International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |

Recommendation System Using Machine Learning

Rohan Lohia, Sumit Bhatnagar

Student (I.T), Maharaja Agrasen Institute of Technology, New Delhi, India

Student (I.T), Maharaja Agrasen Institute of Technology, New Delhi, India

ABSTRACT: Recommendation system plays an important role in Internet world and used the most valuable application of machine learning today's. It has created the collection of many application and the growth for numerous information & implemented various algorithms. The recommendation system has made finding many things that we need in our daily life usage and are getting importance due to their significance in making decision ad providing detailed information about the required product or a items which a user want to see. The recommendation system have been using machine learning algorithms which is suitable for developing a system to the user and approaches a new algorithms.
This paper represents the overview of approaches and techniques generated in recommendation system. There are some methods to evaluate the recommendation system : Collaborative filtering, Content-based filtering, Matrix-factorization, Hybrid methods and knowledge based system. This paper classifies collaborative filtering in two ways: Memory based and Model based recommendation. Hybrid system combine multiple recommendation techniques to enhance the performance of a system.

KEYWORDS: Collaborative Filtering, Content-Based Filtering, Matrix-factorization, Hybrid and Knowledge-Based.

I.INTRODUCTION

Recommendation System is part of daily life where people are making decision of their personal interest. A recommendation system is a machine learning system, we can train our recommendation system by using the training data. It is the subclass of information filtering to predict preferences to the items used by users. It is basically the relation between items and users. Once it trained, we can ask it for make prediction about a new user might give some rate to the items.
The recommendation system might give some interest regarding the items to the user and there are some of the examples which make the system more interesting and systematic:

1. **Content-** (Quora has recommendation for the similar article.)
2. **Things-**(Amazon has recoomendation for the similar things to buy.)
3. **Music-** (It's an example of content-based recommendation.)
4. **People-** (It's an example of Online dating site.)
5. **Search Result-**(Informationretrievel, Web pages.)

One of the important example which is mentioned in this system are movie in movilens, and so on. In this paper different approaches and techniques are used to compare each techniques and methods in proper manner to provide future recommendations.
Here, is the intro of some of the methods which is mentioned in this project:-

1. **Content-Based Filtering:**Content-based filtering (CBF) is one of the most widely used and researched recommendation approaches. CBF is the user modeling process, in which the interests of users from the items that users interacted with 'Items' for instance webpages. It is a technique where individual user profile are taken into account. It enchances the user's interest and predicts whether the user would be inteterest at any particular interested like movies, restaurant or any other things.
2. **Collaborative Filtering:** Collaborative filtering (CF) is mainly based upon the data of the user available to the system. The data available helps to build the user profile and the data available about the item is usd to make the item profile. It is a techniques for predicting unknown preference of people by using known from many users.

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |

3. **Matrix-factorization methods:**It is the most successful latent factor models on matrix factorization. Matrix factorization implies both items and users by vectors of factors inferred from item rating patterns. The most convenient data is high-quality explicit feedback, which includes input by user regarding their interest in products.
4. **Hybrid methods:** Hybrid system associates multiple recommendation approaches together for the better results. It defines that if the given tw recommendation approaches X and Y, a hybrid system XY combines the two in one way.

II.LITERATURE REVIEW

A recommendation system based on collaborative filtering, content-based filtering, matrix-factorization and hybrid approaches. Collaborative filtering takes the data from all the users to shown in product for the user and based on that generates recommendations. Content-based filtering is for individual user profile to taken into account. This system analysis the recommendation techniques as there are many techniques that approaches the process of user and the items which is used in this system.

The evaluation methods, metrices and datasets used; the number of participants in the user studies and several other factors to judge the evaluations in this system. The evaluation methods like user studies, online and offline.
User Studies:- It considered the optimal evaluation and the ouput of this evaluation would be depends on the 'perceived relevance' or 'global satisfaction' which users are asked in the recommendation system. The number of required participants to receive statistically significant results, depends on the number of approaches being evaluated, the number of recommendations being displayed and the variations in the resuts.
Online:- Online evaluation were first used by the online advertising and e-commerce fields. It measures the acceptance rates of recommendation in real-world system. Acceptance rates are basically measured by click-through rates (CTR). The clicked recommendation to display recommendation would be measured in online evaluation.
Offline:- Offline evaluation is basically measure the accuracy of a recommender system which is based on a ground-truth. Or to measure accuracy, we need to often precision at position: n(P@n) which is used to express how many items of the ground-truth are recommended within the top-n recommendations.

