

A Review of AI-Based Resume Screening and Job Matching Systems

¹Arsh Solkar, ²Gaurangi Sonar, ³Vaidehi Talekar, ⁴Santosh Jadhav

^{1,2,3}*Student, BE Information Technology, Department of Information Technology Finolex Academy of Management and Technology, Ratnagiri, India*

⁴*Associate Professor Department of Information Technology, Finolex Academy of Management and Technology, Ratnagiri, India*

Abstract—Recruiters today are facing a challenge. They get many applications that its impossible to review them all. Because of this resume screening has become the automated part of the hiring process. This paper looks at how intelligence and machine learning can be used to screen resumes and match candidates to jobs. We examine methods such as keyword extraction, TF-IDF similarity, document embeddings, Siamese convolutional networks and transformers. We focus on four studies: a systematic review of AI in hiring a TF-IDF semantic matching approach, a deep Siamese network for resume-job matching and a survey of machine learning techniques in resume screening. As the way we represent data gets more advanced—from simple word counts to semantic models—the accuracy of matching improves. Most of the models we studied only do one thing. They don't test their results against hiring decisions. They rely on labels. They don't explain why a candidate gets a certain score. The results on reducing bias are mixed. Issues like transparency and privacy remain unsolved. We believe that a better solution is needed—a system that is integrated, explainable and aware of bias and that includes people in the loop. To address this we introduce the HireSense AI framework. We also suggest directions for research.

Index Terms—Resume Screening, Job Matching, Natural Language Processing, Semantic Similarity, TF-IDF, Siamese Network, Explainable AI, Algorithmic Bias, Recruitment Automation.

I. INTRODUCTION

Hiring today is a process. Every job opening receives online applications. Resumes come in formats, lengths and structures. Reading each resume by hand takes time. Manual screening and keyword matching is inefficient. Gives no insight into a candidate's skills [4]. Companies want to cut costs and treat every candidate the same; these reasons drive the use of Artificial Intelligence [1].

Artificial Intelligence, Machine Learning and Natural Language Processing are now common in the recruitment cycle. They help with sourcing, screening resumes, detecting skills, analyzing job descriptions matching similarity scores, ranking candidates and interviewing [1] [4]. Approaches differ: some use term matching and TF-IDF similarity scores [2]; others use document embeddings, siamese neural networks [3]. Transformers such as BERT [5].

With these advances some problems remain. Some systems perform one task instead of supporting the whole screening process. Real-world recruiting decisions are not fully documented, so models are judged on labels. Learned models can be biased by input data. Are hard to interpret, which lowers recruiter trust [1] [2]. Additionally resumes contain information about candidates [2] [4].

This paper reviews four studies on using Artificial Intelligence, in recruitment and resume–job matching. We compare the methods, datasets and findings of the authors. Point out remaining gaps. Then we discuss research directions and introduce our concept HireSense AI.

II. LITERATURE REVIEW

A. AI in Employee Recruitment: A Systematic Review (Discover Global Society, 2025)

A systematic review was done by Dadaboyev et al. [1] Using the PRISMA methodology. In their study they identified 1,647 papers from the Web of Science database. Selected 49 peer-reviewed papers published from 2018, to 2025. This systematic review organizes the literature into three categories: AI-based candidate identification and screening AI-based interviewing and assessment and ethical challenges and considerations. The studies identify processing of a large number of applications and improved quality of the applicant pool and describe AI technology as "collaborative intelligence " assisting human recruiters rather than replacing them. Far as bias is concerned though the authors state that the results are quite controversial since AI tools based on the hiring history might inherit past bias. Furthermore black-box nature, legal ambiguity and reluctance of human resources professionals are discussed. Despite the scope of review it can be considered relevant. Nevertheless being a review it does not present a matching system, which was actually proposed.

B. TF-IDF Semantic Similarity Job Matching Framework (Information Sciences, 2026)

In their article, Ajjam and Al-Raweshidy [2] put forward a real-time recruitment system that matches resumes to job descriptions. The system uses TF-IDF vectorization, cosine similarity and weights keywords based on domains. They apply a one-to-one matching method, which pairs each job with just one candidate.

