World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(Research Article)

Context aware intelligence resume analyser using deep learning algorithms

Raghupathi Kanala, Meghana Vijaya Raghavan, Gundala Jathin Kumar *, Sushruth Bommagoni and Arun Kumar Onteddu

Department of CSE (Data Science), ACE Engineering College, Hyderabad, Telangana, India.

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(02), 738-746

Publication history: Received on 26 March 2025; revised on 02 May 2025; accepted on 04 May 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.15.2.0588

Abstract

This research introduces an AI-powered resume analyzer that makes it easier for people to evaluate and polish their resumes. Using cutting-edge Natural Language Processing (NLP), deep learning, and powerful Large Language Models (LLMs), the system quickly pulls out key details—like skills, education, work experience, and projects—from resumes in formats like PDF or DOCX. Unlike older resume tools, this one is faster and smarter, offering unique features like a Skill Gap Analysis to see how well your qualifications match a job's needs. It also includes a neural network that ranks your resume's strength and a helpful AI assistant that gives tailored tips to make your resume stand out. Together, these tools simplify the process and help job seekers showcase their best selves with confidence.

Keywords: Natural Language Processing; Deep Learning; Large Language Models; Skill Gap Analysis; Neural Network

1. Introduction

The hiring world has gone digital, but sifting through piles of resumes by hand remains a major roadblock. It's a slow, subjective process that often misses the mark in connecting the right candidates with the right jobs. Thankfully, the rise of Artificial Intelligence (AI) and Natural Language Processing (NLP) is opening doors to smarter, more efficient ways to evaluate candidates.

We're excited to introduce our AI-powered resume analyzer, a system that taps into advanced NLP, deep learning, and cutting-edge Large Language Models (LLMs) to take the hassle out of resume screening. It dives into messy resume formats like PDFs or Word docs, pulling out key details like skills, education, work history, and projects. But it doesn't just stop at organizing data—it goes further, analyzing how a candidate's skills measure up to a job's needs, using a neural network to score how well they fit, and even offering job seekers tailored tips to make their resumes shine.

Unlike traditional tools that just hunt for keywords, our system digs deeper, understanding the context and meaning behind the words to give a fuller picture of each candidate. It uses clever techniques, like Retrieval-Augmented Generation (RAG), to spot even the unspoken skills a job demands. By offering not just evaluations but also practical advice for improvement, it empowers candidates to stand out and helps recruiters make fairer, more informed choices based on solid data.

sizes. By automating repetitive, error-prone tasks, it saves time, cuts costs, and reduces bias, making the hiring process smoother and more effective. In this paper, we'll walk you through how we built this system, share the results of our testing, and explore how it can evolve to keep transforming talent acquisition.

* Corresponding author: Gundala Jathin Kumar

2. Literature review

In recent years, Artificial Intelligence (AI) and Natural Language Processing (NLP) have started transforming the way we handle hiring and resume screening. Platforms like LinkedIn and Indeed offer basic tools that match resumes to jobs using keywords. But these tools often focus on surface-level word matching, missing the deeper meaning behind a candidate's skills and experience.

Researchers have been exploring ways to make this process smarter. For example, Gupta and colleagues in 2020 built a system using Support Vector Machines (SVMs) to classify resumes based on skills and experience. Around the same time, Zhang's team created an NLP-based tool to pull out key candidate details from resumes, aiming to improve how well candidates are matched to jobs. These efforts show promise, but many still struggle to handle diverse resume formats or adapt to different writing styles, and they often don't scale well.

On the commercial side, tools like HireVue and Pymetrics use AI to assess candidates, but they lean more toward analyzing video interviews or personality traits rather than digging into resumes. Meanwhile, resume-building platforms like Resume.io and Zety help job seekers create polished applications but don't offer smart features like spotting skill gaps or predicting how well a candidate fits a role.

Our system stands out by bringing together cutting-edge tools like Retrieval-Augmented Generation (RAG), deep learning (using LSTM models for ranking), and powerful pre-trained Large Language Models (LLMs). It doesn't just read and evaluate resumes—it also pinpoints where a candidate's skills fall short compared to a job's needs and offers practical tips to make their resume stronger. By filling these gaps, our approach aims to make automated hiring smarter, more inclusive, and more helpful for everyone involved.

