Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 4, 2026

SMART RECRUITMENT: CATEGORIZING, AND SELECTING APPROPRIATE AND UNIQUE TALENT USING POWER OF ARTIFICIAL INTELLIGENCE

Fozia Noureen[*1], Safina Soomro[2], Sarvat Naz[3], Faiza Mehreen[4]

[*1]*Department of Software Engineering, Quaid E Awam University of Engineering Science and Technology Nawabsah.*
[2]*Department of Computer Engineering, Sir Syed University of Engineering and Technology Karachi*
[3]*Department of Computer Science, Govt. Girls Degree College Sukkur*
[4]*Shaheed Zulfiqar Ali Bhutto University of Law SZABUL, Karachi*

[*1]engrnoureen@quest.edu.pk, [2]ssoomro@ssuet.edu.pk, [3]ap.sarvat@gmail.com, [4]faiza.mehreen@szabul.edu.pk

DOI: https://doi.org/10.5281/zenodo.19939230

Keywords
Smart Recruitment, Artificial Intelligence, NLP, CNN

Article History
Received: 03 March 2026
Accepted: 13 April 2026
Published: 30 April 2026

Corresponding Author: *
Fozia Noureen

Abstract
The use of artificial intelligence (AI) and machine learning (ML) is transmuting the recruitment process. Present and traditional recruitment procedures are based on screening of resume and rudimentary organized interviews, which primarily unable to provide a complete view of the candidate. In this paper, we propose an AI-driven multimodal recruitment system that modernizes the recruitment process through video, text, and speech-based behavioral analysis of the candidates' performance. Our system allows applicants to apply for vacant job on the portal, answer questionnaires, and undergo video interviews. The data is preprocessed using cleaning, feature extraction, and then transformed by normalization and encoding. NLP and Deep learning techniques are applied to process the text and video based preprocessed data. It provides a detailed assessment report for decision-making. Preliminary experiments show that the proposed system improves the efficiency of the hiring process, eliminates human bias, and offers a more comprehensive evaluation of the candidates than conventional methods.

INTRODUCTION

Human resource management is crucial to an organization's success and recruitment is one of the key functions. Traditionally, recruitment systems have been human-based, interview-driven and rely on human decision-making by managers. However, with the introduction of artificial intelligence (AI) [1], machine learning (ML) and multimedia analytics, there is a transition of recruitment systems to an automated and data-driven approach [2].

Conventional recruitment systems (manual or early computerized systems) tend to focus on applicant screening and tracking. These processes can be biased, inefficient to process large data, not real-time and do not include skills other than text. As organizations scale, these problems lead to increased time-to-hire, inconsistency in decisions and loss of talent [3].

Many of these issues have been overcome by new online hiring systems that use applicant tracking systems (ATS), online assessments and algorithmic screening. These tools improve the efficiency and convenience of the recruitment process but still rely on structured data sources such as resumes and structured questions [4]. They typically do not offer more detailed behavioral and cognitive profiling, such as communication skills, emotional intelligence and body language in interviews.

The recent advances in AI-based systems for hiring overcome these challenges by using multimodal data analysis techniques such as text, speech and video processing. These systems enable automated interviewing, real-time evaluation and profiling through data analysis. However, these methods also pose challenges such as privacy, bias, computational cost and the need for a large amount of human-annotated data for model training [5].

Recruitment systems can be divided into two parts. One side are traditions methods and the other are modern methods as shown in the Figure 1. The traditional recruitment methods have evolved over decades from newspaper, walk-in interview to social media based hiring and jobs through AI Chabot.

