques.

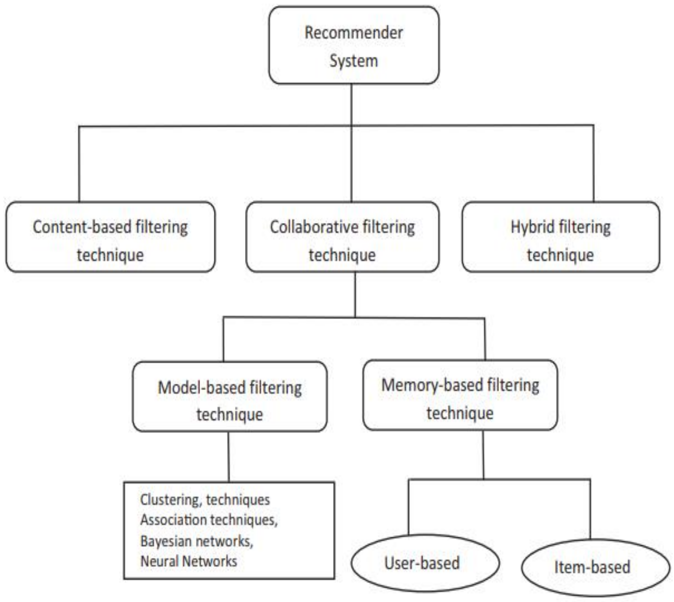


Figure 1.1. Recommendation System Filtering Techniques.

Tele: 9457888087
E-mail address: sachinsirohi47@gmail.com

© 2018 Elixir All rights reserved

A. Collaborative Based Recommendation System

Collaborative filtering Algorithm recommender system ended up a standout amongst the most looked into strategies of recommender systems. in the event that clients had similar interests before, they will likewise have comparative tastes later on. For example, if client an and client B have a buy history that covers firmly and client A has as of late purchased a thing that client B has not yet been, the fundamental method of reasoning is to propose this thing likewise to client B. Collaborative filtering is an area autonomous expectation method for content that can't without much of a stretch and satisfactorily be depicted by metadata, for example, motion pictures and music. Collaborative filtering method works by building a database (client thing network) of inclinations for things by clients. It at that point matches clients with important intrigue and inclinations by computing similitudes between their profiles to make recommendations. Such clients manufacture a gathering called neighborhood. A client gets recommendations to those things that he has not evaluated previously but rather that were at that point decidedly appraised by clients in his neighborhood. Recommendations that are delivered by CF can be of either forecast or recommendation. Forecast is a numerical esteem, Rij, communicating the anticipated score of thing j for the client I, while Recommendation is a rundown of best N things that the client will like the most as appeared in Fig. 1.3. The system of collaborative filtering can be partitioned into two classifications: memory-based and model based.

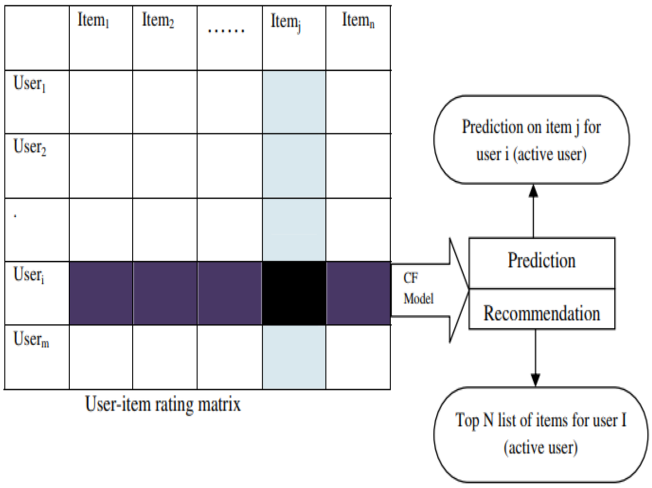


Figure 1.2. Collaborative Filtering process.

The collaborative filtering method recommends items in light of user based approach and item based approach [2].

