Recommender Systems

In this assignment, we have implemented and compared various techniques for building recommender systems like Collaborative Filtering, SVD, and CUR.

Dataset

MovieLens 1M movie ratings. 1 million ratings from 6000 users on 4000 movies

Collaborative Filtering

Basic:

In this technique, recommendations are based on users similar to the user. For a user, 50 nearest neighbors are considered based on correlation coefficient (similarity). The predicted rating of user x and item i is computed using:

$$r_{xi} = \frac{\sum_y S_{xy} \cdot r_{yi}}{\sum_y S_{xy}} \quad y \in N$$ N is set of nearest neighbors

Baseline:

This method is used to effectively handle generous and strict raters appropriately. We come up with a baseline estimate which is used in predictions.

$b_{xi} = u + b_x + b_i$ u is overall mean and b is rating deviation

$$r_{xi} = b_{xi} + \frac{\sum_y S_{xy} (\cdot r_{yi} - b_{yi})}{\sum_y S_{xy}} \quad y \in N$$ N is set of nearest neighbors

Singular Value Decomposition

In this technique, the dimensions of the matrix are reduced from n to say k. This is achieved by decomposing the matrix into 3 parts U, sigma, V. This reduction is achieved by reducing the topics in our case movies into concepts by grouping similar ones into a concept.

$A = U \Sigma V^T$

This is the decomposition.

The concepts and user relation in our data is now used to predict the similarity between two users instead of relying on movies itself.

The correlation between two users is now found using the concepts and is used for collaborative filtering.

User-concept matrix is AV

Using the below properties of the matrices

$UU^T = I$

$VV^T = I$

$AV = U\Sigma$

This is used for calculating the similarity matrix.

CUR Decomposition

In this method, training utility matrix is split into three matrices C,U and R. In the construction of C and R matrices, we use the sampling technique.

$A = CUR$

C - sampled and normalized columns of A

R - sampled and normalized rows of A

U - pseudo inverse of matrix formed by intersection of C and R

Assumptions

- The number of nearest neighbors considered in Collaborative is 50
- The number of rows and columns considered in CUR is 600
- For metrics K = 10 and threshold = 3.5

Metrics

Recommender System Technique	Root Mean Square Error (RMSE)	Precision on top K	Spearman Rank Correlation	Time taken for prediction (secs)
Collaborative	1.149	0.782	0.99	51.19
Collaborative along with Baseline approach	1.097	0.791	0.99	93.58
SVD	1.501	0.724	0.99	101.60
SVD with 90% retained energy	1.503	0.724	0.99	47.79
CUR	0.689	0.96	0.99	94.42
CUR with 90% retained energy	0.683	0.95	0.99	184.30

Conclusions

The basic collaborative filtering does well but is not able to predict ratings for new users/ items. It also can not take into account the generous and tough raters. This is solved using the baseline estimate. The former problem is solved using the SVD technique is another technique which can be used for better predictions as it introduces the idea of concepts. But SVD is a time consuming algorithm, hence we make use of the improved algorithm CUR. CUR works equally well as SVD 98% of the times. This reduces the complexity by a huge amount.

Team Members

Adesh Kumar Pradhan – 2017B3A70960H
Pranavi Marripudi – 2018A7PS0507H
Mereddy Aishwwarya Reddi – 2018A7PS0276H