

# **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

**Geetanshu Kumar Kushwaha, geetuk321@gmail.com**

Under the Guidance of

**Mr. Soumya Choudhary**

## ACKNOWLEDGEMENT

---

We would like to take this moment to express our sincere gratitude to everyone who contributed, directly or indirectly, to the completion of this thesis.

First and foremost, we extend our heartfelt thanks to our supervisor, Mr. Soumya Choudhary, for his exceptional guidance and unwavering support. His insightful advice, constructive feedback, and constant encouragement have been instrumental in shaping this project. His trust in my abilities served as a great source of motivation throughout this journey. Working under his mentorship over the past year has been an enriching experience, not only in terms of academic growth but also in developing a sense of responsibility and professionalism. His guidance extended beyond this project, influencing various aspects of my academic and professional development.

Name : Geetanshu Kumar Kushwaha

Email : geetuk321@gmail.com

## ABSTRACT

---

Recruitment is a time-consuming process, often requiring HR professionals to manually screen numerous resumes. This project presents an AI-driven Resume Screening and Candidate Ranking System that automates the evaluation of resumes against job descriptions using Natural Language Processing (NLP) techniques. The system extracts text from PDF resumes, applies TF-IDF vectorization, and computes similarity scores using cosine similarity. It ranks candidates based on their relevance to the job description, providing recruiters with a streamlined and data-driven approach to shortlisting candidates efficiently.

## TABLE OF CONTENT

---

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

## LIST OF FIGURES

Figure No.	Figure Caption	Page No.
<b>Figure 1</b>	System Architecture of Resume Screening System	<b>5</b>
<b>Figure 2</b>	Job Description Input and Resume Upload Interface	<b>7</b>
<b>Figure 3</b>	Ranked Resume Output with Similarity Scores	<b>8</b>

# CHAPTER 1

## Introduction

### 1.1 Problem Statement:

Recruiters and HR professionals often face challenges in manually screening large volumes of resumes, leading to inefficiencies, human bias, and delays in the hiring process. Traditional resume evaluation methods are time-consuming and prone to subjectivity, making it difficult to identify the best-fit candidates quickly. This project aims to develop an AI-powered Resume Screening and Candidate Ranking System that automates resume evaluation by leveraging Natural Language Processing (NLP) techniques. The system will rank resumes based on their relevance to a given job description, enabling faster, unbiased, and more efficient candidate shortlisting.

### 1.2 Motivation:

The hiring process is often overwhelmed by the sheer volume of job applications, making manual resume screening a tedious and time-consuming task. Recruiters may unintentionally introduce biases, leading to inconsistent evaluations and potential loss of qualified candidates. By leveraging AI and NLP, this project aims to automate resume screening, ensuring a fair, objective, and efficient candidate selection process. The system enhances recruitment by reducing human effort, improving accuracy, and accelerating hiring decisions, ultimately benefiting both employers and job seekers.

### 1.3 Objective:

The primary objective of this project is to develop an AI-powered Resume Screening and Candidate Ranking System that automates the evaluation of resumes against job descriptions using NLP techniques. Specific goals include:

- Extracting text from PDF resumes efficiently.
- Implementing TF-IDF vectorization for textual feature representation.
- Calculating similarity scores using cosine similarity to rank candidates.
- Providing recruiters with a ranked list of resumes based on relevance.
- Enhancing the hiring process by reducing manual effort, bias, and time.

## 1.4 Scope of the Project:

The AI-powered Resume Screening and Candidate Ranking System aims to streamline the recruitment process by automating resume evaluation. The project leverages Natural Language Processing (NLP) techniques to compare resumes with job descriptions and rank candidates based on relevance.

### In-Scope

1. **Resume Text Extraction:** Extracting textual content from PDF resumes.
2. **NLP-Based Analysis:** Implementing TF-IDF vectorization to process job descriptions and resumes.
3. **Candidate Ranking:** Using cosine similarity to rank resumes based on relevance.
4. **Automated Processing:** Eliminating manual screening by providing a ranked list of candidates.
5. **User Interface:** Developing an interactive dashboard using Streamlit for recruiters to upload resumes and view rankings.