1) User-based approach

In user-based approach user plays critical part. On the off chance that specific gathering of users has same taste, at that point they join into one gathering. Proposals are given to user based on assessment of things by different users shape a similar gathering, with whom he/she shares regular inclinations. If the item was positively rated by the community , it will be recommended to the user.

2) Item-based approach

In Item-based approach the items assume a critical part. Recommendations are based on assessment of the items. The framework creates recommendations with items in the area that a client would incline toward.

B. Content Based Recommendation System

In Content-based recommender systems, it manages profiles of the clients that were made toward the start.

A profile has data about the client and his/her taste which is based on how a client rates the things. In the recommendation process, the motor analyzes things that were at that point evaluated by client with things that he/she didn't rate and searches for similitudes. The things which are profoundly appraised will be prescribed to the clients [2].

With regards to content-based recommendation, the accompanying inquiries that emerge:

• How to figure out which items match, or minimum like or good with, a client's advantages?

• Which techniques can be utilized to consequently remove or take in the item portrayals?

C. Hybrid Recommendation System

In Hybrid Recommendation it is a blend of both collaborative approach and content-based approach. With the assistance of Hybrid Recommendation, distinctive kinds of issues can be effortlessly beaten the issue, for example, Cold-Start issue can be taken care of utilizing the Hybrid Recommendations[3]. The blend of methodologies can continue in various ways:

1) How does implementation of algorithm take place?

2)How to utilize some rules of content filtering in collaborative filtering?

3)How to extract rules in the Hybrid Recommendation System?

2. Related Work

S.Pandya [10]. In this paper, the real difficulties, for example, "data sparsity" and "cold start problem" are tended to.To defeat these difficulties, they propose another philosophy by consolidating the clustering algorithm with Eclat Algorithm for better principles age. Right off the bat, they group the rating grid in view of the user similarity. At that point they change over the clustered data into Boolean data and applying Eclat Algorithm on Boolean data productive principles age happens.Finally, in view of guidelines age suggestion happens. Our investigations demonstrate that approach diminish the sparsity level as well as increment the precision of a framework.

Hirdesh Shivhare et al. [11] proposed an integrative technique by combining fuzzy c-implies clustering strategy and genetic algorithm based weighted similarity measure to construct a movie recommendation system. The proposed motion picture recommendation system gives better similitude measurements and quality than the current Movie recommendation system but the computation time which is taken by the proposed recommendation system is more than the existing recommendation system. This problem can be settled by taking the grouped data focuses as an information dataset

Gaurangi Tilak et al.[12] introduced Movie GEN, a specialist framework for motion picture proposal. They executed the framework utilizing machine learning and group investigation based on the basis of hybrid recommendation method. In view of the Support Vector Machine forecast it chooses motion pictures from the dataset, groups the motion pictures and creates inquiries to the clients. In light of the user`s answers, it refines the motion picture set lastly prescribes films to the clients.

Davidson,et al.[13] examined about the recommendation system that YouTube uses to prescribe recordings to its clients in view of the past action of the clients and keeping in mind that doing that, they have likewise talked about and brought up a few difficulties that is looked by the arrangement of YouTube while doing its undertaking.

They gave their test points of interest and assessment system that they utilized for testing and tuning of new algorithms.

H. Chen and A.Chen [14] composed a music recommendation system. Their work process comprises of dissecting the music objects, determining the representative track, separating six highlights from the track. That is the means by which music objects are assembled. So as to comprehend the interests of the clients, their entrance history is breaking down. Suggestion techniques are proposed chiefly in light of the favored degrees of the audience members to the music groups.

Ahmed,et al.[15] exhibited how TV series recommendation can be challenging and not the same as motion picture recommendation.In TV series recommendation,time responsibility issue should be dissected. Other than investigating classification, which includes some additional work TV series recommendation and this paper demonstrated an approach to accomplish that using fuzzy systems.

