Generic Recommendation Engine using Hybrid Filtering Model

Dr Sushmitha Valli,K Abhijith Saralaya

Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology

2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) | 978-1-6654-9781-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/CONECCT55679.2022.9865766

ABSTRACT:

Recommendation system provides the facility to understand a person's taste and find new, desirable content for them automatically based on the pattern between their likes and rating of different items. Recommendation systems are mainly employed in applications such as online market, which works with big data. Performing data mining on big data is a tedious task due to its distributed nature and enormity. There are humanely overwhelming number of items for us to inspect, evaluate and choose from. This poses a huge challenge, since overwhelming the customers with huge catalog of items out of which the major portion of items are unrelated to user preferences.

There is an imminent need for a recommendation system that eases the process of choosing products by the user and thereby enriching the user experience. To overcome this problem, a recommendation system that uses multiple ML algorithms, a hybrid version of content based filtering and collaborative item-item filtering algorithm is implemented so as to achieve better accuracy in recommendations. The project is aimed to result in a generic recommendation engine suitable for using with any type of items irrespective of domain and datasets.

Keywords—user preferences, big data,item-item collaborative filtering

1. INTRODUCTION

Recommender Systems are tools that emerged in the 90s which are commonly defined as software tools and techniques used to provide suggestions for items to users.

Recommendation algorithms are mostly used on e-commerce web sites like Amazon, Flipkart and Myntra where they make use of customer's interests and display a subset of items. Many of these algorithms use only the items that are purchased or viewed by customer previously.

But other attributes like demographic data, favorite items, favorite sellers, artists can also be used for much effective recommend dation. Recommender systems are very good at handling the information overload problem, they provide a customized, personalized set of recommendations for each specific user thereby showing them with content that is relevant to them, thereby easing the amount of effort the users need to exert to filter and find items that they desire. These systems act as means of assisting the social process of using others suggestions, reviews when there is no previous knowledge at the user-side. These systems can either make use of collaborative filtering, content based filtering or hybrid filtering.

People have always relied on other people's suggestions for decision making whenever there are many options in order to make the best decision. In the last decade or so,the amount of digital information has grown in an exponential manner, leading to huge information that is mostly not rated and arranged properly. Information overload is difficulty in understanding an issue and making decisions when one has too much information about that issue, it is generally associated with excessive quantity of information. Information overload generally occurs when a person is exposed to huge and more information than the brain can process at one time." As more and more complex information is taken in by us in a very less amount of time and we happen to have more options laid out in front of us, our brains start to panic, resulting in us losing the ability to make good decisions.

These recommendation algorithms find a set of customers who also purchased a similar subset of items that are also purchased by the user. Then concerned ratings are also considered for filtering. Then finally all these items are aggregated from the previously computed similar set of customers, in the meanwhile all the items that are already previously purchased by the user are eliminated, in turn showing the remaining list of items. There are two types of such algorithms, these are collaborative filtering and cluster models. Other less popular versions are search based methods which focus on finding similar items but not similar customers. Amazon's item-item collaborative filtering algorithm is one such example of this.

2. LITERATURE SURVEY

Hybrid Recommendations: -

In this paper [7] the author details the intricacies of hybrid recommendation systems. Hybrid algorithms are implemented in several ways either by making collaborative-based predictions and content-based separately and then combining them or by adding collaborative-based capabilities to a content-based approach and vice-versa or by unifying the approaches into one model. Several studies that compare the performance of the hybrid with the pure collaborative and content-based methods demonstrated that the hybrid methods can provide more accurate recommendations.

Hybrid recommendation algorithms can be used to overcome problems like cold start problem and knowledge engineering bottleneck, sparsity problem that arise in recommendation scenarios. It demonstrates offline and online similarity computations for better scalability.

Item-to-item collaborative filtering: -

In this paper[1] the author has proposed a collaborative filtering algorithm based on items rather than users. Instead of matching users to similar customers, item-to-item collaborative filtering matches the user's purchased items to similar items, then in turn combining them to generate a recommendations list. A similar-items table is built to find items that customers tend to purchase together. We could build a product-to-product matrix by iterating through all item pairs and computing a similarity metric for each pair. It exposes the drawbacks of user-based filtering like scalability and sparsity. The paper also introduces correlation based similarity.

Content based filtering vs collaborative filtering:-

In this paper[2] the author details the two most popular algorithms already present namely Content based filtering and Collaborative based filtering. The paper discusses the basics of recommendation systems and also gives a highly intuitive overview of how & why recommendation systems work the way they work. Collaborative filtering approaches build a model from a user's past behavior along with decisions that are similar which are made by other users and then these algorithms use that model to predict items that the user may mostly be interested to engage with. Content-based filtering algorithms make use of a series of characteristics of an item in order to recommend additional items with similar properties. Content-based filtering is another approach for recommender systems. These methods are based on user preferences and user's past behavior in addition to description of the item and. The keywords are used to describe an item to indicate the item's characteristics which are used for generating recommendations. These algorithms try to recommend items that are similar to those items that a user liked in the past.

