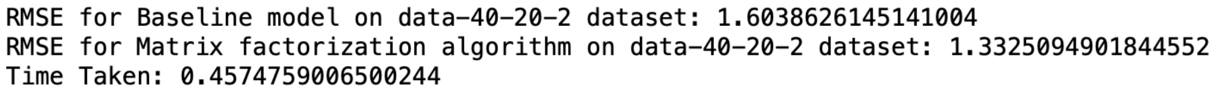
Results For Dataset data-40-20-2 Dataset:



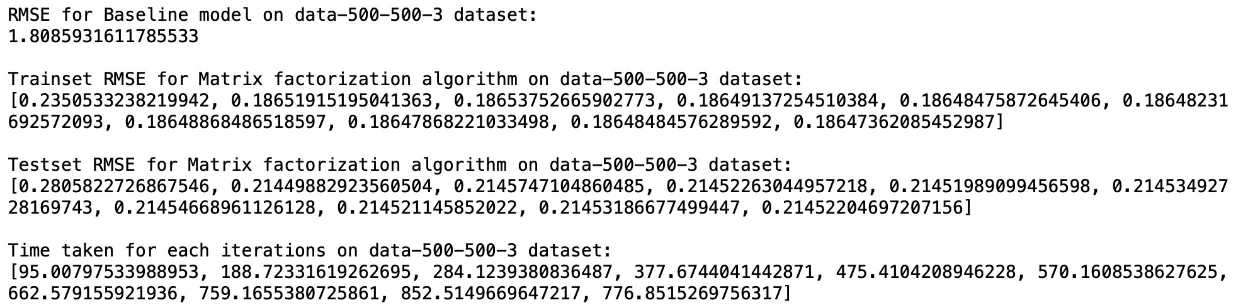
```
RMSE for Baseline model on data-40-20-2 dataset: 1.6038626145141004
RMSE for Matrix factorization algorithm on data-40-20-2 dataset: 1.3325094901844552
Time Taken: 0.4574759006500244
```

We can see that Matrix decomposition algorithm performs better than baseline model.

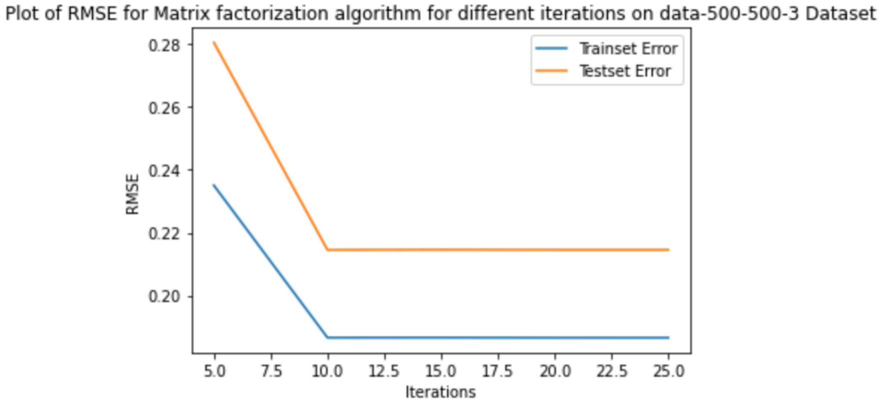
For time saving algorithm was run on multiple notebooks simultaneously.

Part1: Experiments on data-500-500-3 Dataset:

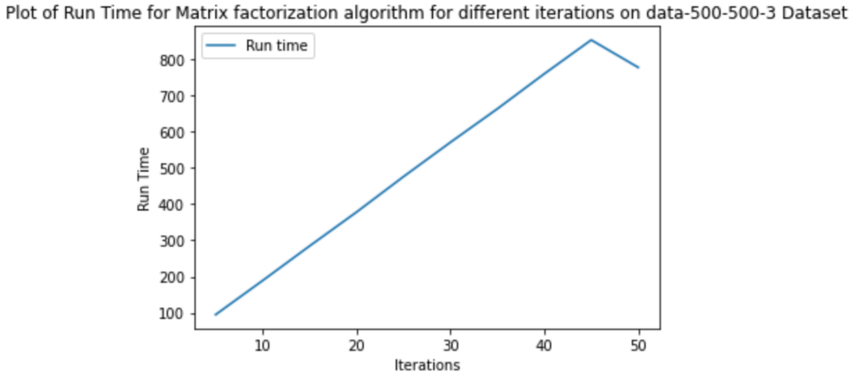
Results for different iterations:



```
RMSE for Baseline model on data-500-500-3 dataset:
1.8085931611785533

Trainset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[0.2350533238219942, 0.18651915195041363, 0.18653752665902773, 0.18649137254510384, 0.18648475872645406, 0.18648231
692572093, 0.18648868486518597, 0.18647868221033498, 0.18648484576289592, 0.18647362085452987]

Testset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[0.2805822726867546, 0.21449882923560504, 0.2145747104860485, 0.21452263044957218, 0.21451989099456598, 0.214534927
28169743, 0.21454668961126128, 0.214521145852022, 0.21453186677499447, 0.21452204697207156]

Time taken for each iterations on data-500-500-3 dataset:
[95.00797533988953, 188.72331619262695, 284.1239380836487, 377.6744041442871, 475.4104208946228, 570.1608538627625,
662.579155921936, 759.1655380725861, 852.5149669647217, 776.8515269756317]
```



