Project Synopsis: AI Resume Analyzer Using RoBERTa

# Project Title

AI-Based Resume Analyzer and Job Description Matcher using RoBERTa and Flask

# Objective

To develop an AI-powered system that analyzes a user’s resume and compares it with a given job description (JD) using semantic similarity. The system provides a match score, identifies missing skills, and suggests improvements — helping job seekers tailor their resumes effectively.

# Problem Statement

Many candidates submit generic resumes without tailoring them to specific job descriptions, reducing their chances of selection. Recruiters use ATS and AI tools to screen resumes based on relevance. A tool that mimics this process can help users prepare better resumes aligned with the job role.

# Features

- Upload resume (PDF)

- Paste job description (JD)

- AI-generated match score (semantic analysis)

- Skill gap analysis (matched and missing skills)

- Smart suggestions to improve resume content

- Model overview with accuracy details

- Modern, responsive web UI (Tailwind CSS)

- Semantic matching using RoBERTa NLP model

- Named Entity Recognition (NER) for skill extraction

# Technology Stack

Frontend: HTML, Tailwind CSS, Jinja Templates

Backend: Python, Flask

PDF Parsing: PyMuPDF

NLP Models: SentenceTransformers (stsb-roberta-large)

Entity Extraction: HuggingFace Transformers (NER Model)

# Methodology

1. Extract resume text using PyMuPDF

2. Convert resume and JD into vector embeddings using RoBERTa

3. Calculate cosine similarity to get match score

4. Extract skills using NER

5. Compare JD keywords with resume skills

6. Provide feedback & suggestions

# Expected Outcomes

- Accurate match score between resume and job description

- Realistic skill gap identification

- Enhanced resume quality for users

- Project-ready code base for final year or AI resume tools

# Target Users

- Final year students

- Career services teams

- Job seekers & professionals

- HR/Recruiting platforms

# Future Scope

- Multi-JD ranking system

- PDF report generation

- Resume section classification

- Integration with LLMs like GPT for rewriting suggestions