The system proposed in this study provides an integrated approach to recruitment by using an end-to-end system. It allows candidates to apply for jobs, complete profile forms, answer questions, and participate in automated video interviews. It also offers data preprocessing (cleaning, feature extraction, and transformation), and multimodal processing (text, speech, and video). Finally, it generates an analytic report for decision-making in candidate selection. This integrated system aims to automate the hiring process, reduce bias and improve hiring accuracy

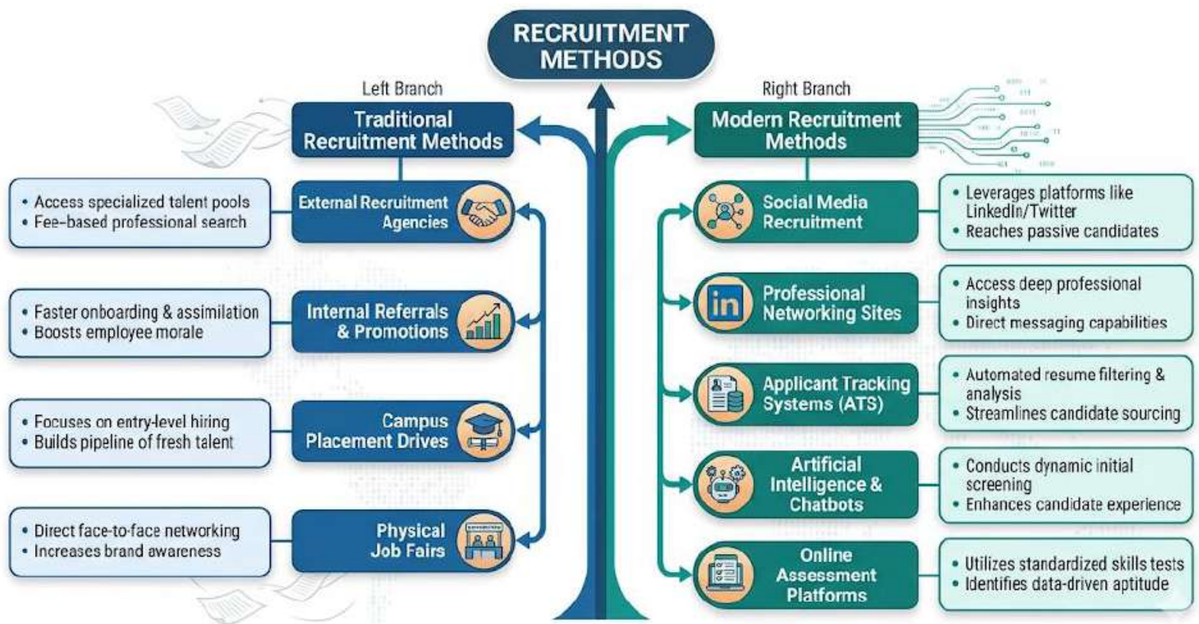

Figure 1: Categories of Traditional and Modern Recruitment Approaches

Literature Review

Recruitment methods have a rich history starting from manual approaches smart artificial Intelligence based methods using various modern platforms. This section provides some key categories of hiring suitable candidates of a specific domain.

Manual Recruitment Methods

Manual recruitment systems are traditional hiring practices, where human recruiters are responsible for screening resumes, conducting interviews and making the final hiring decision. This process is heavily reliant on human judgement, experience and intuition. Manual recruitment systems are inconsistent due to biases, fatigue and the absence of objective selection criteria. This can

lead to biased recruitment, which hampers diversity and performance [6].
However, manual recruitment processes can be used to gain insight into contextual information and candidate responses, which may not be captured by automated processes. Manual processes can be useful in complex decision-making situations, where social intelligence and judgement skills are required. However, manual methods are time-consuming, inefficient and costly, especially when recruiting large numbers of candidates [7]. This has led to the rise of automation.

Application-Based Recruitment Systems (ATS)

Application Tracking System (ATS) [8] is a basic type of automation, designed to streamline the recruitment process. ATS systems help companies to gather, store and organize applicant data. ATS systems typically use keywords and resume scanning to identify candidates for a job.
ATS systems are used for more efficiency and less time consuming. However, they can be problematic due to their use of keywords. For example, candidates may be rejected if their resumes do not include specific keywords, even if they are qualified for the job. Further, ATS may not assess semantic and contextual fit in resumes [9].
ATS does not assess soft skills like communication, teamwork and adaptability. So while ATS systems enhance efficiency and speed, they offer a limited view of fit.

