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

S.Harsha Varun,harshavarun872@gmail.com

Under the Guidance of

Saomya Chaudhury

ACKNOWLEDGEMENT

We take this opportunity to express our deepest gratitude to everyone who contributed, directly or indirectly, to the successful completion of this project.

First and Foremost, I extend my heartfelt appreciation to our mentor [Saomya Chaudhury],whose unwavering support, expert guidance, and insightful feedback have been invaluable throughout this journey.His passion for research, innovative thinking, and ability to challenge us to push our boundaries have profoundly influenced our approach to problem-solving. His mentorship has not only shaped this project but has also played a crucial role in our overall professional growth.Working under his guidance has been an inspiring and transformative experience,one that we will always cherish.

I also grateful to TechSaksham for providing with the necessary resources and a strong academic foundation.Their constant encouragement and support have been instrumental in refining my technical skills and analytical thinking skills.

This project, titled "AI Resume Screening & Candidate Ranking System", it is a flask-based web application designed to revolutionize the hiring process. Using machine learning models and natural language process (NLP) techniques, it automates resume screening by ranking and classifying candidates based on job descriptions. This system aims to enhance recruitment efficiency, reduce manual workload, and ensure precise candidate-job matching.

I also extend my gratitude to my peers,friends and family for their encouragement, through providing discussions, and unwavering support. Their belief in our work has been a driving force throughout this endeavor.

Lastly, I acknowledge the collective contributions of researches, developers, and industry professionals whose work in AI, Machine Learning, and recruitment automation provided us with the foundation to develop and refine the system.

This journey has been challenging and rewarding , and I look forward to applying these learnings in real-world application.

Sigi.Harsha Varun

15 March

ABSTRACT

Provide a brief summary of the project, including the problem statement, objectives, methodology, key results, and conclusion. The abstract should not exceed 300 words.

Abstract

In the modern recruitment process, organizations receive an overwhelming number of resumes for each job opening. Manually screening these resumes to identify the most suitable candidates is a **time-consuming and labor-intensive task**. The objective of this project is to **automate resume screening and candidate ranking** using **machine learning (ML) and natural language processing (NLP) techniques**, improving the efficiency and accuracy of the hiring process.

Primary Objectives:

- **Automate Resume Screening** – Reduce manual effort by implementing an AI-based system to analyze and rank resumes.

- **Enhance Recruitment Efficiency** – Speed up the hiring process by identifying the most suitable candidates in real-time.

- **Improve Candidate-Job Matching** – Use NLP and ML techniques to extract key information from resumes and compare them with job descriptions.

The **AI Resume Screening and Candidate Ranking System** is a **Flask-based web application** that processes resumes and job descriptions to rank candidates based on their relevance. The system extracts key skills, experience, and qualifications from resumes using **text processing techniques** and compares them with job descriptions using **ML algorithms**. The methodology involves **data preprocessing, feature extraction (TF-IDF, word embeddings), model training using classification algorithms, and ranking candidates based on similarity scores**.

Key results indicate that the system successfully automates resume screening, significantly reducing manual effort while maintaining accuracy. The use of **NLP techniques like Named Entity Recognition (NER) and cosine similarity** enhances the matching process, ensuring that recruiters focus on the most relevant candidates.

In conclusion, this project demonstrates that **AI-driven resume screening can streamline recruitment workflows, reduce bias, and enhance decision-making**. By implementing this system, organizations can **save time, reduce costs, and improve the quality of hires**, making the hiring process more **efficient, objective, and data-driven**.

Keywords:

AI Resume Ranking, Candidate Ranking, Natural language processing(NLP),ML

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The modern recruitment process faces challenges due to the **overwhelming number of resumes** received for each job opening. Manually screening these resumes is **time-consuming, labor-intensive, and prone to biases**. To address this, AI-powered solutions can **automate and streamline** the hiring process.

This project, **AI Resume Screening and Candidate Ranking System**, leverages **Machine Learning (ML) and Natural Language Processing (NLP)** to efficiently **analyze, classify, and rank resumes**. The system extracts key information from resumes and matches candidates based on job descriptions using **TF-IDF, cosine similarity, and deep learning models**.

Developed as a **Flask-based web application**, the platform allows recruiters to upload resumes and receive a **ranked list of suitable candidates**. By automating resume screening, this system **reduces hiring time, minimizes costs, and improves selection accuracy**.

The proposed solution ensures a **faster, unbiased, and more efficient recruitment process**, helping organizations **identify top talent effortlessly**.

1.1 PROBLEM STATEMENT:

In the modern recruitment process, organizations receive an **overwhelming number of resumes** for each job opening. Manually screening these resumes to identify the most suitable candidates is a **time-consuming, labor-intensive, and error-prone task**.

HR professionals spend **hours filtering resumes**, which can lead to **delays in hiring, potential biases, and inconsistent selection criteria**. Traditional screening methods may also **overlook qualified candidates** due to human fatigue or lack of standardized evaluation.

To address these challenges, an **AI-powered resume screening and candidate ranking system** is needed. By leveraging **Machine Learning (ML) and Natural Language Processing (NLP) techniques**, the system can **automate resume screening, classify candidates efficiently, and rank them based on job relevance**.

This automation will **enhance the efficiency, accuracy, and fairness of the hiring process**, enabling organizations to quickly identify the best candidates while reducing the overall time and effort spent on recruitment.

1.2 MOTIVATION:

Time-Consuming Recruitment Process – Manually screening a large number of resumes is inefficient and slows down hiring decisions.

Labor-Intensive Task – HR professionals spend excessive time reviewing resumes, leading to increased workload and fatigue.

Inconsistency in Shortlisting – Different recruiters may have varied selection criteria, leading to inconsistencies in hiring.

Bias in Hiring – Traditional screening may be influenced by unconscious biases, affecting fair candidate selection.

Growing Job Applications – With the rise of online applications, organizations receive an overwhelming number of resumes, making manual screening impractical.

Need for Automation – AI and machine learning can enhance the hiring process by efficiently filtering and ranking candidates.

