

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Collaborative Filtering Algorithm over Ecommerce Website Based on User Interest

Nandimalla Yaswanth Reddy¹, Kanchi Sai Srinidhi², Enishetty Manish³

B. Tech, Dept. of CSE, GPREC, Kurnool, Andhra Pradesh, India¹

B. Tech, Dept. of ECE, GNITS, Hyderabad, Telangana, India²

B. Tech, Dept. of CSE, CBIT, Hyderabad, Telangana, India³

ABSTRACT: Collaborative filtering algorithm is one of widely used approaches in daily life, so how to improve the quality and efficiency of collaborative filtering algorithm is an essential problem. Usually, some traditional algorithm focuses on the user rating, while they don't take the user rating differences and user interest into account. However, users who have little rating difference or have a similar interest may be highly similar. In this paper, a collaborative filtering algorithm based on scoring difference and user interest is proposed. Firstly, a rating difference factor is added to the traditional collaborative filtering algorithm, where the most appropriate factor can be obtained by experiments. Secondly, calculate the user's interest by combining the attributes of the items, then further calculate the similarity of personal interest between users. Finally, the user rating differences and interest similarity are weighted to get final item recommendation and score forecast. The experimental results on data set show that the proposed algorithm decreases both Mean Absolute Error and Root Mean Squared Error, and improves the accuracy of the proposed algorithm.

KEYWORDS: collaborative filtering, user similarity, rating differences, user interest

I. INTRODUCTION

With the rapid improvement of Internet technology, cloud computing and smart phone, statistics overload has come to be a chief issue in existence, which make customers infrequently deal with a extensive variety of critical data efficaciously [1]. However, the advice machine has been proven to assist customers clear out undesirable records and make affordable alternatives. Therefore, many large shopping web sites use diverse advice algorithms to offer convenience that allows you to appeal to clients. The examples of advice algorithms are content material-based advice, recommendation of affiliation regulations, Collaborative filtering hints, and many others. [2] Among the above instance, the collaborative filtering set of rules is one of the most broadly used and the most successful algorithms in the recommended subject via a long way [3]. It especially includes object based totally, consumer-based and version-based collaborative filtering algorithms [4]. The idea of person-based collaborative filtering algorithms is to find out the users with similar ratings to the target customers in keeping with historical information and its rating data, and to take them as buddies, then arrange them right into a looked after directory in line with their favourite gadgets. The middle of collaborative filtering algorithm is to calculate the similarity between customers [5], [6]. So Increasing accuracy for the similarity calculation leads to a more powerful and efficient recommendation device. Recently, many scholars have proposed to improve the algorithm primarily based on different views: Kaleli [7] introduced a unique entropy-based totally neighbor choice approach, which assigned a degree of uncertainty for each user, and the method solves the optimization hassle of amassing the most similar entities with a minimum entropy difference inside a neighborhood. Wu et al. [8] proposed the trapezoidal fuzzy scoring version to calculate the similarity between users in the model, which optimized the statistics sparse trouble and reduced the walking time. P Pirasteh et al. [9] proposed a collaborative filtering algorithm that combines consumer sports and content to correctly optimize information sparsity. G Pitsilis used person rankings to calculate a hypothetical trust between users and improves the performance of conventional collaborative filtering algorithms [10]. Considered numerous factors of accept as true with, proposed a agree with version of agree with recommendation, and optimized the recommended bloodless start problem. Considered the effect

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

of venture popularity on advice, and included popularity into the calculation of conventional similarity to enhance the accuracy of advice. To a positive extent, those researches alleviate the problem of sparsity and bloodless start of records to the set of rules, improve the overall performance of the advocated system. However, while calculating the person similarity, the above algorithms don't take person hobby under consideration, and neglect the influence of different score standards between users at the similarity calculation.