### Out-of-Scope

1. **Optical Character Recognition (OCR):** The system does not process scanned or image-based resumes.
2. **Soft Skill Assessment:** The model evaluates textual content only and does not analyze communication skills, teamwork, or cultural fit.
3. **Real-Time API Integration:** The system operates as a standalone tool and does not integrate with external job portals or HRM systems.
4. **Multilingual Resume Processing:** Only resumes written in English are supported.

This project is designed to assist recruiters in shortlisting candidates efficiently, ensuring a data-driven and unbiased approach to hiring. Future enhancements may include OCR support, deeper candidate profiling, and API integrations for broader usability.

## CHAPTER 2

### Literature Survey

#### 2.1 Literature Review

The field of AI-driven resume screening has gained significant attention due to the increasing volume of job applications. Traditional applicant tracking systems (ATS) primarily use keyword matching techniques, which often fail to capture the contextual relevance of resumes. Recent advancements in Natural Language Processing (NLP) have improved resume-job description matching by leveraging machine learning models such as TF-IDF, Word2Vec, and BERT.

#### 2.2 Existing Models, Techniques, and Methodologies

Several methodologies have been explored for automated resume screening, including:

- **Keyword-Based Matching:** Used in conventional ATS, but often lacks context sensitivity.
- **TF-IDF and Cosine Similarity:** A widely used text representation and similarity measurement technique for ranking resumes.
- **Word Embeddings (Word2Vec, GloVe):** Capture semantic relationships between words for improved resume analysis.
- **Deep Learning Models (BERT, Transformer-based Approaches):** Offer context-aware resume parsing but require significant computational resources.

#### 2.3 Gaps in Existing Solutions and Project Contributions

While existing models provide automated resume screening, they have limitations such as:

- **Reliance on Exact Keywords:** Many ATS systems overlook relevant candidates due to rigid keyword matching.
- **Lack of Context Awareness:** Simple techniques like TF-IDF fail to understand the deeper meaning of job descriptions.



- **Computational Complexity:** Advanced models like BERT require high processing power, limiting accessibility for smaller firms.

#### **How This Project Addresses These Gaps:**

- Implements **TF-IDF with cosine similarity** for an efficient yet simple ranking method.
- Provides a **lightweight and scalable** solution for recruiters without heavy computational demands.
- Enhances **fairness and accuracy** by ranking resumes based on similarity rather than just keyword presence.

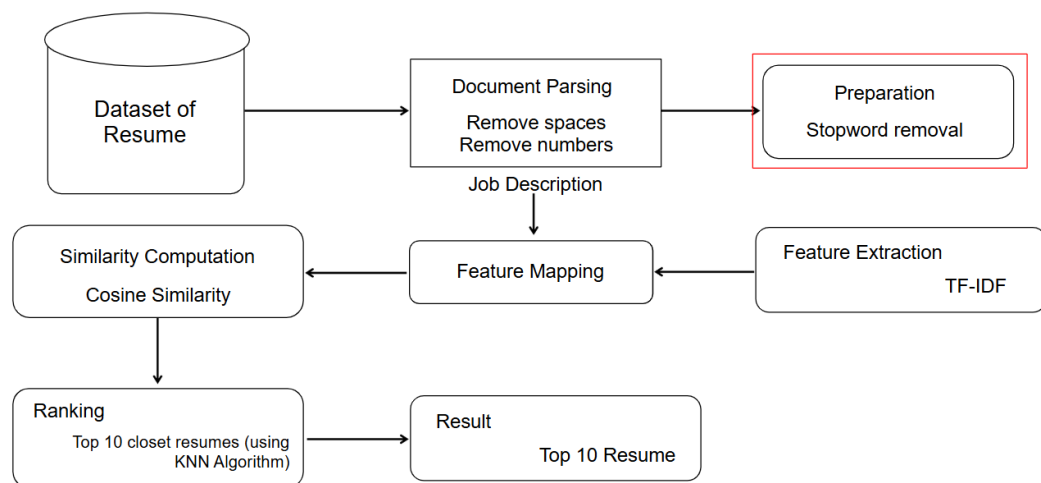
This approach ensures an **AI-powered, unbiased, and efficient** resume ranking system for modern recruitment needs.