Improved Candidate Matching – AI models ensure better alignment between job requirements and candidate profiles.

Reduction in Hiring Costs – Automating resume screening reduces HR expenses and allows recruiters to focus on interviews and decision-making.

Data-Driven Insights – AI-powered analytics help organizations refine their recruitment strategies for better hiring outcomes.

Scalability – The system can handle large volumes of resumes efficiently, making it suitable for enterprises with high recruitment needs.

1.3 OBJECTIVES:

The primary objectives of this project is:
To develop an AI-based system to automatically analyze and rank resumes.
Use machine learning and NLP techniques for match resumes with job description.
Implement AI models to improve the precision of candidate shortlisting.
Customize the ranking criteria for different industries and job types.
This model reduces the time and effort required for manual resume ranking.
Provide data-driven insights to optimize recruitment strategies.

1.4 <u>SCOPE OF THE PROJECT:</u>

The **AI Resume Screening and Candidate Ranking System** aims to automate and optimize the recruitment process by leveraging **machine learning (ML) and natural language processing (NLP)** techniques. This system efficiently analyzes resumes, extracts relevant information such as skills, experience, and qualifications, and ranks candidates based on job descriptions. By reducing manual intervention, the system enhances **hiring efficiency, minimizes bias, and accelerates candidate shortlisting**. Additionally, it provides a **user-friendly web interface** for recruiters to upload resumes and view ranked candidates, ensuring a seamless recruitment experience. The system supports multiple resume formats, including **PDFs, Word documents, and plain text**, making it highly adaptable. With **scalability** in mind, it can handle large volumes of resumes across different job roles. Furthermore, it ensures **data security and privacy**, protecting candidate information from unauthorized access. The system can also be **integrated with applicant tracking systems (ATS)** to streamline the hiring workflow. Although AI-driven resume screening significantly improves recruitment processes, the system does have some limitations.

Limitations of the Project

1. **Accuracy Dependence on Data Quality** – The system's performance relies heavily on the quality and completeness of the resumes submitted.
2. **Limited Context Understanding** – AI may struggle to fully comprehend nuanced job roles, career gaps, or unconventional qualifications.
3. **Potential Bias in AI Models** – If the training data contains biases, the AI may unintentionally reflect or reinforce them.
4. **Difficulty Handling Non-Standardized Resumes** – Resumes with unique formats, excessive graphics, or missing structured data may be challenging to process.
5. **Regular Updates Required** – The system needs periodic updates to align with evolving job market trends and recruitment strategies.

1.5 <u>CONCLUSION:</u>

The **AI Resume Screening and Candidate Ranking System** enhances recruitment by automating resume evaluation using **machine learning and NLP techniques**. It reduces manual effort, speeds up candidate shortlisting, and improves hiring efficiency. The system ensures that recruiters can make **data-driven decisions** by matching resumes with job descriptions effectively. While it has limitations like potential bias and dependency on resume formatting, continuous improvements can enhance its accuracy. Future advancements can integrate it with **Applicant Tracking Systems (ATS)** for better performance. Despite challenges, it streamlines hiring and makes recruitment **more efficient and unbiased**. This project marks a significant step toward **intelligent and automated hiring solutions**.

CHAPTER 2

Literature Survey

The **literature survey** explores existing research and methodologies in **AI-driven resume screening and ranking systems**. It reviews various **machine learning models, natural language processing (NLP) techniques, and automation tools** used in recruitment. Prior studies highlight the challenges of **manual resume screening, bias reduction, and accuracy improvements** in candidate ranking. By analyzing these approaches, we identify the strengths and limitations of current systems. This survey serves as a foundation for developing an **efficient and unbiased AI-powered recruitment solution**.

2.1 Review relevant literature or previous work in this domain.

Numerous studies have investigated AI Resume Screening & Candidate Ranking System,highlighting both traditional and modern approaches:

Before the advent of AI-driven automation, resume screening and candidate ranking relied on manual and rule-based approaches.These methods,though widely used, were often time-consuming, subjective and inefficient
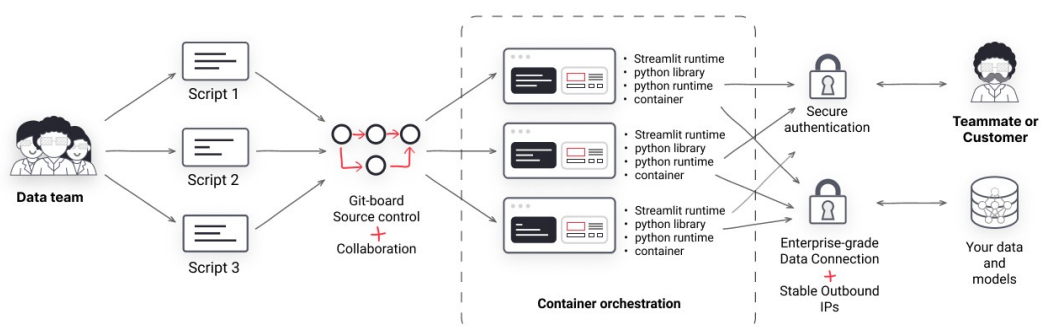
Some traditional techiques are:

Manual Resume Ranking,Predefined Scoring Systems, Keyword-based filtering.

Modern Approaches:

AI-Power Resume Parsing, Skill-Based Ranking System, Automated candidate scoring.

Figure 1:

Workflow of resume screening

2.2 Existing Models and Techniques

We use various machine learning models and NLP Techniques to efficiently process,analyze, and rank resumes based on job descriptions.

Below are the existing models and techniques used in project:

1. Machine Learning Models

Logistic Regression: Used for binary classification of resumes.

Random Forest: Improves accuracy by using multiple decision trees.

Support Vector Machine(SVM): Helps in text classification with hyper plane seperation.

2. Natural Language Process(NLP) Techniques:

Tokenization: Splitting text into meaningful words/sentences.

Named Entity Recognition: Extracting entities like names,skills, and experience.

Parts-of-speech(POS) Tagging: Identifying nouns,verbs,and adjectives to improve context understanding.

