

AI-powered Resume Screening and Ranking System

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning

with

TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Abhijeet Kumar

abhijeetkumarorlaha620@gmail.com

Under the Guidance of

Saomya Choudhary

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

I would like to express my sincere gratitude to **Ms. Saomya Choudhary and P raja sir** for her invaluable guidance, support, and mentorship throughout the development of this project. Her insightful feedback and encouragement played a crucial role in shaping the **AI-powered Resume Screening and Ranking System**.

I also extend my heartfelt appreciation to the **Edunet Foundation** and **AICTE** for providing the platform, resources, and learning opportunities that made this project possible. Their continuous support and commitment to fostering innovation have been instrumental in enhancing my knowledge and technical skills.

This project would not have been possible without the collective efforts and guidance from my mentor and the organizations that nurtured my learning journey. Thank you for your unwavering support and inspiration.

ABSTRACT

Recruiters often spend significant time manually screening resumes, which can be inefficient and prone to biases. This project addresses the challenges associated with traditional resume screening processes by developing an AI-powered system for automated resume screening and candidate ranking. The manual screening of resumes is often a time-consuming, resource-intensive, and subjective process, susceptible to human biases and inconsistencies. These inefficiencies can lead to delays in the hiring process, increased costs, and the potential for overlooking highly qualified candidates. The primary objective of this project was to design and implement a system that leverages the capabilities of artificial intelligence, specifically natural language processing (NLP), to automate the initial screening phase, thereby improving the speed, accuracy, and objectivity of candidate evaluation. The system was developed using Python and the Streamlit framework, providing a user-friendly and accessible solution. The **AI-powered Resume Screening and Ranking System** aims to streamline this process by leveraging **Python** and **Streamlit** to automate resume evaluation based on job descriptions.

The system extracts text from uploaded **PDF resumes** using **PyPDF2** and processes them using **TF-IDF vectorization** to quantify textual similarity. The job description and resumes are transformed into numerical vectors, and cosine similarity is applied to measure relevance. The resumes are then ranked in descending order based on their similarity scores, providing a structured and objective evaluation.

The application features an intuitive **Streamlit** interface, where users can input a job description, upload multiple resumes, and receive ranked results in real time. The approach ensures a faster, unbiased, and data-driven hiring process.

Results demonstrate that the system effectively prioritizes resumes that closely match the job requirements. In conclusion, this AI-driven solution enhances recruitment efficiency, reducing manual effort and improving hiring decisions. Future improvements could include **natural language processing (NLP) enhancements**, and **integration with applicant tracking systems (ATS)** to further refine candidate ranking.

TABLE OF CONTENT

Abstract
Chapter 1. Introduction
1.1 Problem Statement
1.2 Motivation
1.3 Objectives
1.4 Scope of the Project
Chapter 2. Literature Survey
Chapter 3. Proposed Methodology
Chapter 4. Implementation and Results
Chapter 5. Discussion and Conclusion
References

LIST OF TABLES

Table No.	Table Caption	Page No.
1	Introduction	6
1.1	Problem Statement	
1.2	Motivation	
1.3	Objectives	
1.4	Scope of the Project	
2	Literature Survey	7
3	Proposed Methodology	8
4	Implementation and Results	11
5	Discussion and Conclusion	13

CHAPTER 1

Introduction

1.1 Problem Statement:

Recruiters and hiring managers often face significant challenges in manually screening and shortlisting resumes from a large pool of applicants. With the AI-powered Resume Screening and Ranking System, the recruiter simply inputs the job description and uploads the resumes. The system instantly analyze and ranks the resumes based on relevance, allowing the recruiter to focus on top-matching candidates within minutes instead of hours or days

1.2 Motivation:

This project was chosen to address these challenges by leveraging AI and automation to make the hiring process faster, fairer, and more accurate.

Potential Applications & Impact

Recruitment Automation – Organizations can integrate this AI-powered system into their hiring processes to automatically rank resumes based on job descriptions, reducing manual effort.

Time & Cost Efficiency – Companies can significantly cut down the time spent on initial resume screening, allowing recruiters to focus on final interviews and candidate engagement.

1.3 Objective:

The AI-powered Resume Screening and Ranking System is designed to automate and enhance the resume evaluation process. The Develop an AI-based system that automatically extracts, analyzes, and ranks resumes based on their relevance to a given job description. Reduce the time and effort required for recruiters to manually screen and shortlist candidates, making the hiring process faster and more effective and improving the likelihood of selecting the best candidates.

