

Mean Absolute Error: 0.6949234976496896 &

Root Mean Squared Error: 0.8844601266219813

This means that all my predictions are within 0.695 (MAE) range of the actual predictions, on an average, for all the test examples.

The value of RMSE is more than the MAE because it gives more weightage to predictions which are further from the actual rating whereas in the case of MAE, it gives equal weightage to all the predictions however far they may be from the actual rating. So RMSE is a stricter evaluation metric as it penalizes the prediction which are far from the actual rating.

The predictions took a lot of time as the dataset was huge. It had 27555 unique users and when I tried to calculate the weight matrix between each pair of users for all the rated movies, the  $n \times n$  matrix was so huge it would not fit in the memory of the system. Even when I tried to save the intermediate results it was infeasible to save the file in the harddisk.

So the final solution was to calculate the matrix on the fly for each example in the testing file. I optimized it further by calculating the prediction for all the movies for a particular user in each iteration. So I didn't have to calculate for a particular users again if the testing file contained the rating of the same user for different movies.

Still the program takes almost 10 hours to execute. It could be optimized still more by calculating only half of the weight matrix as the  $w(a,i)$  is same as  $w(i,a)$  but I ran out of time.