Online Recruitment Systems

Online recruitment systems are an extension of ATS and use the internet to publish jobs, collect applications and test candidates. They enable applicants to apply from a distance, and overcome barriers to access. Online recruitment systems have transformed the recruitment process by providing convenience, efficiency and flexibility. This can be online tests, surveys and emails. This helps employers to shortlist candidates for interview. However, most online tests still employ static tests and multiple choice questions, which do not test dynamic skills such as communication and response [10]. They may not have interactive features that simulate the workplace environment. So, online recruitment systems are efficient and effective, but they still don't solve the problem of assessing candidates.

AI-Based and Intelligent Recruitment Systems

The rise of artificial intelligence (AI) and machine learning (ML) has made it possible to build smart recruitment systems that can conduct multifaceted analysis of candidates. These systems perform multimodal analysis using natural language processing (NLP), speech recognition and computer vision [11].
Text analysis uses NLP to analyze resumes and written responses [12]. Machine learning to evaluate fit and performance. Likewise, speech analysis, technology can be used to assess vocal characteristics like tone, pitch and emotion, which can be used to predict communication skills [13].
Video analysis has also enhanced the recruitment process through the assessment of facial expressions, gestures and attention. Multimodal analysis has been shown to be effective in predicting success. Such systems often employ deep learning algorithms, such as convolutional neural networks (CNNs), to process facial expressions and other factors [14].
While AI-based hiring systems can be beneficial, they are not without problems. Biases in data can lead to unfair hiring, Privacy, transparency and ethical issues are also significant in the use of AI systems. These systems also need large amounts of data and computing power [15].

Table 1: Comparative Analysis of the Existing Recruitment Methods with Advantages and Limitations

Technique	Description	Advantages	Limitations	Key Features
Manual Recruitment	Traditional human-driven hiring process	Contextual judgment, flexibility	Time-consuming, biased, not scalable	Face-to-face interviews, manual screening
Application-Based Systems (ATS)	Software for managing job applications	Efficient filtering, structured data handling	Keyword dependency, ignores soft skills	Resume parsing, keyword matching
Online Recruitment Systems	Web-based hiring platforms	Accessibility, remote hiring	Limited interaction, static assessments	Online forms, basic tests
AI-Based Recruitment Systems	Intelligent systems using ML/NLP	Automation, data-driven decisions	Bias, privacy concerns, complex implementation	Resume analysis, chatbots
Multimodal Recruitment Systems (Proposed)	Integrates text, speech, and video analysis	Comprehensive evaluation, reduced bias, scalability	High computational cost, data requirements	Video interviews, speech analysis, feature extraction, predictive reporting

Methodology

This project offers an end-to-end, artificial intelligence (AI) based recruiting system that leverages multimodal information processing to automate and improve the candidate assessment process. The proposed approach is an end-to-end system, which collects, preprocesses and analyses the candidate's information using text, speech and video data. The system flows through four stages: data collection, data preprocessing, training, analysis and decision support as shown in Figure 2.

Data Collection

The data collection aims to acquire rich and structured data of candidates via an interactive web-app. Candidates register and log in to the system, choose a suitable job profile, and fill out their profile data, such as education, skills and experience. They then complete a questionnaire that measures their knowledge, problem-solving skills, and writing skills.