1.4 Scope of the Project:

Scope: This system automates resume screening by extracting text from **PDF resumes**, analyzing relevance using **TF-IDF vectorization and cosine similarity**, and ranking candidates based on job descriptions. It provides a **real-time, user-friendly interface** using **Streamlit**, making it suitable for recruiters and HR professionals.

Limitations: The system currently focuses on **text-based resume analysis**, lacking support for **image-based or non-PDF formats**.

CHAPTER 2

Literature Survey

2.1 Review of Relevant Literature or Previous Work

Several AI-driven resume screening systems exist, utilizing natural language processing (NLP), machine learning, and keyword matching to filter candidates. Research has shown that automated screening improves efficiency and reduces bias compared to manual evaluation. However, many existing models struggle with contextual understanding and ranking accuracy.

2.2 Existing Models, Techniques, and Methodologies

Common techniques include TF-IDF vectorization, cosine similarity, BERT-based NLP models, and rule-based keyword matching. Many Applicant Tracking Systems (ATS) use Boolean search and heuristic scoring, but they often fail to prioritize resumes effectively when multiple candidates share similar skills.

2.3 Gaps in Existing Solutions & How This Project Addresses Them

Existing systems often lack precision, struggle with ranking relevance, and rely on simple keyword matching without deep contextual analysis. This project enhances resume screening by using TF-IDF and cosine similarity for more accurate rankings and provides a real-time, user-friendly Streamlit interface for instant results. Future improvements could integrate advanced NLP models for better candidate evaluation.

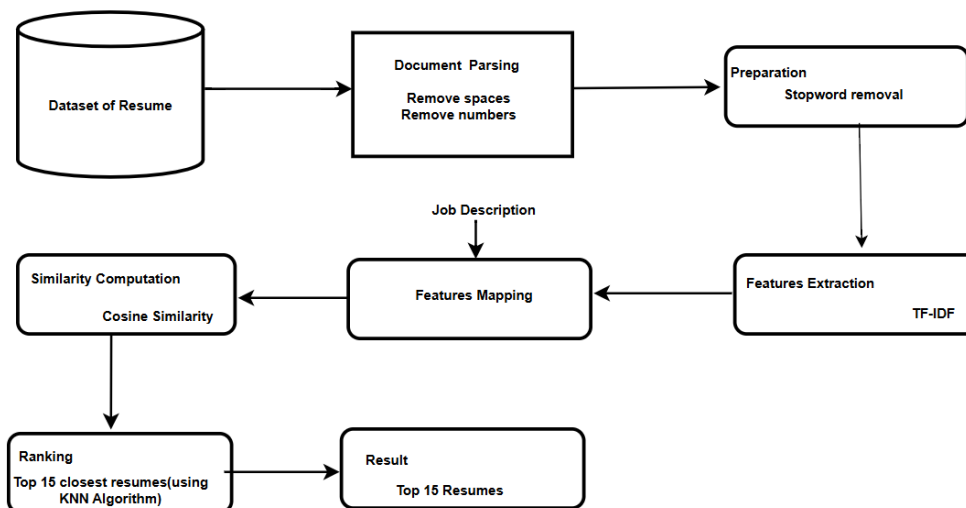
CHAPTER 3

Proposed Methodology

3.1 System Design

AI Resume Screening & Ranking System

System Architecture



1. Dataset of Resumes (Input Layer)

- The system starts with a database of resumes uploaded in PDF format.
- These resumes contain candidate details, including skills, experience, and education.

2. Document Parsing & Preprocessing

- The resumes go through a parsing process where unnecessary elements like extra spaces and numbers are removed.

3. Stopword removal

- It applied to eliminate common words like "the," "and," or "is" that do not contribute to meaningful resume analysis.

4. Feature Extraction (TF-IDF Vectorization)

- The cleaned resume text and job description are transformed into numerical representations using TF-IDF (Term Frequency-Inverse Document Frequency).
- TF-IDF helps quantify the importance of words within each resume compared to the job description.

5.Feature Mapping

- The job description is also vectorized and mapped with all processed resumes.
- This ensures that all resumes are evaluated based on their similarity to the job description rather than arbitrary keyword matching.

