ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct964

Pro-Align: An AI-Powered Resume–Job Description Matching, Optimization and Interview Preparation Framework

Vanjula Tejaswini[1]; Dr. Anjana Devi Bondalapati[2]; Dr. Nagesh Vadaparthi[3]

[1]Under the Esteemed Guidance of
[1;2;3]Department of Information Technology, M V G R College of Engineering, Vizianagarm, India

Publication Date: 2026/10/24

Abstract: The increasing use of automated recruitment systems creates a significant challenge for candidates whose resumes may not sufficiently align with target job descriptions. This paper presents PRO-ALIGN, an AI-powered web-based framework for resume–job description matching, Applicant Tracking System (ATS) compatibility analysis, resume optimization, and personalized interview preparation. The system parses resumes and job descriptions, identifies matched and missing skills, and calculates an interpretable compatibility score using weighted components for skills, experience, projects, education, semantic similarity, and formatting. Semantic similarity is supported through vector embeddings and cosine similarity, with offline and term-frequency fallbacks for robust operation. Generative AI capabilities are used for resume rewriting, improvement recommendations, interview-question generation, and answer evaluation. The platform is implemented with React/Vite, Flask, SQLAlchemy, JWT authentication, SQLite/PostgreSQL, and ReportLab. The proposed framework extends conventional resume matching by connecting analysis with actionable resume improvement and interview preparation in a unified workflow.

Keywords: *Artificial Intelligence, Natural Language Processing, Resume Matching, Job Description, ATS, Semantic Similarity, Cosine Similarity, Generative AI, Resume Optimization, Interview Preparation.*

How to Cite: Vanjula Tejaswini; Dr. Anjana Devi Bondalapati; Dr. Nagesh Vadaparthi (2026) Pro-Align: An AI-Powered Resume–Job Description Matching, Optimization and Interview Preparation Framework. *International Journal of Innovative Science and Research Technology*, 11(10), 1614-1617. https://doi.org/10.38124/ijisrt/26oct964

I. INTRODUCTION

Recruitment has increasingly moved toward digital platforms where organizations receive large numbers of applications for individual job openings. Applicant Tracking Systems (ATS) and automated screening mechanisms are used to identify potentially relevant candidates by examining skills, experience, education, keywords, and other job-related information. The underlying computational problem can be formulated as a matching task between a candidate resume and a target job description.

Traditional keyword-based matching is simple and interpretable, but exact matching may fail when semantically related expressions use different terminology. For example, a resume describing “developed Flask APIs” can be relevant to a job requiring “RESTful backend development” even when the exact phrases differ. Research has therefore explored semantic approaches to resume–job matching. Maheshwary and Misra proposed a deep Siamese-network approach for matching semi-structured resumes and job descriptions, demonstrating the usefulness of learned semantic representations [1]. Roy et al. investigated automated resume recommendation using content-based similarity and k-nearest-neighbor techniques [2]. More recent work has continued to investigate semantic similarity and domain-specific weighting for recruitment matching [3].

However, a practical candidate-oriented system should provide more than a similarity score. Candidates need to understand why their resume does not match a position, which skills are missing, how the resume can be improved, and how they can prepare for the associated interview. PRO-ALIGN addresses this broader requirement by integrating resume parsing, job-description analysis, ATS compatibility scoring, semantic similarity, skill-gap analysis, AI-assisted resume optimization, interview preparation, and report generation.

- *Problem Statement*

Candidates frequently submit resumes without sufficiently adapting them to individual job descriptions. This can reduce visibility in automated screening and makes it difficult for candidates to identify the exact areas that require improvement. The problem addressed in this research is how to analyze the explicit and semantic relationship between a candidate resume and a target job description, quantify compatibility, identify skill gaps,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct964

recommend resume improvements, and provide personalized interview preparation within a single platform.

