CS 7650: Natural Language
Project Guidelines

Instructor: Kartik Goyal
Teaching Assistants: Neelabh Sinha, Tomohiro Sawada, Om M Khare, Ishita Datta

October 25, 2024

1 Description

The course project is an attempt to enable students to work on a practical Natural Language Processing problem of their interest. It will allow students to define an NLP problem of their choice and solve it using their understanding of the topic and learning outcomes from the class.

2 General Instructions

- The project has to be an entirely new problem statement, or novel findings within existing problems that can be deemed as the original work of the team. They should be well motivated, grounded well in existing literature, diligently implemented and well presented. Reproducing or introducing small changes/enhancements to existing works will not be considered a valid project.

- It is to be done in a team of 3–4. Less-sized teams will be permitted for genuine reasons only, and will be graded on the same benchmark as the expectations with teams of size 3–4. Please reach out to the instruction team and seek a prior approval if you plan to do so. Teams greater than size of 4 are strictly prohibited.

- We all discussed in Lecture 1 about the scope of Natural Language Processing. This project has to be in that domain. A project not deemed valid will not be evaluated and 0 credits will be provided for all components.

- It carries 40% of your total grade for this course (in total).

3 Evaluation Components

The evaluation is divided into the following parts:

1. Proposal (5 pts)

2. Literature Survey and Detailed Plan (5 pts)

3. Project Progress Report 1 (5 pts)

4. Project Progress Report 2 (5 pts)

5. Final Report (12.5 pts)

6. In-class presentations (7.5 pts)

3.1   Proposal (5 pts)

The project proposal should describe your project - the problem, the motivation behind it, intended methodology, expected results and a feasibility analysis of the project.

Requirements for the proposal -
There will be Gradescope assignment where you will need to submit the following -

- Full names and GT Email ID (xxx@gatech.edu) of all team members

- Title of your Project

- A 2 page PDF (excluding references) in ACL template following the below mentioned requirements. A standard template will allow us to evaluate all groups fairly without accounting for difference in presentation format.

- The template has detailed instructions on how to use it. Before submitting the document, make sure to remove the review option to generate the final version, and add all your team members as authors.

- Only one member of the team needs to submit the Gradescope assignment. But, all the team members need to be added to the Gradescope submission to consider them for the evaluation. You can follow this guide to do the same.

Deadline: September 4, 2024, 11:59 PM EST

Components to be included in the proposal -

1. Introduction/Background - Provide an introduction to the topic and a brief literature review to establish the background.

2. Motivation - Describe the importance and exciting aspects of the problem statement of your project.

3. Problem definition and originality - Clearly define your problem statement that you are addressing, and discuss how is it different from the existing approaches.

4. Proposed Methodology - Provide an overview of the methods you are trying to use and the aspects of the project that, assuming ideal outcome, would lead to new knowledge.

5. Potential Results and Comparisons - Discuss the ways in which you will evaluate your implementation. You can mention metrics, qualitative outcomes, baselines you will be comparing your method against, etc.

6. Feasibility Analysis - Describe the most difficult aspects of the project and discuss the project's feasibility given the quarter long timeline.

Please note the following -

- Ground your proposal in existing literature. Include at least 5 references cited appropriately in your project proposal.

- You can use illustrations, tables, etc., to better convey your ideas. Appropriately refer the figures and tables that are added in the text. Please cite the source of any external illustrations, tables, etc., in your document.

- It is necessary to use the recommended ACL template. Violation of this will result in zero credits. No hand-written content in the document will be accepted. Changing any format setting of the template (like font size, column width, etc.) is also considered a violation.

- At any point, feel free to come to office hours to discuss the validity, scope, feasibility or anything else related to your project ideas.

Your project proposal will be evaluated on the following criteria -

1. Gradescope Form (1 point) - Filling the Project Title (0.5 pts) and details of your team members correctly (0.5 pts)

2. Motivation (1 point) - How clear is the motivation? Why is it important to work on the identified problem?

3. Originality/Innovation and Definition (1 point) - Is the problem clearly defined? How original and exciting is it?

4. Planning (1 point) - How well are the proposed methodologies planned and described? Are the potential results and comparisons that are defined relevant and holistic? Has the team planned well on how to approach the project through the course of the semester? Is the amount of work planned sufficient for the recommended group size of 3-4?

5. Feasibility Analysis (1 point) - How detailed is the feasibility analysis? Are most potential difficulties identified and discussed effectively? Is there some plan outlined to deal with those?

Details related to other components will be shared in due course of time. We will also share tentative project directions (similar to ones discussed in Lecture 1) and some example project ideas implemented by the students of this course with time.

3.2 Literature Review and Detailed Plan (5 pts)

This section will give you an opportunity to dive deeper into the existing research relevant to your project and formulate a detailed plan for the remainder of the project timeline. The goal is to ensure that the proposed methodology is well grounded in existing literature and that you have clear, actionable steps to proceed.

