### Final Report (50 pt)

PDF write-up, 4 - 8 pages + unlimited references. We will deduct **2pts** if you exceed the page limit.

**Please follow the below format:**

1. Abstract (2)
   1. Link to your video (1)
      1. No points will be awarded if the link is not clickable or does not direct to your public project video (e.g., YouTube, OneDrive, Google Drive, or <https://mediaspace.illinois.edu/>).
   2. Link to Public GitHub Repo (1)
      1. No points will be awarded if the link is not clickable or does not direct to your public project repository.
2. Introduction (3)
   1. A clear, high-level description of what the original paper is about and what is its contribution to the wider research space. **Cite the original paper.** (2)
   2. Scope of reproducibility (1)
      1. How much of the paper were you able to reproduce?
         1. The dataset processing?
         2. The model?
         3. Baselines?
      2. **Note:** We understand that it can be difficult to reproduce research. We will grade very leniently. Just do your best and go as far as you can within limits.
3. Methodology (20)
   1. Environment (2)
      1. Python version (1)
      2. Dependencies/packages needed (1)
   2. Data (5)
      1. Data download instructions
      2. Data descriptions with helpful tables and visualizations
      3. **Please use LLMs to help in writing data preprocessing code**
         1. What was the initial prompt that you used? What was the initial output of the LLM? Validate the LLM response. How correct, relevant, and helpful was the LLM? How many prompts did you use?
   3. Model (5)
      1. Includes link to the original paper’s repo (if applicable)
      2. Includes descriptions of the model (4)
         1. Equations and their explanations (if applicable)
         2. Breakdown of inputs and outputs, techniques used, etc.
      3. Use LLMs to help with the implementation of the model used
         1. What was the initial prompt that you used? What was the initial output of the LLM? Validate the LLM response. How correct, relevant, and helpful was the LLM? How many prompts did you use?
      4. Pretrained model (if applicable)
   4. Training (4)
      1. Includes Hyperparameters (2)
         1. Report at least 3 types of hyperparameters such as learning rate, batch size, hidden size, and dropout.
      2. Includes Computational requirements (1)
         1. Report at least 3 types of requirements such as type of hardware, average runtime for each epoch, total number of trials, GPU hrs used, and # training epochs.
      3. Includes Training Details (1)
         1. Loss functions
         2. **Please use LLMs to help write code for the training loop.**
            1. What was the initial prompt that you used? What was the initial output of the LLM? Validate the LLM response. How correct, relevant, and helpful was the LLM? How many prompts did you use?
   5. Evaluation (4)
      1. **Please use LLMs to help identify and write code for metrics and evaluations.**
         1. What was the initial prompt that you used? What was the initial output of the LLM? Validate the LLM response. How correct, relevant, and helpful was the LLM? How many prompts did you use?
      2. Report the descriptions of metrics you use in the reproduction
4. Results (15)
   1. Includes a table/figures of results (5)
   2. Discuss your results with respect to the hypothesis and results from the original paper (5)
      1. Compare and contrast the results
      2. Explain why the results may be the same or different.
   3. Additional Extensions or Ablations (5)
      1. Use LLMs to help brainstorm at least **one** new extension - new dataset, new loss function, removing a part of the model (ablation study), etc.
         1. What extensions did the LLM come up with? How valid are they?
         2. What were your insights and analysis on making prompts effective for brainstorming / implementation?
      2. Use LLMs to help implement your planned extension/ablation and validate them. Include results and a discussion of the extension/ablation.
5. Discussion (8)
   1. Implications of the experimental results. (4)
      1. Is the original paper reproducible? And if it wasn’t, what factors made it irreproducible?
   2. What was easy? (1)
   3. What was difficult? (1)
   4. Recommendations to the original authors or others who work in this area for improving reproducibility (2)
6. Author Contributions (2)
   1. Please clearly mention the workload distribution between the group members.

### 