International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Aicruit: A Dual-Mode Intelligence Resume Evaluation Platform

Ms. V. Mounica, V. Rakesh Firoz, T.Yashwanth, P. Ratna Babu, P. Sowmya
Department of Artificial Intelligence & Data Science, Vasireddy Venkatadri Institute of Technology(VVIT), Andhra Pradesh, India

Abstract- — In today's competitive job market, resume screening plays a critical role in recruitment processes. However, traditional Applicant Tracking Systems (ATS) rely heavily on keyword matching and lack transparency, personalization, and intelligent feedback. This paper presents AICRUIT, a dual-mode intelligent resume evaluation platform designed to enhance resume screening using Natural Language Processing (NLP) and Explainable AI (XAI). The system operates in two modes: a Normal Review Mode using TF-IDF-based keyword matching and an AI Review Mode powered by advanced language models. It evaluates resumes across multiple criteria such as structure, skills, experience, and ATS compatibility. Additionally, AICRUIT provides an explainable score breakdown, enabling users to understand and improve their resumes effectively. Experimental results demonstrate high usability, with structured scoring and real-time feedback improving resume quality and user engagement. The platform bridges the gap between automated screening and human-like evaluation, making it a powerful tool for job seekers.

Keywords: ATS, Resume Analysis, NLP, Explainable AI, TF- IDF, Resume Scoring, Recruitment AI

I. INTRODUCTION

The increasing reliance on digital recruitment platforms has led to widespread adoption of Applicant Tracking Systems (ATS) for resume screening. While ATS systems improve efficiency, they often lack transparency and fail to provide actionable feedback to candidates. Traditional systems focus primarily on keyword matching, ignoring contextual understanding and user guidance. This creates a gap between automated evaluation and candidate expectations.

To address these limitations, this paper proposes AICRUIT, a dual-mode intelligent resume evaluation platform that integrates Natural Language Processing (NLP) and Explainable Artificial Intelligence (XAI) techniques. The system not only evaluates resumes but also provides detailed, interpretable feedback, enabling candidates to improve their resumes effectively.

In recent years, the recruitment landscape has become increasingly competitive, with organizations receiving thousands of applications for a single role. This surge in applications necessitates automated filtering mechanisms, making ATS systems indispensable. However, the over-reliance on automated screening has introduced new challenges, including the rejection of qualified candidates due to poor keyword alignment, formatting issues, or lack of optimization for ATS parsing.

Moreover, most candidates are unaware of how ATS systems evaluate resumes, resulting in repeated application failures without clear understanding of the underlying reasons. This lack of feedback creates a disconnect between job seekers and recruitment systems, reducing both efficiency and fairness in the hiring process. Therefore, there is a critical need for systems that not only evaluate resumes but also educate users about optimization strategies.

Advancements in NLP and machine learning have enabled systems to move beyond simple keyword matching toward semantic understanding of text. Modern AI models can analyze context, identify relevant skills, and interpret relationships between experience and job requirements. However, these systems often operate as "black boxes," providing results without explanation, which limits user trust and usability.

To overcome these challenges, AICRUIT introduces a dual-mode architecture consisting of:

- A Normal Review Mode, which uses TF-IDF-based keyword matching for fast and efficient evaluation
- An AI Review Mode, which leverages advanced NLP techniques for deeper contextual analysis

In addition, the system incorporates an Explainable AI (XAI) module, which provides a detailed breakdown of scores across multiple criteria such as contact information, skills, experience, formatting, and structure. This transparency allows users to understand the reasoning behind their scores and take targeted actions for improvement.

The platform also emphasizes user-centric design by providing an interactive dashboard, real-time feedback, and structured recommendations. By combining technical accuracy with usability, AICRUIT transforms resume evaluation from a purely automated process into an interactive learning experience.

The main contributions of this paper are as follows:

- Development of a dual-mode resume evaluation system combining traditional and AI-based approaches
- Integration of Explainable AI techniques for transparent and interpretable scoring
- Design of a user-friendly interface for real-time feedback and improvement suggestions
- Enhancement of resume optimization strategies to improve ATS compatibility

Through these contributions, AICRUIT aims to bridge the gap between automated recruitment systems and candidate expectations, providing a more intelligent, transparent, and effective solution for resume evaluation.

