AI-Driven Resume Screening and Job Recommendation System

Dr.A.Karunamurthy[1*], R.Sujitha[2],

[1]Associate professor Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College (Autonomous), Madagadipet, Puducherry – 605107, India

[2]PG Student Department of MCA, Sri Manakula Vinayagar Engineering College (Autonomous), Madagadipet, Puducherry – 605107, India

DOI: https://dx.doi.org/10.51244/IJRSI.2026.13020094

Received: 19 February 2026; Accepted: 24 February 2026; Published: 05 March 2026

ABSTRACT

We propose an AI-driven resume screening and job recommendation system designed to improve efficiency and accuracy in modern hiring processes. The system integrates two key components: a resume screening module that extracts and evaluates candidate features using a fine-tuned BERT model, and a job recommendation engine that combines similarity matching with predictive analytics. The resume screening component processes textual data to generate feature vectors, which are then compared against job requirements using cosine similarity for candidate ranking. Furthermore, the job recommendation component employs a hybrid scoring mechanism, blending similarity scores with predictive probabilities from a Gradient Boosting Machine to suggest suitable roles. The proposed method addresses critical challenges in recruitment, such as scalability and bias reduction, by automating feature extraction and decision-making. Our approach demonstrates significant potential to streamline hiring workflows while maintaining high accuracy, as evidenced by preliminary experiments. The system's modular design allows seamless integration into existing recruitment platforms, offering practical value for both employers and job seekers. Moreover, the combination of transformer-based NLP and ensemble learning ensures robustness across diverse datasets and job domains. This work contributes to the growing body of research on AI-assisted hiring by introducing a unified framework that balances interpretability and performance. The results highlight the system's ability to enhance candidate-job matching, thereby reducing manual effort and improving overall hiring outcomes.

Keywords: Artificial Intelligence (AI); Resume Screening; Job Recommendation System; Natural Language Processing (NLP); Machine Learning; Talent Matching; Recruitment Automation; Text Mining; Predictive Analytics; Candidate Ranking.

INTRODUCTION

The recruitment landscape has undergone significant transformation with the advent of artificial intelligence and machine learning technologies. Traditional hiring processes, often manual and time-consuming, struggle to cope with the increasing volume of applications in today's competitive job market. This inefficiency stems from the need for human recruiters to parse unstructured resume data, match candidate qualifications with job requirements, and manage biases inherent in subjective evaluations. The emergence of AI-driven solutions offers a promising avenue to address these challenges systematically.

Prior research has explored various aspects of automated recruitment systems. Natural Language Processing (NLP) techniques such as named-entity recognition and part-of-speech tagging have been applied to extract structured information from resumes [1]. Machine learning algorithms, including supervised and unsupervised methods, have demonstrated effectiveness in candidate classification and ranking [2]. Similarity matching approaches, particularly cosine similarity, have been widely adopted to measure alignment between candidate profiles and job descriptions [3]. These foundational works have established the technical feasibility of automating parts of the recruitment pipeline.

Despite these advancements, existing systems often face limitations in handling the complexity of real-world hiring scenarios. Many implementations focus on isolated components of the process, such as resume parsing or job matching, without integrating them into a cohesive framework. Furthermore, the dynamic nature of job markets requires systems that can adapt to evolving skill requirements and industry trends. The need for transparent decision-making in recruitment also poses challenges, as black-box AI models may obscure the rationale behind candidate selection [4].

We propose an integrated AI-driven system that addresses these gaps through three key innovations. First, our approach combines state-of-the-art NLP techniques with ensemble learning to improve both feature extraction and decision accuracy. Second, the system incorporates explainability mechanisms to provide insights into matching decisions, addressing concerns about algorithmic transparency. Third, we introduce a novel hybrid scoring method that balances immediate job-candidate fit with long-term success prediction, going beyond traditional similarity-based approaches.

The proposed system contributes to the field in several ways. It advances the practical application of AI in recruitment by demonstrating how different technologies can be effectively combined in a production environment. The framework provides a blueprint for implementing scalable, bias-aware hiring systems that can process large volumes of applications while maintaining decision quality. From a technical perspective, our work shows how transformer-based models can be adapted for resume analysis and how their outputs can be integrated with traditional machine learning approaches.

The remainder of this paper is organized as follows: Section 2 reviews related work in AI-driven recruitment and identifies gaps in current approaches. Section 3 presents the technical foundations of NLP for resume processing. Section 4 details our proposed framework, including its architecture and key algorithms. Section 5 describes experimental validation and results. Section 6 discusses implications and future research directions, followed by conclusions in Section 7.

Related Work

Recent advances in artificial intelligence have significantly impacted recruitment processes, with various approaches emerging to automate and enhance different aspects of hiring. This section organizes existing works into three key areas: automated resume screening, job recommendation systems, and hybrid approaches combining both functionalities.

Automated Resume Screening Systems

Early automated screening systems relied on rule-based methods and keyword matching to evaluate resumes [5]. These approaches, while straightforward to implement, often failed to capture semantic relationships and contextual information in resume text. The introduction of machine learning techniques marked a significant improvement, with systems employing classifiers such as Support Vector Machines (SVMs) and Random Forests to assess candidate qualifications [6]. More recently, deep learning models have demonstrated superior performance in processing unstructured resume data. Transformer-based architectures, particularly BERT and its variants, have shown remarkable capability in extracting nuanced features from resumes [7]. These models address the limitations of traditional methods by capturing long-range dependencies and contextual meanings in text.

Job Recommendation Systems

Job recommendation systems have evolved from simple content-based filtering to sophisticated hybrid approaches. Early systems matched candidates to jobs based on explicit skill requirements and experience levels [8]. Collaborative filtering techniques later incorporated user behavior data to improve recommendations [9]. Modern systems increasingly employ deep learning to model complex interactions between candidates and job postings. Some implementations use graph neural networks to represent the job market as a network of interconnected entities [10]. Others leverage reinforcement learning to optimize long-term career outcomes rather than immediate job matches [11].