3. Existing System

Big job platforms like LinkedIn EasyApply, Indeed, and Glassdoor use keyword-matching tools to sort through resumes. They scan for specific words in a candidate's profile and compare them to job postings, providing a basic way to filter applicants. But these tools often miss the mark because they don't truly understand the meaning or context behind the words. This can lead to great candidates being passed over just because they didn't use the "right" buzzwords, while others slip through by peppering their resumes with the expected terms. Plus, these systems rarely give candidates feedback or tips to make their resumes better.

Researchers have tried to improve things with machine learning. Older systems, like those using Support Vector Machines (SVMs) or strict rule-based approaches, worked on sorting resumes into job categories. While they got better at this, they struggled to handle the huge variety of resume styles, languages, and industries out there. Some commercial tools, like HireVue and Pymetrics, have brought AI into hiring by focusing on things like behavioral assessments or fun, game-like tests to gauge personality and thinking skills. But these don't dive deep into analyzing the actual content of a resume, leaving a gap in smart, automated resume screening.

Then there are tools like Zety, Resume.io, and Canva's resume builder, which help job seekers create slick, professional-looking resumes with templates. They're great for making things look good but don't check if the resume actually fits the job or highlight where it could improve. Our system, though, takes a different approach. It blends deep analysis of resume content, smart skill-matching powered by Retrieval-Augmented Generation (RAG) models, neural network-driven ranking, and personalized suggestions for improvement into one seamless package. By understanding context and offering actionable advice, it tackles the shortcomings of other tools, creating a smarter, more supportive way to evaluate resumes.

4. Proposed System

Our proposed system is a smart, flexible AI tool designed to take the headache out of resume screening by handling everything from parsing to personalized feedback in one go. It starts by tackling resumes in messy formats like PDFs or Word docs. Using a clever document parser and some pattern-matching tricks, it pulls out the important stuff—things like your name, skills, education, work history, and projects—organizing it neatly so the rest of the system can work smoothly, no matter how freeform or jumbled the original resume was.

Next up, the Skill Gap Analysis steps in to see how well a candidate's skills match a job's needs. It uses a powerful text-to-text model (called BART-large) to dig into the job description and pick out both obvious and hidden skills the employer is looking for. Then, it compares those to the skills listed in the resume, spotting what matches and what's missing. This gives a clear picture of how close a candidate is to the job's expectations and points out areas where they might want to level up.

The Resume Ranking part is where things get really cool. It uses a lightweight neural network with Long Short-Term Memory (LSTM) layers to give each resume a score based on two big factors: how many skills match the job (which counts for 70% of the score, because skills are king in most industries) and how well the candidate's years of experience line up with what the job asks for (worth 30%). These scores make it easy for recruiters to sort through a stack of resumes and decide who to call in for an interview.

Finally, the Resume Enhancement Assistant acts like a personal coach. It looks at the resume and the skill gaps, then offers tailored advice—like suggesting missing technical skills to add, rephrasing vague project descriptions, emphasizing key achievements, or tweaking the layout to make everything pop. These tips help make the resume clearer, stronger, and more likely to catch a recruiter's eye. The system's modular setup means each piece works on its own but fits together perfectly, creating a seamless, intelligent tool that makes resume evaluation easier, fairer, and more helpful for everyone.

5. Methodology

The proposed AI-powered resume analyzer system is composed of four core modules: Resume Parsing, Skill Gap Analysis, Resume Ranking, and Resume Enhancement Assistant. These modules work together to automate resume evaluation, gap identification, candidate ranking, and improvement suggestions. The system is deployed through a user-friendly Streamlit-based web interface, ensuring accessibility and interactivity.

5.1. Resume Parsing

The first step involves parsing unstructured resume files provided in formats like PDF and DOCX. Using document processing libraries such as pdfplumber and python-docx, the system extracts textual content while preserving its structure. Regular expressions are applied to identify and extract key fields such as Name, Email, Phone Number, Skills, Education, Work Experience, and Projects. This structured data is critical for downstream analysis as it ensures consistency and relevance. Special preprocessing techniques are employed to handle noisy formatting, missing fields, and multi-line data entries, thus making the parser robust to variations across different resumes.