For testing the authors used 660 job listings from Glassdoor for data science roles 962 resumes from Kaggle covering 25 domains and also created a synthetic dataset. Their semantic matching method worked better than a keyword overlap approach in every domain. For example in tests for software engineers the semantic method scored 0.74 while the keyword method only reached 0.35. They also did a ranking-based test with 82 resumes. For Data Science roles Precision@10 was 0.400 and for Hadoop roles it was 0.598. The TF-IDF model did better than two models: MiniLM sentence-transformer, which scored 0.262 and 0.188 and the keyword overlap baseline, which

scored 0.078 and 0.004. The authors chose TF-IDF because it's easy to understand, doesn't cost much and has a risk of bias. They also talked about privacy mentioning how they removed protected attributes and used anonymization methods. These steps follow GDPR and CCPA rules. However they pointed out some problems. The system only works in English so it's language-dependent. It also struggles with sparse texts. They didn't have recruiter feedback to compare against and they admit TF-IDF itself might still carry some bias. They said fairness testing should be done at work.

C. Resume–Job Matching via Deep Siamese Network (WWW Companion, 2018)

The matching problem is presented as a task to learn a similarity metric according to Maheshwary and Misra [3]. The resumes and job descriptions are turned into 200-dimensional embeddings [8] and then given to two weight-sharing convolutional networks that use contrastive loss. It is assumed that matching pairs of resumes and job descriptions are close together in the space while non-matching pairs are far apart. The number of resumes is 1,314. The number of job descriptions is 3,809, which together produce 5,005,026 pairs. Labeling is done semi-automatically: a portion of the data is annotated manually by interns and label propagation is applied to the remaining data. The siamese network achieved accuracies of 81.34% and 83.19%, on two test sets whereas a plain CNN reached 73.31% and 74.68% and a TF-IDF method reached 60.46% and 61.23%. Thus the experiment shows that the learned semantic representation outperforms term-based methods. However the siamese network behaves like a box; some labels have been propagated without validation by recruiters and fairness has not been examined in this work.

D. Machine Learning for Resume Analysis and Job Description Matching (IJRITCC, 2024)

Modak et al. [4] Give a look at how machine learning and natural language processing are used in resume analysis. They walk through the process starting with resume parsing then moving to skill extraction, job description analysis and finally matching algorithms. The methods they cover range from tools like regular expressions and ontology-based parsing to more advanced ones such as CNN-CRF segmentation [10] Hidden Markov Models (HMM) for extracting information support vector machines (SVM) and ensemble learning. They also include BERT-based matching [5] as an approach. The paper outlines challenges in the field. These include problems related to bias and fairness in automated systems, the need for explainability in decisions handling unstructured and constantly changing data, measuring skills and culture fit, maintaining a good balance between human judgment and machine automation, protecting privacy and security and adapting to the fast-changing job market. Looking ahead the authors suggest research directions like creating synthetic data and improving feature extraction techniques. While the paper is helpful, in organizing and classifying the methods it doesn't provide a way to compare them directly. Also, some of the performance metrics used don't have benchmarks making it harder to assess how well the methods actually work.

III. COMPARATIVE ANALYSIS

The four studies approach AI in recruitment from different angles: a broad systematic review [1], a lightweight and interpretable similarity framework [2], a deep-learning matching model [3] and a task-oriented survey [4]. Table I summarizes their methods, data, results and limitations.

TABLE I. COMPARISON OF THE REVIEWED STUDIES

Study	Approach	Data	Key Results	Limitations
Dadaboyev et al. [1] (2025)	Systematic review (PRISMA) of AI in recruitment; three themes	49 articles from Web of Science, 2018–2025	Efficiency and candidate-quality gains; bias evidence mixed; need for legal and ethical frameworks	Single database; no system proposed or tested
Ajjam & Al-Raweshidy [2] (2026)	TF-IDF + cosine similarity, domain keyword weighting, greedy one-to-one matching	Glassdoor 660 postings; Kaggle 962 resumes; simulated set	Precision@10 of 0.400 (Data Science) and 0.598 (Hadoop) vs 0.262 / 0.188 for SBERT-MiniLM	English only; no recruiter ground truth; fairness untested; relevance defined by resume domain
Maheshwary & Misra [3] (2018)	Doc2Vec embeddings with siamese CNN and contrastive loss	1,314 resumes, 3,809 JDs, 5,005,026 pairs	Accuracy 81.34% / 83.19% vs 73.31% / 74.68% for CNN	Black box; semi-automatic labels; no fairness analysis
Modak et al. [4] (2024)	Survey of parsing, skill extraction, similarity and matching with ML/NLP	Literature survey (various datasets)	Taxonomy of tasks; list of open challenges and future directions	Descriptive; no unified benchmark or new system

