

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

SADIYA KOUSAR, kousars266@gmail.com

Under the Guidance of
Mr.Saomya Chaudhury

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **AICTE (All India Council for Technical Education)** for providing the opportunity to work on this impactful project, **AI-Powered Resume Screening and Ranking System**. Their support and encouragement have been pivotal throughout the course of this project.

I am immensely grateful to my project mentor Saomya Chaudhury, for their continuous guidance, constructive feedback, and valuable insights. Their expertise in the field of Artificial Intelligence and Machine Learning helped shape this project and provided me with the knowledge and resources needed to succeed.

A special thanks to all the faculty members and technical experts who provided their support, suggestions, and encouragement. Their assistance in resolving challenges during the development process has been invaluable.

Lastly, I would like to thank each and everyone for their unwavering support, patience, and encouragement throughout this project.

ABSTRACT

The evolving technology is creating many chances of employment for many.

Nowadays to apply for any job the most essential document is a resume. A resume tells a lot about the person's achievements and skill sets in all walks of life. The person applying for the job highlights the strong points and skill sets required for the company.

Multinational organizations receive thousands of emails from such people who send their resumes for them to apply for a certain post. Now the real challenge is to know which resume is to be sorted and shortlisted according to the constraints. One method is to manually check and sort the resume. Now, this method is the most time consuming and also can lead to a lot of errors because of human interventions. Also, humans cannot keep on working continuously. Hence there is a problem of less efficiency as well. The proposed system focusses on finding the required skill set by scanning the document or the resume and sort according to the skill sets which is a specified constraint of the organization. Machine learning for recruiting is an emerging category of HR technology, designed to reduce or even remove time consuming activities like manual screening resumes

TABLE OF CONTENT

Abstract	
Chapter 1. Introduction	
1.1 Problem Statement	1
1.2 Motivation	2
1.3 Objectives.....	6
1.4. Scope of the Project	8
Chapter 2. Literature Survey	11
Chapter 3. Proposed Methodology	18
Chapter 4. Implementation and Results	22
Chapter 5. Discussion and Conclusion	25
References	27

CHAPTER 1

Introduction

1.1 Problem Statement:

With an increasing volume of job applications, human recruiters often face challenges such as:

- Difficulty in shortlisting resumes quickly.
- Potential human bias during resume evaluation.
- Time-intensive process of manually reviewing resumes.

This project aims to create an AI-based solution that can automate and streamline the process of resume screening, saving time and providing more objective results.

1.2 Motivation:

Why This Project Was Chosen

The decision to choose the **AI-Powered Resume Screening and Ranking System** stemmed from both the growing relevance of AI in the recruitment industry and a desire to tackle real-world challenges that many HR departments face today. In an era where companies are inundated with job applications, manually reviewing resumes is both time-consuming and prone to human error or bias. Additionally, with the increasing emphasis on diversity and efficiency in hiring, it's clear that automation and data-driven decision-making could offer valuable solutions.

This project was chosen because it presents an opportunity to apply cutting-edge AI and Natural Language Processing (NLP) technologies to streamline the hiring process. It combines several aspects of AI, including machine learning, text mining, and natural language understanding, providing a hands-on experience in solving complex, real-world problems. By automating resume screening, this project aligns with my interests in both AI and human resources, while also aiming to create a tangible solution that can improve hiring practices in organizations.

1.2 Motivation:

Why This Project Was Chosen

The decision to choose the **AI-Powered Resume Screening and Ranking System** stemmed from both the growing relevance of AI in the recruitment industry and a desire to tackle real-world challenges that many HR departments face today. In an era where companies are inundated with job applications, manually reviewing resumes is both time-consuming and prone to human error or bias. Additionally, with the increasing emphasis on diversity and efficiency in hiring, it's clear that automation and data-driven decision-making could offer valuable solutions.

This project was chosen because it presents an opportunity to apply cutting-edge AI and Natural Language Processing (NLP) technologies to streamline the hiring process. It combines several aspects of AI, including machine learning, text mining, and natural language understanding, providing a hands-on experience in solving complex, real-world problems. By automating resume screening, this project aligns with my interests in both AI and human resources, while also aiming to create a tangible solution that can improve hiring practices in organizations.

Potential Applications:

1. Streamlining Recruitment Processes

The primary application of the AI-powered resume screening and ranking system is to automate the labor-intensive and time-consuming process of sorting through thousands of job applications. By automatically ranking resumes based on relevance to job descriptions, HR teams can focus on reviewing only the most qualified candidates, significantly reducing the time spent on initial screenings.

2. Reducing Human Bias

Manual resume screening can often lead to unconscious biases based on factors such as gender, age, ethnicity, or education. By implementing an AI-driven system, the process becomes more objective, as the algorithms are focused purely on the skills, experience, and qualifications outlined in the resumes. This leads to fairer, more inclusive hiring practices and can help organizations attract a diverse talent pool.

