Given 3 users with ratings…

u1:	1	3
u2:	2	4
u3:	1	4

A. $Sim_{corr}(u1, u2) > Sim_{corr}(u1, u3)$
B. $Sim_{corr}(u1, u2) = Sim_{corr}(u1, u3)$
C. $Sim_{corr}(u1, u2) < Sim_{corr}(u1, u3)$

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer B

With Pearson correlation, similarity is computed using vectors that are centered around the mean and the normalized. Therefore, after normalization all three vectors are the same, namely 1/sqrt(2) * (-1,1).

Item-based collaborative filtering addresses better the cold-start problem because …

A. usually there are fewer items than users
B. it uses ratings from items with similar content
C. item similarities can be pre-computed
D. none of the above

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer D

If a user has no ratings, also item-based collaborative filtering cannot make a prediction on the user's rating of an unknown item.

If the item in question has not ratings, not similar items can be determined and no predication can be made as well.

For a user that has not done any ratings, which method can make a prediction?

A. User-based collaborative RS
B. Item-based collaborative RS
C. Content-based RS
D. None of the above

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer D

User-base and item-based collaborative filtering cannot make predictions without previous ratings of the users.

Since the user has not rated any item, also content-based filtering is not applicable.

For an item that has not received any ratings, which method can make a prediction?

A. User-based collaborative RS
B. Item-based collaborative RS
C. Content-based RS
D. None of the above

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer C

Based on the content, similat items can be found, even if the item has not yet received a rating.

Question

How does matrix factorization address the issue of missing ratings?

A. It uses regularization of the rating matrix
B. It performs gradient descent only for existing ratings
C. It sets missing ratings to zero
D. It maps ratings into a lower-dimensional space

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer B

In gradient descent only data on existing ratings (user-item pairs) are exploited.

Question

With matrix factorization one estimates ratings of unrated items

A. By retrieving the corresponding item from a user vector
B. By retrieving the corresponding item from an item vector
C. By computing the product of a user and an item vector
D. By looking up the rating in an approximation of the original rating matrix

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer C

For estimating a new rating the latent representations of the corresponding user and item are used to compute an estimation. In principle, one could also consider Answer D as correct, but this would imply that all ratings are precomputed, which would be a waste of resources, if only few ratings need to be known.

Question

Which of the following matrices is not sparse?

A. The matrix W

B. The matrix $\hat{R}$

C. The matrix RW

D. More than one of the above are sparse

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer A

Only the matrix W is sparse. The two other matrices are the same, and contain rating estimations for all items. Therefore, they are not sparse.

Question

If users rate about 1% of items, and W has sparsity 1%, how many multiplications are needed to recommend the top-k elements?

A. $|I| * |I| * 0.0001$

B. $|I| * |I| * k * 0.01$

C. $|U| * |I| * k * 0.0001$

D. $|U| * |I| * 0.0001$

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer A

For finding the top k ratings, for each non-rated item of the user (about |I|), the rating vector of the user has to be multiplied with the corresponding item vector in W. The user rating vector has about |I| * 0.01 non-zero entries, as well as the item vector. Computing the scalar product of a dense vector takes |I| multiplictaions. Therefore, in total |I|^2 * 0.00001 multiplications will be performed.

Question

Assume in top-1 retrieval recommendation 1 is (2, 3, 1) and recommendation 2 is (2, 1, 3)

A. RMSE(rec 1) < RMSE(rec 2) and DCG(rec 1) > DCG(rec 2)
B. RMSE(rec 1) = RMSE(rec 2) and DCG(rec 1) > DCG(rec 2)
C. RMSE(rec 1) < RMSE(rec 2) and DCG(rec 1) = DCG(rec 2)
D. RMSE(rec 1) = RMSE(rec 2) and DCG(rec 1) = DCG(rec 2)

Remark: the optimal method would rank (3,2,1)

©2023, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Answer C

Since DCG considers only the top-k items, in this case the first item, the measure is independent of the ordering of the non top-k items. RMSE will give a better score to rec 1, since the higher rated item 3 is ranked higher in the result.