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

Prathamesh Chikankar, prathu3322@gmail.com

Under the Guidance of

Saomya Chaudhury

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep sense of gratitude to all individuals who helped me directly or indirectly during this project work.

Firstly, I would like to thank my supervisor, Saomya Chaudhury, for being a great mentor and adviser. Their guidance, encouragement, and insightful critiques were a source of inspiration and innovation that helped in the successful completion of this project. Their confidence in my work was the driving force that kept me motivated. It has been a privilege working under their supervision. Their invaluable lessons and discussions have not only aided in this project but have also contributed to my overall professional development.

I also extend my sincere gratitude to TechSaksham, Microsoft, and SAP for providing me with this learning opportunity.

ABSTRACT

This project presents an AI-powered Resume Screening and Ranking System designed to streamline the recruitment process. Employers often receive a large number of resumes for each job posting, making manual screening time-consuming and inefficient. This project leverages Natural Language Processing (NLP) and Machine Learning (ML) to automate the resume screening process by extracting text, analyzing relevance, and ranking candidates based on job description compatibility.

The system extracts textual information from resumes, converts it into numerical representations using TF-IDF vectorization, and applies cosine similarity to rank resumes against a given job description. The system is built using Python, Pandas, Streamlit, and PyPDF2, making it user-friendly and efficient.

The results demonstrate the system's ability to accurately filter and rank resumes, significantly reducing the time and effort required for initial candidate screening. The project provides a foundation for further enhancements, such as deep learning-based ranking models and skill-based classification.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1   Problem Statement:

Recruitment processes involve manually reviewing hundreds of resumes, which is time-consuming and inefficient. Automating resume screening using AI can significantly reduce effort and improve candidate selection quality.
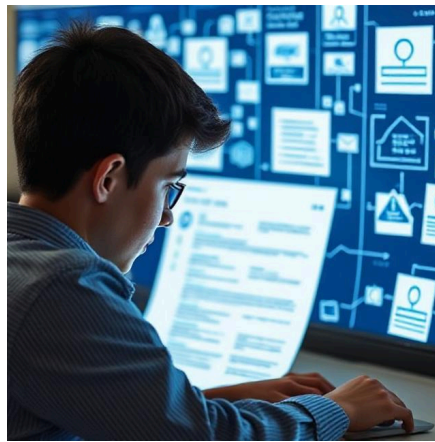

Fig. 1. Resume Screening

1.2   Motivation:

The increasing demand for efficient hiring solutions motivates the development of an automated resume screening system. This project provides a practical application of AI in recruitment, enhancing HR efficiency and candidate matching accuracy.

1.3   Objective:

- Extract information from resumes (PDF format)
- Convert text into numerical representations using NLP
- Rank resumes based on job description similarity
- Develop an interactive web-based user interface for easy access

1.4   Scope of the Project:

- Automates initial resume screening using AI techniques
- Supports text-based analysis but does not include advanced skill validation
- Designed for small to medium-scale recruitment use cases

CHAPTER 2

Literature Survey

A literature survey provides an overview of existing research, methodologies, and technologies related to AI-powered resume screening. This section explores traditional methods, recent advancements, and key research papers that contribute to automated hiring solutions.

2.1 Traditional Resume Screening Methods

Traditional resume screening is a manual process where recruiters review each resume individually. This process involves:

- **Keyword Matching:** Recruiters scan for specific keywords related to job requirements.
- **Experience and Qualification Filtering:** Candidates are shortlisted based on academic qualifications, skills, and past experience.
- **Human Bias:** Decision-making can be influenced by cognitive biases, leading to subjective hiring.
- **Time-Consuming Nature:** Reviewing hundreds of resumes manually is inefficient.

These limitations highlight the need for automated AI-powered resume screening systems.

2.2 AI and NLP in Resume Screening

With advancements in **Artificial Intelligence (AI) and Natural Language Processing (NLP)**, automated resume screening has become more efficient. AI-based approaches address the challenges of traditional screening by:

- **Text Parsing and Entity Recognition:** Extracting relevant details (name, contact, skills, experience) using NLP techniques.
- **Vectorization Techniques (TF-IDF, Word Embeddings):** Converting unstructured text into numerical representations for similarity comparison.
- **Machine Learning Models:** Using classification and ranking models to assess resume-job fit.