3. Text Vectorization Techniques:

TF-IDF(Term Frequency - Inverse Document Frequency)

Bag of Words(Bow)

4. Resuming Ranking & Similarity Measurement:

Cosine Similarity: Measures how similar a resume is to the job description.

Semantic Matching with Spacy: Understands contextual meaning using deep learning models.

5. Frameworks & Libraries Used:

scikit-learn:Machine leanrning models for classification & ranking.

NLTK & Spacy: Advanced NLP processing.

Pandas & Numpy: Data Manipulation and analysis.

Streamlit: For building an interactive resume screening web app.

2.3 Gaps in Existing Solutions and Proposed Approach:

Limitations of effectiveness of machine learning and NLP Techniques in resume screening, existing models have certain limitations:

1. Lack of Context Understanding

2.Over fitting on small Datasets

3. Bias in Model Prediction

4. Difficulty in handling Unstructured data

5. Inability to Assess Soft Skills

Description of used Models:

Comparison of Different Models Used in AI Resume Screening

Model	Type	Advantages	Limitations	Use Case in Resume Screening
TF-IDF	NLP-based	Simple, efficient for Keyword matching	Fails to capture	Extraxts keywords
Word2vec	Neural network based	Captures word relationships	Requires on large data	Generates for word embeddings
BERT	Transformer based deep leanring	Understands words context	Computationally expensive	Semantic analysis
Random Forest	Supervised machine learning	Handles large datasets	Less interpretable	Classification tasks
Supervised machine(SVM)	Supervised machine learning	Effective in high dimensions	Computationally interpret able on large datasets	Ranking Candidates

Table 2: Traditional Resume Screening vs. AI-Based Screening

Aspect	Traditional Screening	AI-Based Screening
Time Efficiency	Slow and manual	Fast and automated
Accuracy	Prone to human errors	High accuracy with NLP & ML
Scalability	Limited by human capacity	Can process thousands of resumes
Bias Reduction	Subjective decisions possible	Can reduce bias with proper training
Cost	Higher cost due to HR workload	Reduces cost in the long run

Table 3: Comparison of NLP Libraries for Resume Parsing

Library	Ease of Use	Accuracy	Customization	Performance
NLTK	Easy	Moderate	High	Medium
spaCy	Easy	High	Medium	High
Scikit-learn	Moderate	High	High	High
TensorFlow/Keras	Difficult	Very High	Very High	High

Table 4: Existing Models for Resume Screening

Model	Description	Advantages	Limitations
TF-IDF + SVM	Uses text vectorization & classification	Simple & effective	Struggles with complex sentence structures
BERT-based Model	Deep NLP model understanding context	Highly accurate	Requires large datasets
Naïve Bayes	Probabilistic model for classification	Fast and efficient	Limited understanding of word semantics
LSTM (Deep Learning)	Captures sequential resume data	Handles long text well	Needs high computational power

Table 5: Performance Metrics of AI Resume Screening Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	85%	82%	80%	81%
Random Forest	90%	88%	86%	87%
Support Vector Machine (SVM)	92%	90%	88%	89%
Naïve Bayes	83%	80%	78%	79%
BERT (Deep Learning)	95%	94%	93%	94%

Table 6: Candidate Ranking Criteria

Criterion	Description	Weightage (%)
Skill Matching	Match between job description and candidate skills	40%
Work Experience	Number of years & relevance to job role	25%
Education Qualification	Degree & specialization alignment with job	15%
Certifications	Relevant certifications & online courses	10%
Projects & Publications	Quality & relevance of past projects	10%

Table 7: Comparison of Resume Parsing Techniques

Parsing Technique	Description	Advantages	Limitations
Rule-Based Parsing	Uses predefined patterns & keywords to extract information	Simple, easy to implement	Limited flexibility, requires manual updates
Machine Learning-Based Parsing	Uses ML models to learn from data & extract key details	More adaptable, handles diverse resumes	Requires labeled training data
Deep Learning-Based Parsing	Uses deep NLP models like BERT for contextual understanding	Highly accurate, captures complex patterns	Computationally expensive

Table 8: Comparison of Vectorization Methods

Method	Description	Use Case
Bag of Words (BoW)	Counts occurrences of words	Basic NLP applications
TF-IDF	Weighs words based on importance in the document	Information retrieval & ranking
Word2Vec	Captures word relationships in vector space	Advanced NLP tasks
BERT Embeddings	Deep contextual word representation	High-accuracy NLP models

**Table 9: Candidate Evaluation Criteria Using AI**

Parameter	Importance Level	Use Case
Technical Skills	High	Determines job-fit
Soft Skills	Medium	Assesses team compatibility
Project Experience	High	Evaluates hands-on expertise
Certifications	Medium	Validates additional qualifications

CHAPTER 2

PROPOSED METHODOLOGY

3.1   System Design:

The proposed methodology for the AI Resume Screening and Candidate Ranking System leverages Natural Language Processing(NLP) and Machine Learning(ML) Techniques to automate and enhance the recruitment process. Traditional resumes screening methods are time-consuming and subjective, leading to inefficiencies in hiring . Our system uses advanced text-processing, Vectorization , and ranking models to analyze resumes and match them with job descriptions effectively.