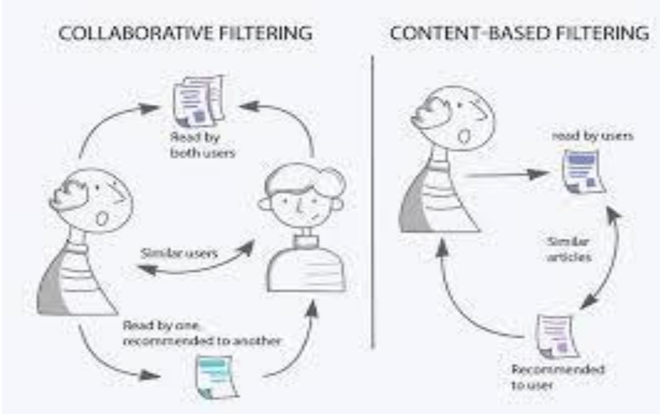


Fig 1: Collaborative filtering & content based filtering

3. PROPOSED SYSTEM

The system makes use of two already present-tested recommendation algorithms mainly content based and collaborative item-item filtering algorithms. Behavioral events like search terms used, navigation-history, items bought, items rated and many other data points derived from user input are captured for feature extraction.

The whole system is dependent on item-item similarities. This similarity computation is very expensive and so this similarity computation is done periodically or whenever a new item is introduced into the catalog. Two items similarity is directly proportional to the similarity score of these items given by the respective algorithm used to compute the result and so for two items there are two similarity scores one from content based recommendation and other from item-item collaborative filtering algorithm. These two scores are combined into a hybrid score which mandates the system on what to recommend to the user. Content based score denotes the item-item similarity based on the item-attributes only.

Collaborative based score denotes the similarity based on user ratings.

This hybrid score is calculated as:

hybridScore=contentScore+item-itemCollaborativeScore*(avgRatingOftheTwoItems)

Higher the hybrid score, higher the similarity between the two items.

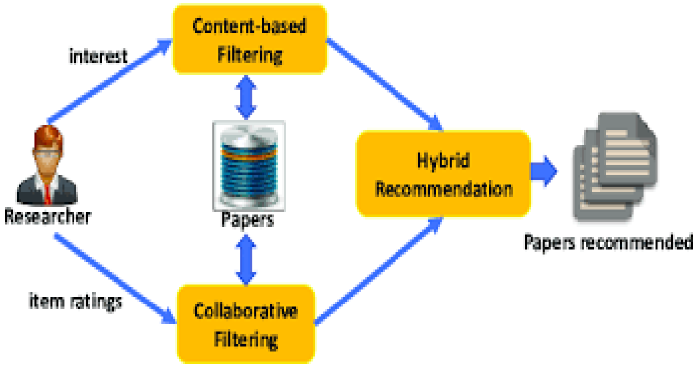


Fig 2: Hybrid recommendations

Firstly, content based recommendation is applied on the whole dataset of the items resulting in a subset of similar items, similar to the items the user already used/rated. These subsets of similar items are then subjected to the item-item collaborative filtering algorithm and the corresponding scores are computed. Then finally, the hybrid scores of these items are computed and sorted accordingly and then the top-N items are recommended to the user.

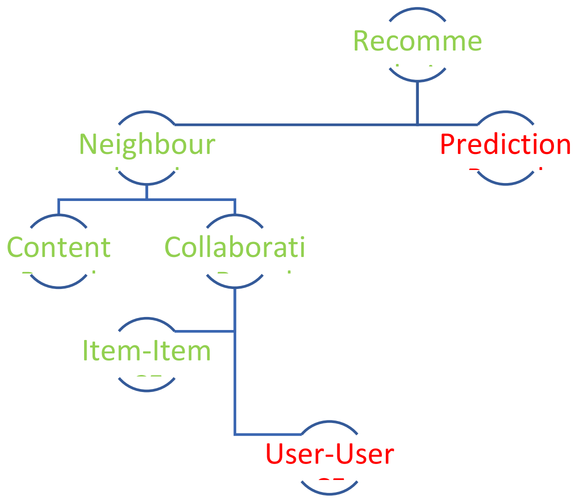


Fig 3: Approaches considered for the proposed system

Green colour denotes the techniques used in the proposed system,where red denotes techniques that are not injected in the system.

Reasons to have chosen Item-Item CF over User-User CF:

- Works well with item-centric approach of the project
- Scales relatively better with increasing number of users
- Less space requirements than User-User CF
- Far easier to handle changing user preferences

Items can be recommended in two ways:

1. By Item-Top similar items of just the given item are recommended.
2. By User -All the similar items of the user-rated items are aggregated and then recommended.