3. Improving Candidate Experience

For job applicants, an AI-powered system can provide quicker feedback and transparency, as resumes are processed and ranked faster. This enhances the candidate experience by ensuring that resumes are evaluated based on their actual merits, rather than potentially being overlooked due to the sheer volume of applications or human error.

4. Enhancing HR Decision-Making

The system doesn't just filter resumes; it ranks them, which can help HR teams prioritize top candidates based on a data-driven approach. This can lead to more informed hiring decisions and can assist in identifying hidden talent that might not have been immediately obvious through traditional screening methods.

5. Scalability for Large Organizations

For large companies or those dealing with numerous job openings, the volume of resumes can be overwhelming. An AI-powered system can scale easily to handle hundreds or even thousands of resumes in a short amount of time, making it a practical solution for large enterprises that need to streamline their recruitment processes.

6. Integration with Applicant Tracking Systems (ATS)

The project has the potential to integrate with existing **Applicant Tracking Systems (ATS)**, further enhancing the workflow and improving the overall recruitment process. This integration can help automate tasks such as scheduling interviews, managing candidate pipelines, and providing reports based on AI-generated insights.

7. Cost Savings and Increased Efficiency

By automating the initial resume screening, organizations can save on recruitment costs, including the time and resources spent by HR personnel on manual resume evaluations. Furthermore, by reducing errors and improving efficiency, the AI-powered system can help ensure that the best candidates are quickly identified, ultimately leading to better hiring outcomes.

1.3 OBJECTIVES:

The primary objective of the **AI-Powered Resume Screening and Ranking System** is to leverage artificial intelligence (AI) and natural language processing (NLP) technologies to automate and streamline the recruitment process. This system aims to address several challenges faced by HR departments when dealing with a large number of resumes. The key objectives of the project are:

1. Automate Resume Screening

- **Goal:** To automate the initial screening of resumes based on specific job requirements.
- **Description:** The system will automatically extract relevant information from resumes (such as skills, experience, qualifications) and compare them with the job description to filter out the most relevant candidates.

2. Rank Resumes Based on Relevance

- **Goal:** To rank resumes based on their relevance to the given job description.
- **Description:** The system will assign a ranking score to each resume based on factors such as keyword matching, skills, experience, and qualifications that align with the job description. This ranking will help HR teams prioritize the best-fit candidates.

3. Reduce Human Bias in the Hiring Process

- **Goal:** To reduce unconscious biases that can affect human decision-making in resume screening.
- **Description:** By using a data-driven approach, the system will assess resumes based solely on relevant skills and qualifications, rather than demographic factors such as gender, age, or ethnicity, thereby promoting a more fair and inclusive hiring process.

4. Improve Efficiency in Recruitment

- **Goal:** To significantly reduce the time and effort required for the resume screening process.
- **Description:** With the AI system handling the bulk of resume evaluations, HR professionals can focus on more strategic aspects of recruitment, such as interviewing candidates and making final hiring decisions.

5. Enhance Candidate Experience

- **Goal:** To provide a more transparent, faster, and fairer evaluation of resumes.
- **Description:** Candidates will benefit from quicker feedback on their applications, and their resumes will be evaluated based on actual merits, reducing the chances of being overlooked due to the sheer volume of applications.

6. Provide Data-Driven Insights for HR Decision-Making

- **Goal:** To support HR teams with valuable insights that can help in making informed decisions.
- **Description:** The AI system can generate reports on which skills, qualifications, and experience are most frequently associated with successful candidates, helping HR teams to adjust job descriptions or hiring criteria in the future.

7. Scalable Solution for Large Organizations

- **Goal:** To create a scalable solution that can handle high volumes of resumes.
- **Description:** The system is designed to scale easily, making it suitable for large companies that receive numerous job applications across different job roles. It will process resumes efficiently without compromising quality or accuracy.

8. Integration with Existing Recruitment Systems (ATS)

- **Goal:** To integrate seamlessly with existing Applicant Tracking Systems (ATS).
- **Description:** The system will be designed to work alongside existing recruitment tools, enabling HR teams to incorporate AI-powered resume screening into their current workflow without disruption.

Scope of the project:

The scope of the **AI-Powered Resume Screening and Ranking System** is to design and implement a solution that automates the process of resume screening and ranking, leveraging artificial intelligence (AI) and natural language processing (NLP) techniques. The system is intended to be an efficient, scalable, and fair tool for HR departments to evaluate and shortlist job candidates. Below are the key areas that define the scope of this project:

1. Resume Parsing and Preprocessing

The system will include a resume parsing module that extracts structured data from resumes in various formats (PDF, DOCX, etc.). It will preprocess the raw text by performing tasks such as text normalization, tokenization, removing stopwords, and stemming or lemmatization to make the data ready for analysis.