Figure 2:
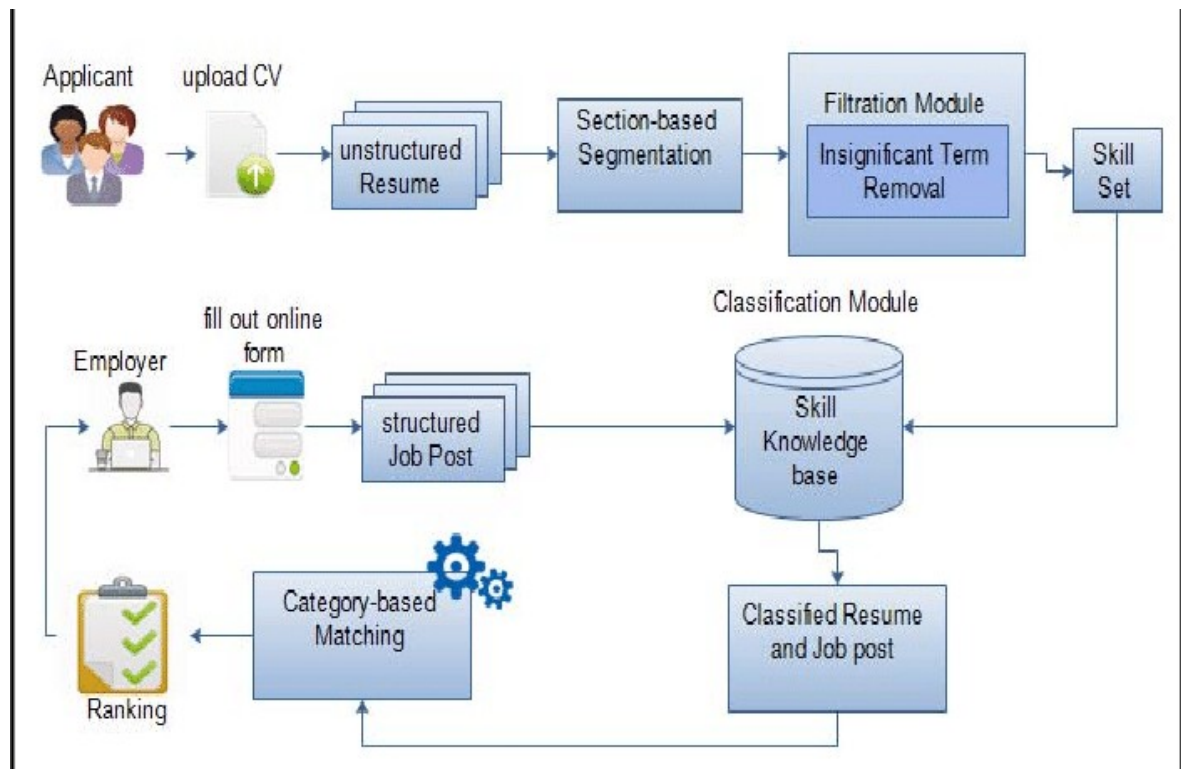
System Architecture Diagram.

Figure 3:
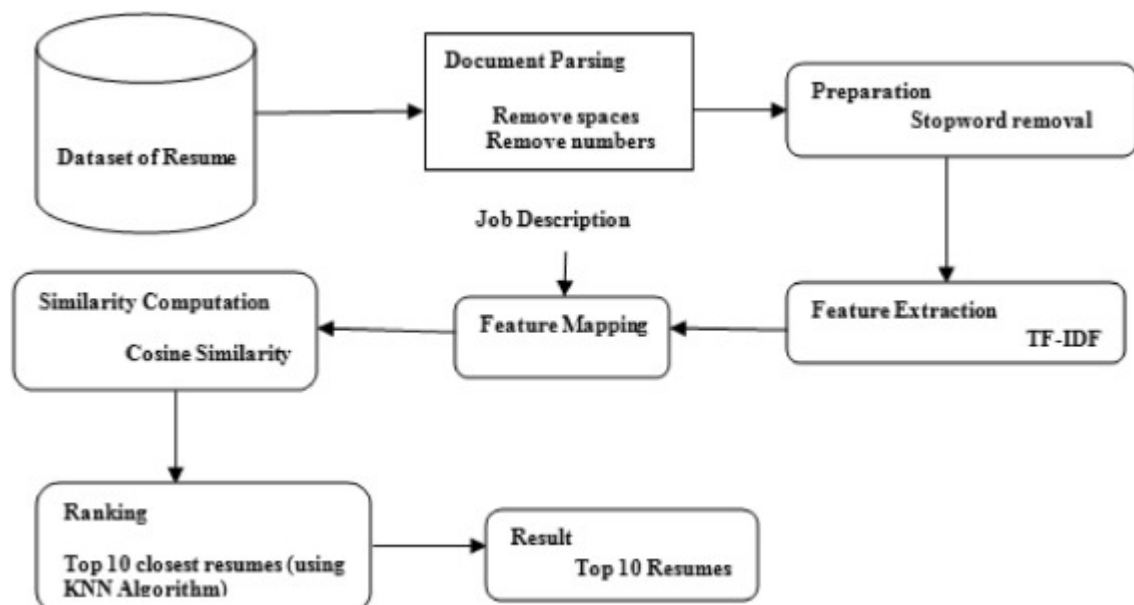
Data flow diagram.
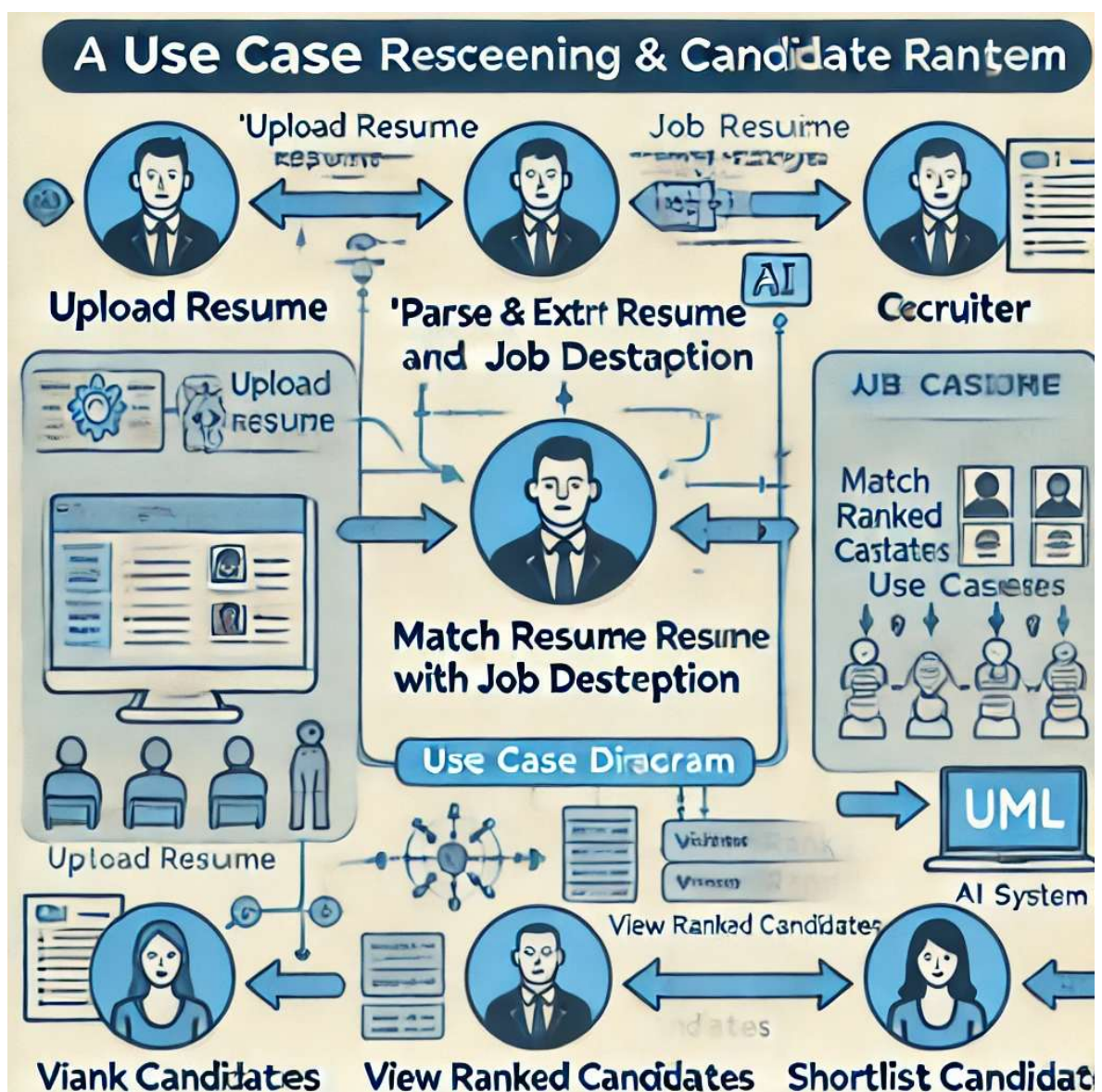
1.System architecture :

