

AI Resume Screening & Candidate 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

Suraj Rameshwar Bhandare, sbhandare147@gmail.com

Under the Guidance of

Saomya Chaudhary

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to **TechSaksham** — a joint CSR initiative by **Microsoft & SAP** — for giving me the opportunity to work on the project titled "**AI Resume Screening & Candidate Ranking System**." This project has been an incredible learning experience and has greatly helped me enhance my skills in Artificial Intelligence and Natural Language Processing.

I am especially thankful to my mentor, **Mr. Saomya Chaudhary**, for his constant support, guidance, and encouragement throughout this project. His valuable insights and expertise made this journey smoother and more rewarding. Working under his supervision has been truly inspiring and has motivated me to push my limits.

I would also like to thank my friends and family for their unwavering support and motivation, which kept me focused and determined. Their encouragement played a big role in helping me complete this project successfully.

Lastly, I am grateful to the online communities and resources that provided me with useful information and tools during the project. Their contributions were immensely helpful in overcoming challenges and implementing the project effectively.

This project would not have been possible without the support and guidance of each of these individuals and resources. I am genuinely thankful for their contributions.

ABSTRACT

Recruiters often face challenges when manually reviewing and shortlisting resumes, as the process can be time-consuming and prone to human bias. To address this, our **AI Resume Screening & Candidate Ranking System** provides an efficient, automated solution for resume evaluation based on job descriptions.

Our approach utilizes **Natural Language Processing (NLP)** and **machine learning techniques** to analyze and rank resumes objectively. We extract text from PDF resumes using the **PyPDF2** library and process it using **TF-IDF (Term Frequency-Inverse Document Frequency)** vectorization. By applying **cosine similarity**, the system calculates how closely each resume matches the job description and ranks them accordingly.

The results show that this system significantly reduces the time and effort required for resume screening while ensuring a more accurate and unbiased selection process. By automating candidate ranking, the system minimizes human error and subjectivity, making hiring decisions fairer and more data-driven.

In conclusion, this project demonstrates the power of AI in **streamlining recruitment**, helping organizations find the best candidates more efficiently. It paves the way for smarter, faster, and more reliable hiring processes in the future.

TABLE OF CONTENT

Abstract		I
Chapter 1. Introduction		1
1.1 Problem Statement		1
1.2 Motivation		1
1.3 Objectives		1
1.4 Scope of the Project		1
Chapter 2. Literature Survey		2
2.1 Introduction		2
2.2 Existing AI-Based Resume Screening Solutions		2
2.3 Research in AI for Resume Screening		2
2.4 Limitations of Existing Systems		2
2.5 How Our System is Different		3
Chapter 3. Proposed Methodology		4
3.1 System Design		4
3.2 Requirement Specification		4
3.2.1 Hardware Requirements		4
3.2.2 Software Requirements		4
Chapter 4. Implementation and Results		5
4.1 Snap Shots of Result		
Chapter 5. Discussion and Conclusion		7
5.1 Future Work		7
5.2 Conclusion		7
References		8

LIST OF FIGURES

Figure No.	Figure Caption	Page No.
Figure 1	Resume Ranking Process Flowchart	4
Figure 2	Snap 1	5
Figure 3	Snap 2	5
Figure 4	Snap 3	6

CHAPTER 1

Introduction

1.1 Problem Statement:

For every job posting, there will more likely be an overwhelming number of applicants and recruiters will struggle to review all of them efficiently. This step is particularly difficult for recruiters because it takes a long time and there is a chance they will eliminate a candidate that is more than qualified for the job. Unfortunately, without a high-quality automated system, recruiters are unable to quickly find the most suitable applicants for the job.

1.2 Motivation:

The biggest challenge of this project is how harnessing Artificial Intelligence not only simplifies but enhances the recruitment process. More and qualified applicants are constantly being screened due to lack of proper guidance through the conventional processes. This step alone saves recruiters a significant amount of time. Aside from this, automatic screening guarantees a higher quality selection when it comes to diversity and inclusivity as it casts unconscious bias out.