5.2. Embedding Generation and Vector Storage

After parsing, the extracted content is passed through an Embedding Model to generate dense vector representations. These embeddings capture semantic information from the text, going beyond simple keyword extraction. The vectors are stored in a Vector Database, allowing the system to perform efficient similarity search operations later. This embedding step is crucial for enabling the Retrieval-Augmented Generation (RAG) Pipeline to fetch relevant context dynamically when answering user queries or extracting additional structured insights.

5.3. RAG-Based Information Retrieval

The RAG pipeline forms the core of intelligent data retrieval in the system. It uses the stored embeddings to retrieve relevant portions of the resume or associated data points when a user query is issued or when additional context is needed for analysis. The RAG model combines retrieved knowledge with a pre-trained language model (like BART-large) to generate contextualized outputs. For Skill Gap Analysis, the RAG pipeline helps extract required job skills from job descriptions and matches them semantically to the candidate's parsed skills. This hybrid approach ensures that even implicit or non-directly stated skills are identified accurately, providing a deeper understanding of candidate-job alignment.

5.4. AI-Based Evaluation Models

Once the resume content is parsed and relevant skills are extracted, specialized AI models perform deeper evaluation. Three major functionalities are implemented:

- Skill Matching: A set-based analysis identifies matched and missing skills by comparing candidate skills with extracted job requirement skills.
- Resume Ranking: A lightweight LSTM-based neural network model predicts a ranking score for each resume. Inputs to the model include encoded skill matches and years of experience. The model is trained using a custom
- dataset where resumes are labeled with scores based on job fit, allowing it to generalize across unseen resumes.

- Resume Enhancement: Using simple NLP heuristics and best-practice guidelines, the system suggests specific improvements to resumes, such as adding measurable achievements, reordering sections for impact, highlighting key skills, and correcting grammatical inconsistencies.

5.5. User Interaction and Results Generation

Finally, the processed outputs are presented to the user via a Streamlit-based web interface. Users can upload resumes, provide job descriptions, and interact with the system through a chatbot interface. The system displays extracted data, skill match results, resume ranking scores, and personalized enhancement suggestions. The modular design allows the user to use these features individually or as an integrated evaluation package. Generated results are dynamically updated based on user inputs, offering a highly interactive and responsive user experience.

5.6. System Architecture

The architecture of the proposed resume analyzer system is designed for modularity, scalability, and efficient resume evaluation. It integrates several AI-driven components to enable parsing, skill gap analysis, ranking, and enhancement suggestions within a unified framework.

