

User-based Collaborative Filtering

Focus on the similarities between users and recommend items that similar users like (Su and Khoshgoftaar, 2009)

Advantage

Recommendations based on historical behaviors of similar users are closer to user interests.
Suitable for scenarios with a stable user base: such as social platforms and subscription services.

Disadvantage

Sparsity: Since the proportion of items purchased by users is very low, the recommendation system based on the nearest neighbor algorithm may not be able to make effective recommendations, resulting in low accuracy.

Scalability: As the number of users and items grows, the computational complexity of the nearest neighbor algorithm increases, making the recommendation system difficult to scale (Sarwar et al., 2001).

Implementation

1. Construct a user-item rating matrix.
2. Calculate user similarity (such as cosine similarity):

$$\text{Sim}(u, v) = \frac{\sum_{i \in I} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in I} (r_{v,i} - \bar{r}_v)^2}}$$

3. Select the K most similar users, calculate the weighted average rating, and recommend the items with the highest rating.

Item-based Collaborative Filtering

Focus on the similarity between items and recommend items similar to the items the user likes (Ajaegbu, 2021)

Advantage

Efficient calculation: the number of items is usually smaller than the number of users, so the similarity calculation is faster.

More stable recommendation results: because the similarity of items is not likely to fluctuate over time(Ko et al., 2022).

Disadvantage

Cold start of new items: New items do not have enough user ratings to calculate similarity.

Recommendations lack personalization: focus more on item similarity rather than user personality(Ajaegbu, 2021).

Implementation

1. Construct a user-item rating matrix.
2. Calculate user similarity (such as cosine similarity):

$$\text{Sim}(i, j) = \frac{\sum_{u \in U} r_{u,i} \cdot r_{u,j}}{\sqrt{\sum_{u \in U} r_{u,i}^2} \cdot \sqrt{\sum_{u \in U} r_{u,j}^2}}$$

3. Select items that users have rated and recommend the most similar items

Evaluate the performance of a recommendation system

The simplest way is RMSE (Root Mean Squared Error).

It is calculated by taking the square root of the mean squared error (MSE) by dividing the sum of the squares of the differences between the actual and predicted grades by the total number of predicted grades(Ko et al., 2022).

Other common qualitative evaluation indicators of recommendation systems include Precision, Recall, Accuracy, F-Measure, ROC Curve, and AUC (Area Under the Curve)(Ko et al., 2022).