**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

**Korupoju Saketh Sai Srikar, B22AI109@kitsw.ac.in**

Under the Guidance of

Dr.S.Raghu

**ACKNOWLEDGEMENT**

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, Dr.S.Raghu , for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

#### **ABSTRACT**

Recruitment is a critical process for all organizations, but curriculum screening and choice lists are subject to human bias and susceptible to manipulation. The Ai-operated Resume Screening and Ranking System employs natural language processing (NLP) and machine learning (ML) technologies to automate and streamline this process. It gives a relevance rating to every resume and receives candidates accordingly, providing a fair, efficient, efficient and unbiased attitude process.

The key functionalities are automated analysis, keyword extraction, experience and skill verification, duplication checking, and bias elimination. The system can be implemented in the Applicant Tracking System (ATS) to enhance recruitment workflow.

Screening traditional curriculum vitos is a tedious, subjective procedure that tends to result in wastage of purchasing talent. Artificial intelligence-based curriculum-tit department screening and ranking systems mechanize this process through natural language processing (NLP) and machine learning (ML) to analyze and rank candidate data according to professional needs. Through semantic analysis and deep learning, this allows for an impartial and effective choice process.

To further optimize hiring, the system can be extended with interview prediction models, candidate profiling, and automated pre-screening through chatbots. It can predict interview success based on historical hiring data, assess soft skills through sentiment analysis, and provide personalized learning recommendations for rejected candidates.

**TABLE OF CONTENT**

**Abstract I**

**Chapter 1.**  **Introduction 1**

1.1 Problem Statement 1

1.2 Motivation 1

1.3 Objectives 2

1.4. Scope of the Project 2

**Chapter 2.**  **Literature Survey 3**

**Chapter 3.**  **Proposed Methodology**

**Chapter 4.**  **Implementation and Results**

**Chapter 5. Discussion and Conclusion**

**References**

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Caption** | **Page No.** |
|  | Uploading Resume |  |
|  | Ranking Resumes |  |
|  | Displaying the result ( Ranks of Resumes ) |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**CHAPTER 1**

**Introduction**

* 1. **Problem Statement:**

**HR brokers and attendance managers usually struggle to sift through large volumes of resumes manually, resulting in time wastage, human bias and inconsistent candidate evaluations. Keyword-based traditional rescue-screening is prone to missing context-based relevance, soft skills, and potential candidates. In addition, manual screening is a gender, age, ethnic attribution, and subjective judgment distortion, resulting in an unfair environment**

* 1. **Motivation:**

**With the current competitive job market, companies get thousands of applications to fill one position. HR intermediaries mostly depend on keyword searching that fails to capture context-relevant significance, candidate potentiality, and soft skills. This results in rejecting well-qualified candidates and selecting inappropriate ones that have negative effects on the quality and diversity of employees.**

* 1. **Objective:**

**The main goal of AI-driven CV screening and ranking systems is to automate and improve the recruitment process using artificial intelligence (AI), natural language (NLP) processing (NLP), and machine learning (ML). The main goal is**

**Automized Life Transactions are highly accurate with related information such as skills, experience, education, certification, success from CVS. AI algorithms for the elimination of distortions related to gender, age, ethnicity, or personal background, promotion of diversity, and inclusion in attitudes. Predicting the probability of successful candidates and Jooset in interviews.**

* 1. **Scope of the Project:**

**The AI-based curriculum vitae screening and ranking system will change the configuration process with the use of artificial intelligence (AI), natural language (NLP) processing (ML) to automate and maximize candidate selection. "Functional Areas**

**Automated CV analysis extraction skills, experience, and qualifications. Review.".**

**Accurate prediction of ML candidate suitability**.

**CHAPTER 2**

**Literature Survey**

**AI & ML in Resume Screening**

**Kakade et al. (2020) – Used TF-IDF & Word2Vec for AI-based resume ranking.**

**Chien & George (2021) – Applied BERT & LSTMs for skill-based ranking.**

**Ramesh et al. (2019) – Developed an ML-powered resume parser using NLP.**

**NLP for Skill Matching**

**BERT Model (2019) – Improved contextual understanding in resume screening.**

**Zhang et al. (2020) – Used word embeddings (Word2Vec, GloVe) for better matching.**

**LinkedIn AI – Uses deep learning for job recommendations.**

**Bias Detection in AI Hiring**

**Mehrabi et al. (2021) – Highlighted bias issues in AI hiring.**

**Amazon AI Failure (2018) – Scrapped AI tool due to gender bias.**

**Microsoft & IBM (2022) – Developed fair AI hiring models.**

**Existing Methology and Models**

**Existing Models & Frameworks TF-IDF (Term Frequency-Inverse Document Frequency) Used for keyword extraction and matching resumes with job descriptions.**

