GRAPH DATA MINING: CAPSTONE PROJECT

For the capstone project, you are to investigate an **advanced topic in graph data mining**. This includes identifying a **research question**, performing a **literature review**, and proposing, implementing, and evaluating a **solution**.

You are to perform the following tasks:

1. Identify a **research question** in graph data mining that was not covered in the class.
    - You may choose a particular **application of graph data mining** (e.g., *biological networks*, *financial networks*), a particular **graph mining task** (e.g., *frequent subgraph mining*, *link prediction*), or another **topic related to graph data mining** (e.g., *multilayer networks*, *network resilience*).
    - Your research question may be related to a topic covered in the class. For example, you may not choose community detection in graphs, but you may choose community detection in *dynamic* graphs.
2. Identify at least three **research papers** that could be relevant to solving your research question. Search for highly cited papers published in top tier conferences (e.g., KDD, ICDM, ICDE) or high-impact journals using Google Scholar.
3. Write a one-paragraph **summary** of each relevant research paper.
4. Propose a **solution** for your research question and implement a **prototype** in Python.
5. Evaluate your proposed solution and report the **results** obtained.

Note that you will be graded based on your ability to investigate an advanced topic in graph data mining, and not based on the novelty of your proposed solution.

Submission Details

Submit a ZIP folder containing the following materials:

- A text file with the names of the members of the team and the contributions of each member to the project.
- Python code of your proposed solution and README file with detailed instructions on how to run your code. Also include any data sets used to evaluate your code.
- Relevant research papers (as PDF files).
- Project report (2-3 pages, as a PDF file) with the following information:
    - Description of research question, including why solving this question is important.
    - Summaries of relevant research papers.
    - Description of proposed solution.
    - Description of methodology used to evaluate the proposed solution.
    - Discussion of results obtained.

Grading Rubric

Implementation	40 pts.
Project report • Description of research question • Summaries of research papers • Description of proposed solution • Description of evaluation methodology • Discussion of results	 15 pts. 15 pts. 10 pts. 10 pts. 10 pts.