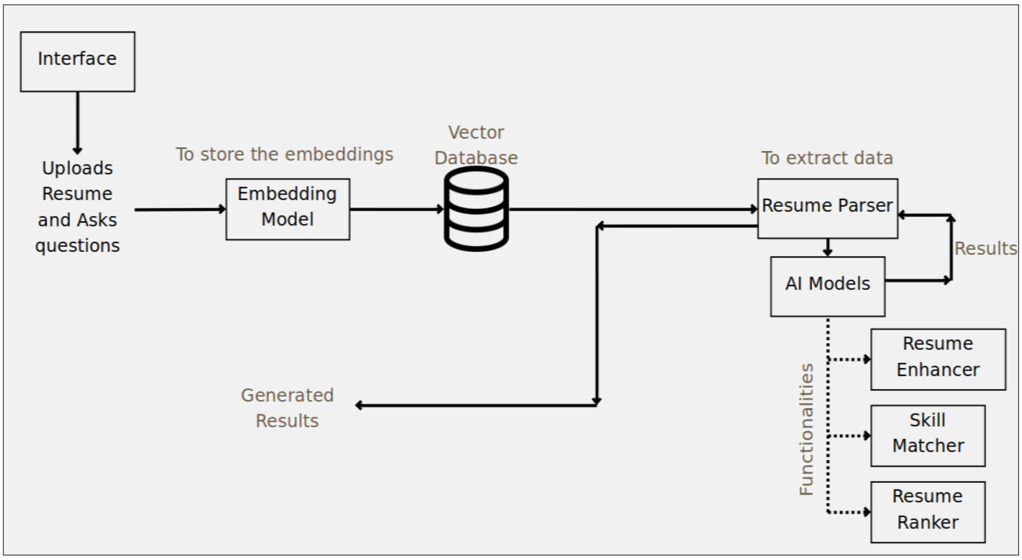

Figure 1 System Architecture

5.7. Sequence Diagram of Workflow

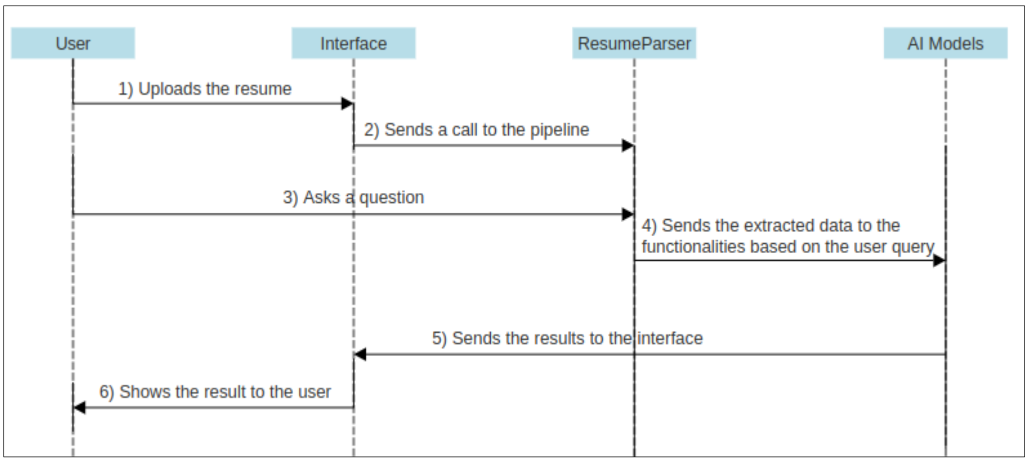

Figure 2 Sequence Diagram of Resume Analyzer Workflow

6. Results and Discussion

The proposed AI-powered resume analyzer system was evaluated through extensive testing on various sample resumes and job descriptions. The system's performance was assessed based on multiple modules: resume parsing accuracy, skill gap analysis precision, resume ranking effectiveness, and the quality of enhancement suggestions. Both qualitative and quantitative analyses were conducted to demonstrate the system's capabilities.

6.1. Resume Parsing Results

The Resume Parsing module was tested on a diverse set of resume formats, including PDFs and DOCX files with varying structures and templates. The parser successfully extracted structured data such as names, skills, education, work experience, and project details. Minor errors were observed mainly in resumes with highly unconventional layouts, which can be further minimized by improving parsing heuristics.

6.2. Skill Gap Analysis

The Skill Gap Analysis module utilized a pre-trained BART-large model to extract skill requirements from provided job descriptions. It then compared these with candidate skills to identify matches and gaps.

For instance, when evaluating a resume against a data scientist job description, the system accurately identified "Python" as a matched skill and missing skills such as "SQL" and "Deep Learning"

