BITS Pilani Hyderabad Campus

CS F469 IR Assignment - 3

Deadline: November 2, 2019

This assignment is aimed at implementing and comparing various techniques for building a Recommender System taught in class.

The assignment can be done in groups of 4. Please refrain from changing the groups for the assignment and the ones coming in the near future.

Programming Languages:

The assignment can be implemented in any programming language of your choice. You are expected to code the core functionality of the algorithm. Matrix multiplication, eigenvectors and eigenvalues can be found using the built in libraries.

Expectations

- Successful implementation of the techniques on a reasonably sized dataset.
- Ensuring generous raters and strict raters are handled appropriately.
- Explaining the results in the design document and clearly stating all the information.
- Mention how many rows and columns were used for CUR, how many neighbours considered in case of Collaborative Filtering and the other essential details in the design document.
- Fill in the following table:

Recommender System Technique	**Root Mean Squared Error (RMSE)**	**Mean Average Error (MAE)**	**Time taken for prediction**
Collaborative			
Collaborative along with Baseline approach			

SVD			
SVD with 90 per cent energy			
CUR			
CUR with 90 per cent energy			
Latent factor model			

- Experiment with different number of latent factors in the latent factor model.

References

- http://infolab.stanford.edu/~ullman/mmds/ch11.pdf
- Latent Factor Model: https://www.youtube.com/watch?v=E8aMcwmqsTg

Deliverables

The final submission must contain the following documents:

1. **Design Document**: This document should contain the description of the application's architecture along with the major data structures used in the project. It should contain the report on the different distance measures used for the problem. Precision and Recall, if possible, should also be calculated. Running for all the preprocessing should be mentioned. Also, mention the running time of the retrieval and the search.
2. **Code**: The code should be well commented.
3. **Documentation**: All the classes, functions and modules of the code must be documented.
4. **README**: The README file should describe the procedure to compile and run your code for various datasets.

For any queries, contact Mr. Bharat Malhan (f20160050@hyderabad.bits-pilani.ac.in)

Submission Guidelines:

All deliverables must be zipped and submitted in CMS latest by deadline.

You are expected to demo your application and present your results as per the schedule that will be made available.

Evaluation Criteria

- Implementation: 20M
- Design Document: 5M
- Filling the table: 15M
- Handling generous raters and strict raters: 5M
- Viva: 5M

It should be noted that all the assignments would be run through a plagiarism detector and any form of plagiarism will not be tolerated and shall be brought to the notice of AUGSD/AGSRD. The final decision lies in the hands of the instructor and only one submission per group would be allowed for one assignment.