Park, Hong, and Cho [16] proposed a recommendation framework which is customized where clients' inclination is reflected by Bayesian Networks. The parameters are found out from a dataset though the structure of the Bayesian Network was worked by a specialist. The framework they proposed works by gathering setting data, for example, area, time, and climate condition. It additionally breaks down client ask for from the cell phone to surmise the most supported thing with the goal that it can give a proper office by indicating it in the guide.

Baatarjav, Phithakkitnukoon, and Dantu [17] presented a group recommendation framework for the prominent social network Facebook understanding the issue clients experience to discover right gatherings. They utilized a blend of hierarchical clustering technique and decision tree. They worked on understanding groups by analyzing the member profiles.

Xiaoyingwang and Chengliang Wang [18] they develop an online business recommendation framework which depends on enhanced community oriented separating procedures. In the system algorithm, we incorporate the review topics into user-based collaborative filtering algorithm, use LDA demonstrate in view of Spark to create audit points dissemination, and afterward build up client comparability by consolidating client inclinations likeness and client rating similitude computed the positioning of the item.

Marcelo de Campos Niero, et al. [19] presents different comparability measures to ascertain the likeness between different clusters. The comes about demonstrate that sets of groups with a large difference in the number of data samples are more sensitive to errors, the number of mixtures of an external model affects the discriminative power of distance measures, and the number of estimated parameters affects the speaker discrimination. All trials are performed on a selection from TIMIT corpus and the diarization undertaking database utilized as a part of the 2002 NIST Speaker Recognition Evaluation.

Tessy Badriyah et al. [20] This exploration builds up a hybrid recommendation framework for internet business that executes Content-based Filtering and Collaborative Filtering techniques,which will process the similarities of item portrayal and client profile. In analyze comes about, it was discovered that the recommendation has closeness with item portrayal and the inclination of client profile with the normal of exactness esteem is 67.5% and review esteem is 71.47%.

Chao Ma et al.[21] built up another question recommendation framework in view of hunt objective move Graphs. Exploratory inquiry is an undeniably critical action for Web searchers. Be that as it may, the ebb and flow look framework can't give adequate help to exploratory inquiry. In this way, they made top to bottom investigation for exploratory pursuit procedures, and found that there are considerable measures of inquiry objective move wonders in exploratory hunt. In view of this reality, they have planned another inquiry recommendation strategy to help exploratory hunt. Right off the bat, as per the behavioral attributes of searchers in the hunt objective move forms, every one of the questions submitted in the inquiry objective move forms are separated from web crawler logs utilizing machine learning. And after that they have utilized the inquiries to fabricate a hunt objective move diagram; at long last, the arbitrary walk calculation is utilized to get the inquiry recommendations in the pursuit objective move chart

3. Proposed Methodology

The research problem is a crucial part of any research activity. If nature of the problem is clear that it is very easy to solve the problem.

The objective of this examination work is to create item recommendation system in light of metric investigation of item portrayals. The recommendation system discovers items like the questions. Similitude is characterized as a base separation between an inquiry and reactions. We additionally endeavor to boost the separation between applicable reactions. The most extreme separation between applicable reactions enables clients to discover important garments with greatest distinction inside a base separation to inquiry.