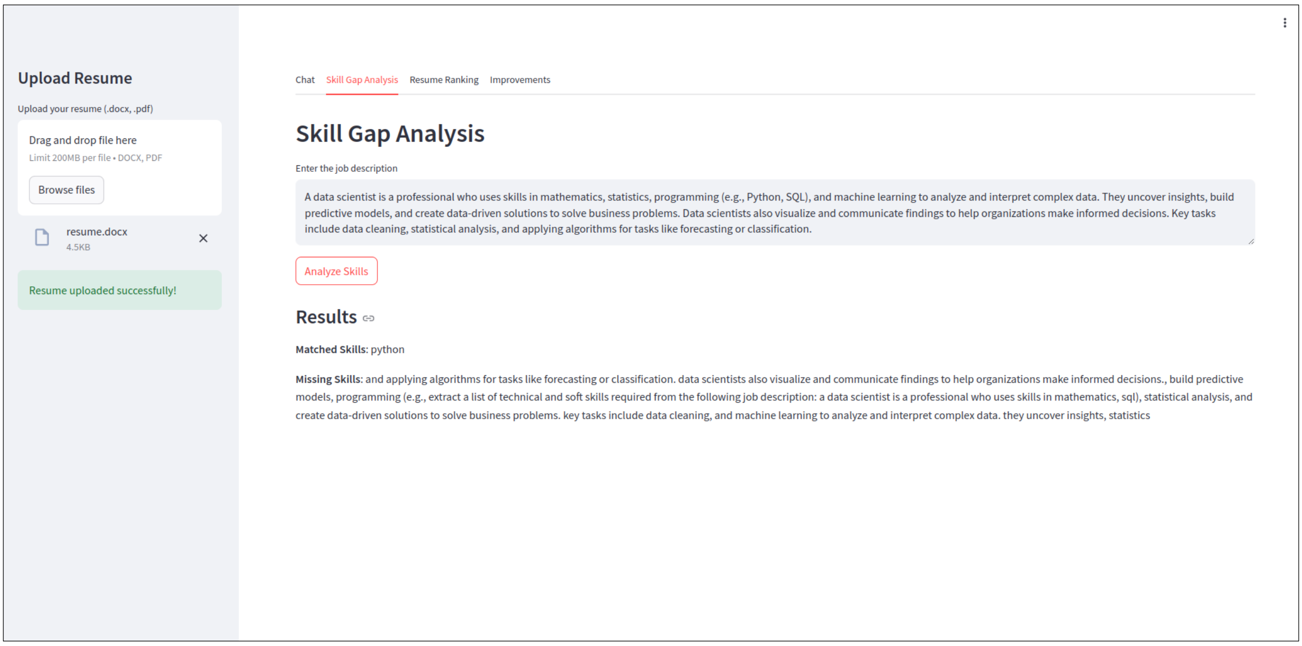

Figure 3 Skill GAP Analysis Result Interface

6.3. Resume Ranking Model

The Resume Ranking module, powered by a lightweight LSTM-based neural network, was trained using a synthesized dataset.

The model's training progress is illustrated by the loss curve shown in Figure 4, where the training loss steadily decreases across epochs, demonstrating successful convergence. The interface is described in **Figure 5**.