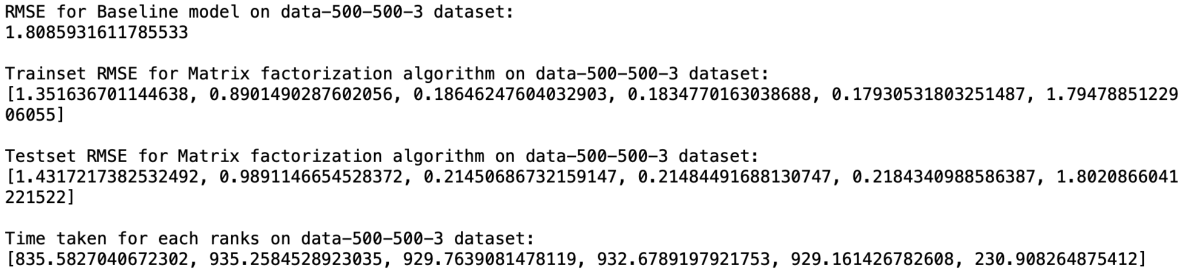
Plot of RMSE for Matrix factorization algorithm for different iterations on data-500-500-3 Dataset



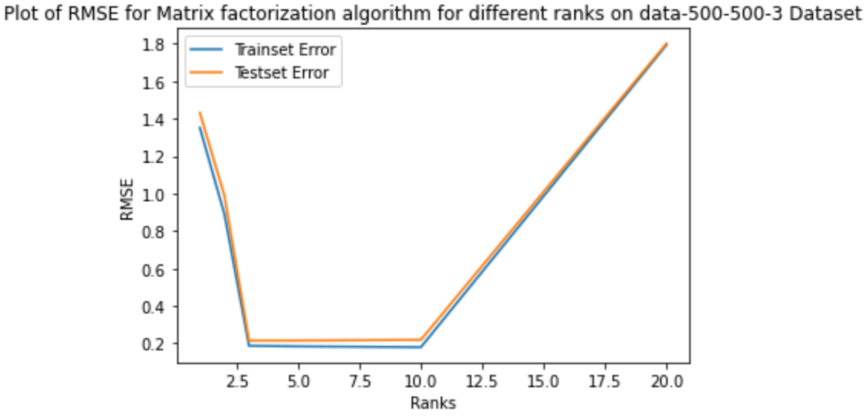
Plot of Run Time for Matrix factorization algorithm for different iterations on data-500-500-3 Dataset

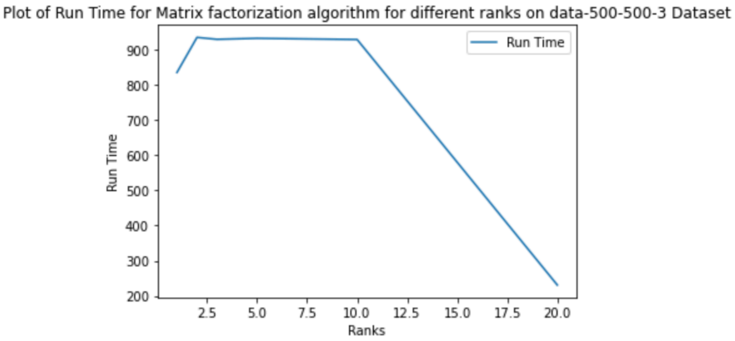
Above results shows that Matrix decomposition model outperforms the baseline model. From the plot we can see that error rate decreases with increase in the number of iterations, and also model converges at 15 iterations. Error on test set is higher compared to training set due to prediction on unseen data. Model doesn't overfit with the increased number of iterations.
Run time of the model increases almost linearly with higher iterations as shown in the plot.

