



User-Based Collaborative Filtering

Difficulty Level : Hard • Last Updated : 16 Jul, 2020

Read Discuss Practice

User-Based Collaborative Filtering is a technique used to predict the items that a user might like on the basis of ratings given to that item by the other users who have similar taste with that of the target user.

Many websites use collaborative filtering for building their recommendation system.

Steps for User-Based Collaborative Filtering:

Step 1: Finding the similarity of users to the target user U.

Similarity for any two users 'a' and 'b' can be calculated from the given formula,

$$Sim(a, b) = \frac{\sum_p (r_{ap} - \bar{r}_a)(r_{bp} - \bar{r}_b)}{\sqrt{\sum_p (r_{ap} - \bar{r}_a)^2} \sqrt{\sum_p (r_{bp} - \bar{r}_b)^2}}$$

r_{up} : rating of user u against item p
 p : items

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !

Step 2: Prediction of missing rating of an item

Now, the target user might be very similar to some users and may not be much similar to the others. Hence, the ratings given to a particular item by the more similar users should be given more weightage than those given by less similar users and so on.

This problem can be solved by using a weighted average approach. In this approach, you multiply the rating of each user with a similarity factor calculated using the above mention formula.

The missing rating can be calculated as,

$$r_{up} = \bar{r}_u + \frac{\sum_{i \in users} sim(u,i) * r_{ip}}{\sum_{i \in users} |sim(u,i)|}$$

Example: Consider a matrix which shows four users *Alice*, *U1*, *U2* and *U3* rating on different news apps. The rating range is from 1 to 5 on the basis of users likability of the news app. The '?' indicates that the app has not been rated by the user.

Name	Inshorts(I1)	HT(I2)	NYT(I3)	TOI(I4)	BBC(I5)
Alice	5	4	1	4	?
U1	3	1	2	3	3
U2	4	3	4	3	5
U3	3	3	1	5	4

Step 1: Calculating the similarity between Alice and all the other users

At first we calculate the averages of the ratings of all the user excluding I5 as it is not rated by Alice. Therefore, we calculate the average as ,

$$\bar{r}_i = \frac{\sum_p r_{ip}}{\sum p}$$

Therefore, we have

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !

$$\begin{aligned}\bar{r}_{Alice} &= 3.5 \\ \bar{r}_{U1} &= 2.25 \\ \bar{r}_{U2} &= 3.5 \\ \bar{r}_{U3} &= 3\end{aligned}$$

and calculate the new ratings as,

$$r'_{ip} = r_{ip} - \bar{r}_i$$

Hence, we get the following matrix,

Name	Inshorts(I1)	HT(I2)	NYT(I3)	TOI(I4)
Alice	1.5	0.5	-2.5	0.5
U1	0.75	-1.25	-0.25	0.75
U2	0.5	-0.5	0.5	-0.5
U3	0	0	-2	2

Now, we calculate the similarity between Alice and all the other users.

$$Sim(Alice, U1) = \frac{((1.5*0.75)+(0.5*-1.25)+(-2.5*-0.25)+(.5*0.75))}{\sqrt{(1.5^2+0.5^2+2.5^2+0.5^2)}\sqrt{(0.75^2+1.25^2+0.25^2+0.75^2)}} = 0.301$$

$$Sim(Alice, U2) = \frac{((1.5*0.25)+(0.5*-0.5)+(-2.5*0.5)+(.5*-0.5))}{\sqrt{(1.5^2+0.5^2+2.5^2+0.5^2)}\sqrt{(0.5^2+0.5^2+0.5^2+0.5^2)}} = -0.33$$

$$Sim(Alice, U3) = \frac{((1.5*0)+(0.5*0)+(-2.5*-2)+(.5*2))}{\sqrt{(1.5^2+0.5^2+2.5^2+0.5^2)}\sqrt{(0^2+0^2+2^2+2^2)}} = 0.707$$

Step 2: Predicting the rating of the app not rated by Alice

Now, we predict Alice's rating for BBC News App,

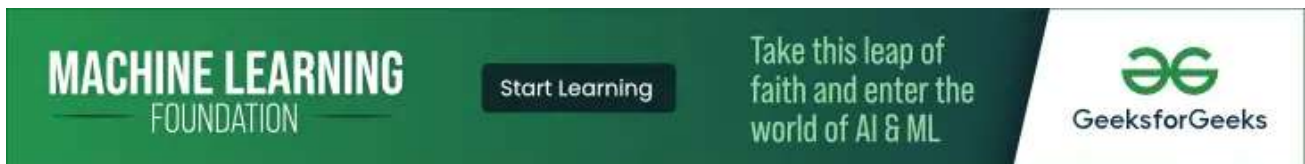
We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !

$$r_{(Alice,I5)} = \bar{r}_{Alice} + \frac{(sim(Alice,U1)*(r_{U1,I5}-\bar{r}_{U1}))+(sim(Alice,U2)*(r_{U2,I5}-\bar{r}_{U2}))+(sim(Alice,U3)*(r_{U3,I5}-\bar{r}_{U3}))}{|r_{U1,I5}|+|r_{U2,I5}|+|r_{U3,I5}|}$$

$$r_{(Alice,I5)} = 3.5 + \frac{(0.301*0.75)+(-0.33*1.5)+(0.707*1)}{|0.301|+|-0.33|+|0.707|} = 3.83$$

Hence, with the help of a small example we tried to understand the working of User-based collaborative filtering.



DSA Data Structures Algorithms Interview Preparation Data Science Topic-wise Practice

Previous

Next

Related Articles

1. Item-to-Item Based Collaborative Filtering
2. Collaborative Filtering – ML
3. Collaborative Learning – Federated Learning

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !

6. Skin Cancer Detection using TensorFlow
7. Dogecoin Price Prediction with Machine Learning
8. Loan Eligibility prediction using Machine Learning Models in Python
9. What is Standardization in Machine Learning
10. Measures Their Categorization and Computation in Data Mining

Article Contributed By :



rachitgupta1997

@rachitgupta1997

Vote for difficulty

Current difficulty : [Hard](#)

Easy	Normal	Medium	Hard	Expert
------	--------	--------	------	--------

Article Tags : [data mining](#), [data-science](#), [Machine Learning](#)

Practice Tags : [Data Mining](#), [Machine Learning](#)

Improve Article

Report Issue

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !

Company

About Us
Careers
In Media
Contact Us
Privacy Policy
Copyright Policy
Advertise with us

News

Top News
Technology
Work & Career
Business
Finance
Lifestyle
Knowledge

Web Development

Web Tutorials
Django Tutorial
HTML
JavaScript
Bootstrap
ReactJS
NodeJS

Learn

DSA
Algorithms
Data Structures
SDE Cheat Sheet
Machine learning
CS Subjects
Video Tutorials
Courses

Languages

Python
Java
CPP
Golang
C#
SQL
Kotlin

Contribute

Write an Article
Improve an Article
Pick Topics to Write
Write Interview Experience
Internships
Video Internship

© 2022 GeeksforGeeks. All Rights Reserved.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Got It !