IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13327

AI RESUME ANALYZER

Sowmiya C[1], Dr. K. Santhi[2]

Student, Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore.[1]

Professor, Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore.[2]

Abstract: The rapid growth of digital recruitment platforms has intensified the need for automated resume screening systems that efficiently evaluate candidate profiles and match them with suitable job roles. However, many existing Applicant Tracking Systems (ATS) operate as opaque, proprietary solutions, limiting transparency and accessibility for job seekers. This paper presents an NLP-driven automated resume analysis and job matching framework designed to democratize access to career optimization tools. The proposed system integrates classical Natural Language Processing (NLP) techniques—including tokenization, lemmatization, Named Entity Recognition (NER), TF-IDF vectorization, and cosine similarity—to evaluate resume quality, simulate ATS compatibility scoring, and recommend relevant job opportunities.

The framework employs a hybrid scoring model combining keyword relevance, section completeness, skill alignment, and formatting compliance to generate interpretable resume scores. Additionally, a similarity-based ranking algorithm matches resumes with structured job descriptions stored in a PostgreSQL database. The system is designed for transparency, determinism, and reproducibility, ensuring explainable outputs suitable for academic and practical deployment. Experimental evaluation demonstrates that the proposed method provides consistent, scalable, and computationally efficient job recommendations while maintaining interpretability—a critical requirement for ethical AI deployment in recruitment technologies.

Keywords: Natural Language Processing (NLP), Applicant Tracking System (ATS), Resume Scoring, Job Recommendation, Skill Gap Analysis, TF-IDF Vectorization, Cosine Similarity, Named Entity Recognition (NER), Machine Learning**.**

I. INTRODUCTION

The recruitment landscape has undergone significant transformation due to digitization and large-scale online job portals. Organizations increasingly rely on Applicant Tracking Systems (ATS) to automatically filter, rank, and shortlist resumes before human evaluation. While these systems enhance efficiency, they often lack transparency and may disadvantage candidates unfamiliar with ATS optimization techniques. Consequently, there is a growing need for accessible, explainable, and technically sound systems that provide candidates with actionable insights into resume quality and job compatibility.

Natural Language Processing (NLP) has emerged as a powerful tool for extracting structured information from unstructured text data, making it well-suited for resume analysis. Resumes are inherently semi-structured documents containing skills, work experience, education, and certifications in free-text formats. Extracting meaningful insights from such documents requires robust linguistic preprocessing, entity recognition, and semantic similarity computation.

This paper proposes an NLP-driven framework for automated resume evaluation and job matching that simulates ATS-style screening mechanisms. Unlike black-box AI solutions, the proposed approach emphasizes deterministic scoring algorithms and explainable outputs. By combining keyword relevance analysis, structural validation, and vector-space similarity modelling, the system generates an ATS compatibility score and ranked job recommendations. The implementation leverages Python-based NLP libraries and a PostgreSQL-backed structured job repository, ensuring scalability and maintainability. The primary objective is to democratize career success by providing a transparent, technically grounded platform that enables users to iteratively improve their resumes and align them with market demands.

II. LITERATURE REVIEW

Earlier studies have explored resume screening using keyword-based systems. Sharma and Verma (2020) proposed an automated resume parser that extracted candidate details but lacked scoring and recommendation features. Singh et al.

(2021) developed a machine learning-based job matching system using TF-IDF similarity; however, the system was not explainable to users.

Kumar and Rao (2022) introduced an Applicant Tracking System (ATS) scoring model using Natural Language Processing techniques. Their work demonstrated how resume structure, keyword relevance, and skill alignment could be used to evaluate candidate suitability. However, many such systems are proprietary and do not provide accessible or user-friendly tools for job seekers.

Recent research has explored advanced Natural Language Processing (NLP) techniques such as Named Entity Recognition and semantic similarity methods for analyzing resume content and matching it with job requirements. NER helps extract key entities from resumes, including technical skills, educational qualifications, organizations, and job titles. These extracted entities are compared with job descriptions using semantic similarity techniques to evaluate how well a candidate's profile matches a specific role. Such approaches improve the accuracy of resume parsing and job matching compared to traditional keyword-based filtering systems.

