

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...

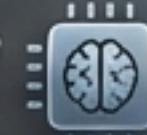
The 'Ghost Job' Problem: Why Your Application is Ignored

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

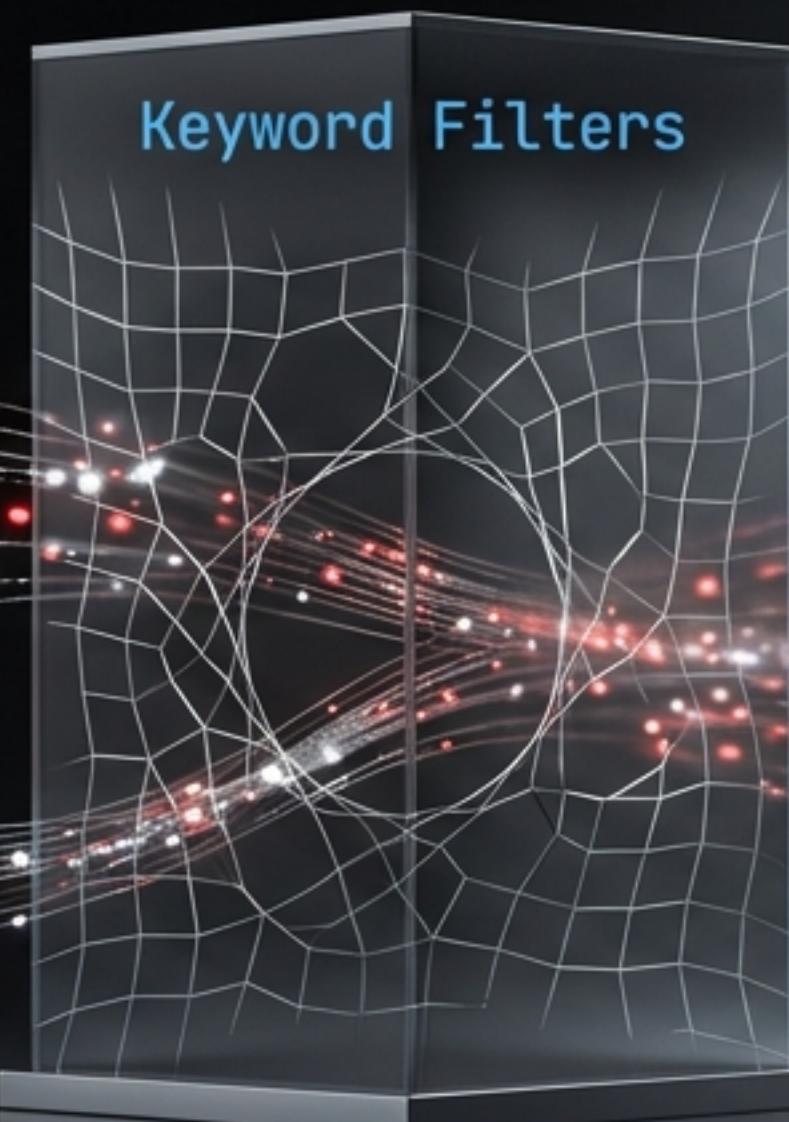


FBI.gov - Jan 5, 2024

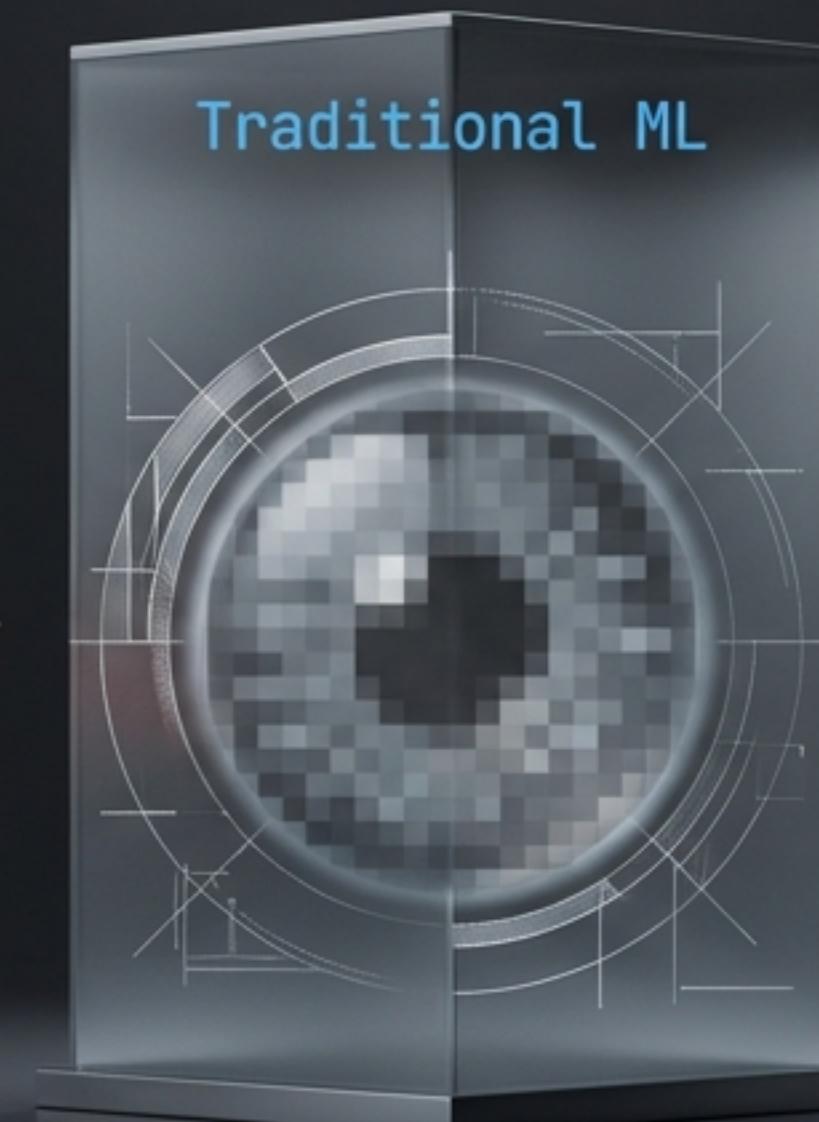
Victims report significant monetary losses through advance-fee fraud and identity theft...

Widespread fraud. Financial impact. Eroding confidence.