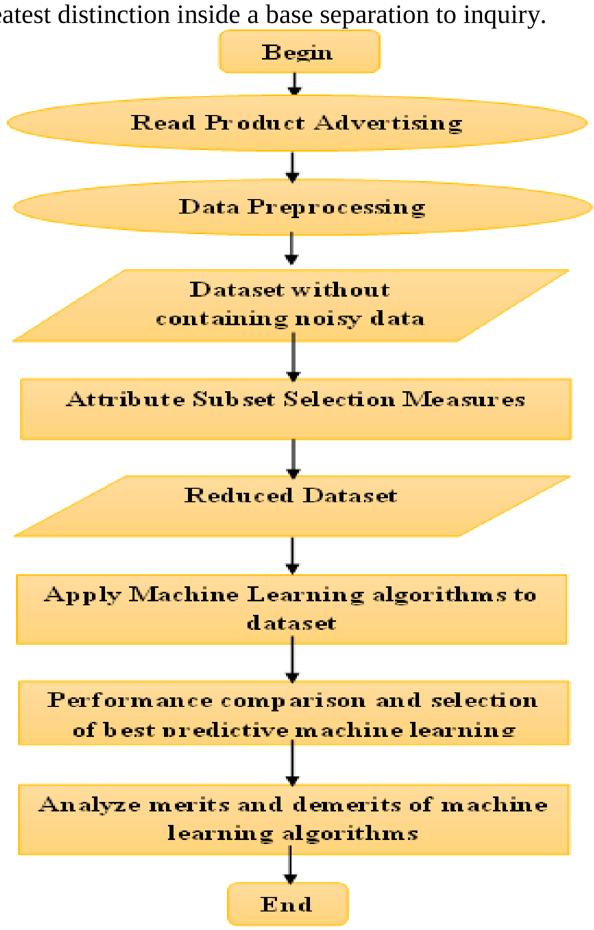


Figure 1.3. The Overall System Design for analyzing product recommendation System.

We propose the speculation that item depictions (highlights) have diverse significance (distinctive weight). It isn't important to consider all highlights to suggest client's comparable items. Most likely individuals take a gander at a few highlights and, in view of them, settle on their choices to pick items. In figure 1.3 we show an overall system design for analyzing product recommendation system.

Machine Learning Algorithms

We apply Bow and Tf-idf algorithm to the product data set. During implementation the methodology steps are applied one by one. This work compares the performance of two algorithms on the same dataset.

1. Bag of Words (BoW) algorithm

A bag-of-words model, or BoW for short, is a method for removing highlights from content for use in displaying, for example, with machine learning algorithms. The approach is extremely basic and adaptable, and can be utilized as a part of a horde of courses for separating highlights from documents. A bag-of-words is a portrayal of content that depicts the event of words inside a record. It includes two things:

- A vocabulary of known words.
- A measure of the presence of known words.

It is known as a "bag" of words, on the grounds that any data about the request or structure of words in the report is disposed of. The model is just worried about whether known words happen in the archive, not where in the record. You can think about the bag-of-words (BoW) display as a machine which takes as info an arrangement of archives and yield a table containing the recurrence include of each word each report. Let's us suppose we have only one document containing only one sentence: **I love dogs.**

Let's us apply BoW

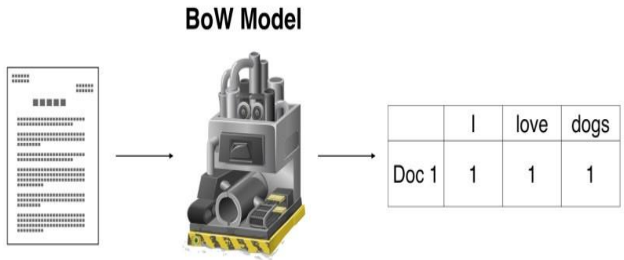


Figure 1.4. Bag-of-Words apply to One Document.

The BoW yields a table wherein each line compares to an archive and every section speaks to a one of a kind word. The passages are the include of each word each document. We can't generally get any knowledge from the table, since we're just thinking of one as archive. Every section in the table is 1, since the primary record clearly contains every interesting word in the principal document. The BoW demonstrate was made as a way to think about an arrangement of reports. There truly is no point on the off chance that you just have a solitary record.

Algorithm:

a. Collect Data.
b. Design the vocabulary.
c. Create Document vectors.
d. Managing vocabulary.
e. Scoring words.
f. Word hashing.

2. Term Frequency inverse document Frequency (Tf-idf)

The BoW model is a perfectly acceptable model to convert raw text to numbers. However, if our purpose is to identify signature words in a document, there is a better transformation that we can apply.

Tf-idf is shorthand for term frequency – inverse document frequency. So, two things: **term frequency** and **inverse document frequency**.