➢ *Objectives*

The objectives of PRO-ALIGN are: (1) automatically extract meaningful information from uploaded resumes; (2) process job descriptions and identify relevant skills and requirements; (3) compare resume content with job requirements using explicit and semantic matching; (4) calculate an interpretable ATS compatibility score; (5) identify missing skills; (6) provide AI-assisted resume improvements; (7) generate role-specific interview questions and evaluate answers; and (8) produce a downloadable analysis report.

II. RELATED WORK

Research in automated recruitment has evolved from keyword filtering toward machine learning and semantic matching. Maheshwary and Misra introduced a deep Siamese network for resume–job matching and showed that learned representations can capture semantic relationships between the two document types [1]. Roy, Chowdhary, and Bhatia proposed an automated resume recommendation approach involving classification and content-based recommendation, including cosine similarity and k-NN [2]. Gugnani and Misra explored implicit skill extraction using document embeddings and its application to job recommendation [4]. Transformer-based language models such as BERT [5] and Sentence-BERT [6] have further enabled contextual text representations. These studies motivate the semantic matching component of PRO-ALIGN.

The research gap is that many approaches focus primarily on ranking, recommendation, or similarity. PRO-ALIGN extends the workflow toward candidate development by combining matching, interpretable scoring, skill-gap identification, resume optimization, interview coaching, and reporting.

➢ *Proposed System*

PRO-ALIGN is organized as a web-based client–server architecture. The React frontend provides authentication, resume upload, job-description input, analysis dashboards, visualization, resume optimization, interview coaching, history, and report access. The Flask backend exposes REST APIs and coordinates parsing, matching, scoring, AI services, persistence, and report generation. SQLAlchemy provides database abstraction and supports SQLite for development and PostgreSQL for production.

The overall workflow is: Resume + Job Description → Parsing → Keyword Matching → Semantic Similarity → ATS Scoring → Skill-Gap Analysis → Recommendations → Resume Optimization → Interview Preparation → PDF Report.

➢ *System Architecture*

The architecture contains five logical layers. The presentation layer is implemented with React, Vite, Tailwind CSS, Recharts, and Lucide React. The API layer is implemented with Flask and modular route services. The intelligence layer includes resume parsing, JD parsing, keyword matching, experience and education analysis, embeddings, ATS scoring, optimization, and interview services. The data layer uses SQLAlchemy with SQLite/PostgreSQL. The reporting layer generates PDF outputs using ReportLab.

JWT-based authentication protects user-specific operations. The system also supports graceful degradation for semantic embedding generation: an external embedding service can be used when configured, an offline SentenceTransformer model can be used when available, and term-frequency cosine similarity can act as a lightweight fallback.

➢ *Resume and Job Description Processing*

The resume parser converts supported PDF, DOCX, and TXT documents into machine-readable text and extracts available profile information. The job-description parser identifies role requirements and relevant keywords. Extracted content is normalized before being supplied to the matching engine. This preprocessing step is important because inconsistent casing, spacing, document formatting, and file-specific representations can otherwise reduce matching quality.

➢ *Keyword Matching*

For a set of job-description keywords $K = \{k1, k2, \ldots, kn\}$, the system determines whether each keyword occurs in the resume text. The skill score is calculated as the proportion of matched keywords.

$$S_{skill} = (N_{matched} / N_{total}) \times 100$$

➢ *Semantic Similarity*

Keyword matching alone cannot fully capture semantic relationships. PRO-ALIGN therefore supports vector-based similarity. Let VR represent the resume embedding and VJ represent the job-description embedding. Cosine similarity measures the angular relationship between the two vectors and provides a normalized measure of semantic alignment. The similarity result is transformed into a percentage for integration into the ATS scoring model. The implementation supports SentenceTransformer-based embeddings and a term-frequency cosine fallback when embedding services are unavailable.

$$CosSim(VR,VJ) = (VR \cdot VJ) / (||VR|| \, ||VJ||)$$

➢ *ATS Scoring Model*

A central component of PRO-ALIGN is an interpretable weighted scoring model. The implemented weights are: Skills Matching 40%, Experience Alignment 20%, Projects 15%, Education 10%, Semantic Similarity 10%, and Formatting 5%. The overall score is computed from normalized component scores.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct964