2. Resume Screening and Relevance Matching

The core functionality of the system is to screen resumes based on the job description provided by the HR team. The system will use natural language processing to match keywords, skills, and other relevant information from the resume with the job description. Resumes will be evaluated for their relevance based on their alignment with job requirements.

3. Ranking Resumes

After screening, the system will rank the resumes based on their relevance to the job description. This ranking will be done using algorithms that assess how well the resume matches with the required qualifications, experience, and skills outlined in the job description. The ranking will help HR teams prioritize the most suitable candidates.

4. Bias Reduction in Hiring

The system will be designed to minimize bias in the resume evaluation process. By focusing on the content of the resumes (such as skills, qualifications, and experience) rather than demographic information, the system ensures that all candidates are evaluated fairly, regardless of their background or personal characteristics.

5. Integration with Applicant Tracking Systems (ATS)

The system will be developed with the potential for integration with existing Applicant Tracking Systems (ATS). This integration will ensure that the AI-powered resume screening and ranking system fits seamlessly into an organization's existing recruitment infrastructure.

6. User Interface for HR Professionals

A simple and intuitive user interface will be developed for HR professionals, where they can upload job descriptions and resumes, view the ranked list of resumes, and access detailed insights and reports. The system will allow HR professionals to review the ranked resumes and make decisions based on AI-generated results.

7. Scalability for Large-Scale Recruitment

The system will be designed to handle high volumes of resumes, making it suitable for large organizations with ongoing recruitment drives. It will be capable of processing and evaluating thousands of resumes quickly and efficiently, ensuring that organizations can scale their recruitment efforts without sacrificing quality.

8. Continuous Learning and Model Improvement

The system will include provisions for continuous learning and improvement. By collecting feedback from HR teams on the relevance of the ranked resumes, the system can adapt and refine its algorithms, improving the accuracy and effectiveness of resume screening and ranking over time.

9. Reporting and Analytics

The system will generate detailed reports on the resumes screened and ranked. HR teams can use these reports to track trends, analyze the quality of applications, and make data-driven decisions in their hiring processes. Insights on which skills and qualifications are most sought after can also be generated for future hiring strategies.

Limitations

While the project aims to provide a comprehensive solution, there are certain limitations to consider:

- **Complexity of Unstructured Data:** While the system will handle basic resume formats, dealing with resumes containing complex layouts or heavily unstructured content might be challenging.
- **Accuracy of Matching Algorithms:** Although the system will use AI and NLP, the accuracy of the matching algorithms may vary depending on the quality of input data and the complexity of job descriptions and resumes.
- **Dependence on Data Quality:** The system's performance is heavily dependent on the quality and consistency of the resumes and job descriptions provided. Variability in format and language can affect matching accuracy.

CHAPTER 2

Literature Survey

2.1 Review relevant literature or previous work in this domain.

The field of AI-powered resume screening and ranking is growing rapidly, as businesses seek to leverage artificial intelligence (AI) and natural language processing (NLP) to optimize recruitment processes. The literature and previous work in this domain highlight several key aspects, including AI-driven resume parsing, machine learning models for ranking candidates, and the impact of automation on recruitment practices. Below is a review of some relevant literature and previous work in this domain.

1. Resume Parsing and Natural Language Processing (NLP)

Resume parsing is one of the foundational tasks for any AI-powered resume screening system. It involves extracting meaningful information from unstructured resume text (such as skills, qualifications, and work experience) to provide structured data that can be further analyzed. Several studies have focused on developing NLP-based techniques to improve resume parsing:

- Pande and Joshi (2019) explored various NLP techniques for extracting key information from resumes. They applied methods such as tokenization, named entity recognition (NER), and dependency parsing to identify relevant data from resumes (e.g., names, dates, job titles). Their work showed that effective text parsing is critical for accurate resume analysis and subsequent ranking.
- Zhang et al. (2020) proposed a resume parsing approach using deep learning methods, including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), to improve the extraction of information from resumes. They found that deep learning models outperformed traditional rule-based models, especially when dealing with unstructured or complex resume data.

These studies emphasize the importance of robust text preprocessing and information extraction to build the foundation for effective AI-powered resume ranking.