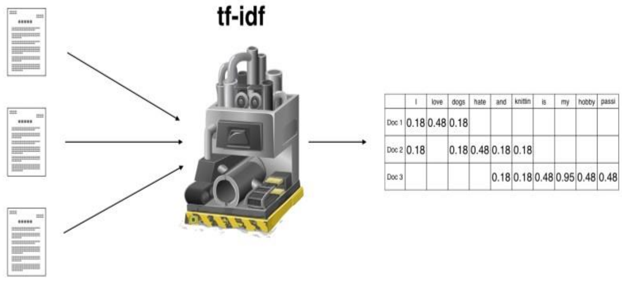


Figure 1.5. Tf-idf Model.

Inverse document frequency (idf): For a word to be considered a signature word of a document, it shouldn't appear that often in the other documents. Thus, a signature word's document frequency must be low, meaning its inverse document frequency must be high. The idf is usually calculated as

$$\mathrm{idf}(W) = \log \frac{\#(\text{documents})}{\#(\text{documents containing word W})}$$

The Tf-idf is the product of these two frequencies. For a word to have high Tf-idf in a document, it must appear a lot of times in said document and must be absent in the other documents. It must be a signature word of the document.

	I	love	dogs	hate	and	knitting	is	my	hobby	passion
Doc 1	0.18	**0.48**	0.18							
Doc 2	0.18		0.18	**0.48**	0.18	0.18				
Doc 3					0.18	0.18	**0.48**	**0.95**	**0.48**	**0.48**

Figure 1.6. Output of Tf-idf Model.

4. Implementation Result

In this paper, Python language and Anaconda tool is used in order to implement different techniques on different datasets.

1. Data Analysis and Interpretation

1.1 Snapshot of amazon product dataset

In this study, we used the cloth dataset of amazon company, which has been taken from the kaggle. Table 1.1 describe the feature details of the dataset.

Table 1.1. Features Description.

Variable Name	Represented As
Brand	Text
Color	Text
Product type name	Text
Title	Text
Price	Integer
Brand	Text
Asin (product id)	text

2. Recommendation Result

2.1 Using BoW algorithm

This algorithm takes two value as an input. First is product id of the item that you clicked and second is the number of similar products that you want to recommend to the end user.

The query is bag_of_word_model (12566,20). When we execute this query the result is displayed in the heat map form with recommended image. The result is displayed in figure 1.7.

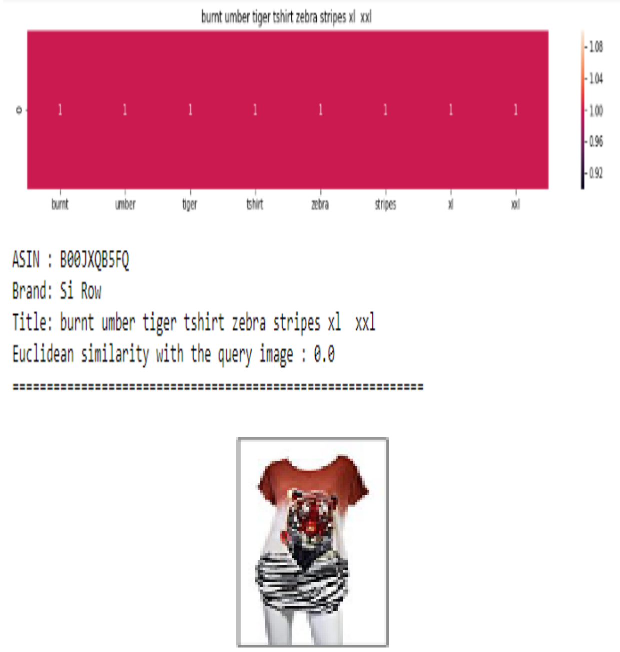
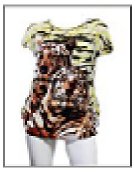


Figure 1.7. Result of Bag of Words Algorithm.

Recommended Results of the query according to the minimum Euclidean distance.

Table 1.2. Recommended Result of Bag of Words Algorithm.

Euclidean Similarity with the query image	Recommended Product
1.73205080757	
2.44948974278	
2.64575131106	
3.0	
3.0	
3.0	

3.16227766017	
3.16227766017	
3.16227766017	
3.16227766017	

2.2 Using Tf-idf algorithm

First it calculates the cosine similarity of the product id with all the remaining product id store in the corpus and maintain the distance matrix. Then it returns the list of products that has minimum distance.