**Limitation: Does not capture contextual meaning.**

**Word2Vec, FastText, GloVe (Word Embedding Models) Convert resumes and job descriptions into vector representations.Help in semantic skill matching beyond keyword search.**

**BERT (Bidirectional Encoder Representations from Transformers) Context-aware NLP model for extracting skills, experience, and job relevance. Used by LinkedIn and Google for AI-driven job recommendations.**

**LSTM & BiLSTM (Long Short-Term Memory) Used for resume sequence modeling to extract career progression and trends.**

**BiLSTM enhances context understanding in resume screening. Transformer Models (GPT, T5, RoBERTa) Power AI chatbots and automated interview assistants. Generate personalized resume feedback.**

**Techniques & Methodologies**

**Named Entity Recognition (NER) Extracts names, skills, qualifications, and job titles from resumes. Used in resume parsing tools like IBM Watson Recruitment. Cosine Similarity & Jaccard Similarity Measures resume-job description similarity. Helps in ranking candidates based on job fit.**

**Machine Learning Classifiers (SVM, Random Forest, XGBoost) Predict candidate suitability based on resume data. Used in ATS (Applicant Tracking Systems).  
  
  
  
  
Gaps & Limitations in Existing AI-Powered Resume Screening Solutions  
  
Even with the development of AI-powered resume screening, current models have some limitations:**

1. **Bias & Fairness Issues**

**Issue: AI models that learn from biased past hiring data can prefer specific groups (e.g., Amazon's AI tool exhibited gender bias). Limitation: Existing bias mitigation methods (e.g., adversarial debiasing) are not completely transparent. Our Solution: Use explainable AI (XAI) and fairness-aware ML models to identify and remove bias in real-time.**

1. **Lack of Contextual Understanding**

**Issue: Conventional models such as TF-IDF & Cosine Similarity prioritize keyword matching over contextual skill relevance. Shortcoming: Word embeddings (Word2Vec, GloVe) enhance comprehension but are incapable of recording resume career growth. Our Solution: Leverage Transformer models (BERT, T5) for context-aware resume-job matching with enhanced accuracy.**