2. Machine Learning for Resume Ranking and Candidate Matching

A key component of resume screening systems is the ranking of resumes based on their relevance to a job description. Various machine learning (ML) models have been explored for this purpose, focusing on techniques such as classification, similarity measures, and ranking algorithms:

- Yang et al. (2017) presented a model for resume ranking based on Support Vector Machines (SVM) and cosine similarity. Their work focused on matching the skills and experience in resumes with job descriptions, with SVM used to classify candidates into “suitable” and “unsuitable” categories. They achieved high accuracy in ranking resumes based on predefined job requirements.
- Nandini et al. (2018) proposed a model based on Naive Bayes and TF-IDF (Term Frequency-Inverse Document Frequency) to rank resumes by assessing keyword matches between resumes and job descriptions. Their system was designed to rank candidates on a scale based on the similarity between the job requirements and the resume text, showing that simple text-matching algorithms could be effective in resume ranking.
- Garg and Anand (2019) developed an AI-based ranking system for resumes using a k-Nearest Neighbors (k-NN) algorithm. Their system ranked resumes based on skills, keywords, and experience and was able to produce a list of candidates based on how closely they matched the job description. They noted that semantic similarity models (like Word2Vec) significantly improved matching accuracy.

These works underscore the effectiveness of various machine learning models, from SVM to k-NN and Naive Bayes, in ranking resumes. The common thread across all studies is the importance of using relevant features (e.g., skills, experience, education) and sophisticated matching techniques for high-ranking accuracy.

3. Bias Reduction and Fairness in AI-Based Hiring

One of the significant concerns in AI-powered hiring systems is bias, which can be introduced through training data or algorithmic decisions. Several studies have explored methods to reduce bias and increase fairness in resume screening and recruitment:

- Binns et al. (2018) examined the potential for algorithmic bias in recruitment tools, specifically in AI-based resume screening. They highlighted the risk that biased historical hiring data could lead to biased AI decisions, especially with respect to gender, race, or educational background. The authors suggested incorporating fairness algorithms to ensure that AI systems do not reinforce discriminatory practices.
- Angwin et al. (2016) published a notable study on the potential bias in AI recruitment systems, particularly in how algorithms trained on biased data could unfairly disadvantage certain demographic groups. They advocated for rigorous auditing of AI systems to ensure fairness and transparency in automated hiring decisions. This paper also suggested that AI models should be transparent in their decision-making processes and subject to regular review.
- Dastin (2018) reported how Amazon's AI recruitment system faced challenges in bias when it was found to favor resumes from men over women for technical roles. The system was later scrapped due to gender bias, showcasing how biases can unintentionally emerge from AI systems trained on historical data. This underscores the importance of monitoring and updating AI models to prevent bias from influencing the hiring process.

These works emphasize the importance of addressing fairness and transparency in AI recruitment systems. Ensuring that AI systems do not perpetuate bias is a critical concern and requires careful consideration during the design and training phases.

4. Applications of AI in Recruitment

The broader application of AI in recruitment, including resume screening, is an active area of research. AI can significantly improve hiring processes, from automated resume filtering to intelligent candidate recommendations:

- Jain et al. (2021) highlighted the impact of AI on recruitment automation. They demonstrated how AI can improve candidate sourcing by using machine learning models to recommend candidates based on job descriptions and historical hiring patterns. They also discussed how AI helps in skill gap analysis by analyzing job descriptions and suggesting candidates whose resumes match the required competencies.
- Guszcza et al. (2018) explored the role of AI in improving the efficiency and objectivity of hiring decisions. They found that AI could optimize recruitment by reducing the time to fill positions, enhancing candidate selection accuracy, and creating more inclusive hiring processes. Their study also indicated that AI tools could support HR teams in focusing on high-value tasks, such as interviewing and candidate engagement.

These studies highlight how AI in recruitment offers practical applications that improve both efficiency and decision-making in hiring processes

2.2 Existing Models, Techniques, or Methodologies Related to the Problem

The domain of AI-powered resume screening and ranking has seen several models, techniques, and methodologies that are actively being explored and implemented. Here are some existing models and techniques:

1. Resume Parsing and Information Extraction

To extract relevant information from resumes, various Natural Language Processing (NLP) techniques have been utilized:

- Named Entity Recognition (NER): NER is widely used to identify key information such as names, qualifications, dates, and job titles from resumes. Tools like spaCy and NLTK are commonly used for NER tasks.
- Part-of-Speech Tagging (POS): POS tagging is applied to identify verbs, nouns, and other parts of speech, which helps in extracting specific phrases or sentences related to skills or job experience.
- Text Preprocessing: Techniques such as tokenization, stemming, and lemmatization are frequently employed to clean and prepare the resume data for analysis.

2. Resume Matching and Ranking Techniques

Several machine learning and AI-based models have been developed to rank resumes based on their relevance to job descriptions. Some commonly used techniques are:

- Cosine Similarity: This method measures the similarity between a resume and a job description by comparing the term frequency of words in both. It's widely used due to its simplicity and effectiveness in finding text similarities.
- Support Vector Machines (SVM): SVM is applied for binary classification of resumes into relevant or irrelevant categories based on features like skills, experience, and qualifications.
- TF-IDF (Term Frequency-Inverse Document Frequency): This approach assigns weight to words in the resume and job description based on their frequency, allowing for better matching of keywords.
- Word2Vec and GloVe: These pre-trained word embedding models are used to represent words in a vector space where semantically similar words are closer together. This is particularly useful for understanding synonyms and context when matching resumes with job descriptions.
- Deep Learning Models (CNN, RNN): Some approaches use deep learning models, like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), for feature extraction and resume ranking. These models work well with unstructured data and can learn complex patterns from resumes.