Represents the overall structure of the system, including data input, processing layers and output.
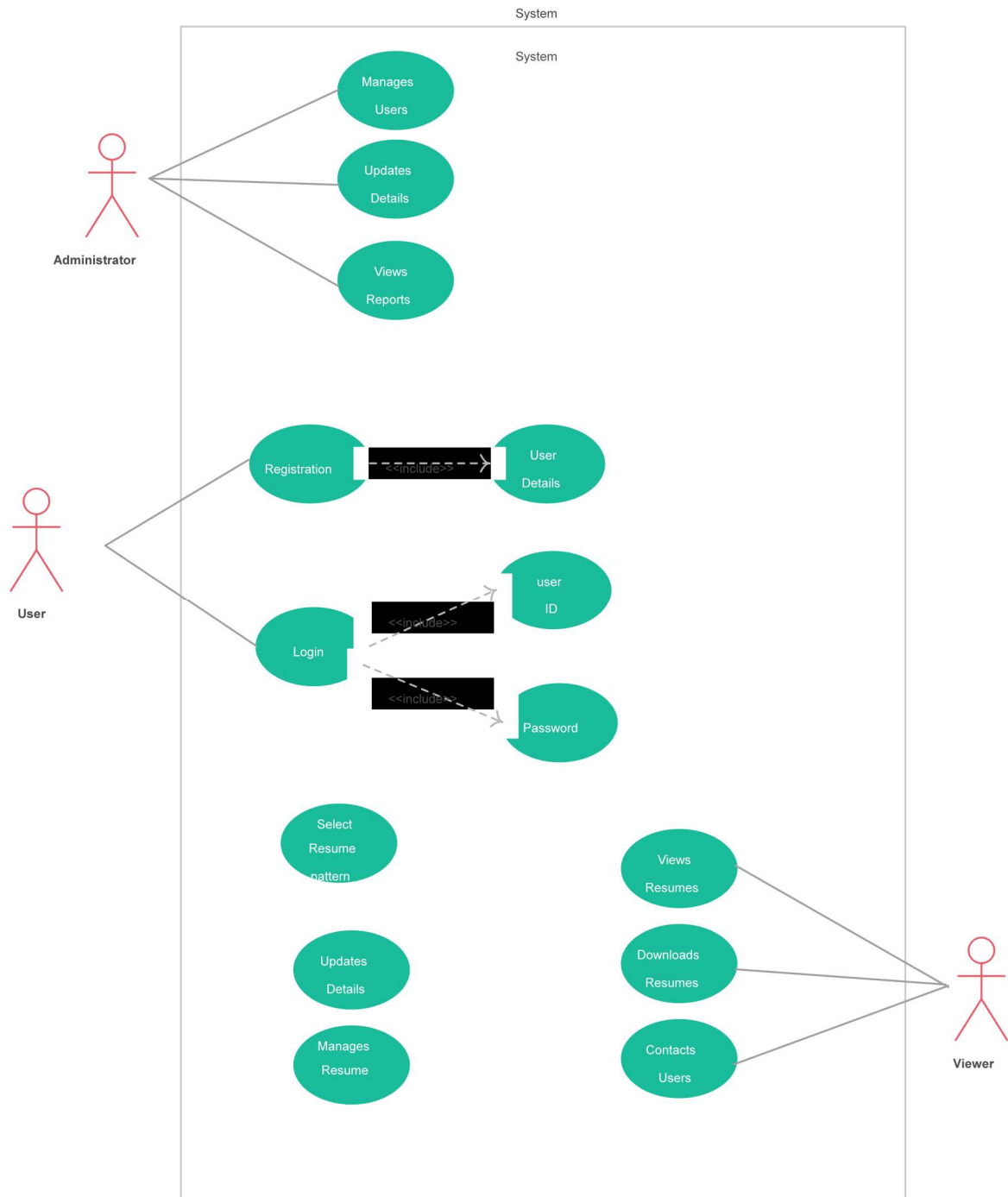
2.Data Flow diagrams:

Illustrates how resumes and job descriptions are processed through various stages.