There are many advantages of using a hybrid system over choosing a specific algorithm only. As combining multiple systems facilitates us to eliminate disadvantages in one system by complimenting them with advantages of the other systems.

3.2 HIGH LEVEL ARCHITECTURE

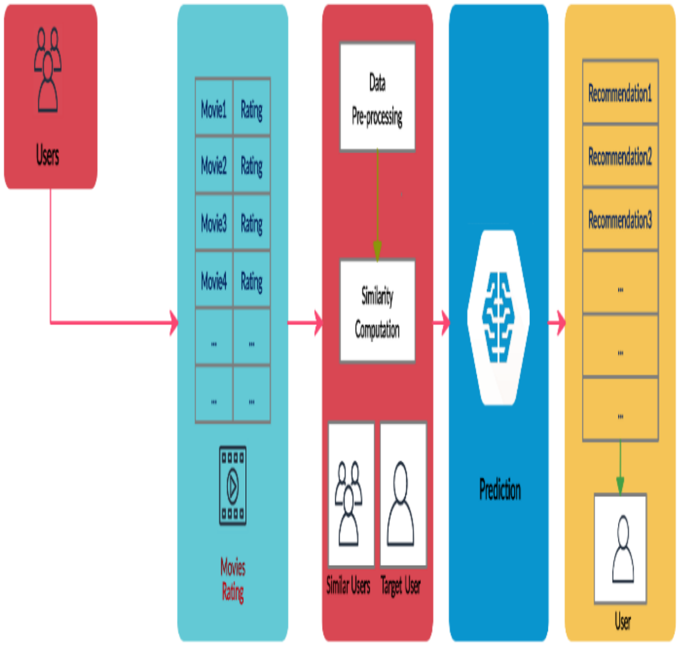


Fig 4: Proposed System Architecture

3.3 PROCESS

The above figure represents the high-level architecture of the project. Here users can perform four key operations. They are:

1.Get similar items.

2. Get user recommendations.

3.Adding rating to an item by a user.

4.Searching for an item by a user.

1.If the user has used some item and just wants the recommendations to be based upon just that one item, then the user can use up this functionality by supplying just the identifier (title or Id) of that product only.

2.If the user wants recommendations based on his whole user profile, he can just choose this option resulting in recommendations that take note of all his ratings comprehensively.

3.If the user wants to add ratings to an item, then he/she can do so by inputting the id and rating to be allocated. This activity is then captured by the system in subsequent similarity matrix computations and the user recommendations for the user gets altered accordingly.

4.If the user wants to search for an item in the item catalog, then it can be done so by inputting the search term along with userId.The search results for each user differ even for a same search term as this results also factor in the user's profile along with search term provided at that instant.

Structure Of Item:

- Each item is stored as vector of user-ratings
- Stored in the format of python dictionary Eg: HarryPotterBook:{Bhaskar : 5,RajShekar: 2.5,Naveen: 4,Priyanka: 3}
- Stored in such a way as a user will not rate all products available in the store, saving space.

Item-Item similarity score computation:

- Similarity score is computed using Cosine similarity

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

- Here A and B are item vectors.
- We can use the Cosine Similarity algorithm to work out the similarity between two things. We might then use the computed similarity as part of a recommendation query.
- Higher the similarity score, greater the similarity.
- Score ranges from 0 to 1.

Item and User ratings :

Table 1.Item and its user ratings

Item/User	RajShekar	Priyanka	Naveen	Bhaskar
HarryPotter	2	3	4	5
Marvel Comics	4	4	1	1
Narnia	2	4	3	4

Naveen and Bhaskar seem to like HarryPotter and Narnia,whereas not like Marvel Comics at all.This pattern is captured by the system by making use of similarity matrices to recommend items to the users in the future.

Similarity Matrix :

Table 2. Similarity matrix generated using cosine similarity

Book	HarryPotter	Marvel Comics	Narnia
HarryPotter	1	0.68	0.98
Marvel Comics	0.68	1	0.56
Narnia	0.98	0.56	1

Since HarryPotter is more similar to Narnia the score (0.98) is more for them,whereas HarryPotter is not so similar to Marvel Comics ,the score(0.68) is relatively less when compared to the former.

Ranking items by aggregating similarity scores per user:

- For each item that the user has rated ,individual recommendations are generated.
- All these recommendations are clubbed together,resulting in a final list of items to be recommended.
- Top N items from the final list are shown to the user.

Technical Optimizations:

- Missing data like summary and author details for a book are fetched from public books API s.
- Some ratings in the dataset are padded with a constant value to reduce bizarre scenarios.
- HashTables are used in the form of python dictionaries to optimize similarity score computation time.