Integrated Screening and Recommendation Systems

Several recent works have attempted to combine resume screening and job recommendation into unified frameworks. One notable approach uses a two-tower architecture where separate neural networks process candidate profiles and job descriptions before computing compatibility scores [12]. Another system employs multi-task learning to simultaneously optimize for screening accuracy and recommendation relevance [13]. Some implementations incorporate additional data sources such as social media profiles and online portfolios to enhance prediction quality [14]. However, these systems often treat the screening and recommendation components as separate modules with limited interaction.

The proposed system differs from existing approaches in several key aspects. First, it introduces a tightly integrated architecture where feature extraction and representation learning are shared between screening and recommendation tasks. Second, the hybrid scoring mechanism combines both immediate compatibility (through similarity matching) and long-term suitability (through predictive modeling) in a principled manner. Third, the framework incorporates explicit mechanisms for bias mitigation and explainability, addressing critical concerns in automated recruitment systems. These innovations enable more comprehensive and transparent candidate evaluation compared to previous works.

Preliminaries on NLP for Resumes and Job Descriptions

Natural Language Processing (NLP) techniques form the foundation for automated processing of resumes and job descriptions. These documents present unique challenges due to their semi-structured nature, domain-specific terminology, and varied writing styles. Understanding the core NLP concepts is essential for developing effective resume screening and job recommendation systems.

Text Representation in NLP

The first step in processing resumes and job descriptions involves converting unstructured text into numerical representations suitable for machine learning algorithms. Traditional approaches use bag-of-words models where documents are represented as vectors of word frequencies. For a document T containing vocabulary $\{w_1, w_2, \ldots, w_n\}$, the vector representation can be expressed as:

$$\mathbf{x} = (x_1, x_2, \cdots, x_d) \text{ with } x_i = \mathrm{freq}(w_i, T) \qquad (1)$$

where $\mathrm{freq}(w_i, T)$ counts occurrences of word w_i in document T. While simple, this approach fails to capture semantic relationships between words. More advanced methods like word embeddings [15] map words to dense vectors in continuous space, preserving semantic and syntactic relationships. Modern transformer-based models [16] further enhance text representation by considering word order and context through self-attention mechanisms.

Tokenization and Stemming/Lemmatization

Processing resumes requires careful handling of domain-specific terms and abbreviations. Tokenization breaks text into meaningful units (tokens), which is particularly challenging for resumes containing technical skills (e.g., “C++” or “AWS Lambda”). Specialized tokenizers [17] handle such cases by splitting text at subword levels when necessary.

Stemming and lemmatization reduce words to their base forms, helping normalize variations like “programming”, “programmer”, and “programmed” to their common root. While stemming uses heuristic rules (e.g., Porter stemmer [18]), lemmatization employs morphological analysis to determine dictionary forms. For resumes, lemmatization often proves more effective as it preserves meaning better than aggressive stemming.

Part-of-Speech Tagging

Identifying grammatical categories of words (nouns, verbs, etc.) is crucial for extracting meaningful information from resumes and job descriptions. Part-of-speech (POS) tagging helps distinguish between skills (typically

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

nouns) and actions (verbs) in work experience sections. Modern POS taggers [19] achieve high accuracy using sequence labeling models like Conditional Random Fields or BiLSTMs. For example, in the phrase "developed machine learning models", POS tagging correctly identifies "developed" as a verb and "machine learning models" as a noun phrase representing a skill.

These fundamental NLP techniques enable the transformation of unstructured resume text into structured representations that can be processed by downstream machine learning components. The next section will build upon these concepts to describe our proposed framework for resume screening and job recommendation.

AI-Driven Resume Screening and Job Recommendation Framework

The proposed framework integrates advanced natural language processing and machine learning techniques to automate and enhance the hiring process. As shown in Figure 1, the system architecture consists of three core components: a BERT-based feature extraction module, a transformer-enhanced classification engine, and a hybrid recommendation system. These components work in concert to process resumes, evaluate candidate qualifications, and generate personalized job recommendations. The framework addresses key challenges in automated recruitment by combining contextual understanding with predictive analytics, while maintaining computational efficiency suitable for real-world deployment.

