IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

JOB APPLICATION SCAM DETECTION USING ARTIFICIAL INTELLIGENCE

[1]Prof.Vaishnavi C S, [2]Ananya S V , [3]Eshwari K M, [4]Hampritha M, [5]Meghana R [1]Dept.of Information Science & Engineering, Jain Institute of Technology, Davangere, VTU, Davangere, Karnataka, India.

Abstract:

Online hiring portals have become prime targets for fraudulent actors who post counterfeit job vacancies to steal personal data or extort money from vulnerable job seekers. This paper reviews twelve peer-reviewed studies spanning 2021 to 2026, tracing how detection methodologies have progressed over that period. Early works leaned on classical machine learning classifiers, whereas newer contributions exploit transformer-based encoders and behavioral fingerprinting of scam perpetrators. Building on these observations, a hybrid detection architecture is proposed—one that pairs BERT-style semantic embeddings with structured attribute scoring. Cross-study evidence indicates that modern AI-oriented pipelines consistently achieve higher discrimination accuracy and adapt more effectively to novel fraud patterns than their rulebased counterparts.

Index Terms — Employment Fraud Detection, Machine Learning, Deep Learning Models, Natural Language Processing, Scam Identification, Cybersecurity, Transformer-Based Models, Bidirectional LSTM, Fake Job Listings, AI-Based Systems

I. Introduction

Over the past decade, the internet has fundamentally altered the way employers source candidates and how job seekers pursue career opportunities. Platforms such as LinkedIn, Indeed, and Naukri have aggregated millions of listings in one place, reducing the friction that once made job searching a slow, paper-heavy activity. For companies, this means faster access to a wider talent pool; for applicants, it means the ability to apply to dozens of positions within a single afternoon. By most measures, this shift has been beneficial—yet it has simultaneously introduced a class of threat largely absent from traditional hiring channels.

Fraudulent job advertisements exploit the trust that users place in these platforms. Scammers fabricate vacancies at recognizable companies, invent non-existent staffing agencies, and craft listings designed to appear indistinguishable from genuine postings. Their objectives vary: some seek to harvest identity documents and contact details for downstream misuse; others deceive applicants into paying registration fees, training costs, or equipment deposits that are never refunded. In either case, the personal and financial harm to victims can be substantial, and reputational damage to affected platforms—even when they are not directly at fault—erodes user confidence in online recruitment as a whole.

The scale of contemporary job portals makes manual moderation an inadequate response. Thousands of new listings are posted every hour across major platforms, and the sophistication of modern scam content—often indistinguishable in tone and format from genuine offers—means that even trained reviewers miss a significant proportion of deceptive entries. These realities have made automated, AI-driven detection a practical necessity rather than an optional enhancement.

Research into automated scam detection has advanced considerably since early keyword-filtering approaches. Initial systems flagged postings based on the presence of suspicious phrases or anomalous salary figures, but their fixed rule sets were quickly circumvented as scammers adapted their wording. Machine learning classifiers introduced statistical decision-making, improving generalization but still falling short

when confronted with nuanced or context-dependent deception. More recently, deep learning architectures—particularly transformer-based language models—have demonstrated the ability to encode subtle semantic relationships within advertisement text, enabling detection of fraud signals that resist surface-level analysis.

This paper surveys twelve studies published between 2021 and 2026 to map the development of fraud detection techniques across that period, and proposes a composite detection framework integrating multiple complementary modeling strategies. The goal is to provide both a structured account of where the field currently stands and a practical architecture that practitioners can build upon. Subsequent sections cover the reviewed literature, the proposed system design, experimental findings from reviewed studies, a discussion of remaining challenges, and directions for future investigation.

II. Literature Review

A review of published research reveals steady, measurable gains in the ability to flag deceptive job advertisements. Studies from 2026 broadened the analytical scope by examining the behavioral signatures of scam orchestrators—patterns such as manufactured time pressure, psychological coercion tactics, and reuse of fraudulent content templates—rather than relying solely on the surface text of individual postings [1, 2].

Research appearing in 2025 highlighted the dominance of transformer architectures, with BERT-derived models proving especially capable of producing rich, context-sensitive text representations. Hybrid systems merging classical classifiers with neural layers recorded benchmark accuracy figures above 95%. Several groups in this period also introduced synthetic scam samples generated by large language models to better probe detection robustness [3–7].

Work from 2024 demonstrated that deep convolutional and recurrent models outperformed conventional baselines in extracting layered representations from advertisement corpora [8], while 2023 publications explored ensemble and sequencelabeling strategies, with Bidirectional LSTM architectures showing notable strength in capturing dependencies across lengthy passages of text [9, 10].

Contributions from 2021 and 2022 laid the empirical groundwork using well-established algorithms including Logistic Regression, Decision Trees, and Random Forests. Although computationally straightforward and easy to interpret, these classifiers struggled with subtle semantic signals present in more elaborate fraudulent listings [11, 12]. Taken together, this body of literature traces a clear trajectory from relatively transparent but limited models toward adaptive, meaning-aware architectures better equipped to handle increasingly sophisticated deception strategies.