From the literature, it is evident that there is a need for an explainable and user-friendly AI-based resume analyzer that not only evaluates resumes but also provides detailed feedback, job recommendations, and skill gap analysis. Such a system can help candidates improve their resumes step by step and increase their chances of securing suitable job opportunities in a competitive job market.

III. PROBLEM STATEMENT

In the modern recruitment process, companies receive a large number of resumes for every job opening. Due to this high volume, recruiters often use Applicant Tracking Systems (ATS) or manual screening methods to filter candidates. However, manual resume evaluation is time-consuming, inconsistent, and may lead to biased shortlisting. Many qualified candidates fail to get selected because their resumes are poorly structured, missing important keywords, or not aligned with job requirements.

Most job seekers are unaware of ATS rules and do not know whether their resume matches industry expectations. Existing resume screening tools mainly focus on basic keyword filtering and do not provide detailed feedback on missing skills, formatting issues, or section completeness. As a result, candidates do not receive proper guidance to improve their resumes and increase their chances of selection.

Although some automated resume analysis systems exist, many of them are proprietary black-box solutions that lack transparency and personalization. They do not clearly explain why a resume is rejected or what improvements are needed. Therefore, there is a need for an intelligent, automated, and user-friendly system that can analyze resumes, generate an overall score, recommend suitable job roles, and suggest missing skills and relevant courses.

The proposed AI-Based Virtual Resume Analyzer addresses this problem by using Artificial Intelligence and Natural Language Processing techniques to evaluate resume quality and provide actionable recommendations. This system helps job seekers optimize their resumes step by step and improves their employability in a competitive job market.

IV. METHODOLOGY

The proposed system follows a modular pipeline architecture consisting of Resume Ingestion, NLP Processing, ATS Scoring, Job Matching, and Recommendation Generation. Each stage is designed for interpretability and computational efficiency.

A. RESUME INGESTION AND TEXT EXTRACTION

Resumes are uploaded in PDF format and processed using a text extraction engine to convert document content into machine-readable text. Preprocessing steps include:

- Lowercasing
- Removal of punctuation and special characters
- Stop word elimination
- Tokenization
- Lemmatization

These preprocessing steps ensure normalization of linguistic variations (e.g., "developed," "developing," → "develop"), improving matching consistency.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13327

B. NAMED ENTITY RECOGNITION AND SKILL EXTRACTION

To identify key resume components, Named Entity Recognition (NER) is applied using spacy-based models. The system extracts:

- Technical skills
- Job titles
- Educational qualifications
- Organizations
- Experience duration indicators

Extracted skills are mapped to a standardized skill taxonomy stored in PostgreSQL. Synonym normalization (e.g., "JS" → "JavaScript") enhances matching accuracy.

C. ATS COMPATIBILITY SCORING MODEL

The ATS simulation model evaluates resumes using a weighted multi-factor scoring approach:

$$\text{ATS Score} = (\text{Keyword Match} \times 0.5) + (\text{Section Completeness} \times 0.2) + (\text{Skill Relevance} \times 0.2) + (\text{Formatting Compliance} \times 0.1)$$

- **Keyword Match**: Measures overlap between resume tokens and job-role keyword sets.
- **Section Completeness**: Detects presence of mandatory sections (Skills, Experience, Education).
- **Skill Relevance**: Evaluates alignment between extracted skills and industry-defined skill requirements.
- **Formatting Compliance**: Penalizes ATS-unfriendly structures such as excessive tables or multi-column layouts.

This structured scoring approach ensures deterministic, interpretable results rather than opaque AI predictions.

D. JOB MATCHING VIA VECTOR SPACE MODELLING

For job recommendation, the system employs TF-IDF (Term Frequency–Inverse Document Frequency) vectorization to represent both resume content and job descriptions in a high-dimensional vector space.

Cosine similarity is then computed to quantify semantic closeness:

$$\text{Similarity} = \frac{\text{A} \cdot \text{B}}{|| A || \ \ || B ||}$$

Where:

- A= Resume TF-IDF vector
- B= Job description TF-IDF vector

To enhance relevance, the final job match score is computed as:

$\text{Final Job Score} = (\text{Cosine Similarity} \times 0.7) + (\text{Skill Match Percentage} \times 0.3)$

Jobs are ranked in descending order, and the top N roles are recommended.

