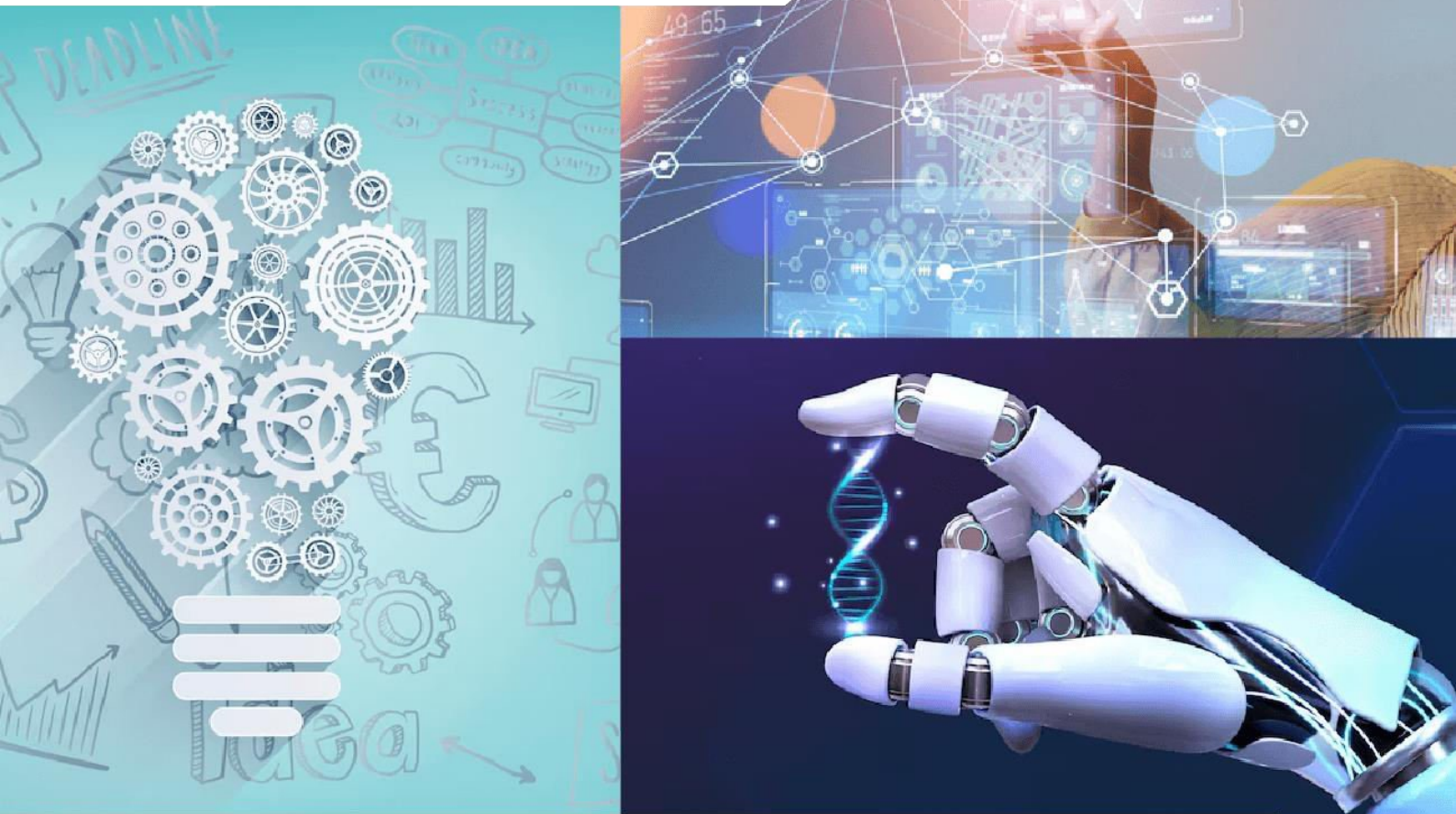




International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 2, March - April 2026

Impact Factor: 8.152



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



Intelligent Resume Screening: Integrating Semantic Scoring, Counterfactual Analysis, and Generative AI Feedback