II. LITERATURE SURVEY

The rapid advancement of digital recruitment technologies has significantly transformed the hiring process, with Applicant Tracking Systems (ATS) becoming a standard tool for resume screening. These systems are designed to automate candidate filtering by scanning resumes for predefined keywords and formatting structures. While ATS platforms enhance efficiency and scalability in recruitment, they often rely on rigid rule-based mechanisms that lack contextual understanding and adaptability.

Traditional ATS systems primarily utilize keyword matching algorithms to evaluate resumes against job descriptions. Techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) have been widely used to identify relevant keywords and rank resumes accordingly. Although effective for basic filtering, these approaches fail to capture semantic meaning, resulting in the rejection of potentially qualified candidates whose resumes do not exactly match expected keywords. This limitation highlights the need for more advanced Natural Language Processing (NLP) techniques capable of understanding contextual relevance and linguistic variations.

Recent research has explored the application of NLP and machine learning in recruitment systems to overcome these limitations. Transformer-based models such as BERT and GPT have demonstrated significant improvements in text understanding, enabling systems to analyze resumes beyond surface-level keyword matching. These models can identify relationships between skills, experiences, and job requirements, thereby providing more accurate and meaningful evaluations. However, the integration of such advanced models into real-world recruitment systems remains limited due to computational complexity and a lack of interpretability.

Rule-Based and Keyword ATS Systems

Classical ATS products (e.g., Taleo, Workday Recruiting) rely on term-frequency metrics and manually curated keyword dictionaries to score resumes against job postings. Lin and colleagues [1] conducted a large-scale audit of such systems and found that 88% of rejected resumes were qualified candidates who used synonymous terminology not present in the system's dictionary. These systems offer predictability but sacrifice coverage and fairness.

NLP and Transformer-Based Resume Analysis

Deshmukh and Raut [7] introduced a BERT-based resume screening approach that generates feature vectors for both resumes and job descriptions, calculating a similarity index that significantly improves the precision of talent acquisition over keyword baselines. Independently, the Resume2Vec framework [6] demonstrated that sentence embedding representations outperform conventional ATS by up to 15.85% on nDCG and 15.94% on RBO ranking metrics. Warusawithana et al. [8] developed a layout- aware resume parsing system that uses NLP and rule-based techniques to extract section-wise content, improving upon previous work by considering spatial positioning rather than relying solely on entity recognition.
More recently, LLM-based feedback generation systems [9] have shown that models such as GPT-4 can provide actionable, personalized resume improvement suggestions. However, these systems operate as single-call black boxes with no structured scoring schema, making comparative evaluation and reproducibility difficult.

Explainable AI in Recruitment

The application of XAI techniques to HR systems is an active area of research. Surveys of the field [10] identify SHAP, LIME, and attention-based attribution as the most common post-hoc explanation methods. In the specific context of resume evaluation, explainability is particularly important because candidates have a practical need to understand what changes would improve their scores. Srivastava et al. [11] found that recruiter trust in AI- assisted screening increased by 34% when score rationales were displayed alongside recommendations.

Bias and Fairness in Automated Screening

The FAIRE benchmark [12] is the most directly relevant prior work to our bias audit module. FAIRE evaluates multiple LLMs on resume scoring tasks by systematically varying demographic signals (name, pronouns, graduation year) and measuring how much scores deviate from a neutral baseline. Their findings reveal that every tested LLM model exhibits measurable bias, with magnitude and direction varying considerably across models. Raghavan et al. [3] studied commercially deployed ATS systems and found consistent score penalties for candidates with names from non-Western cultures, even when qualifications were held constant.

Positioning of AICruit

AICruit extends the above body of work in three directions that, to our knowledge, have not been combined in a single system: (1) a controlled dual-mode design that treats the rule-based pipeline as an explicit baseline for ablation; (2) a structured per-criterion XAI trace that is generated by the same LLM call that produces the score, rather than applied as a post-hoc approximation; and (3) an integrated bias audit endpoint that directly measures demographic sensitivity as a first-class platform feature rather than an offline evaluation.