II. RELATED WORK

Existing System

With the rapid improvement of Internet technology, cloud computing and smart cellphone, records overload has become a chief problem in life, which make users infrequently take care of a extensive form of critical statistics correctly. However, the advice machine has been proven to assist users filter out unwanted facts and make affordable picks. Therefore, many large buying web sites use diverse advice algorithms to provide comfort in an effort to attract clients. The examples of recommendation algorithms are content-based totally recommendation, advice of affiliation policies, now not inclusive of the Collaborative Rules & Mining

Proposed System

In collaborative filtering suggestions, the collaborative filtering algorithm is one of the most widely used and the most successful algorithms in the recommended field by using a long way . It especially consists of item based totally, person-primarily based and model-based collaborative filtering algorithms . The idea of person-primarily based collaborative filtering algorithms is to discover the customers with comparable scores to the target customers according to ancient facts and its rating statistics, and to take them as associates, then prepare them into a looked after listing consistent with their favourite objects. The core of collaborative filtering algorithm is to calculate the similarity between users . So Increasing accuracy for the similarity calculation leads to a extra powerful and green advice machine.

III. IMPLEMENTATION

User-based nearest-neighbor collaborative filtering

Example

A database of ratings of the current user, Alice, and some other users is given:

| | Item1 | Item2 | Item3 | Item4 | Item5 |
|-------|-------|-------|-------|-------|-------|
| Alice | 5 | 3 | 4 | 4 | ? |
| User1 | 3 | 1 | 2 | 3 | 3 |
| User2 | 4 | 3 | 4 | 3 | 5 |
| User3 | 3 | 3 | 1 | 5 | 4 |
| User4 | 1 | 5 | 5 | 2 | 1 |

Table:-1 Determine whether Alice will like or dislike *Item5*, which Alice has not yet rated or seen

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Some first questions

How do we measure similarity?

How many neighbors should we consider?

How do we generate a prediction from the neighbors' ratings?

| | Item1 | Item2 | Item3 | Item4 | Item5 |
|-------|-------|-------|-------|-------|-------|
| Alice | 5 | 3 | 4 | 4 | ? |
| User1 | 3 | 1 | 2 | 3 | 3 |
| User2 | 4 | 3 | 4 | 3 | ? |
| User3 | 3 | 3 | 1 | 5 | 4 |
| User4 | 1 | 5 | 5 | 2 | 1 |

Table:-2 Measuring user similarity

• A popular similarity measure in user-based CF: **Pearson correlation**

- a, b : users

- $r_{a,p}$: rating of user a for item p

- P : set of items rated both by a and b

| | Item1 | Item2 | Item3 | Item4 | Item5 | |
|-------|-------|-------|-------|-------|-------|-------------|
| Alice | 5 | 3 | 4 | 4 | ? | |
| User1 | 3 | 1 | 2 | 3 | 3 | sim = 0,85 |
| User2 | 4 | 3 | 4 | 3 | 5 | sim = 0,00 |
| User3 | 3 | 3 | 1 | 5 | 4 | sim = 0,70 |
| User4 | 1 | 5 | 5 | 2 | 1 | sim = -0,79 |

Table :- 3 Pearson correlation

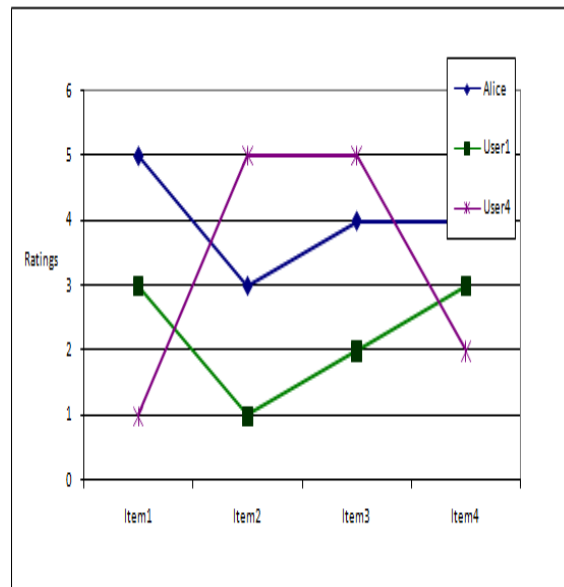
International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Takes differences in rating behavior into account



