Instructor: Alina Vereshchaka

Project 5 -- Recommendation Systems

Deadline: August 14, Fri, 11:59pm
Hard Deadline: August 17, Mon, 11:59pm

Description

The main goal of our Project 5 is to build a collaborative filtering recommendation system using traditional ML approach and the deep learning method. We will test these algorithms on MovieLens datasets.

Datasets

For this project we will work with MovieLens dataset. This dataset is often used for the benchmarking.
MovieLens (small 100k): https://grouplens.org/datasets/movielens

Tasks

Part I: ML Method for Recommendations [40 points]

1. Preprocess the dataset. Highlight main statistics about it.
2. Choose any machine learning method used for recommendation system, e.g.
   a. SVD
   b. Nearest Neighbor
3. Given a completely new user in the system, return a list of recommended movies. Reason your choice.
4. Given an existing user ID and other attributes as an input, return a list of recommended movies (including movie's names). This has to be based on the expected rating by that user.
5. Discuss the results.

Part II: Deep Learning Methods for Recommendations [60 points]

1. Prepare the dataset for training.
2. Choose any deep learning method used for recommendation system, e.g.
   a. Softmax deep neural network
   b. Wide and deep learning ([pdf](#))
3. Describe the model, hyperparameters, accuracy/loss, etc. Show the list of recommended movies.
4. Discuss the results of how a deep learning method works for recommendation problems. Compare with the results from Part 1.

Submit the Project

- Submit at **UBLearns > Assignments**
- The code of your implementations should be written in Python. You can submit multiple files, but they all need to have a clear name
- All project files should be packed in a ZIP file named **TEAMMATE#1_UBIT_TEAMMATE#2_UBIT_project5.zip** (e.g. **avereshc_neelamra_project5.zip**).
- Your Jupyter notebook should be saved with the results. If you are submitting python scripts, after extracting the ZIP file and executing command python main.py in the first level directory, all the generated results and plots you used in your report should appear printed out in a clear manner.
- In your report include the answers to questions for each part. You can complete the report in a separate pdf file or in Jupyter notebook along with your code.
- Include all the references that have been used to complete the project.

Important Information

This project can be done in a team of up to two people. The standing policy of the Department is that all students involved in an academic integrity violation (e.g. plagiarism in any way, shape, or form) will receive an F grade for the course. Refer to the [Academic Integrity website](#) for more information.

Late Days Policy

You can use up to 5 late days throughout the course that can be applied to any project. You don't have to inform the instructor, the late submission will be tracked in UBlearns. If you work in teams the late days used will be subtracted from both partners. E.g. you have 4 late days and your partner has 3 days left. If you submit one day after the due date, you will have 3 days and your partner will have 2 days left.

Important Dates

August 14, 11:59pm - Project 5 is Due

August 17, 11:59pm - Hard Deadline