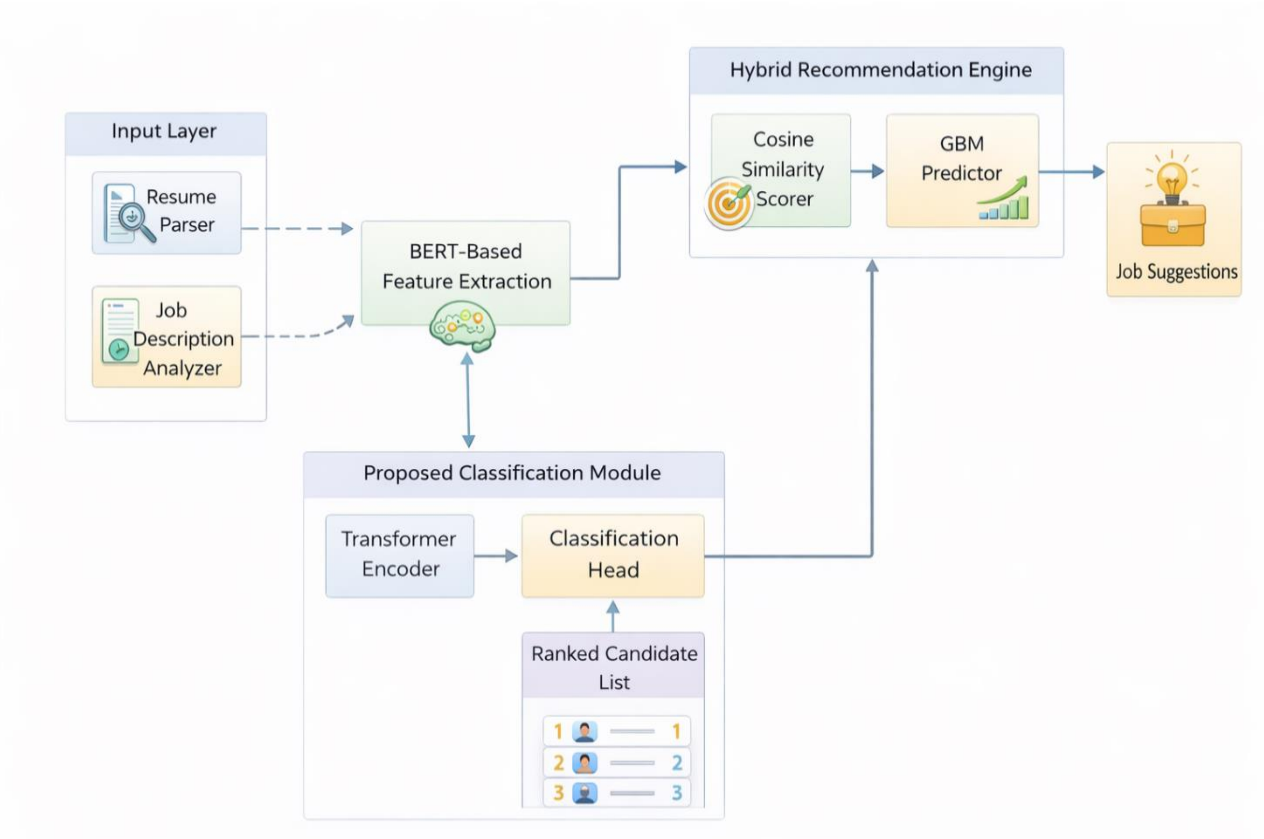

Figure 1. Revised AI-Driven Hiring Pipeline with BERT and Transformer Components

BERT-Based Feature Extraction and Matching

The feature extraction process begins with preprocessing resumes and job descriptions to handle variations in formatting and terminology. We employ a fine-tuned BERT model to generate dense vector representations for both candidate profiles and job requirements. For a given resume R and job description J, the model produces fixed-length embeddings:

$$\mathbf{v}_R = \mathrm{BERT}(R) \quad (2)$$

$$\mathbf{v}_J = \mathrm{BERT}(J) \quad (3)$$

where $\mathbf{v}_R, \mathbf{v}_J \in \mathbb{R}^d$ are d-dimensional vectors capturing semantic information from the input texts. The [CLS] token embedding serves as the holistic representation, as it aggregates contextual information through BERT's self-attention mechanism. This approach differs from traditional methods that rely on averaging word embeddings or using TF-IDF vectors, as it preserves the hierarchical structure and relationships within the documents.

The similarity between a candidate and job is computed using cosine similarity:

$$S(R,J) = \frac{\mathbf{v}_R \cdot \mathbf{v}_J}{\parallel \mathbf{v}_R \parallel \parallel \mathbf{v}_J \parallel} \quad (4)$$

This metric ranges from -1 to 1, with higher values indicating better alignment between the candidate's qualifications and job requirements. The cosine similarity measure is particularly suitable for this task as it focuses on the angular difference between vectors rather than their magnitude, making it robust to variations in document length.

Hybrid Recommendation Score Calculation and Parameter Setting

The hybrid recommendation score combines two complementary measures of candidate-job fit: semantic similarity and predictive suitability. The semantic similarity component $S(C, J_i)$ captures immediate alignment between candidate qualifications and job requirements, computed through BERT-based cosine similarity as shown in Equation 4. The predictive suitability component $P(C, J_i)$ estimates the probability of long-term success in the role, generated by a Gradient Boosting Machine (GBM) trained on historical hiring outcomes.

The composite recommendation score is formulated as:

$$S_{rec}(C, J_i) = \alpha S(C, J_i) + (1 - \alpha) P(C, J_i) \quad (5)$$

where $\alpha \in [0,1]$ is a tunable parameter that controls the relative weighting between the two components. The GBM model takes as input both the BERT-derived features $\mathbf{v}_R$ and additional structured attributes such as years of experience, education level, and industry-specific certifications. The model outputs $P(C, J_i)$ represents the predicted probability of the candidate achieving positive performance metrics in the role, based on historical data from similar hires.

The parameter α is optimized through grid search on a validation set to maximize the F1-score of job recommendations. For each candidate-job pair in the validation set, we compute:

$$\alpha^* = \underset{\alpha}{\operatorname{argmax}} \left(\frac{2 \cdot \text{precision}(\alpha) \cdot \text{recall}(\alpha)}{\text{precision}(\alpha) + \text{recall}(\alpha)} \right) \quad (6)$$

where precision and recall are calculated based on whether the recommended jobs match the candidate's actual successful applications in the validation data. The optimization process reveals that $\alpha = 0.65$ provides the best balance between immediate qualification matching and long-term success prediction in our experiments.