Requirements for Submission:

- A 2 page PDF (excluding references) on top of your existing proposal document in the ACL template. So the overall document will be proposal length + 2 pages + references.

- Only one team member should submit, but ensure all team members are added to the submission. You can reference the guide shared in the project proposal guidelines on how to add members to a Gradescope submission.

- In your PDF, near the authors section, add the group number for your group assigned by the instructional team.

- A document extending the proposal with literature review and detailed plan as described below.

- Clearly labeled pages in Gradescope for each section.

- All group members added to the Gradescope submission.

- Write your own content. Refrain from copy-pasting text from existing works or from AI text generators. Any project flagged for any of the two will be penalized severely.

Deadline: September 19, 2024, 11:59 PM EST

Components to include in your document:

1. **Proposal Amendments (optional)**:

    - Incorporate any feedback provided by the instructor in your initial project proposal. The revised proposal should highlight changes made to improve the project scope, methodology, or problem definition, based on the feedback.

    - Add a small paragraph later highlighting what changes you made to the earlier proposal document.

    - If you think there isn't sufficient time to incorporate this after the proposal feedback release and before the deadline of this component, you can skip this. Discuss it later with your TA coordinator.

2. **Expanded Literature Review**:

    - Provide a detailed review of all key papers and works relevant to your project. Each team should critically analyze the research literature, focusing on works related to the methodology, datasets, or models you're using.

- Discuss at least 8 related work (cited appropriately), approximately 2 papers per team member.

- Recommended conferences to refer are ACL, EMNLP, NAACL, EACL, COLING, NeurIPS, ICLR, ICML. But, feel free to explore outside this. We have only provided this to give a starting point if you are looking for one, you are not bound to it. Also look at other conferences if you work falls at the intersection of NLP and any other application (like vision).

- Don't just summarize the existing works. Critically analyze the existing works, discuss limitations in those and specify how your project will aim to address those, or why your project does not cover that scope.

3. **Detailed Plan, Timeline and Milestones**:

- Provide a detailed timeline with weekly or bi-weekly objectives for your project, from the current stage to the final submission. Week 6 is the first week after the submission of this plan. So, the timeline should start with Week 6 till the end, i.e., Week 15. The timeline should reflect realistic expectations.

- For each milestone, indicate the expected outcome.

- An example timeline is provided in Table 1.

- For a submission milestone (checkpoint 1, 2, final), you are expected to complete everything in the timeline up to the week before that submission milestone. For example, checkpoint 1 falls in Week 8, so we expect everything in Week 6 and 7 completed and added to the checkpoint report.

4. **Contribution Table (optional)**:

- A table describing the contribution of each team member towards this component of the submission.

- This is optional. If the team mutually agree that the contribution is equivalent by each team member, they can skip this. Marginal imbalances are acceptable.

- No points deducted if the contributions are equivalent and mutually agreed. However, we might penalize some team members individually if their contribution is not significant.

Your deliverables for the future checkpoints will be the objectives outlined till the week previous to the week when the checkpoints are due.

Week	Objectives	Expected Outcome
6	Objective 1 description	Outcome 1 description
	Objective 2 description	Outcome 2 description
7	Objective 1 description	Outcome 1 description
	Objective 2 description	Outcome 2 description
...	...	...
15	Objective 1 description	Outcome 1 description
	Objective 2 description	Outcome 2 description

Table 1: Example Project Timeline (consider Week 6 to be the week of Sep 23, consistent with the course page)

Grading Criteria:

- Literature Review Quality (3 pts): How well does the literature review connect to your problem statement and proposed methodology? Is it comprehensive and critical, or does it simply describe references? Are there discussions on how the existing work deals with the described limitations of existing work?

- Plan (2 pts): Does the timeline reflect a clear understanding of the project requirements, responsibilities and workload? Are the milestones realistic and achievable within the given timeline?

3.3   Project Progress Report 1 (5 pts)

The first progress report is intended to ensure that your team is on track with the project and making steady progress. This report will provide a detailed update on the work completed so far, the methods implemented, and any challenges encountered.

Requirements for the submission:

- A 1-2 page PDF (excluding references) following the ACL template, building upon your proposal and literature review submissions (total PDF = proposal length + literature review length + 1-2 pages) on Gradescope.

- Only one team member should submit the report to Gradescope, but ensure that all team members are added to the submission.

- Remember, your document will be evaluated on everything included upto the week prior to the submission of progress report.

Deadline: October 9, 2024, 11:59 PM EST

Components to include in the document:

- **Progress Update**:

  - Provide a detailed update on what you planned to achieve at the beginning of this phase as per the original plan submitted, and what you have accomplished so far (up till the week prior to the deadline of this document).
  - Describe what was your plan before this phase, and where currently stands.

