AI-Powered Resume Screening Tool

Project Overview

This project is an intelligent, AI-powered resume screening tool designed to automate and improve the recruitment process by ranking candidate resumes based on how well they match a given job description. Leveraging advanced Natural Language Processing (NLP) techniques, it extracts and analyzes text from resumes and compares them semantically with the target job description to provide meaningful similarity scores.

Features

- PDF Resume Text Extraction: Uses pdfplumber for robust extraction of text content from PDF resumes, handling various layouts and formats.

- Semantic Similarity Matching: Utilizes SentenceTransformers (BERT embeddings) to compute meaning-aware similarity between resumes and job descriptions, capturing synonyms and related concepts beyond exact keyword matches.

- Interactive Web Interface: Built with Streamlit, allowing users to upload multiple resumes and paste a job description to instantly see ranked candidate matches with detailed scores.

- Downloadable Results: Option to export resume rankings as a CSV file for recruiter convenience.

- Cleaned and Normalized Text Processing: Ensures accurate comparisons through preprocessing steps like lowercasing and removal of special characters.

Tech Stack

- Python 3

- pdfplumber (PDF text extraction)

- sentence-transformers (semantic NLP embeddings)

- pandas (data handling)

- scikit-learn (for optional traditional ML methods)

- Streamlit (interactive web app)

How It Works

1. Upload multiple resume PDF files.

2. Paste or input a job description in the provided text box.

3. The tool extracts text from each resume, cleans and preprocesses it.

4. Both job description and resumes are transformed into semantic vector embeddings.

5. Cosine similarity scores are calculated between the job description and each resume.

6. Resumes are ranked and displayed in descending order of matching scores.

7. Users can download the ranking results as a CSV file.

Getting Started

Prerequisites

- Python 3.x installed
- Install required Python packages:

bash

pip install pdfplumber sentence-transformers pandas scikit-learn streamlit

Running the Web App

1. Clone or download the repo.
2. Navigate to the project folder in terminal.
3. Run the Streamlit app:

bash

streamlit run app.py

4. Open your browser at the URL shown (typically http://localhost:8501).
5. Upload resumes and enter the job description.
6. View ranked results instantly.

Project Benefits

- Demonstrates practical data science skills in NLP, embedding techniques, and user interface design.
- Provides an end-to-end pipeline from data extraction, processing, matching to visualization.
- Makes a professional project suitable for internship/job portfolios focused on recruitment, HR tech, and ML applications.

Future Enhancements

- Add keyword/skill extraction and highlighting to explain match scores.
- Support for other resume file formats (DOCX, TXT).
- Cloud deployment for online access and sharing.
- Integration with applicant tracking systems (ATS).
- Advanced UI features like resume previews and candidate profiles.

About

Name: Dhruv Meena

Email: dhruv123meena@gmail.com

Institute : Indian Institute of Technology Patna