ATS = 0.40S_skill + 0.20S_exp + 0.15S_proj + 0.10S_edu + 0.10S_sem + 0.05S_format

Table 1 ATS Scoring Model

Component	Weight
Skills Matching	40%
Experience	20%
Projects	15%
Education	10%
Semantic Similarity	10%
Formatting	5%

➢ *Experience, Project and Education Analysis*

The experience analyzer detects common expressions such as "3 years experience," "5+ years," and similar patterns from job descriptions. Candidate experience is compared against the inferred requirement. Project matching examines project descriptions against relevant job keywords, which is particularly useful for students and candidates with limited formal employment history. Education analysis maps common academic levels such as Bachelor, Master, and Doctorate to broad levels and compares them with the job requirement.

➢ *Skill-Gap Analysis*

The system explicitly identifies skills appearing in the job description but not in the resume. For example, if a job requires Python, Flask, Docker, AWS, and SQL while a resume contains Python, Flask, and SQL, Docker and AWS are reported as missing skills. This output provides an actionable explanation of why the candidate may have a lower compatibility score.

➢ *AI-Based Resume Optimization*

The Resume Optimizer can target summary, experience, projects, and skills. Generative AI can improve wording, make statements more action-oriented, and suggest ways to incorporate relevant qualifications. The design principle is that the system should not fabricate credentials. A missing skill should be presented as a recommendation unless the candidate genuinely possesses it. Local fallback logic can provide rule-based improvements when external AI services are unavailable.

➢ *Interview Preparation*

The Interview Coach connects resume analysis to interview readiness. Questions can be generated using the candidate profile, matched skills, missing skills, and target job description. The workflow can include technical, behavioral, role-specific, and HR questions. Candidate answers can then be evaluated using AI-assisted feedback, suggested answer structure, STAR guidance, and coaching tips.

➢ *Security and Data Management*

JWT authentication is used to protect user-specific API operations. A successful login produces a token that is presented with protected requests and validated by the backend. SQLAlchemy models establish relationships among users, resumes, and analysis reports. Report deletion is restricted to resources belonging to the authenticated user. Production deployments should additionally apply HTTPS, secure secret management, controlled data retention, and appropriate privacy policies.

➢ *Algorithm*

Algorithm 1 summarizes the analysis pipeline. First, the system parses the resume and job description. It then extracts keywords, computes skill overlap, identifies experience and education requirements, calculates project alignment, generates embeddings, computes semantic similarity, evaluates formatting heuristics, and combines the normalized components using the weighted ATS formula. Finally, it generates recommendations, interview material, stores the analysis, and optionally creates a PDF report.

- *Algorithm 1: PRO-ALIGN Resume–JD Analysis*

✓ Input: Resume R and Job Description J.
✓ Parse R and J; normalize extracted text.
✓ Extract job keywords K; compute matched and missing skills.
✓ Calculate skill, experience, project, education, semantic, and formatting scores.
✓ Generate embeddings and calculate cosine similarity.
✓ Compute the weighted ATS score.
✓ Generate recommendations and interview questions.
✓ Store the analysis and optionally generate the PDF report.
✓ Return ATS score, score breakdown, skill gaps, recommendations, and interview material.

➢ *Experimental Design*

A rigorous evaluation should use a labeled collection of resumes and corresponding job descriptions covering different technical domains and experience levels. Each pair should be assessed using keyword match, semantic similarity, overall ATS score, missing-skill count, and recommendation quality. Human evaluators or expert recruiters can provide reference compatibility judgments. The system can then be compared with baseline approaches such as TF-IDF cosine similarity and embedding-only similarity.

