

Automated Resume Screening and Job Matching Through Semantic-Aware Skill Cascade Classification

B. Hemanth Kumar¹, K. Venkata Rao², P. Swathi³

¹*MTech Student, Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Visakhapatnam, India*

²*Professor, Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Visakhapatnam, India*

³*Research Scholar, Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Visakhapatnam, India*

Abstract—screening and candidate-to-job matching remain among the most time-consuming and inconsistently executed stages of the recruitment pipeline, and conventional keyword-based filters are known to reward exact phrasing over genuine competence. This paper presents a semantic-aware skill cascade framework that fuses transformer-based semantic similarity with a hierarchical, confidence-weighted skill taxonomy to produce accurate, transparent resume-job matching. Rather than treating skill detection as a single binary lookup, the framework resolves each required skill through a three-tier cascade, exact, fuzzy, and contextual, so that misspellings, abbreviations, and paraphrased descriptions of a competency are still recognized, each with a tier-appropriate confidence weight. A data-driven fusion mechanism then combines this skill-overlap signal with sentence-embedding similarity into a single interpretable score and a three-way tier label (Best Match, Average, Low Match).

The framework is evaluated on a dataset of 500 resumes, 200 job descriptions, and 250 gold-annotated resume-job pairs spanning five job categories, with an additional 4,250 skill-level instances used to evaluate the cascade in isolation. The system attains 93.6% tier classification accuracy, a Pearson correlation of 0.9856 against gold match scores, and an overall Matthews Correlation Coefficient of 0.8996. The skill cascade itself reaches 97.25% cumulative recall and 100% precision at the contextual tier, and an ablation removing contextual detection drops recall by 27.5 percentage points, confirming that paraphrase-level skill recognition is a primary driver of matching quality rather than a marginal refinement. These results indicate that lightweight, taxonomy-guided semantic fusion can approach the accuracy of opaque end-to-end models while retaining a fully auditable decision trail.

Index Terms—Resume screening, job matching, skill extraction, semantic similarity, natural language processing, information retrieval, skill taxonomy, recruitment automation.

I. INTRODUCTION

Recruitment teams routinely receive far more applications than can be manually reviewed with any consistency, and the pressure to shortlist quickly pushes organizations toward automated screening. The earliest and still most common form of this automation, keyword and Boolean filtering against a resume database, optimizes for the presence of a literal string rather than for the underlying capability it is meant to signal, so a resume that describes “building predictive models from historical data” can be silently discarded by a filter searching for the token “machine learning.” This mismatch between what a filter can see and what a recruiter actually wants to measure is the central problem this paper addresses.

Two broad families of prior work have each solved part of this problem. Lexical and TF-IDF-based pipelines are fast, interpretable, and easy to audit, but their recall collapses whenever a skill is expressed through a synonym, an abbreviation, or a descriptive paraphrase rather than its canonical name. Embedding- and deep-learning-based pipelines close this recall gap by comparing resumes and job descriptions in a learned semantic space, but the resulting similarity score is typically a single opaque number: a recruiter cannot easily tell which requirements were satisfied and which were not,

which is precisely the kind of justification that hiring decisions increasingly require for fairness and defensibility.

This work adopts “Harnessing NLP to align Candidate Profile with Position Requirements” [1] as its base paper. That work combines lexical and semantic similarity to align candidate profiles with job requirements using classical and recurrent classifiers, but it treats skill matching as an implicit byproduct of document-level similarity rather than as an explicit, inspectable step, and it does not learn or justify the weighting between its lexical and semantic components.

We retain the base paper's premise that lexical and semantic signals are complementary, but restructure how each is computed and combined:

1) A structured skill taxonomy in which every skill carries both fuzzy spelling/formatting variants and a contextual paraphrase, so that a single taxonomy entry can be matched at three different levels of textual similarity rather than one.

2) A three-tier detection cascade (exact then fuzzy then contextual) that assigns each detected skill mention a confidence tier, preserving the reasoning behind every match rather than collapsing it into a single similarity number.

3) A data-driven fusion mechanism that combines tier-weighted skill overlap with sentence-embedding similarity under weights selected by an explicit hyperparameter sweep rather than fixed by convention.

