

AI-Powered Resume Ranker – Project Report

Introduction

In today's fast-paced and highly competitive job market, recruiters are often overwhelmed by the volume of resumes received for a single job posting. Manually reviewing each resume is not only inefficient and time-consuming but also subject to human bias and inconsistency. To solve this problem, the AI-Powered Resume Ranker was developed—a web-based intelligent screening system that automates the evaluation and ranking of resumes against a job description. This system leverages Natural Language Processing (NLP) and Machine Learning (ML) to ensure fair, consistent, and data-driven shortlisting of candidates. It is designed to support HR professionals and recruitment teams in accelerating the hiring process and enhancing selection accuracy.

Abstract

This project focuses on transforming the traditional resume screening process through intelligent automation. The AI-Powered Resume Ranker uses a combination of NLP techniques and scoring algorithms to assess how well a candidate's resume matches a given job description. The evaluation process considers several key factors:

- TF-IDF similarity between the resume and job description
- Keyword presence and partial matching
- Work experience extraction via date recognition
- Academic qualifications and institution identification
- Detection of both technical and soft skill.

The tool is built with a user-friendly Flask web interface where HR personnel can upload resumes in PDF format and input the job description. The system then generates a ranked list of candidates based on a composite score and provides an Excel report summarizing the analysis. Additionally, the backend architecture is designed to support API integration with external HR or ATS platforms, enabling seamless integration into existing recruitment workflows.

System Workflow

1. Text Extraction and Preprocessing

Resumes are processed using PyPDF2 to extract raw text. This text is then cleaned and normalized using SpaCy, which includes steps like lowercasing, removal of punctuation and stop words, and lemmatization. The same preprocessing is applied to the job description to ensure consistency.

2. Scoring Mechanism

Each resume is scored across five dimensions:

- TF-IDF Similarity (25%): Measures cosine similarity between the job description and the resume content.
- Keyword Matching (20%): Detects important keywords from the job description and searches for exact or approximate matches in resumes using fuzzy logic.
- Experience Scoring (20%): Identifies date ranges and calculates total years of professional experience.
- Education Scoring (15%): Looks for degree names and prominent institution names to evaluate academic background.

- Skill Scoring (20%): Compares listed skills in the resume with a curated set of required skills from the job description.

All scores are aggregated into a final composite score, which is used to rank the resumes in descending order of relevance.

Technologies and Tools

The application stack includes:

- Python – Core programming language for data handling and backend logic
- Flask – Lightweight web framework for building the interactive web interface
- PyPDF2 – For extracting text from PDF files
- SpaCy – Used for advanced NLP preprocessing
- Scikit-learn – For TF-IDF vectorization and similarity measurement
- FuzzyWuzzy – Enables approximate keyword matching using fuzzy logic
- Pandas – Efficient handling of structured resume data
- XlsxWriter – Generates downloadable Excel reports for HR teams

Key Features

- Upload and analyze multiple PDF resumes at once
- Input job descriptions directly on the web app
- Intelligent keyword and skill extraction using NLP
- Multi-factor scoring based on relevance, experience, education, and skills
- Real-time ranking of resumes with transparency in scores
- Clean and responsive web interface for easy usability
- Excel report generation with detailed candidate rankings
- API-ready backend for future integration with ATS systems.

Conclusion

The AI-Powered Resume Ranker introduces an intelligent and scalable solution to one of the most repetitive tasks in recruitment—initial resume screening. By automating the shortlisting process with data-driven techniques, it ensures every resume is evaluated fairly and efficiently. The system significantly reduces recruiter workload, improves shortlisting accuracy, and enhances the overall hiring experience.