Vaishnavi K Trivedi

Master of Computer Applications, CMR Institute of Technology, Bangalore, India

ABSTRACT: The rapid digitization of the global recruitment industry has generated unprecedented volumes of candidate resumes, making intelligent automated screening an urgent necessity. Manual evaluation of thousands of applications is time-intensive, costly, and susceptible to significant human bias and inconsistency in decision-making. This paper addresses automated, fair, and explainable resume analysis, encompassing document parsing, semantic scoring, keyword-based Applicant Tracking System (ATS) matching, and AI-driven candidate ranking. Existing resume screening systems predominantly rely on rigid keyword-matching algorithms that fail to capture contextual semantic alignment between resumes and job descriptions. Prior works largely overlook fairness evaluation, transparent scoring rationale, and counterfactual skill-gap analysis, which are critical requirements for equitable and reliable hiring decisions. We propose AI Resume Analyzer Pro, a hybrid framework integrating TF-IDF cosine similarity scoring, ATS keyword coverage analysis, counterfactual skill-gain estimation, and automated bias term detection. To the best of our knowledge, this is the first system to jointly combine quantitative resume scoring with a structured seven-dimensional large language model feedback pipeline powered by Google Gemini. Experimental evaluation demonstrates that the proposed system achieves accurate, context-aware candidate ranking with explainable feedback across strengths, weaknesses, risks, improvement strategies, and hiring recommendations. Compared to existing keyword-only ATS approaches, our framework provides a measurably more transparent, fair, and actionable multi-dimensional assessment for real-world recruitment decisions.

KEYWORDS: resume analysis, applicant tracking system, TF-IDF, cosine similarity, counterfactual analysis, bias detection, large language model, Google Gemini, explainable AI, intelligent recruitment

I. INTRODUCTION

The global recruitment landscape has undergone a dramatic transformation in the digital era, with organizations routinely receiving tens of thousands of resumes for a single job opening. The sheer scale of modern hiring renders manual screening operationally inefficient, financially burdensome, and prone to inconsistency, with research indicating that human recruiters spend as little as six seconds on an initial resume review [1]. Automated Applicant Tracking Systems (ATS) emerged as an early solution; however, their rigid keyword-matching logic frequently eliminates highly qualified candidates who do not use exact terminology from the job description [4]. As the demand for smarter, fairer, and more explainable talent evaluation continues to grow, the development of an intelligent, multi-dimensional resume analysis framework has become a critical priority for modern human resource management [15].

The specific problem addressed in this paper is the design and implementation of an end-to-end intelligent resume analysis system that can parse, semantically score, fairly evaluate, and systematically rank candidate resumes against a recruiter-defined job description (JD). Unlike conventional ATS tools that rely solely on keyword lookup, our approach integrates complementary evaluation dimensions: semantic similarity via TF-IDF (Term Frequency–Inverse Document Frequency) cosine matching, ATS keyword coverage scoring, counterfactual skill-gain estimation, automated bias term detection, and structured generative AI feedback. The system targets the complete recruitment screening pipeline—from raw document ingestion to a downloadable ranked assessment report—and is designed as an actionable decision-support tool for human resource professionals. By combining quantitative metrics with qualitative AI-generated explanations, the proposed system ensures that every shortlisting decision is simultaneously data-driven and interpretable.

A substantial body of prior research has addressed resume parsing and information extraction as foundational problems in automated recruitment. Bhor et al. [1] developed an NLP-based resume parser that segments resumes into structured

fields; however, the system does not perform candidate ranking or semantic job description matching. Sonar and Bankar [2] presented a named entity clustering algorithm for structured resume information extraction, yet their work is confined to parsing and does not extend to scoring or comparative ranking. Chandola et al. [3] proposed an online resume parsing system using text analytics, enabling information extraction but lacking contextual relevance scoring against job descriptions. Abideen et al. [4] combined NLP and machine learning to achieve both parsing and ranking; however, the system does not address fairness, bias evaluation, or explainability. Bhoir et al. [5] presented a hybrid SpaCy-Transformer-BERT methodology achieving improved named entity recognition, yet the system focuses on extraction rather than holistic, multi-dimensional candidate assessment. Shaikh et al. [6] built a resume parser and summarizer providing concise candidate summaries