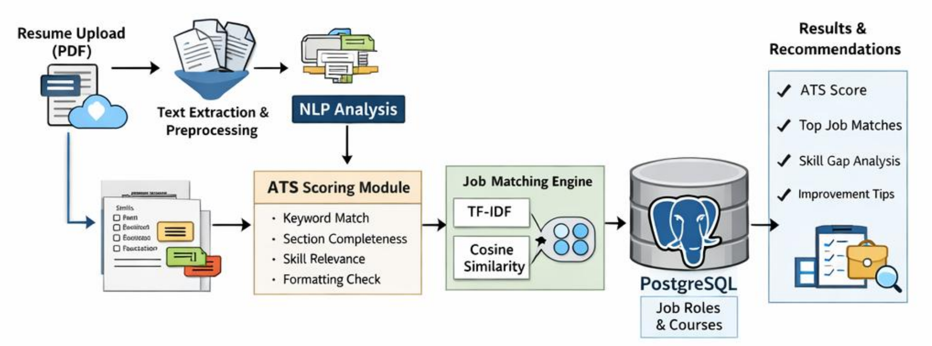


Figure 1: System Architecture of the NLP-Driven Resume Analyzer.

E. DATABASE ARCHITECTURE

A PostgreSQL relational database stores:

- User profiles
- Resume metadata

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.13327

- Extracted skill entities
- Structured job descriptions
- Match results

Normalization and indexing are applied to ensure query efficiency and scalability. Version control of resumes enables comparative score analysis across updates.

F. EXPLAINABILITY AND TRANSPARENCY

Unlike deep learning-based black-box systems, the proposed framework prioritizes explainability. For each recommendation, the system provides:

- Matched keywords
- Missing skills
- Section-wise ATS scores
- Improvement suggestions

This ensures users understand how scores are derived and how to enhance resume performance.

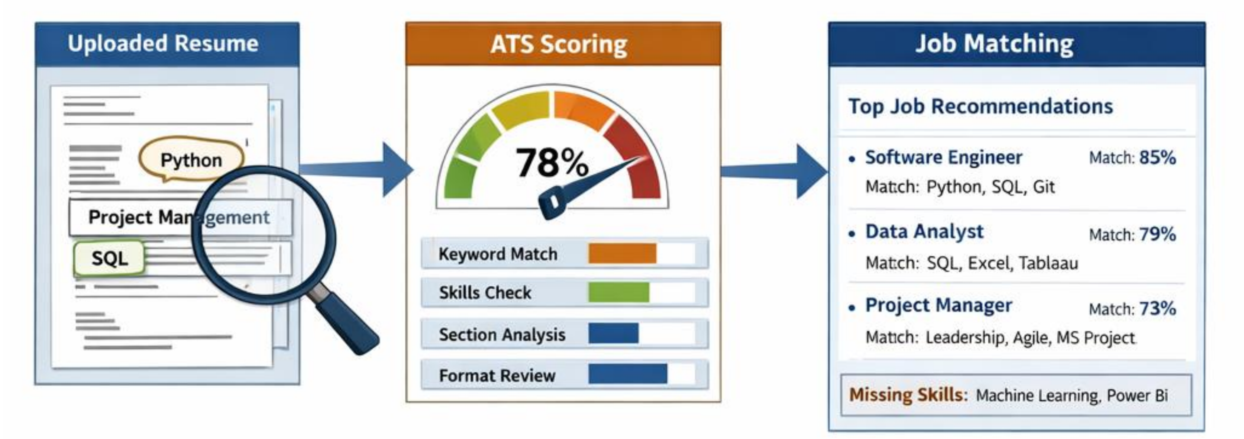
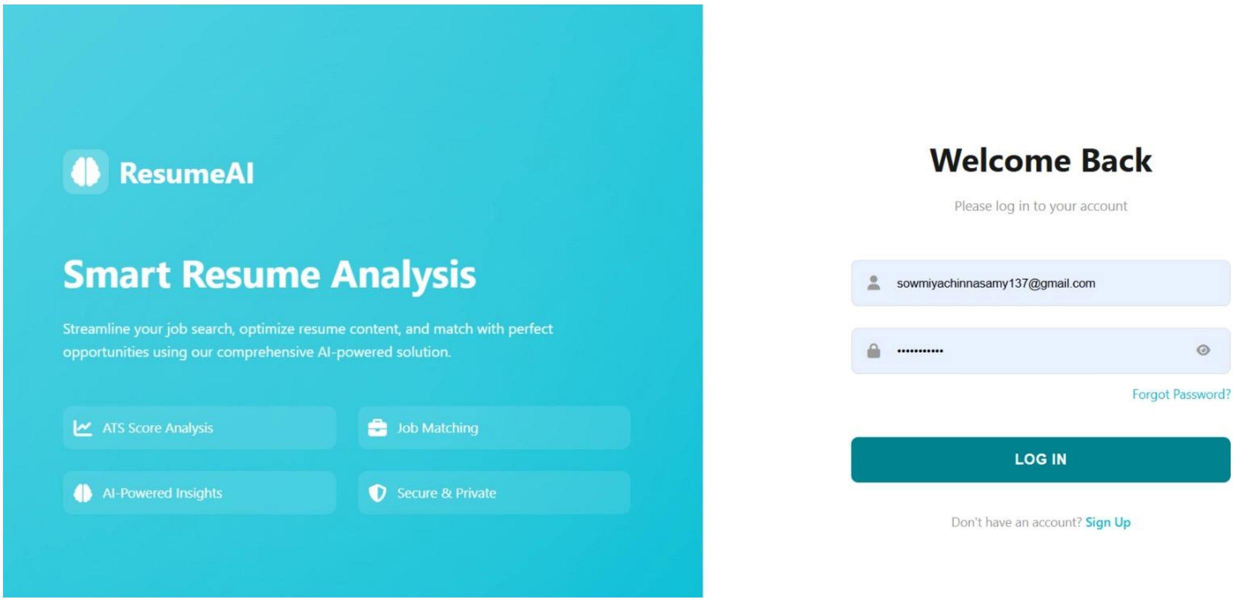


