

# Project Proposal

CS 490: Natural Language Processing · Spring 2026

Due on: 03/27/2026 @ 11:59 PM (AoE)

## Assignment Instructions

Your group (2–4 people) must submit a project proposal of at most **4 double-column pages** detailing your final project idea. Please include all member names in the document. Use the ACL template and submit via Gradescope, ensuring all group members are added to the submission.

**Goal.** Your proposal must demonstrate that the project is (i) well-motivated and well-defined, and (ii) **feasible within the course timeframe and available resources**. Avoid planning for extensive data collection, and do not propose pretraining large models from scratch.

Your proposal must include:

- **Introduction:** Define the problem and clearly explain the motivation for tackling it.
- **Related Work:** Cite the papers that describe the specific methods or architectures you plan to apply, as well as the baseline models you intend to compare against.
- **Proposed Approach:** Detail your core modeling idea. Highlight what is novel compared to existing baselines and what insights you expect to gain.
- **Experimental Plan:** Describe your setup, including:
  - **Datasets:** Which datasets will you use? Provide citations and license constraints. If annotating data, specify your pipeline, estimated size, estimated person-hours, and a contingency plan. *Keep manual annotation minimal.*
  - **Baselines & Metrics:** What are your baselines, and how will you evaluate success?
  - **Training Details:** Expected hyperparameters (batch size, epochs, etc.) and tooling (PyTorch, HuggingFace, etc.). You may provide an initial estimate of the hyperparameters since you may need to tune them to find the best values.
- **Compute & Resource Estimate (Required):**
  - **Hardware:** Specify your setup (e.g., CPU, GPU, RAM information).
  - **Budget:** Estimate total GPU/CPU hours (number of runs × runtime per run).
  - **Scale constraints:** Specify the model size (e.g., fine-tuning a 1B parameter model) and justify its feasibility. **Pretraining large foundation models is out of scope.**
  - **Resource plan:** Where will you run this (personal machine, free tiers, RCAC cluster)?

## Project Scope

A successful project must be appropriately scoped for the timeline. Generally, this includes a **clear problem to solve**, at least one **strong baseline**, a **reproducible experiment or well-documented engineering build** using readily available data, and a **meaningful extension** (e.g., new analysis, modeling idea, robustness evaluation, or a functional application).

### Examples of appropriate scope:

- Build a conversational agent that solves a specific problem, such as travel planning, via chat-based interaction with human users.
- Fine-tune a small LLM to evaluate indirect meaning or pragmatics in dialogue, comparing targeted prompting strategies against strong baselines.
- Compare standard models on culturally grounded datasets (e.g., MPDD, MELD) to analyze robustness, domain shift, or social biases.
- Build a small, focused dataset (e.g., 200–1,000 annotated examples of social context in text) to test a specific hypothesis, running controlled experiments on model failures.

### Examples to avoid (too large / unfeasible):

- Scraping or annotating millions of examples without a concrete timeline.
- Pretraining multi-billion parameter language models from scratch.
- Relying on proprietary data or compute without guaranteed access.

## Project Tracks

You may choose to:

- **Develop something novel:** e.g., fine-tune a model on a newly collected, highly focused dataset.
- **Adapt existing work:** e.g., apply a known architecture to a new domain or modify its training paradigm. *Note: This requires a meaningful extension beyond the original paper to receive full credit.*

## Grading Rubric (10 points total)

- **(3 pts)** Is the problem you aim to solve well-defined?
- **(2 pts)** Is the project well-motivated?
- **(3 pts)** Are the proposed methods reasonable and **feasible**? Are the dataset and compute plans clear?
- **(2 pts)** Is the proposal well-written and formatted?

**Late day policy.** You may use late days for the proposal. For group submissions, we do not sum the late days; rather, we take the maximum days late for the assignment itself. For example, a proposal submitted 2 days late deducts exactly 2 late days from each group member’s allowance. If *any* member has insufficient late days, the standard late penalty applies to the group submission.