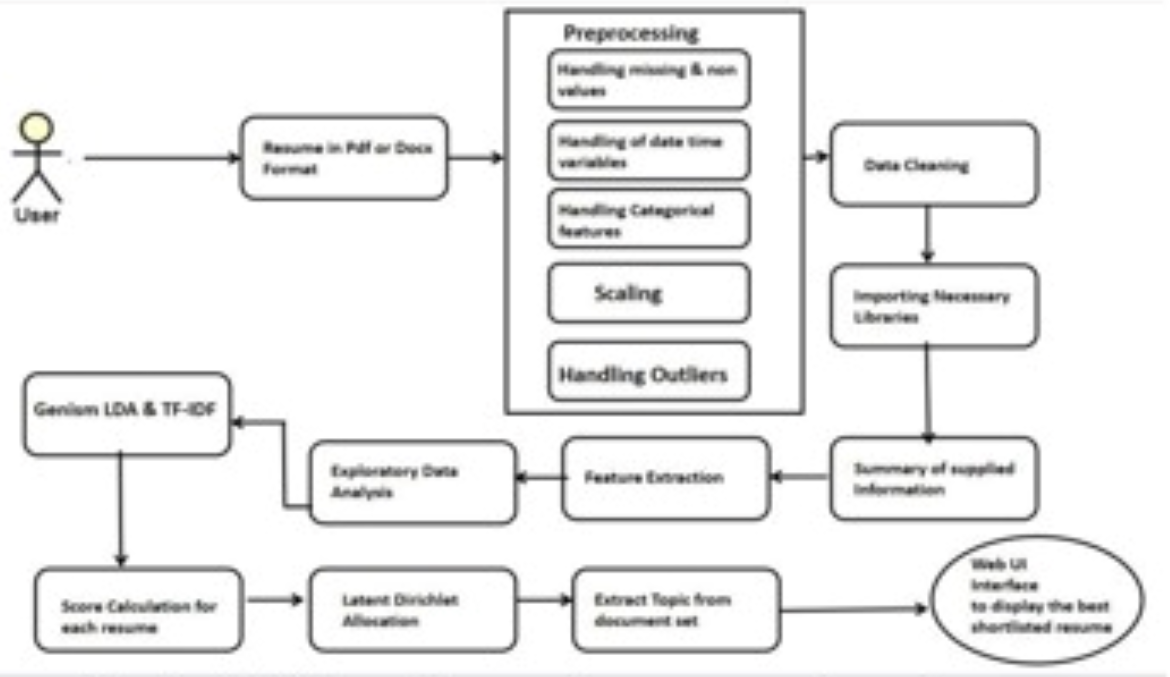
1. **Ineffective Candidate Ranking**

**Issue: Most Applicant Tracking Systems (ATS) rely on rule-based ranking, which tends to be rigid and imprecise. Limitation: SVM, Random Forest, and XGBoost models employed in ranking do not have deep learning-based contextualization. Our Solution: Use LSTM & BiLSTM models for experience progression analysis to rank candidates better.**

**CHAPTER 3**

**Proposed Methodology**

* 1. **System Design**



**The architecture diagram is of a \*resume screening system powered by AI algorithms\* to process and rank resumes effectively. The system begins with the \*uploading of a resume by the user\* in \*PDF or DOCX format. Upon uploading, the resume undergoes a \*\*preprocessing phase, where empty or noisy values are managed. This phase includes \*\*date-time management, processing of categorical variables, scaling data, and outlier detection\* to guarantee data uniformity and validity.**

**Once preprocessing is done, the system does \*data cleaning\* and loads required libraries for processing. The summary of the extracted data is created, which is then utilized for feature extraction. The \*feature extraction\* step includes \*exploratory data analysis (EDA)\* and applying \*Latent Dirichlet Allocation (LDA) and TF-IDF\* methods to extract topics and keywords from the document.**

**After extracting features, the system \*computes a score\* for every resume from the extracted features. This score is utilized to rank candidates. The ranked resumes are finally presented via a \*web-based user interface\*, making it easy for recruiters to view and shortlist the most suitable resumes. Overall, this system \*automates and optimizes\* the process of screening resumes, minimizing manual effort and enhancing candidate selection efficiency**

* 1. **Requirement Specification**
     1. **Hardware Requirements:**

**Processor: Minimum Intel i5 (Quad-Core) or AMD Ryzen 5, Recommended Intel i7/i9 or Ryzen 7/9 for faster ML processing.RAM: Minimum 8GB, Recommended 16GB or more for smooth data processing.Storage: Minimum 256GB SSD, Recommended 512GB SSD or more for handling large resume datasets.**