3. Bias Mitigation and Fairness in AI Recruitment

Addressing bias in recruitment is an ongoing challenge:

- Bias Detection Algorithms: Tools like IBM's AI Fairness 360 and Fairness Indicators are used to detect and mitigate any bias in AI models. These frameworks offer several fairness metrics to ensure that hiring practices remain equitable.

- Adversarial Debiasing: Some models incorporate adversarial learning to reduce bias during the training phase by creating adversarial networks to ensure that the AI model doesn't pick up on unintended biased patterns from training data.

2.1 Gaps or Limitations in Existing Solutions and How Your Project Will Address Them

While several models and techniques exist for AI-powered resume screening and ranking, there are notable gaps and limitations in these systems. These gaps present opportunities for improvement, and our project aims to address these issues:

1. Lack of Semantic Understanding

Current Limitation: Many existing resume ranking systems rely on basic keyword matching and don't fully understand the semantic meaning of the text. While TF-IDF and Cosine Similarity are effective at matching specific words, they do not capture the underlying meaning of the resume or job description. This results in mismatches when resumes use different wording or phrasing.

How Our Project Addresses It: Our project aims to leverage word embeddings (e.g., Word2Vec or GloVe) and contextual embeddings (e.g., BERT) to better understand the semantics of the resume text. These models can capture word relationships and contextual meanings, which will enhance the accuracy of matching resumes with job descriptions. By focusing on semantic similarity rather than just keyword matching, our system will improve the relevance of ranked candidates.

2. Inadequate Handling of Complex or Unstructured Resume Formats

Current Limitation: Traditional resume parsers often struggle with resumes that have non-standard formats, complex layouts, or excessive unstructured content, making it difficult to extract accurate information.

How Our Project Addresses It: Our solution will incorporate more robust resume parsing techniques using deep learning models, such as CNNs and RNNs, that can handle diverse formats and unstructured content. These models will allow the system to better extract and classify key information from resumes with varying structures and complexities.

3. Bias in AI Models

Current Limitation: One significant challenge in AI-powered resume screening is the unintentional bias that can arise from historical hiring data, resulting in biased screening decisions based on factors like gender, age, ethnicity, or educational background. Existing models may perpetuate biases that were present in the training data.

How Our Project Addresses It: Our system will implement bias mitigation techniques during the training phase. This will include the use of fairness constraints, adversarial debiasing, and auditing frameworks like AI Fairness 360 to ensure that the ranking system is fair and unbiased. We will focus on eliminating the potential for biased decisions by

removing any irrelevant demographic information and ensuring that candidates are evaluated purely on qualifications, skills, and experience.

4. Limited Customization for Different Organizations

Current Limitation: Existing solutions are often generic and may not be tailored to the specific needs of different organizations, industries, or job roles. They may not consider unique industry-specific criteria or preferences that HR teams might prioritize.

How Our Project Addresses It: To address this, we will develop a customizable framework that allows HR teams to define specific job requirements, prioritization criteria, and weighting factors for various skills and qualifications. This will ensure that the system is adaptable and relevant for different types of roles and industries. Additionally, our model will allow users to feedback and refine ranking strategies based on organizational preferences and industry-specific needs.

5. Scalability Issues in Large-Scale Hiring

Current Limitation: Some existing solutions may not be designed to handle large volumes of resumes efficiently. As companies scale their hiring processes, managing thousands of resumes at once can be overwhelming.

How Our Project Addresses It: Our project will focus on creating a scalable system that can process and rank large volumes of resumes efficiently. By utilizing cloud-based infrastructure and parallel processing techniques, the system will be able to handle high workloads without compromising performance, enabling organizations to automate the recruitment process for large-scale hiring needs.