Results for different Ranks:



```
RMSE for Baseline model on data-500-500-3 dataset:
1.8085931611785533

Trainset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[1.351636701144638, 0.8901490287602056, 0.18646247604032903, 0.1834770163038688, 0.17930531803251487, 1.79478851229
06055]

Testset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[1.4317217382532492, 0.9891146654528372, 0.21450686732159147, 0.21484491688130747, 0.2184340988586387, 1.8020866041
221522]

Time taken for each ranks on data-500-500-3 dataset:
[835.5827040672302, 935.2584528923035, 929.7639081478119, 932.6789197921753, 929.161426782608, 230.908264875412]
```



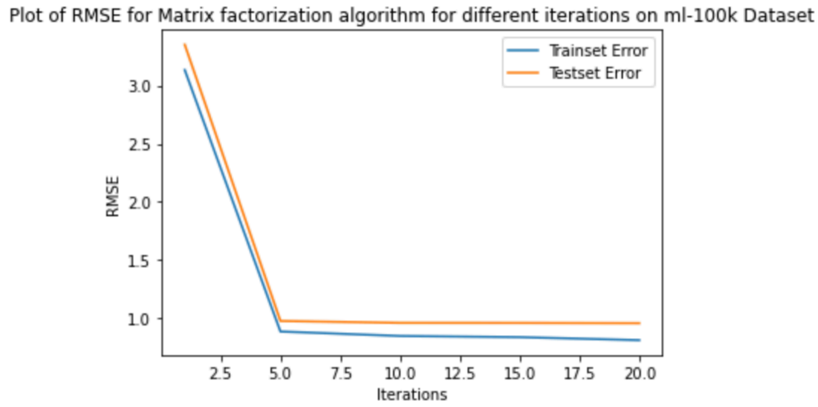


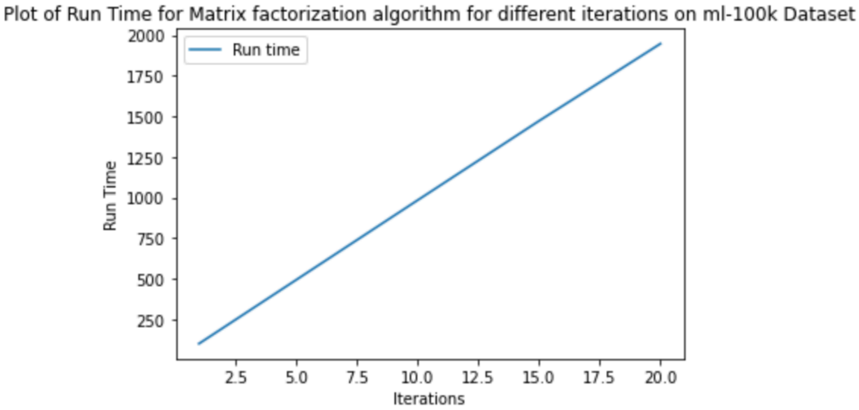
From the above results we can see that error decreases with the increase in rank, and coverages at rank 3. The large spike in at rank 20 is due to termination of algorithm based on $\|u^{-} - u^{-\,old}\| < 0.01$ following condition which relates to the less run time. Since normally run time increases with increase in rank, which is reflected until rank 10.

Part2: Experiments on ml-100k Dataset:

Results for different iterations:

```
RMSE for Baseline model on ml-100k dataset:
0.987077661785431

Trainset RMSE for Matrix factorization algorithm on ml-100k dataset:
[3.1346149379000994, 0.8856999821064543, 0.8475142178717391, 0.8371820833241175, 0.8114342626193253]

Testset RMSE for Matrix factorization algorithm on ml-100k dataset:
[3.3523981254162605, 0.9764807805268304, 0.960245923231339, 0.9596497376779604, 0.9572013363790863]

Time taken for each iterations on ml-100k dataset:
[97.9582929611206, 491.2160429954529, 982.4216120243073, 1471.5200171470642, 1947.568200826645]
```