Graph:-1 Works well in usual domains, compared with alternative measures, such as cosine similarity

Making predictions

- A common prediction function:

$$pred(a,p) = \bar{r}_a + \frac{\sum_{b \in N} sim(a,b) * (r_{b,p} - \bar{r}_b)}{\sum_{b \in N} sim(a,b)}$$

- Calculate, whether the neighbors' ratings for the unseen item i are higher or lower than their average
- Combine the rating differences – use the similarity with a as a weight
- Add/subtract the neighbors' bias from the active user's average and use this as a prediction

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

| | Batman Begins | Alice in Wonderland | Dumb and Dumber | Equilibrium |
|--------|---------------|---------------------|-----------------|-------------|
| User A | 4 | ? | 3 | 5 |
| User B | ? | 5 | 4 | ? |
| User C | 5 | 4 | 2 | ? |
| User D | 2 | 4 | ? | 3 |
| User E | 3 | 4 | 5 | ? |

Table :- 3 Vales of the users

Neighborhood of 2 (A and D)

C's mean = 3.667

S(C,A) = 0.832

S(C,D) = -0.515

$$= 3.667 + \frac{(0.832) * (5 - 4) + (-0.515) * (3 - 3)}{|0.832| + |-0.515|}$$

$$= 3.667 + \frac{0.832}{1.347}$$

$$= 4.284$$

Admin

Admin is the main user of our application, after login admin can add/delete products for user's shopping. Admin will collect the user's reviews and ratings and monitor those data. For Rating prediction admin will collaborative filter model.

User

User is the end user of our application, and we can also consider e-commerce user. User can perform search products, buy products, and submit rating and review for products. This data will deliver to admin, and this data will be our dataset to our application. User can check the ratings and reviews of the products.

