National University of Computer & Emerging Sciences, Karachi
Spring 2023 (School of Computing)
Assignment # 1

Course Code: CS-4053/AI-4006	**Course Name:** Recommender Systems
Course Instructor: Syed Zain Ul Hassan	
Open Date: February 18, 2023	**Deadline:** March 12, 2023 (11:30 PM)

Read carefully:

- The assignment consists of three parts.
- You can find the max points for each task and rubrics for assessment in the supplementary file.
- The submission instructions are also provided in the supplementary file.
- You will be given zero points in case any corrupt files or empty folders are submitted. You are expected to be careful while submitting your assignment.
- You will be given -1 weightage if your work is found to be copied. This will be unarguable.

Part 1

For the following ratings given to items by the users, perform the tasks below:

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	?	3	?	3	4
User 2	4	?	?	2	?
User 3	?	?	3	?	?
User 4	3	?	4	?	3
User 5	4	3	?	4	4

a) Find nearest neighbors of user 1 for *k=2.*
b) Apply user-based Collaborative Filtering to predict **R**(U1, I3) (read: rating for user 1 and item 3).
c) Apply item-based Collaborative Filtering to predict **R**(U3, I2) and **R**(U4, I4).

For all of these tasks, use Pearson Correlation Coefficient as similarity measure and mean-centered prediction function.

d) Say we had an option of using mode-centered prediction function instead of using mean. What do you think would be the effect of using mode as a baseline instead of average? Can using mode ever improve the performance of recommender system?

Part 2

Find and download the file “ml-latest-small.zip” containing the movie ratings dataset in the course folder on Google Drive. Perform the following tasks:

a) Write a Python (or Java) program to implement a user-based Collaborative Filtering Recommender System that takes a movie name as input and recommends top-5 movies to the current user. Use Cosine Similarity for finding neighborhoods and mean-centered prediction function (refer to Lecture 2 slides for context).
b) Write a Python (or Java) program to implement a Content-based Recommender System that takes a movie name as input and recommends top-3 movies to the user. Use genres as features and the provided ratings for vectorization of all profiles (refer to Lecture 3 slides for context).

Note: You are allowed to use basic libraries such as Pandas and Sci-kit. Using utility functions for Cosine Similarity and other numerical operations is also allowed.

Part 3

Mention any two of the Youtube channels that you are currently subscribed to. Be aware that this data may be used in our course at a later stage but without any mention of your identity.