- **Method Description**:

  - Describe precisely and in-depth the methodologies implemented, dataset analysis, experiments conducted so far.
  - Describe any difficulties, roadblocks, or unexpected challenges encountered so far.
  - Discuss the steps taken or being considered to take in order to overcome these challenges. Also, discuss why you believe these steps will solve the issues.

- **Experimental Results and Analysis**:

  - Initial results and discussions which demonstrate that your project is progressing as expected towards an expected solution.
  - Use tables, figures, or plots (if applicable) to present intermediate results, comparisons, or visualizations of your progress.
  - Discuss your results and what you infer from those. Also outline if it was what you expected, or if different, how does it place you in the larger context of the project.

- **Next Steps**:

  - A detailed overview of steps you are planning to achieve by the second checkpoint. This needs to be more concrete and well-defined than those provided in the initial plan.
  - Indicate any changes in the milestones or deadlines from your original plan and describe why those adjustments were made.

- **Contribution Table (optional)**:

  - A table describing the contribution of each team member towards this component of the submission.
  - This is optional. If the team mutually agrees that the contribution is equivalent by each team member, they can skip this. Marginal imbalances are acceptable.
  - No points will be deducted if the contributions are equivalent and mutually agreed. However, we might penalize some team members individually if their contribution is not significant.

Grading Criteria:

- Progress (1pt): Has the team accomplished everything according to the plan so far? If not, are the reasons justifiable?

- Method Description (1.5 pts): How clear and precise are the methods that have been implemented so far described? Are any challenges encountered so far (if any) described properly? Does the team discuss how they addressed it, or how they are planning to do that in future?

- Experimental Results and Analysis (1.5 pts): How are the results outlined and discussed? What does the team infer from those results? Have they highlighted any useful insights which help them determine where they stand presently in relation to the project goal?

- Next Steps (1 pt): Does the team have a clear path forward (at least up till the next checkpoint) based on the progress so far? How clearly and precisely have they defined that path?

3.4   Project Progress Report 2 (5 pts)

The second progress report is intended to ensure that your team has made substantial progress and has adhered to the plan of action you provided. This report will provide a detailed update on the work completed so far, the methods implemented, the initial observations, and any challenges encountered.

Requirements for the submission:

- A 1-2 page PDF (excluding references) following the ACL template, building upon your proposal, literature review and progress report 1 submissions (total PDF = proposal length + literature review length + progress report 1 length + 1-2 pages) on Gradescope.

- Only one team member should submit the report to Gradescope, but ensure that all team members are added to the submission.

- Remember, your document will be evaluated on everything included upto the week prior to the submission of progress report.

Deadline: November 4, 2024, 11:59 PM EST

Components to include in the document:

- **Progress Update and adherence to plan**:

  - Provide a detailed update on what you planned to achieve at the beginning of this phase as per the original plan submitted, and what you have accomplished so far (up till the week prior to the deadline of this document).
  - Report if your group was able to stick to the proposed plan of action. If not, then explicitly state what went wrong if you were not able to adhere to the proposed plan of action. Remember, you will not lose marks if you provide a convincing reason as to why your group deviated from the plan of action.

- **Method Description**:

  - Describe precisely and in-depth the methodologies implemented, dataset analysis, experiments conducted so far.
  - Describe any difficulties, roadblocks, or unexpected challenges encountered so far.
  - Discuss the steps taken or being considered to take in order to overcome these challenges. Also, discuss why you believe these steps will solve the issues.

- **Experimental Results and Analysis**:

  - Initial results and discussions which demonstrate that your project is progressing as expected towards an expected solution.
  - Use tables, figures, or plots (if applicable) to present intermediate results, comparisons, or visualizations of your progress.
  - Discuss your results and what you infer from those. Also outline if it was what you expected, or if different, how does it place you in the larger context of the project.

- **Next Steps**:

    - A detailed overview of steps you are planning to achieve by the second checkpoint. This needs to be more concrete and well-defined than those provided in the initial plan.

    - Indicate any changes in the milestones or deadlines from your original plan and describe why those adjustments were made.

- **Contribution Table (optional)**:

    - A table describing the contribution of each team member towards this component of the submission.

    - This is optional. If the team mutually agrees that the contribution is equivalent by each team member, they can skip this. Marginal imbalances are acceptable.

    - No points will be deducted if the contributions are equivalent and mutually agreed. However, we might penalize some team members individually if their contribution is not significant.

Grading Criteria:

- Progress (1pt): Has the team accomplished everything according to the plan so far? If not, are the reasons justifiable?

- Method Description (1.5 pts): How clear and precise are the methods that have been implemented so far described? Are any challenges encountered so far (if any) described properly? Does the team discuss how they addressed it, or how they are planning to do that in future?

- Experimental Results and Analysis (1.5 pts): How are the results outlined and discussed? What does the team infer from those results? Have they highlighted any useful insights which help them determine where they stand presently in relation to the project goal?

- Next Steps (1 pt): Does the team have a clear path forward (at least up till the next checkpoint) based on the progress so far? How clearly and precisely have they defined that path?