Complete Instructions for AIE241 NLP Final Course Project

Project Overview

In this project, your **team (out of 4 students)** will select a Natural Language Processing (NLP) task inspired by projects from Stanford's **CS224N** course (2024). You will improve upon the original approach by proposing and implementing an innovative custom method. Your goal is to enhance the original results, using the same dataset for a fair comparison. The project is worth **10 marks** and will be evaluated based on innovation, implementation, analysis, report quality, and teamwork.

Project Selection and Innovation

- **Objective**: Choose an NLP project from CS224N (e.g., machine translation, sentiment analysis, text classification) and propose an innovative improvement to the original method. Your innovation should aim to enhance performance, efficiency, or applicability.
- **Dataset**: Use the same dataset as the original project to ensure a direct comparison between the baseline (original method) and your custom approach.
- **Inspiration**: Refer to sample reports from ***CS224N 2024 Projects.***
- **Consultation**: Discuss your idea and proposed innovation with TAs to ensure suitability.

Report Requirements

- **Length**: 6-8 pages (excluding references)
- **Format**: Use the provided LaTeX template (Overleaf link). Points will be deducted for using a different template.
- **Style**: Write as an NLP research paper, understandable to peers.
- **Sections**:
    1. **Key Information** (Required):
        - Project title
        - Team member names and emails
        - Original CS224N project reference
    2. **Abstract** (Required):
        - <300 words summarizing the problem, your approach, and key findings
    3. **Introduction** (Required):
        - Explain the problem, its significance, and your innovation
    4. **Related Work** (Required):
        - Contextualize your work within existing research
    5. **Approach**:
        - Detail your custom method and how it differs from the baseline
    6. **Experiments**:
        - **Data**: Describe the dataset (same as original)
        - **Evaluation**: Use the same metrics as the original for comparison
        - **Details**: Report setup for baseline and your approach

    - **Results**: Compare your results to the baseline
7. **Analysis**:
    - Provide qualitative insights (e.g., error analysis, visualizations)
8. **Conclusion** (Required):
    - Summarize findings, limitations, and future work
9. **Team Contributions** (Required):
    - Describe each member's role (1-2 sentences per person)
10. **References** (Required):
    - Use BibTeX
11. **Optional Appendix**: Extra details (not graded)

Submission Instructions for Your Project

Your submission must include two main components: a **project report** and a **code zip file**. Follow the guidelines below carefully to prepare and submit your work correctly.

What to Submit

1. **Project Report**:
    - Submit a PDF file created using the provided LaTeX template.
    - The report should clearly present your problem, solution, and supporting details as outlined in the assignment requirements.
    - **File Name**: Name your file in the format ***TeamName_ProjectReport.pdf*** (e.g., ***TeamBeta_ProjectReport.pdf***).
2. **Code**:
    - Submit a zip file containing all the code you wrote or adapted for the assignment.
    - **Do not include**: Datasets, model checkpoints, or pre-existing libraries/packages.
    - **File Name**: Name your file in the format TeamName_Code.zip (e.g., TeamBeta_Code.zip).
    - **File Size Limit**: The zip file must not exceed 1MB.

How to Submit

- **Team Submission**: Only **one submission per team** is needed. Make sure all team members are tagged in the submission to receive credit.

Important Notes

- **Deadline**: Submit by **May 26, 2025**
- **Late Submission Policy**:
    - Late submissions will receive a **60% deduction** unless you provide official documentation (e.g., medical or university-approved reasons).
    - Submissions that are more than **one week late** will receive **zero marks**.

- **File Naming**: Use the specified naming conventions to avoid confusion during grading.
- **Code Purpose**: The code is collected for reference (e.g., to verify originality) and will not be directly graded unless specified otherwise.

Checklist Before Submission

- ☐ The project report is in PDF format and follows the provided LaTeX template.
- ☐ The report is named correctly: ***TeamName_ProjectReport.pdf.***
- ☐ Code is zipped and named correctly: TeamName_Code.zip.
- ☐ The code zip file excludes datasets, checkpoints, and pre-existing packages.
- ☐ All team members are tagged in the submission.

Grading Policy (10 Marks Total)

- **Innovation and Originality (3 marks)**:
  - 3/3: Highly novel, well-justified improvement
  - 2/3: Solid but less groundbreaking
  - 1/3: Minimal innovation
  - 0/3: No significant contribution
- **Implementation and Experimentation (3 marks)**:
  - 3/3: Flawless implementation and comparison
  - 2/3: Minor errors
  - 1/3: Significant issues
  - 0/3: No functional results
- **Analysis and Insight (2 marks)**:
  - 2/2: Deep, clear analysis
  - 1/2: Basic analysis
  - 0/2: Minimal analysis
- **Report Quality (1 mark)**:
  - 1/1: Clear, professional writing
  - 0.5/1: Readable but flawed
  - 0/1: Poorly written
- **Team Contributions (1 mark)**:
  - 1/1: Clear, equitable roles
  - 0.5/1: Uneven or vague
  - 0/1: Missing details
- **Holistic Evaluation**: Overall quality may adjust final scores.

Publication Opportunity

Top-performing teams with exceptional innovation and results may be selected to refine their work for submission to a top-tier conference like **ACL**. This is a rare opportunity for undergraduates to gain research recognition.

- **Process**:
  - Submit your initial report by the deadline.
  - Top teams will be mentored to iterate and submit to a conference during the summer.
- **Goal**: Publish your work with guidance, a significant achievement for undergraduates.

This project combines technical rigor with creativity, offering a chance to contribute to cutting-edge NLP research. Aim high, your innovation could make a real impact!

Best of Luck!