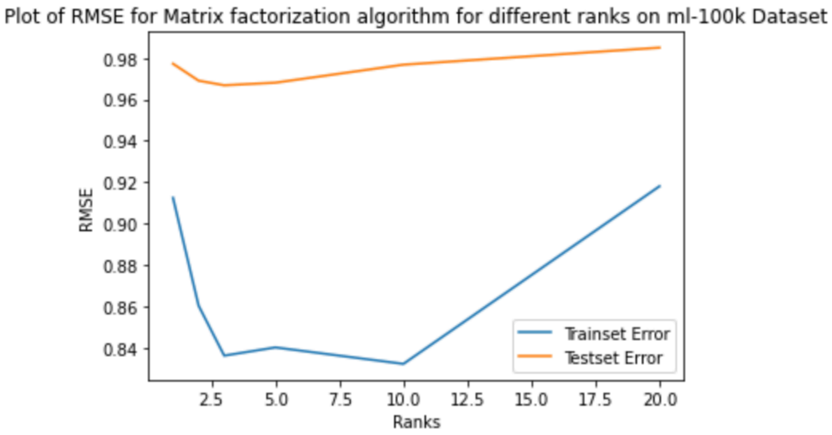


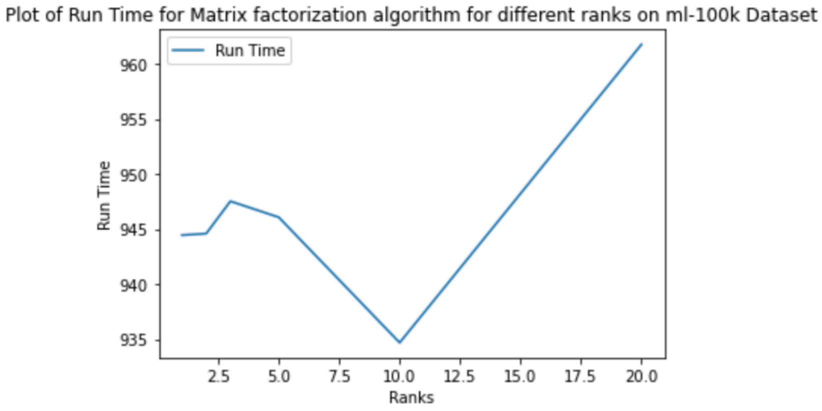


Above results shows that Matrix decomposition model performs better the baseline model. From the plot we can see that error rate decreases with increase in the number of iterations, and also model converges at 10 iterations. Error on test set is higher compared to training set due to prediction on unseen data. Model doesn't overfit with the increased number of iterations.
Run time of the model increases almost linearly with higher iterations as shown in the plot.

Results for different Ranks:

```
RMSE for Baseline model on ml-100k dataset:
0.987077661785431

Trainset RMSE for Matrix factorization algorithm on ml-100k dataset:
[0.9124310093663421, 0.8602739410091855, 0.8361557031762931, 0.8401142798960715, 0.8321635916508645, 0.918029674117
959]

Testset RMSE for Matrix factorization algorithm on ml-100k dataset:
[0.9772580111982806, 0.9690775491187654, 0.9668531204629545, 0.9681123906224746, 0.9768082803871406, 0.985033818237
7717]

Time taken for each ranks on ml-100k dataset:
[944.4395041465759, 944.579163312912, 947.5134961605072, 946.0610370635986, 934.6771709918976, 961.7496240139008]
```



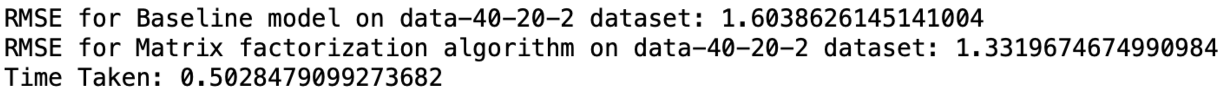


From the above results we can see that error decreases with the increase in rank, and coverages at rank 10. The large spike in at rank 20 is due to termination of algorithm based on $\|\bar{u} - \bar{u}^{old}\| < 0.01$ following condition which relates to the less run time. Since normally run time increases with increase in rank, which is reflected until rank 10.

Extra Credits:

Matrix U and V were initialized to the results obtained after one iteration of EM SVD algorithm.

