NLP Course Project

Objective

This project aims to encourage you to explore, implement, and evaluate state-of-the-art NLP technologies, including but not limited to Large Language Models (LLMs), transformers, retrieval-augmented generation (RAG), multimodal models, and other recent advancements in NLP.

You must work in teams of **2-3 members**.

Project Guidelines

The project **must include one** of the following:

1. **Research-Oriented Work** – A project that investigates a novel idea, technique, or approach in NLP, possibly supported by experiments, comparisons, or analysis.
2. **Problem-Solving Approach** – A project that identifies a real-world problem or gap and proposes a solution using NLP methods. The solution should demonstrate practical impact.

Project Scope

You may select your own topic within the domain of NLP, but the project must focus on modern techniques such as:

- **Large Language Models (LLMs):** Fine-tuning, prompt engineering, adaptation to new tasks, or efficiency improvements.
- **Text Generation & Summarization:** Abstract and extractive methods, controllable text generation, factual consistency.
- **Conversational AI:** Dialogue systems, chatbot development, reinforcement learningbased dialogue agents.
- **Information Retrieval & Question Answering:** Open-domain QA, retrieval-augmented generation (RAG), semantic search.
- **Multimodal NLP:** NLP models that integrate text with images, video, or audio.
- **Low-Resource & Multilingual NLP:** Training or fine-tuning models for underrepresented languages (e.g. Urdu).
- **Efficient NLP:** Model compression, distillation, and deployment of large models on resource-constrained devices.

Deliverables & Timeline

1. **Project Proposal with group members names (14 September 2025) – 5%**
- A 1-2 page proposal outlining:

- Research problem and motivation
- NLP techniques to be used
- Datasets and tools to be used
- Evaluation metrics
- **Approval is required before proceeding.**

2. **Literature Review (28 September 2025) – 10%**
- Summary of recent research (10+ papers) related to the chosen topic. Use Google Scholar to search for papers from high-quality journals and conferences.
- Key findings and gaps in existing work.

3. **Final Project Report with Code submission (7 November 2025) – 85%**
- Comprehensive report including:
- Abstract
- Introduction & problem statement
- Literature review
- Methodology
- Implementation details
- Results & discussion
- Conclusion & future work

   Project Evaluation
- Live demonstration of results.

Evaluation Criteria
- **Originality & Innovation (20%)** – Novelty of the idea, uniqueness of approach.
- **Technical Depth (25%)** – Complexity of the implementation, correctness of techniques used.
- **Implementation & Results (30%)** – Performance of the system, effectiveness of the proposed method.
- **Report Quality (15%)** – Clarity, organization, and completeness of the final report.

Tools & Resources
You are encouraged to use:
- **LLMs & APIs:** OpenAI, Hugging Face Transformers, LLaMA, etc.
- **Frameworks:** PyTorch, TensorFlow, LangChain, LangGraph, Sentence Transformers, etc.
- **Datasets:** Hugging Face Datasets, custom datasets, Google Dataset search etc.
- **Evaluation Metrics:** BLEU, ROUGE, perplexity, accuracy, F1-score.

Submission Guidelines

All deliverables should be submitted on Google classroom