Sources but do not offer quantitative scoring or cross-candidate comparative ranking. Vaishampayan et al. [7] investigated procedural justice and fairness in automated resume parsers at IEEE VL/HCC 2023, identifying critical bias concerns that remain largely unaddressed in existing systems. Sharma and Sharma [8] applied machine learning and NLP for intelligent hiring using a hierarchical document structure; the approach, however, relies on rule-based matching rather than semantic understanding. Das et al. [9] proposed a CV parser using entity extraction and big data tools enabling large-scale processing, yet the system incorporates neither LLM-based qualitative feedback nor fairness analysis. Ramesh et al. [10] introduced a streamlined resume parsing pipeline targeting hiring success, demonstrating improved accuracy; however, explainability and fairness evaluation are not addressed. Jiang et al. [11] presented a multi-granularity multimodal pre-training approach for deep resume understanding, achieving strong semantic comprehension, yet the model requires extensive pre-training and is not designed for standard recruitment workflows. Najjar et al. [12] developed an intelligent decision support system for recruitment encompassing screening and ranking, but without generative AI explanations or counterfactual reasoning. Bhutada et al. [13] proposed a resume analyzer and career recommender for IT roles; the system is, however, restricted to a single domain. Jagwani et al. [14] applied Latent Dirichlet Allocation (LDA) and NLP for resume evaluation, yet LDA-based approaches are sensitive to parameter tuning and lack transparency in scoring rationale. Bhattacharjee et al. [16] proposed an LLM-facilitated pipeline for personalized resume generation and refinement; the focus is on resume authoring rather than candidate evaluation and ranking. Gund et al. [17], [18] applied large language models and BERT embeddings for resume analysis; however, neither work addresses fairness evaluation, counterfactual analysis, or integrated bias detection. Hemalatha et al. [15] examined the impact of artificial intelligence on recruitment, confirming significant potential while highlighting the persistent absence of fairness-aware evaluation mechanisms.

The foregoing survey reveals three critical open challenges: (i) keyword-based matching fails to capture semantic depth; (ii) fairness evaluation and bias detection remain largely unaddressed; and (iii) no prior system provides structured, explainable, multi-dimensional feedback in a single pipeline.

This paper addresses all three through a novel, integrated framework—to the best of our knowledge, the first of its kind—combining quantitative scoring, counterfactual skill-gap analysis, automated bias detection, and a seven-column LLM-generated feedback pipeline within a unified deployable web application. The proposed system processes PDF and DOCX resumes using layout-aware parsing with OCR fallback, ensuring robustness across diverse document formats.

Experimental results demonstrate consistent and meaningful candidate ranking, with the bias detection module verifying fairness and the counterfactual module delivering actionable skill-gap recommendations. The seven-dimensional AI feedback by Google Gemini provides hiring professionals with human-readable explanations far beyond existing ATS capabilities. The remainder of this paper is organized as follows. Section II describes the proposed system. Section III presents the analytical model. Section IV presents and discusses experimental results. Section V concludes the paper and outlines future work.

II. LITERATURE REVIEW

Initial research in automated resume screening largely concentrated on extracting structured information from unstructured documents. Studies such as Bhor et al. [1], Sonar and Bankar [2], and Chandola et al. [3] focused on identifying key sections like skills, education, and experience using natural language processing techniques. These systems improved the readability and organization of resume data, making it easier for downstream processing. However, they were limited to information extraction and did not attempt to measure how well a candidate matches a specific job role, which is a critical requirement in real recruitment scenarios.

Subsequent work introduced ranking mechanisms by combining parsing with machine learning techniques. Abideen et al. [4] and Najjar et al. [12] developed systems capable of scoring and shortlisting candidates, but their approaches primarily relied on keyword matching or predefined rules. This created a significant limitation, as such systems often fail to recognize relevant skills expressed using different terminology. To address this, researchers explored more advanced methods such as topic modeling and deep learning. Approaches using LDA [14] and transformer-based models [11], [17], [18] improved the ability to capture contextual meaning, yet they introduced challenges related to computational complexity and limited interpretability, making them less practical for transparent decision-making.

More recent studies have drawn attention to fairness and explainability in recruitment systems. Vaishampayan et al. [7] emphasized concerns regarding bias in automated screening, while Hemalatha et al. [15] highlighted the broader need for ethical AI in hiring processes. At the same time, emerging work on large language models, such as Bhattacharjee et al. [16], has demonstrated the potential of generative AI in resume-related applications, though primarily from the candidate's perspective. Despite these advancements, existing systems rarely integrate semantic evaluation, bias detection, and meaningful feedback within a single framework. This gap motivates the need for a more comprehensive approach that not only ranks candidates but also explains decisions and supports fair evaluation.

