

AI Based Resume Screening and Candidate Matching - A Review of Methods, Applications, and Responsible Use

Dr. Pawar Santosh Barku, Prof. Mukta Jadhav, Shivani Pratapsingh Rawat

Vidya Niketan Institute of Management and Computer Application, Bota, Sangamner.

Abstract: *Organizations receive large volumes of applications, while applicants often struggle to identify roles that fit their experience. Artificial intelligence (AI) can support these processes by extracting structured information from resumes, comparing candidate profiles with job requirements, and ranking or recommending potential matches. This review explains the main stages and methods used in AI-based resume screening and candidate matching, from document parsing and skill extraction to semantic retrieval and ranking. It summarizes potential benefits, including faster review and more consistent application of job-related criteria, and examines limitations involving biased historical data, weak proxies for job performance, inaccessible or incomplete resumes, privacy, explainability, and overreliance on automated scores. The literature indicates that algorithmic tools do not guarantee objective or fair decisions: outcomes depend on data, design choices, validation, and how employers use model outputs. Responsible systems should be evaluated for job relevance and subgroup performance, provide meaningful human review and candidate accommodations, and be monitored after deployment. AI is therefore best treated as decision support within a documented hiring process, rather than as an autonomous substitute for recruiter judgment.*

Keywords: Artificial Intelligence, Resume Screening, Candidate Matching, Natural Language Processing, Recruitment, Algorithmic Fairness, Explainable AI

I. INTRODUCTION

Recruitment combines information retrieval with consequential judgments about people. A resume contains evidence of skills, education, work history, and accomplishments, while a job description expresses an employer's requirements in a mixture of structured and informal language. When application volumes are high, recruiters may use applicant tracking systems and AI-enabled tools to organize applications, find relevant profiles, and support shortlisting. These tools range from simple keyword filters to machine-learning models and large language models (LLMs).

AI-based screening is not one single technology. It may identify and extract information from resumes, classify applications against minimum requirements, estimate similarity to a job description, rank applicants, or recommend candidates to recruiters. Matching may also operate in the other direction by recommending jobs to job seekers. Early work on resume-job matching used information extraction and semantic similarity; later systems have applied neural representations to compare resumes and job descriptions [1]. The central challenge is to represent candidate evidence and job requirements accurately enough to support a useful match without turning a similarity score into an unsupported prediction of a person's worth or future performance.

This paper is a focused narrative review of research and guidance on resume screening, candidate-job matching, fairness, and responsible deployment. It is organized around system methods, potential value, limitations, and practical safeguards. It does not claim to be a systematic review or a comparative benchmark of commercial products.



II. AI METHODS IN RESUME SCREENING AND MATCHING

2.1 Resume parsing and information extraction

A screening pipeline usually begins by converting a resume file into usable text. Optical character recognition may be needed for scanned documents. Parsing methods then identify fields such as contact details, employment dates, job titles, education, certifications, and skills. Rule-based patterns can reliably capture predictable formats, while natural language processing (NLP) methods such as named-entity recognition and contextual language models can help identify information expressed in varied ways. Errors at this stage—such as misreading dates, missing a column, or confusing a skill with a job title—can flow into later scoring and ranking.

2.2 Representing skills and requirements

A job description can be decomposed into required and preferred skills, experience, credentials, location, and other constraints. Resume phrases may be normalized to a shared skill vocabulary so that “customer support” and “client service,” for example, can be compared. Keyword and taxonomy approaches are relatively transparent but may miss transferable skills, synonyms, and context. They can also reward applicants who repeat phrases from an advertisement even when the underlying experience is limited.

2.3 Similarity, retrieval, and ranking

Traditional systems compare term counts or weighted vectors, using methods such as TF-IDF and cosine similarity. These methods are efficient and interpretable at a basic level, but they are sensitive to wording. Semantic embedding models represent text in a vector space and can retrieve resumes whose language differs while the underlying meaning is similar. Neural matching approaches have demonstrated the use of paired resume and job-description representations for ranking likely matches [1]. In larger systems, a two-stage design can first retrieve a broad set of candidates and then apply a more detailed ranking model to that set.