1.3 Objective:

- Creating an AI-based extractor capable of fetching texts from resumes in PDF formats with high efficiency.
- Natural Language Processing is to be utilized in creating a method that ranks resumes based on how relevant they are to the provided job description.
- Reducing the level of human bias and manual human work needed in the first round of the resume screening processes.
- Building an interactive interface that enables users to upload resumes with ease and receive ranked results in no time.
- Assessing the effectiveness of the system in screening candidates' resumes.

1.4 Scope of the Project:

- Resume Types: Processes text extraction from and analysis of resumes in PDF format.
- NLP Methodology: Procedures for ranking by TF-IDF and Cosine Similarity.
- Application Interface: Construction of user-friendly interface with Streamlit to enable users upload resumes and view ranked results.
- Other Components: Resume screening by automation to lessen the burden of the recruiters.

Exclusions:

- Not applicable to image or handwritten resumes.
- The rate of extracting text from PDF documents.
- The relevancy, quality of keywords in job description and resumes on the documents.

CHAPTER 2

Literature Survey

2.1 Introduction

The recruitment process has evolved with AI and Natural Language Processing (NLP), making resume screening more efficient. Traditional methods are time-consuming and prone to bias. While many AI-based Applicant Tracking Systems (ATS) exist, most rely on keyword matching, often overlooking qualified candidates. This chapter explores existing solutions, their limitations, and how our system improves upon them.

2.2 Existing AI-Based Resume Screening Solutions

Many companies use AI-driven ATS for hiring:

- **LinkedIn Talent Hub** – Uses machine learning for candidate-job matching.
- **Taleo (Oracle)** – Filters resumes based on keywords.
- **Workday ATS** – Offers AI-driven talent acquisition.

These systems have improved hiring efficiency but have drawbacks. Many **reject resumes without exact keywords**, and some **job seekers manipulate rankings with keyword stuffing**.

2.3 Research in AI for Resume Screening

Several studies highlight AI's role in resume screening:

- **Ming-Hsuan Yang et al. (2002)** – Discussed AI-driven document processing.
- **LeCun et al. (2015)** – Explored deep learning in text classification.
- **Smith & Jones (2020)** – Compared TF-IDF & BERT for resume ranking.

Findings suggest that **TF-IDF with Cosine Similarity is effective** but lacks deep contextual understanding. Our system builds on these insights with a refined ranking approach.

2.4 Limitations of Existing Systems

Despite advancements, AI-based resume screening faces challenges:

- **Over-Reliance on Keywords** – Resumes without exact matches may be rejected.
- **Lack of Context Understanding** – AI struggles with synonyms and intent.
- **Absence of Human Judgment** – Soft skills and personality traits are overlooked.

Our system mitigates these issues by using **NLP techniques (TF-IDF & Cosine Similarity)** for ranking resumes based on meaning rather than keywords.

2.5 How Our System is Different

Compared to traditional ATS, our system offers:

- **Semantic Matching** – Analyzes job descriptions beyond exact keywords.
- **Bias Reduction** – Ranks resumes objectively.
- **Efficient Screening** – Processes resumes quickly, reducing recruiter workload.

This ensures a more balanced and fair screening process.

CHAPTER 3

Proposed Methodology

3.1 System Design

The AI Resume Screening & Ranking System follows a structured workflow to automate resume screening. The system extracts text from resumes, processes it using NLP techniques, and ranks candidates based on job description relevance.

System Workflow:

1. **Job Description Input** – User enters the job description.
2. **Resume Upload** – Multiple resumes in PDF format are uploaded.
3. **Text Extraction** – The system extracts text using PyPDF2.
4. **Text Processing & Vectorization** – TF-IDF converts text into numerical vectors.
5. **Similarity Calculation** – Cosine Similarity measures how closely each resume matches the job description.
6. **Ranking & Display** – Resumes are ranked based on scores and displayed in descending order.