6. Similarity Computation

- Cosine similarity is applied to compare the job description vector with each resume vector.
- The higher the similarity score, the better the match between the resume and the job description.

7. Ranking (KNN Algorithm)

- The system selects the top 10 resumes that are closest to the job description using the K-Nearest Neighbors (KNN) algorithm.
- This ensures that only the most relevant candidates are prioritized.

8.Result Display (Output Layer)

- The system presents the top-ranked resumes with their respective similarity scores in a structured format.
- The recruiter can now quickly review the best-matching resumes and proceed with the hiring process efficiently.

3.2Requirement Specification- The implementation of the AI-powered Resume Screening & Ranking System requires specific hardware and software components to ensure smooth execution.

3.2.1 Hardware Requirements:

- **Processor-** Intel Core i5 12th Gen
- **RAM-** Minimum 8GB RAM Efficient for fast processing
- **Storage** – At least 10 GB enough free space for storing resumes and logs

3.2.2 Software Requirements:

Operating System

- Windows 10/11, macOS, or Linux

Programming Language

- Python 3.x (Core language for AI processing)

Frameworks & Libraries

- Streamlit – For building the web interface
- PyPDF2 – For extracting text from PDF resumes
- Pandas – For handling resume data
- Scikit-learn – For TF-IDF vectorization & cosine similarity
- NumPy – For numerical computations

Development Tools

- Jupyter Notebook / VS Code / PyCharm – For coding & debugging
- GitHub/Git – For version control

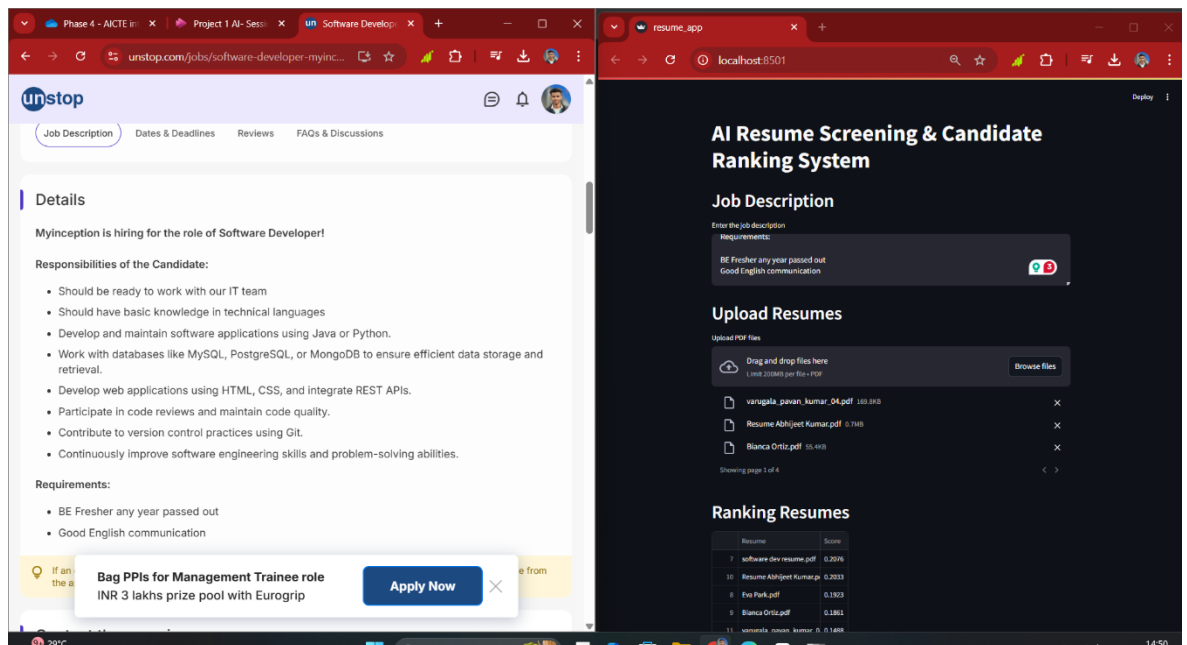
Deployment (Optional)