2.3 Existing AI-Based Resume Screening Systems

Several AI-powered platforms have been developed for automated resume screening. Some notable systems include:

Platform	Description
HireVue & Pymetrics	Utilize AI-driven assessments and predictive analytics for candidate ranking.
LinkedIn Talent Insights	Uses AI to analyze candidate profiles and recommend suitable matches.
IBM Watson Recruitment	Implements NLP and deep learning to evaluate resumes against job descriptions.

Table 1. Comparison of different Platforms

A comparison of these systems shows that while they provide high accuracy, many are expensive and require complex integration.

2.4 Comparison of Manual vs. Automated Screening

Feature	Manual Screening	AI-Based Screening
Speed	Slow	Fast
Bias	High (human bias)	Reduced bias
Accuracy	Subjective	Data-driven
Scalability	Limited	Scalable

Table 2. Manual Screening vs AI-Based Screening

This comparison demonstrates the advantages of AI-powered systems in recruitment.

2.5 Research Gaps and Future Scope

Despite advancements in AI-based resume screening, some gaps remain:

- **Contextual Understanding:** Many AI models struggle to understand the meaning behind job descriptions and resumes.
- **Bias in AI Models:** Training data can introduce bias, leading to unfair candidate rankings.
- **Multi-Format Support:** Many systems fail to process resumes in formats other than PDF (e.g., DOCX, LinkedIn profiles).

Addressing these challenges will lead to more robust and fair AI-driven hiring solutions.

CHAPTER 3

Proposed Methodology

3.1   System Design

- **Input:** Job description and multiple resumes (PDF format)
- **Processing:** Text extraction, TF-IDF vectorization, Cosine Similarity calculation
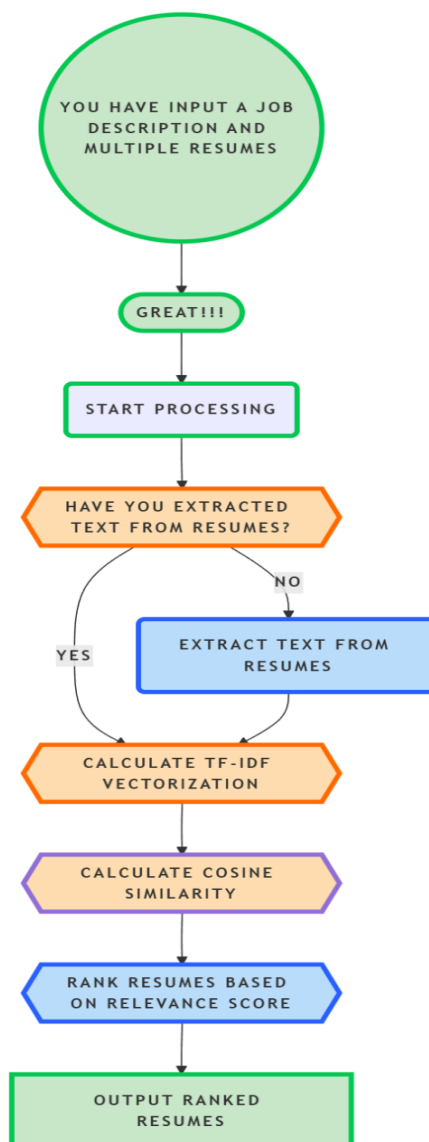- **Output:** Ranked resumes based on relevance score

Fig. 2. Flowchart of AI-Powered Resume Screening and Ranking System

3.2 Requirement Specification

3.2.1 Hardware Requirements:

- Minimum 4GB RAM
- Processor: Intel i5 or higher
- Storage: Minimum 10GB free space