III. Methodology A. Proposed Framework Overview

The system developed in this study is an AI-driven pipeline designed to distinguish authentic job postings from fabricated ones distributed through digital hiring channels. The architecture integrates several mutually reinforcing technical components to maximize classification reliability while keeping false negatives low.

B. Core System Modules

1) Data Acquisition: Job-related records are collected from publicly accessible hiring platforms and standardized open datasets to form a balanced corpus containing both genuine and fraudulent samples.

2) Preprocessing Pipeline: Raw inputs go through noise elimination and normalization steps. These include tokenization, text case standardization, and imputation of missing attribute values to produce a clean representation suitable for analysis.

3) Feature Engineering: Discriminative attributes are extracted from job titles, listing descriptions, and supplementary fields such as company profile data and stated compensation. Markers closely associated with fraud—advance payment demands, artificial urgency signals—are explicitly encoded as binary or ordinal features. The model core integrates BERT for context-rich embeddings, a Bidirectional LSTM for directional sequence modeling, and a Random Forest classifier for the final binary decision.

4) Classification Engine: The inference layer consolidates outputs from each trained component to determine whether a given listing represents a legitimate opportunity or a deceptive one.

5) Real-Time Alert System: When the ensemble assigns a high fraud probability to a listing, an automated notification is sent to designated stakeholders, enabling timely intervention before potential harm reaches applicants.

C. End-to-End Workflow

Processing begins with ingestion of raw posting data, which is immediately cleaned and normalized. Feature extraction follows, drawing on both free-text content and structured metadata. The enriched representation passes through the composite model ensemble, yielding a classification output surfaced to end users; when warranted, the alert mechanism is activated simultaneously.

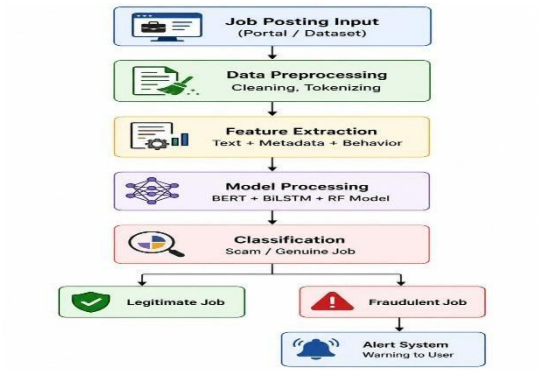

Fig 3.1: System Flow Diagram

IV. Results and Discussion

Comparing results across the twelve reviewed studies reveals a marked upward trend in detection performance. Classical supervised approaches—Logistic Regression and Decision Trees—returned moderate outcomes in the 75–85% accuracy range, constrained chiefly by their reliance on hand-engineered feature sets that could not capture deeper semantic cues.

Ensemble methods, particularly Random Forest, lifted performance toward the 85–90% bracket by combining multiple decision boundaries. Deep learning brought a more pronounced jump: Bidirectional LSTM models captured long-range textual relationships that shallower architectures missed, pushing accuracy to roughly 98%. Transformer-based models—principally BERT-family variants—improved results further by encoding document-level semantic structure, reliably exceeding 95% accuracy. The strongest outcomes emerged from hybrid configurations pairing neural feature extractors with ensemble classifiers, delivering stable results in the 96–99% range.

Collectively, these figures confirm that data-driven, intelligent detection architectures have raised both the precision and robustness of job scam identification substantially above what earlier generations of tools could achieve, with state-of-the-art systems showing considerably better resilience to adaptive fraudsters.

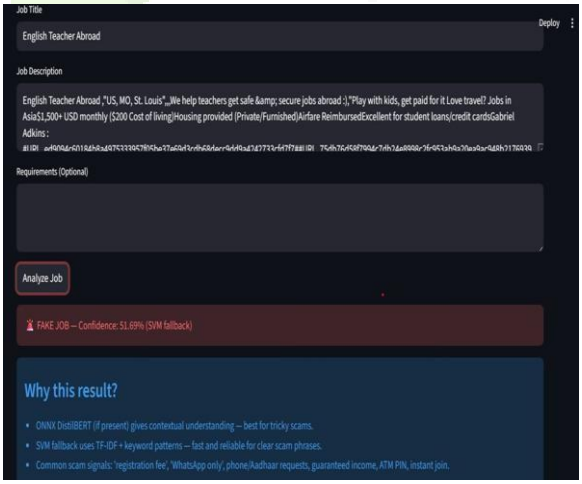

Fig 4.1: Job Analysis Illustration

This illustration shows an AI-powered review interface applied to a sample advertisement for an "English Teacher Abroad" position. The listing displays multiple features commonly flagged in deceptive postings: a salary significantly above market rate, subsidized housing, covered travel expenses, and descriptions of the workload as exceptionally light.