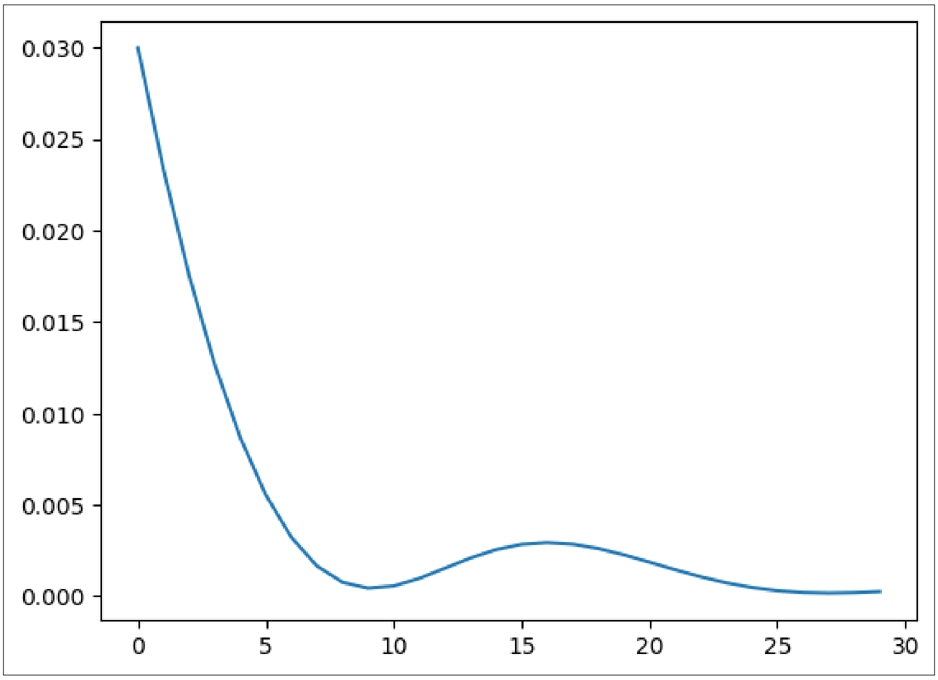

Figure 4 Training Loss Curve for Resume Ranking Model

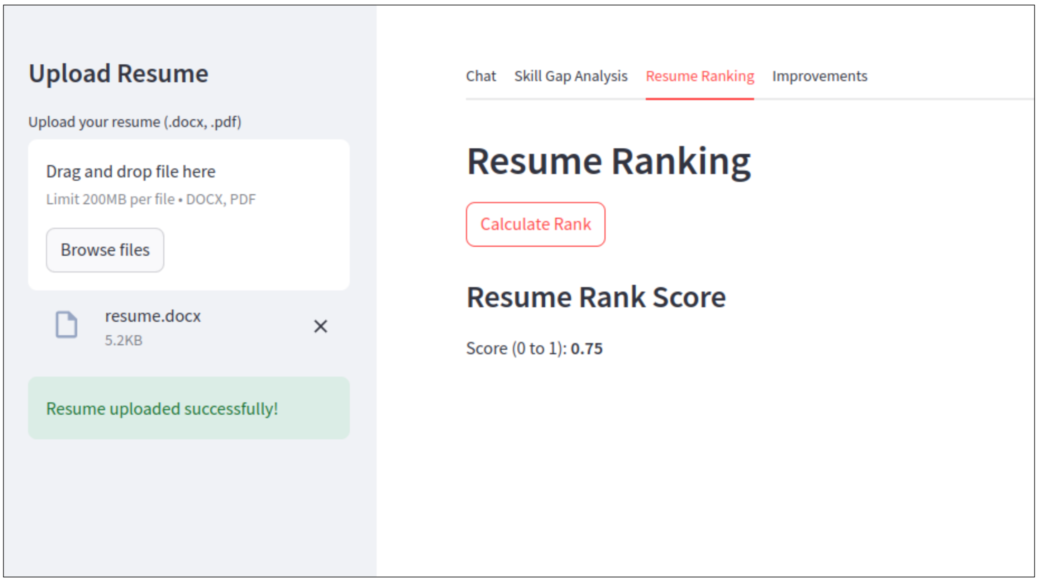

Figure 5 Ranking Resume Result Interface

6.4. Resume Enhancement Suggestions

The Resume Enhancement Assistant provided personalized recommendations for improving resumes.

For example, suggestions included adding specific measurable achievements, highlighting missing skills, and improving project descriptions. An example output is shown in Figure 6, where actionable improvements were suggested for a sample resume.