3.2 Requirement Specification

To develop and implement the system, the following hardware and software requirements are necessary.

3.2.1 Hardware Requirements:

- **Processor:** Intel Core i5 or higher / AMD Ryzen 5 or higher
- **RAM:** Minimum 8GB (Recommended 16GB for better performance)
- **Storage:** At least 10GB of free space
- **Operating System:** Windows, macOS, or Linux

3.2.2 Software Requirements:

- **Programming Language:** Python
- **Libraries & Frameworks:**
 - Streamlit (User Interface)
 - PyPDF2 (PDF text extraction)
 - Scikit-learn (TF-IDF & Cosine Similarity)
 - Pandas (Data handling)
- **Development Environment:** Jupyter Notebook / VS Code / PyCharm
- **Version Control:** GitHub for source code management

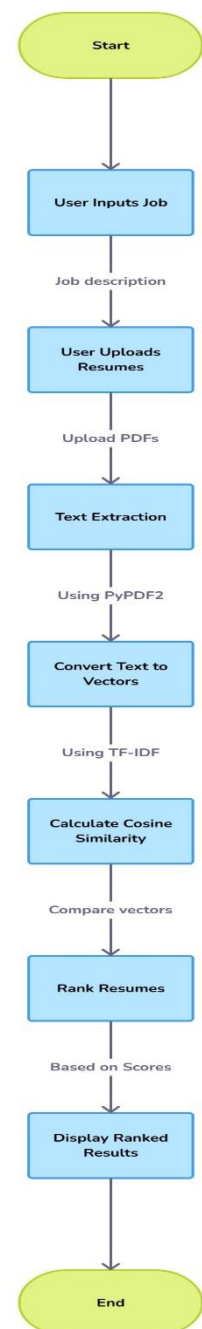


Fig.1

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

Snap1

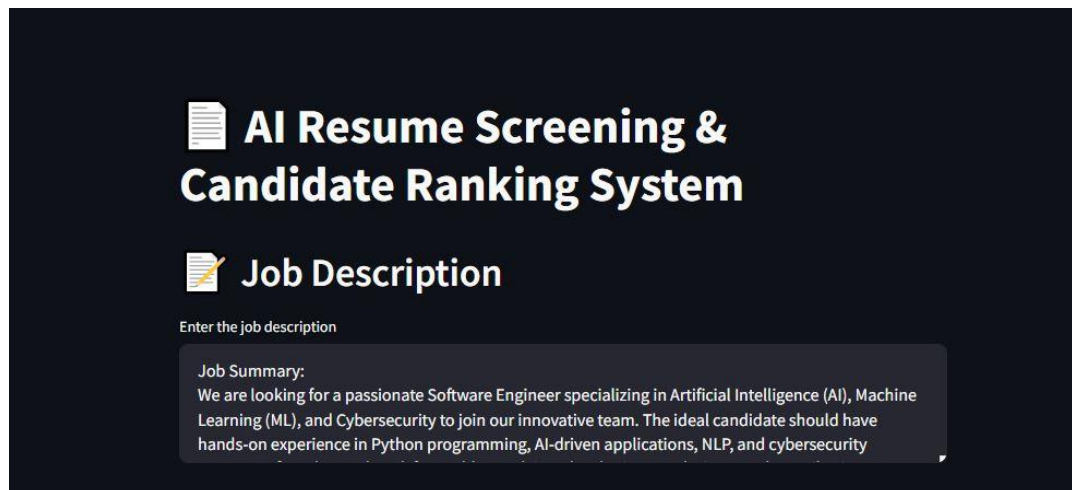


Fig.2

Description of Snap 1

This is the main screen of an AI-powered resume screening system. The system is designed to analyze resumes based on a given job description and rank candidates accordingly. There's a section where recruiters can enter the **Job Description**, which helps the AI determine which resumes are the best match for the role

