

# CSE4/510: Applied Deep Learning Summer 2020

**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 two different datasets.

### Datasets

For this project we will work with two datasets -- MovieLens and any of your choice.

1. MovieLens (small 100k): <https://grouplens.org/datasets/movielens>  
MovieLens is a rating dataset from the MovieLens website, which has been collected over some period. This dataset is often used for the benchmarking.
2. Any of the following:
  - a. [Book Crossing](#) -- collected by Cai-Nicolas in a 4-week crawl from the Book-Crossing community.
  - b. [Million Song Dataset](#) -- collection of audio features and metadata for a million contemporary popular music tracks.
  - c. [Dating Agency](#) -- contains 17,359,346 anonymous ratings of 168,791 profiles made by 135,359 LibimSeTi users.
  - d. [Anime Recommendations Database](#) -- contains information on user preference data from 73,516 users on 12,294 anime. Each user is able to add anime to their completed list and give it a rating and this data set is a compilation of those ratings.
  - e. [Chicago Entree](#) -- contains a record of user interactions with the Entree Chicago restaurant recommendation system.

## Tasks

### Part I: ML Method for Recommendations [40 points]

1. Preprocess the datasets. Highlight main statistics about them.
2. Choose any deep learning method used for recommendation system, e.g.
  - a. SVD
  - b. Nearest Neighbor
3. Given an input a user ID and other attributes, return a list of recommended items, based on the expected rating by that user.
4. Discuss the results.

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

1. Preprocess the datasets.
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.
4. Discuss the results of how deep learning methods work for recommendation problems.

## 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 UBLearn. 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