3.2.2 Software Requirements:

- Python
- Jupyter Notebook / Google Colab
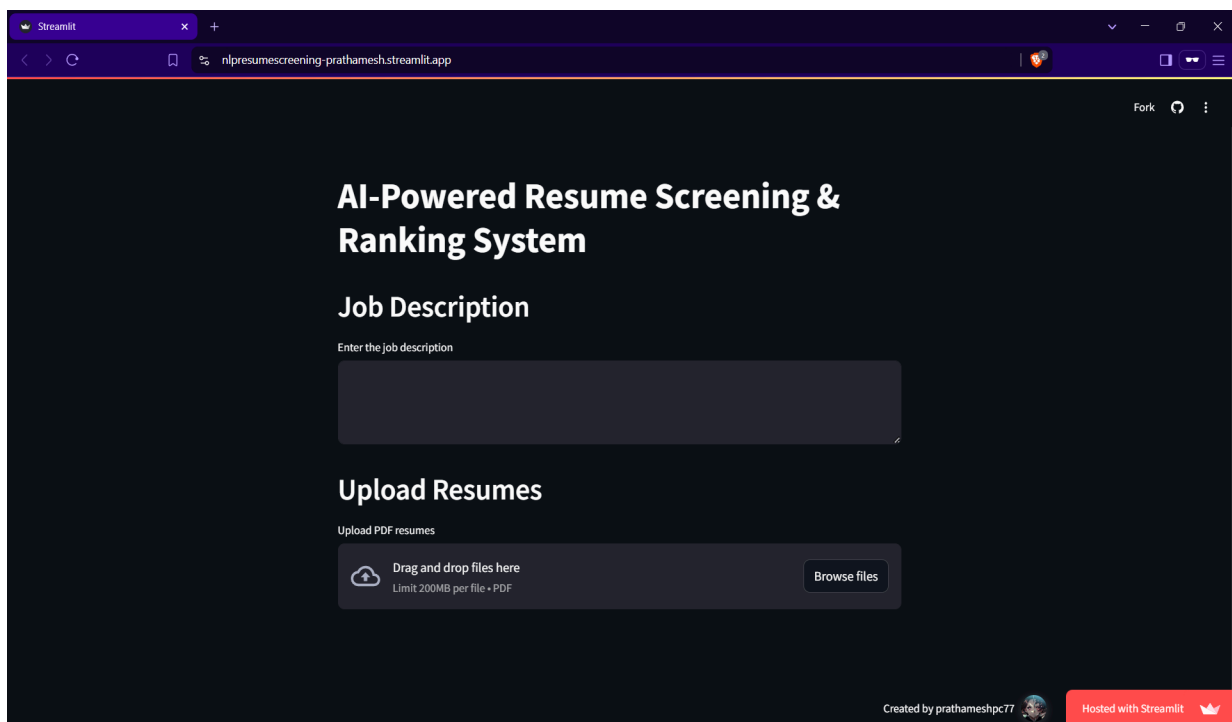- Libraries: Pandas, Scikit-learn, PyPDF2, Streamlit

Fig. 3. Softwares + Libraries used

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

- **Figure 1: Streamlit UI – Homepage**

This screenshot shows the initial UI of the AI Resume Screening & Ranking System built with Streamlit. The interface provides an intuitive design where users can input a job description and upload multiple resumes for evaluation.

- **Figure 2: Job Description Entry & Resume Upload**

Here, the user enters the job description in the provided text area and uploads multiple resumes in PDF format. This step is crucial as it allows the system to compare resumes against the given job description using NLP techniques.

