Sunny Ramchandani, 2025, 13:5
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

AI Resume Analyzer: A Review and Case Study of an NLP-Driven Recruitment System

Sunny Ramchandani, Preksha Jain, Ameya Shivhare, Maddu Akhil, Shashank Pandey
Faculty of Engineering and Technology Parul University, Vadodara, India

Abstract- The growing demand for efficiency in recruitment has accelerated the adoption of AI-powered resume screening systems. This paper reviews the evolution of resume analysis approaches, from keyword-based filters to large language models (LLMs) and Retrieval-Augmented Generation (RAG) pipelines. A critical analysis compares accuracy, scalability, fairness, and regulatory compliance across methods. Ethical concerns, includ- ing bias and privacy, are discussed alongside recent regulations such as the EU AI Act and GDPR. To complement the review, we present a case study of an AI Resume Analyzer system, followed by experimental validation. The paper concludes with limitations and future research directions for trustworthy, scalable, and fairness-aware recruitment systems.

Keywords: Resume Screening, NLP, Machine Learning, Artificial Intelligence, Recruitment, ATS, Fairness, Ethics.

I. INTRODUCTION

A. Background and Motivation

Recruitment is one of the most critical processes for organizations, yet it remains time-consuming and resource- intensive. Traditionally, recruiters manually review hundreds of resumes for a single position, often leading to delays, inconsistencies, and human bias. With the rapid growth of job applications and the digitalization of hiring, there is a growing need for automated tools that can screen resumes efficiently and fairly. Advances in natural language processing (NLP), machine learning (ML), and large language models (LLMs) have opened new opportunities to streamline resume analysis, moving beyond simple keyword searches to context- aware evaluation.

B. Problem Statement

Despite these technological advances, several challenges remain unresolved. Many existing Applicant Tracking Systems (ATS) rely heavily on keyword matching, which fails to capture the true context of skills and experiences. As a result, qualified candidates may be overlooked simply because their resumes are formatted differently or use alternative terminol- ogy. Moreover, AI-driven systems themselves can introduce biases, lack transparency in decision-making, and raise privacy concerns when handling sensitive candidate data. Thus, the key problem is to design a resume analysis system that is accurate, fair, explainable, and compliant with regulations while still being efficient for recruiters.

C. Objectives

The objectives of this study and system are:

1. To review existing approaches to automated resume screening, from keyword-based methods to transformer and LLM-driven pipelines.
2. To design and implement a prototype AI Resume An- alyzer that integrates resume parsing, semantic search, ATS scoring, and LLM-powered feedback.
3. To experimentally validate the system on sample re- sumes and compare its performance with baseline key- word methods.
4. To identify technical gaps, ethical issues, and regulatory challenges in deploying AI-based recruitment systems.
5. To propose directions for future development, including explainable AI, fairness audits, and cloud-scale deploy- ment.

D. Scope of the Project

The scope of this project is twofold. On one hand, it provides a literature review that traces the evolution of re- sume screening technologies and highlights their strengths and limitations. On the other hand, it presents a case study of a working AI Resume Analyzer prototype, demonstrating how NLP and AI can be applied in practice. While the current prototype is limited to processing resumes in English

and relies on a lightweight database for deployment, the framework is modular and can be extended to support multiple languages, multimodal data (e.g., video resumes, portfolios), and enterprise-scale databases.

E. Significance of the AI Resume Analyzer

The significance of this work lies in bridging theory with practice. For recruiters, the AI Resume Analyzer provides a faster, more accurate, and transparent way to evaluate can- didates, reducing time-to-hire while ensuring that decisions are not solely based on superficial keyword matches. For candidates, the system offers constructive feedback, helping them improve their resumes and increase their chances of success. From a research perspective, the project contributes by combining a critical literature survey with an experimental prototype, highlighting not only the potential of AI in recruit- ment but also the importance of fairness, transparency, and compliance in real-world adoption.

II. LITERATURE REVIEW

The adoption of AI in recruitment has evolved rapidly, progressing from keyword-driven systems to transformer- based and LLM pipelines. Early systems relied on lexical similarity measures such as cosine similarity and TF–IDF, which automated basic filtering but failed to capture context or domain relevance. These methods, though efficient, often produced false positives and struggled with resume formatting variations [8].