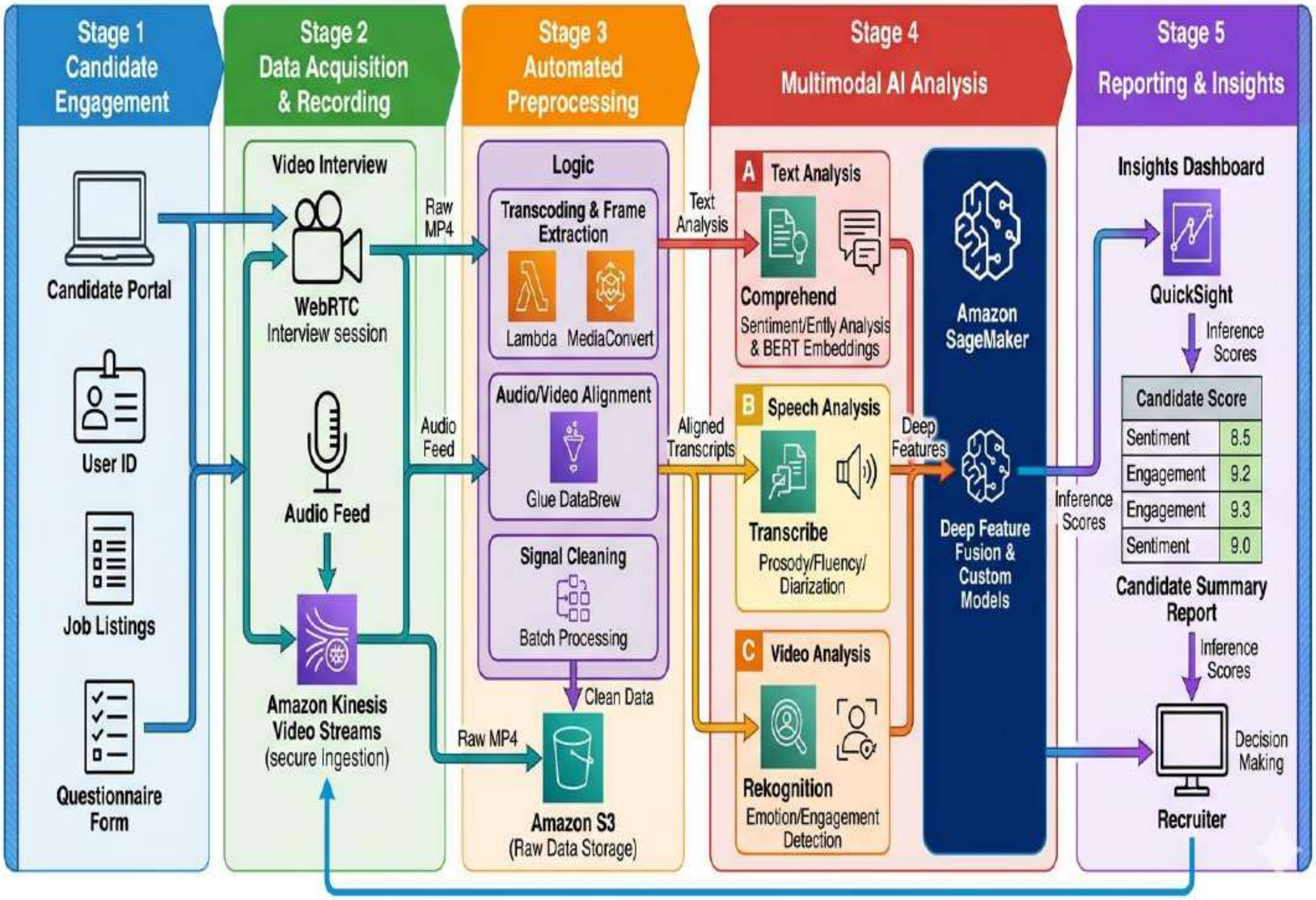

Figure 2: System Architecture of Proposed Model

Data Preprocessing

To enrich the assessment, an automated video interview is performed in which candidates are asked predetermined or dynamic questions. This process yields multimodal data, including text, speech and video. More specifically, the system gathers three types of data: (i) textual data from candidate resumes and questionnaire, (ii) speech data from candidate's audio responses, and (iii) video data from facial and body movements. This multimodal data is the basis for further analysis and allows for a comprehensive evaluation of candidate skills.

The raw data is then preprocessed to guarantee quality, uniformity and compatibility with machine learning algorithms. For text data, preprocessing involves tokenization, stop word removal, text normalization and lemmatization to convert linguistic data to a standard form. Stop-words and other irrelevant data are removed to enhance semantic relevance.

In the case of audio data, there are segmentation, de-noising, and amplitude normalization techniques. Silence removal and voice activity detection are used to remove irrelevant speech segments. Likewise, video data is preprocessed by extracting frames, normalizing resolution, and stabilizing the video for computer vision processing.