Recommended evaluation categories:

- Software Engineering (Python/Java/React)
- Data Science and Machine Learning
- Cloud Computing
- AI/NLP
- Web Development

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct964

➢ *Evaluation Metrics*

For classification-oriented skill extraction or relevance decisions, precision, recall, and F1-score can be used. If human evaluators provide reference compatibility scores, mean absolute error can quantify the difference between human judgment and the PRO-ALIGN score. These metrics should be computed on an independently prepared evaluation dataset; application-generated scores alone should not be presented as validated accuracy.

III. EXPECTED RESULTS AND DISCUSSION

The expected benefit of the hybrid approach is that explicit keyword overlap remains interpretable while semantic similarity can capture related terminology. The platform also converts the analysis into actions. Instead of ending with "ATS score = 62%," the user receives matched skills, missing skills, recommendations, resume improvements, and interview questions. This creates a candidate-development loop: Analyze → Identify Gaps → Optimize → Prepare → Evaluate.

IV. LIMITATIONS

The current system has several limitations. ATS scoring is an approximation and is not equivalent to any specific commercial ATS. Resume parsing quality can vary with complex layouts, tables, graphics, and unconventional formatting. Skill extraction can miss specialized or newly emerging terminology. Semantic similarity can identify related text without proving that a candidate truly satisfies a requirement. Generative AI can produce unsuitable suggestions and therefore requires user validation. Larger datasets and human evaluation are required before making claims about predictive validity or recruitment accuracy.

V. FUTURE WORK

Future work can investigate transformer-based matching models such as BERT, RoBERTa, and Sentence-BERT; domain-specific embeddings; richer skill ontologies and knowledge graphs; recruiter dashboards for multi-resume ranking; learning-resource recommendations for missing skills; institutional placement analytics; and multilingual resume/JD support. A future research version should also evaluate fairness, privacy, calibration, explainability, and robustness across industries and candidate populations.

➢ *Ethical Considerations*

AI-assisted recruitment should be treated as decision support rather than autonomous hiring. Potential concerns include algorithmic bias, inaccurate parsing, over-reliance on scores, privacy of candidate documents, and unequal access to AI services. Human review should remain part of final recruitment decisions. Candidate data should be protected through access control, secure storage, and appropriate retention policies.

VI. CONCLUSION

This paper presented PRO-ALIGN, an integrated AI-powered framework for resume–job description matching, ATS compatibility analysis, resume optimization, and interview preparation. The system combines keyword matching, semantic similarity, experience, project, education, and formatting analysis in an interpretable weighted scoring model. It then extends the result into actionable skill-gap analysis, AI-assisted resume improvement, personalized interview preparation, and reporting. The platform builds upon prior research in automated resume matching and semantic similarity while emphasizing an end-to-end candidate-development workflow. Future empirical work should establish a labeled dataset, compare the system with strong NLP baselines, and conduct human evaluation before making claims about real-world recruitment effectiveness.

REFERENCES

[1]. S. Maheshwary and H. Misra, "Matching Resumes to Jobs via Deep Siamese Network," in Companion Proceedings of The Web Conference 2018, Lyon, France, 2018, pp. 87–88, doi: 10.1145/3184558.3186942.

[2]. P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," Procedia Computer Science, vol. 167, pp. 2318–2327, 2020, doi: 10.1016/j.procs.2020.03.284.

[3]. M.-H. Ajjam and H. S. Al-Raweshidy, "AI-driven semantic similarity-based job matching framework for recruitment systems," Information Sciences, vol. 724, Art. no. 122728, 2026, doi: 10.1016/j.ins.2025.122728.

[4]. A. Gugnani and H. Misra, "Implicit Skills Extraction Using Document Embedding and Its Use in Job Recommendation," Proc. AAAI Conf. Artif. Intell., vol. 34, 2020, doi: 10.1609/aaai.v34i08.7038.

[5]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. NAACL-HLT, 2019.

[6]. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in Proc. EMNLP-IJCNLP, 2019.

[7]. A. Vaswani et al., "Attention Is All You Need," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[8]. C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Cambridge, U.K.: Cambridge Univ. Press, 2008.