3.4 ALGORITHM:

- The program contains two processes running simultaneously.These are:

1. On User Demand Process
2. Background Periodic Process

- The background periodic process affects the On User Demand Process directly.
- Both these processes work hand in hand together to yield fine tuned recommendations.

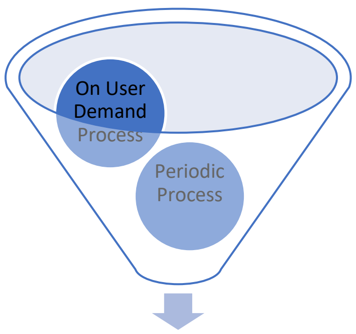


Fig 5: Processes in the system

On User Demand Process:

- It starts execution when the users select any feature in the system and start interacting with the system.
- In this process,the various content based and collaborative similarity matrices are fetched which are generated by the periodic process.
- These similarity matrices are taken up and based on the appropriate feature selected by the user,the corresponding items are retrieved.
- Top N similar items to a given item/user are filtered using CFR Score from all the products.
- These top N items are sorted using their corresponding CBR score.
- Then out of them,top K items are filtered and sorted using hybrid score computed.
- Then these sub items are shown as recommendations on the screen.

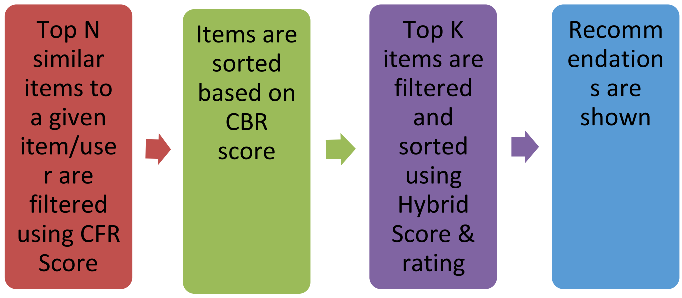


Fig 6: Steps involved in User Demand Process

- With every step execution ,the no. of candidate items for recommendations are reduced by an order of 10.

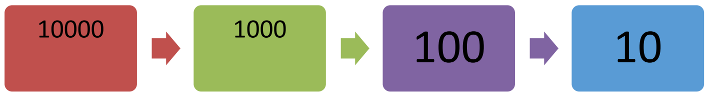


Fig 7. Order of shrinking of recommendations with each steps

Periodic Process:

- This process starts its execution when the server is first started on.
- It runs once in a while, like once in a day as it is a very computationally expensive process.
- It is responsible for generating similarity scores and updating similarity matrices accordingly.
- These similarity scores are computed with the help of item vectors which are generated inturn using the underlying dataset.

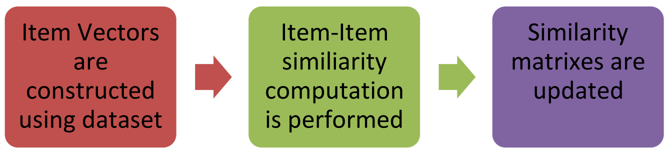


Fig 8: Flow of periodic process

PseudoCode:

```
def getItemRecommendations(userId,num_of_items=100):
    user=getUser(userId)
    sim_items={}
    for itemId in user['ratings'].keys():
```

```
for itemId,cbr_score,cfr_score,matches,avg_rating in
getHybridSimilarItemsForAItem(itemId):
        if itemId in sim_items:
sim_items[itemId]+=cbr_score+cfr_score*matches*avg_ratin
g
else:
sim_items[itemId]=cbr_score+cfr_score*matches*avg_rating
   sim_items_list=[]
   for itemId,score in sim_items.items():
      sim_items_list.append((itemId,score):
   sim_items_list.sort(key=lambda x: (x[1]),reverse=True
   return sim_items_list[:num_of_items]
end procedure
```

getItemRecommendations(userId): *It generates item-recommendations to users.*
The function ***getUser(userId)*** *gets the user profile of the user with specified userId.*
The function get ***HybridSimilarItemsForAItem(itemId)*** *gets the most similar items to the item identified by itemId based on the hybrid scores computed internally beforehand.*

4. RESULTS

DataSet used: Book-Crossing Dataset
(http://www2.informatik.uni-freiburg.de/~cziegler/BX/)

The Book-Crossing dataset comprises 3 tables.
1.BX-Users
This file contains data about the users. Note that user IDs are anonymized and map to integers. Demographic data is provided like location and age. Otherwise, these fields contain NULL-values.
Entries-2,60,000

2.BX-Books
Books are identified by their respective ISBN. Invalid ISBNs are removed from the dataset. The file contains columns (`Book-Title`, `Book-Author`, `Year-Of-Publication`, `Publisher`,'Genre','Description'), obtained from Amazon Web Services. In the case of several authors, only the first author details are provided.
Entries-15,450