Model Training and analysis

Next, features are extracted from each modality. For text, features include word frequencies, semantic embedding, and text polarity. Auditory features are pitch, tone, speech speed, and Melfrequency cepstral coefficients (MFCCs). Visual features include facial features, facial expressions, and body motion. These features are then encoded and normalized to produce numerical representations, allowing for effective input to machine learning systems.

This data is then analyzed by dedicated machine learning algorithms for each media type. For text, natural language processing techniques are used to measure semantic similarity, skills and sentiment. Methods such as TF-IDF vectorization and word embedding are used to represent text, and similarity scores to compare candidates' responses to the skills required for the job.

Acoustic analysis is used to evaluate communication skills. Machine learning models analyze features such as pitch variability, tone steadiness and speech rate to assess confidence and clarity. Techniques such as statistical analysis and pattern recognition are applied to identify speech anomalies, pauses and emotional tones.

The fusion of multimodal data is used to combine the predictions of these models. In this paper, we use late fusion, in which the predictions or scores of each modality are used to create a representation of the candidate's performance. This method makes the system more robust by using multiple information sources.

Decision Support System

The final step in the process is to generate a report that can be used for hiring decisions. This step uses the scores of text, speech and video to determine the overall score for the candidates. This is used to make the final decision, which can be "Hired", "Not Hired" or "Needs Reevaluation".

In addition to the binary decision, the system also provides detailed information about candidates' performance on several aspects, such as technical competence, communication skills and personality. The report is augmented with summary statistics and visualizations for clarity and ease of understanding.

To reduce human biases, the decision support system is based on objective assessment metrics and evidence-based insights. However, it also enables human recruiters to review and verify the outcomes to ensure a combination of automation and human decision-making. This blended decision-making approach improves the accuracy, objectivity and speed of recruitment.

Results and discussions

After processing three types of data (text, speech and vide) using various presented in the methodology section, the results has been obtained based on the set parameters. The system used NLP for semantic and sentiment analytics of text data. For voice data, pitch, fluency and tone has been considered to calculate speech score of the candidate. Moreover, confidence, facial expression and the engagement of the interviewer have been used to calculate the video score. Multimodal fusion has used to compute a final score of the candidate and weighted mean was used to produce the output and to reduce bias of the modality. The system has evaluated 50 candidates applied for a software engineer job. A final computed score chart of the candidates has been presented in the Figure 3. It can be seen that most of the candidates have achieved more score in the video analysis. They may be appeared with preparation in this domain. Overall frequency of the selected and rejected candidates of the system has been presented in Table 2. On the assessment of 100 candidates, the system observed a maximum value of the score is 90% and the minimum is 50%. The success rate of the candidates is only 40%. Performance of the candidates at individual level is depicted in the Figure 4. When the threshold value is set to 75%, only 40% candidates have been selected based on the overs score of text, voice and video analysis. When the criteria is set to 70%, the success rate is observes at 76%. The system facilitates the recruiters to select the candidates on a specific criteria.

