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

Course Code: CS-4053	**Course Name:** Recommender Systems
Course Instructor: Syed Zain Ul Hassan	
Open Date: October 6, 2023	**Deadline:** October 21, 2023 (11:30 PM)

Instructions:

- The assignment consists of three (3) 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

Consider the following ratings given to items by the users and perform the given tasks:

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

a) Find nearest neighbors of *User 1* and *User 4* for *k=1.*
b) Apply item-based Collaborative Filtering to predict **R**(U1, I3), **R**(U3, I1) and **R**(U4, I4).

Use Pearson Correlation Coefficient as similarity measure and mean-centered prediction function.

c) Using only your human judgement, state which one of these users may be the least reliable in terms of their provided ratings. Give a reason for your selection.
d) Say that instead of mean we find the median and use it in the prediction function. Will that change make our predictions more accurate in any case? Justify your opinion with an argument.

Part 2

Find and download the file "*ml-latest-small.zip*" containing the movie ratings dataset in the course folder on Google Drive. Use this data and write a Python (or Java) program to implement an item-based Collaborative Filtering Recommender System. It should take a movie name as input and recommends top-3 movies to the current user. Use Adjusted Cosine Similarity for finding neighborhoods and mean-centered prediction function *(refer to Lecture 2 slides for context)*.

Part 3

Download the Python notebook given with this assignment. It contains sample data and starter code. Your task is to implement the functions to find the likelihood, prior and posterior. Then with the help of these functions, predict rating for active user-item using User-based Naïve Bayes Collaborative Filtering.

Note: You are allowed to import and use basic libraries such as Sci-kit in your functions.

- EOF -