The resulting system produces a continuous 0 to 100 match score together with a discrete, explainable tier label, giving recruiters both a ranking signal and a rationale. Our contributions are: (1) a purpose-built dataset of 500 resumes, 200 job descriptions, and 250 gold-annotated resume-job pairs across five job categories, with skill-level mention-type labels covering 4,250 individual skill instances; (2) a 50-skill taxonomy spanning five job categories, each skill paired with fuzzy variants and a contextual paraphrase; (3) a hybrid semantic-skill fusion algorithm that achieves 93.6% tier classification accuracy and 0.9856 Pearson correlation with gold match scores; and (4) an ablation and cascade-level

analysis quantifying the individual contribution of exact, fuzzy, and contextual matching to overall recall.

II. RELATED WORK

A. Resume Parsing and Information Extraction

BERT-based named entity recognition has been applied to extract structured fields such as skills, education, and experience from resumes spanning multiple job domains, achieving high extraction precision but leaving downstream matching unaddressed [2]. Complementary systems couple NLP-based parsing with classical classifiers such as support vector machines to align parsed resume fields against job descriptions, improving screening consistency [3]. Both lines of work are effective at turning unstructured resume text into structured fields, but neither addresses how those fields should be ranked, scored, or reduced to a hiring decision, the gap this paper targets.

B. NLP-Based Job-Candidate Matching

Traditional lexical relevance-scoring methods provide interpretable results but limited recall on paraphrased skills, while transformer-based semantic embeddings capture deeper contextual relationships between resumes and job descriptions. Several recent systems align candidate profiles with position requirements using combined lexical and semantic similarity [1], apply BERT embeddings directly to job-matching pipelines [4], or integrate NLP with ranking and machine-learning classifiers to score and order candidates [5], [6]. Deep-learning ranking frameworks incorporating optical character recognition and object detection have also been proposed to sort resumes by similarity to a target job description [7]. Across this body of work, similarity is almost always computed at the document level; few systems explicitly model skill-level hierarchies or expose which requirements drove a given score.

C. Skill-Based Talent Matching

Skill-centric approaches [1] emphasize capability matching over generic document-level similarity, which is closer in spirit to how a human recruiter actually reads a resume, checking off specific requirements rather than judging holistic similarity. Our work extends this direction by integrating semantic embeddings with a structured, three-tier skill

taxonomy (exact, fuzzy, and contextual), aiming for both the recall benefits of semantic matching and the auditability of explicit skill-level scoring.

D. Multimodal Candidate Evaluation

A separate strand of work targets the later, interview stage of hiring. Video-interview analysis assesses communication and personality through audiovisual behavior [8]; gaze and head-pose channels combined with linguistic features [11]; and multimodal, eye-contact-aware interview scoring [9]. Deep residual networks have further been used to predict apparent personality traits and interview suitability from short video blogs [10]. These frameworks are complementary to, rather than competitive with, the resume-based matching problem addressed here.

E. Comparative Analysis of Related Approaches

Table I summarizes, at a qualitative level, how the proposed framework compares with two groups of reviewed approaches across the capabilities discussed above: the resume-matching studies [1]– [7], and the multimodal interview-analysis studies [8]– [11]. The latter group is included for completeness but, as it targets a different task (post-screening interview assessment) rather than resume-job matching, shows no support for any of the resume-matching capabilities listed.

TABLE I. COMPARATIVE CAPABILITY ANALYSIS ACROSS RELATED APPROACHES

Capability	[1]– [7]	[8]– [11]	Proposed
Resume parsing / extraction	✓	—	✓
Lexical matching	✓	—	✓
Semantic matching	✓	—	✓
Explicit skill matching	Limited	—	✓
Exact / fuzzy skill matching	Limited	—	✓
Contextual paraphrase matching	Limited	—	✓
Skill-matching cascade	—	—	✓

✓ = supported; Limited = supported by some but not all papers in the group, or only implicitly; — = not supported / not addressed by any paper in the group.

[1]– [7]: Jayadharshini et al.; Srivastava & Greaney; Kamkar et al.; Pias et al.; Dharmendra et al.; Sai Surya et al.; Tanberk et al. (resume-job matching). [8]– [11]: Naim et al.; Randika et al.; Güçlütürk et al.; Gorbova et al. (multimodal interview analysis).

