

Assignment 1 – Advanced Topics in Recommender Systems

Part 1 – Stochastic Gradient Descent:

- 1) Write the objective of a regression model with global bias, user bias, item bias and L2 regularization.
- 2) Write the update step for each parameter.
- 3) Write a pseudo code for the algorithm.
- 4) What hyper-parameters do you need to tune?
- 5) Explain how would you work with the validation set and how would you check for convergence?
- 6) How would you train the last \ best model?
- 7) Implement an SGD solution for the model and train it using the training and validation data. Explain the main work items you had to take.
- 8) What is the RMSE, MAE, R^2 and MPR of your model based on the validation set?
- 9) Submit the test result file according to the following instructions:
 - a. The name of the file should be SGD and the student ID numbers separated with an underline. i.e: **SGD_<id1>_<id2>_<id3>.csv**
 - b. The file content should be CSV in the same order as the test file you received and using the following format:
User_ID_Alias, Movie_ID_Alias, Rating
 - c. The report and the results should be submitted by one of the group members to Moodle.

Part 2 – Alternating Least Squares:

- 1) Write the objective of a regression model with global bias, user bias, item bias and L2 regularization. Is there any difference from the SGD objective?
- 2) Write the update step for each parameter.
- 3) Write a pseudo code for the algorithm.
- 4) What hyper-parameters do you need to tune?
- 5) Explain how would you work with the validation set and how would you check for convergence?
- 6) Implement an ALS solution for the model and train it using the training and validation data. Explain the main work items you had to take.
- 7) What is the RMSE, MAE, R^2 and MPR of your model based on the validation set?
- 8) Compare the ALS and SGD solutions in terms of implementation, training and quality.
- 9) Submit the test result file according to the following instructions:
 - a. The name of the file should be ALS and the student ID numbers separated with an underline. i.e: **ALS_<id1>_<id2>_<id3>.csv**
 - b. The file content should be CSV in the same order as the test file you received and using the following format:
User_ID_Alias, Movie_ID_Alias, Rating
 - c. The report and the results should be submitted by one of the group members to Moodle.

*submission notes below.

- The assignment should be submitted in groups of 2-3 students after signing up on Moodle.
- The due date is 7/12/22 at 23:59.
- The train and validation files contain a list of movie ratings by users. The test file is in the same format, but you are not given the ratings. The ratings in the test file will be filled using the recommender system you built.
- The final submission files are the following:
 - PDF according to the assignment instructions.
 - Code file.
 - Test results – using SGD.
 - Test results – using ALS.

*submit 4 separate files (don't use a zip file).
- Add the group members' names and ID numbers to the PDF.