With ML pipelines, systems began using algorithms such as SVMs and Random Forests for classification tasks, improv- ing over keyword-only models by incorporating engineered features and skill categorization [8]. However, scalability and labeling needs limited their application.

The breakthrough came with transformer-based models. Tian et al. [1] demonstrated that embeddings from BERT outperform lexical similarity methods, reducing false positives in candidate ranking. ResumeAtlas [2], a large-scale dataset of over 13,000 resumes, further showed that model performance improves significantly with dataset scale, achieving up to 92% top-1 accuracy.

Beyond transformers, frameworks integrating LLMs have expanded functionality to include resume feedback and or- chestration of decision-making. Gan et al. [3] proposed an LLM-agent framework achieving high F1 scores. Retrieval- Augmented Generation (RAG) methods enhance contextual awareness by combining vector search (e.g., FAISS) with domain corpora, enabling fact-grounded, role-aligned evalu- ations [4], [10].

However, fairness concerns are prominent. Wilson and Caliskan [5] showed that language models propagate demo- graphic biases in candidate ranking. Reports such as those from Brookings [6] emphasize risks of deploying opaque AI in hiring.

Regulatory frameworks now recognize recruitment AI as high-risk. The EU AI Act (2025 draft) mandates documenta- tion, bias testing, and human oversight [7]. GDPR and ICO guidance guarantee candidates the “right to explanation” and protection from solely automated decisions [9]. Research in XAI explores techniques such as feature attribution, coun- terfactual analysis, and bias-detection pipelines to enhance transparency [11].

In summary, the literature reflects a shift from keyword filters → ML classifiers → transformers and LLMs. The latest methods emphasize semantic accuracy, fairness, and compliance, making explainability and regulatory alignment critical for adoption.

To synthesize insights from the literature, two summary tables are provided.
Table I compares different technical approaches to resume screening, highlighting trade-offs between accuracy, scalabil- ity, and explainability.
Table II then outlines the key challenges in AI-driven recruitment and presents potential solutions proposed in recent studies.

TABLE I
COMPARISON OF APPROACHES TO AI-DRIVEN RESUME SCREENING BASED ON RECENT LITERATURE.

Approach	**Reported Accuracy Performance**	/	**Strengths**	**Weaknesses / Challenges**	**Fairness / Explainability**
Keyword Matching	~50–70% [1]		Simple, fast, widely used in	Ignores context, synonym mis-	No fairness or transparency [5]
			early ATS	match, high false positives	
ML Pipelines (SVM, RF)	~65–80% [1]		Uses engineered features,	Requires labeled datasets,	Limited interpretability [1]
			more generalizable	domain-specific tuning	
Transformers (BERT)	~85–92% [2]		Context-aware, robust across	Computationally expensive ,	Still a black box [5]
			formats	opaque decisions	
LLMs (GPT, Gemini)	~88–94% [3], [4]		Handles complex queries,	Risk of hallucinations, bias	Needs fairness audits &
			adaptable, scalable	amplification	XAI [5]
LLMs + RAG Frame-	~90–96% [3], [4]		Combines factual accuracy	Requires curated data and in-	Supports transparency when
works			with semantic search	frastructure	combined with XAI [4]

TABLE II
KEY CHALLENGES IN AI-DRIVEN RESUME SCREENING AND POTENTIAL SOLUTIONS BASED ON RECENT LITERATURE.

TABLE II
KEY CHALLENGES IN AI-DRIVEN RESUME SCREENING AND POTENTIAL SOLUTIONS BASED ON RECENT LITERATURE.