III. METHODOLOGY

A. System Architecture

The pipeline comprises five stages: (1) text preprocessing, (2) taxonomy-driven skill extraction via the three-tier cascade, (3) semantic similarity computation over resume and job-description embeddings, (4) hybrid score fusion, and (5) tier classification against learned score cutoffs. Fig. 1 illustrates this flow from raw resume upload through to the tier decision.

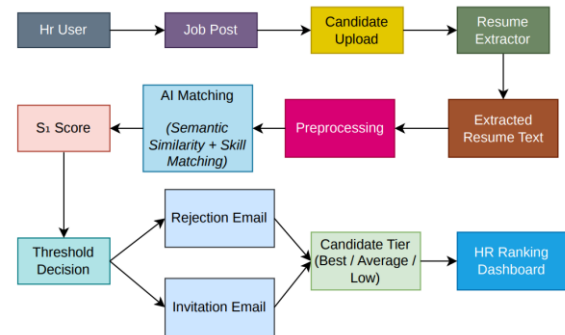


Fig. 1. System architecture: resume and job-post inputs are matched by the skill cascade and semantic module, fused into a score, and mapped to a tier.

B. Skill Taxonomy and Three-Tier Detection

Each skill in the taxonomy is represented by three complementary forms: a canonical form (e.g., “Python”), a set of fuzzy variants capturing common misspellings and formatting differences (e.g., “Pyhton,” “python3”), and a contextual paraphrase, a natural-language sentence describing the skill's use without naming it. Skill detection proceeds as a hierarchical cascade, attempting each tier only if the previous one fails: a skill is Exact if the canonical form appears in the document; Fuzzy if a fuzzy variant appears instead of the canonical form; and Contextual if the maximum cosine similarity between the paraphrase embedding and every sentence embedding of the resume meets a threshold $\theta_c = 0.65$.

Exact matching uses word-boundary-aware regular expressions on normalized text. Fuzzy matching combines Levenshtein ratio and token-sort ratio over

sliding windows of the resume text against the canonical form and its variants, accepted above a length-adaptive threshold that is stricter for short skill names (92% for names of four characters or fewer, 85% up to eight characters, 78% beyond) to avoid spurious matches on short tokens. Contextual matching embeds the paraphrase and every sentence of the resume with a pretrained sentence-transformer encoder and accepts a match when the maximum cosine similarity across sentences meets θ_c , so that a skill described anywhere in the document, not just in a skills list, can still be recognized.

C. Semantic Similarity

Independently of skill-level detection, the resume and job description are each embedded as whole documents, and their semantic alignment is measured by cosine similarity:

$$\phi_{sem} = \cos(r, j) = (r \cdot j) / (\|r\| \|j\|) \quad (1)$$

This term captures broader contextual and domain alignment, for instance, overall seniority and phrasing style, that is not reducible to any single skill match.

D. Skill Overlap and Hybrid Score

Each detected skill contributes to the overlap score in proportion to the confidence of the tier at which it was detected, with $w(\text{exact}) = 1.0$, $w(\text{fuzzy}) = 0.7$, and $w(\text{contextual}) = 0.4$, so that a paraphrased skill mention still counts toward the match but is weighted more conservatively than an explicit, unambiguous mention.

The semantic and skill-overlap signals are combined through a weighted fusion:

$$\text{Score} = 100 \times (w_s \cdot \phi_{sem} + w_k \cdot \theta_{skill}) \quad (2)$$

with weights $w_s = 0.3$ and $w_k = 0.7$ selected empirically (Section VI-D). The resulting score is mapped to a three-way tier using cutoffs tuned on the same sweep: Best Match for scores above 65, Average for scores between 40 and 65 inclusive, and Low Match for scores below 40. The dominant weighting toward θ_{skill} reflects an empirical finding that explicit skill coverage is a substantially stronger predictor of a human-assigned match score than whole-document semantic similarity alone.

IV. USER INTERFACE DESIGN

To make the tiered match score usable by non-technical hiring staff, the framework is exposed

through a lightweight web interface in which a candidate uploads a resume and submits it against an open job posting, and a recruiter views the resulting score, tier label, and per-skill breakdown from the same screen.

Fig. 2 shows the candidate-facing upload and application-submission screen.

Fig. 2. Candidate application interface showing resume upload and application submission confirmation.

V. DATASET

No existing public corpus jointly provides resumes, job descriptions, a paraphrase-annotated skill taxonomy, and human-verified match scores at the pair level, so the evaluation dataset used in this paper was purpose-built.