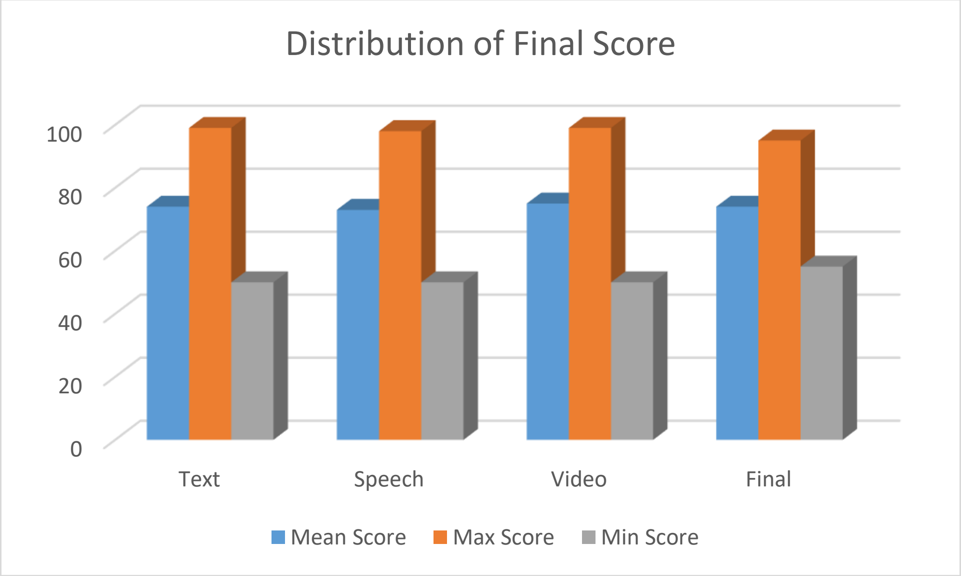

Figure 3: Computed Score of the Candidates at parametric Level

Table 2: Performance of the Total Candidates Applied for a Software Engineer job

Metric	Text	Speech	Video	Final
Mean Score	74	73	75	74
Max Score	99	98	99	95
Min Score	50	50	50	55
Selected Candidates	-	-	-	40%
Rejected Candidates	-	-	-	60%

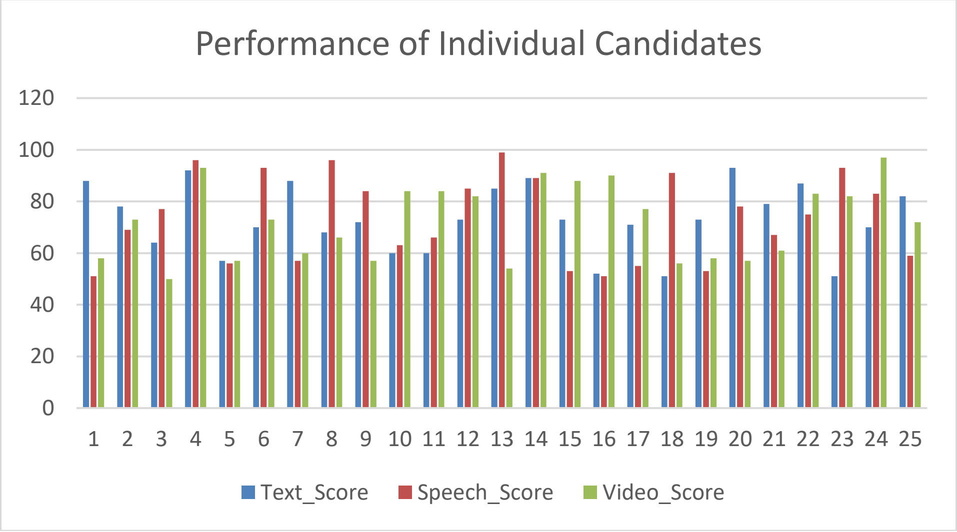

Figure 4: Performance of the Candidates at Individual Level

Table 3: Confusion Matrix

	Predicted Selected	Predicted Rejected
Actual Selected	18	2
Actual Rejected	3	27

After calculating a final aggregated score, a binary classification based on selected and rejected has been applied to assess the efficiency of proposed multimodal recruitment approach as shown in Table 3. It can be seen that the model predicts the positive class with only 2% errors. The model achieved 90% accuracy that shows that most of the candidates have appropriately classified and 85.7% precision to show the relevancy of the candidates to the applied job. The model achieved 90% recall that imitates a robust ability of the system for the selection of most suitable candidates and 87.8% f1 score indicates the satisfactory balance between recall and precision.
Using multimodal data integration enhances the quality and depth of candidate assessments. Text analysis is effective in capturing domain knowledge and intelligence through semantic similarity and sentiment analysis, ensuring candidates have the technical skills for the job. However, text is unable to capture communication and behavioral traits.
The analysis of speech helps with this by measuring tone, pitch and fluency, which are all good predictors of confidence and communicative skills. Those with steady tone and fluency are more effective communicators in the real world, so this mode is critical for roles that require good communication skills.
Video is also a very important mode that captures non-verbal cues. For example, facial recognition and eye tracking can demonstrate levels of candidate confidence, attention and excitement. This can show information that can't be detected in the text or speech, such as anxiety or boredom. The classification accuracy (90%) shows that the multimodal system is accurate. The confusion matrix also shows that there is a good true positive to false positive ratio. Crucially, the multimodal system is less biased as it uses several features to rate the candidate rather than a single subjective feature.

