Project Title: Implementing Membership Inference Attacks on Synthetic Graphs

Project Description:
In this project, students will select one of the provided research papers on membership inference attacks (MIAs) on machine learning or generative models and implement its approach using a synthetic graph dataset that will be supplied. The core goal is to understand the underlying attack methodology detailed in the chosen paper and adapt its key inference techniques in the context of graph data. Students will explore how membership information, whether a graph was used during the training of a model can be inferred by analyzing the model's behavior or output distributions.

Project Objectives:

- **Literature Review:** Read and summarize a selected paper, identifying its attack mechanism, key hypotheses, and evaluation metrics.
- **Implementation:** Adapt the attack method (e.g., through shadow models, gradient-based features, or distribution comparison) described in the paper.
- **Experimentation:** Apply the implemented method to the provided synthetic graph data. Analyze how effectively your approach can distinguish between nodes used during training and those that are not.
- **Evaluation and Analysis:** Compare your experimental results with the claims or baseline results given in the paper and discuss their impact on the attack's success.

Project Tasks:

1. **Paper Selection & Review:**
   - Choose one paper from the provided list.
   - Write a brief literature review summarizing the key attack strategy, assumptions, and results.
2. **Design & Planning:**
   - Outline the attack's methodology and define the experimental setup using the provided synthetic graph.
   - Determine the evaluation metrics (e.g., AUC, accuracy, precision, recall) that you will use to measure your implementation's performance.
3. **Implementation:**
   - Code the membership inference attack as described in the paper using a programming language such as Python.
   - Utilize appropriate libraries (e.g., NetworkX for graph handling, PyTorch/TensorFlow for model implementation, scikit-learn for classifier training) to facilitate your work.
   - Reproduce the key steps of the attack: data preparation, model training (or simulation of the target model's behavior on the synthetic graph), and membership inference.
4. **Experiments & Analysis:**
   - Run experiments on the synthetic graph dataset.
   - Collect results and compare them with the expected outcomes or baseline performance mentioned in the paper.
   - Analyze the impact of different hyperparameters or settings (e.g., number of training samples, noise levels, or sampling steps) on your attack's effectiveness.

5. **Documentation & Presentation:**
    - Write a project report documenting your approach, implementation details, experimental results, and a discussion on the challenges faced and lessons learned.
    - Prepare a short presentation to share your findings with the class.

Deliverables:

- A project report (max 10 pages) including background, methodology, experimental setup, results, analysis, and conclusions.
- Source code of your implementation, organized and commented.
- A brief presentation (5–10 minutes) summarizing your work and main findings.

Evaluation Criteria:

- Depth of literature review and clarity in summarizing the chosen paper.
- Correctness and completeness of the implementation based on the paper's methodology.
- Rigor in experimental analysis and insightful discussion on the results.
- Quality of documentation and presentation.