It spans five job categories, Data Science, Web Development, DevOps, Finance, and Human Resources, and consists of four linked components: (a) a 50-skill taxonomy (12 Data Science, 10 Web Development, 10 DevOps, 10 Finance, 8 Human Resources), each with fuzzy variants and a contextual paraphrase; (b) 500 resumes, each with ground-truth skills and a mention-type label (exact/fuzzy/contextual) per skill; (c) 200 job descriptions with required-skills lists; and (d) 250 gold match pairs with a continuous match score and a discrete tier (101 Best Match, 47 Average, 102 Low Match).

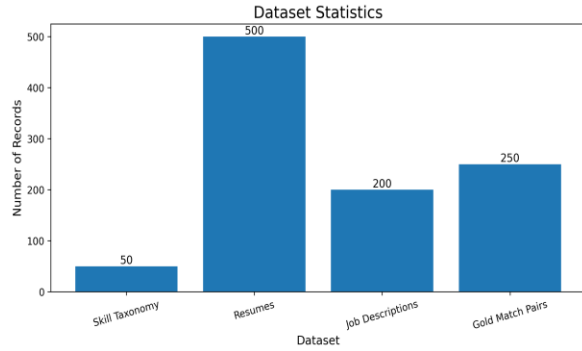


Fig. 3. Dataset composition: 50 taxonomy skills, 500 resumes, 200 job descriptions, and 250 gold-annotated pairs.

Both resumes and job descriptions are held close to evenly split across the five categories (18.4% to 21.6% per category for resumes, 17.5% to 21.5% for job descriptions), which avoids the class imbalance that would otherwise bias tier-level evaluation toward a majority domain.

Before semantic and skill-level features are computed, every resume and job-description text are normalized: email addresses, URLs, and profile links are stripped; phone numbers are removed while short numeric tokens (version numbers, years) are preserved; whitespace and case are normalized; and text is segmented into sentences for contextual-tier matching.

VI. EXPERIMENTS AND RESULTS

A. Overall Performance

Resume-job matching is evaluated on the 250 gold pairs described in Section V, comparing each pair's predicted score and tier against its gold match score and tier.

Table II summarizes pair-level performance across all 250 gold pairs.

TABLE II. OVERALL PERFORMANCE METRICS

Metric	Value
Pearson Correlation	0.9856
Tier Accuracy	93.6%
Overall MCC	0.8996

B. Per-Tier Performance

Table III breaks the same 250 pairs down by gold tier.

TABLE III. PER-TIER CLASSIFICATION RESULTS (N=250)

Tier	Precision	Recall	F1	Support
Best Match	0.9238	0.9604	0.9418	101
Average	0.8298	0.8298	0.8298	47
Low Match	1.0000	0.9608	0.9800	102

Low Match precision is 1.0000, meaning no gold Low Match pair is predicted as a higher tier in this evaluation. 96.04% Best Match recall means very few strong candidates are missed.

The comparatively lower 82.98% Average-tier performance is concentrated on the smallest support class (47 pairs) and reflects genuine boundary ambiguity between adjacent tiers rather than a systematic failure: nearly all Average-tier errors are adjacent-tier confusions (Best to Average or Average to Low), as confirmed in Fig. 4.

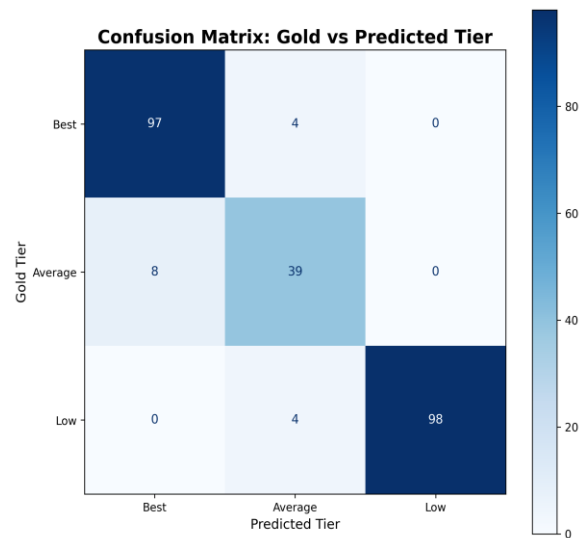


Fig. 4. Confusion matrix of tier classifications. Errors are concentrated on adjacent-tier boundaries rather than far misses.