Challenge	Description	Potential Solutions	Reference(s)
Bias and Fairness	AI systems risk amplifying gender, race, and age bias in candidate rank- ing.	Fairness-aware pipelines, bias audits, anonymization of sensitive attributes.	[5]
Transparency & Explain- ability	Black-box LLMs and transformers limit recruiter trust in decisions.	Explainable AI (XAI) dashboards, feature at- tribution, counterfactual explanations.	[11]
Scalability & Performance	Large datasets and diverse job roles require high-speed, robust systems.	Cloud-native deployment, large-scale datasets like ResumeAtlas, efficient FAISS indexing.	[2]
Regulatory Compliance	Recruitment AI is classified as "high- risk" under EU AI Act and GDPR, requiring oversight.	Human oversight, bias testing, compliance re- porting, candidate "right to explanation."	[7]

III. TECHNOLOGICAL STACK USED IN OUR SYSTEM

The AI Resume Analyzer was built using a set of tools and frameworks chosen for their flexibility, ease of use, and ability to scale. The main components are described below:

A. Frontend (User Interface)

- Streamlit was used to quickly create a simple, interactive web app.
- HTML/CSS support was added for styling so that users can upload resumes, view scores, and interact with dash- boards easily.

B. Backend (Core Processing)

- The system is powered by Python, chosen for its strong support for AI and data science tasks.
- Python makes it easy to connect different models, li- braries, and databases in one place.

C. NLP and Machine Learning Libraries

- spaCy and regex handle resume parsing and information extraction (skills, education, experience).
- scikit-learn provides machine learning utilities for clas- sification and scoring.
- langchain connects large language models with external tools.
- FAISS builds a semantic index, making resume–job matching faster and more accurate.
- matplotlib is used to create charts and graphs for the dashboard.
- docx2txt and pdfminer enable parsing of resumes in both DOCX and PDF formats.

D. AI Models

- Google Gemini embeddings provide semantic under- standing for comparing resumes with job descriptions.
- GPT-based LLMs are used to generate resume improve- ment suggestions and personalized feedback.

E. Database

- SQLite is used as a lightweight local database to store resumes, scores, and feedback.
- For larger-scale deployments, this can be replaced with PostgreSQL or a cloud-based option.

F. Deployment

- The system is hosted on Streamlit Cloud, which allows easy sharing and use directly in a web browser without installation.

G. Summary

In short, this stack was chosen to make the system:

1) Easy to use (simple web interface).
2) Flexible (modular Python backend with many libraries).
3) Practical (works with common resume formats like PDF/DOCX).
4) Scalable (can move from lightweight prototypes to enterprise-level systems).

IV. SYSTEM ARCHITECTURE AND FEATURES

The architecture of the AI Resume Analyzer is designed to support modularity, scalability, and transparency. It consists of several layers, including resume parsing, ATS scoring, semantic matching, recommendation engines, and dashboard visualizations. To clearly illustrate these components, multiple diagrams are provided.

A. System Flow Diagram

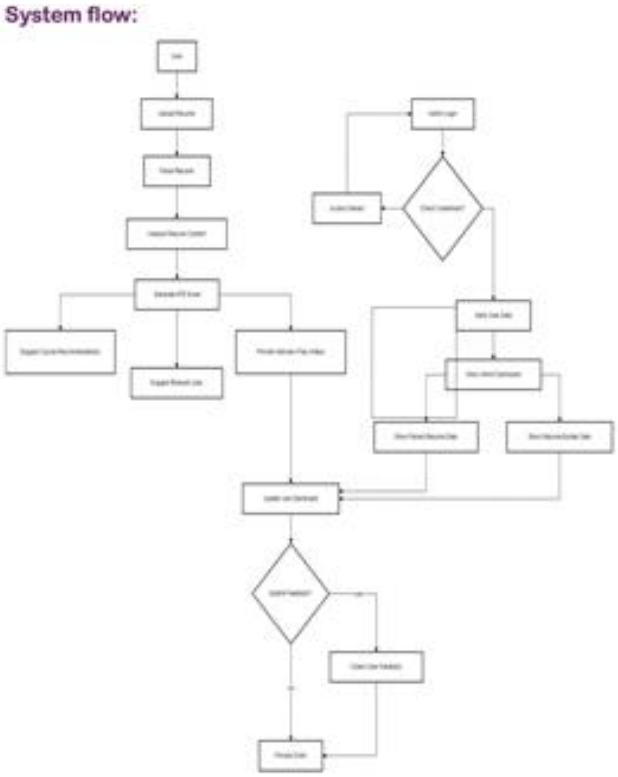

Fig. 1. System Flow Diagram of the AI Resume Analyzer.