3.BX-Book-Ratings
Contains the book rating information. Ratings are expressed on a scale from 1-10, higher values denoting higher appreciation.
Entries- 10,48,574

4.1 Item Recommendations For A User

User with an user-id 9 is presented below,we can notice that the user is religious by the fact that the user has used 'Testament' and 'Beloved'.

```
Select any one
1.Get user recommendations
2.Get similar items
3.Print user Info
4.Get item info
5.Add Rating to a Item
6.Search for item
7.Exit

1
Enter UserId
9
User Detail
id:- 9
location:- germantown, tennessee, usa
age:- nan

-----------Previously Used Items----------
The Testament 3
Beloved (Plume Contemporary Fiction) 4
Our Dumb Century: The Onion Presents 100 Years of Headlines from America's Finest News
Insights regarding recommendations

Focal Points of CBR recommendation are-
Fiction,John Grisham,Toni Morrison,notFound,Juvenile Fiction,

Focal Points of CFR recommendation are-
Fiction,notFound,Juvenile Fiction,James Patterson,Family & Relationships,

Focal Points of hybrid recommendation are-
Fiction,Juvenile Fiction,David Baldacci,notFound,Elizabeth Berg,
Convergence score= 0.9128846280565871
------------- CBR Users --------------
    The Testament   482
    Beloved: A Novel (Plume Contemporary Fiction)   9928
    Beloved   2865
    Beloved   9016
    Jazz (Plume Contemporary Fiction)   11255
    The Testament   2039
    Sula   1671
    The Lesson of Her Death   8118
    The Firm   195
    The Partner   917
```

Observation: Directly similar books of type novels and of fiction and juvenile fiction genre are recommended similar to users previously used items.

```
-------------- CFR Users ---------------
    The Handmaid's Tale : A Novel   2459
    Pop Goes the Weasel   2322
    Piercing the Darkness   9205
    Inferno (Mentor)   3357
    Eclipse Bay   9815
    The Return of the King (The Lord of the Rings, Part 3)
    Tuck Everlasting   5214
    The Summons   2871
    Standing in the Rainbow : A Novel   1030
    Black Like Me   3079
-------------- END ---------------

-------------- Hybrid Users ---------------
    Pop Goes the Weasel   2322
    Saving Faith   4972
    Secret History : A Novel   4605
    The Last Suppers   3944
    The Pull of the Moon   8616
    Field of Thirteen   2935
    Circle of Three: A Novel   793
    American Psycho (Vintage Contemporaries)   7160
    Mother Earth Father Sky   10839
    Good Omens   696
-------------- END ---------------
```

Hybrid recommendations are generated by including both CBR and CFR recommendations.Mostly religious and fiction novels are recommended as we can see that “Saving Faith” , “The Last Supper” and “Good Omens ” are listed in Hybrid recommendations.

4.2 Getting similar items

```
Select any one
1.Get user recommendations
2.Get similar items
3.Print user Info
4.Get item info
5.Add Rating to a Item
6.Search for item
7.Exit

2
Enter ItemId
101
Item Info
id:- 101
isbn:- 0345404793
title:- Protect and Defend
author:- Richard North Patterson
pub-year:- 2001.0
publisher:- Ballantine Books
category:- Fiction
description:- When the Chief Justice of the Supreme Court drops dead following the inauguration, President

------------- CBR Items --------------
   Protect and Defend   0.9999999999999999
   The Final Judgement   0.1610585905467913
   The Final Judgment   0.1420584879801597
   No Safe Place   0.1345698826019582
   Let Me Call You Sweetheart   0.1117218333581819
   LET ME CALL YOU SWEETHEART   0.1084970523745949
   Private Screening   0.1039168263723954
   The Undertaker's Widow   0.1036851050621961
   The Laws of Our Fathers   0.1034970870559240
   Firestorm (Johansen, Iris)   0.0956006972553726
------------- END --------------
```

The item that the user selected for recommending similar items deals with court-trial genre.
The CBR recommendations also seem to recommend titles relating to court genres like “Protect and Defend”,”The Final Judgement”,”The Laws Of Our Fathers”.

```
------------- CFR Items --------------
   The Beach House   54
   Tara Road   2436
   E Is for Evidence: A Kinsey Millhone Mystery (Kinsey Millhone Mysteries (Paperback))   2807
   Executive Orders (Jack Ryan Novels)   7552
   Private Screening   99
   Ruthless.Com (Tom Clancy's Power Plays (Paperback))   5645
   Legacy of Silence   8177
   RUSSIA HOUSE, THE   251
   The Scorpio Illusion   7441
   GEMINI CONTENDERS   10237
------------- END --------------

------------- Hybrid Items --------------
   Executive Orders (Jack Ryan Novels)   7552
   Private Screening   99
   Legacy of Silence   8177
   GEMINI CONTENDERS   10237
   Vanished   5880
   Secrecy   3216
   Escape the Night   3107
   The House of the Spirits   569
   Beloved   2865
   The Best Laid Plans   2585
------------- END --------------
```

It can be noticed that hybrid recommendations seem to retrieve titles that are indirectly related to court trial and biography genres whereas CBR and CFR are only able to capture very direct mappings.