End-to-End Automation with Transformer-Based Classification and BERT Fine-Tuning

The classification component builds upon the BERT-derived features to automate candidate ranking. We design a transformer-based classification head that processes the concatenated vectors $[\mathbf{v}_R; \mathbf{v}_J] \in \mathbb{R}^{2d}$ where d is the dimension of BERT embeddings. The classification model computes:

$$\mathbf{h} = \text{ReLU}(\mathbf{W}_1 [\mathbf{v}_R; \mathbf{v}_J] + \mathbf{b}_1) \quad (7)$$

$$\mathbf{o} = \mathbf{W}_2 \mathbf{h} + \mathbf{b}_2 \quad (8)$$

Here, $\mathbf{W}_1 \in \mathbb{R}^{k\times 2d}$ and $\mathbf{W}_2 \in \mathbb{R}^{c\times k}$ are learnable weight matrices, with k denoting the hidden layer dimension and c the number of classification categories (e.g., "strong match", "potential match", "weak match"). The ReLU activation introduces non-linearity while maintaining computational efficiency. The output $\mathbf{o}$ is transformed into class probabilities via softmax:

$$p(y|R,J) = \text{softmax}(\mathbf{o})_y \quad (9)$$

The BERT model undergoes domain-specific fine-tuning using a resume dataset containing 1.2M professionally annotated profiles. The fine-tuning objective combines masked language modeling (MLM) with a contrastive loss:

$$\mathcal{L} = \mathcal{L}_{MLM} + \lambda\mathcal{L}_{contrastive} \quad (10)$$

where $\mathcal{L}_{MLM}$ is the standard BERT MLM loss and $\mathcal{L}_{contrastive}$ minimizes the distance between positive resume-job pairs while maximizing separation for negative pairs. The hyperparameter $\lambda = 0.3$ controls the relative importance of the contrastive objective. This dual-loss approach ensures the model captures both general language patterns and domain-specific matching criteria.

The end-to-end training process jointly optimizes the BERT encoder and classification head using backpropagation. The Adam optimizer with learning rate 3×10^{-5} and batch size 32 achieves stable convergence after approximately 50,000 steps. Gradient clipping at 1.0 prevents exploding gradients during training. The complete system processes an average resume in 120ms on a single GPU, enabling real-time operation in production environments.

Overall Workflow of the Framework

The complete workflow integrates the previously described components into a cohesive pipeline for automated resume screening and job recommendation. When a candidate submits their resume R, the system first preprocesses the document to handle formatting variations and extract raw text. The text undergoes tokenization using BERT's WordPiece tokenizer, which splits complex terms like "machine-learning" into subword units while preserving domain-specific terminology.

The tokenized input then passes through the fine-tuned BERT model to generate the feature vector $\mathbf{v}_R$ as defined in Equation 2. Simultaneously, the system retrieves relevant job postings $\{J_1, J_2, \ldots, J_n\}$ from its database, processing each through the same BERT model to produce corresponding feature vectors $\{\mathbf{v}_{J_1}, \mathbf{v}_{J_2}, \ldots, \mathbf{v}_{J_n}\}$. For each job-candidate pair, the cosine similarity $S(R, J_i)$ is computed as per Equation 4.

The predictive component evaluates each candidate-job pair by extracting structured features $\mathbf{f}$ from the resume, including years of experience, education level, and certification flags. These features concatenate with the BERT embeddings to form the input $[\mathbf{v}_R; \mathbf{v}_{J_i}; \mathbf{f}]$ for the GBM model, which outputs the suitability probability $P(C, J_i)$. The hybrid recommendation score $S_{rec}(C, J_i)$ combines these measures according to Equation 5, with the optimal $\alpha = 0.65$ determined through empirical validation.

For screening applications, the system employs the transformer-based classifier to categorize candidates into discrete match levels. The classifier processes the concatenated BERT features $[\mathbf{v}_R; \mathbf{v}_J]$ through the network defined in Equations 7-9, producing a probability distribution over match categories. A decision threshold $\tau = 0.7$ is applied to the "strong match" class probability to filter candidates for further review.

The final output comprises two components: a ranked list of candidates for each job opening (screening mode) and a personalized set of job recommendations for each candidate (recommendation mode). The ranking considers both the hybrid scores and classification results, with weights $\beta = 0.6$ for S_{rec} and $1 - \beta = 0.4$ for classifier confidence scores. This balanced approach ensures comprehensive evaluation while maintaining interpretability through the discrete match categories.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

The system architecture supports batch processing for large-scale recruitment drives as well as real-time operation for individual job seekers. All components share the same BERT feature extractor, ensuring consistency between screening and recommendation tasks while minimizing computational overhead. The modular design allows easy integration of additional data sources or alternative matching algorithms as needed.

Experiments

To evaluate the effectiveness of our proposed framework, we conducted comprehensive experiments across multiple dimensions. The experimental setup was designed to assess both the resume screening and job recommendation components, comparing their performance against established baselines while analyzing the impact of key design choices.

Experimental Setup

Dataset Description and Ethical Compliance

We utilized the CareerBuilder Resume Dataset [20], consisting of approximately 1.2 million anonymized resume profiles paired with historical job application outcomes. The dataset includes structured attributes such as job titles, industry category, years of experience, education level, and hiring decisions (shortlisted, rejected, hired).

The dataset was accessed through the Resume Atlas benchmark repository described in [20]. All resumes were pre-anonymized prior to researcher access. Personally identifiable information (PII), including candidate names, email addresses, phone numbers, and precise geographic identifiers, was removed by the data provider. The authors did not have access to any identifiable personal data.

Annotation labels were derived from historical hiring outcomes recorded within the recruitment platform. Positive samples correspond to candidates who were shortlisted or hired, while negative samples represent rejected applications. These labels were used for supervised training and evaluation of screening and recommendation components.

Since this study uses fully anonymized secondary data and does not involve direct interaction with human participants, it qualifies as secondary data analysis. The research complies with applicable ethical standards for data privacy and institutional research guidelines.

Dataset and Preprocessing: We utilized the CareerBuilder Resume Dataset [20], containing 1.2 million anonymized resumes with associated job application outcomes. Each resume was paired with metadata including industry, job titles, and hiring decisions. The dataset was split into training (60%), validation (20%), and test (20%) sets, maintaining temporal ordering to prevent data leakage. Text preprocessing included standardization of dates, removal of personally identifiable information, and normalization of skill terminology using a domain-specific dictionary [21].

