

Fake Job Listing Detection

Deep Learning & Agentic Generative AI

MILESTONE 1 REPORT

The Current Landscape: A Crisis of Trust



FTC Warns of Surge in Fake Job Scams

NYTimes.com - Oct 24, 2023

Agency reports record complaints as scammers exploit remote work trends...



Thousands Fall Victim to LinkedIn Phishing Attacks

TechCrunch - Nov 15, 2023

Malicious actors impersonate recruiters to steal personal data and credentials...



The 'Ghost Job' Problem: Why Your Application is Ignored

Forbes - Dec 1, 2023

Many listings are fake or inactive, used for market research or data collection...



How Scammers Use AI to Create Convincing Job Posts

Wired - Dec 12, 2023

Generative AI tools are making it easier to create highly targeted and believable fraudulent listings...



New Report: Financial Losses from Job Fraud Hit Record High

FBI.gov - Jan 5, 2024

Victims report significant monetary losses through advance-fee fraud and identity theft...

Widespread fraud. Financial impact. Eroding confidence.

Project Context & Importance

Why Fraud Detection Matters



Financial Loss



Identity Theft



Reputational Damage

Stakeholders



Job Seekers



Employers



Platforms

Current Status of Job Scams



Increasing Volume



Higher Sophistication



Low Detection Rate

Scope & Boundaries



In-Scope

- Text Analysis
- Salary Verification
- Domain Checks



Out-of-Scope

- Legal Action
- Dark Web Monitoring
- Physical Verification

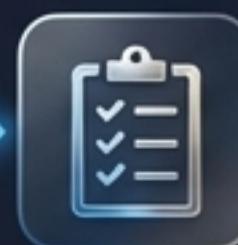
Project Analysis Pipeline



Data Collection



Model Training



Testing & Validation



Deployment

Addressing the urgent need for advanced fraud detection mechanisms.

THE DETECTION LANDSCAPE: FIVE APPROACHES

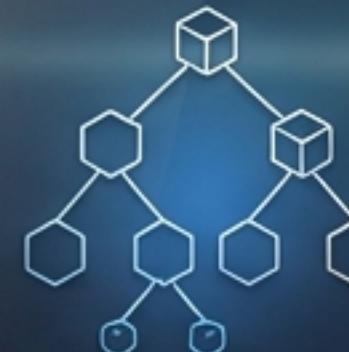
01

RULE-BASED



02

CLASSICAL ML



03

DEEP LEARNING



04

TRANSFORMERS



05

EXPLAINABLE AI



Keyword Filtering.
Simple matching,
easily bypassed.

Naive Bayes /
Random Forest.
Uses TF-IDF features.

CNN / LSTM.
Better context,
black box.

BERT / RoBERTa.
State-of-the-art text
understanding.

LIME / SHAP.
Feature weights, not
user explanations.

Performance Leaderboard

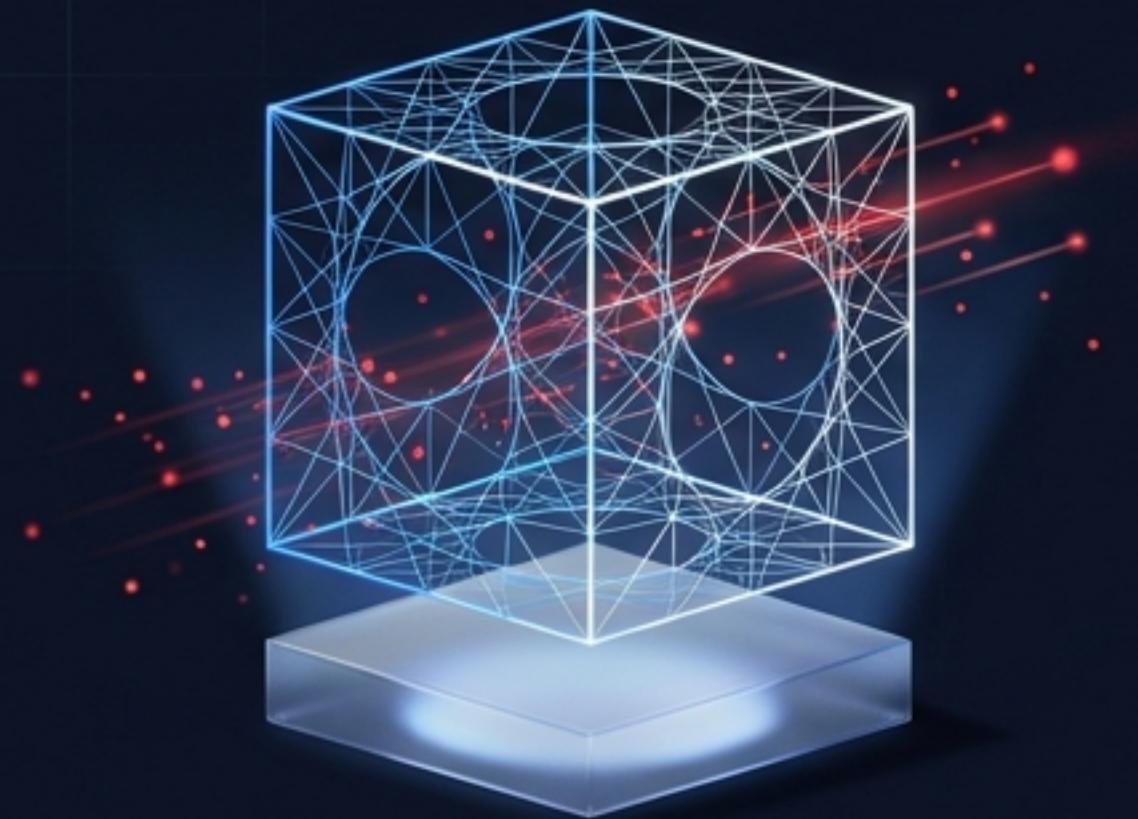
Transformers define the current ceiling for text accuracy