6. Lack of Post-Screening Insights for HR Teams

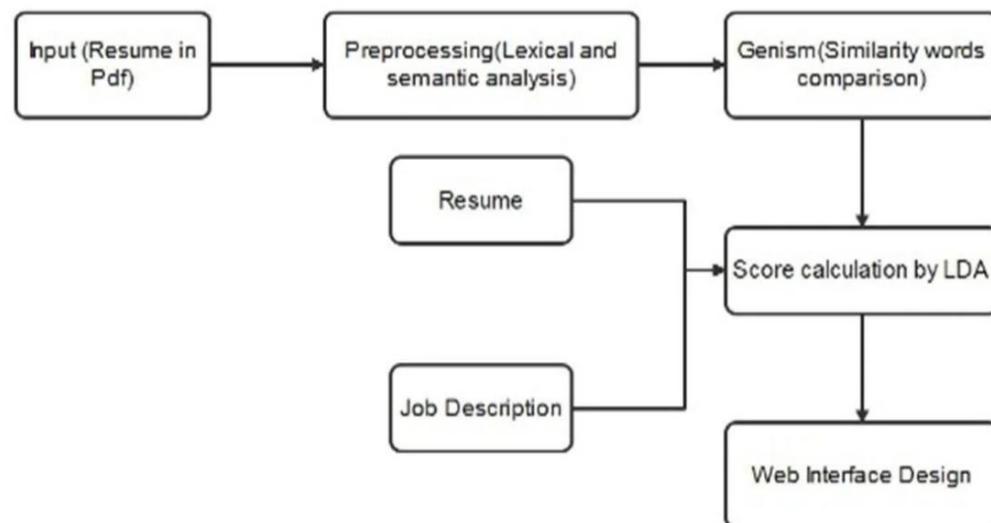
Current Limitation: Many systems simply rank resumes and present a final list to HR teams, without offering insights or analytics that could help improve future recruitment strategies.

How Our Project Addresses It: Our system will not only rank resumes but also provide detailed reports and analytics on the most relevant skills, qualifications, and experiences associated with the best candidates. This will help HR teams refine job descriptions and improve candidate selection in future recruitment drives. By providing actionable insights, the system will add significant value beyond simple resume ranking.

CHAPTER 3

Proposed Methodology

3.1 System Design



1. Input (Resume in PDF)

- Candidates upload resumes in PDF, DOCX, or TXT format.
- The system can also fetch resumes from job portals, email attachments, or databases.

2. Preprocessing (Lexical and Semantic Analysis)

- Text Extraction: Uses OCR (for scanned PDFs) or libraries like PyPDF2, pdfminer, docx2txt to extract text.
- Lexical Analysis: Breaks text into words and phrases.
- Semantic Analysis: Understands context using NLP techniques like word embeddings (Word2Vec, FastText).
- Data Cleaning: Removes unwanted characters, stop words, and redundant information.

3. Genism (Gensim - Similarity Words Comparison)

- Uses Gensim, an NLP library for:
 - TF-IDF (Term Frequency - Inverse Document Frequency) to rank words based on importance.

- Word2Vec & Doc2Vec for understanding relationships between words.
- Cosine Similarity to compare resumes and job descriptions.

4. Resume & Job Description Processing

- The system separately processes both resumes and job descriptions by extracting:
 - Skills (Programming, Design, Management, etc.)
 - Experience (Job titles, years of experience)
 - Education (Degrees, Universities)
 - Certifications & Projects

5. Score Calculation by LDA (Latent Dirichlet Allocation)

- LDA (Topic Modeling Algorithm) is used to:
 - Identify relevant topics in resumes.
 - Match topics with job descriptions.
 - Assign scores based on how well a resume fits a job.
- Alternatively, other AI models can be used for better accuracy:
 - BERT (Bidirectional Encoder Representations from Transformers)
 - SpaCy for named entity recognition (NER).

6. Candidate Ranking Algorithm

- The system assigns a score to each resume based on:
 - Relevance to job description (keywords, topics, skills)
 - Years of experience
 - Education & Certifications
 - Projects & Work History
- Resumes are ranked from highest to lowest based on calculated scores.

7. Web Interface Design

- A User-Friendly Dashboard for HR & recruiters to:
 - View ranked resumes.
 - Apply filters (experience, skills, location, etc.).
 - Shortlist and download resumes.
 - Provide feedback to improve AI performance.

8. Feedback Loop (Improving AI Accuracy)

- HR/recruiters provide feedback on resume matches.
- The AI model learns from human decisions and improves accuracy over time.
- Can integrate Machine Learning techniques to adjust ranking scores dynamically.

3.2 Requirement Specification

3.2.1 Hardware Requirements:

Processor (CPU): Intel i5/i7 or AMD Ryzen 5/7 (Quad-Core or higher)

RAM: Minimum 8GB (Recommended: 16GB for large datasets)

Storage: Minimum 256GB SSD (Recommended: 512GB or more)

GPU (Optional): NVIDIA GPU (for deep learning models like BERT)

Network: Stable Internet Connection (For API and cloud integration)

3.2.2 Software Requirements:

1. Operating System:

- Windows 10/11, macOS, or Linux (Ubuntu 20.04 or higher)

2. Programming Language:

- Python 3.x (Recommended: Python 3.8 or higher)