A model score should be interpreted narrowly: it estimates alignment to the selected criteria under the model and data used. It does not establish that a candidate will succeed, that an omitted candidate is unqualified, or that the ranking is fair. Recruiters should be able to see which job-related evidence contributed to a recommendation and check whether the system has overlooked relevant experience.

2.4 Generative AI and human review

LLMs can summarize resumes, extract structured details, compare evidence with a job description, and draft explanations. Their flexible language handling can reduce the need for rigid templates, but outputs can be inconsistent, overly confident, or unsupported by the source document. A practical design should require evidence-linked summaries, preserve the original resume for verification, and route uncertain or incomplete cases to a person. The model should not invent qualifications or infer sensitive characteristics from names, photos, addresses, or writing style.

III. POTENTIAL BENEFITS AND APPLICATIONS

AI can reduce the manual effort required to locate relevant applications in a large pool and help recruiters apply declared criteria more consistently. Candidate-job recommendation tools can also help applicants find openings that are difficult to discover through exact keyword search. Semantic retrieval may identify relevant experience described with different terminology, including adjacent skills or career transitions. These benefits depend on well-defined job requirements, reliable parsing, and a workflow that allows recruiters to correct errors.

Automation can support routine organization, but it does not remove the need for careful job analysis. A model trained to reproduce previous hiring decisions may reproduce prior selection patterns, including patterns that were unrelated to job performance or shaped by unequal opportunity. Raghavan and colleagues’ analysis of commercial hiring assessments emphasizes the importance of examining how systems are developed, validated, and audited, rather than relying on general claims that algorithms are objective [2].



IV. CHALLENGES AND LIMITATIONS

4.1 Bias in data, labels, and proxies

Historical hiring and employment data may encode unequal access, past discrimination, or organizational preferences that are not valid measures of ability. A model trained on past hires can learn to favor profiles resembling those previously selected. Even when protected attributes are excluded, features such as career gaps, institution names, location, or wording can act as proxies. Bias can enter at multiple points in the machine-learning lifecycle, including data collection, problem formulation, measurement, model development, and deployment [3, 4]. Removing demographic fields alone is therefore not a complete fairness strategy.

4.2 Validity and the meaning of a match

A system can be technically accurate at reproducing its labels while still making poor hiring recommendations if those labels are weak. Past recruiter decisions, employee performance ratings, or retention records may be noisy and influenced by workplace context. A high similarity score may also reward resume style or document length rather than relevant capability. Employers should specify what each model output means, validate it against job-related evidence, and avoid presenting a rank as a prediction of overall candidate quality.

4.3 Explainability and applicant experience

Applicants may not know that automated tools are involved or how their information is used. In four studies, Lavanchy and colleagues found that participants viewed algorithm-driven hiring procedures as less fair than human-only or algorithm-assisted human procedures, partly because algorithms were perceived as less able to recognize a candidate's individuality [5]. This evidence supports clear communication and a process for correcting inaccurate information. A score without an understandable reason or route for review can undermine trust even when the tool is technically useful.

4.4 Privacy, accessibility, and security

Resumes contain personal information and can include sensitive details unrelated to job requirements. Employers should limit collection to necessary information, define retention periods, restrict access, and assess vendors' data handling. Resume formats and application pathways must accommodate assistive technologies and disability-related needs. The U.S. Equal Employment Opportunity Commission (EEOC) warns that algorithmic tools may screen out applicants with disabilities and describes the need for reasonable accommodations and safeguards [6].

4.5 Distribution shift and automation bias

Job requirements, language, labor markets, and applicant populations change. A system that worked on one occupation or historical dataset may not perform well for another role or later hiring cycle. Recruiters can also give undue weight to a machine-generated ranking, especially when the system's limitations are unclear.

Human oversight is meaningful only when reviewers can inspect evidence, challenge recommendations, and override them without penalty.

