Product Recommendation using Machine Learning Algorithm - A Better Appoarch

Prof. Shivganga Gavhane
of (SPPU) Computer Engineering,
Dr. D. Y. Patil Institute of Engineering Management
& Research, Akurdi ,
Pune, India

Jayesh Patil
of (SPPU) Computer Engineering,
Dr. D. Y .Patil Institute of Engineering Management
& Research, Akurdi ,
Pune, India.

Harshal Kadwe
of (SPPU) Computer Engineering,
Dr. D. Y. Patil Institute of Engineering Management
& Research, Akurdi ,
Pune, India

Prajwal Thakhre
of (SPPU) Computer Engineering,
Dr. D. Y .Patil Institute of Engineering Management
& Research, Akurdi ,
Pune, India.

Sushovan Manna
of (SPPU) Computer Engineering,
Dr. D. Y. Patil Institute of Engineering Management
& Research, Akurdi ,
Pune, India

Abstract— **Being in the era of advance technology, people are more open minded and rely heavily on modern applications for buying accessories, watching movies and stuffs related to their daily needs. Due to this perpetual rise in online shopping demand and watching movies , the organizations are relying on machine learning based technologies which are helping them in seeking the actual targeted users with less efforts as compared to earlier methods of advertisement. We are developing this Technology which helps us to understand the requirements and gives recommendation for the product searched by the user by comparing their previous history. This model compares various machine learning algorithms for recommendation of various product buying pattern by users and gives more accurate result related to search. It also carries all buying history of user, so it is easy for the model to predict the best option to buy and recommend the product.**

Keywords:- Machine learning, recommendation systems, Supervised , Unsupervised Learning, K-means, Collaborative Filtering.

I.   INTRODUCTION

In the modern world, the Web has become omnipresent. Used to shop, watch movies, listen to music, or communicate with friends. A huge number of people's activities-Internet users-open up the possibility of collecting information from thousands, millions, and even billions of people. It is an invaluable opportunity for those who are looking for this data. It helps them to describe collective intelligence within the society being observed-their attitudes, interests, and world view. These are valuable insights for the marketing market, for example the marketplace. They show whether their sales tactics are working for a given social group, and sometimes they help to find the best way to reach a given target group. The Internet, in turn, enables you to monitor network users ' various activities without affecting their intentions.

Information are at the disposal from different social groups around the world. The knowledge base assessment is the biggest problem confronted by scientists and application developers based on data mining. People cannot effectively perform complicated and tedious calculations on it because of its size. The machine learning problem comes with support. Statistical methods can be one of the fundamental techniques in machine learning: regression and study of association. More advanced methods are problems associated with learning neural networks or fuzzy logic. The designer creates a recommendation algorithm and the computer determines the conclusion related to the properties of this set on its basis, acting on a given set of data. There are great opportunities for such systems. Recommendation systems are growing in popularity. The development of the research on their subject has somehow forced now a days to improve the efficiency of recommendation system.

II.   MACHINE LEARNING

Machine Learning is widely known for use of algorithms and technique to develop recommendation system now a days. Machine learning is application that provides systems with the ability to learn and improve automatically from experience of explicit coding. Machine learning focusers on creating computer programs that are able to access information and use it to learn on their own. This learning process can be divided into two parts known as Supervised and unsupervised learning. According to their model used in machine learning it can give higher precision output to the users. Machine learning can be classified by the following four major categories.

A. Classification - This addresses issues relevant to class assignment to each of the products being evaluated. The issue of image recognition can be an instance of such a task.

B. Regression - This is an estimate of the object's actual value. A good example of such a function may be an effort on the basis of economic factors to calculate the price of securities.

C. Ranking - It is liable for a particular standard for sorting objects. The most common function of this sort is the returning websites of the search engine that satisfies the user's request.

D. Clustering - This addresses issues relating to the separation of items into some homogeneous categories. In social networks, these algorithms can be used.

III. TYPES OF MACHINE LEARNING ALGORITHM

The following types of machine learning algorithms can be differentiated due to different situations of availability of training data, test data and analysis of teaching methods.

A) Supervised learning - This algorithm collects training data in the case of instructor education in which the expected performance quality from the input data is known. It's one of the most popular methods for education.

B) Unsupervised learning - In contrast to the teaching method with the teacher, the algorithm receives training data that does not take into account which output value should be obtained from the input data. In this scenario, the assessment of the extent to which the algorithm has mastered the training data can be troublesome.

C) Semi-supervised learning - Training data, while partially supervised, consist of samples having the expected initial value as well as samples that do not have it. This method is popular when the input data is easy to obtain, but the output data is much more expensive.

D) Reinforcement learning - The training and testing phases are combined in a reinforcement approach. The learned algorithm, by interacting with the environment, collects data. He receives, depending on the action taken, a reward or penalty. The purpose of this method is to maximize the reward for the learned algorithm

IV. WORKING WITH ALGORITHMS

1. K-means

The k-mean is a method for grouping which is not parametric. According to their similarity, it distributes the objects into k clusters. In this article, using the Euclidean distance[1], this similarity is calculated. We took graded and unrated films as binary to measure the Euclidean distance. Each cluster has a centroid that is the mean of all the cluster items. Both entities in a cluster travel in the direction of the centroid and in each iteration the centroid is modified. The loop continues on until a stage of equilibrium occurs when the centroid stops changing. By following this approach, we are reducing the search space resulting in a reduced complexity of computations[2]. These calculations are performed off-line, which in terms of time complexity helps the classification to be effective.