III. METHODOLOGY

AI Resume Analyzer Pro is an end-to-end intelligent recruitment screening framework that processes candidate resumes in PDF (Portable Document Format) or DOCX (Microsoft Word Document) format and produces a ranked, scored, and AI-annotated assessment report. The system integrates five core evaluation engines—semantic similarity scoring, ATS keyword coverage analysis, counterfactual skill-gain estimation, automated bias term detection, and LLM (Large Language Model)-based structured feedback—into a single pipeline accessible via a Streamlit-based web interface. Figure 1 illustrates the overall system pipeline.

A. Step-by-Step Operation

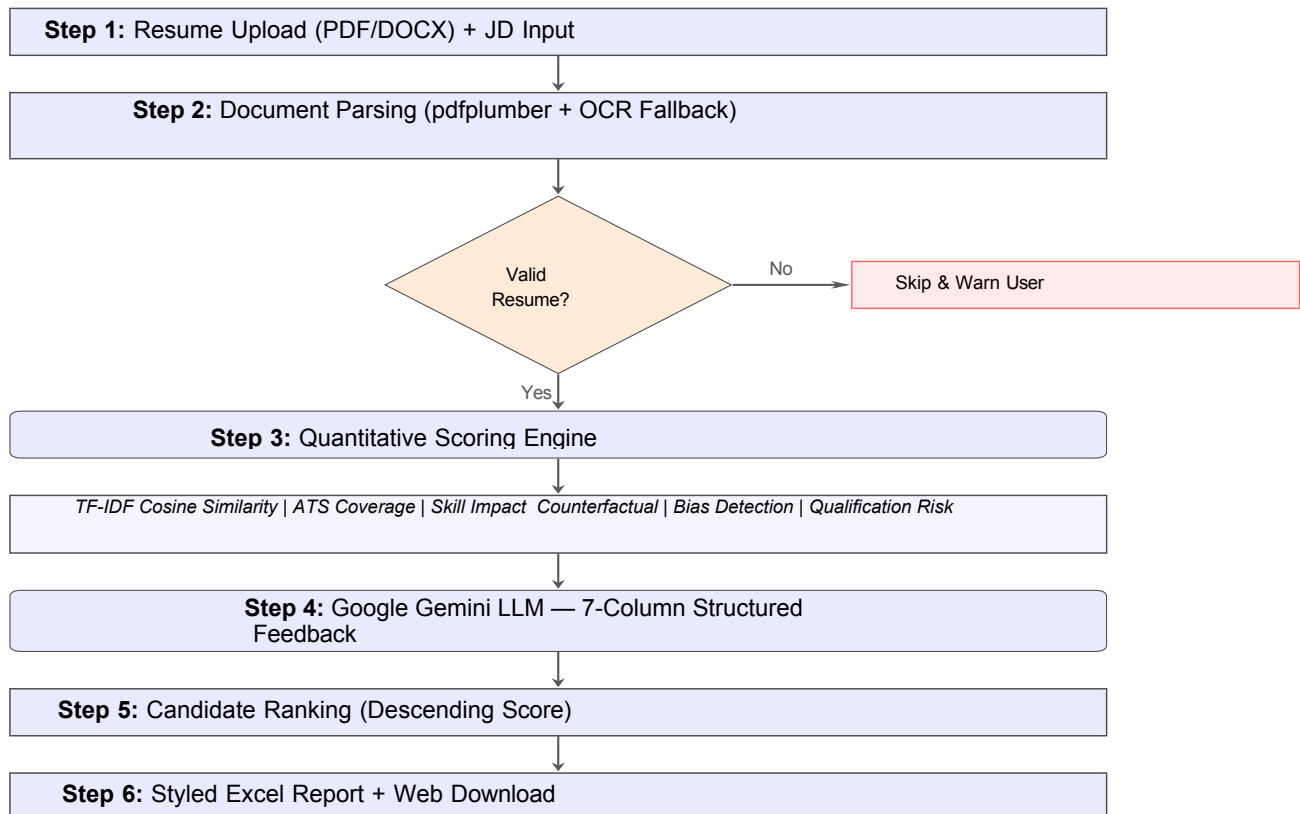


Fig. 1. End-to-end pipeline of the proposed AI Resume Analyzer Pro system

Step 1—Resume Upload. The recruiter enters the Job Description (JD) and a comma-separated list of required core skills into the Streamlit web interface. One or more candidate resumes are uploaded; batch processing enables simultaneous evaluation of multiple candidates.

