Preliminary Analysis Report – Fake Job Postings

1 | Objective

Fraudulent job ads harm both candidates and platform trust. This analysis identifies patterns distinguishing real and fake postings and demonstrates a transparent, first-pass model to flag suspicious listings for prioritization. We aim to provide an explainable model so Trust & Safety staff can see why a specific ad is flagged.

2 | Methodology

Dataset:
17,880 unique job postings labelled fraudulent (1 = fake, 0 = real).

Cleaning:
- Removed duplicate rows and columns.
- Filled missing text fields with empty strings.

Features Engineered:
- Text lengths for: title, company_profile, description, requirements, benefits.
- Binary flags: telecommuting, has_company_logo, has_questions.

Exploratory Data Analysis Visuals:
1. Class balance (real vs. fake).
2. Text-length kernel density estimation (KDE) plots for description and requirements.
3. Correlation heatmap (text lengths + binary flags vs. fraudulent).
4. Word clouds for real vs. fake descriptions.

Prototype Model:

- Pipeline: TF-IDF (10,000 unigrams/bigrams) → Logistic Regression
  - Hyperparameters: class_weight='balanced', max_iter=1000.

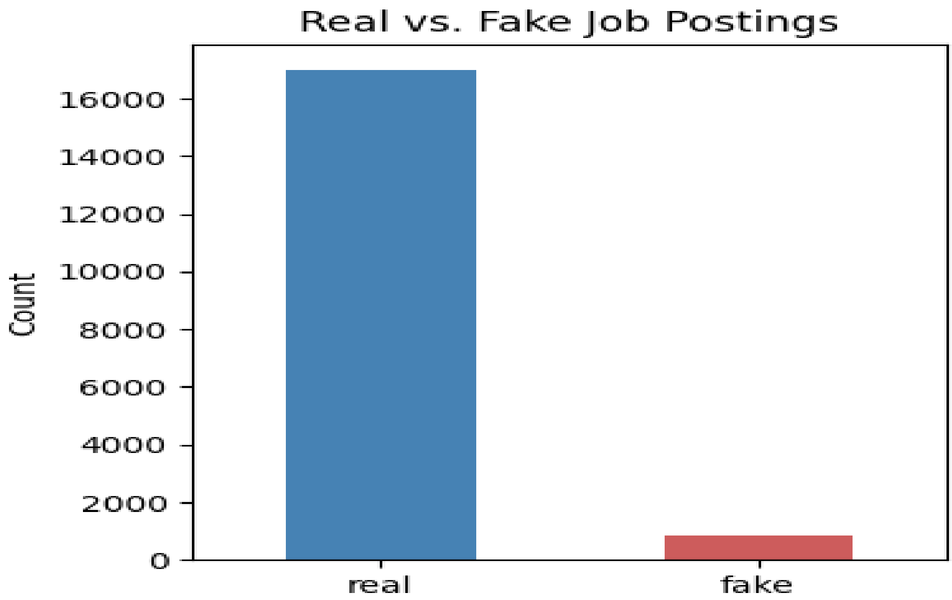
- Performance:
  - Accuracy: 0.96
  - Precision (fraud class): 0.89
  - Recall (fraud class): 0.82
  - F1-score: 0.85
  - ROC AUC: 0.97

- Explainability: SHAP waterfall plot generated for a high-risk ad (model probability = 92% fraud).

Outputs:
- Figures: PNG files Figures 1–5.

