Carleton University COMP5801H/4900A: Course Project Requirements

Winter Term 2026

1 Overview

- **Requirement:** This is the Course Project outline with the requirements for the course COMP5801H/4900A - Generative AI and Large Language Models delivered during Winter 2026 at Carleton University. The project will be worth **40%** of your final grade.

- **Collaboration:** If you are a graduate student enrolled in COMP5801H, the project should be completed individually. If you are an undergraduate student enrolled in COMP4900A, this should be completed in a group of two.

- **Researching Papers:** Use **Google Scholar or DBLP** to find relevant literature. Most papers are available for free. As a student, you should not pay for downloads; if a free link is unavailable on Scholar, access it via the University's library proxy or campus internet.

2 Project Options

Option A: Empirical Evaluation

- Pick a problem that interests you.

- Implement and experiment with several Generative AI techniques to tackle this problem.

Option B: Algorithm Design

- Identify a problem for which there are no satisfying approaches.

- Develop a Generative AI technique to tackle this problem.

- Analyze performance theoretically and/or empirically.

Option C: Dataset / Simulator / Benchmark Design

- Identify a problem lacking datasets or benchmarks for Generative AI evaluation.

- Collect a dataset or design a new benchmark.

- Demonstrate baseline Generative AI algorithm performance on your dataset.

- *Reference:* See NeurIPS datasets and benchmarks track for examples.

Option D: Theoretical Analysis

- Identify a Generative AI problem or technique where properties (complexity, performance) are not well understood.

- Analyze the properties of this problem or technique.

Option E: Surveying a Generative AI Subfield

- Picking this option requires a **strong** justification.

- The project proposal is **mandatory**.

- Identify a Generative AI subfield for which there are several publications but no good survey article summarizing the key breakthroughs in that subfield.

- Provide a comprehensive survey of important publications in this field.

- Provide some ideas for new work in this area, and provide your perspectives for how you think this subfield will evolve in the near future.

- *Reference:* See IJCAI survey track for examples.

3   Proposal (optional except if you choose Option E, no mark)

- Proposal submission is optional if you pick Options A, B, C or D, but **encouraged** for the purposes of receiving feedback on relevance, feasibility, and quality.

- If you choose Option E, then the project proposal is **mandatory**.

- **Submission:** Submit electronically on Brightspace website by **Mar 1 (11:59 pm)**.

- **Length:** At most one page (excluding references).

- **Format:** Use the JMLR format: `https://www.jmlr.org/format/format.html`

- **Requirement:** Clearly state which option you have selected.

Option Specific Requirements:

Option A (Empirical evaluation):

- What is the problem?
- What Generative AI techniques do you plan to experiment with?
- Cite 4 to 8 related papers that you plan to review.

Option B (Algorithm design):

- What is the problem?

- Why are there no satisfying approaches?
- What is the intuition behind the new technique that you plan to develop?
- Cite 4 to 8 related papers that you plan to review.

Option C (Dataset/Simulator/Benchmark design):

- What is the Generative AI problem for which there is a lack of datasets or benchmarks?
- What dataset or benchmark do you plan to design?
- Cite 4 to 8 related papers that you plan to review.

Option D (Theoretical analysis):

- What is the problem or technique that you plan to analyze?
- What properties would you like to analyze/prove about this problem or technique?
- Cite 4 to 8 related papers that you plan to review.

Option E (Surveying a Generative AI subfield):

- Which sub-area of Generative AI are you proposing to survey?
- Why does the field require a survey of this topic? Why cannot a Generative AI tool prepare the survey for you? Are there any insights that you would provide that a Generative AI tool is unlikely to provide? **Justify this strongly**.
- Cite at-least 20 related papers that you plan to review.

4   Submission Files

Every submission should include a **Final Report** using the guidelines below. If you picked Option A, Option B or Option C, you should also include your **implementation files**.

5   Final Report

- **Deadline: April 8 (11:59 pm)** via Brightspace.

- **Length:** At most 8 pages (excluding references).

- **Content:** Explain the big picture and any necessary detail.

- **Format:** JMLR format.

6 Suggested Structure for the Report

Option A (Empirical evaluation):

1. **Introduction**
   - What is the problem?
   - Why is it an important problem?
2. **Techniques to tackle the problem**
   - Brief review of previous work (the 4-8 papers read)
   - Brief description of the techniques chosen and why
3. **Empirical evaluation**
   - Compare empirically for complexity, performance, ease of use, etc.
4. **Conclusion**
   - What is the best technique?
   - Is any technique good enough to declare the problem solved?
   - What future research do you recommend?

Option B (Algorithm design):

1. **Introduction**
   - What is the problem?
   - Why can't existing techniques effectively tackle this problem?
   - What is the intuition behind the new technique developed?
2. **Techniques to tackle the problem**
   - Brief review of previous work (the 4-8 papers read)
   - Describe the technique that you developed
   - Brief description of the existing techniques for comparison
3. **Evaluation**
   - Analyze and compare (empirically or theoretically) your approach
4. **Conclusion**
   - Can your new technique effectively tackle the problem?
   - What future research do you recommend?

Option C (Dataset/Benchmark):

1. **Introduction**
   - What is the problem?
   - Why aren't existing datasets/benchmarks sufficient?

2. Proposed Dataset or Benchmark

- Describe the proposed dataset or benchmark
- Describe the unique properties of the dataset

3. Evaluation

- Brief description of baseline techniques to be evaluated
- Compare baseline techniques with your dataset/benchmark

4. Conclusion

- Most important weaknesses of existing baselines highlighted
- What future research do you recommend?

Option D (Theoretical analysis):

1. Introduction

- What is the problem or technique?
- What properties did you analyze/prove?

2. Analysis

- Brief survey of previous work (the 4-8 papers read)
- Describe the analysis performed

3. Conclusion

- What have you discovered about the technique?
- What future research do you recommend?

Option E (Survey analysis):

1. Introduction

- What is the subfield of Generative AI that you are writing the survey for?
- Why is this survey article required?
- Is a Generative AI tool likely to autonomously generate such an article?
- Are there any new insights this article is providing that a Generative AI tool is unlikely to provide?

2. Survey

- Elaborate survey of different literature (at-least 20 papers).
- Your additional insights and comments on (at-least a subset of) the papers.

3. Future work

- Can you provide some new "low-hanging fruits" for research in this area?
- What new research is required in this field?
- What results are new research in this area likely to obtain?