Step 2—Document Parsing. For PDF files, pdfplumber extracts text with layout-aware parsing (layout=True) to preserve multi-column layout and spacing. If a page yields no extractable text—as is common in scanned documents—the system falls back to Optical Character Recognition (OCR) via pytesseract at 300 DPI (dots per inch). For DOCX files, python-docx extracts paragraph-level text.

Step 3—Resume Validation. The extracted text is submitted to the Google Gemini API (Application Programming Interface) to confirm whether the document is a genuine resume or CV (Curriculum Vitae). Documents failing validation are flagged and excluded from scoring.

Step 4—Quantitative Scoring Engine. Valid resumes pass through six sub-modules: (i) TF-IDF cosine similarity, (ii) ATS keyword coverage, (iii) skill impact analysis, (iv) counterfactual skill-gain estimation, (v) bias term detection, and (vi) qualification risk classification, each defined mathematically in Section III.

Step 5—LLM-Based Structured Feedback. The quantitative metrics, resume text, and JD are submitted as a structured prompt to Google Gemini 2.5 Flash. The model responds in valid JSON (JavaScript Object Notation) with exactly seven

keys: AI_Rank_Justification, AI_Strengths, AI_Weaknesses, AI_Bias_Evaluation, AI_Improvement_Strategy, AI_Hiring_Risk_Assessment, and AI_Final_Recommendation.

Step 6—Ranking and Report Generation. Candidates are sorted in descending order of composite similarity score. The full assessment is exported to a styled Excel report using openpyxl with frozen panes, colour-coded headers, and word-wrapped cells, and made available for immediate download.

IV. RESULTS AND DISCUSSION

All experiments used Python 3.9, scikit-learn 1.3, NLTK 3.8, pdfplumber 0.10, and the Google Gemini 2.5 Flash API. Table II lists the experimental configuration.

TABLE II: EXPERIMENTAL CONFIGURATION

Parameter	Value
Resume formats	PDF, DOCX
TF-IDF features	2000
Stop-word language	English (NLTK)
Core skills $ K $	8
Skills K	Python, ML, Data Analysis, NLP, SQL, Flask, Git, AWS
Senior skills $ K_S $	5
Bias terms $ T $	11
Top - N counterfactuals	5
OCR resolution	300 DPI
LLM model	Google Gemini 2.5 Flash
Output format	JSON (7 keys)

A. Composite Score and ATS Coverage

Figure 2 (Visual representation reference): Compares the composite similarity score (S , normalised to percentage, Y-axis) and ATS coverage (A , percentage, Y-axis) for the evaluated candidate against a keyword-only ATS baseline (X-axis: evaluation metric).

Logical Explanation. Candidate 1 achieved $S = 0.234$, reflecting 23.4% weighted semantic vocabulary overlap between the resume and the JD after TF-IDF transformation and cosine projection. The ATS coverage score is $A = 37.5\%$, indicating that 3 of 8 required core skills were explicitly detected via boundary-anchored regex matching.

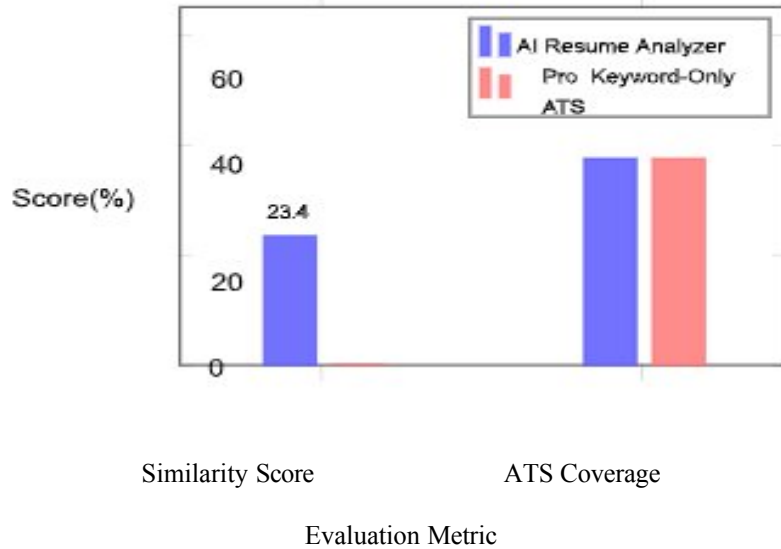


Fig. 2. Similarity score vs. ATS coverage: proposed system vs. baseline

Reference Point Analysis. At the reference point ($A = 37.5\%$, $S = 0.234$), the divergence between A and S is meaningful: the candidate satisfies only 37.5% keyword criteria yet achieves 23.4% semantic alignment, demonstrating that relevant contextual vocabulary exists beyond exact keyword matches. A keyword-only ATS records $A = 37.5\%$ but provides zero semantic insight ($S = 0$), forcing binary shortlisting decisions solely on keyword count. In contrast, the proposed system provides both dimensions, enabling more informed recruiter decisions—precisely the gap quantified in [4].