IV. EXPERIMENTAL RESULTS

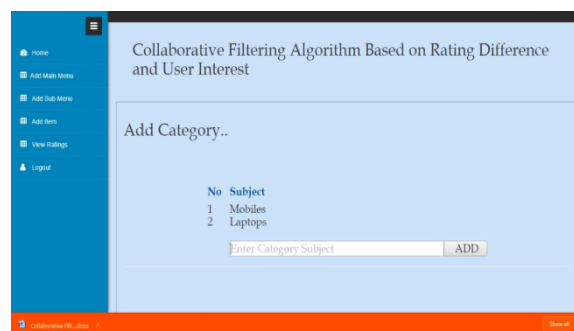


Fig: - 1 New Products into application

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

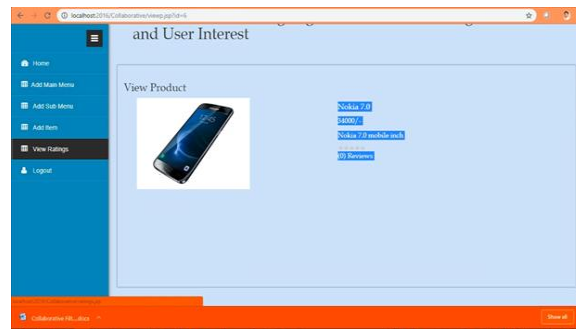


Fig: - 2 View Products

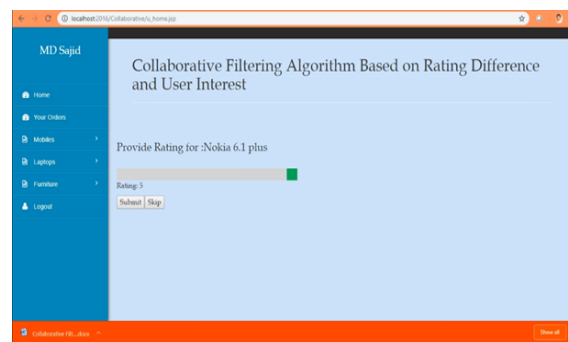


Fig: - 3 Products Rating

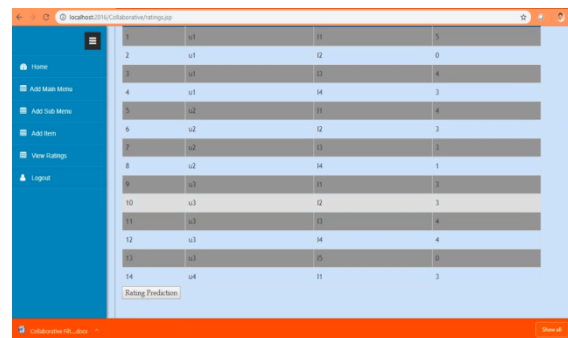


Fig: - 4 Rating Predictions

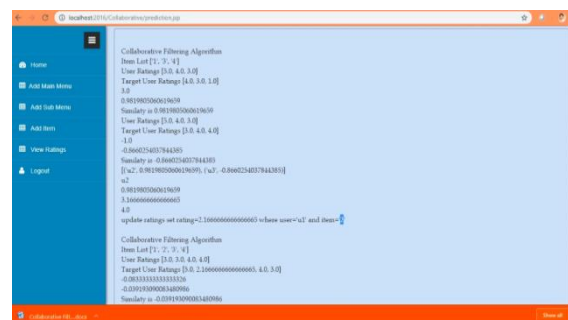


Fig: - 5 Collaborative Filtering Algorithm

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

V. CONCLUSION

Through numerous experimental schemes, this paper analyzes the overall performance of collaborative filtering advice set of rules primarily based on person score distinction and user interest. Firstly, the layout ideas and algorithm steps are added, and then score difference aspect and person hobby are taken into the conventional similarity set of rules. Secondly, numerous parameters of the algorithm are decided via numerous experiments. Finally, the progressed algorithm proposed in this paper is compared with the conventional one. The stepped forward set of rules proposed now not handiest improves the accuracy in popular situations, but also generates a better result under the condition of sparse data. The disadvantage of the experiment is that the experimental parameters might not attain the excellent results below exceptional datasets. So, the way to make the encouraged algorithm automatically regulate the parameters in line with specific records units is the following path of efforts.

REFERENCES

- [1] Choi K, Suh Y. A new similarity function for selecting neighbors for each target item in collaborative filtering[J]. Knowledge-Based Systems, 2013, 37(1):146-153.
- [2] Zhang X, Li Y. Use of collaborative recommendations for web search: an exploratory user study[M]. Sage Publications, Inc. 2008.
- [3] Park Y, Park S, Jung W, et al. Reversed CF: A fast collaborative filtering algorithm using a k -nearest neighbor graph[J]. Expert Systems with Applications, 2015, 42(8):4022-4028.
- [4] Zhang J, Lin Y, Lin M, et al. An effective collaborative filtering algorithm based on user preference clustering[J]. Applied Intelligence, 2016, 45(2):230-240.
- [5] Wen J H, Shu S. Improved Collaborative Filtering Recommendation Algorithm of Similarity Measure[J]. Computer Science, 2014, 41(5):68-71.
- [6] Kaleli C. An entropy-based neighbor selection approach for collaborative filtering[J]. Knowledge-Based Systems, 2014, 56(C):273-280.
- [7] Kaleli C. An entropy-based neighbor selection approach for collaborative filtering[J]. Knowledge-Based Systems, 2014, 56(C):273-280.
- [8] Yi-Tao W U, Zhang X M, Wang X M, et al. User fuzzy similaritybased collaborative filtering recommendation algorithm[J]. Journal on Communications, 2016.
- [9] Pirasteh P, Jung J J, Hwang D. Item-Based Collaborative Filtering with Attribute Correlation: A Case Study on Movie Recommendation[M]// Intelligent Information and Database Systems. Springer International Publishing, 2014:245-252.
- [10] Pitsilis G, Knapkog S J. Social Trust as a solution to address sparsity-inherent problems of Recommender systems[J]. Computer Science, 2012.