

CS4248 Natural Language Processing

Final Group Project

Guidelines on Project Proposal

The project proposal of each group should be at most 1 page. Please provide the following information in the proposal:

Project group ID (e.g., G01, G02, etc.), group members' names (5), and the pre-defined project topic chosen.

Brief description of the approach / method / algorithm to be used.

What is the uniqueness of your proposed method (if any)? How does it compare to prior methods?

Any additional and publicly available training data to be used? Do they require any manual annotation?

Any external (supporting) code to be used?

Breakdown of work: The responsibility of each group member.

Schedule/timeline of the project.

Grading Scheme

The breakdown of the final project marks is as follows:

Content (60%): Is the work valid and technically sound? Is the quantity of work carried out sufficient for the project? How well does the implemented system perform? Is relevant work in the literature cited and compared to?

Presentation (30%): Is the work presented in a clear, logical, and coherent manner? (including both oral presentation and final project report)

Novelty (10%): Does the work carried out have any novel or unique aspects?

Your final project will be graded based on the actual work done and effort put in by your project group, and the results achieved by your group.

Materials to Submit by Project Due Date

Each final project report is in the form of an ACL conference-style short paper. The main content of the paper is limited to 4 pages, with the references taking up a maximum of 1 additional page. You must use the ACL conference paper style files available at the following URL (latex or Microsoft WORD):

<https://github.com/acl-org/acl-style-files>

The page limit must be strictly adhered to. Marks will be deducted if the main content of your paper exceeds the 4-page limit, or the references exceed 1 page, or the font size is not Times Roman 11-point font.

Each group will give a 15-minute oral presentation in a video recording. Each group member must give some part of the oral presentation. Each project group must submit its project report, presentation slides, and video recording. Each project group must also submit the source code implemented in the project, all external (supporting) code and data, and instructions on how to run their code.

You are allowed to use external open-source code or libraries in your project, but you must clearly state in your project report which parts of your project are your own contributions and implemented by your group members, and which parts are implemented by others. You must provide citations to relevant prior works and papers in your project report.

Note that closed-source large language models and API calls (like ChatGPT and GPT-4o) are not allowed in your project. Open-source pre-trained word embeddings (like word2vec) are allowed

If you have any questions about what is allowed or not allowed, please check with the course lecturer.