Authors: Sankalp Kulkarni, Aaditya Rathi, Akshat Oke, Rishi Poddar IDs: 2020A7PS1097H, 2020A7PS2191H, 2020A7PS0284H, 2020A7PS1195H

Recommendation Systems

The task is to compare various techniques used in implementing Recommender Systems on the basis of their errors using Root Mean Square Error, Precision on top K and Spearman Rank Correlation. Also compare their overall running time and prediction time.

Dataset

The Dataset used was taken from MovieLens. These files contain 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users who joined MovieLens in 2000.

This was divided into training and testing dataset of 80-20 percent.

Collaborative Filtering

For user based collaborative filtering, we follow these steps to find the ratings for an unrated movie for a given user:

1. Find the top k similar users for the given user. The similarity is cosine similarity between the movie vectors of two users.
2. For a movie, calculate the average of the ratings given by the above top k users.

We used k=15, that is, the top 15 similar users are taken into consideration for calculating the movie ratings.

Collaborative Filtering with Baseline Approach

Rating is predicted using the global mean and the movie deviation in addition to the collaborative filtering value from above.

$$\text{Rating} = c.f. + \mu_{global} + b_{movie}$$

where $c.f.$ is collaborative filtering prediction and b_{movie} is the deviation of the movie.

$$b_{movie} = \mu_{movie} - \mu_{global}$$

SVD

The singular value decomposition decomposes a matrix A into $U \times \Sigma \times V^T$ such that Σ is a diagonal matrix of the eignvalues of AA^T.

SVD with 90% retention

We retain 90% of the sum of squares of the diagonal matrix Σ and make other elements to zero. The following condition then holds true:

$$\Sigma[i] \geq \Sigma[i+1] \; \forall i \in \{0, \min(m,n)\}$$

CUR Decomposition

Similar to SVD, the matrix A is decomposed so that

$$A = C \times U \times R$$

where

- C has r randomly selected columns of A
- R has r randomly selected rows of A
- U is the pseudo-inverse of the intersection of C and R (=W)

r was selected to be $3,000$ rows/columns.

CUR with 90% retention

To find the pseudo-inverse of W, we take the SVD decomposition

$$W = X \times \Sigma \times Y^T$$

Now take the pseudo-inverse of Σ, which is just the reciprocals of all non-zero elements (since it is diagonal matrix) after retaining 90% energy.
Then,

$$U = Y \times \frac{1}{\Sigma}^2 \times X^T$$

Metrics

Training

Technique	RMSE	Precision on top 4	Spearman Rank Coefficient	Time taken
Collaborative	2.33	0.998	0.935	24.03
Collaborative with baseline	0.87	0.992	0.937	24.78
SVD	7.24e-15	0.9994	0.9999	41.34
SVD (90%)	0.8	0.905	0.9998	38.244
CUR	0.65	0.997	0.999	20.65
CUR (90%)	3.72	0.998	0.993	20.85

Testing

Technique	RMSE	Precision on top 4	Spearman Rank Coefficient
Collaborative	2.51	0.85	0.916
Collaborative with baseline	0.939	0.95	0.9233
SVD	3.75	0.978	0.999
SVD (90%)	3.66	0.983	0.9998
CUR	3.81	0.973	0.999
CUR (90%)	3.76	0.984	0.993

To account for generous and strict raters, we normalise all users by subtracting the mean of each user from their ratings.