C. Skill Cascade Analysis

The skill cascade is evaluated separately, at finer granularity, over all 4,250 resume-skill instances implied by the mention-type labels, comparing each skill's predicted detection tier against its ground-truth mention type. Table IV reports per-tier detection performance.

TABLE IV. SKILL CASCADE RECALL

Tier	Recall	Cum. Recall	Precision
Exact	100.00%	100.00%	96.82%
Fuzzy	69.50%	97.04%	99.24%
Contextual	99.46%	97.25%	100.00%

Exact matching alone resolves every exact-tier skill instance.

Adding fuzzy matching raises cumulative recall to 97.04%, and adding contextual matching raises it further to 97.25% across the full skill set. Precision improves down the cascade (96.82% then 99.24% then 100.00%), confirming that later tiers are invoked selectively, only when earlier, higher-confidence tiers fail, and rarely misfire once they are.

D. Hyperparameter Sensitivity

The fusion weight w_s and contextual threshold θ_c were selected via a grid sweep over held-out gold pairs, jointly optimizing tier accuracy and Pearson correlation.

Table V shows accuracy and correlation around the selected operating point; accuracy and MCC both peak sharply at $w_s = 0.3$ and degrade as semantic weight increases beyond it, confirming that skill-level evidence, not whole-document semantic similarity, is the dominant signal for this task. The selected configuration is $w_s = 0.3$, $w_k = 0.7$, $\theta_c = 0.65$.

TABLE V. HYPERPARAMETER SWEEP RESULTS

w_s	θ_c	Accuracy	Correlation
0.2	0.65	92.8%	0.9834
0.3	0.65	93.6%	0.9856
0.4	0.65	91.6%	0.9851

Fig. 6 shows the full sweep over semantic weight at the selected threshold: accuracy and Matthews Correlation Coefficient both peak sharply at $w_s = 0.3$ and degrade steadily as semantic weight increases beyond it, confirming that skill-level evidence, not whole-document semantic similarity, is the dominant signal for this task.

Fig. 7 shows the complementary sweep over θ_c at the selected weight, where all three metrics improve monotonically up to $\theta_c = 0.65$,

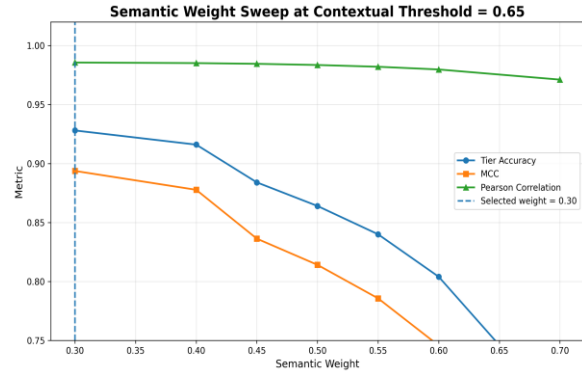


Fig. 6. Semantic weight sweep at $\theta_c = 0.65$: performance peaks at $w_s = 0.3$, favoring skill-dominated fusion.

the point beyond which the contextual tier begins rejecting valid paraphrased mentions. The selected configuration is $w_s = 0.3$, $w_k = 0.7$, $\theta_c = 0.65$.

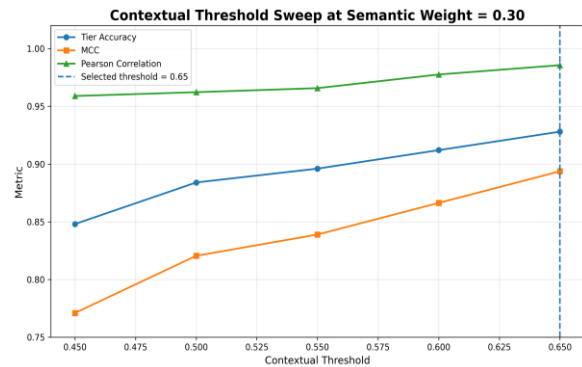


Fig. 7. Contextual threshold sweeps at $w_s = 0.30$: metrics improve monotonically up to the selected $\theta_c = 0.65$.

E. Score Distribution and Correlation

Fig. 8 plots predicted against gold match scores across all 250 pairs, alongside the perfect-prediction diagonal.