4.3 Adding ratings to items by users

The user with user id 9 as of now has a profile incling to religious category.
Let us try adding ratings to science fiction titles like Star Wars by user id 9.This act intends to make the user more of a Star Wars person.

```
Select any one
1.Get user recommendations
2.Get similar items
3.Print user Info
4.Get item info
5.Add Rating to a Item
6.Search for item
7.Exit

5
Enter userId,bookId,rating seperated by spaces
9 2231 8
CBR: Before Rating modified {'Location': 'germantown, tennessee, usa', 'Age': nan, 'ratings': {3: 0, 4: 6, 5: 0}}
CBR: Added Rating {'Location': 'germantown, tennessee, usa', 'Age': nan, 'ratings': {3: 0, 4: 6, 5: 0, 2231: 8}}

Select any one
1.Get user recommendations
2.Get similar items
3.Print user Info
4.Get item info
5.Add Rating to a Item
6.Search for item
7.Exit

5
Enter userId,bookId,rating seperated by spaces
9 1835 9
CBR: Before Rating modified {'Location': 'germantown, tennessee, usa', 'Age': nan, 'ratings': {3: 0, 4: 6, 5: 0, 2231: 8}}
CBR: Added Rating {'Location': 'germantown, tennessee, usa', 'Age': nan, 'ratings': {3: 0, 4: 6, 5: 0, 2231: 8, 1835: 9}}
```

Now trying to get recommendations for user with user-id 9 will result as in next shown:

```
1
Enter UserId
9
User Detail
id:- 9
location:- germantown, tennessee, usa
age:- nan

-----------Previously Used Items----------
The Testament 3
Beloved (Plume Contemporary Fiction) 4
Our Dumb Century: The Onion Presents 100 Years of Headlines from America'
Heir to the Empire (Star Wars: The Thrawn Trilogy, Vol. 1) 2231
Jedi Search (Star Wars: The Jedi Academy Trilogy, Vol. 1) 1835
Insights regarding recommendations

Focal Points of CBR recommendation are-
Fiction,notFound,Juvenile Fiction,John Grisham,Toni Morrison,

Focal Points of CFR recommendation are-
Fiction,notFound,Stephen King,Juvenile Fiction,James Patterson,

Focal Points of hybrid recommendation are-
Fiction,Juvenile Fiction,notFound,Frank Herbert,Elizabeth Berg,
Convergence score= 0.900950081682385
------------- CBR Users --------------
    The Testament   482
    Star Wars: Shadows of the Empire   2433
    Dark Apprentice (Star Wars: The Jedi Academy Trilogy, Vol. 2)   11918
    Return of the Jedi (Star Wars)   667
    Champions of the Force (Star Wars: The Jedi Academy Trilogy, Vol. 3)   11920
    Beloved: A Novel (Plume Contemporary Fiction)   9928
    Empire Strikes Back Wars   1308
    Star Wars Episode 1: The Phantom Menace   12769
    The Last Command (Star Wars: The Thrawn Trilogy, Vol. 3)   4821
    Dark Force Rising (Star Wars: The Thrawn Trilogy, Vol. 2)   3226
------------- END --------------
```

Now users get recommendations tending to fiction genre especially movies from the Star Wars franchise that the user has not yet used.

```
-------------- CFR Users ---------------
    Piercing the Darkness   9205
    Jack &amp; Jill (Alex Cross Novels)   589
    Tuck Everlasting   5214
    Inferno (Mentor)   3357
    Dark Tide I: Onslaught (Star Wars: The New Jedi Order, Book 2
    The Talisman   4333
    Guilty Pleasures (Anita Blake Vampire Hunter (Paperback))   2
    The Crucible: A Play in Four Acts (Penguin Plays)   9521
    The Handmaid's Tale : A Novel   2459
    Speaker for the Dead (Ender Wiggins Saga (Paperback))   3827
-------------- END ---------------

-------------- Hybrid Users ---------------
    Pop Goes the Weasel   2322
    The Summoning God (The Anasazi Mysteries, Book 2)   12004
    Saving Faith   4972
    The Best Laid Plans   2585
    Secret History : A Novel   4605
    Dune (Remembering Tomorrow)   3618
    Circle of Three: A Novel   793
    The Last Suppers   3944
    The Pull of the Moon   8616
    Star Wars: The Truce at Bakura (Star Wars (Random House Paper
-------------- END ---------------
```

Users get recommendations tending to fiction and some other unexplored genres that relate to Action using Hybrid Algorithm.