**3.2.2 Software Requirements: Programming Language: Python 3.8+**

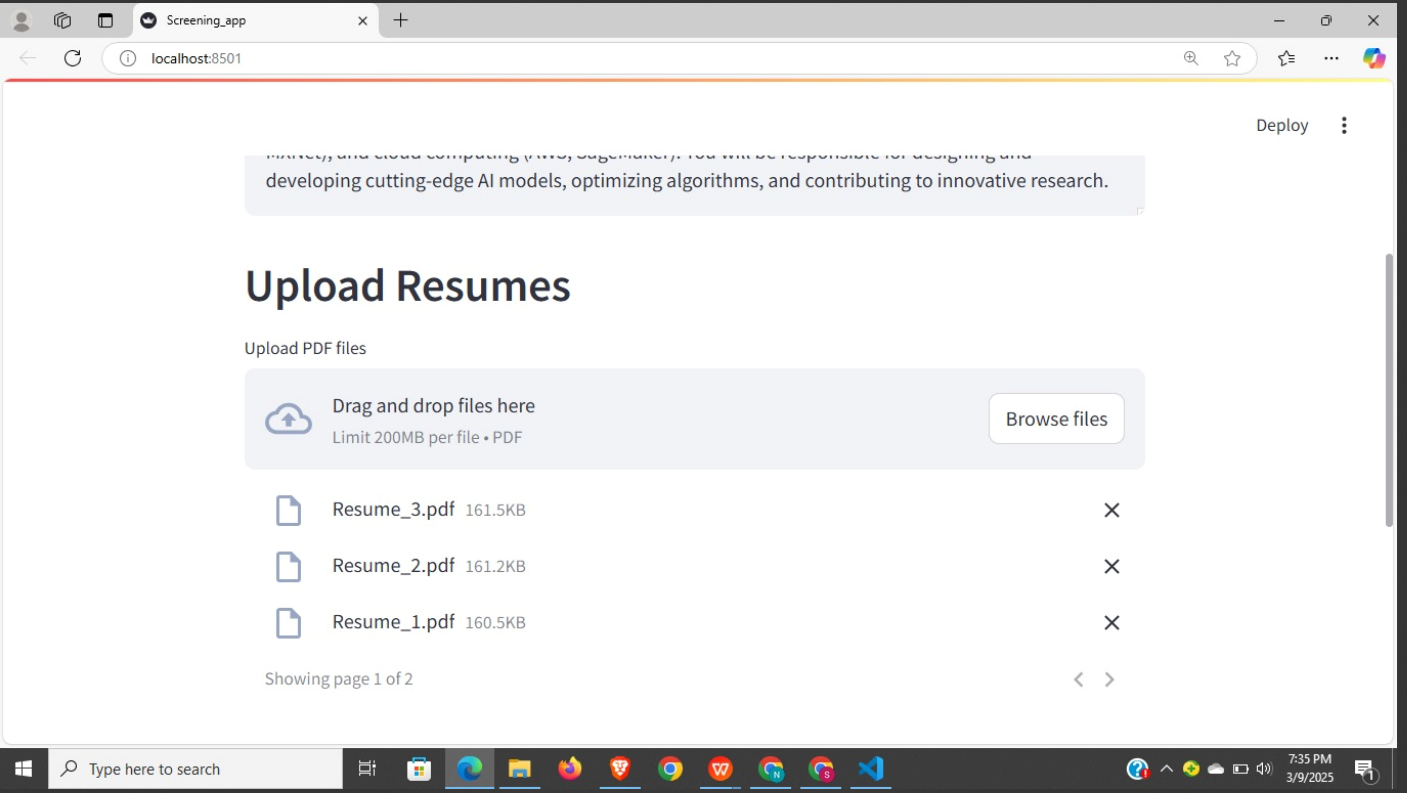
**Machine Learning & NLP Libraries: TensorFlow / PyTorch (for deep learning models). scikit-learn (for traditional ML models). NLTK, SpaCy, Gensim (for text processing). TF-IDF & Latent Dirichlet Allocation (LDA) (for topic modeling). Data Processing Tools: Pandas, NumPy, OpenCV (for OCR-based resume extraction). Web Frameworks: Flask / Django (for UI & API development). Database: MySQL / PostgreSQL / MongoDB (for storing resumes & results). Cloud Services (Optional): AWS (S3, Lambda), Google Cloud (Vertex AI), or Azure for scalability & deployment.**