Upon processing, the system returns a fraud probability estimate of 51.69%. Primary semantic analysis is handled by a DistilBERT encoder, with a secondary SVM model activated when confidence in the lead prediction falls below a set threshold. The output explanation lists specific risk indicators identified during analysis: requests involving advance payment or processing charges, contact routed through informal messaging applications such as WhatsApp, offers of unusually high pay, and language designed to pressure applicants into acting immediately.

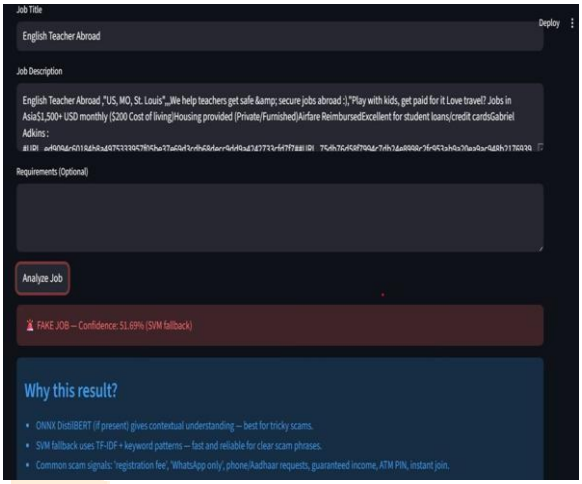

Fig 4.2: Feedback Illustration

A second evaluation of the same English Teacher posting—viewed through a feedback lens—returns a similar risk profile. The system flags requests for application or processing fees, communication conducted via informal channels rather than professional platforms, demands for government-issued identity documents including Aadhaar, above-market salary promises, and recruitment messages containing urgency-driven language intended to compress decision time.

V. Discussion

The shift from handcrafted feature extraction toward deep contextual understanding mirrors a broader escalation in the sophistication of fraudulent job advertisements. Simpler classifiers were adequate when scams were straightforward and their textual signals were easy to isolate, but they lost ground as fraudsters learned to mimic legitimate listing formats more closely. Their core limitation was an inability to parse meaning at the sentence and document level rather than at the level of individual keywords or phrases.

Neural and transformer-based architectures overcame much of that limitation by encoding semantic relationships within text, allowing models to identify fraud indicators that are distributional rather than lexical in nature. More recent research has moved further still, incorporating behavioral analysis that examines the back-and-forth between scammers and their targets to uncover communication patterns not visible from the listing text alone.

Notwithstanding these gains, several practical constraints remain. High-capacity neural models carry substantial computational costs, depend on large annotated training sets, and introduce deployment complexity that can be prohibitive in resource-limited settings. Latency in inference pipelines also limits applicability in genuinely real-time screening scenarios.

A practical path forward involves combining lightweight interpretable classifiers with compact neural feature extractors, trading some peak performance for reduced resource requirements. Research priorities should include model compression techniques, multilingual extension to cover non-English hiring markets, and improved explainability tools to build trust among operators and end users alike.

VI. Conclusion

Fake job postings have grown into a serious problem on digital recruitment platforms, causing harm to both individual applicants and the employers whose reputations are sometimes exploited. This review of twelve studies published from 2021 through 2026 documents the field's progression from basic statistical classifiers to sophisticated AI-driven systems with substantially improved detection capability and generalization.

The proposed framework brings together transformer-based text encoding, sequential deep learning, and behavioral profiling within a unified detection pipeline. Evidence from the surveyed literature consistently shows that ensemble and hybrid approaches outperform single-method systems in accuracy and stability, particularly when evaluated against diverse or out-of-distribution samples.

Recommended directions for future work include improving model transparency to support audit and regulatory requirements, broadening language coverage for global applicability, and stress-testing systems at the scale typical of major commercial job boards. Progress on these fronts will be essential for deploying effective fraud detection in real-world hiring infrastructure.

Acknowledgment

The authors would like to thank their respective institutions for supporting this research and providing access to academic resources required for the literature review and system development.

References

[1] A. Pitumpe and A. Rahmati, "Anansi: Scalable characterization of job scams," 2026.
[2] G. Anagha et al., "Modeling behavioral signals in job scams," 2026.
[3] K. Taneja et al., "Fraud-BERT: Transformer-based recruitment fraud detection," 2025.
[4] S. S. Sanisetty et al., "Fraudulent job post detection using ML and BERT," IEEE Conference, 2025.
[5] A. J. Veliyath et al., "Fake job detection using statistical and NLP methods," IEEE Conference, 2025.
[6] "Real or fake job posting prediction using ML," IEEE Conference, 2025.
[7] X. W. Tan et al., "Generative AI framework for scam detection," 2025.
[8] N. Akram et al., "Online recruitment fraud detection using deep learning," IEEE Access, 2024.
[9] A. Pillai, "Detecting fake job postings using Bi-LSTM," 2023.
[10] "Crowdsourcing-based job fraud detection using ML," 2023.
[11] M. Naudé et al., "Machine learning approach to detecting fraudulent job types," 2022.
[12] C. Prashanth et al., "Online fake job advert detection using ML," IEEE Conference, 2022.