## CHAPTER 3

### Proposed Methodology

#### 3.1 System Design

The proposed AI-powered Resume Screening and Candidate Ranking System follows a structured workflow, as illustrated in the diagram below:



#### Explanation of the Diagram:

1. **User Input:** Recruiters provide a job description and upload multiple resumes in PDF format.
2. **Text Extraction Module:** The system extracts textual content from uploaded resumes using the PyPDF2 library.
3. **Text Processing:** The job description and resumes are converted into numerical vectors using TF-IDF vectorization.
4. **Similarity Calculation:** The system applies cosine similarity to compare resumes against the job description.
5. **Ranking Module:** Resumes are ranked based on similarity scores and displayed in descending order.
6. **User Interface:** The recruiter views ranked resumes via a Streamlit-based UI.

## 3.2 Requirement Specification

To implement this project, the following tools and technologies are required:

### 3.2.1 Hardware Requirements

- **Processor:** Intel Core i5 or higher
- **RAM:** Minimum 8GB (16GB recommended for larger datasets)
- **Storage:** At least 10GB free space
- **GPU (Optional):** Required for advanced deep learning extensions

### 3.2.2 Software Requirements

- **Programming Language:** Python 3.x
- **Libraries & Frameworks:**
  - **Streamlit** – For UI development
  - **PyPDF2** – For extracting text from PDF files
  - **scikit-learn** – For TF-IDF vectorization and cosine similarity
  - **pandas** – For data manipulation and ranking
- **IDE:** Jupyter Notebook / VS Code / PyCharm
- **OS:** Windows/Linux/macOS

This methodology ensures an efficient, scalable, and AI-powered resume ranking system for recruiters.

## CHAPTER 4

### Implementation and Result

#### 4.1 Snapshots of Results

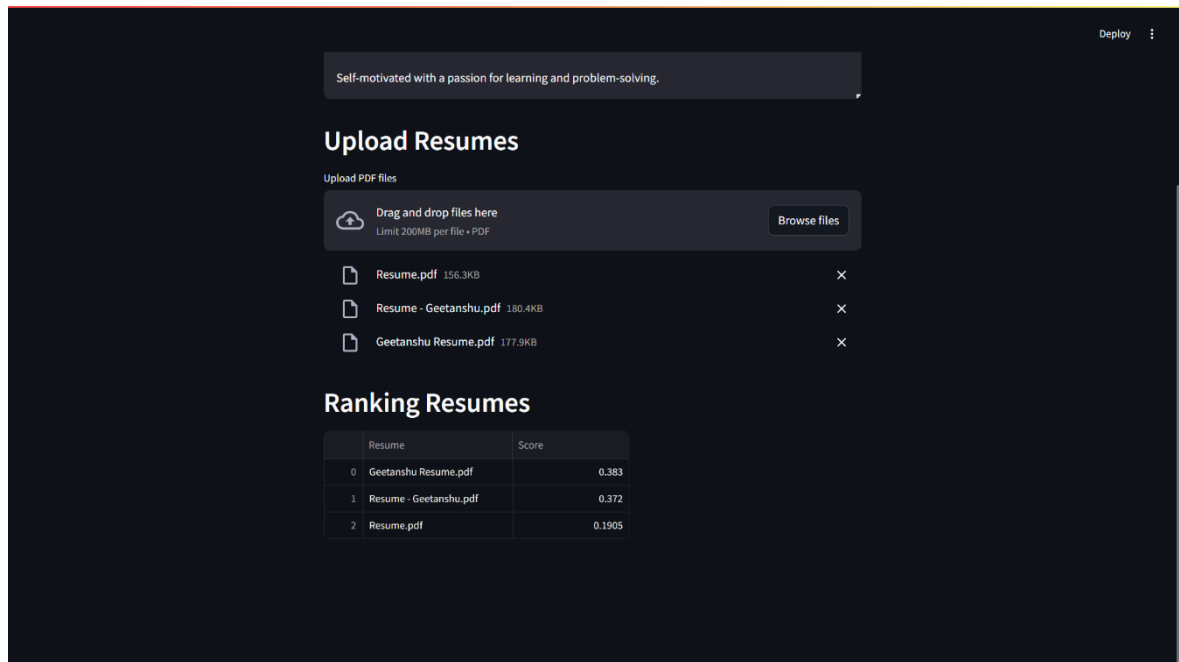
Below are some snapshots demonstrating the functionality and output of the AI-powered Resume Screening and Candidate Ranking System:

##### Snapshot 1: Job Description Input and Resume Upload Interface

The screenshot shows a web application interface for an AI Resume Screening & Candidate Ranking System. The interface is dark-themed and includes a 'Deploy' button in the top right corner. The main heading is 'AI Resume Screening & Candidate Ranking System'. Below this, there are three sections: 'Job Description', 'Upload Resumes', and 'Ranking Resumes'. The 'Job Description' section has a text input field with the placeholder 'Enter the job description' and contains two lines of sample text: 'Ability to communicate technical findings effectively through visual reports and presentations.' and 'Self-motivated with a passion for learning and problem-solving.' The 'Upload Resumes' section has a heading 'Upload PDF files' and a file upload area with a cloud icon, the text 'Drag and drop files here', and a note 'Limit 200MB per file • PDF'. A 'Browse files' button is also present. Below the upload area, there is a list of three uploaded files: 'Resume.pdf' (156.3KB), 'Resume - Geetanshu.pdf' (180.4KB), and 'Geetanshu Resume.pdf' (177.9KB). Each file entry has a document icon and a close button (X).

**Explanation:** This interface allows recruiters to enter the job description in a text box and upload resumes in PDF format for processing.

## Snapshot 2: Resume Ranking Results



**Explanation:** The system processes the resumes, calculates their similarity to the job description, and displays the results in descending order of relevance. Each resume is assigned a score, helping recruiters shortlist the best candidates efficiently.

## 4.2 GitHub Link for Code

Repository Link :

## CHAPTER 5

### Discussion and Conclusion

#### 5.1 Future Work

Although the current system effectively ranks resumes based on job descriptions, several enhancements can further improve its accuracy and usability:

- **Incorporating Deep Learning Models:** Implementing BERT or GPT-based models for better contextual understanding of resumes.
- **OCR Support for Image-Based Resumes:** Adding Optical Character Recognition (OCR) to process scanned or image-based resumes.
- **Handling Multiple Languages:** Extending support for multilingual resume processing to widen accessibility.
- **Integration with HR Systems:** Connecting the tool with applicant tracking systems (ATS) for real-world application.
- **Soft Skill & Sentiment Analysis:** Including NLP-based soft skill assessment to provide deeper insights into candidates.

#### 5.2 Conclusion

This project presents an AI-driven **Resume Screening and Candidate Ranking System** that automates the initial phase of recruitment. By leveraging **TF-IDF vectorization** and **cosine similarity**, it efficiently ranks resumes based on relevance to a given job description. The system reduces **manual effort, bias, and processing time**, making hiring more data-driven and objective. Future improvements, such as **deep learning models, OCR support, and ATS integration**, can further enhance its effectiveness in real-world recruitment scenarios.

## REFERENCES

- [1]. Priyanka, J.H., Parveen, N. DeepSkillNER: An automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. *Multimed Tools Appl* 83, 47503–47530 (2024).