Model Approach

Rule-Based (Keyword Filter)	<div style="width: 10%;"> </div>	Accuracy: Low F1: Low
Classical ML (Random Forest)	<div style="width: 97%;"> </div>	Accuracy: 97% F1: 0.82
Deep Learning (BiLSTM)	<div style="width: 97%;"> </div>	Accuracy: 97% F1: 0.83
Transformer (BERT)	<div style="width: 98%;"> </div>	Accuracy: 98% F1: 0.88
Transformer (RoBERTa)	<div style="width: 98.5%;"> </div>	 Best Baseline Accuracy: 98.5% F1: 0.91

Why Current Defenses Fail

Keyword Filters



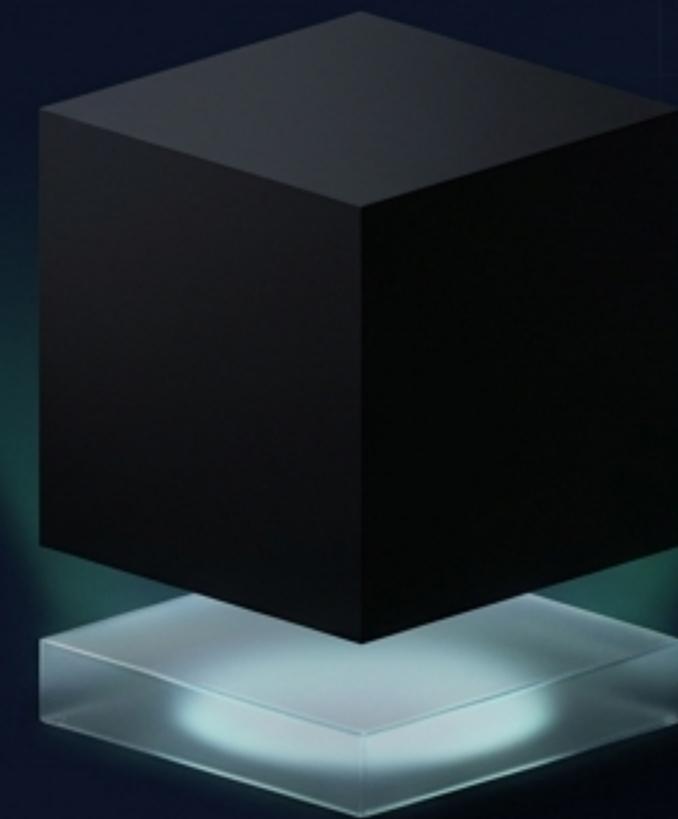
Easily Bypassed

Traditional ML



Misses Context

Standard Deep Learning



The Black Box

We need verification, not just classification.

The Blind Spot: Why High Accuracy Isn't Enough

Critical weaknesses in existing approaches.



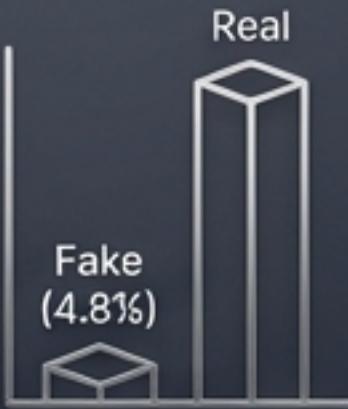
NO METADATA VERIFICATION

Models ignore email domains, missing info, and salary ranges.



BLACK BOX PREDICTIONS

Outputs a score without explanation. Hard to trust.



POOR CLASS IMBALANCE HANDLING

Only ~4.8% of listings are fake; models miss fraud cases.



NO HUMAN-READABLE EXPLANATION

LIME/SHAP provide numbers, not plain-language reports.

Gap Analysis

Fractures in current defenses

Gap 1: No multi-step verification
(Domain/Salary ignored)

Gap 2: No narrative explanation
(Just scores)

Gap 3: Metadata is ignored
(Email/Location missing)

Gap 4: Class imbalance unresolved
(Low recall)



Bridging the Gaps: The Proposed System

RoBERTa Classifier

Fixes: Text Understanding

Fine-tuned transformer achieves ~98.5% accuracy.