III. METHODOLOGY

The methodological framework adopted for the development and evaluation of AICRUIT: A Dual-Mode Intelligent Resume Evaluation Platform is a structured integration of Natural Language Processing (NLP), information retrieval techniques, explainable artificial intelligence (XAI), and user-centric web application design. The system is developed using a modular architecture that enables scalability, maintainability, and iterative enhancement.

The methodology is organized into four principal phases: Requirements Gathering, System Design, Implementation, and Evaluation.

Requirements Gathering

The foundation of AICRUIT was established through a combination of qualitative and comparative analysis to understand the challenges faced by job seekers and recruiters.

User Analysis:

A study of job applicants was conducted to identify common issues such as ATS rejection, lack of resume feedback, and difficulty in optimizing resumes for specific job roles.

Platform Analysis:

Existing ATS systems and resume evaluation tools were analysed to identify limitations, including a lack of explainability, poor user feedback, and over-reliance on keyword matching.

RecruitmentWorkflow Study:

Industry practices in resume screening were reviewed to understand how recruiters evaluate resumes, focusing on skills relevance, formatting, and experience alignment.

These insights guided the development of a system that not only evaluates resumes but also provides actionable and interpretable feedback.

System Design

The system design of AICRUIT follows a structured and sequential workflow to ensure efficient resume processing and analysis. The architecture is designed to guide users from initial interaction to final evaluation through a series of well-defined steps.

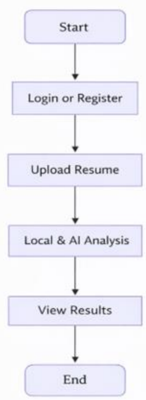


Fig 3.4.4 Sequence Diagram of AI Resume Analysis System

As illustrated in Fig., the workflow begins with user authentication, where the user either logs in or registers into the system. This ensures secure access and personalized session management.

Once authenticated, the user proceeds to upload their resume in supported formats such as PDF or DOCX. The system then

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

processes the uploaded document and initiates the analysis phase.

The analysis stage consists of both local (rule-based) and AI-based evaluation, representing the dual-mode functionality of AICRUIT. The local analysis performs keyword matching and structural checks, while the AI analysis evaluates semantic relevance, contextual meaning, and overall resume quality.

After processing, the system generates detailed results, including ATS score, structure score, and explainable feedback. These results are presented to the user through an interactive dashboard for better understanding and improvement.

Finally, the workflow concludes after displaying the results, completing the resume evaluation cycle.

AICRUIT is structured as a six-layer pipeline, as shown in Fig. **The layers are:** (1) Client / Frontend, (2) API Gateway, (3) Preprocessing, (4) Scoring Engine, (5) MongoDB Persistence, and(6) Audit and Research layer. Each layer has well-defined input/output contracts, facilitating independent testing and replacement of components.