Together the studies show a sequence in how resumes and job descriptions are represented namely: keyword matching, TF-IDF, Doc2Vec or learned embeddings and CNN or transformer-based models [2]–[5]. In [3] the performance of learned semantics is better than TF-IDF. In [2] the opposite is true when comparing it to a pre-trained sentence transformer. This difference is not surprising because the studies used datasets and also had different ways of defining "relevance." Relevance in [2] was based on the field of the resume. In [3] it was based on what recruiters thought. Also in [3] some of the labels were created using label propagation. Accuracy, Precision@10 and raw cosine scores could not be compared directly.

Three points come to mind. First no representation works the same for all datasets. So combining semantic methods makes sense. Second there is a tradeoff between being able to understand the model and how accurate it's. TF-IDF gives understanding but siamese and transformer models offer better accuracy [2] [3]. Third [1] and [2] look at fairness and privacy in their work. Also, in [2] checking for fairness is something that will be done in the future.

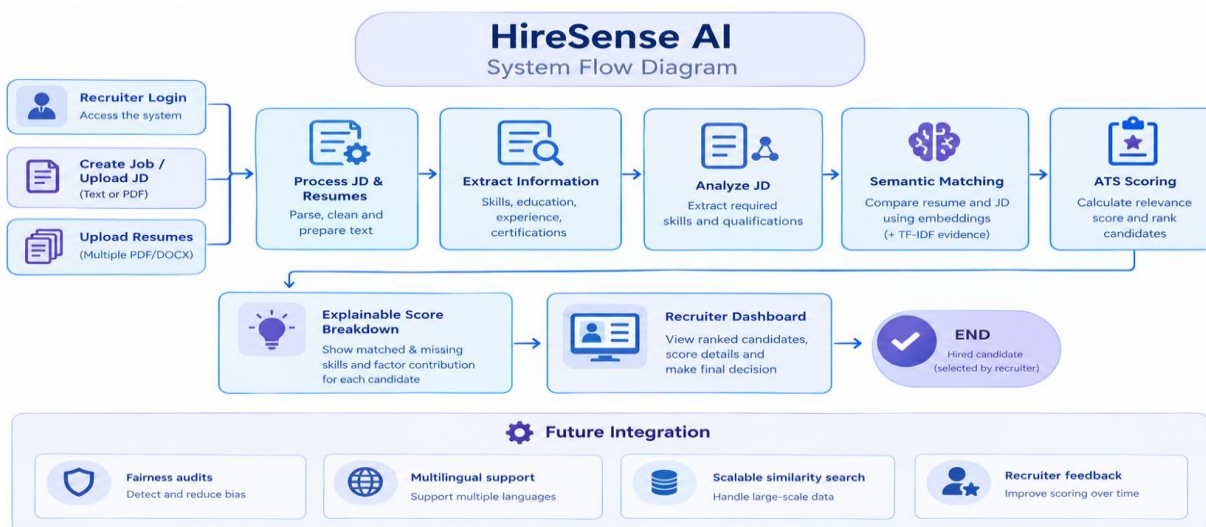
IV. CHALLENGES AND RESEARCH GAP

The reviewed article showcases how AI can be used for resume screening, while at the same time demonstrating the persistent problems it encounters.

- Lack of ground truth. Final selection results are rarely available, limiting the process to verification using proxy labels, propagated labels, or domain labels [2], [3]. There is no common benchmark with recruiter-approved labels.
- Bias and fairness. Historical data-based models can perpetuate biases, and even TF-IDF can privilege terms related to certain social groups [1], [2], [4]. There is no clear evidence on whether AI can help alleviate such issues [1].
- Explainability. Scoring based on a black-box model hinders recruiters' trust, and reluctance to use the system is well known [1], [4]. There are few systems explaining what causes a specific score.
- Fragmented solutions. Resume parsing, skills identification, and ranking are usually researched separately, but not as part of one pipeline [4].
- Dynamic and unstructured data. Resumes have different formats and structures, and job requirements and terminology can change quickly [2], [4].
- Human judgment and cultural fit. Cultural fit and soft skills cannot be easily quantified, and over-reliance on automatic processing devalues the role of human judgment [1], [4].
- Privacy and security. Resumes contain personal data, thus require anonymization and proper access control [2], [4].
- Context. System presented in [2] works only with English language.