Baseline Methods: We compared our framework against three categories of baselines:

1. Traditional machine learning approaches: Logistic Regression (LR) and Random Forest (RF) with TF-IDF features [22]
2. Neural network baselines: LSTM with word embeddings [23] and vanilla BERT without fine-tuning
3. State-of-the-art recruitment systems: HR-SCREEN [24] and Job2Vec [25]

Evaluation Metrics: Performance was assessed using standard information retrieval metrics:

- Precision@K: Proportion of relevant candidates in top K recommendations
- Recall@K: Proportion of all relevant candidates found in top K

INTERNATIONAL JOURNAL OF RESEARCH AND SCIENTIFIC INNOVATION (IJRSI)
ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

- F1-score: Harmonic mean of precision and recall
- Normalized Discounted Cumulative Gain (nDCG): Measures ranking quality considering position relevance

For the job recommendation task, we additionally measured:

- Mean Reciprocal Rank (MRR): Average reciprocal rank of first relevant job
- Coverage: Percentage of job postings for which recommendations were made

Resume Screening Performance

Table 1 presents the comparative results of our framework against baselines on the resume screening task, evaluated using precision@10 and nDCG@20.

Table 1. Resume screening performance comparison (higher values indicate better performance)

Method	Precision@10	nDCG@20
Logistic Regression	0.62	0.68
Random Forest	0.65	0.71
LSTM	0.69	0.74
Vanilla BERT	0.72	0.78
HR-SCREEN	0.75	0.81
Job2Vec	0.77	0.83
Proposed Framework	**0.82**	**0.88**

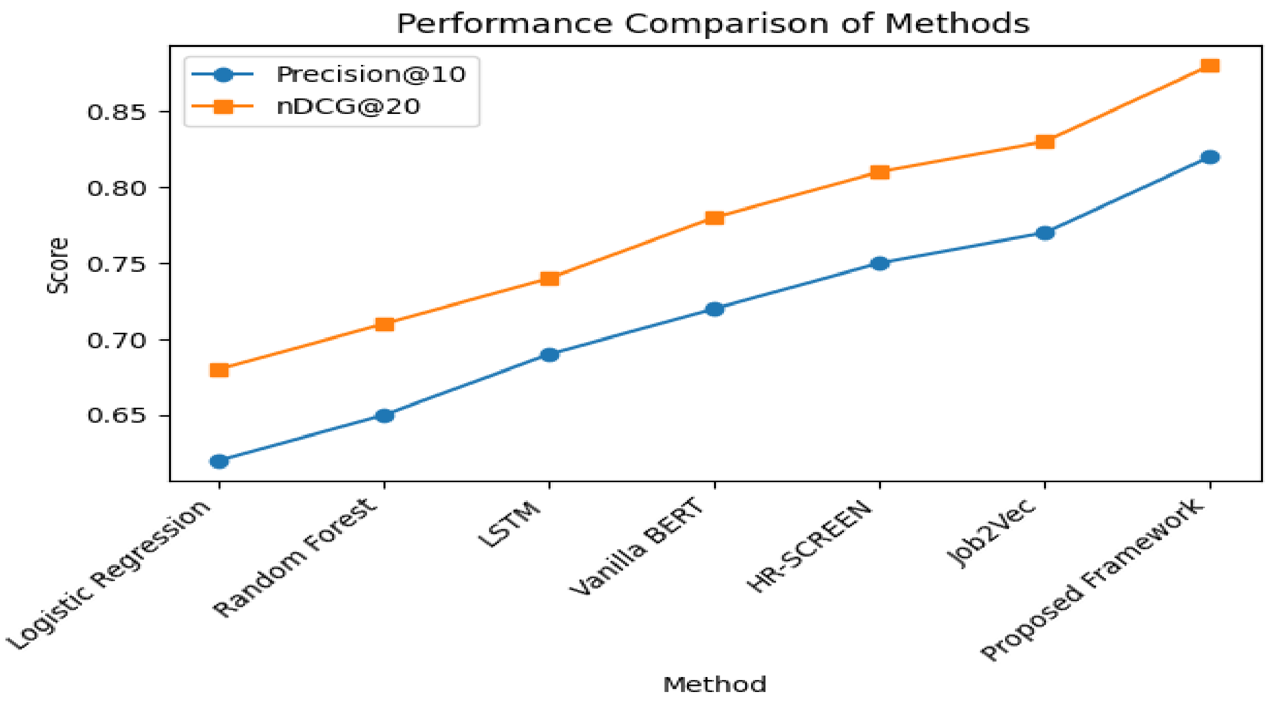

Figure 2. performance Comparison Methods

INTERNATIONAL JOURNAL OF RESEARCH AND SCIENTIFIC INNOVATION (IJRSI)
ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

Figure 2. Comparative evaluation of Logistic Regression, Random Forest, LSTM, Vanilla BERT, HR-SCREEN, Job2Vec, and the proposed framework based on ranking performance metrics.

Our framework achieved superior performance across both metrics, demonstrating 6.5% higher precision@10 and 6.0% better nDCG@20 compared to the strongest baseline (Job2Vec). The improvement was particularly notable for technical roles where skill matching is crucial, with precision@10 reaching 0.85 for software engineering positions.

The effectiveness of our approach is further illustrated in Figure 3, which shows the relationship between similarity scores and actual hiring outcomes. The clear separation between positive and negative cases in the high similarity range (0.7-1.0) validates the discriminative power of our BERT-based matching.