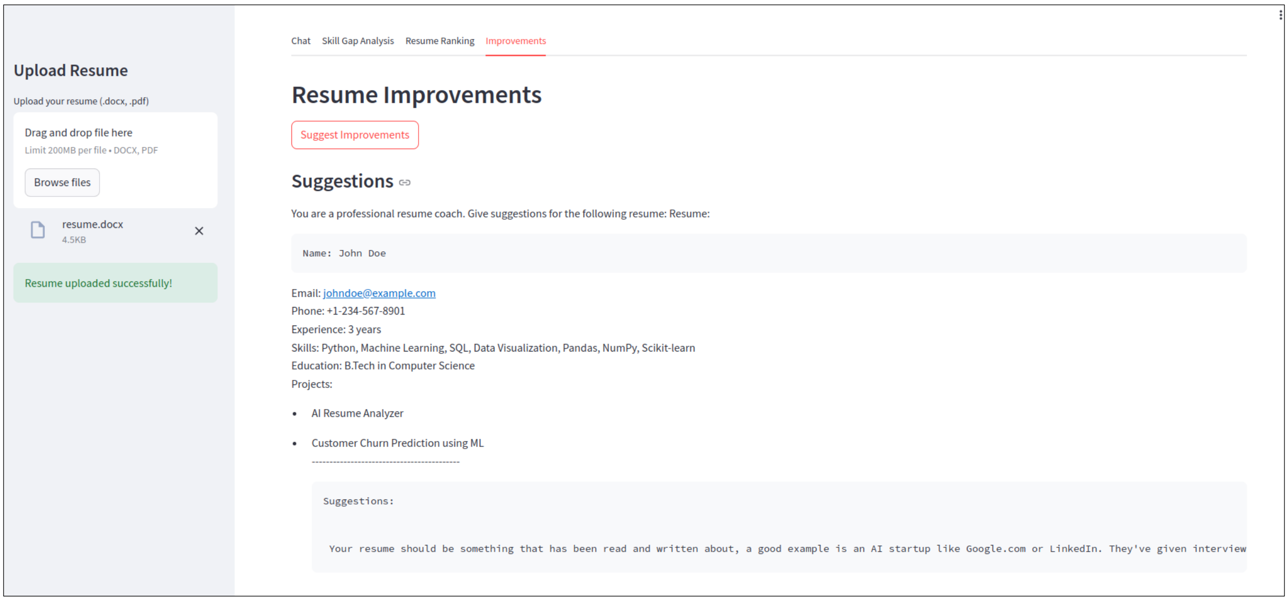

Figure 6 Resume Improvement Suggestions Interface

6.5. Chatbot Assistant Interaction

The system also incorporated a chatbot interface where users could interactively ask about their resume's strengths, missing skills, or overall ranking.

The chatbot demonstrated quick and contextually accurate responses, enhancing user experience and engagement, as seen in Figure 7.

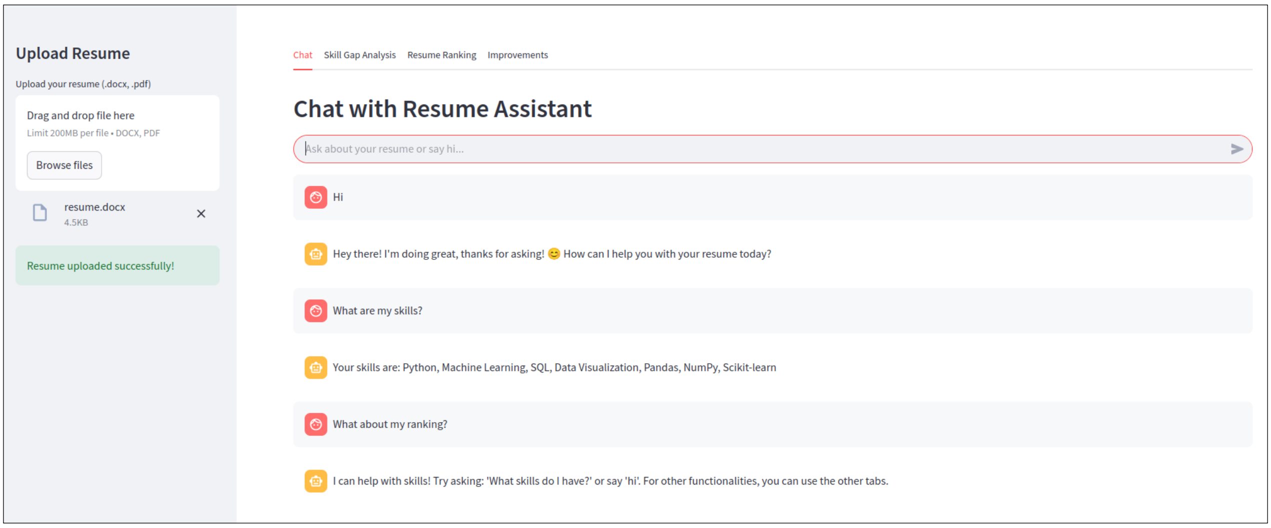

Figure 7 Chat Interaction with Resume Assistant

7. Conclusion

This research presents the design and development of an AI-powered resume analyzer system that automates resume parsing, skill gap analysis, candidate ranking, and resume enhancement suggestions using advanced Natural Language Processing (NLP), deep learning models, and pre-trained Large Language Models (LLMs).