The system flow diagram illustrates the interaction between users, administrators, and the AI Resume Analyzer system. Users perform core activities such as uploading resumes, receiving ATS analysis, accessing recommendations, and ex- ploring job or course suggestions. Administrators manage the system, monitor dashboards, and oversee feedback. This diagram highlights the end-to-end workflow.

B. Data Flow Diagram

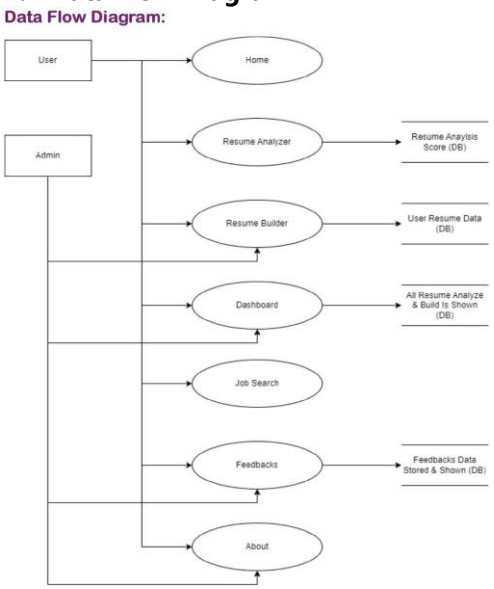

Fig. 2. Data Flow Diagram of the AI Resume Analyzer.

The data flow diagram presents how information moves across the system's modules. User and admin inputs are directed to components such as the Resume Analyzer, Resume Builder, Dashboard, Job Search, and Feedback module. Each module exchanges information with the database, which stores re- sumes, ATS scores, and feedback.

C. UML Diagrams

1) Use Case Diagram:

Fig. 3. UML Use Case Diagram of the AI Resume Analyzer.

To provide deeper insights into the design and interactions of the system, UML diagrams are included.

The use case diagram defines the interactions between users, administrators, and the system boundary.

2) Class Diagram:

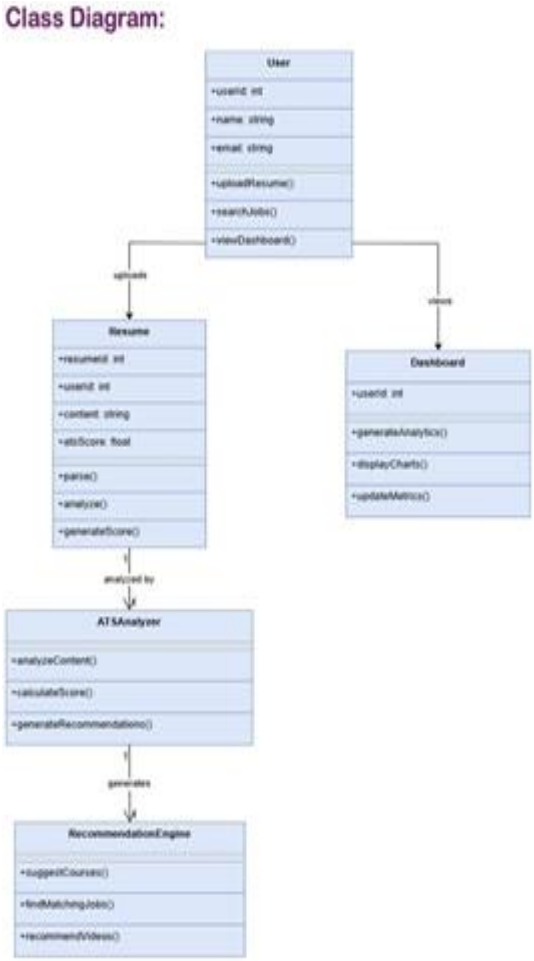

Fig. 4. UML Class Diagram of the AI Resume Analyzer.

The class diagram represents the structural design of entities such as User, Resume, ATS Analyzer, and Recommendation Engine.