The query is Tf-idf (12566,20). When we execute this query the result is displayed in the heat map form with recommended image. The result is displayed in figure 1.8.

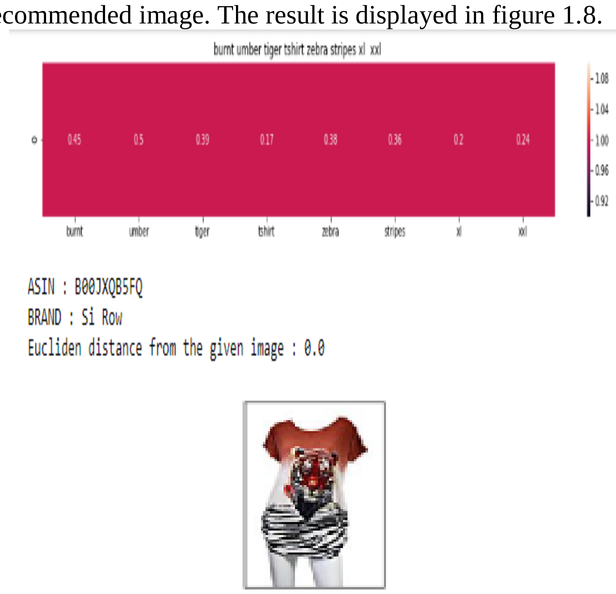


Figure 1.8. Result for Original query according to Tf-idf.

Recommended Results of the query according to the minimum Euclidean distance.

Table 1.3. Recommended Result of Tf-idf Algorithm.

Euclidean Similarity with the query image	Recommended Product
0.753633191245	

0.935764394377	
1.00007496145	
1.02321555246	
1.02321555246	
1.0319918463	
1.21068436704	
1.21216838107	
1.21979064028	
1.220684966	

2.3 Performance Comparison

The motivation behind this work is the performance shown by BoW and Tf-idf techniques. As Table 1.2 shows the accuracy details of BoW whereas Table 1.3 shows the improved performance of Tf-idf model. This shows an effective impact of Tf-idf algorithm in the prediction results of recommendation system. Bag-of-words is simple to generate but far from perfect. If we count all words equally, then some words end up being emphasized more than we need. Tf-idf is a simple twist on top of bag-of-words. It is clearly visible in table 1.3 that the results are calculated using Tf-idf algorithm are more accurately recommended as comparison to the result 1.2 which are calculated by BoW.

5. Conclusion and Future Scope

Recommendation systems have been an important in E-commerce on the web for the customer to suggest items what they would be interested. With the increasing number of users and items, recommendation systems encounter the main shortcoming: data sparsity and data scalability problems, which bring out the poor quality of prediction and the inefficient time consuming.

In this paper, I have proposed item-based collaborative filtering approach applying dimension reduction to improve the predictive accuracy and recommendation quality in overcoming the existing limitations. By reducing the noise of dimensional data, it focuses on typical and popular items to compute the similarity between them and to predict the most similar items to users. The experimental results show that this approach makes a considerable impact on providing better accuracy of prediction and much faster execution time in comparison with traditional UBCF and IBCF. It results in improving the quality of recommendation system using collaborative filtering.

The potential limitation would use this approach with dataset widely consisting of not enough ratings by users, expecting less accuracy. Therefore, to overcome this challenge, I propose an approach to mix both explicit and implicit ratings to alleviate the data sparsity problem further in this aspect.

