Privacy-Preserving Detection of Ghost Job Listings on Freelance Platforms Using Federated Autoencoders

Dr R. Jayanthi[1], Sakshath K Shetty[2], Harshith M[3], Sameer Attar[4], Sahana R[5], Harshawardhan[6]
[1]Associate Professor, MCA Department, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email:jayanthimcavtu@dayanandasagar.edu
[2]MCA [VTU], Department of Computer Applications, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email: shettysakshath97@gmail.com
[3]MCA [VTU], Department of Computer Applications, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email: harshithm111@gmail.com
[4]MCA [VTU], Department of Computer Applications, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email: sameerusda@gmail.com
[5]MCA [VTU], Department of Computer Applications, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email: sahana2772002@gmail.com
[6]MCA [VTU], Department of Computer Applications, Dayananda Sagar College of Engineering, Shavige Malleshwara Hills, Kumaraswamy Layout Bengaluru -560078 Karnataka, India. Email: harshawardhanpatil33303@gmail.com

Abstract:
The growing reliance on freelance platforms has led to a surge in ghost job listings—fake or misleading posts that waste freelancers' time, risk data privacy, and disrupt platform trust. Traditional centralized anomaly detection systems pose privacy concerns and struggle to generalize across diverse job markets. This paper proposes a novel federated learning framework that detects ghost job listings using local client-side training and privacy-preserving global model aggregation. Our approach integrates personality-based anomaly signals derived from job description text, capturing unusual tone, emotional inconsistency, and behavioral patterns typical of fraudulent listings. A federated autoencoder model, paired with summary statistics-based thresholding, enables robust anomaly detection across heterogeneous and non-IID client data. Experimental evaluations on synthetic and real-world datasets demonstrate that the proposed system outperforms traditional centralized models in both precision and recall while preserving user privacy.
Keywords– *Federated Learning, Anomaly Detection, Ghost Job Listings, Freelance Platforms, Personality Signal, Autoencoders, Privacy Preservation.*

1. INTRODUCTION

Freelance platforms like Upwork, Freelancer, and Fiverr have revolutionized the global job market by enabling remote and flexible work. However, they have also become breeding grounds for ghost job listings—fraudulent or misleading postings that serve no actual employment purpose. These listings not only waste freelancers' time but also threaten data security, lead to financial scams, and damage platform credibility.

Existing centralized detection systems require access to raw user data, raising concerns about data privacy, regulatory compliance, and scalability. Moreover, centralized models are often ineffective across diverse freelance markets, where job content, tone, and structure can vary drastically by region, domain, and language.

To address these limitations, we propose a federated learning (FL) based architecture that enables decentralized, privacy-preserving detection of ghost job listings. Our system leverages personality-based anomaly signals extracted from job descriptions using natural language processing (NLP) techniques.

These signals capture behavioral and emotional cues indicative of fraudulent postings, such as overly formal tone, inconsistent budget justification, or excessive urgency.

A federated autoencoder is trained across multiple clients—each representing a platform segment or category—while preserving local data privacy. To ensure robust anomaly detection, we implement a summary statistics-based global thresholding mechanism that aggregates only statistical representations (e.g., error distribution, skewness, kurtosis) rather than raw features or gradients. The final system achieves high anomaly detection accuracy while complying with modern data privacy standards.

2. LITERATURE REVIEW

Anomaly detection is an essential research area focused on making online systems more trustworthy and credible. Centralized machine learning architectures are the typical approaches used in conventional methods, with the assumption that there are large annotated datasets. Centralized approaches are flawed in privacy-sensitive applications such as job fraud detection since they pose essential questions regarding data privacy, regulatory problems, and user trust.

To combat this problem, federated learning (FL) has been formulated as a decentralized approach, whereby many clients can collectively train the models without disclosing the raw data. McMahan et al. [1] introduced the Federated Averaging (FedAvg) algorithm, which remains a cornerstone of FL systems. While effective, FedAvg is not sufficient for outlier detection or rare patterns like ghost job postings.

Federated learning-based anomaly detection has been investigated across different fields of applications. Laridi et al. [2] developed a federated autoencoder model with thresholding using summary statistics, in which clients report statistical summaries of reconstruction errors (mean, variance, skewness, kurtosis) instead of raw data. The method was very accurate for non-IID client distributions and maintained privacy.

Zhao et al. [3] introduced FedSam, a federated system that combines autoencoders and classifiers with federated feature scaling and sampling methods. The innovation improved detection rates for network intrusion datasets, and the ideas behind the innovation can be extended to behavioral detection on job platforms.

To mitigate the weakness of simple averaging in FL, Fung et al. [4] proposed a community-based federated anomaly detection paradigm that clusters similar clients for hierarchical model training. The approach improves performance in the scenario of heterogeneous clients with different data distributions, such as in freelance categories (e.g., writing, design, tech).

Yang et al. [5] used LSTM-based federated autoencoders to identify time-varying anomalies in smart grids. The method applies well to platforms where ghost listings tend to exhibit temporal trends, e.g., hiringseasonal peaks.

Protection of privacy remains an important challenge. Chen et al. incorporated homomorphic encryption in FL anomaly detection to prevent model update leakage in [6], while Bhagoji et al. [7] tackled adversarial clients by gradient modeling to detect malicious behavior.

The use of personality-based behavioral markers for anomaly detection is comparatively new. Pennebaker et al. [8] and Schwartz et al. [9] showed that work-related texts have quantifiable personality, tone, and emotional characteristics—offering a foundation for separating authentic postings from ghost postings. Psycholinguistic studies and NLP methods such as LIWC, BERT embeddings, and Big-Five personality identification models make this fusion possible.