- Heroku / AWS / Google Cloud – If deploying the application online

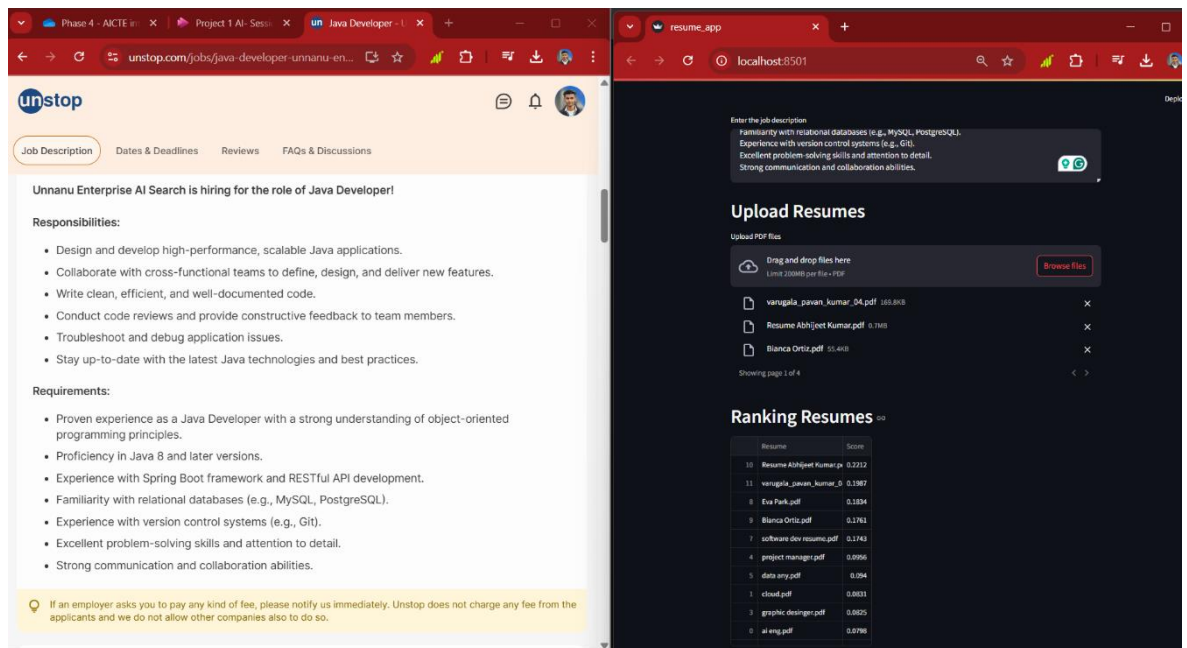
CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:



in the given picture we copy the job description from unstop for the software developer position and paste in my streamlit application and upload multiple resume to see which resume most relevant with the given job vacancy.



In this screenshot we simply copy the job role of java developer responsibilities and their requirements and then I paste in into my generated web application and then upload multiple resume after that we can see the where our resume ranking in the list

4.2 GitHub Link for Code:

Link- <https://github.com/iamrony1101?tab=repositories>

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

- **Advanced NLP Models** – Implement **BERT, GPT-based models, or Word2Vec** for a deeper understanding of resume content beyond keyword matching.
- **Support for Multiple File Formats** – Extend compatibility to **DOCX, TXT, and image-based resumes (OCR integration)**.
- **Context-Aware Ranking** – Improve ranking by considering **experience level, job roles, and skill relevance**.
- **Integration with ATS (Applicant Tracking Systems)** – Allow seamless **integration with HR tools** to automate end-to-end recruitment.
- **Real-Time Feedback for Job Seekers** – Provide **automated suggestions** to candidates on improving their resumes based on job descriptions..

5.2 Conclusion:

The **AI-powered Resume Screening & Ranking System** successfully automates and optimizes the recruitment process by **reducing manual effort, improving efficiency, and ensuring fair candidate evaluation**. By leveraging **Python, Streamlit, TF-IDF, and cosine similarity**, the system enables **real-time resume ranking**, helping recruiters shortlist the most relevant candidates quickly. This project demonstrates the **power of AI in recruitment** and lays the foundation for **future advancements in automated hiring systems**. With continued improvements, such as **advanced NLP techniques and ATS integration**, this system can revolutionize the way companies **identify and hire top talent** in a **faster, fairer, and more data-driven manner**.

REFERENCES

Salton, G., & McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. McGraw-Hill.

Upadhyay, A., & Khandelwal, K. (2018). *Applying AI for Bias-Free Hiring: A Review*. *Journal of HR Analytics*, 3(2), 109–123.