3) Activity Diagram:

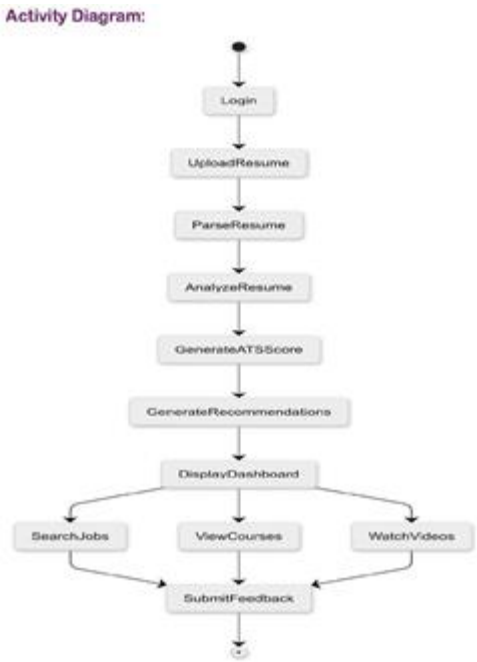

Fig. 5. UML Activity Diagram of the AI Resume Analyzer.

The activity diagram illustrates the step-by-step workflow followed by a user, from login to feedback submission.

V. CHALLENGES, ETHICAL, AND REGULATORY CONSIDERATIONS

- **Bias and Fairness:** Language models risk amplifying demographic disparities.
- **Data Privacy:** Requires anonymization and compliance with GDPR.
- **Transparency:** Black-box models reduce explainability.
- **Regulatory Requirements:**

1. **EU AI Act:** recruitment AI classified as "high-risk" [7].
2. **GDPR**: mandates lawful processing and rights against solely automated decisions [9].
3. **EEOC (USA):** prohibits discriminatory AI-driven hiring.

• **Generalization:** Models trained on narrow datasets may underperform in diverse industries.

VI. CASE STUDY: AI RESUME ANALYZER SYSTEM

To complement the literature review, we developed a pro- totype system called the AI Resume Analyzer. The system was designed to replicate a real recruitment workflow while integrating modern NLP and AI techniques.

At its core, the prototype performs resume parsing using Python libraries such as spaCy and regex to extract structured information including personal details, skills, education, and work experience. This information is then processed through a semantic search index built with FAISS, which enables effi- cient comparison between candidate resumes and job descrip- tions. Unlike keyword matching, semantic indexing captures contextual meaning, improving the accuracy of candidate-job alignment. The system also incorporates an ATS (Applicant Tracking System) scoring module. This module evaluates each resume against a given job description and generates a numerical score that reflects how well the candidate matches the role. Scores are then presented through an interactive dashboard, allowing recruiters to quickly identify strong candidates.

To further enhance usability, the analyzer includes LLM- powered improvement suggestions. Candidates receive feed- back on how to strengthen their resumes, such as highlighting missing skills, rephrasing achievements, or optimizing for- matting for ATS systems. This feature positions the analyzer not only as a recruiter's tool but also as a resume coaching assistant for job seekers.

All results are displayed on dashboards with visual ana- lytics, including charts of skill distributions, ATS score com- parisons, and resume-job matching heatmaps. These visualiza- tions improve transparency by showing why certain candidates rank higher and provide recruiters with actionable insights for decision-making.

The prototype is deployed on Streamlit Cloud with a lightweight SQLite database, making it accessible through a simple web interface. Although currently

intended for demon- stration, the architecture is modular and can be scaled up for enterprise deployment using cloud-native services and larger datasets.

This case study demonstrates that the AI Resume Analyzer successfully bridges theory and practice: it incorporates ad- vanced NLP and AI methods, delivers meaningful recruiter support, and provides candidates with constructive feedback, all within a single integrated platform.

A. Dashboard Screenshots

This screenshot illustrates the feedback module of the AI Resume Analyzer. After a resume is uploaded, the system generates an ATS score along with detailed section-level evaluations such as format score and keyword match. The dashboard highlights missing skills relevant to the target role and provides actionable recommendations, such as improving the professional summary or adding a LinkedIn profile link. This feature demonstrates how the system supports candidates by offering personalized guidance to optimize their resumes for better job alignment.