The system demonstrated its capability to efficiently extract structured information from resumes, assess candidate-job fit, rank resumes based on objective criteria, and offer actionable feedback to improve resume quality.

Compared to conventional screening tools, the proposed system significantly reduces manual effort, enhances screening accuracy, and improves the overall candidate experience. By integrating multiple AI techniques within a modular and scalable framework, the system provides a comprehensive solution for modern recruitment challenges.

Experimental results indicate that the system is effective in parsing diverse resume formats, accurately performing skill gap analysis, and generating reliable ranking scores. The platform can be easily deployed in various recruitment environments to support faster and more informed hiring decisions.

Compliance with ethical standards

Disclosure of conflict of interest

There is no conflict of interest.

References

[1] Lewis, P., Oguz, B., Rinott, R., Riedel, S., & Augenstein, I. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv preprint arXiv:2005.11401. https://arxiv.org/abs/2005.11401.

[2] Google, "Gemini 1.5 Technical Report," Google DeepMind, Feb. 2024. [Online]. Available: https://deepmind.google/technologies/gemini/

[3] Lee, M. Y. (2023). Building Multimodal AI Chatbots. arXiv preprint arXiv:2305.03512. https://arxiv.org/abs/2305.03512

[4] Google Cloud. (n.d.). What is Retrieval-Augmented Generation (RAG)? https://cloud.google.com/use-cases/retrieval-augmented-generation

[5] ACS Publications. (2024). Using Retrieval Augmented Generative AI Chatbots to Support and Enhance Chemistry Education. https://pubs.acs.org/doi/10.1021/acs.jchemed.4c00765

[6] ResearchGate. (2024). Bring Retrieval Augmented Generation to Google Gemini via External API: An Evaluation with BIG-Bench Dataset. https://www.researchgate.net/publication/380486834

[7] ResearchGate. (2024). Enhancing PDF Information Retrieval Through a Gemini Pro LLM-Powered Chatbot. https://www.researchgate.net/publication/388992399

[8] Kumar, A., Gupta, R. (2022). AI-Based Multi-Modal Chatbot Interactions for Enhanced User Engagement. https://rjpn.org/ijcspub/papers/IJCSP24C1126.pdf

Author's short biography

Raghupathi Kanala

I am Raghupathi Kanala, an Assistant Professor with an academic background in M.Tech (CSE) and pursuing (PhD) in the field of Computer Science and Engineering. With over 20 years of professional experience, my research interests are primarily focused on Big Data. I have dedicated my career to exploring the vast potential of data-driven technologies and their applications in solving complex real-world problems.

Meghana Vijaya Raghavan

I am Meghana Vijaya Raghavan and I am an undergraduate student pursuing a B.Tech in Computer Science and Engineering (Data Science). Her research interests include Data Science, Machine Learning, Deep Learning, and Computer Vision. She is passionate about developing innovative solutions using data-driven approaches to solve real-world problems.

Gundala Jathin Kumar I am Gundala Jathin Kumar and I am currently pursuing a B.Tech in Computer Science and Engineering (Data Science). His research interests primarily lie in the field of Machine Learning, with a focus on developing intelligent systems for real-world applications. As an undergraduate researcher, he is actively exploring various machine learning techniques, aiming to contribute innovative solutions to complex problems	
Sushruth Bommagoni I am Sushruth Bommagoni, undergraduate student pursuing a B.Tech in Computer Science and Engineering (Data Science). My research interests include Big Data Analytics, focusing on efficient data processing and analysis. I am passionate about exploring data-driven solutions to real-world challenges. My academic journey has helped me develop skills in machine learning, data mining, and statistical analysis.	
Arun Kumar Onteddu I, Arun Kumar Onteddu, am an undergraduate student pursuing a B.Tech in Computer Science and Engineering (Data Science). My research interests focus on Artificial Intelligence, particularly in machine learning and deep learning. I am passionate about developing intelligent systems that enhance human-computer interactions. My goal is to contribute innovative solutions that bridge theoretical AI research with real-world applications