III.CONTENT

This section explains the machine learning process and its method in various types of recommendation system. It also analysis the related applications and algorithms of recommender techniques in recommendation system.
The methods which is used in this projects are:

1. Collaborative filtering
2. Content-Based Filtering
3. Matrix-Factorization
4. Hybrid

Collaborative Filtering

In this section, we are discussing about the user and the item which are stored in the historical data of the user available to the system. The heart of collaborative filtering is the ability to find people similar to you or items similar which you have liked.
The first step in collaborative filtering is measuring the similarity between things or people. Sparsity also introduce some computional challenges, so under the hood, we end up using structures like sparse arrays that avoid storing all the empty space in this recommender.
The collaborative filtering uses two based filter which are:

- User-based Collaborative Filtering
- Item-based Collaborative Filtering

User-based Collaborative Filtering: It is a technique used to predict the items that a user might rating given to that items by the other users who have similar data with the target user.
User-> item rating matrix
User-> user similarity matrix
Looks up similar users
Candidate generation

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |

Candidate scoring
Candidate filtering

Item-based Collaborative Filtering: It is basically a technique which is used for similarity between items could be better than similarities in peoples. One is that, item tend to be of a more permanent nature than people. Another important advantage to building item similarity is that, you usually have a far fewer item deal with than people.
There are some of the equation in collaborative filtering which is used in recommendation system.

Cosine Equation:

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

Pearson Similarity (item-based):

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[\, n\sum x^2 - (\sum x)^2 \,][\, n\sum y^2 - (\sum y)^2 \,]}}$$

Mean Squared Difference:

MSDSim (x,y) = 1/MSD (x,y) +1

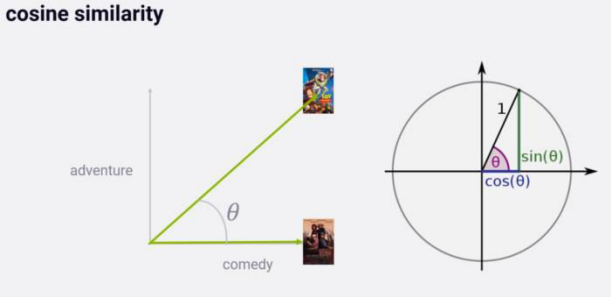
Content-Based Filtering:
Content-based filtering is the one who widely used and researched recommendation approaches. The interests of users are infer from the user that interacted with Items and items are usually textual for instace webpages. It represents the comparison between the content contained in the items with the content of items of user's interest.

Example of movie attributes: The movielens dataset doesn't give us much to work with but one things it does tell us which movie genres each movie belongs to. For every movie, we are given a list of genres like science friction, romance, etc. So, that's a pretty simple example of doing content-based filtering, we'll just recommend movies that have similar genres and similar years to the movies each user liked.

Cosine Similarity Graph:

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |



The Cosine of 45 degree is about 0.7, so we could say that the cosine similarity score between toy story and grumpy old men is 0.7.

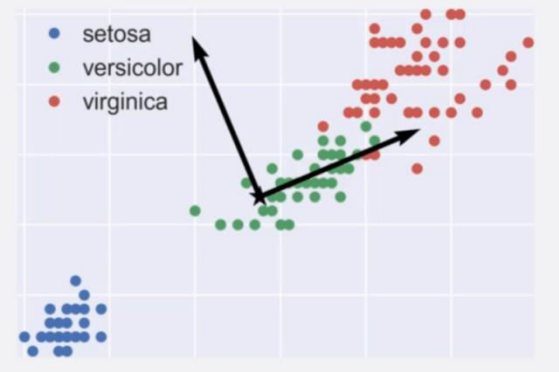
Matrix-Factorization Methods:

It is the most successful method which is based on matrix factorization. In this methods, we generally factorize the items and users by vectors from item rating patterns. High correspondence between item and user factors leads to a recommendation. These methods have combining good scalability with predictive accuracy.

Principal component analysis (PCA) :- It is usually described as a dimensional reduction problem. That's we take data that exists in many dimension, like all of the movies a user might rate into a smaller set of dimension that can accurately describe a movie, such as its genres.

$$R = M\Sigma U^T$$

We describe our training data in terms of smaller matrices that are factors of the ratings we want to predict. There is also that sigma matrix in the middle here that we haven't talked about that we need, it's just a simple diagonal matrix that only serves to scale the values we end up with into the proper scale.



Hybrid Methods:

Hybrid system are the combination systems, that is merge the multiple recommendation approaches together for better result. The basic definition of Hybrid recommender states that ' the two recommendation approaches X and Y combine to form XY in such a way that the limitations of one are overcome by the other'. One of the most popular examples of a Hybrid system is NetFlix. In the hybrid recommender system we make use of all three created recommender system to optimize our algorithm is such a way that it recommends more accurate items or movies to the user.