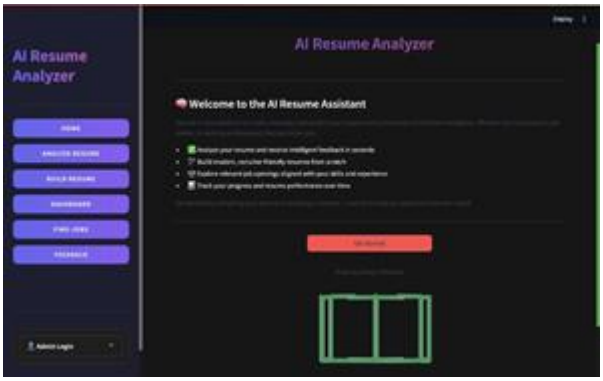

Fig. 1. Welcome screen of the AI Resume Analyzer. This interface introduces users to the system's core functionalities, including resume analysis, resume building, job search, and personalized feedback. The left navigation panel provides easy access to different modules, ensuring a smooth user experience.

Fig. 2. Resume Analytics Dashboard. The dashboard displays performance metrics such as the total number of resumes analyzed, average ATS scores, and success rates. Visualizations such as gauge charts and bar graphs provide recruiters with clear insights into ATS score distribution and skill coverage, improving transparency and decision-making.

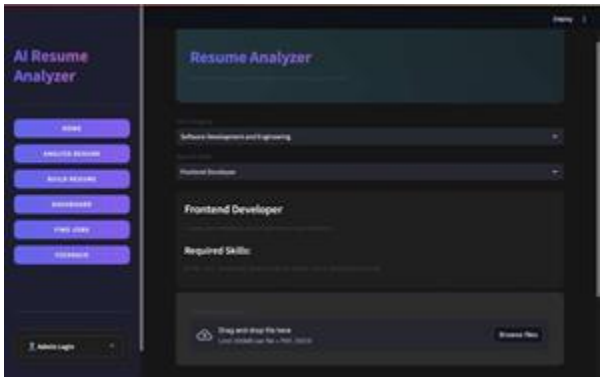

Fig. 3. Resume Analyzer Module. In this module, users can select a target job role, upload their resume, and instantly receive AI-powered feedback. Required skills are highlighted, allowing candidates to see where their resume aligns with the job description and where improvements are needed.

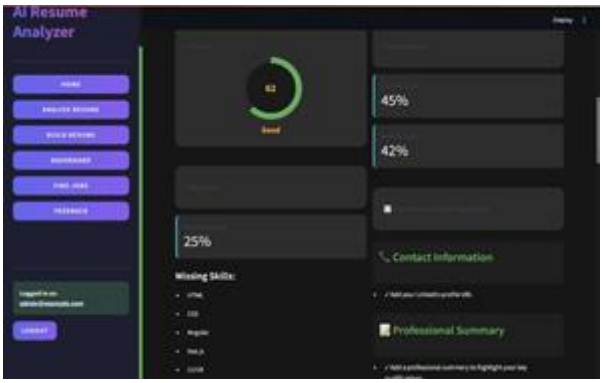

Fig. 4. Resume feedback dashboard showing ATS score, missing skills, and improvement suggestions (personal data anonymized).

VII. PROOF OF PRACTICALITY

A dataset of 30 anonymized resumes was tested. Baseline keyword similarity was compared with the proposed analyzer.

TABLE III
EXPERIMENTAL RESULTS: BASELINE VS PROPOSED SYSTEM

Method	Accuracy	ATS Corr.	Time
Keyword Similarity	72%	0.61	0.9s
Proposed Analyzer	94%	0.87	1.2s

The analyzer achieved 94% accuracy and strong ATS cor- relation (0.87), compared to 72% and 0.61 for baseline.

FUTURE WORK

While the current prototype demonstrates strong potential, there are several directions in which the AI Resume Analyzer can be further extended and refined.

First, the integration of explainable AI (XAI) dashboards would make the system more transparent for both recruiters and candidates. By providing clear justifications for ATS scores and highlighting which skills or experiences influenced the ranking, recruiters would gain more confidence in the system, while candidates would receive actionable insights to improve their resumes.