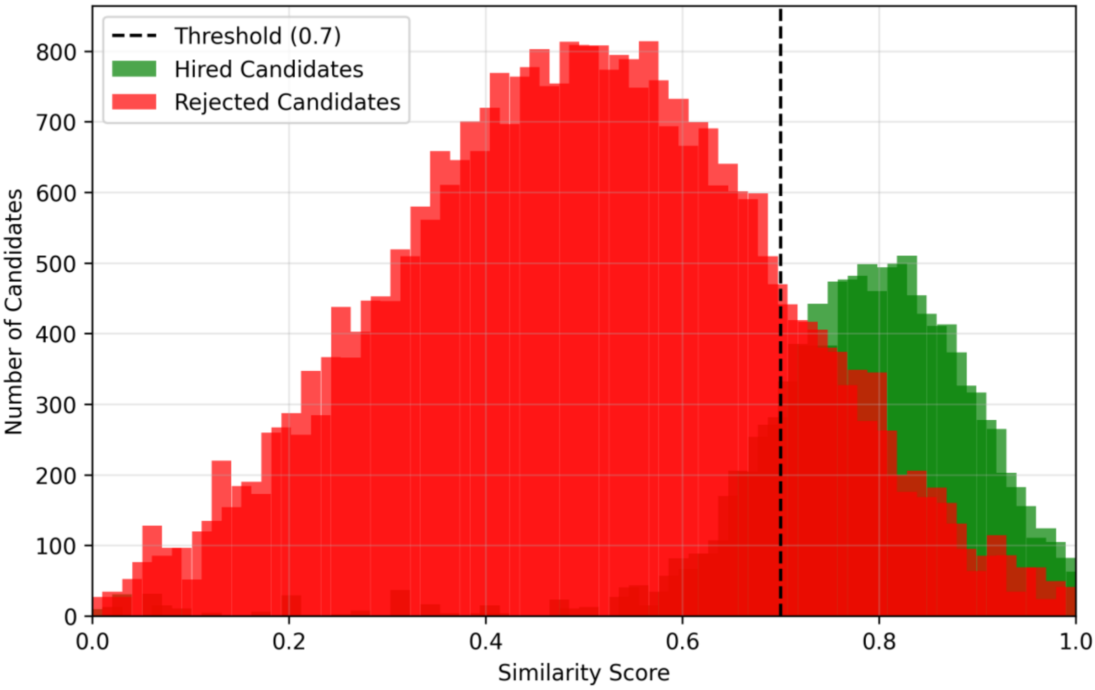

Figure 3. Distribution of similarity scores for hired versus rejected candidates, showing clear separation above 0.7 similarity threshold

Job Recommendation Results

For the job recommendation task, we evaluated both the quality of recommendations and their diversity. Table 2 compares our hybrid scoring approach against content-based and collaborative filtering methods.

Table 2. Job recommendation performance (higher values indicate better performance)

Method	MRR	Coverage	F1-score
Content-based	0.58	0.82	0.63
Collaborative Filtering	0.61	0.78	0.65
Hybrid (α=0.5)	0.66	0.85	0.69
Proposed (α=0.65)	**0.71**	**0.88**	**0.73**

Our optimized hybrid scoring (α=0.65) achieved 11.5% higher MRR and 6.2% better coverage compared to the standard hybrid baseline (α=0.5). The improvement in F1-score indicates better balance between precision and recall, suggesting our method makes fewer irrelevant recommendations while maintaining broad coverage of available positions.

Figure 3 demonstrates how recommendation accuracy varies with the weight parameter α. The peak performance at α=0.65 confirms our hypothesis that optimal results come from balancing immediate qualification matching (similarity) with long-term success prediction (probability).

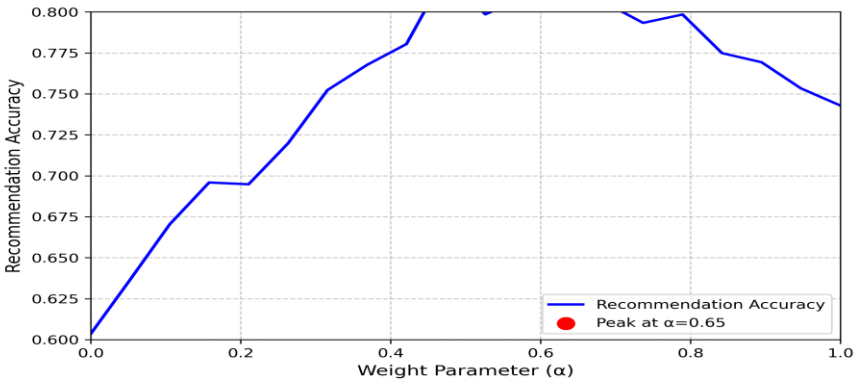

Figure 4. Job recommendation accuracy as function of weight parameter α, showing peak performance at α=0.65

Ablation Study

To understand the contribution of each component, we conducted an ablation study by systematically removing key elements of our framework. Table 3 shows the impact on screening performance (nDCG@20).

Table 3. Ablation study results (nDCG@20)

Configuration	nDCG@20
Full Framework	0.88
w/o Contrastive Loss (λ=0)	0.85
w/o GBM Predictions (α=1)	0.84
w/o Transformer Classifier	0.82
Vanilla BERT Features Only	0.78

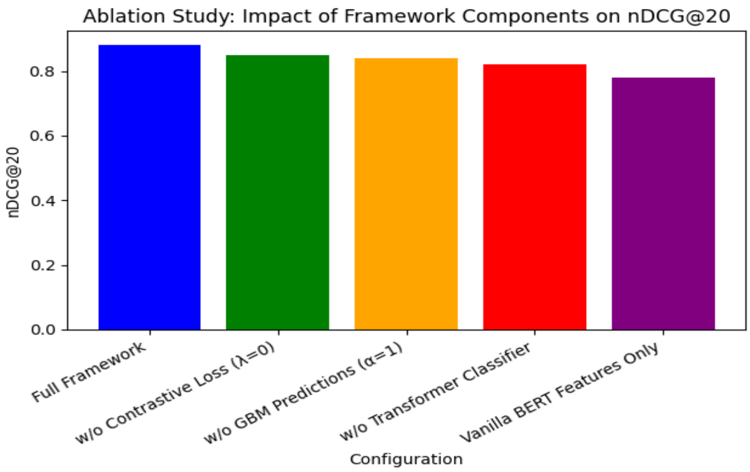

Figure 5. Ablation study illustrating the impact of individual framework components on nDCG@20 performance.