Figure 2: *ATS Scoring and Job Matching Process.*

V. USER INTERFACE



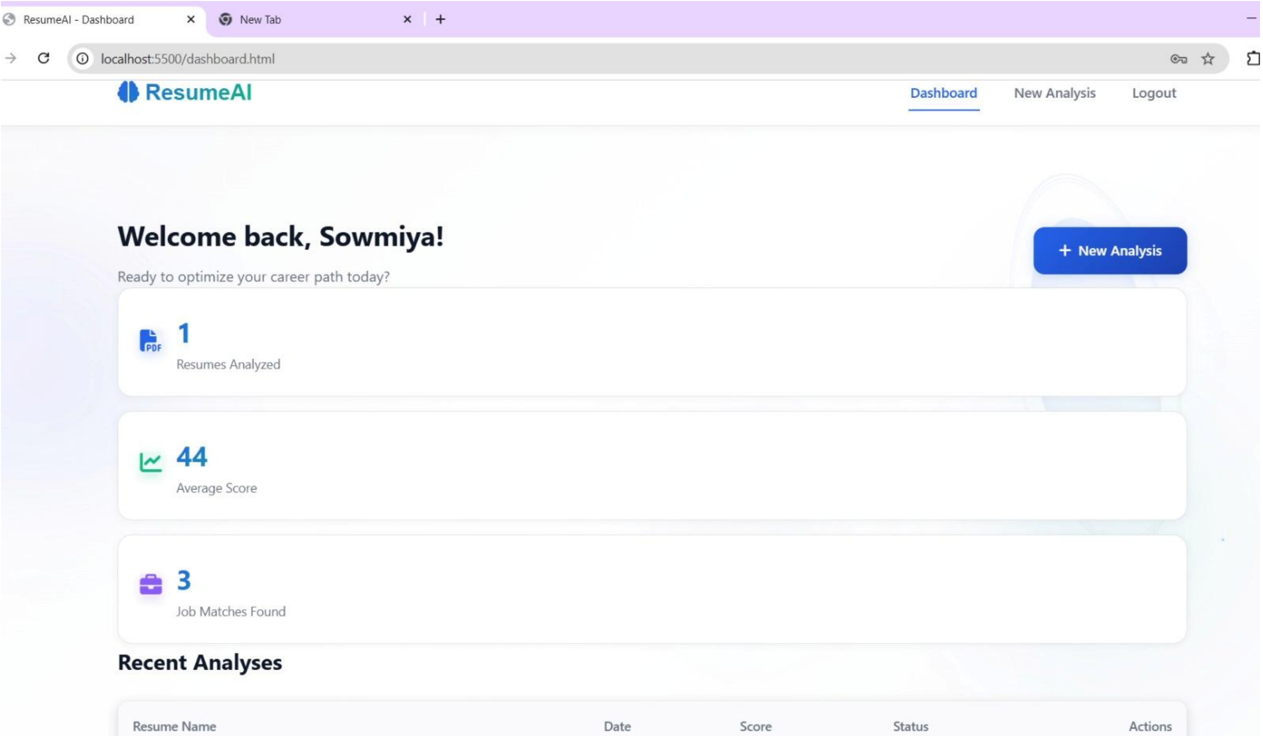
[Figure 3: Resume AI User Login Interface]