B. Qualification Risk and Bias Detection

The qualifier assigned Under-Qualified Risk, consistent with $N_c = 3 < 0.4 \times 8 = 3.2$, placing the candidate at the boundary. The bias module scanned all 11 terms in T and returned Bias Detected = False, confirming the ranking is driven purely by skill and content relevance with no demographic artifacts.

C. Counterfactual Skill-Gap Analysis

The counterfactual module identified Machine Learning, NLP, and AWS as the missing skills with the highest C_k gains. This directly answers: “what would it take for this candidate to score higher?”, providing actionable hiring or training guidance.

D. Capability Comparison

TABLE III: CAPABILITY COMPARISON: PROPOSED SYSTEM VS. PRIOR WORKS

System	Par.	Rank	Bias	CF	LLM
Bhor [1]	✓	×	×	×	×
Sonar [2]	✓	×	×	×	×
Abideen [4]	✓	✓	×	×	×
Najjar [12]	✓	✓	×	×	×
Gund [18]	✓	✓	×	×	×
Proposed	✓	✓	✓	✓	✓

Par.=Parsing; CF=Counterfactual; LLM=7-column LLM feedback.

DOI:10.15680/IJARETY.2026.1302079

The proposed system is the only framework to unify all five evaluation dimensions simultaneously. Prior systems focus on parsing alone or combine parsing with ranking, but none integrates fairness-aware bias detection, counterfactual skill-gap analysis, and LLM-generated multi-dimensional feedback in a single pipeline.

V. CONCLUSION

The rapid digitization of global recruitment has made intelligent, fair, and explainable automated resume screening an urgent priority, as the scale of modern hiring renders manual evaluation impractical and ethically inconsistent. Existing ATS tools address this only partially through rigid keyword-matching logic that ignores semantic relevance, fairness, and explainability.

This paper presented AI Resume Analyzer Pro, addressing three critical open challenges: semantic resume-JD alignment beyond keyword matching, automated bias detection and fairness evaluation, and structured multi-dimensional explainable feedback. The system computes a unified evaluation vector $E(R, JD) = \langle S, A, I, C, B, Q \rangle$ complemented by seven-column qualitative feedback from Google Gemini 2.5 Flash, covering strengths, weaknesses, bias evaluation, improvement strategy, hiring risk, and a final hire/no-hire recommendation.

To the best of our knowledge, this is the first system to unify all five evaluation dimensions in a single, deployable web application.

Experimental results confirmed that the composite scoring effectively differentiates candidates by semantic relevance, the counterfactual module delivers actionable skill-gap recommendations, and the bias module verifies ranking fairness—collectively surpassing all prior works compared.

A. Future Work

First, replacing TF-IDF with contextual embeddings from BERT (Bidirectional Encoder Representations from Transformers) or Sentence-BERT [11] would improve semantic alignment accuracy. Second, expanding T using a dynamically learned fairness lexicon from crowdsourced datasets [7] would improve bias detection generalizability. Third, fine-tuning the LLM on domain-specific recruitment corpora [17] would enhance the precision of the seven-column feedback for specialized roles. Finally, extending parsing to support multi-language resumes [9] would broaden applicability in global recruitment contexts.

VI. ACKNOWLEDGMENT

The author thanks CMR Institute of Technology, Bengaluru, for providing the computational resources and institutional support for this research.