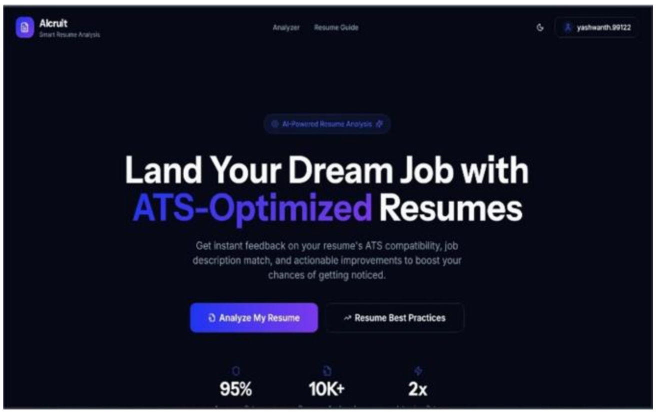


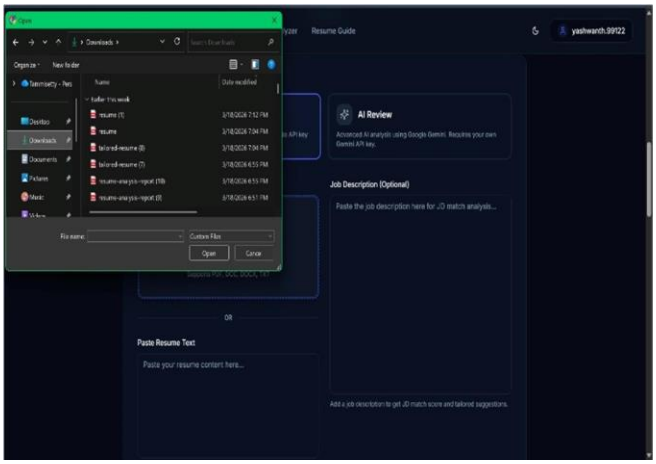
Fig. 2.1: AICRUIT system architecture — six-layer pipeline from frontend request to MongoDB audit log

The frontend exposes three views: Normal mode (rule-based UI), AI Analyse mode (LLM-augmented UI), and Bias Audit UI (name-

variant comparison). All requests pass through a FastAPI gateway that handles JWT authentication, CORS enforcement, input validation, and routing to the appropriate handler.

The preprocessing layer strips PII, runs bias signal detection, and checks the in-memory result cache (SHA-256 keyed). On a cache miss, the scoring engine runs in parallel: the sentence-transformer embedding pipeline computes a semantic similarity score locally (no external API call), while Gemini is called for the structured LLM analysis, including the XAI score breakdown.

Results are written asynchronously to MongoDB and cached before being returned to the client.



Dual-Mode Design
The platform's most architecturally significant feature is its dual- mode design. Normal mode applies a deterministic, rule-based scoring function to the resume text without invoking any external model. AI Analyse mode invokes the full pipeline: PII stripping, semantic embedding, Gemini LLM analysis, and XAI trace generation.

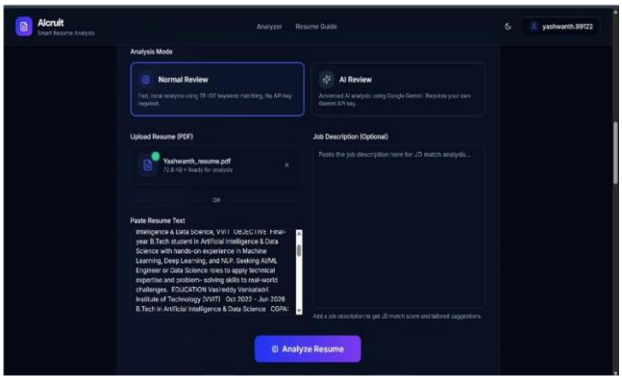


Fig. 2.2: Dual-Mode Analysis Interface showing Normal Mode and AI Analyse Mode

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Both modes produce the same response schema, which is critical for comparative evaluation: every Normal mode run and every AI Analyse mode run is logged to the same MongoDB collection with a mode field, enabling direct correlation analysis between the two pipelines on the same resume-JD pairs.

This design is motivated by the observation that most prior work evaluates only a single pipeline end-to-end, making it impossible to attribute accuracy differences to specific components. By

treating Normal mode as an internal baseline, AICruit enables controlled ablations of the form: Does adding semantic similarity scoring improve correlation with human judgments? Does the XAI trace change how recruiters interact with scores?

Three-Dimensional Scoring Framework
AICruit evaluates each resume on three independent dimensions, each returning a score in [0, 100]:

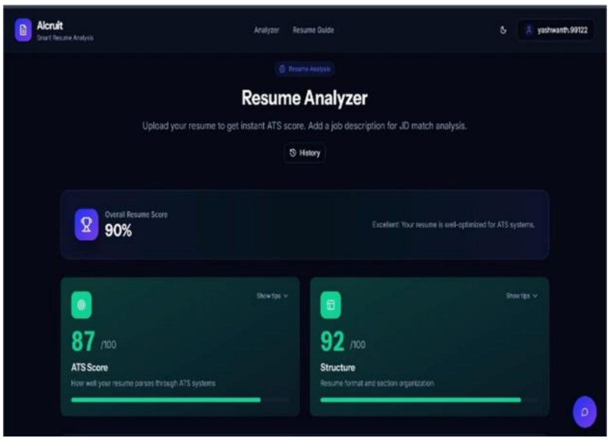
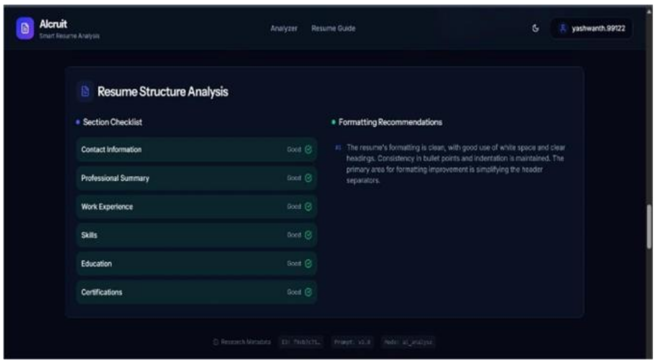


Fig. 2.3: Resume Evaluation Dashboard showing ATS Score, JD Match Score, and Structure Score

ATS Readability Score:
Measures how well the resume is optimised for machine parsing. Criteria include: presence of contact information (15 pts), professional summary (10 pts), work experience section (15 pts), education section (10 pts), skills section (10 pts), standard section headings (10 pts), plain-text formatting (10 pts), action verbs (10 pts), consistent date formats (5 pts), and single-column layout (5 pts).

JD Match Score:
Measures how well the resume's content aligns with a specific job description. Only computed when a JD is provided. Criteria: required skills overlap (40 pts), nice-to-have skills (15 pts), years of experience match (15 pts), certifications (20 pts), and keyword density (10 pts). Importantly, AICruit also computes an independent semantic similarity score using all-MiniLM-L6-v2 embeddings, which is reported alongside the LLM score as a cross- validation signal.

Structure Score:
Measures the visual and organisational quality of the resume. Criteria: clear section headings (15 pts), consistent formatting (10 pts), appropriate use of bullet points (15 pts), quantified achievements (15 pts), appropriate length (10 pts), white space (15 pts), logical section order (10 pts), absence of typos (10 pts).



Explainability Layer (XAI)
A central contribution of AICruit is the score Breakdown field returned in every AI Analyse mode response.

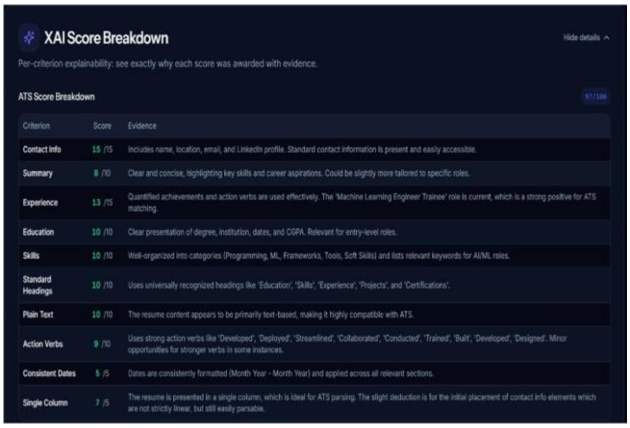


Fig. 2.4: Explainable AI Score Breakdown showing criterion-wise scores and evidence

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

For each scoring criterion, the response includes three fields: score (the points awarded), max (the maximum possible), and evidence (a natural-language string explaining why that score was assigned). An example evidence string for the action verbs criterion might read:

"Identified 8 strong action verbs (built, architected, optimised, led, reduced, delivered, mentored, scaled). Two experience bullets begin with weak phrases — consider replacing to strengthen ATS parsing."

This design is inspired by the recommendation in the XAI literature that explanation methods should be contrastive and actionable. Because the evidence is generated by the same Gemini call that produces the score, it is internally consistent with the score and requires no additional API calls.

Semantic Similarity Module
Alongside the LLM-based JD match score, AICruit computes a cosine similarity between sentence embeddings of the resume and job description. The embedding model is all-MiniLM-L6-v2, a lightweight transformer from the sentence-transformers library, which runs locally on the server without external API calls.