Algorithm k-mean clustering

Input: Here we are taking product as movies. users X, films m, ranking r, μ suggested amount of movies, k price.

Output: Suggested film R.

1. Start Algorithm.
2. Choose centroids by random selection.
3. Calculate Euclidean distance (E_d) from k centroids for X.
4. Allocate X to the E_d based k^{th} cluster.
5. Update centroid with (summation(k_i) of 1 to p)/p for each cluster, where p is the number of members in k_i cluster
6. Repeat step 3 to step 5 until the centroid (t+1) is in the centroid.
7. Choose movies shown s, unknown films s.
8. Select films seen s, unseen films s.
9. Search for similarity (Si)
10. Select the user with the highest Si.
11. Select the lowest Si and the smallest Si and i^{th} client.
12. Calculate weight W (me') where the recommendations for e / m'
13. Return top μ weight.
14. End for loop.
15. End.

2. Collaborative Filtering

There can be many users who must be having the same pattern of rating an item as the user intended. This similar pattern of their ratings with the user guides the Collaborative Filtering[3][4][5].The notion behind the Collaborative Filtering is the recommendation of an item based on the preferences of like-minded users.

Algorithm 2. Collaborative Filtering

Input: Here we are taking product as movies. users X, movies m, rating r, Number of movies to be recommended(μ).

Output: Recommended movies R.

Algorithm Collaborative Filtering

1. For all users do
2. Select seen movies s, unseen movies s'
3. Find similarity (Si) w.r.t s, where i = 1 to n.
4. Select highest Si user
5. Select m' ∈ s of user obtained in step 4 and s' of ith user.
6. Calculate weight $W(m_{e'})$ where e ∈ m'
7. Return top μ weight recommendations.
8. End for

In this algorithm, the notations used have the following meaning : Si represents common movies between user i and other users. weight(me') = rating of particular $movie_e$ / max rating.

V. RESULTS WITH WORKING ALGORITHMS

We are now illustrating the analysis of the experiments carried out and comparing all the state-of - the-art methods described above. For testing their precision, we used the 10 K, 50 K and 100 K MovieLens database. The sparsity of the data set varies. The 100 K MovieLens database, for instance,

has 100 K scores, 943 participants, and 19 different genres of 1682 films. Depending on precise tests, the application of these algorithms is shown. We transform 30 percent of the user's watched movies into unseen movies for each trial user and implement the above mentioned algorithms. Of the total number recommendations (T), the ones which are also present in the converted movies are the correct recommendations(tc).
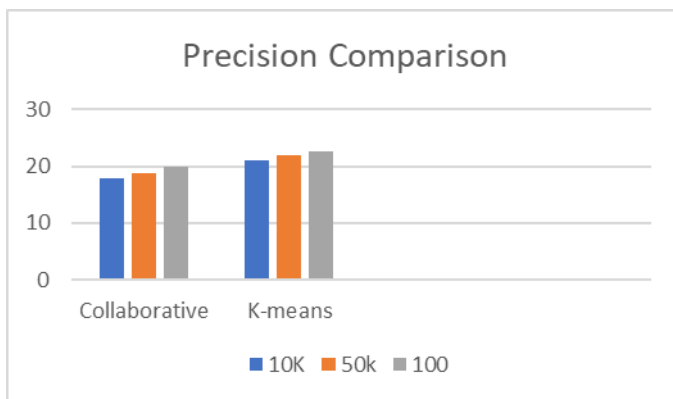
Precision = $(\Sigma tc / \Sigma T) * 100$

For all the experiments, we are taking value of $\mu = 5$ and value

of $k = 10$.

Algorithm/Size	10K	50K	100K
Collaborative	17.97	18.69	19.95
K-means	21.05	21.93	22.67

Table 1. Precision of different algorithms

VI. CONCLUSION

With regard to their precision rates, all the algorithms described in this paper are compared. Throughout different versions of the MovieLens database, this comprehensive analysis shows the strength and vulnerability of each of them. Thus we compare two supervised and unsupervised algorithm and our results have shown and this paper is compatible with K-means giving the best accuracy out of all these methods. So to develop our product recommendation system we can also use K-means algorithm having more precision value.

VII. REFERENCES

[1] Prasad M. and Singh A. A Novel Hybrid AntColony Optimization Approach to Terminal Assignment Problem.ACM, AICTC '16,Article 29 August 12-13,2016.

[2] Campos P.G., Bellogín A., Díez F., and Chavarriaga J.E. Simple Time-Biased KNN-based recommendations, ACM, 978-1-4503-0258-6, 2010.

[3] Tewari A.S., Kumar A., and Barman A.G.,. Book Recommendation System Based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining, IEEE, 9781-4799-2572-8, pp 500-503, 2014.

[4] Wanaskar U.H., Vij S.R., and Mukhopadhyay D. A Hybrid Web-Recommendation System Based on the Improved Association Rule Mining Algorithm, Journal of Software Engineering and Applications, 6, pp 396-404 2013

[5] Laishram A., Sahu S.P., Padmanabhan V., and Udgata S. K. Collaborative Filtering, Matrix Factorization and Population Based Search. The Nexus Unveiled. ICONIP ,Part III, LNCS 9949, pp. 352–361, 2016.