Project Overview

This project focuses on practical work with LLMs. You may work individually or in groups of up to three. Each group will select an NLP task, clearly define the project objective, prepare an appropriate dataset for use with LLMs, and apply one or more LLM adaptation techniques. The project requires completing the full dataflow: dataset preparation, tokenizer and embedding preparation, model adaptation, evaluation and reporting. The emphasis is on understanding how LLMs are adapted and applied in practice.

As part of the project, you are required to compare multiple adaptation strategies for a single LLM *or* compare two or more LLMs using the same adaptation technique. Adaptation methods may include, for example:

- Full fine-tuning/Instruction-tuning
- Parameter-Efficient Fine-Tuning (PEFT) techniques such as LoRA, Adapters, Prefix-Tuning, BitFit
- Prompt Tuning (soft prompts)
- Prompt Engineering (instruction prompts, few-shot prompting, structured prompts)

Your comparison should highlight performance differences, efficiency trade-offs (memory, training time, number of trainable parameters), and practical considerations rather than focusing solely on raw accuracy.

Note: You are not allowed to reuse the same task or project topic from any of your other courses.

Recommended Tasks

You may choose *any* NLP task (Except tasks used in your other courses). Examples include:

Classification Tasks

- Sentiment or emotion classification
- Hate-speech or toxicity detection
- Fake news or misinformation detection

Generative & Reasoning Tasks

- Summarization
- Dialogue or response generation
- Hallucination detection or truthfulness evaluation

Information Extraction / QA

- Named entity extraction
- Question answering (extractive or closed-book)

You are encouraged to work on a low-resource language.

You may also propose other tasks of their choice, as long as they fit the learning objectives of the course and are approved by the instructor.

Expected Deliverables

Your project submission consists of three required components:

A. Code Repository

Your repository must include:

- Data preparation and embeddings
- Training or fine-tuning scripts
- PEFT or prompt-tuning implementation
- Evaluation scripts
- requirements.txt or environment file
- A clear README with instructions to run the project

Your code should be clean, documented, and reproducible.

B. Written Report

Your report must include:

1. Introduction & Dataset

- Description of the selected NLP task
- Dataset source, preparation steps, basic statistics

2. Model & Adaptation Method

- LLM(s) chosen, with a short description
- Adaptation techniques used and justification

3. Training / Fine-Tuning Process

- Implementation and training details (hyperparameters, preprocessing, embeddings preparation, hardware used)

4. Evaluation

- Quantitative metrics and qualitative examples

5. Comparative Analysis

- Compare:
  - Different adaptation methods on the same LLM, or
  - Different LLMs using the same/different adaptation methods
  - Your results with related benchmark papers or baselines if exists
- Provide a clear justification for the observed performance differences, strengths, and limitations

---

C. Final Presentation

You must deliver a 15-minute presentation summarizing:

- The task and dataset
- LLM(s) and adaptation method(s) used
- Key implementation decisions
- Evaluation results
- Comparative analysis
- Main insights, limitations, and conclusions