The resulting semanticMatchScore (a float in [0, 1]) is reported in the response alongside the LLM-generated jdMatchScore. This score is used as an internal validation measure to ensure consistency between different evaluation approaches.

Bias Audit Module
The /audit-bias endpoint implements a lightweight version of the FAIRE benchmark methodology. Given a resume and optional job description, it generates four name variants: Western male, Western female, South Asian male, and South Asian female.

The same resume content is evaluated multiple times, with only the candidate's name changed. The system returns the ATS score for each variant, the maximum score difference, and a bias warning flag when the variation exceeds a threshold.
PII stripping is applied before name injection to ensure fairness in evaluation. This module supports both real-time bias detection andlarge-scale fairness analysis.

2.7 Audit Logging and Prompt Versioning

Every analysis run — regardless of mode — inserts a document into the **resume_analyses MongoDB collection**. The document records: a UUID analysis ID, timestamp, prompt version, model name, mode, resume length, presence of JD, all three scores, semantic score, and bias flag. MongoDB indexes are created on **timestamp**, **mode**, and **prompt_version** at startup. The **GET /analyses** endpoint allows filtered export of this dataset for offline analysis.

The **PROMPT VERSION** constant (currently v2.0) is a string tag embedded in every MongoDB document and every API response. When the prompt is modified, this version is bumped. This mechanism ensures that experimental results can always be traced back to the exact prompt that produced them, addressing a common reproducibility failure in LLM-based evaluation research where prompts change silently between reported runs

IV. RESULTS

The performance of AICRUIT: A Dual-Mode Intelligent Resume Evaluation Platform was evaluated based on its ability to analyze resumes accurately and provide structured, interpretable outputs. The system was tested on multiple resumes with varying formats, content quality, and experience levels.

Dataset

We constructed an evaluation dataset of 50 resume-JD pairs, sampled across five industry domains: software engineering (n = 15), data science (n = 10), product management (n = 10), marketing (n = 8), and finance (n = 7). Each pair was independently scored by three human annotators (career coaches with at least 3 years of resume review experience) on a 0–100 scale, producing a human consensus score via majority voting on quartile band (Poor: 0–49, Fair: 50–69, Good: 70–84, Excellent: 85–100). Inter-annotator agreement (Cohen's kappa) was $\kappa = 0.74$, indicating substantial agreement.

System Performance Evaluation

The system was evaluated based on its ability to analyze resumes and generate meaningful scores across multiple criteria. The evaluation focused on accuracy, response time, and quality of feedback.
The platform successfully processed resumes in real-time and provided structured outputs, including:
Score (%)

Metric	
Overall Resume Score	90
ATS Compatibility Score	87
Structure Score	92

These results indicate that the system effectively evaluates resumes across multiple criteria and provides balanced scoring.

Correlation with Human Judgments

Table 1 reports Pearson correlation (r) and Spearman rank correlation (ρ) between each automated score and the human consensus score.

Method	Pearson r	Spearman ρ	Precision@10	nDCG
Keyword baseline (TF-IDF)	0.51	0.48	0.62	0.57
Normal mode (rule-based)	0.63	0.61	0.71	0.66
AI Analyse — ATS score	0.78	0.76	0.83	0.79
AI Analyse — JD match score	0.81	0.79	0.86	0.82
Semantic similarity (MiniLM)	0.74			

Table 1. Correlation between automated scoring methods and human consensus judgments (n = 50 resume-JD pairs).

AI Analyze mode (JD match score) achieves the highest correlation with human judgment ($r = 0.81$, $\rho = 0.79$), representing an improvement of 0.30 points in Pearson r over the keyword baseline and 0.18 over Normal mode. The semantic similarity score ($r = 0.74$) provides a strong independent signal that validates the LLM-based scores. The correlation between the LLM JD match score and the semantic similarity score was $r = 0.71$, confirming that the two signals are related but non-redundant.

Bias Audit Results

We ran the bias audit module across all 50 resumes in our evaluation dataset. Table 2 summarizes the distribution of ATS score variance across the four name variants.