In this we choose SVD++ algorithm as a base then apply it to all the k-neighbours and then arrange all the recommendation in descending order of the ratings of the movies and choose top-n recommendations to recommend to the user.

IV.METHODOLOGY

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |

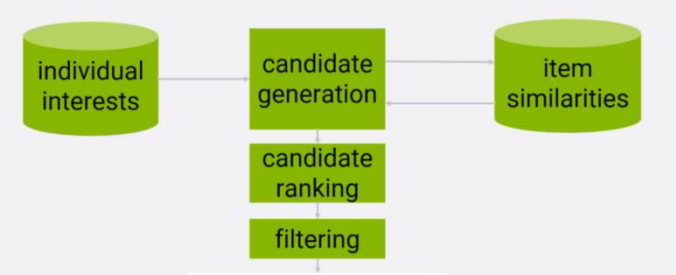
"Various Types of Recoomendation System" has been created using a combination of some basic equations like MAE (Mean Absolute Error), RMSE (Root Mean Square Error), KNN (Knowlegde nearest neighbor), SVD (Singular Value Decomposition), ARHR (Average Reciprocal Hit Rate).

V.IMPLEMENTATION

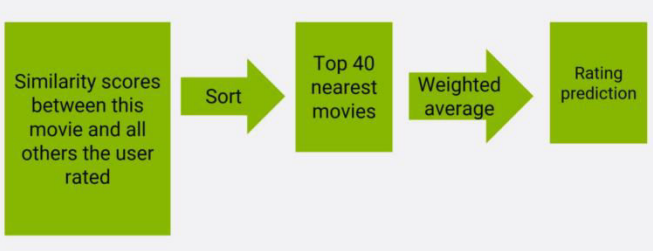
There are many algorithms to implement on this project. Few of them is important in this project because the implementation in project is basically how it works and exhibit on the applications. The algorithms are : Top-n recommender, KNN recommender, SVD Matrix recommender, SVD++ Hybrid recommender and so on. Some are the equations which is used in it like: MAE, RMSE, HitRate, Coverage, Diversity, Novelty, etc.. but we are talking about the main algo which is develop the recommendation system to the next level.

Top-N Recommender:

Top-N recommender system are everywhere from online portals to video portals. They provide users with a ranked list of N items, which will likely be interested in order to views and get the products.
One of the example is amazon.com where the amazon sites has lots of items list which the user want to buy it or not. For the recommender algo, there is a list of items which the items rating is good, average or bad. But it's all up to the user which product the user want to buy.



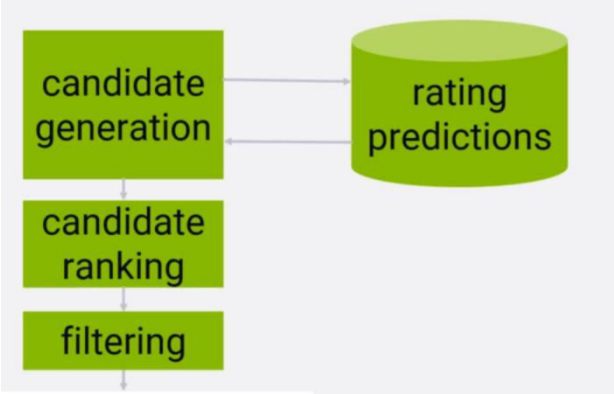
KNN Recommender:
KNN (Knowlegde Nearest Neighbor) recommender system is a machine learning algorithm to find the clusters of similar user based on ratings and predicting using the average rating. It works by finding the distance between a query and all the data which is closed to the query then votes for the most frequent label or average the label.



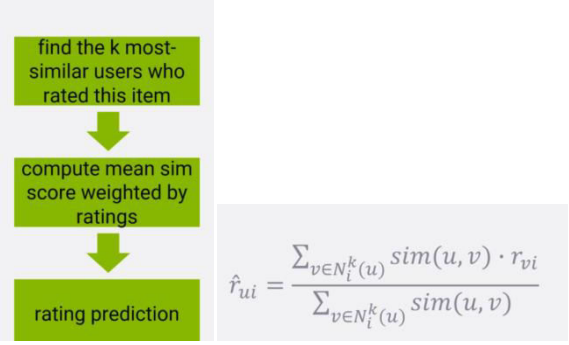
It's just selecting some number of things that are close to the things you're interested in, that's it is neighbors and predicting something about that item based on the properties of its neighbors is known as K-nearest neighbors.