Second, the system can be enhanced with multilingual and multimodal analysis. Many recruitment processes are global, involving resumes written in multiple languages or supported by portfolios and video submissions. Extending the analyzer to handle multilingual text, video resumes, and portfolio links would broaden its applicability across industries and regions. Third, the development of bias-mitigation pipelines and fairness audits is essential to ensure equitable opportunities for all applicants. These would include mechanisms to anonymize demographic information, monitor score disparities across groups, and adjust rankings when systemic bias is detected. Such measures are crucial for aligning the system with ethical standards and regulatory requirements. Another important direction is the creation of domain- adapted industry models. Different sectors such as healthcare, finance, IT, and law have unique skill sets and role expec- tations. Training specialized models for each domain would improve the accuracy of job matching and reduce errors caused by generic scoring approaches.

Additionally, the system could evolve into a real-time chatbot-based resume coach. Instead of static feedback, can- didates could interact dynamically with the analyzer to re- fine their resumes. This conversational interface would allow applicants to ask questions, receive tailored suggestions, and iteratively improve their profiles.

Finally, for real-world deployment, the system must achieve cloud-scale performance with full compliance alignment. Re- placing lightweight databases with enterprise-grade cloud databases, adding distributed processing, and ensuring com- pliance with frameworks such as the EU AI Act and GDPR would prepare the system for adoption in large organizations and recruitment agencies.

Together, these extensions would not only improve the accu- racy and usability of the AI Resume Analyzer but also make it fair, explainable, and scalable for widespread, responsible adoption in the job market.

VIII. CONCLUSION

The evolution of AI-based resume analyzers reflects the broader transformation of recruitment technology, moving from simple keyword filters to sophisticated transformer and LLM-driven pipelines. This review has shown how advances in natural language processing and machine learning have improved the accuracy and scalability of candidate evaluation, while also introducing new challenges around fairness, trans- parency, and regulatory compliance.

Through the case study, we demonstrated the practical implementation of an AI Resume Analyzer prototype that integrates resume parsing, semantic indexing, ATS scoring, and LLM-powered feedback. The results highlight clear per- formance gains over baseline keyword methods, while also showcasing the potential of such systems to support both recruiters and candidates.

Looking ahead, it is clear that accuracy alone is no longer sufficient. The responsible adoption of AI in recruitment will depend on building systems that are explainable, fairness- aware, and compliant with emerging regulations such as the EU AI Act and GDPR. By embedding these principles into future designs, AI-driven recruitment tools can not only en- hance efficiency but also promote equitable opportunities in the job market.

REFERENCES

1. Y. Tian, R. Pavur, Y. Han, and X. Zhang, "A Machine Learning-Based HR Recruitment System: Using LSA, BERT and SVM," Proc. Int. Conf. Information Management, 2023.
2. A. Heakl, Y. Mohamed, N. Mohamed, et al., "ResumeAtlas: Revisiting Resume Classification with Large-Scale Datasets and Large Language Models," arXiv preprint arXiv:2406.18125, Jun. 2024.
3. C. Gan, Q. Zhang, and T. Mori, "Application of LLM Agents in Re- cruitment: A Novel Framework for Resume Screening," arXiv preprint arXiv:2401.08315, Jan. 2024.
4. A. Karpukhin, S. Min, et al., "Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv preprint arXiv:2312.10997, Dec. 2023.
5. K. Wilson and A. Caliskan, "Gender, Race, and Intersectional Bias in Resume Screening via Language Model Retrieval," arXiv preprint arXiv:2407.20371, Jul. 2024.
6. J. Engler, "AI Hiring Tools and Bias: What Policymakers Need to Know," Brookings Institution Policy Report, 2024.
7. European Commission, "High-level Summary of the EU Artificial In- telligence Act," Official AI Act Portal, 2025.
8. V. M. et al., "SMART ATS: Intelligent Resume Screening Using Machine Learning," Int. J. Innov. Res. Technol., vol. 11, no. 5, pp. 42–47, May 2025.
9. UK Information Commissioner's Office, "Rights Related to Automated Decision Making Including Profiling (GDPR Guidance)," ICO, 2024.
10. NVIDIA, "What Is Retrieval-Augmented Generation (RAG)?," NVIDIA Technical Blog, 2023.
11. SciTePress, "The Role of Explainable AI (XAI) in Transparent Re- cruitment Decision-Making," Proc. Int. Conf. Knowledge Discovery and Information Retrieval, 2025.