

Title: Job Match Genie

Overall Goal:

The goal of this project is to develop an NLP-based system that helps students streamline their job application process by automating key tasks. The system will classify resumes into job categories, generate personalized cover letters, suggest relevant interview questions, and optionally simulate mock interviews. The value for students is a more efficient, personalized, and supportive job application experience, saving time while helping them prepare effectively for interviews and improve their chances of landing the job.

Scope:

We aim to develop a limited prototype of the system, focusing on:

- Classifying resumes into job categories.
- Cover Letter Generation: Automatically generating personalized cover letters.
- Interview Question Suggestion: Providing tailored interview questions based on the resume.
- The project will be developed as a scalable prototype, with the core functionalities focused on fine-tuning small LLMs for each NLP task.
- Mock Interview (Optional): Simulating mock interviews to provide feedback.

Team Members and NLP Tasks:

VXD240002 - Dacharla Venkata Abhiram

IXB240007 - Indrani Borra

KXP240029 - Krishna Tejaswini Paleti

SXP2400449 - Sai Charan Palvai

Task 1: Integrating NLP methods for resume classification using fine-tuned LLMs (e.g., DistilBERT or MiniLM). This involves model fine-tuning, preprocessing resumes, and integrating the classification pipeline.

Task 2: Implementing cover letter generation using a fine-tuned LLM (e.g., T5 Small or DistilGPT-2). The task will include designing prompts, fine-tuning the model with Kaggle cover letter datasets, and generating personalized content based on resume input.

Task 3: Developing the interview question suggestion feature. The team member will extract key skills and job roles from resumes and design a prompt-based approach to generate relevant questions for interviews based on the resume data.

Task 4: Optional mock interview simulation. This member will integrate the mock interview feature, which involves generating responses based on resume content and evaluating candidate answers using NLP techniques like response relevance scoring.

Data Sources and Approaches:

Data Sources: Kaggle datasets of resumes, cover letters, and interview questions. Additional data may be gathered from LinkedIn or Glassdoor for augmenting the training dataset.

Approach: Fine-tuning small LLMs such as DistilBERT, MiniLM, T5 Small, and DistilGPT-2 for each specific task. Zero-shot learning will be employed for resume classification, and prompt-based learning will be used for generating cover letters and interview questions.

Conclusion:

This project aims to automate and improve the mentioned tasks using efficient NLP techniques. By streamlining the resume classification process, generating cover letters, and providing relevant interview questions, the system will save time for recruiters and enhance the hiring experience for candidates. Each team member will focus on a distinct NLP task to ensure a well-rounded, efficient development process.