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

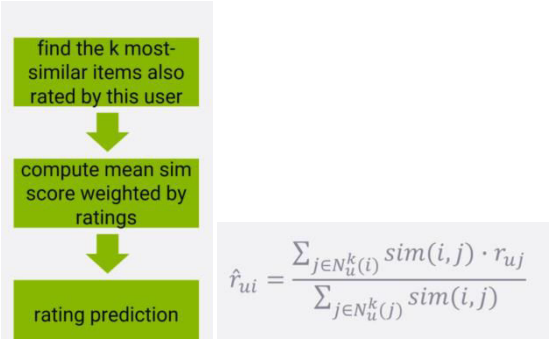
| Volume 8, Issue 1, January 2021 |



User-based KNN:-



Item-based KNN:-



SVD Matrix Recommender:

It is a real or complex matrix that generalizes the eigen decomposition of a square normal matrix to any m*n. All SVD is doing running PCA on both the users and the items, and giving us back the matrices we need that are factors of the rating matrix we want.The Singular Value Decomposition (SVD) , a method from linear algebra that has been generally used as a dimensionality reduction technique in machine learning. SVD is a matrix factorization technique which reduces number of features of a dataset by reducing the space dimension from N-dimension to K-dimension (where K<N). In the context of the recommender system, the SVD is used as a collaborative filtering techniques. It uses a matrix structure where each row represents a user, ad each column represents an item. The elements of this matrix are the ratings that are given to items by users.

There a lots of machine learning techniques that can do that sort of things such as stochastic gradient descent or SGD for shoot. Basically it just keeps iterating at some given learning rate until it arrives at a minimum error values.

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 8, Issue 1, January 2021 |

$$R = M\Sigma U^T$$

$$R_{Bob,Empire\ Strikes\ Back} = M_{Bob} \cdot U^T_{Empire\ Strikes\ Back}$$

When we say, we are doing SVD recommendations, it is not really SVD because you can't do real SVD with missing data. It's an SVD-inspired algorithm that was invented for the Netflix prize, but it's not really pure SVD.
The important point on this: you can think of all ratings for a set of users & items as a matrix R, and that matrix R can be factored into smaller matrices that describe general categories of users and items that can be multiplied together.

SVD++ Hybrid Recommender:
In this hybid recommender system, we make use of all three created recommender system to optimize our algorithm is such a way that it recommends more accurate items or movies to the user. In this we choose SVD++ algorithm as a base then apply it all the K-neighbous and then arrange all the recommendation in descending order of the ratings of the movies and choose top-n recommendation to recommend to the user.
Hybrid system are the combinational system that merge the two or more recommendation technique in such a way that the limitations of the recommendation approaches are beat by the other. Netflix is the most popular hybrid recommender based on collaborative and content-based approach.The basic definitionof Hybrid recommender states that, "Given two recommendation approachesX and Y, a Hybrid system XY combines the two in such a way that thelimitations of one are overcome by the other."

VI.CONCLUSION

In this paper, we studied and compare the various recommendation approaches developed. The focus of this study is KNN recommender, Top-N recommender, SVD and Hybrid recommender system. Several recommendation system have been expect are based on collaborative filtering, content-based filtering, matrix-factorization and hybrid method. Most of them have been able to resolve the problem while providing improved recommendation. However, due to information, it is required to work on this research area to explore and provide new methods that can provide recommendation in a wide range of application while considering the quality and privacy aspects. These system work on individual user's ratings, hence limiting the choice to explore more. While our recommendation system which is based on many methods approach computes the connection between different users and the product/items which upon their ratings, like movie in movielens. It is a web application that allows users to rate movies as well as recommends many other things also. Thus, the current recommendation system needs enchance for present and future requirements of better recommendation qualities.

REFERENCES

[1] C. Pan, W. Li Research Paper recommendation with topic analysis In Computer Design and Application IEEE, 4 (2010).
[2] J.A. Konstan, J. Riedl Recommendation systems: from algorithm to user experience User Model User-Adapt Interact, 22 (2012).
[3] Pazzani, M. A Framework for Collaborative, Content-based and Demographic Filtering. Department Of Information and Computer Science. University Of California, Irvine. 1999.
[4] Hill, W., Stead, L., Rosenstein, M., and Furnas, G. Recommending and Evaluating Choices in a Virtual Community of Use. In Proceedings of CHI'95.
[5] G. Adomavicius, A. Tuzhilin Toward the next generation of recommender system. A survey of the state-of-the-art and possible extensions IEEE Trans Knowl Data Eng.
[6] R. Burke Hybrid recommender systems P.Brusilovsky, A. Kobsa, W. Nejdl.
[7] J.L Herlocker, J.A Konstan, L.G Terveen, J.T. Riedl Evaluating collaborative filtering recommender system,22.
[8] https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning
[9] https://www.udemy.com/course/building-recommender-systems-with-machine-learning-and-ai/
[10]https://www.geeksforgeeks.org/python-implementation-of-movie-recommender-system/