3.Use Case diagram:

Shows interactions between users.

System

System

Manages Users

Updates Details

Views Reports

Administrator

Registration <<include>> User Details

User

Login <<include>> user ID

<<include>> Password

Select Resume pattern

Updates Details

Manages Resume

Views Resumes

Downloads Resumes

Contacts Users

Viewer

USE CASE DIAGRAM OF AI RSUME SCREENING AND CANDIDATE RANKING SYSTEM

Requirement Specification

The **requirement specification** outlines the functional and non-functional requirements necessary for the development and implementation of the **AI Resume Screening and Candidate Ranking System**.

3.1 Software Requirements:

The software requirements define the tools, frameworks, and technologies used for building, deploying, and running the AI-based resume screening system. These requirements ensure the system operates efficiently and provides accurate candidate ranking.
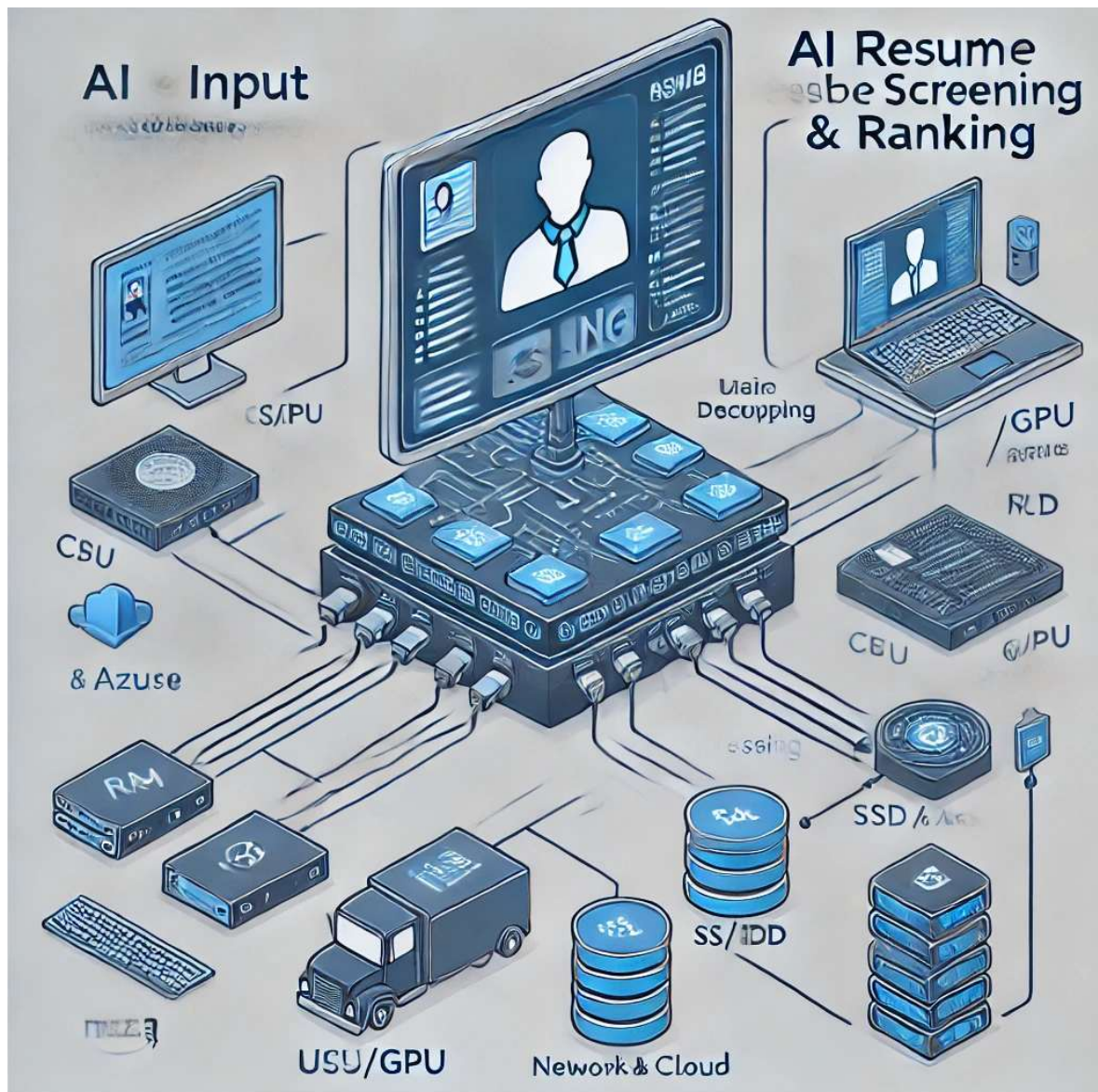
Figure 5:
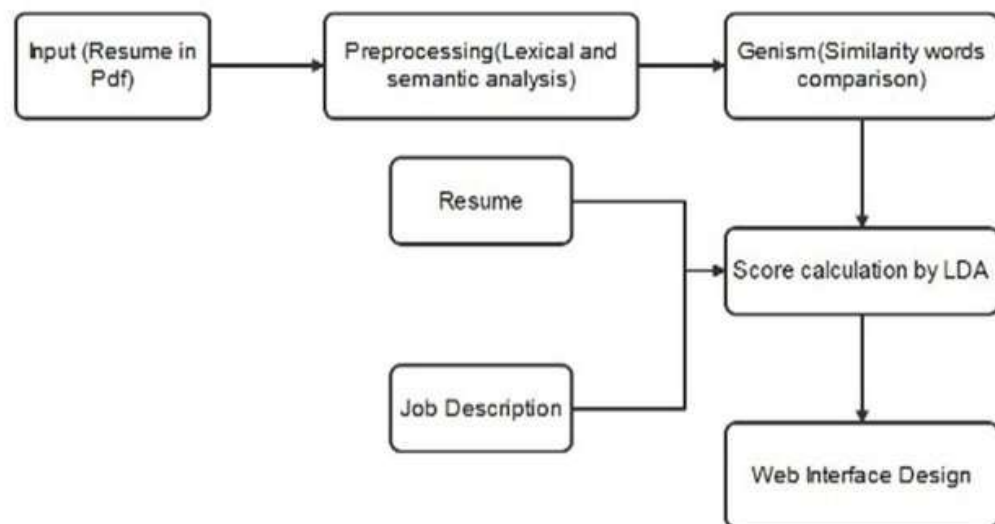

Software requirement diagram using stack

3.2 Hardware requirements:

The hardware requirements define the necessary computing resources for developing, deploying, and running the AI-based resume screening and ranking system efficiently. These requirements vary depending on whether the system is hosted **locally** or **on the cloud**.