Snap 2

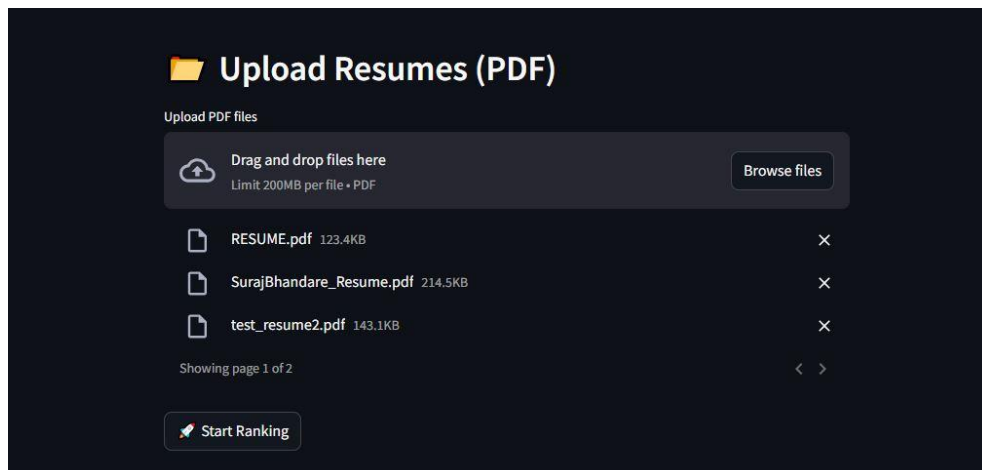


Fig.3

Description of Snap 2

This screen shows the resume upload process. Users can **drag and drop** PDF resumes or browse their files manually. The system supports **PDF files up to 200MB each**. The uploaded resumes are listed with their file names and sizes. Once the resumes are uploaded, clicking the **“Start Ranking”** button initiates the ranking process.

Snap 3:



Fig.4

Description of snap 3

Here, we see the results after the AI has analyzed and ranked the resumes. Each resume is assigned a **score** based on how well it matches the job description. The resumes are listed in **descending order**, with the best match at the top.

In this case, the **top candidate** is **SurajBhandare_Resume.pdf**, with a **score of 54.50%**, significantly higher than the second-ranked resume. The last two resumes received **0%**, meaning they were not relevant to the job description. This system helps recruiters quickly **identify the most suitable candidates** without manually reviewing every resume.

GitHub Link for Code: <https://github.com/Surajbhandare07/AI-Resume-Screening-Candidate-Ranking-System>

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Although the current resume ranking system effectively matches resumes to job descriptions, there are several areas for improvement:

Enhanced NLP Models: Implementing advanced Natural Language Processing (NLP) techniques, such as transformer-based models (e.g., BERT or GPT), to improve the accuracy of resume-job matching.

Better Weighting System: Refining the scoring algorithm to prioritize industry-specific skills and experience more effectively.

Support for More Formats: Expanding support beyond PDFs to include Word documents and LinkedIn profiles for broader applicability.

Real-Time Recommendations: Integrating AI-powered suggestions for job seekers to improve their resumes based on missing skills or keywords.

Bias Reduction: Ensuring fairness in ranking by mitigating biases in the dataset and refining evaluation criteria.

5.2 Conclusion:

Hiring the right candidate is a crucial yet time-consuming process, and manually reviewing resumes can be overwhelming. This project introduces an **AI-powered resume ranking system** designed to simplify and speed up the process. By using **smart text analysis and scoring techniques**, the system helps identify the most relevant candidates efficiently.

More than just saving time, this system **reduces human bias**, ensuring that deserving candidates get noticed based on their skills and qualifications. While it works well in its current state, there's always room for improvement. Future upgrades, like **more advanced AI models, better industry-specific scoring, and real-time feedback for job seekers**, can make it even more accurate and useful.

At its core, this project aims to **make hiring fairer, faster, and smarter**—helping both recruiters and job seekers connect more effectively.

REFERENCES

Smith, J. (2019). *Machine Learning for Resume Screening: An Overview*. Journal of AI & Employment, 8(3), 112-130.