Thus, there is a possibility to develop a comprehensive solution integrating parsing, skills extraction, and hybrid semantic matching, including an explanation of the score, built-in bias and privacy controls, leaving the final decision to the recruiter.

V. PROPOSED SYSTEM FRAMEWORK



VI. FUTURE DIRECTION

There are ways to make AI-based resume screening better. First it is important to look at how combining keyword (TF-IDF)-based scoring with sentence embedding similarity works. This would help explain the method and make the second method more meaningful. Second having score explanation that looks at matched skills missing skills, experience and education can help recruiters understand and question the score. It is clear from the research that this kind of openness is linked to trust and fairness [1].

Checking for fairness should be done through tests, not guesses. Subgroup audits and fake demographic information used in [2] should be part of the process. Also removing details before scoring is important. In addition the way we check things should move towards datasets that are rated by recruiters and use measures like Precision and Recall.

To make things work better with people we need methods that reduce the number of features and find similar items quickly [2]. Other things to consider are supporting languages keeping the list of skills up to date with new words and following privacy rules like not keeping data longer than needed. All these ideas should be used while thinking about the intelligence view, from [1].

VII. CONCLUSION

AI and machine learning can make resume screening faster and more consistent. The studies reviewed here show gains in matching quality as representations moved from keywords to learned semantics [1]–[4]. At the time most approaches solve a single task. They depend on labels. They explain their scores poorly. They address fairness and privacy in part. The evidence that AI reduces hiring bias remains mixed.

This review therefore points to the need for an integrated, explainable and bias-aware screening approach. The same approach must involve AI supporting recruiters. The same approach must not replace recruiters. The proposed HireSense AI framework sketches one approach. Future work, on recruiter-labelled evaluation, fairness auditing, multilingual support and scalability can help make AI-based resume screening more transparent, practical and trustworthy.

ACKNOWLEDGMENT

The authors thank their project guide and faculty members for their guidance and support during this review, and acknowledge the research papers that helped them study existing AI-based recruitment and resume-matching approaches and identify the research gaps discussed here.

REFERENCES

- [1] S. M. U. Dadaboyev, J. Abdullayeva, N. Abbosova, A. Suleymenova, and K. Mamadjanova, “Role of artificial intelligence in employee recruitment: systematic review and future research

directions,” Discover Global Society, vol. 3, art. no. 99, 2025, doi: 10.1007/s44282-025-00246-w.

- [2] 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.
- [3] S. Maheshwary and H. Misra, “Matching resumes to jobs via deep siamese network,” in *Companion Proc. The Web Conference 2018 (WWW '18 Companion)*, Lyon, France, Apr. 2018, doi: 10.1145/3184558.3186942.
- [4] S. Modak, P. Shinde, A. Tiwari, and S. Nalamwar, “A review of resume analysis and job description matching using machine learning,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 12, no. 2, pp. 247–250, Feb. 2024.
- [5] V. Bhatia, P. Rawat, A. Kumar, and R. R. Shah, “End-to-end resume parsing and finding candidates for a job description using BERT,” arXiv:1910.03089, 2019.
- [6] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence embeddings using siamese BERT-networks,” arXiv:1908.10084, 2019.
- [7] 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, pp. 4171–4186.
- [8] Q. V. Le and T. Mikolov, “Distributed representations of sentences and documents,” in *Proc. ICML*, 2014.
- [9] K. Kurita, N. Vyas, A. Pareek, A. W. Black, and Y. Tsvetkov, “Measuring bias in contextualized word representations,” arXiv:1906.07337, 2019.
- [10] C. Ayishathahira, C. Sreejith, and C. Raseek, “Combination of neural networks and conditional random fields for efficient resume parsing,” in *Proc. Int. CET Conf. Control, Commun. Comput. (IC4)*, IEEE, 2018, pp. 388–393.
- [11] P. K. Roy, S. S. Chowdhary, and R. Bhatia, “A machine learning approach for automation of resume recommendation system,” *Procedia Comput. Sci.*, vol. 167, pp. 2318–2327, 2020.