Agentic Verification

Fixes: Multi-step Reasoning

Checks company domain, email pattern, and salary.

Generative AI Layer

Fixes: Human-Readable Report

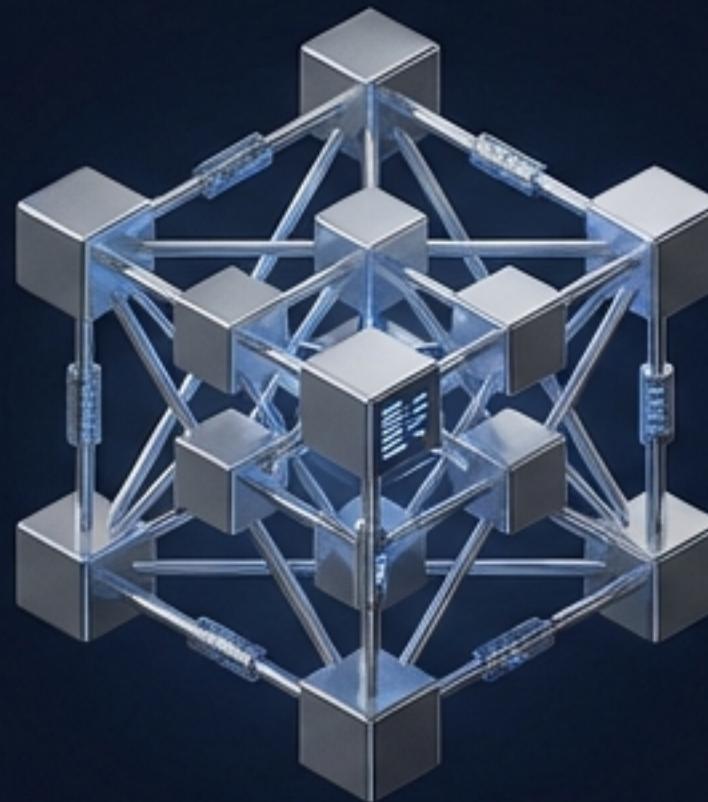
Produces a plain-language fraud report.

Text + Metadata Pipeline

Fixes: Ignoring Metadata

Combines text signals and structured data.

A Hybrid Approach



Transformer Models



Agentic Verification



Trust & Explainability

Combining RoBERTa for text understanding with Agents for fact-checking.

The Agentic Workflow



Multi-step reasoning to cross-check suspicious attributes.

Not Just a Score. A Story.

Generative AI provides structured, plain-language warnings.



Key Differentiators

Feature	Traditional ML	Deep Learning	Proposed System
Context Understanding	Low	High	High 
Metadata Verification	No 	No 	Yes 
Multi-step Reasoning	No 	No 	Yes 
Structured Explanation	No 	Limited	Yes 
Agent-Based System	No 	No 	Yes 

Target Objectives: Robust Multi-Step Verification & Use Case



Raw Data



Multi-Step Verification
(Domain, Salary, Metadata)



Robust Prediction
Model



Real-world Use Case
(e.g., Fraud Detection)



95%



Target Accuracy

F1-Score

>0.90



Verification Tools Integrated

3+



Benchmark Dataset

EMSCAD

The Builders (Milestone 1)

Arun Dutta

Literature Review & Gap Analysis

- State-of-the-art methods (SOTA)
- Identify existing dataset limitations

Hritik Roshan Maurya

Problem Framing & System Architecture

- Define problem scope and objectives
- Architectural diagram & module breakdown

Vivek Bajaj

Data Pipeline & Deep Learning Workflow

- Data cleaning, preprocessing, & augmentation
- Model selection, training, & hyperparameter tuning

Vishwas Mehta

Fraud Pattern Analysis & Domain Research

- Analyze common fraud tactics
- Consult with domain experts for feature engineering

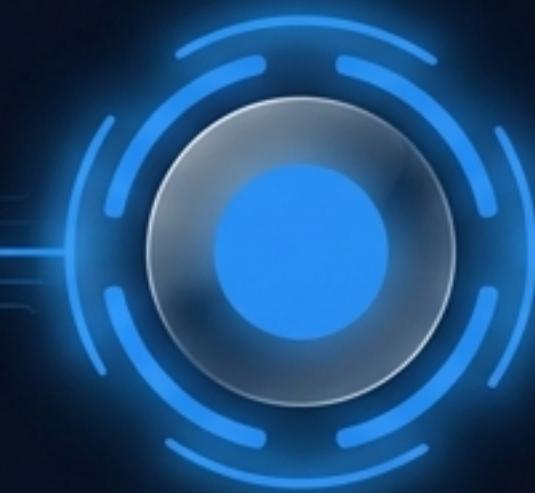
Status & Next Steps

Research &
Architecture



Milestone 1
Complete

Prototype
Development



In Progress

Integration
& Testing



Upcoming

Moving from architecture to implementation.