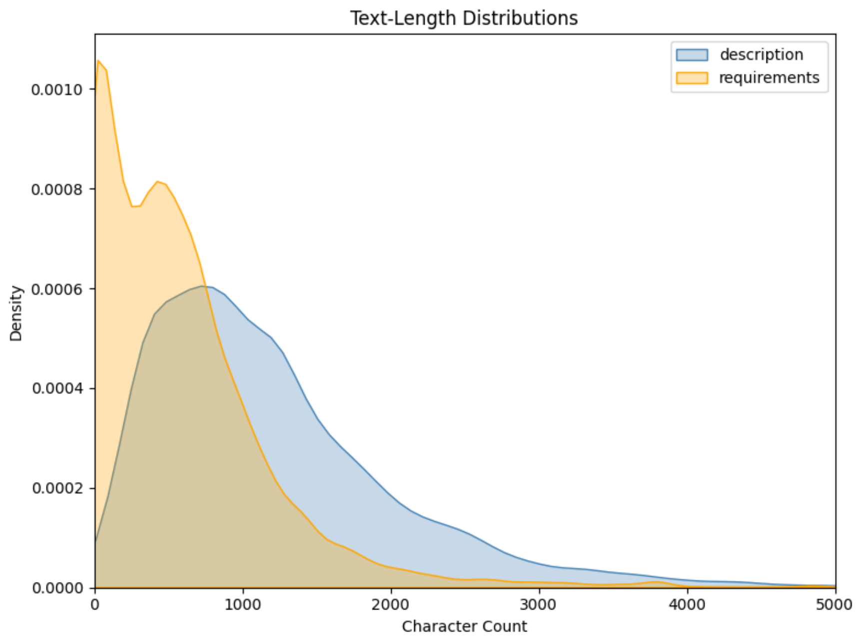
3 | Results & Visuals

Figure 1 – Class Balance



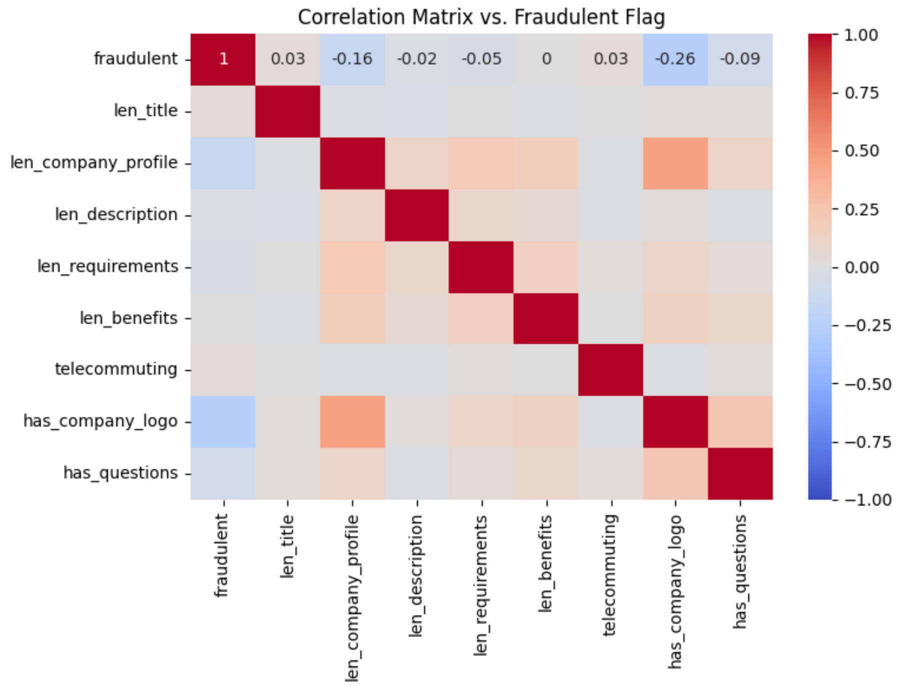
Severe imbalance: 17,199 real (94%) vs. 1,074 fake (6%).

Figure 2 – Text-Length Distributions



Legitimate posts use longer descriptions (median ≈850 chars) vs. requirements (≈550 chars). Fraudulent ads cluster at shorter lengths (e.g., fake descriptions median = 620 chars).

Figure 3 – Correlation Matrix



Key fraud indicators:

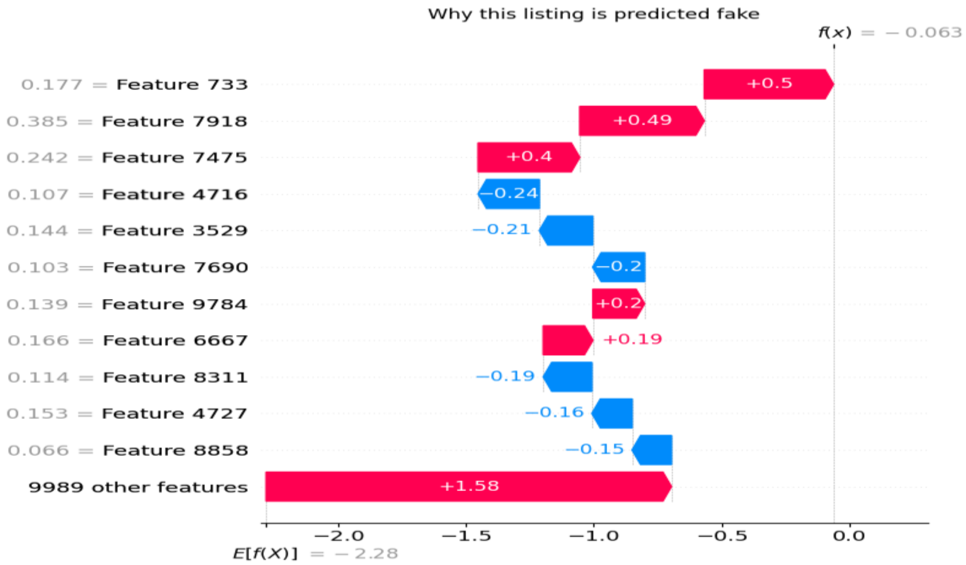
- has_company_logo: −0.26 (missing logo = strong red flag).
- len_company_profile: −0.16 (short/empty profiles = risk).
- telecommuting: +0.03 (weak link to fraud).

Figure 4 – Word Clouds

- Real ads: Role-specific terms (engineer, analyst, marketing).

Fake ads: Generic phrases (project, provide, solution), suggesting templated content.

Figure 5 – SHAP Waterfall



Model probability: 92% fraud.

- Top risks: No logo (+0.50), profile <150 chars (+0.49), phrases like "work from home" (+0.40).
- Mitigating factors: Technical keywords (−0.24), salary range (−0.19).

4 | Key Insights

1. Logo absence = critical signal: 26% correlation with fraud easy to screen in real time.
2. Content depth matters: Fake posts use shorter, vague descriptions (median 620 vs. 850 characters for real posts).
3. Language genericity: Fraudulent posts rely on boilerplate phrases (TF-IDF weights confirm this).
4. Explainability builds trust: SHAP clarifies model logic, enabling faster moderator decisions.
5. Strong baseline performance: Simple logistic regression achieves 96% accuracy (recall = 82% for fraud), with room for improvement.

5 | Recommendations

- Deploy hybrid screening:
  - Tier 1: Rule-based filter (logo check + profile length < 200 characters) to flag 40% of fraud instantly.
  - Tier 2: ML model for nuanced cases (e.g., posts with logos but generic text).
- Prioritize recall: Adjust thresholds to capture ≥ 90% of fraud (even with lower precision) to reduce victim harm.
- Enhance transparency: Display SHAP's top 3 risk phrases to moderators (e.g., "application fee").
- Future features:
  - Add readability scores (e.g., Flesch-Kincaid) and spelling-error rates.
  - Use NER to detect mismatches (e.g., company name ≠ profile).
- Monitor evolving tactics: Retrain monthly with new scam patterns (e.g., post-COVID remote work lures).

6 | Business Impact

Cost of inaction

- Current fraud volume: 1,074 fake ads in the dataset.
- Assumptions
  - Each scam ad reaches 10 applicants; 20 % (≈ 2 users) fall victim.
  - Average identity-theft remediation cost: $1,300 per victim (FTC 2023).
- Projected annual liability
  - 1,074 ads × 2 victims × $1,300 = $2.8 million

Return on prevention

- Cutting fraud in half saves ≈ $1.4 million in avoided losses.
- Manual review cost: $10 per flagged ad × 1 074 ads = $10,740
- Net savings: $1.4 M – $10.7 k ≈ $1.39 million (≈ 130 × ROI)

Customer retention

- 20 % of scam victims are likely to abandon the platform.
- Preventing 2 148 potential victims retains
  - 2,148 users × 20 % × $50 lifetime value = $21,480 in future revenue.

A lightweight fraud-detection pipeline pays for itself within days, mitigates $2.8 M in annual liability, and preserves both revenue and brand trust.