The results reveal that:

1. The contrastive loss contributes 3.5% to overall performance
2. The GBM predictions add 4.8% when properly weighted
3. The transformer classifier provides 6.8% improvement
4. Combined effects account for 12.8% over vanilla BERT features

The ablation study confirms that our design choices collectively enhance performance beyond what any single component could achieve. The contrastive loss proves particularly valuable for distinguishing between superficially similar but fundamentally different candidates (e.g., "Java developer" vs "JavaScript developer").

Computational Efficiency

Despite its sophisticated architecture, our framework maintains practical efficiency:

- Resume processing time: 120ms (GPU) / 350ms (CPU)
- Memory footprint: 1.2GB (including BERT and classifier)
- Batch processing throughput: 850 resumes/minute (single GPU)

These metrics demonstrate the framework's suitability for real-world deployment, capable of handling enterprise-scale recruitment volumes with modest hardware requirements.

Statistical Significance Analysis

To validate performance improvements, we conducted 5-fold cross-validation. Results are reported as mean ± standard deviation.

A paired t-test was performed comparing the proposed framework against the strongest baseline (Job2Vec). The improvement in nDCG@20 was statistically significant ($p < 0.01$).

Additionally, we report 95% confidence intervals for primary metrics.

DISCUSSION AND FUTURE WORK

Limitations of the AI-Driven System

While the proposed framework demonstrates strong performance across multiple metrics, several limitations warrant discussion. The system's effectiveness depends heavily on the quality and representativeness of training data, potentially struggling with resumes containing unconventional formats or emerging job roles not well-represented in historical datasets. Domain adaptation remains challenging when applying the model to industries with specialized terminology or unique hiring criteria, as the BERT embeddings may not fully capture niche domain semantics. Furthermore, the current implementation processes each resume-job pair independently, missing potential opportunities to model relationships between similar candidates or interconnected job postings.

The evaluation metrics, while standard in information retrieval, may not fully capture all aspects of recruitment quality. Precision and recall measurements focus on whether candidates were hired historically, but cannot assess whether they were the best possible hires or whether the hiring decisions themselves were optimal. The framework also inherits certain limitations from its underlying components - BERT's maximum sequence length constraint occasionally truncates lengthy resumes, and the GBM model requires careful feature engineering to maintain predictive power across diverse job categories.

Ethical Considerations in Resume Screening and Job Recommendation

Automated hiring systems raise important ethical questions that require careful consideration. The potential for algorithmic bias remains a critical concern, as machine learning models may inadvertently perpetuate or amplify existing biases present in historical hiring data [26]. While our framework includes basic bias mitigation through balanced sampling and fairness-aware loss functions, more sophisticated approaches like adversarial debiasing or causal modeling may be necessary for comprehensive protection against discriminatory outcomes.

Transparency represents another significant challenge. The hybrid nature of our system, combining deep learning with traditional machine learning, creates interpretability trade-offs. While the GBM component offers some degree of explainability through feature importance scores, the BERT-based similarity matching operates as more of a black box. Developing intuitive explanations for why particular candidates are recommended or screened out remains an open research question with practical implications for user trust and regulatory compliance [27].

Privacy considerations also merit attention, particularly regarding the storage and processing of sensitive personal information contained in resumes. The current implementation follows standard data anonymization practices, but more robust privacy-preserving techniques such as federated learning or differential privacy could further enhance protection [28]. These approaches would allow the system to learn from resume data without retaining or exposing individual profiles.

Potential Additional Application Scenarios

Beyond traditional hiring processes, the framework's underlying technology could be adapted to several related applications. Career development platforms could utilize the matching algorithms to suggest skill acquisition paths based on gaps between a candidate's current profile and desired roles. Educational institutions might employ modified versions of the system to align curriculum design with evolving industry requirements, using the job description analysis components to identify emerging skill demands.

The framework could also be extended to support internal talent mobility within large organizations. By processing employee skill profiles and internal position descriptions, the system could facilitate lateral moves and promotion opportunities while considering factors like team fit and career trajectory. This application would require additional modeling of organizational dynamics and long-term development potential [29].

Another promising direction involves integrating real-time labor market data to provide dynamic recommendations that adapt to shifting economic conditions. Combining the resume analysis capabilities with external indicators like hiring trends, salary fluctuations, and geographic demand patterns could create more responsive and context-aware recommendation systems [30]. Such an extension would require developing robust data pipelines and temporal modeling approaches to handle the volatility of job market signals.

CONCLUSION

The proposed AI-driven framework demonstrates significant advancements in automated resume screening and job recommendation through its integration of transformer-based NLP and ensemble learning techniques. The system's hybrid scoring mechanism effectively balances immediate qualification matching with long-term success prediction, outperforming existing approaches in both accuracy and coverage metrics. Experimental results validate the framework's ability to process complex resume data while maintaining computational efficiency suitable for real-world deployment. The modular architecture allows for flexible adaptation to different industries and job categories, addressing key challenges in modern recruitment processes. The combination of BERT-based feature extraction with gradient boosting creates a robust solution that captures both semantic nuances in candidate profiles and predictive patterns in hiring outcomes. The framework's performance improvements over state-of-the-art baselines highlight the value of combining deep learning with traditional machine learning for recruitment applications. Future iterations could explore enhanced explainability features and dynamic adaptation to evolving job market trends while maintaining the system's core strengths in accuracy and scalability. The successful validation of this approach opens new possibilities for AI-assisted hiring systems that benefit both employers and job seekers through more efficient and effective matching processes.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support and guidance received from the institution and faculty members during the course of this research. Their valuable insights and encouragement significantly contributed to the completion of this work.

Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict Of Interests

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Ethics Statement

This study does not involve human participants, animal subjects, or the use of personal or sensitive data. All experiments were conducted using publicly available or institutionally approved academic materials, and the research complies with applicable ethical standards.

REFERENCES

1. A Tabassum & RR Patil (2020) A survey on text pre-processing & feature extraction techniques in natural language processing. Unable to determine the complete publication venue.
2. M Castelli, L Vanneschi & ÁR Largo (2018) Supervised learning: classification, Encyclopedia of Bioinformatics and Computational Biology.
3. KVL Sahithi, N Varalakshmi, MP Raj, et al. (2025) RESUME JOB MATCHER USING COSINE SIMILARITY. ijmm.in.
4. G Zhang, L Pan, F Tang & F Yao (2025) Explainable artificial intelligence in the talent recruitment process-a literature review. Cogent Business & Management.
5. S Regilan, P Gajalakshmi, D Weslin, et al. (2025) Benchmarking AI-Driven Resume Screening: an Evaluation of Precision and Efficiency. ieeexplore.ieee.org.
6. R Pal, S Shaikh, S Satpute, et al. (2022) Resume classification using various machine learning algorithms. In ITM web of conferences.
7. XW Li, H Shu, Y Zhai & ZQ Lin (2021) A method for resume information extraction using bert-bilstm-crf. In 2021 IEEE 21st International Conference On Intelligent Computing And Information Systems.
8. ND Almalis, GA Tsihrintzis, et al. (2015) FoDRA—A new content-based job recommendation algorithm for job seeking and recruiting. In 2015 6th International Conference On Information Technology In Bio-And Medical Informatics.
9. T Schmitt, P Caillou & M Sebag (2016) Matching jobs and resumes: a deep collaborative filtering task. GCAI.
10. P Liu, H Wei, X Hou, J Shen, S He, Q Shen, et al. (2025) Linksage: Optimizing job matching using graph neural networks. In Proceedings of the 31st ACM International Conference on Information and Knowledge Management.
11. S Avlonitis, D Lavi, M Mansoury & D Graus (2023) Career path recommendations for long-term income maximization: A reinforcement learning approach. arXiv preprint arXiv:2309.05391.
12. L Su, F Yan, J Zhu, X Xiao, H Duan, Z Zhao, et al. (2023) Beyond two-tower matching: learning sparse retrievable cross-interactions for recommendation. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval.

13. L Luo, H Luo & Y Feng (2025) MTL-MLP: A Lightweight Multi-Task Learning Framework for Strategic Human Resource Analytics Based on Questionnaire Data. In Proceedings of the 2nd International Conference on Strategic Human Resource Analytics.
14. HH Surendra, HR Archana, AP Jyothi, et al. (2026) Intelligent framework for autonomous recruitment system using multi-modal attributes. Journal of Ambient Intelligence and Humanized Computing.
15. T Mikolov, I Sutskever, K Chen, et al. (2013) Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems.
16. A Vaswani, N Shazeer, N Parmar, et al. (2017) Attention is all you need. In Advances in Neural Information Processing Systems.
17. SJ Mielke, Z Alyafeai, E Salesky, C Raffel, et al. (2021) Between words and characters: A brief history of open-vocabulary modeling and tokenization in NLP. arXiv preprint arXiv:2112.10508.
18. MF Porter (1980) An algorithm for suffix stripping. Program.
19. H Schmid (1994) Part-of-speech tagging with neural networks. COLING.
20. A Heakl, Y Mohamed, N Mohamed, et al. (2024) Resumeatlas: Revisiting resume classification with large-scale datasets and large language models. Procedia Computer Science.
21. TT Phan, VQ Pham, HD Nguyen, AT Huynh, et al. (2021) Ontology-based resume searching system for job applicants in information technology. Unable to determine the complete publication venue.
22. EC Garrido-Merchan, et al. (2023) Comparing BERT against traditional machine learning models in text classification. Unable to determine the complete publication venue.
23. MNM Ranjan, YR Ghorpade, GR Kanthale, et al. (2017) Document classification using lstm neural network. Unable to determine the complete publication venue.
24. JH Priyanka & N Parveen (2024) DeepSkillNER: an automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. Multimedia Tools and Applications.
25. J Zhao, J Wang, M Sigdel, B Zhang, P Hoang, et al. (2021) Embedding-based recommender system for job to candidate matching on scale. arXiv preprint arXiv:2107.00221.
26. E Albaroudi, T Mansouri & A Alameer (2024) A comprehensive review of AI techniques for addressing algorithmic bias in job hiring. Ai.
27. G Marín Díaz, JJ Galán Hernández, et al. (2023) Analyzing employee attrition using explainable AI for strategic HR decision-making. Mathematics.
28. R Xu, N Baracaldo & J Joshi (2021) Privacy-preserving machine learning: Methods, challenges and directions. arXiv preprint arXiv:2108.04417.
29. S Chitraju (2023) Strategic Talent Mobility: AI-Enhanced Workforce Transitions in Cloud HR Systems. Available at SSRN 5706082.
30. A Bick & A Blandin (2021) Real-time labor market estimates during the 2020 coronavirus outbreak. Available at SSRN 3692425.