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

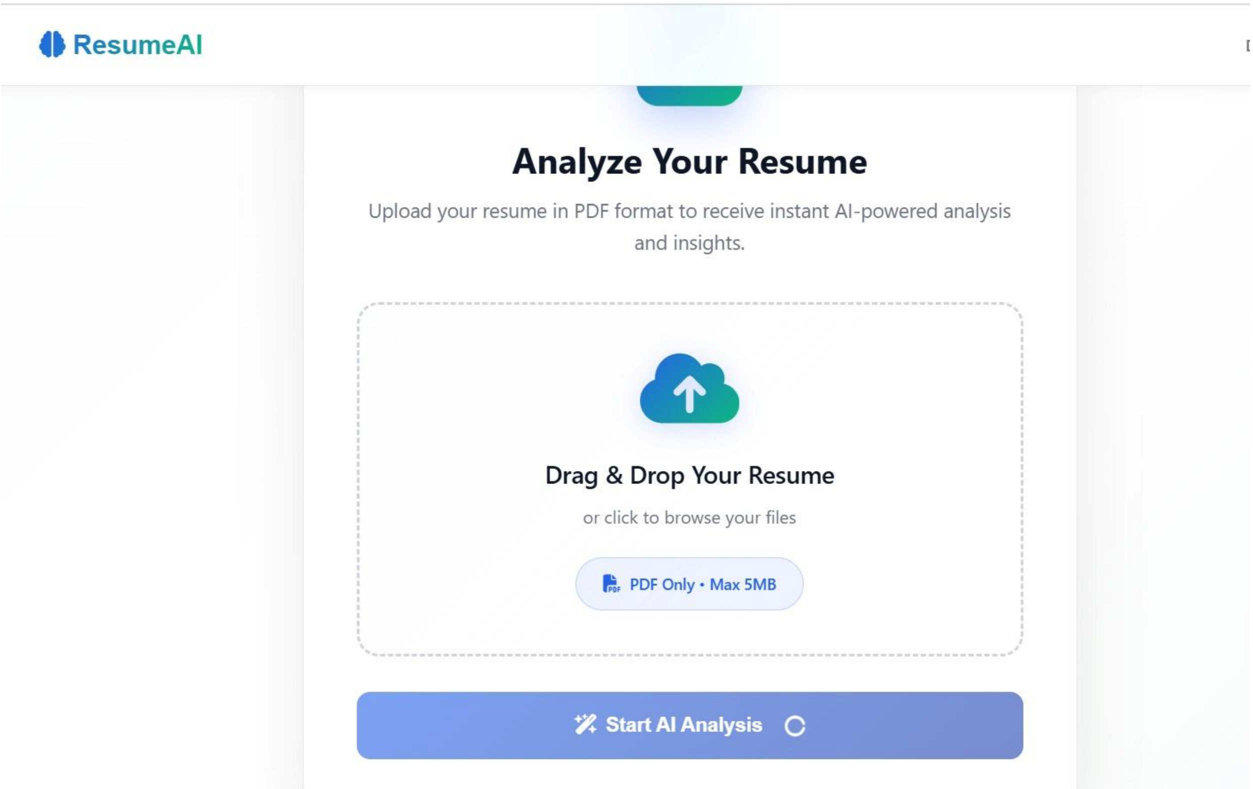
International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13327