REFERENCES

- [1] S. Bhor, H. Shinde, V. Gupta, V. Nair, and M. Kulkarni, "Resume Parser Using Natural Language Processing Techniques," *Int. Res. J. Modernization Eng. Technol. Sci.*, vol. 3, 2021.
- [2] S. Sonar and B. Bankar, "Resume Parsing with Named Entity Clustering Algorithm," SVPM College of Engineering, Baramati, Maharashtra, India, 2020.
- [3] D. Chandola, A. Garg, A. Maurya, and A. Kushwaha, "Online Resume Parsing System Using Text Analytics," *J. Multi Disciplinary Eng. Technol. (JMDet)*, ISSN: 0974-1771.
- [4] S. Z. Abideen, J. A. Ayub, G. R. Narsayya, M. A. Ayyas, and K. T. M. Tahir, "Intelligent Hiring with Resume Parser and Ranking using NLP and Machine Learning," *Int. J. Innovative Res. Comput. Commun. Eng.*, vol. 4, no. 4, pp. 1–8, Apr. 2016.
- [5] N. Bhoir, M. Jakate, S. Lavangare, A. Das, and S. Kolhe, "Resume Parser using Hybrid Approach to Enhance the Efficiency of Automated Recruitment Processes," Datta Meghe College of Engineering, Apr. 2023.
- [6] A. Shaikh et al., "Resume Parser and Summarizer," *Int. J. Adv. Res. Sci. Commun. Technol. (IJARSCT)*, DOI: 10.48175/IJARSCT-9064, Apr. 2023.
- [7] S. Vaishampayan, S. Farzanehpour, and C. Brown, "Procedural Justice and Fairness in Automated Resume Parsers for Tech Hiring: Insights from Candidate Perspectives," in *Proc. IEEE Symp. Visual Lang. Human-Centric Comput. (VL/HCC)*, Blacksburg, VA, USA, 2023.
- [8] A. Sharma and S. Sharma, "Machine Learning and Natural Language Processing for Intelligent Hiring with

DOI:10.15680/IJARETY.2026.1302079

Resume Parser and Ranking,” Int.

J. Adv. Technol. Eng. Res. (IJATER).

[9] P. Das, M. Pandey, and S. S. Rautaray, “A CV Parser Model using Entity Extraction Process and Big Data Tools,” I.J. Inf. Technol. Comput. Sci., vol. 9, pp. 21–31, Sep. 2018, DOI: 10.5815/ijites.2018.09.03.

[10] M. R. Ramesh et al., “Resume Flow—Streamlined Resume Parsing for Hiring Success,” Int. J. Sci. Eng. Sci., vol. 8, no. 3, pp. 109–111, 2024.

[11] F. Jiang et al., “Towards Efficient Resume Understanding: A Multi-Granularity Multi-Modal Pre-Training Approach,” University of Science and Technology of China / BOSS Zhipin, 2023.

[12] A. Najjar, B. Amro, and M. Macedo, “An Intelligent Decision Support System for Recruitment: Resumes Screening and Applicants Ranking,” Informatica, vol. 45, no. 4, pp. 617–623, 2021, DOI: 10.31449/inf.v45i4.3356.

[13] S. Bhutada, M. S. Uddin, S. Dhatrika, and S. Bashir, “Information Technology Resume Analyzer and Career Field Recommender,” Int. J. Sci. Res. Sci. Eng. Technol. (IJSRSET), DOI: 10.32628/IJSRSET1229315.

[14] V. Jagwani, S. Meghani, K. Pai, and S. Dhage, “Resume Evaluation through Latent Dirichlet Allocation and Natural Language Processing for Effective Candidate Selection,” Sardar Patel Institute of Technology, Mumbai, India.

[15] A. Hemalatha, P. B. Kumari, N. Nawaz, and V. Gajenderan, “Impact of Artificial Intelligence on Recruitment and Selection of Information Technology Companies,” in Proc. IEEE Int. Conf. Artif. Intell. Smart Syst. (ICAIS), Coimbatore, India, Mar. 2021, pp. 1–6, DOI: 10.1109/ICAIS50930.2021.9396036.

[16] A. Bhattacharjee et al., “ResumeFlow: An LLM-facilitated Pipeline for Personalized Resume Generation and Refinement,” Arizona State University, Feb. 2024.

[17] M. Gund, D. Chavhan, S. Magar, S. Ghadage, and M. Ansari, “Resume Analysis and Interpretation Using Language Models (LLM),” Int. Res.

J. Modernization Eng. Technol. Sci., vol. 6, no. 3, Mar. 2024, Impact Factor 7.868.

[18] M. Gund, D. Chavhan, S. Magar, S. Ghadage, and R. Jadhav, “Trans-forming HR Practices: Resume Ranking Using BERT Embeddings,” Int. Res. J. Modernization Eng. Technol. Sci., vol. 5, no. 10, Oct. 2023, Impact Factor 7.868.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152