Why Current Defenses Fail



Easily Bypassed



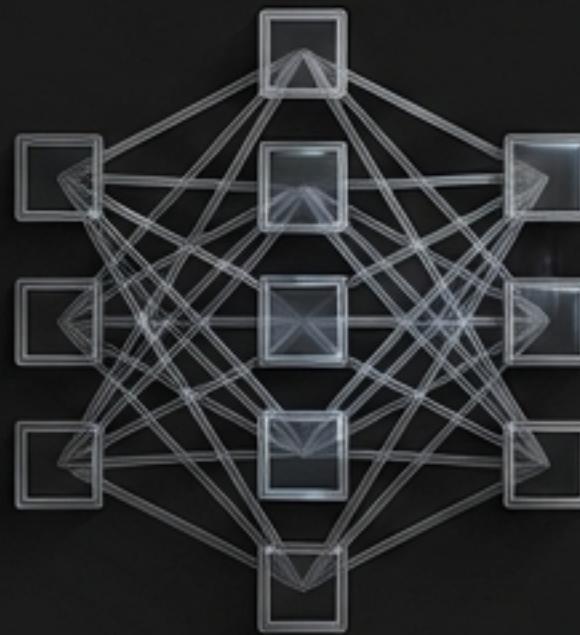
Misses Context



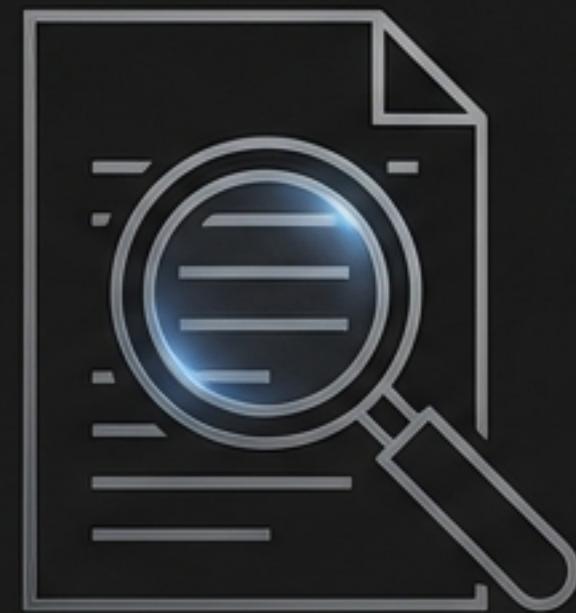
The Black Box

We need verification, not just classification.

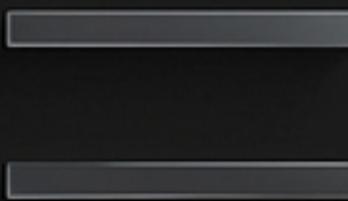
A Hybrid Approach



Transformer
Models



Agentic
Verification



Trust &
Explainability

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

Project Context & Importance.

Why Fraud Detection Matters



Financial Loss



Identity Theft



Reputational Damage

Stakeholders



Job Seekers



Employers

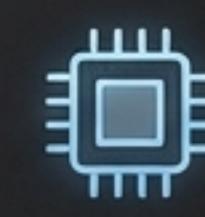


Platforms

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

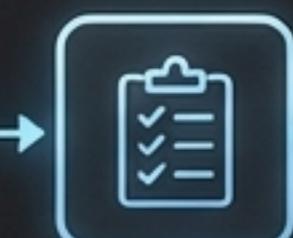
Project Analysis



Data Collection



Model Training



Testing & Validation



Deployment

Addressing the urgent need for advanced fraud detection mechanisms.

The Agentic Workflow



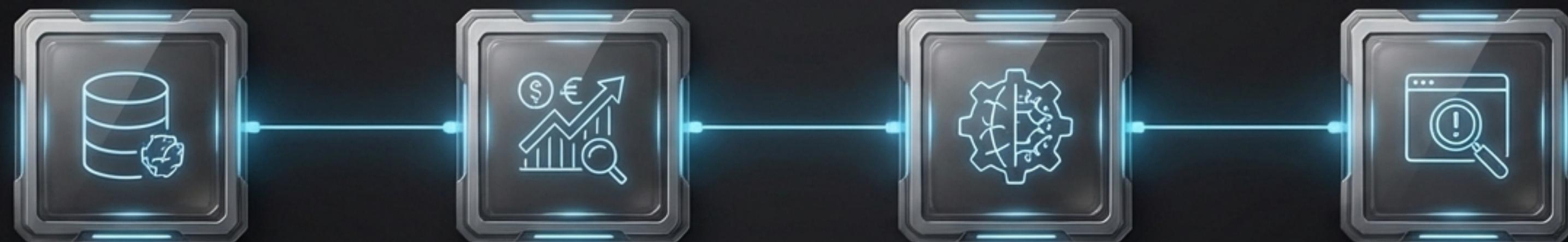
Multi-step reasoning to cross-check suspicious attributes.

Not Just a Score. A Story.

Generative AI provides structured, plain-language warnings.



Target Objectives: Robust Multi-Step Verification & Use Case



Raw Data

Multi-Step Verification
(Domain, Salary, Metadata)

Robust Prediction Model

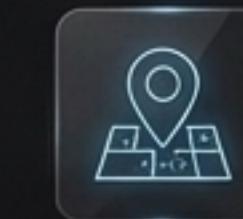
Real-world Use Case
(e.g., Fraud Detection)



Target Accuracy
95%



F1-Score
>0.90



3+
Verification Tools
Integrated



EMSCAD
Benchmark Dataset

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.