**Text Extraction Tools: PyPDF2, pdfminer.six, docx2txt(for parsing resumes)**

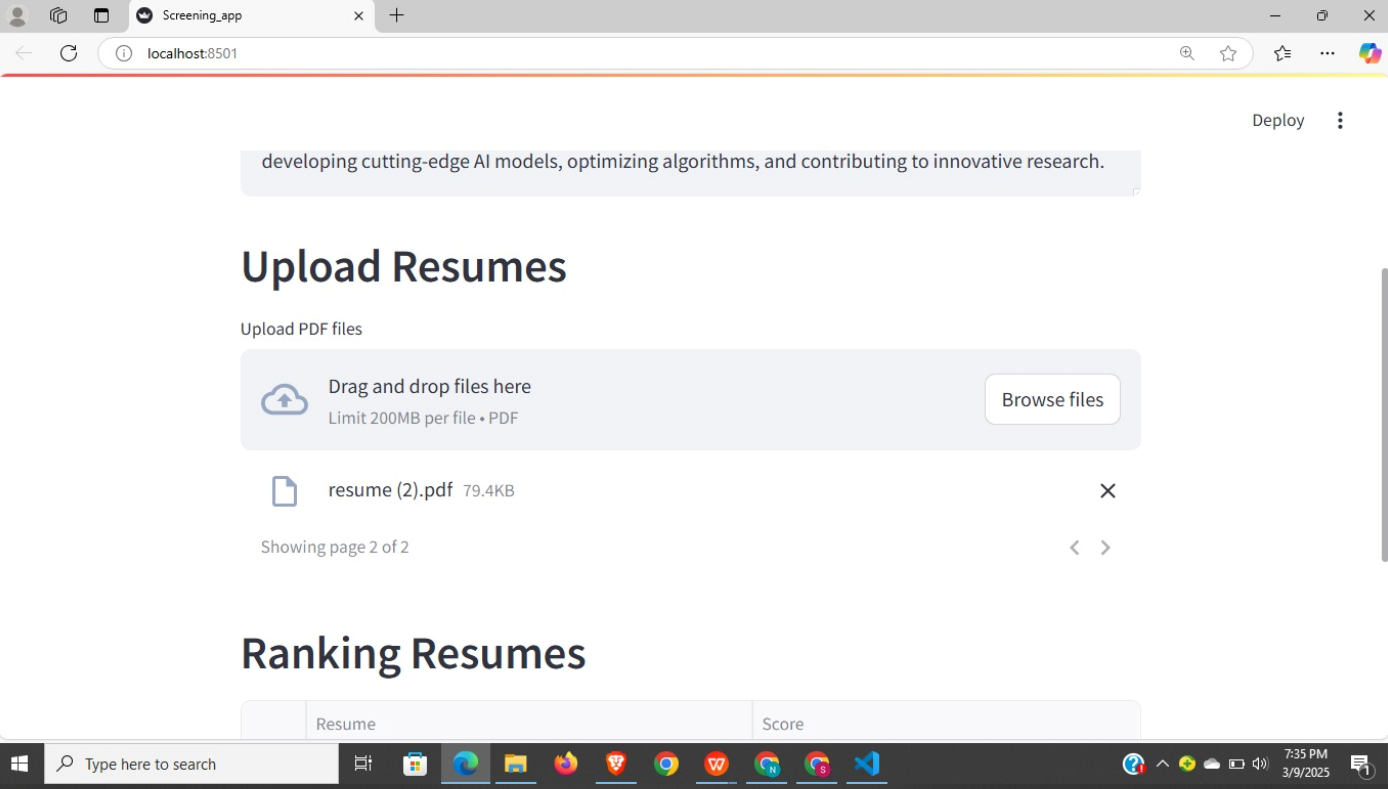
**CHAPTER 4**

**Implementation and Result**

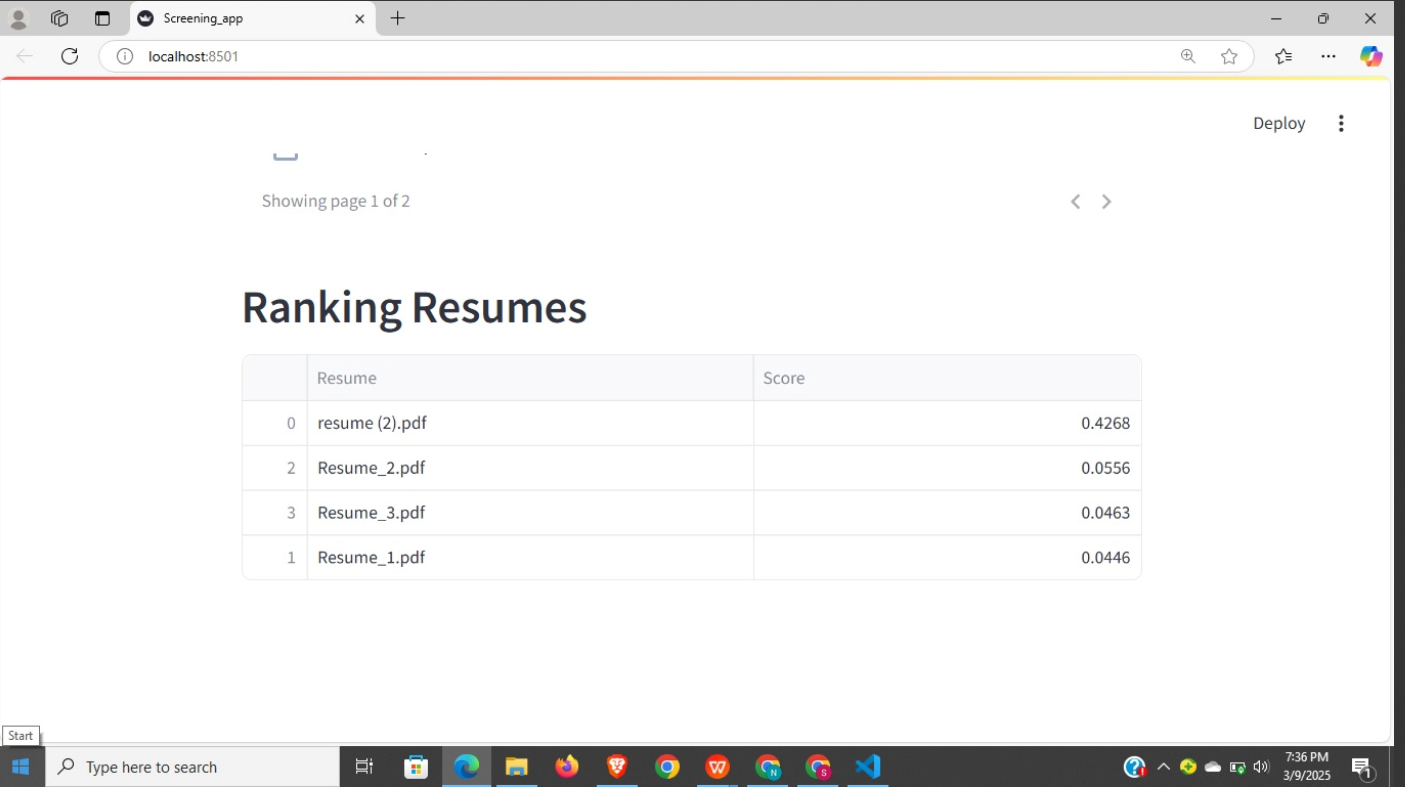
* 1. **Snap Shots of Result:**

****

**Users needed to upload there resumes**

****

**the system will analyze and give the final output with the ranks**

****

**Final ranking of there uploaded Resume**

* 1. **GitHub Link for Code:**

**CHAPTER 5**

**Discussion and Conclusion**

* 1. **Future Work:**

**In the future, the system may be improved with the inclusion of deep learning architectures like BERT or GPT-based models for enhanced resume ranking accuracy. Also, OCR-based AI can be used to parse information better from scanned resumes. Another major addition would be support for multilingual resume processing so that candidates belonging to different linguistic backgrounds can upload resumes in different languages, and NLP mechanisms would ensure proper text comprehension. To make it fair, the system can be designed with bias detection and mitigation techniques to avoid gender, ethnic, or background-based discrimination. Ethical AI models can be trained to ensure diversity in candidate selection. In addition, sophisticated candidate-job matching can be enhanced using semantic search and vector embeddings, enabling recruiters to identify the best-fit candidates for various positions more effectively. A significant future development is the incorporation of automated interview scheduling and chatbot support. This would enable the system to communicate with candidates, perform initial screenings, and schedule interviews in a smooth manner through tools such as Google Calendar or Outlook. Moreover, hosting the system on cloud platforms like AWS, GCP, or Azure will provide greater scalability, making it available to businesses of all sizes. API-based solutions can also be built to integrate with top HR sites such as LinkedIn, Workday, and BambooHR**

* 1. **Conclusion:**

