Dataset Names - Details - Links ….

Work Done till now -

1. Performed Matrix Factorization method of Collaborative filtering on 5-core Amazon Reviews datasets (subsets) - datasets containing items of 1 category . 2 different methods were checked - ALS and SGD, with regularization and various choices for the hyperparameters. MSE is the cost function used
2. We see that SGD performs better on these datasets - much better learning curve for all hyperparameter choices.

Challenges ahead -

1. Currently, we have made progress on the problem of Matrix Completion of the User-Item ratings matrix, where using the above techniques we are able to get
2. Currently, we see slow convergence for even moderately sized datasets. We need to find a way to parallelize the solution or use more compute power
3. We want to run our solution on larger datasets where many more categories of products are present.
4. The cold start problem of new user/ new item needs to be addressed.
5. We will try to find a way of recommending existing and new users, top-N recommendations once we have decomposed(factorized) the particular User-Item rating Matrix.