Fake Job Listing Detection using Deep Learning and Agentic Generative AI

1. Introduction and Background

The rapid expansion of online recruitment platforms has significantly improved accessibility to employment opportunities. However, this digital growth has also enabled the proliferation of fraudulent job listings designed to exploit job seekers through advance-fee scams, phishing attacks, and identity theft. These fake job postings are often carefully crafted to resemble legitimate advertisements, making them difficult to detect using traditional keyword-based filtering or manual moderation approaches. As a result, there is a critical need for an intelligent, scalable, and explainable AI-driven solution capable of detecting fraudulent listings with high accuracy and reliability.

2. Problem Statement

The objective of this project is to design and develop an end-to-end Deep Learning–based agentic fraud detection system capable of identifying fake job listings from online recruitment platforms. Given a job posting consisting of:

- Job description text
- Company profile information
- Salary and employment details
- Location and contact metadata

The system must:

1. Predict the probability of the listing being fraudulent.
2. Provide a structured and interpretable explanation supporting the decision.

3. Perform multi-step reasoning by validating suspicious attributes using auxiliary verification modules.

The system moves beyond simple binary classification by incorporating evidence-based reasoning and autonomous verification capabilities.

3. Motivation

Current fraud detection systems largely rely on:

- Rule-based filtering
- Manual moderation
- Surface-level text classification

These approaches fail to capture complex linguistic deception patterns and contextual inconsistencies. Additionally, most existing models operate as black boxes, offering little interpretability to users. An intelligent, explainable, and agent-driven framework can:

- Reduce financial and emotional harm to job seekers
- Assist recruitment platforms in automated moderation
- Improve trust in digital hiring ecosystems

4. Proposed Solution

The proposed system integrates three major AI components:

4.1 Transformer-Based Fraud Classification

A pre-trained transformer model (e.g., roberta-base or similar architecture) will analyze:

- Linguistic patterns
- Semantic inconsistencies
- Contextual anomalies

The model outputs a fraud probability score.

4.2 Agentic Verification Framework

An intelligent decision-making agent orchestrates multiple tools:

- Metadata anomaly detector (suspicious email domains, unrealistic salaries, missing company details)

- Company validation module (cross-verification of website domain and organization details)
- Consistency checks across structured and unstructured inputs

The agent aggregates evidence from multiple sources before arriving at a final fraud risk score.

4.3 Generative AI Explanation Layer

A Generative AI component synthesizes:

- Model prediction scores
- Identified anomaly signals
- Verification results

It produces a structured, human-readable explanation highlighting:

- Deceptive language cues
- Suspicious metadata
- Validation mismatches

This enhances transparency and user trust.

5. System Architecture Overview

The following flow chart provides a high-level overview of the system architecture:

Component	Function	Output
Input	Job posting data	Raw data
Transformer Classifier	Initial fraud probability prediction	Probability score
Agent Controller	Orchestrates verification tools	Decisions for verification
Anomaly Detection + Verification Tools	Checks for suspicious attributes and external data	Evidence signals
Evidence Aggregation	Combines all evidence for a final risk score	Final fraud risk score
Generative Explanation	Creates a human-readable report	Structured explanation

Component	Function	Output
Final Fraud Report	Presents prediction and explanation	Final outcome

The architecture ensures both predictive accuracy and interpretability.

6. Expected Outcomes

The project aims to deliver:

- A high-accuracy fake job detection model
- An evidence-driven fraud reasoning system
- Explainable AI outputs for end users
- A deployable prototype suitable for integration with recruitment platforms

7. Novelty of the Approach

The novelty of the proposed system lies in:

- Combining Deep Learning classification with an agentic decision-making framework
- Incorporating multi-tool verification instead of single-model prediction
- Generating structured explanations through Generative AI
- Providing interpretable and actionable fraud assessments