- **Figure 3: Resume Ranking & Scoring Output**
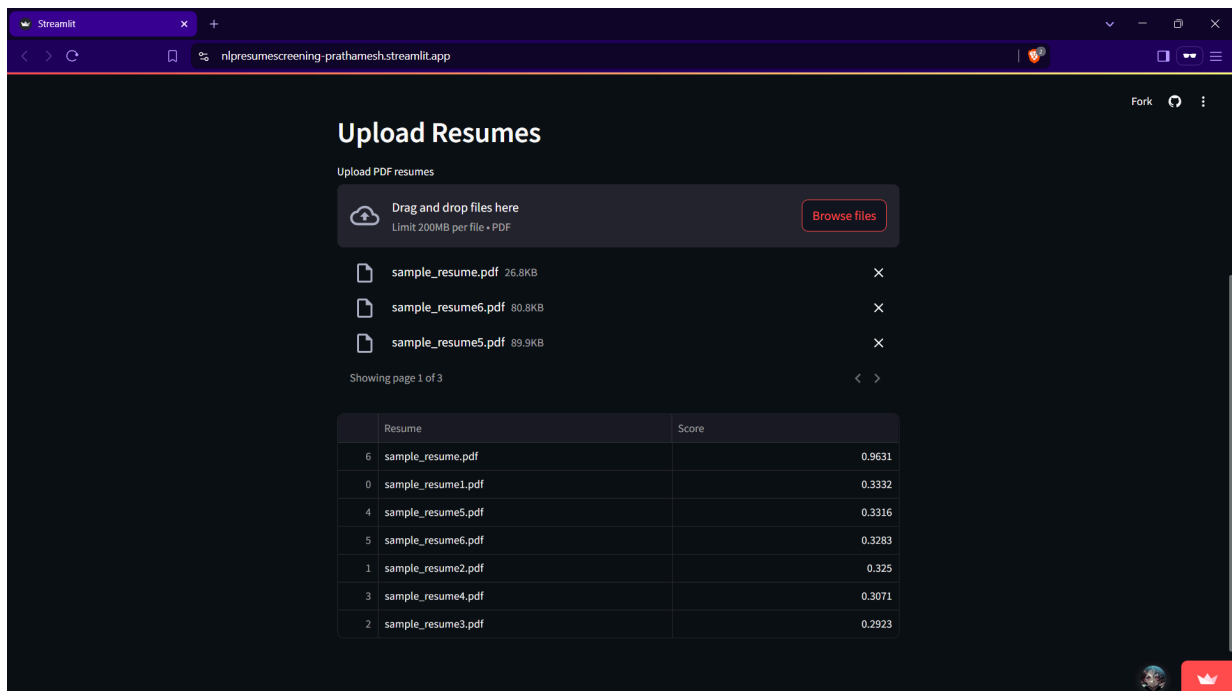
This screenshot displays the calculated scores of the uploaded resumes, generated using TF-IDF and Cosine Similarity. Users can hover over the table to access a CSV download option, which allows them to save and manually rank resumes based on the scores.

4.2   GitHub Link for Code:

- **https://github.com/PrathameshPC77/nlp_resume_screening**
- **Deployment Link: https://nlpresumescreening-prathamesh.streamlit.app/**

CHAPTER 5

Discussion and Conclusion

5.1   Future Work:

- Integration with Deep Learning models (BERT/GPT-based ranking)
- Addition of Skill-based filtering
- Expanding system to support multiple resume formats (DOCX, TXT, etc.)

5.2   Conclusion:

This project successfully demonstrates an AI-powered resume ranking system that automates initial screening using NLP techniques. The implementation significantly reduces recruitment efforts and enhances efficiency. Future advancements will further improve accuracy and usability.

REFERENCES

[1]. **TechSaksham AICTE Internship Sessions** – AI: Transformative Learning. (2024). [Online]. Available through TechSaksham Learning Portal.

[2]. **YouTube Tutorials on AI-powered Resume Screening** – Various video resources on NLP, TF-IDF, Cosine Similarity, and Resume Parsing. (2024). [Online].

Available at:
[https://youtube.com/playlist?list=PLHmESvg2HbgeQUQZJBMWEXnHe8SjDmPGK].

[3]. **Python Official Documentation** – Python Libraries for NLP and Text Processing. (2024). [Online].

Available at: https://docs.python.org/3/

[4]. **scikit-learn Library** – Machine Learning Toolkit. (2024). [Online].

Available at: https://scikit-learn.org/

[5]. **Streamlit Official Documentation** – Web Framework for Data Apps. (2024). [Online].

Available at: https://docs.streamlit.io/

[6]. **ChatGPT Assistance** – AI-powered guidance and technical clarifications for implementing Resume Screening System. (2024).