Conclusion and Future Work

Finally, the proposed approach combines multimodal data capture, preprocessing, and machine learning analysis to provide a holistic recruitment tool. The system provides a strong assessment of candidates by processing audio, video and text input data that changes the weaknesses of traditional recruitment approaches and enabling effective fruitful decision support. The system accelerates the hiring process, enhances decision-making and offers scalable hiring practices. Overall, the proposed system is better than traditional hiring systems because it provides a holistic, unbiased and scalable approach that is well suited to the current hiring landscape. Future research will aim to increase model accuracy, resolve ethical issues and include explainable AI to improve decision-making transparency.

REFERENCES

[1] Yanamala, K. K. R. (2021). Integration of AI with traditional recruitment methods. Journal of Advanced Computing Systems, 1(1), 1-7.
[2] Bakalopulos, I. (2024). The Evolution of Recruitment Markets: a comparative analysis of traditional vs. modern approaches and the impact of Generation Z.
[3] Sarkar, S. K., & Sen, A. (2026). Traditional vs. Modern Recruitment Methods. In Emerging Technologies for Recruitment Strategy and Practice (pp. 59-84). IGI Global Scientific Publishing.

[5] Gaike, V. V., Kulkarni, M., Shastri, A., Sabale, S. K., Nanwatkar, R., & Watvisave, D. (2024). A Comparative Study of Traditional and Modern Recruitment Techniques: The Role of Psychological Assessments in Enhancing Hiring Quality in Maharashtra EV Field. Library of Progress-Library Science, Information Technology & Computer, 44(3).

[6] Cuggia, M., Besana, P., & Glasspool, D. (2011). Comparing semi-automatic systems for recruitment of patients to clinical trials. International journal of medical informatics, 80(6), 371-388.

[7] Diego, L. A. C. (2015). Automatic methods for assisted recruitment (Doctoral dissertation, Université d'Avignon).

[8] Al-Qassem, A. H., Agha, K., Vij, M., Elrehail, H., & Agarwal, R. (2023, March). Leading talent management: Empirical investigation on applicant tracking system (ATS) on e-recruitment performance. In 2023 International Conference on Business Analytics for Technology and Security (ICBATS) (pp. 1-5). IEEE.

[9] Alzhrani, A. M. (2020). The Effectiveness of E-Recruitment Software Over Other Online-Based Recruitment Methods. Global Journal of Economics & Business, 8(2).

[10] Rosoiu, O., & Popescu, C. (2016). E-recruiting platforms: Features that influence the efficiency of online recruitment systems. Informatica Economica, 20(2), 46.

[11] Yanamala, K. K. R. (2021). Integration of AI with traditional recruitment methods. Journal of Advanced Computing Systems, 1(1), 1-7.

[12] MR, N., & BS, S. (2024). Enhancing User Sentiment Analysis in Video Interviews: Leveraging NLP and Facial Emotion Recognition for Comprehensive Analysis. Grenze International Journal of Engineering & Technology (GIJET), 10.

[13] Cherian, E., Singh, J., Naveen, S., Al-Nussairi, A. K. J., Alanssari, A. I., & Faris, N. (2025, August). Multimodal Learning-Driven Recruitment Screening System for Talent Acquisition Automation. In 2025 International Conference on Recent Innovation in Science Engineering and Technology (ICRISET) (pp. 1-6). IEEE.

[14] Muhammad, S. (2025, April). Enhancing Recruiter Outreach: Predicting Job Ad Quality with Advanced Machine Learning and NLP Models. In Artificial Intelligence and Applications (Vol. 3, No. 4, pp. 443-458).

[15] Albassam, W. A. (2023). The power of artificial intelligence in recruitment: An analytical review of current AI-based recruitment strategies. International Journal of Professional Business Review: Int. J. Prof. Bus. Rev., 8(6), 4.