References

1. M.H.Esfahani and F.K.Alhan, “New Hybrid Recommendation System Based On C-Means Clustering Method”,Fifth IEEE conference of information and Knowledge Technology2014, 28-30 May 2013, pp145-149.
2. Q.Wang, W.Cao and Y.Liu, “A Novel Clustering based Collaborative Filtering Recommendation system Algorithm”, An Advanced Technologies, Embedded and Multimedia for Human-centric Computing, Springer journal 2014,13 November 2013, pp 673-680.
3. R.Burke,“Hybrid Recommender Systems: Survey and Experiments”, Springer journal on user modeling and User Adapted Interaction, September 2002, pp 331-370.
4. Ricci F., Rokach L., Shapira B. Introduction to Recommender Systems (Tutorial) / Intelligent Information Access IIA08 / Cagliari, Italy. – 9-11 December 2008. – P.1-20
5. Ge, Yong, et al. An energy-efficient mobile recommender system / Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining / ACM. – 2010. – P. 983-991
6. Leskovec J., Rajaraman A., Ullman J. Mining of Massive Datasets / Cambridge University Press /ACM. – 2014. – P. 307-341
7. Adomavicius G., Tuzhilin A. Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions / IEEE Trans. on Data and Knowledge Engineering 17:6 / USA. – 2005. – P. 734–749
8. Ge, Yong, et al. An energy-efficient mobile recommender system / Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining / ACM. – 2010. – P. 983-991
9. Melville P., Sindhwani V. Recommender Systems / Encyclopedia of Machine Learning / Springer. – 2010. – P. 829-838.
10. S. Pandya and N. Joshi. A Novel Hybrid based Recommendation System based on Clustering and Association Mining