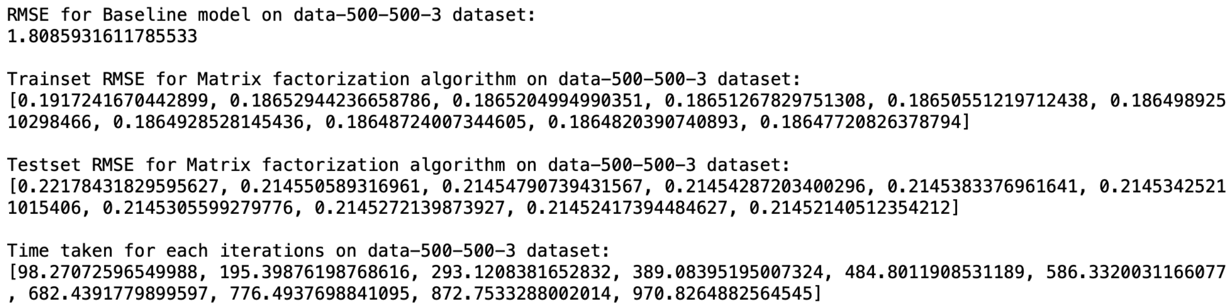
Results For Dataset data-40-20-2 Dataset:



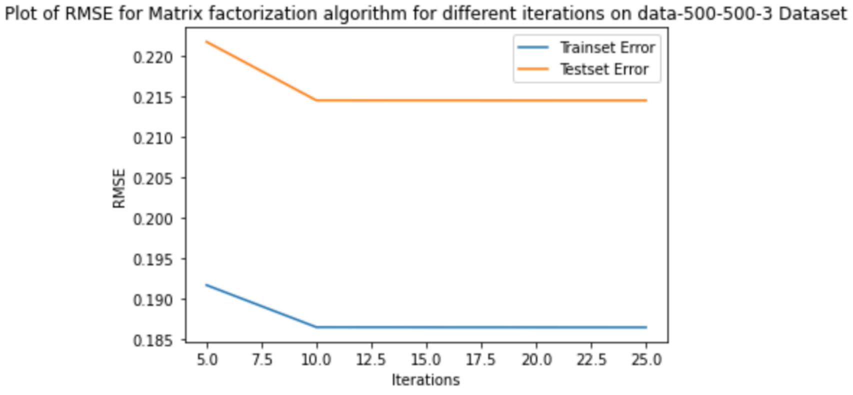
```
RMSE for Baseline model on data-40-20-2 dataset: 1.6038626145141004
RMSE for Matrix factorization algorithm on data-40-20-2 dataset: 1.3319674674990984
Time Taken: 0.5028479099273682
```

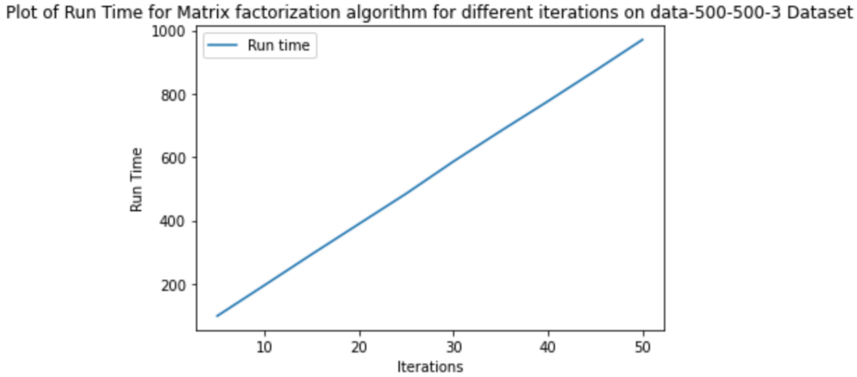
We can see that Matrix decomposition algorithm performs better than baseline model. But does make much difference from the normal random initialization on small dataset.

Results for different iterations on data-500-500-3 Dataset:



```
RMSE for Baseline model on data-500-500-3 dataset:
1.8085931611785533

Trainset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[0.1917241670442899, 0.18652944236658786, 0.1865204994990351, 0.18651267829751308, 0.18650551219712438, 0.186498925
10298466, 0.1864928528145436, 0.18648724007344605, 0.1864820390740893, 0.18647720826378794]

Testset RMSE for Matrix factorization algorithm on data-500-500-3 dataset:
[0.22178431829595627, 0.214550589316961, 0.21454790739431567, 0.21454287203400296, 0.2145383376961641, 0.2145342521
1015406, 0.2145305599279776, 0.2145272139873927, 0.21452417394484627, 0.21452140512354212]

Time taken for each iterations on data-500-500-3 dataset:
[98.27072596549988, 195.39876198768616, 293.1208381652832, 389.08395195007324, 484.8011908531189, 586.3320031166077
, 682.4391779899597, 776.4937698841095, 872.7533288002014, 970.8264882564545]
```





Results shows that the matrix decomposition algorithm the EM SVD Initialization converges faster at iteration 10 compared to the model with normal random initialization. And also we see lower initial error rates when compared to the other model. But the outcome for both models is same.