Figure 6:

Hardware Requirement Diagram

Figure 7:
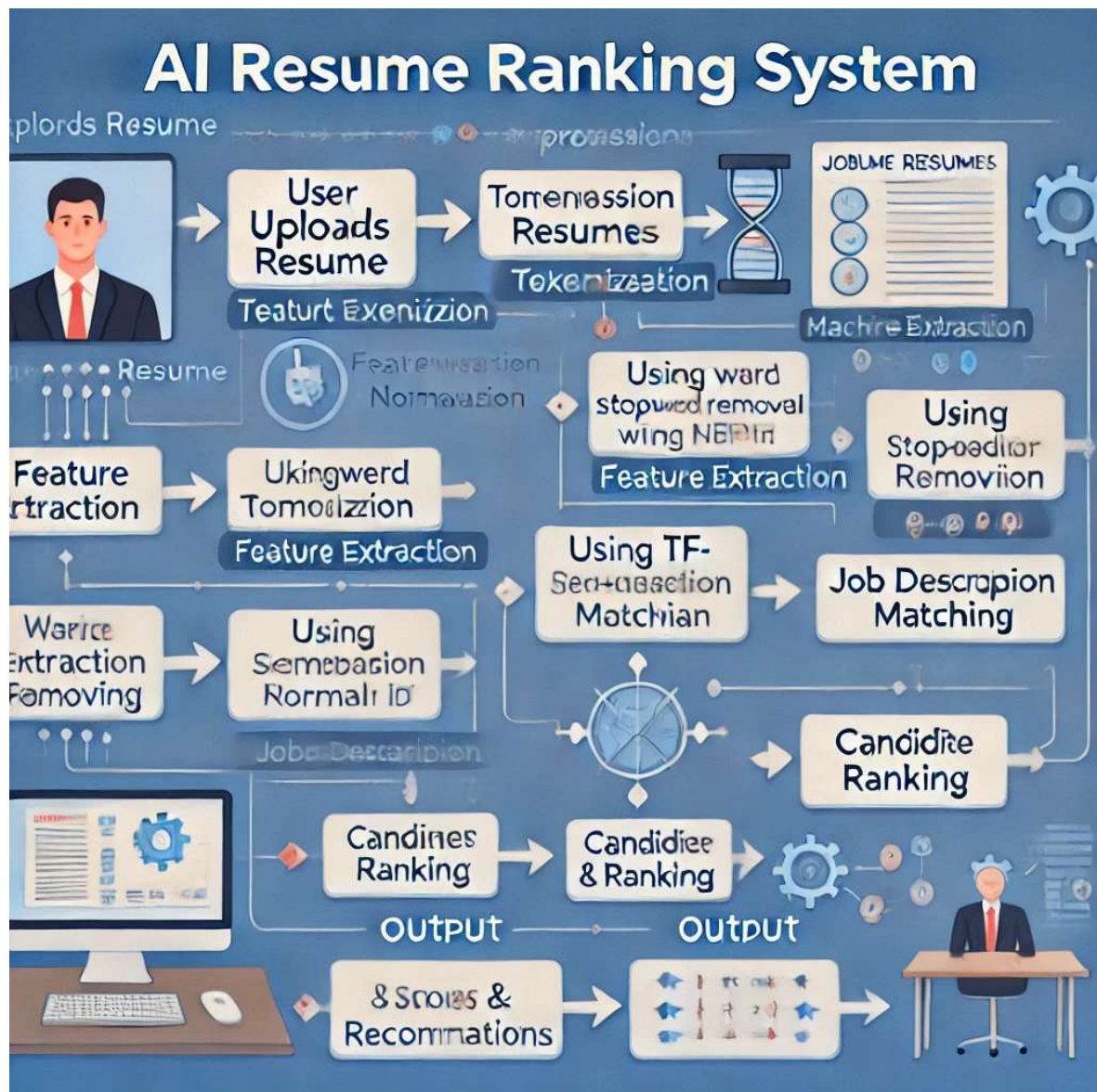
The Flow Chart of AI Resume Ranking System

CHAPTER 4

IMPLEMENTATION AND RESULT

A diagram showcasing how different components interact in the system, including:

- **Input Processing:** Uploading resumes and job descriptions.

- **Feature Extraction:** Using NLP techniques (TF-IDF, word embeddings).

- **Machine Learning Models:** Classifying and ranking resumes.