[Figure 4: Resume Upload and AI Analysis Page]



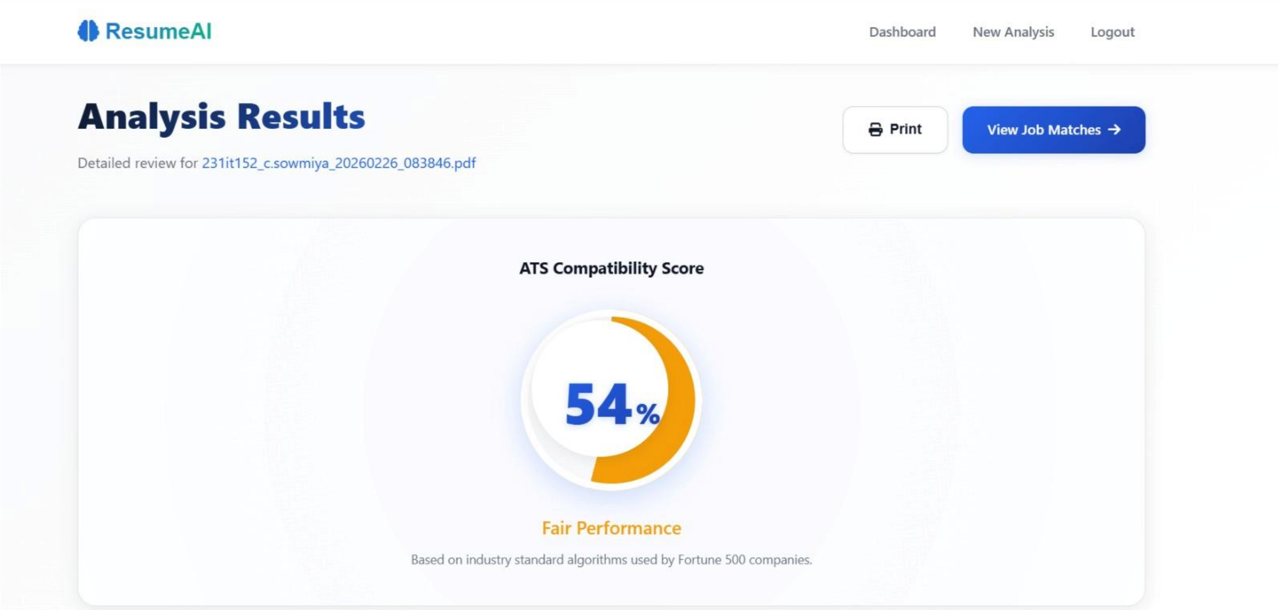
[Figure 5: Resume Analysis Results with ATS Compatibility Score]