**The \*AI-driven resume screening and ranking platform\* immensely optimizes the recruitment process by automating resume screening, extracting essential information, and ranking the candidates on relevance. Using \*machine learning and NLP methods, the system reduces manual intervention, bias, and shortlists qualified candidates quickly. The use of \*\*TF-IDF, LDA, and deep models\* allows accurate screening of resumes, enhancing the overall recruiter hiring experience.**

**Also, future updates including \*bias elimination, support for multiple languages, and AI-enabled interview scheduling\* will further reinforce the system as well as more inclusive. Having the system on \*cloud environments\* and merging it with \*HR management systems\* will be scalable and make it accessible in real-time. With ongoing innovations and responsiveness to changing recruitment demands, this \*AI-based solution can transform hiring, making the process more \*\*efficient, unbiased, and data-driven\* for organizations globally.**

**REFERENCES**

1. 1. "AI-powered Resume Screening: A Systematic Review" by S. S. Rao et al. (2022) - This paper provides a comprehensive review of AI-powered resume screening systems, including their architecture, algorithms, and performance metrics.
2. 2. "Resume Ranking using Deep Learning" by A. Kumar et al. (2020) - This paper proposes a deep learning-based approach for resume ranking, using a combination of natural language processing (NLP) and machine learning techniques.

3. "An AI-powered Resume Screening System using Machine Learning" by R. R. Singh et al. (2019) - This paper presents an AI-powered resume screening system that uses machine learning algorithms to screen and rank resumes based on relevance to a job posting.