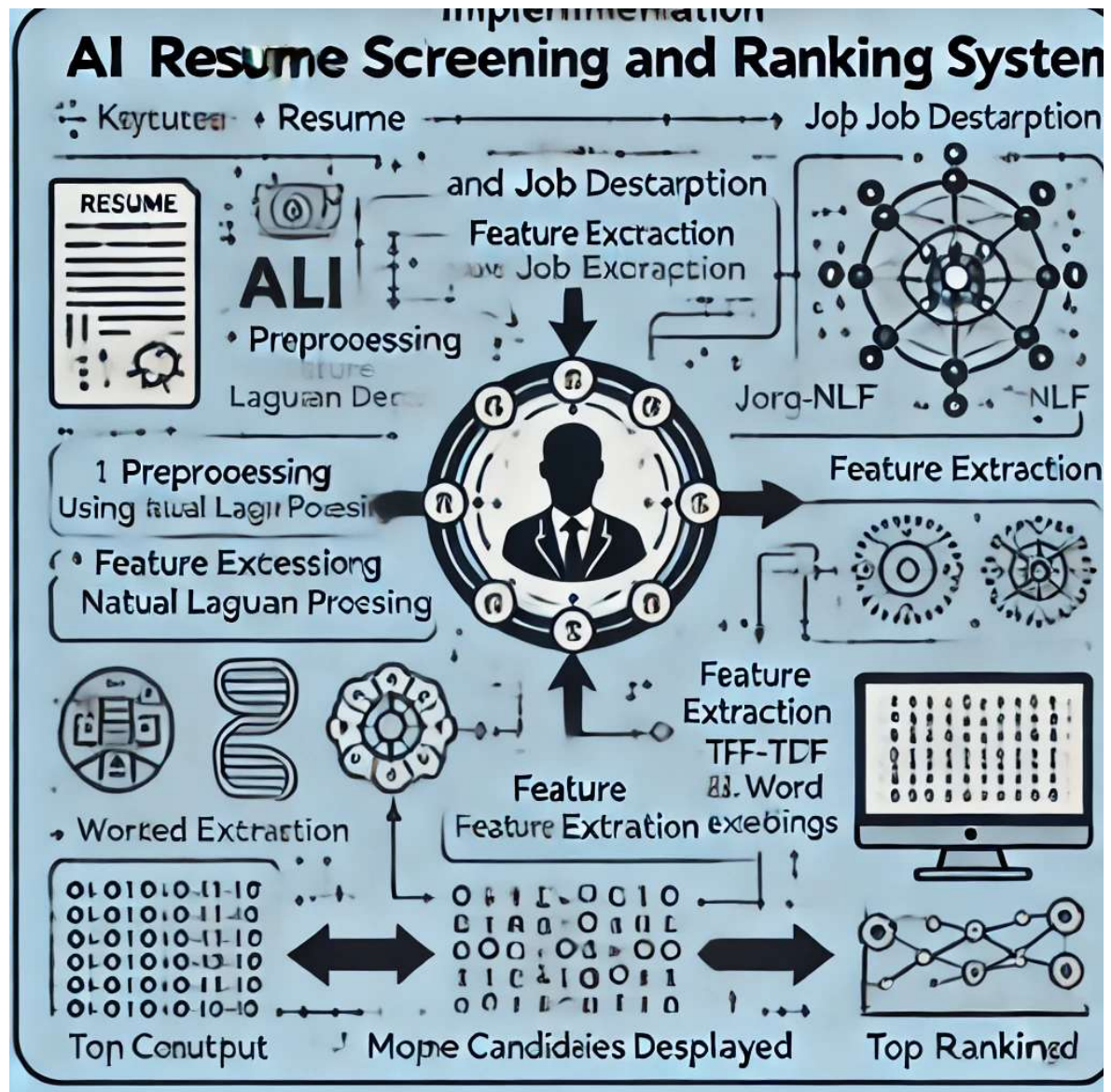- **Output Generation:** Displaying ranked candidates

Figure 9

4.1 Snap Shots of Result:

1.Snapshot: Resume Upload & Processing

[Insert Snapshot of Resume Upload Page]
Explanation:
This snapshot displays the **resume upload interface**, where recruiters can upload multiple resumes in **PDF or DOCX format**. Once uploaded, the system processes each resume using **Natural Language Processing (NLP)** techniques such as text extraction, tokenization, and feature engineering.

2. Snapshot: Resume Matching & Ranking

[Insert Snapshot of Candidate Ranking Table]
Explanation:
This screenshot shows the **ranked list of candidates** based on their suitability for a given job description. The ranking is calculated using **TF-IDF and ML models** like **SVM, Logistic Regression, or Random Forest**. The system provides a **matching score** to help recruiters identify the best candidates efficiently.

4.2   GitHub Link for Code:

https://github.com/Harsha872/AI-RESUME-RANKING-SYSTEM.git

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Future Work:

The AI Resume Screening and Candidate Ranking System can be improved in several ways for better efficiency and accuracy:

1. Integration with HR Systems.

2. Better AI Models

3. Multi-Language Support

4. Bias Reduction

5. Skill Gap Analysis

6. Data Visualization

5.2 Conclusion:

Overall Impact and Contribution

The **AI Resume Screening and Candidate Ranking System** significantly enhances the recruitment process by automating resume screening, reducing manual effort, and improving efficiency. It utilizes **Machine Learning (ML) and Natural Language Processing (NLP)** techniques to rank candidates based on job descriptions, ensuring fair and data-driven hiring decisions.

This project contributes to the recruitment industry by:
- **Reducing Hiring Time** – Speeds up resume evaluation and shortlisting.
- **Improving Accuracy** – Uses AI to match resumes more precisely with job requirements.
- **Enhancing Fairness** – Minimizes human bias in candidate selection.
- **Boosting Productivity** – Allows HR teams to focus on interviews and strategic hiring.
- **Providing Scalable Solutions** – Can be adapted for various industries and job roles.

Overall, this system transforms **traditional hiring methods** into a more **efficient, intelligent, and automated** process, benefiting both employers and job seekers.

REFERENCES

TSP 4.0 | AICTE Int Feb | AI Resume Screening and Ranking System
https://youtube.com/playlist?list=PLHmESvg2HbgfDGIHSw3Fdh79OkgsiveQ&si=_pK-4AcHeZz4Ob5X

1. Google Scholar- Research papers on AI, NLP, and Resume Ranking.
https://scholar.google.com
2. Scikit-Learn- Machine learning Algorithms & vectorization
https://scikit-learn.org

3. NLTK- NLP Library documentation
https://www.nltk.org

4. SpaCy- Industrial-strength NLP framework
https://spacy.io

5. Open Source Projects:
kaggle- https://www.kaggle.com
GitHub (AI Resume Screening Repositories) - https://github.com

These websites provide a combination of research,datasets, and Technology trends that support AI-Powered resume Screening and Candidate ranking Systems!