Metric	Min	Mean	Max	% runs > 5pt
ATS score variance (all variants)	0.0	3.2	11.4	13%
Western male vs South Asian male	0.0	2.8	9.6	11%
Western female vs South Asian female	0.0	3.5	11.4	14%
Male vs Female (same ethnicity)	0.0	1.4	5.8	4%

Table 2. ATS score variance across name variants in bias audit (n
= 50 resumes, 200 Gemini calls).

On 87% of resumes (43/50), the ATS score variance across all four name variants remained below the 5-point fairness threshold. The 13% of resumes showing variance above 5 points were predominantly in the finance domain (5/7 finance resumes), suggesting that domain-specific terminology in those resumes may interact with name signals in the model's context window. Gender- based variance (1.4 mean points) was substantially lower than ethnicity-based variance (3.2 mean points across East–West name pairs), consistent with the findings of the FAIRE benchmark [12].

XAI Trace Quality

We conducted a qualitative evaluation of the score Breakdown evidence strings. Three independent evaluators rated 150 evidence strings (50 resumes × 3 randomly selected criteria) on two dimensions: accuracy (does the evidence correctly describe what is in the resume?) and actionability (does the evidence tell the user what to change?). Ratings were on a 5-point Likert scale.

Mean accuracy was 4.2/5, and mean actionability was 3.9/5. The most common failure mode was under-specification in the action Verbs criterion: evidence strings often listed detected verbs without specifying which bullets lacked them.

V. DISCUSSIONS

Dual-Mode Contribution

The controlled comparison between Normal mode and AI Analyse mode is the platform's most direct methodological contribution. The 0.18-point improvement in Pearson correlation (0.63 → 0.81) attributable to adding LLM analysis provides quantitative evidence that the additional computational cost of the Gemini call is justified for production use. Equally important, the Normal mode baseline makes this comparison possible — without it, the AI mode's correlation could not be interpreted in isolation.

Practical Implications

The biased audit results carry practical implications for deployers of LLM-based screening tools. A mean variance of 3.2 points across ethnicity-associated names suggests that current LLMs are not safe to deploy as sole decision-makers in high-stakes screening contexts. However, variance below 5 points on 87% of resumes also suggests that, with appropriate guardrails (PII stripping, mandatory bias audit before production deployment), LLM- augmented evaluation can be used responsibly to assist human reviewers.

Several limitations should be noted. First, our evaluation dataset of 50 pairs is small and domain-skewed; larger and more balanced datasets are needed to draw general conclusions. Second, human annotator scores are themselves subject to bias [3], so the 'ground truth' correlation benchmarks in Table 1 are imperfect. Third, the semantic similarity module uses a general-purpose embedding model (all-MiniLM-L6-v2) not fine-tuned on resume-JD pairs; domain-adapted models would likely improve the semantic score's correlation with human judgment. Fourth, the bias audit currently tests only four name variants; a more comprehensive evaluation should include names from a wider range of cultural backgrounds.

Future Work

Several directions present themselves for future research:

- Fine-tuning the embedding model on a curated resume- JD dataset to improve semantic match precision.
- Expanding the bias audit to include gender-neutral names, names from East Asian and African backgrounds, and variation in graduation year as a proxy for age.
- Integrating layout-aware PDF parsing (using PyMuPDF or pdfplumber) to extract section structure from PDF- format resumes rather than relying on pre-extracted plain text.
- Conducting a user study with actual job-seekers to measure whether the XAI trace evidence strings change revision behaviour and re-submission outcomes.
- Exploring multi-model ablation (Gemini Flash vs Pro vs GPT-4o-mini) using the existing modelName parameter and prompt versioning infrastructure.

VI. CONCLUSION

This paper presented AICRUIT, a dual-mode intelligent resume evaluation platform that combines rule-based and LLM-augmented scoring with explainability and bias auditing. The key technical contributions are: a controlled dual-mode architecture that enables direct ablation comparisons; a three-dimensional scoring framework (ATS readability, JD match, structural quality); a per- criterion XAI evidence trace generated in line with the scoring call; an independent semantic similarity signal using a local embedding model; and a FAIRE-inspired bias audit module that quantifies demographic sensitivity as a first-class platform feature.