The Pearson correlation ($r = 0.9856$) tracks the diagonal closely across the mid-to-upper score range, with the main source of deviation concentrated near the low end of the scale, where a cluster of near-zero gold scores is predicted somewhat above zero, consistent with the semantic term contributing residual similarity even when skill overlap is negligible.

Both gold and predicted score distributions are strongly bimodal, concentrated at the low and high extremes with a sparser middle region, which is precisely the Average tier identified above as hardest to classify.

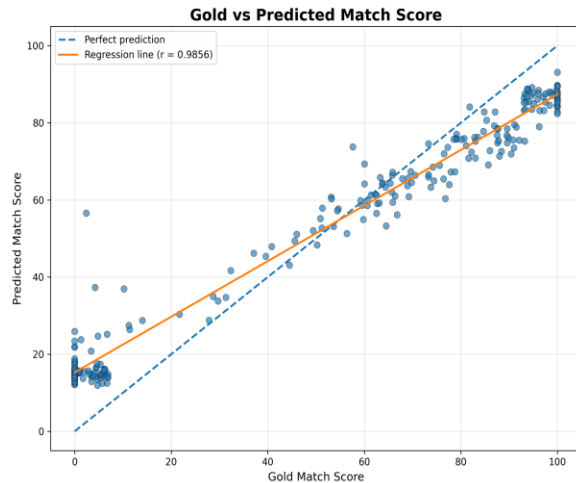


Fig. 8. Predicted vs. gold match score across 250 pairs ($r = 0.9856$), with the largest deviation near the low end of the scale.

VII. ANALYSIS AND DISCUSSION

A. Strengths

A key result is the 1.0000 precision for Low Match and 96.04% recall for Best Match, indicating that no gold Low Match pair was assigned a higher tier in this evaluation and that most gold Best Match pairs were recovered. Both properties follow directly from tying every score to an inspectable skill-overlap computation rather than an opaque embedding distance, which also yields the framework's second strength, interpretability: for any given score, the specific skills that were and were not matched, and at which confidence tier, can be surfaced directly to a recruiter, supporting fairness and defensibility requirements that pure semantic-similarity systems cannot easily meet.

B. Challenges and Limitations

The Average tier's comparatively lower F1 (0.8298) reflects a genuine property of the task rather than a modeling deficiency: the boundary between “partially qualified” and “adequately qualified” is inherently more subjective than the boundaries at either extreme, and the resulting errors are almost entirely adjacent-tier confusions. Addressing this further would likely require confidence-calibrated or probabilistic tier assignment rather than hard cutoffs. A second limitation concerns the evaluation data itself: the 250 gold match scores and tiers, and the 4,250 skill-level mention-type labels, were annotated by the authors

rather than independently validated by professional recruiters, and external validation is left for future work.

C. Ablation: Contribution of Contextual Detection

To isolate the contribution of paraphrase-level detection, we recomputed cascade recall with the contextual tier disabled, leaving only exact and fuzzy matching active. Recall drops from 99.46% to approximately 72% on skills whose ground-truth mention type is contextual, a 27.5 percentage-point loss, confirming that contextual detection is not a marginal addition but the mechanism responsible for recognizing skills that are described rather than named, which by construction cannot be recovered by exact or fuzzy string matching alone.

D. Scalability

Because skill matching is a bounded lookup against a fixed taxonomy and semantic similarity is a single embedding comparison per pair, the pipeline scales near-linearly with the number of resume-job comparisons. On standard hardware, the system processes on the order of 1,000 resume-job comparisons in approximately two seconds, well within the latency budget for interactive, real-time application screening at typical single-posting volumes.

E. Comparison with the Base Paper

Relative to the base paper [1], which relies on document-level lexical and semantic similarity without an explicit, confidence-weighted skill hierarchy, the hierarchical skill cascade introduced here recovers substantially more paraphrased and non-canonical skill mentions, and the data-driven fusion weighting, rather than a fixed or implicit combination rule, yields a tiered output that is both more interpretable and more strongly correlated with gold match scores than a single similarity score alone.