4.4 Searching for item

Searching for ‘star’ by user with user id - 9

```
6
Enter userId and search term
9 star
[(10340, 8, 7.150621439177368), (1834, 4, 20.80748515720196), (3955, 4, 17.39229643305076), (
------------- Search results --------------
    Star Wars: Tales from the Mos Eisley Cantina (Star Wars (Random House Paperback))   10340
    Star Wars: The Truce at Bakura (Star Wars (Random House Paperback))   1834
    Couplehood   3955
    Star Wars: Shadows of the Empire   2433
    Pleading Guilty   11078
    Sweet Revenge   9233
    Coyote Waits (Joe Leaphorn/Jim Chee Novels)   68
    Secrets of the Morning (Cutler)   5300
    A Man in Full   164
    Pleading Guilty   1782
------------- END --------------
```

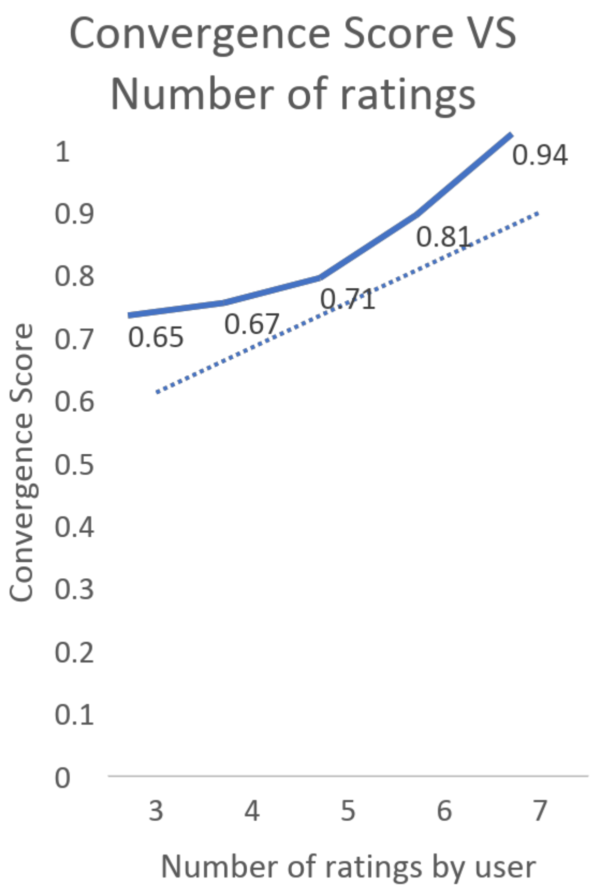
Since the user has previously rated items relating to the Star Wars franchise, the search term star results in search results including Star Wars titles.

Now let’s search the same term ‘star’ using the user id 56

```
6
Enter userId and search term
56 star
[(9904, 4, 8.42001397021088), (5300, 4, 7.020579754982894), (3955, 4, 2.91165182[illegible])
------------- Search results --------------
    Empty Promises   9904
    Secrets of the Morning (Cutler)   5300
    Couplehood   3955
    Enter Whining   9678
    The Queen of the Damned (Vampire Chronicles (Paperback))   1248
    Dragonfly in Amber   577
    Pop Goes the Weasel   4542
    The Vampire Lestat (Vampire Chronicles, Book II)   317
    The Eight   2852
    Harmful Intent   3843
------------- END --------------
```

Even though the same search term is used, different search results are shown.The user with user-id 56 is more of a Vampire and fictional stories person and so the star term results in titles tending to vampire category titles that relate to star entity.

4.5 Performance Analysis



Graph1: Convergence score trend

Convergence score is a quantitative measure denoting the factor of convergence of content based and collaborative filtering recommendations into hybrid recommendations.
It ranges from 0 to 1. Higher the score, higher the accuracy of recommendations.
It is observed that as the no. of previously used items by a user increases the convergence score tends to approach 1 and the algorithm seems to yield accurate recommendations.

5. CONCLUSION

Generalized approach to recommendations resulted in the system working well with wide range of domains and datasets. The system proposed was able to mix functionality of two popular recommendation algorithms with complementing features thereby making the recommendations work better even with less data and reduced the response time for generating recommendations considerably alongside working smoothly across different domains and datasets.
Items available in the catalog are prone to increase exponentially with more and more users and providers getting interconnected on a daily basis in large numbers. And so the energy and time to be spent on the platform to choose a item by the user gets increased thereby reducing the quality of user experience, ultimately leading to the event where the user doesn't use any item.And so there is an ever-growing need of recommendation systems that are better suitable to different domains and datasets ranging from music,movies,shopping data.There is an ever-growing need for recommendation systems that are better suitable to different domains.