V. RESPONSIBLE DESIGN AND DEPLOYMENT

Organizations should begin with a job analysis and document the minimum qualifications and preferred evidence for each role. The model should be evaluated on data representative of the roles and applicant populations where it will be used. Evaluation should measure both job-related performance and error patterns across relevant groups, including intersectional groups where sample sizes permit. Metrics should be selected for the decision context; no single fairness measure can establish that a hiring process is fair.

Before deployment, employers should test the full workflow, not only the model in isolation. Audits should include resume parsing errors, ranking stability under equivalent resume wording, false exclusions, accessibility, and subgroup outcomes. After deployment, organizations should monitor changes in data, job criteria, and outcomes; record versions and decisions; and provide a process for candidate questions, correction, accommodation, and human reconsideration. Vendors should provide sufficient documentation for employers to assess intended use, validation, limitations, and data practices.



Legal duties depend on jurisdiction and use. The EU Artificial Intelligence Act identifies AI systems intended for recruitment or selection—including systems that analyze or filter applications and evaluate candidates—as high-risk in Annex III [7]. In the United States, existing employment discrimination and disability laws continue to apply when employers use software or AI in hiring; EEOC guidance highlights disability-related risks and accommodations [6]. These references are not a substitute for jurisdiction-specific legal advice, and requirements may change as implementation proceeds.

VI. FUTURE DIRECTIONS

Future research should compare keyword, embedding, and LLM-based matching on transparent and representative datasets, with shared definitions of relevance and job-related validity. Evaluation should include the quality of retrieved candidates, subgroup error rates, ranking stability, and the ability of recruiters to detect system mistakes. More work is also needed on how applicants understand and respond to AI-assisted hiring and whether disclosure, explanations, and review mechanisms improve perceived fairness.

Promising design directions include evidence-grounded explanations, uncertainty-aware ranking, skill-based matching that recognizes transferable experience, and privacy-preserving evaluation. Researchers and employers should distinguish between matching assistance and automated selection, since the latter carries greater consequences for applicants. Systems should be tested in real workflows with affected stakeholders, and results should be reported with enough detail to support independent scrutiny.

VII. CONCLUSION

AI-based resume screening and candidate matching can help organize large application pools and connect skills with job requirements. NLP, semantic embeddings, and ranking models can move beyond exact keyword comparison, but they also introduce risks that are easy to hide behind a single score. Data quality, valid criteria, accessibility, transparency, and human review determine whether a tool supports sound decisions or scales existing errors. The evidence reviewed here supports using AI as a carefully monitored decision-support aid, with meaningful recruiter accountability and safeguards for applicants, rather than as an automatic gatekeeper.

REFERENCES

- [1] S. Maheshwary and H. Misra, "Matching Resumes to Jobs via Deep Siamese Network," Companion Proceedings of The Web Conference 2018, pp. 87–88, 2018. <https://doi.org/10.1145/3184558.3186942>
- [2] M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices," Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 2020. <https://doi.org/10.1145/3351095.3372828>
- [3] A. Köchling and M. C. Wehner, "Discriminated by an Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of HR Recruitment and HR Development," Business Research, vol. 13, pp. 795–848, 2020. <https://doi.org/10.1007/s40685-020-00134-w>
- [4] H. Suresh and J. V. Guttag, "A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle," Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, 2021. <https://doi.org/10.1145/3465416.3483305>
- [5] M. Lavanchy, P. Reichert, J. Narayanan, and K. Savani, "Applicants' Fairness Perceptions of Algorithm-Driven Hiring Procedures," Journal of Business Ethics, vol. 188, pp. 125–150, 2023. <https://doi.org/10.1007/s10551-022-05320-w>
- [6] U.S. Equal Employment Opportunity Commission, "The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees," May 2022. <https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>
- [7] European Parliament and Council of the European Union, Regulation (EU) 2024/1689 (Artificial Intelligence Act), Annex III, §4(a), 2024. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