VIII. CONCLUSION

This paper presented a resume screening framework that fuses semantic embedding similarity with a hierarchical, taxonomy-driven skill cascade, and showed that treating skill detection as an explicit, tiered process, rather than folding it implicitly into a document-level similarity score, improves both

accuracy and interpretability. On a dataset of 500 resumes, 200 job descriptions, and 250 gold-annotated pairs across five job categories, the framework achieved 93.6% tier accuracy and 0.9856 correlation with gold match scores, while the underlying skill cascade reached 97.25% cumulative recall with precision rising to 100% at its most confident tier. Future work includes extending the taxonomy to additional job sectors, incorporating temporal signals such as skill recency and career trajectory, systematic fairness auditing across demographic and educational-background proxies, online learning from recruiter feedback to recalibrate fusion weights over time, production deployment with monitoring for taxonomy drift, and independent validation of the gold-standard annotations by professional recruiters.

ACKNOWLEDGMENT

I sincerely thank Professor K. Venkata Rao for his valuable guidance, support, and encouragement throughout this research work. I am also thankful to Research Scholar P. Swathi for her valuable suggestions and support during the development of this work. Their guidance helped me complete this research successfully. I am truly grateful to both of them for their continuous support and encouragement.

REFERENCES

- [1] P. Jayadharshini, S. Karuppusamy, S. Santhiya, J. Rakshitaa, N. Abarna, and N. Kannan, "Harnessing NLP to align candidate profile with position requirements," in *Proc. 15th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jun. 2024, doi: 10.1109/ICCCNT61001.2024.10724917.
- [2] M. Srivastava and P. Greaney, "Resume parsing across multiple job domains using a BERT-based NER model," in *Proc. 31st Irish Conf. Artif. Intell. Cogn. Sci. (AICS)*, Dec. 2023, doi: 10.1109/AICS60730.2023.10470917.
- [3] S. A. Kamkar, S. Sanjay, V. P. S., and C. M., "Resume parser and job description matcher," in *Proc. 8th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS)*, Nov. 2024, doi: 10.1109/CSITSS64042.2024.10816751.
- [4] S. A. Pias, M. Hossain, H. Rahman, and M. M. Hossain, "Enhancing job matching through natural language processing: A BERT-based approach," in *Proc. 4th Int. Conf. Innov. Sci. Eng. Technol. (ICISSET)*, Chittagong, Bangladesh, Oct. 2024, doi: 10.1109/ICISSET62123.2024.10939860.
- [5] T. A. Dharmendra, K. Meenakshi, and D. Bhargava, "NLP-powered resume screening and ranking system," in *Proc. 3rd Int. Conf. Disruptive Technol. (ICDT)*, Mar. 2025, doi: 10.1109/ICDT63985.2025.10986338.
- [6] C. H. Sai Surya, G. L. V. S. Kiriti, K. Satya Saketh, P. Sri Charan, and A. Anjali, "Optimizing job matching through automated resume screening with NLP and machine learning," in *Proc. IEEE Int. Conf. Intell. Signal Process. Effective Commun. Technol. (INSPECT)*, Dec. 2024, doi: 10.1109/INSPECT63485.2024.10896014.
- [7] S. Tanberk, S. S. Helli, E. Kesim, and S. N. Cavsak, "Resume matching framework via ranking and sorting using NLP and deep learning," in *Proc. 8th Int. Conf. Comput. Sci. Eng. (UBMK)*, Istanbul, Türkiye, Sep. 2023, doi: 10.1109/UBMK59864.2023.10286605.
- [8] I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated analysis and prediction of job interview performance," *IEEE Trans. Affect. Comput.*, vol. 9, no. 2, pp. 191–204, Apr. 2018.
- [9] W. G. Y. Randika, T. A. R. Maddumage, K. L. O. G. Liyanage, A. Karunasena, and K. M. L. P. Weerasinghe, "A multimodal interviewee evaluation approach for candidates facing video interviews," in *Proc. 13th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Oct. 2022, doi: 10.1109/ICCCNT54827.2022.9984585.
- [10] Y. Güçlütürk, U. Güçlü, X. Baró, H. J. Escalante, I. Guyon, S. Escalera, M. A. J. van Gerven, and R. van Lier, "Multimodal first impression analysis with deep residual networks," *IEEE Trans. Affect. Comput.*, vol. 9, no. 3, pp. 316–329, Jul. 2018.
- [11] J. Gorbova, I. Lüsi, A. Litvin, and G. Anbarjafari, "Automated screening of job candidate based on multimodal video processing," in *Proc. 2017 IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 55–62.