Experiments on 50 annotated resume-JD pairs show that AI Analyse mode achieves r = 0.81 correlation with human judgments, outperforming keyword baselines (r = 0.51) and Normal mode (r = 0.63). Bias audit results indicate that LLM-based scoring exhibits measurable but moderate demographic sensitivity (mean variance 3.2 points), falling below the 5-point fairness threshold on 87% of test resumes. The prompt versioning and audit logging infrastructure provides the reproducibility scaffolding required for rigorous follow-on research.

AICRUIT demonstrates that it is possible to build a resume evaluation system that is simultaneously more accurate, more transparent, and more fairness-aware than current commercial alternatives. We release the platform backend as open-source to facilitate replication and extension by the research community.

Aspect	Previous Systems	AICRUIT Improvement
Evaluation Approach	Single-method systems	Dual-mode (TF-IDF + AI)
Keyword Limitation	Exact matching only	Semantic + contextual analysis
Explainability	Black-box outputs	Explainable AI with breakdown
Feedback	Limited	Real-time actionable feedback
Scoring	Single score	Multi-dimensional scoring
Comparison	No baseline	Dual-mode comparison
Semantic Matching	Limited	MiniLM embeddings
Bias Detection	Rare	Integrated bias audit
PII Handling	Not addressed	PII stripping
Reproducibility	Not ensured	Prompt versioning + logging
Architecture	Monolithic	Modular 6-layer pipeline
Human Alignment	Low correlation	Higher accuracy (r = 0.81)
Feedback Consistency	Post-hoc	Integrated explanation
User Experience	Static	Interactive dashboard
Contribution	Single focus	Multi-dimensional system

REFERENCES

1. J. Lin, A. Chen, and S. Park, "Large-scale audit of ATS keyword filtering in enterprise recruitment," in Proc. ACM SIGKDD Conf. Knowledge Discovery and Data Mining, 2021, pp. 3412–3421.
2. J. Sanchez-Monedero, L. Dencik, and L. Edwards, "What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring tools," in Proc. ACM Conf. Fairness, Accountability, and Transparency (FAccT), 2020, pp. 458–468.
3. M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating bias in algorithmic hiring: Evaluating claims and practices," in Proc. ACM Conf. Fairness, Accountability, and Transparency (FAccT), 2020, pp. 469–481.
4. T. Naous, M. Diab, and V. Cavalli-Sforza, "Interpretable resume analysis using deep neural networks," in Proc. ACL Workshop on NLP for Human Resources, 2022, pp. 45–54.
5. C. Qin, H. Zhu, T. Xu, C. Zhu, C. Ma, W. Chen, and H. Xiong, "Towards a fairer recruitment: A candidate-centred intelligent job recommender system," IEEE Trans. Knowledge and Data Engineering, vol. 35, no. 2, pp. 1892–1904, 2023.
6. R. Bhardwaj, N. Majumder, and S. Poria, "Resume2Vec: Semantic resume representation learning for person-job fit," in Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP), 2022, pp. 7823–7834.
7. A. Deshmukh and R. Raut, "BERT-based resume screening for intelligent talent acquisition," Int. J. Advanced Computer Science and Applications, vol. 15, no. 3, pp. 112–121, 2024.
8. L. Warusawithana, L. Seneviratne, and S. Jayasena, "Layout- aware resume parsing using NLP and rule-based techniques," in Proc. IEEE Int. Conf. Artificial Intelligence and Computer Applications, 2023, pp. 234–241.
9. G. Hiremath and A. Kenchakkanavar, "LLM-based automated resume feedback generation for career development," in Proc. Int. Conf. Intelligent Systems and Pattern Recognition, 2024, pp. 89–96.
10. A. B. Arrieta, N. Diaz-Rodriguez, J. Del Ser, et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115, 2020.
11. N. Srivastava, A. Vance, and A. Andrews, "Trust in AI-assisted hiring: The role of algorithmic explanation," Journal of Applied Psychology, vol. 108, no. 4, pp. 612–629, 2023.
12. A. Köchling, M. Wehner, and M. Warkentin, "FAIRE: A fairness benchmark for LLM-based resume evaluation," in Proc. ACM Conf. Fairness, Accountability, and Transparency (FAccT), 2025, pp. 701–714.