11. Hirdesh Shivhare, Anshul Gupta and Shalki Sharma (2015), "Recommender system using fuzzy c-means clustering and genetic algorithm based weighted similarity measure", IEEE International Conference on Computer, Communication and Control.
12. Eyrun A. Eyjolfsdottir, Gaurangi Tilak, Nan Li (2008), "MovieGEN: A Movie Recommendation System", 2008 Conference Proceedings
13. Davidson, James, et al. "The YouTube Video Recommendation System." [4] Chen, Hung-Chen, and Arbee L. P. Chen. "A music recommendation system based on music data grouping and user interests." Proceedings of the tenth international conference on Information and knowledge management - CIKM01, 2001,
14. Ahmed, Muyeed, et al. "TV Series Recommendation Using Fuzzy Inference System, K-Means Clustering and Adaptive Neuro Fuzzy Inference System." 2017, pp. 1512–1519.
15. Park, Moon-Hee, et al. "Location-Based Recommendation System Using Bayesian User"s Preference Model in Mobile Devices." Ubiquitous Intelligence and Computing Lecture Notes in Computer Science, pp. 1130–1139.,
16. Huang, Yao-Chang, and Shyh-Kang Jenor. "An audio recommendation system based on audio signature description scheme in MPEG-7 Audio." 2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763),
17. Baatarjav, Enkh-Amgalan, et al. "Group Recommendation System for Facebook." On the Move to Meaningful Internet Systems: OTM 2008 Workshops Lecture Notes in Computer Science, 2008, pp. 211–219.
18. Xiaoying wang and Chengliang Wang "Recommendation system of e-commerce based on improved collaborative filtering algorithm." 2017 8th International Conference on Software Engineering and Service Science (ICSESS), pp. 332-335.
19. Marcelo de Campos Niero, et al. "A Comparison of Distance Measures for clustering in Speaker Diarization." 2014 International Telecommunications Symposium (ITS), pp. 1-5.
20. Tessy Badriyah et al. "A Hybrid Recommendation System for E-Commerce based on Product Description and User Profile." 2017 The Seventh International Conference on Innovative Computing Technology (INTECH 2017), pp. 95-100.
21. Chao Ma et al. "A New Query Recommendation Method Supporting Exploratory Search Based on Search Goal Shift Graphs." IEEE Transactions on Knowledge and Data Engineering 2018, pp. 1-13.
22. Deepa Shankar et al. "A Novel Online Product Recommendation System Based on Face Recognition and Emotion Detection." 2015 International Conference on Control,Instrumentation, Communication and Computational Technologies (lCCICCT), pp. 329-334.
23. Li-Hua et al. "A Simplified Method for Improving the Performance of Product Recommendation with Sparse Data." 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST 2017), pp. 318-323.
24. Rahul Katarya et al. "An effective collaborative movie recommender system with cuckoo search." Egyptian Informatics Journal 18 (2017) 105–112
25. Shreya Agarwal et al**. "**An improved approach for movie recommendation system." International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017), pp. 336-342.
26. T. Sangeetha et al. "Aspects Based opinion mining from online reviews for product recommendation," 2017 International Conference on Computational Intelligence in Data Science(ICCIDS), pp. 1-5.
27. Julian Villegas and Shoma Saito "Assisting System for Grocery Shopping Navigation and Product Recommendation," IEEE 6th Global Conference on Consumer Electronics (GCCE 2017), pp. 1-4.
28. Anwesh Marwade et al. "Augmenting E-Commerce Product Recommendations by Analyzing Customer Personality," 9th International Conference on Computational Intelligence and Communication Networks 2017, pp. 174-180.
29. Basit Mehmood Khan "Collaborative Filtering based Online Recommendation Systems: A Survey," International Conference on Information and Communication Technologies (ICICT) 2017, pp. 125-130.
30. Hermawan Prasetyo "Comparison of Distance and Dissimilarity Measures for Clustering Data with Mix Attribute Types," 1st International Conference on Infonnation Technology, Computer and Electrical Engineering (ICITACEE) 2014, pp. 276-280.
31. Hemant Kumar "Comprehensive Review on Supervised Machine Learning Algorithms," International Conference on Machine learning and Data Science 2017, pp. 37-43.
32. John O. Awoyemi "Credit card fraud detection using Machine Learning Techniques: A Comparative Analysis," International Conference on Computing Networking and Informatics (ICCNI) 2017, pp. 1-9.
33. Pengfei He "CR-IIA:Collaborative Recommendation Algorithm Based on Implicit Item,"
34. Jaydeep Chakraborty "Data Extraction and Integration for Scholar Recommendation System," 12th IEEE International Conference on Semantic Computing 2018, pp. 397-402.
35. Suvarn Sharma "Data Preprocessing Algorithm for Web Structure Mining," Fifth International Conference on Eco-Friendly Computing and Communication Systems (ICECCS-2016), pp. 94-98.
36. Makbule Gulcin Ozsoy "From Word Embeddings to Item Recommendation,"
37. Muyeed Ahmed "Movie Recommendation System Using Clustering and Pattern Recognition Network," pp. 143-147.
38. Noel Lopes et al. "Novel Trends in Scaling Up Machine Learning Algorithms," 16th IEEE International Conference on Machine Learning and Applications 2017, pp. 632-636.
39. Bishwajit Purkaystha "Product Recommendation: A Deep Learning Factorization Method Using Separate Learners," 20th International Conference of Computer and Information Technology (ICCIT), 22-24 December, 2017, pp. 1-5.
40. Junnan Chen "Product recommendation system for small online retailers using association rule mining," Internation conference on innovative Design and manufacturing 2014, pp. 71-77.
41. Kaveri Roy "PRODUCT RECOMMENDATIONS USING DATA MINING AND MACHINE LEARNING ALGORITHMS," ARPN Journal of Engineering and Applied Sciences 2017, pp. 5475-5483.
42. F.O. Isinkaye "Recommendation systems: Principles, methods and evaluation," Egyptian Informatics Journal (2015) 16, pp. 261–273.
43. P. Sukumar "Review on Modern Data Preprocessing Techniques in Web Usage Mining (WUM)," International

Conference on Computational Systems and Information Systems for Sustainable Solutions 2016, pp. 64-69.
44. R. Ananthakrishna, S. Chaudhuri, and V. Ganti, Eliminating Fuzzy Duplicates in Data Warehouses. VLDB, pages 586-597, 2002.
45. A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios. Duplicate Record Detection: A Survey, IEEE TKDE, 19(1):1-16, 2007.
46. Oard DW, Kim J. Implicit feedback for recommender systems. In: Proceedings of 5th DELOS workshop on filtering and collaborative filtering; 1998. p. 31–6.