Data Structures – Lab Project [WEIGHT: 25% of the lab weight]

- This project can be done in groups. Maximum allowed size of the group is three persons. Evaluation will be done on individual basis. Your contribution, concepts and viva performance will define your marks in the project besides quality of submissions.
- Submission will be in two forms: source code and technical report.
- **Code:** Code must be formally documented and basic principals of writing a good software program should be adhered. You may reuse your own (or any of the group members code from labs) but cannot take any piece of code from any other source.
- **Report:** You are required to submit a group project report using Latex (Overleaf recommended, as presented in the tutorial). MS Word reports will carry zero weight. Format of the report must be adhered. Your report should contain problem statement, data analysis of given data, performance of different algorithms on the given data in terms of average search time (theoretical as well as physical time. Report size should range from 6 to 8 pages. PDF file of the report needs to be uploaded over LMS with the source code by one of the group members.
- Any suspicious activity violating honor code (available over LMS) will be referred to the disciplinary committee. In case of any confusion, please feel free to talk to the instructor before submission.
- **Deadline:** Complete projects are due on January 31, 2021, 11:59PM. Individual and group viva will take place after the submission as scheduled by the course team.

TASK

Your task is to build a search engine for the given data (using Python). Data is composed of objects (records). Each object has three attributes: movie ID, average rating and number of votes to review each movie. Your task is populate the data in your search engine based on the average rating. You are required to compare performance of:

a) Hashing (quadratic probing and separate chaining)
b) 2-3-4 trees ☺
c) AVL trees

You can use the search engine to answer questions like:

a) Which one is the most popular movie?
b) Which one is the least popular movie?
c) How many movies have same ratings?
d) …

The dataset has been borrowed from https://www.imdb.com