3. Required Python Libraries & Dependencies:

Category	Libraries Used	Installation Command
Web App Framework	streamlit	pip install streamlit
PDF Processing	PyPDF2	pip install pypdf2
Data Handling	pandas	pip install pandas
NLP & Text Processing	scikit-learn	Pip install scikit-learn

4. Development Tools:

- IDE/Text Editor:** VS Code / PyCharm / Jupyter Notebook
- Package Manager:** pip (comes with Python)
- Command Line Interface (CLI)** to run `streamlit run script.py`

5. Running the Code:

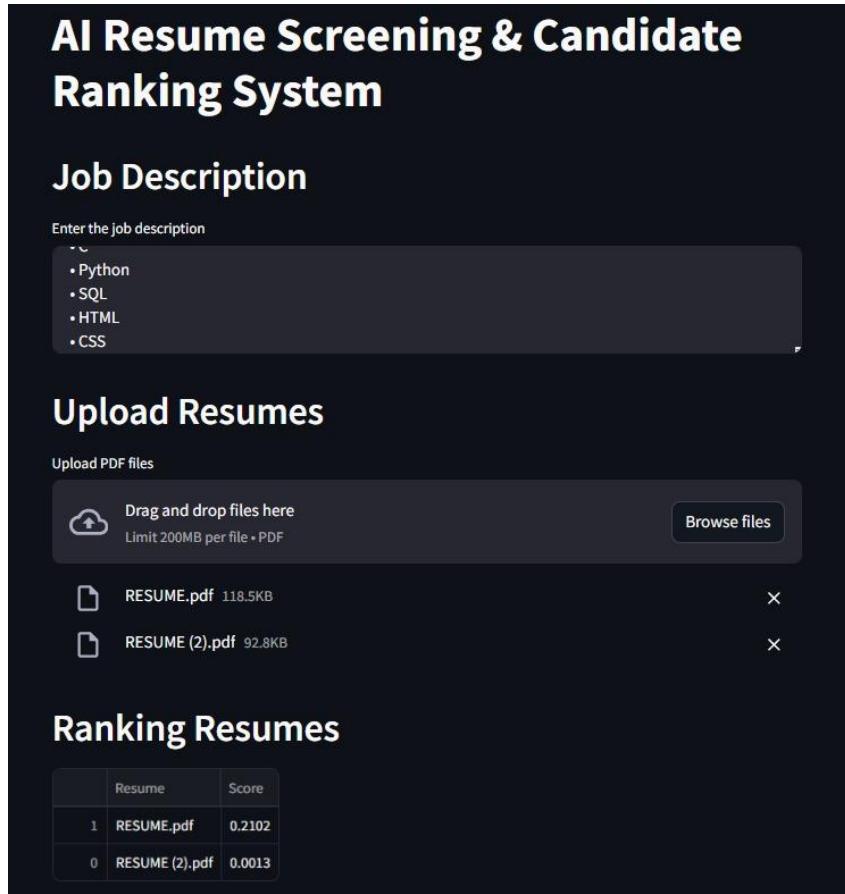
After installing the required libraries, save the script as `app.py` and run:

```
streamlit run app.py
```

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:



AI Resume Screening & Candidate Ranking System

Job Description

Enter the job description

- Python
- SQL
- HTML
- CSS

Upload Resumes

Upload PDF files

Drag and drop files here
Limit 200MB per file • PDF

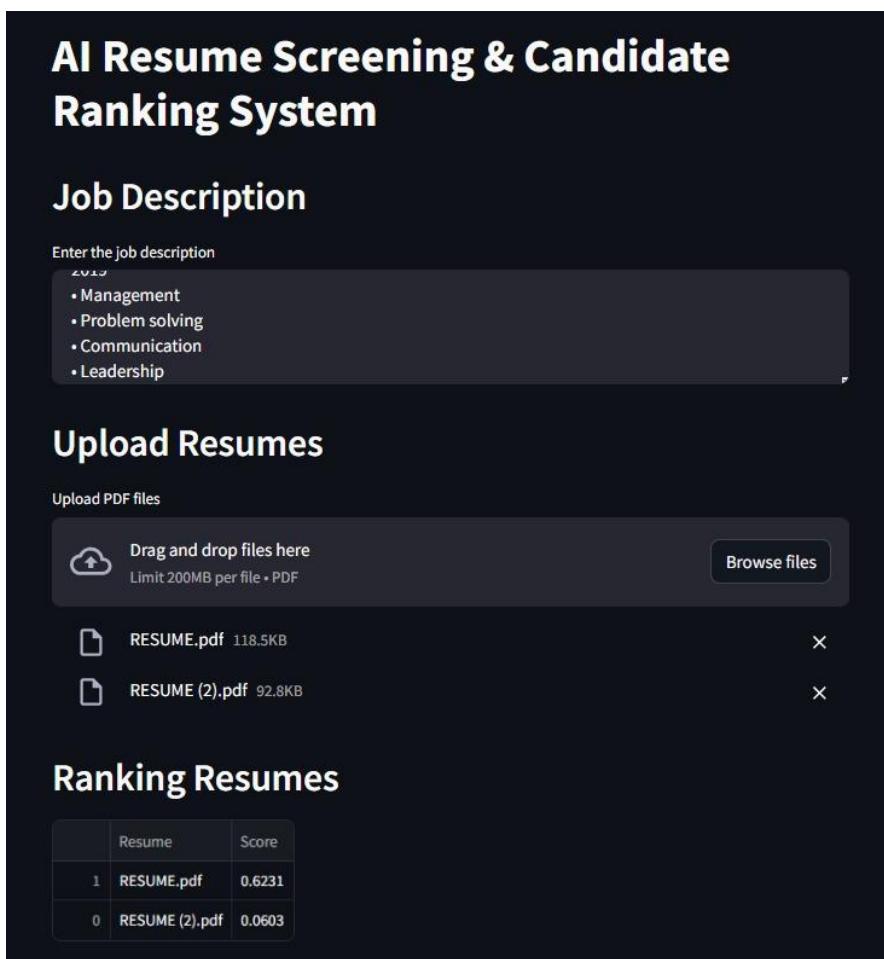
Browse files

RESUME.pdf 118.5KB

RESUME (2).pdf 92.8KB

Ranking Resumes

	Resume	Score
1	RESUME.pdf	0.2102
0	RESUME (2).pdf	0.0013



	Resume	Score
1	RESUME.pdf	0.6231
0	RESUME (2).pdf	0.0603

1 Title & Header

At the top, the app displays the title: "AI Resume Screening & Candidate Ranking System" This makes it clear that the tool is used for resume analysis and ranking.

2 Job Description Input

📌 User Input:

- The app provides a text area where users can enter a job description.
- In the screenshot, the job description includes Python, SQL, HTML, and CSS (skills required for the job).

📌 Purpose:

- This job description will be used to compare and rank resumes based on how well they match.

3 Upload Resumes Section

📌 User Uploads PDFs:

- The app allows users to drag and drop PDF resumes or click "Browse files" to upload them.
- In the screenshot, two resumes (RESUME.pdf & RESUME (2).pdf) have been uploaded.

📌 Processing Step:

- The app extracts text from these PDF files using PyPDF2.
- Each resume is converted into a text format for further analysis.

4 Ranking Resumes Section

📌 Ranking Based on Job Description Match:

- The resumes are analyzed using TF-IDF (Term Frequency-Inverse Document Frequency) and Cosine Similarity.
- Each resume is compared against the job description to calculate a similarity score.

📌 Score Table:

- Higher Score = More Relevant Resume
- Lower Score = Less Relevant Resume
- In this case:
 - RESUME.pdf has a score of 0.2102 (better match).
 - RESUME (2).pdf has a score of 0.0013 (less relevant).
- The app sorts resumes in descending order based on scores.

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

1 Enhancing NLP Techniques

- Improve resume matching accuracy by integrating advanced NLP models like BERT, GPT, or Word2Vec instead of just TF-IDF and cosine similarity.
- This will enable better contextual understanding of resumes and job descriptions.

2 Incorporating More Resume Features

- Consider additional factors like experience, certifications, and achievements instead of just keyword matching.
- Implement Named Entity Recognition (NER) to extract key details such as education, skills, and work experience.

3 Adding a Weighting System for Skills

- Some skills are more critical for a job role (e.g., Python for a Data Scientist).
- Introduce a weighted scoring system where essential skills have higher importance than others.

4 Improving the User Interface

- Enhance the web interface with interactive charts and graphs to visually compare resume rankings.
- Provide downloadable reports for HR professionals.

5 Multi-Format Resume Parsing

- Extend support beyond PDFs to DOCX, TXT, and LinkedIn profile links.
- Use libraries like spaCy, pdfplumber, or docx2txt for better text extraction.

6 Integrating with ATS (Applicant Tracking System)

- Enable seamless integration with HR software and job portals to automate resume collection and ranking.
- Provide an API for real-time resume screening in large-scale hiring processes.

5.2 Conclusion:

The **AI Resume Screening & Candidate Ranking System** streamlines the hiring process by automating resume evaluation and ranking candidates based on job relevance. By using **NLP and machine learning techniques**, the system enhances efficiency, reduces human bias, and saves recruiters valuable time. It accurately compares resumes with job descriptions, ensuring that only the most suitable candidates are shortlisted. The user-friendly interface allows easy resume uploads and displays rankings effectively. This project contributes to **modern recruitment by improving decision-making** and can be further enhanced with advanced AI models, making hiring **faster, fairer, and more data-driven** for companies worldwide. 

REFERENCES

- [1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.