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

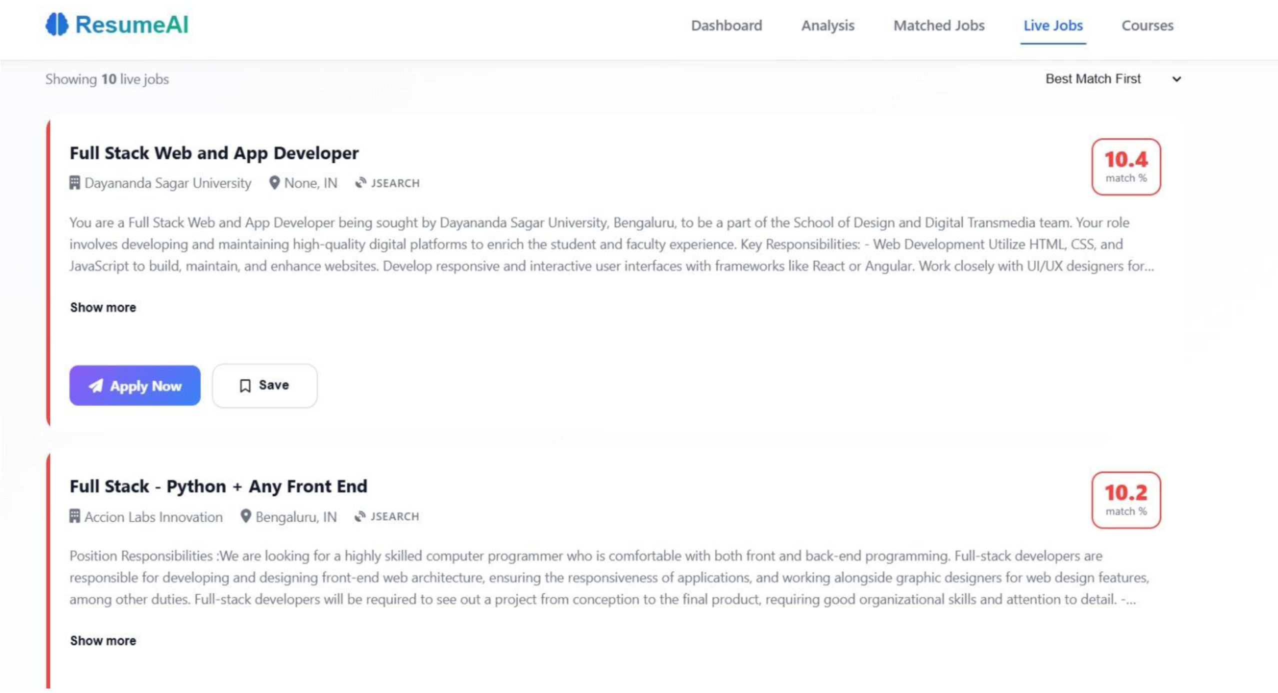
International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13327



[Figure 6: Resume AI User Dashboard]



[Figure 7: Live Job Matching and Recommendation Page]

VI. CONCLUSION

This research presents a transparent and computationally efficient NLP-driven framework for automated resume analysis and job matching. By integrating classical NLP techniques with structured scoring algorithms, the system simulates ATS behavior while maintaining interpretability and fairness. The hybrid scoring mechanism—combining keyword relevance, structural validation, and vector-space similarity—enables deterministic and reproducible evaluation outcomes.

The implementation demonstrates that lightweight NLP approaches such as TF-IDF and cosine similarity can achieve scalable and effective job-role matching without reliance on opaque deep learning architectures. By democratizing access to ATS-style evaluation tools, the proposed system empowers candidates to optimize resumes, identify skill gaps, and align their profiles with industry demands.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⨳ Peer-reviewed & Refereed journal ⨳ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13327

Future work may incorporate contextual embeddings (e.g., transformer-based models), multilingual support, and bias-mitigation strategies to further enhance fairness and robustness. Ultimately, this work contributes toward accessible, explainable AI solutions in recruitment technology, promoting equitable opportunities in the digital hiring ecosystem.

VII. FUTURE WORK

In the future, the AI-Based Virtual Resume Analyzer can be enhanced in several ways. The resume analysis accuracy can be improved by using advanced deep learning models such as BERT or transformer-based embeddings to understand resume content more effectively. The system can also be extended to support multiple resume formats like DOCX and multilingual resumes for wider usability.

Additional features such as real-time job portal integration, personalized career guidance, and interview preparation suggestions can further improve user experience. The skill gap analysis can be strengthened by including industry-specific certification recommendations and trending job market skills.

Moreover, bias detection and fairness mechanisms can be incorporated to ensure ethical and unbiased resume evaluation. Expanding the job description database and adding more role-specific scoring models will make the system more accurate, reliable, and beneficial for job seekers.

REFERENCES

[1]. C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press, 2008.
[2]. D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. (Draft). Pearson, 2023.
[3]. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2013.
[4]. G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.
[5]. J. Ramos, "Using TF-IDF to determine word relevance in document queries," in *Proceedings of the First Instructional Conference on Machine Learning*, 2003.
[6]. M. F. Porter, "An algorithm for suffix stripping," *Program,* vol. 14, no. 3, pp. 130–137, 1980.
[7]. M. Honnibal and I. Montani, "spaCy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing," 2017.
[8]. S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol, CA, USA: O'Reilly Media, 2009.
[9]. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[10]. PostgreSQL Global Development Group, "PostgreSQL Documentation," 2024. [Online]. Available: https://www.postgresql.org/docs/