6. FUTURE WORK

User-User following subsystems can be implemented thereby making use of human intelligence along with machine intelligence. Trends in the overall system can be detected thereby amplifying the quality of recommendations.
Item embeddings can be used in composition with the hybrid system for even better similarity score computation.
Computation of items and users can be done remotely on a Hadoop server thereby reducing the initial load time and increasing performance of the overall system with incoming stream of user ratings and new product registrations.
A questionnaire can be shown to the user periodically depending upon the user's changing preferences of the items to better estimate the user's present likings in the user profile generated.

REFERENCES:

[1] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl Item-Based Collaborative Filtering Recommendation - 2001.

[2] Isinkaye, F.O.; Folajimi, Y.O.; Ojokoh, B.A Isinkaye, F.O.; Folajimi, Y.O.; Ojokoh, B.A Recommendation systems: Principles, methods and evaluation- -2015.

[3] Greg Linden, Brent Smith, and Jeremy York-Recommendations Item-to-Item Collaborative Filtering- 2001

[4] F. Ricci, L. Rokach, B. Shapira, Recommender Systems Handbook, Springer US, Boston, MA, 2011, Ch. Introduction to Recommender Systems Handbook, pp. 1–35.

[5] J.B. Schafer, J.A. Konstan, and J. Reidl, "E-Commerce Recommendation Applications," Data Mining and Knowledge Discovery, Kluwer Academic, 2001.

[6] K. Goldberg et al., "Eigentaste: A Constant Time Collaborative Filtering Algorithm," Information Retrieval J., vol. 4, no. 2, July 2001.

[7] Hongyan Liu, Jun He, Tingting Wang, Wenting Song, Xiaoyang Du, "Combining user preferences and user opinions for accurate recommendation", Elsevier Journal- Electronic Commerce Research and Applications 12 (2013)

[8] Adomavicius, G., and Tuzhilin, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17, 6, 2005, 734–749.

[9] Adomavicius, G., and Kwon, Y. New recommendation techniques for multi-criteria rating systems. IEEE Intelligent Systems, 22, 3, 2007, 48–55.

[10]Adomavicius, G., Manouselis, N., and Kwon, Y. Multi-criteria recommender systems. In F. Ricci, L. Rokach, and P. B. Kantor (eds.), Recommender Systems Handbook: A Complete Guide for Research Scientists and Practitioners, Springer, New York, NY, 2010.

[11] Aggarwal, C., and Zhai, C. Mining Text Data. Springer, New York, NY, 2012.

[12] Alexander, P., and Patrick, P. Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the Seventh Conference on International Language Resources and Evaluation, Valetta, Malta, 2010.

[13] Ali, K., and Van Stam, W. TIVo: Making show recommendations using a distributed collaborative filtering architecture. In Proceedings of the Tenth ACM SIGKDDInternational Conference on Knowledge Discovery and Data Mining, Seattle, WA,2004, 394–401.

[14] Benamara, F., Cesarano, C., Picariello, A., Reforgiato, D., and Subrahmanian, V.Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In Proceeding of International Conference on Weblogs and social media, March 2007, San Jose, CA, 2007,1–7.

[15] Bennett, J., and Lanningm, S. The Netflix prize. In Proceedings of 2007 KDD Cup and Workshop, San Jose, CA, 2007.

[16] Billsus, D., and Pazzani, M. J. Learning collaborative information filters. In Proceedings of the Fifteenth International Conference on Machine Learning,Morgan Kaufmann, San Francisco, CA, 1998, 46–54.

[17] Breese, J. S., Heckerman, D., and Kadie, C. Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the Fourteenth Conference onUncertainty in Artificial Intelligence, Morgan Kaufmann, San Francisco, CA,1998, 43–52.

[18] Christiane, F. WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA, 1998.

[19] Ding, X., and Liu, B. The utility of linguistic rules in opinion mining. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, Netherlands, July 23–27, ACM Press, New York, NY, 2007, 811–812.

[20] Dong, Z., and Dong, Q. HowNet Knowledge Database. 2010. Available at www.keenage.com/html/c_index.html. Gamon, M., Aue, A., Corston-Oliver, S., and Ringger, E. Pulse: Mining customer opinions from free text. In Proceedings of the 6th International Symposium on Intelligent Data Analysis, Madrid, Spain, September 8–10, 2004, Lecture Notes in Computer Science, 3646, Springer Verlag, New York, NY, 2005, 121–132.

[21] Ge, Y., Xiong, H., Tuzhilin, A., Xiao, K., Gruteser, M., and Pazzani, M. J. An energyefficient mobile recommender system. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, July 25–28, 2010, ACM Press, New York, NY, 2010, 899–907.