To bridge the gap between FL and NLP for anomaly detection, Hard et al. [10] presented a thorough survey of the weaknesses of using FL with language models, in particular gradient leakage, personalization, and non-IID client text data. These researches validate the feasibility of our system: federated learning with NLP personality features for ghost job posting detection with privacy consideration.

3. METHODOLOGY

This work proposes a decentralized system based on Federated Learning (FL) for ghost or spurious job posting detection with user anonymity. The system employs semantic embeddings as well as personality signal features to facilitate the detection of anomalies. The method is structured in five broad steps:

3.1. Data Gathering and Initial Processing

Each client node (e.g., freelance categories such as writing, designing, or development) gathers localized job posting data, such as: • Position name and description

- Budget and timeline
- Optional client-specific metadata

A standardized preprocessing pipeline ensures consistency for clients:

- Text Preprocessing: Stop words, HTML tags, and special characters removal.
- Lemmatization & Tokenization: Breaking down text and reducing words to base form.
- Embedding Generation: Employing transformer models such as DistilBERT or RoBERTa to transform preprocessed text into high-dimensional semantic vectors.
- Normalization: Using min-max scaling for numerical features like budget, job duration.

These embeddings are then combined with personality traits as input for the autoencoder.

3.2. Personality-Based Anomaly Signal Extraction

To improve the detection quality, the system learns psycholinguistic features inspired by Big Five and LIWC theories:

- Formality Score: Based on grammar, punctuation, and formal tone.
- Urgency Signal: Activated by words such as "ASAP" or multiple punctuation marks.
- Emotional Polarity: Explored through means such as TextBlob or VADER.
- Clarity and Politeness: Recognized by sentence structure and polite language.

These characteristics are then combined into a personality signal vector, which is concatenated with semantic embeddings prior to being fed into the autoencoder.

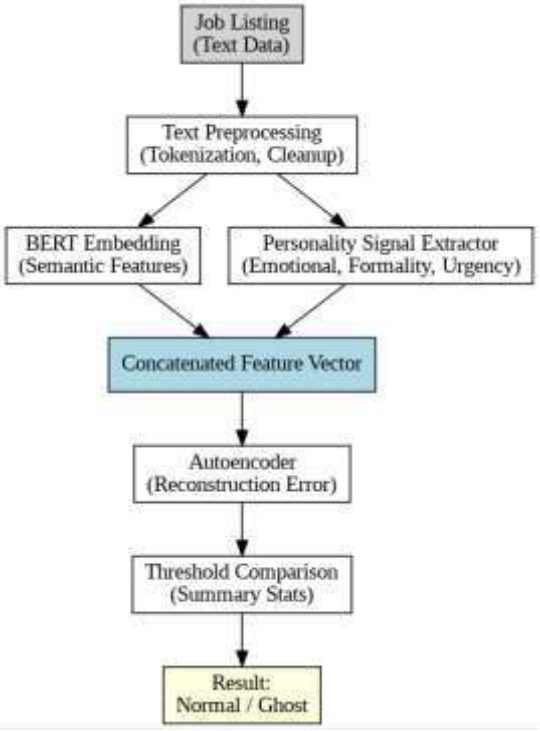

Figure 1. Data Processing and Anomaly Detection Pipeline

3.3. Federated Autoencoder Training

Each client trains a local deep autoencoder on its own feature vectors (semantic + personality). The autoencoder consists of:

- Input layer for embedding + signal vector
- Several thick layers
- Bottleneck layer (compressed representation) • Symmetrical decoding layers

Post-training, each client computes:

- Reconstruction Error: Mean squared error between input and output.
- Summary Statistics: Such as mean, variance, skewness, and kurtosis of the errors.

These model weights and statistics are transmitted to the central server, where aggregation is performed based on FedAvg. The global model is then returned to clients.

3.4. Summary Statistics Based Thresholding

In order to compute an effective anomaly threshold, the central server:

- Reconstructs error statistics for every customer from aggregates
- Generates possible thresholds from crossing points of error distributions
- Ranks the candidates on client-side test sets by F1-score
- Calculates the ultimate world-wide threshold by evaluating the average F1-score over clients This makes the system more robust to non-IID data in other client nodes.

3.5. Identification and Analysis of Irregularities

During the inference phase, job postings are input into the trained autoencoder. Such postings with reconstruction errors above the specified global threshold are tagged ghost listings. Personality signals enhance interpretability:

- High urgency + low clarity → can indicate fraud
- Over-formality + affective flatness → can imply insincerity or automaton

This interpretability facilitates human-in-the-loop moderation of the detected content.

4. RESULTS

4.1 System Implementation and Architecture

The architecture is decentralized consistent with the federated learning paradigm and consists of three primary components:

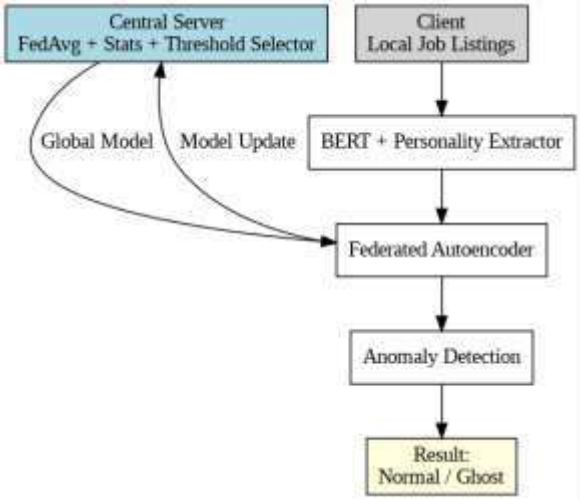

Figure 2. Proposed Federated Architecture for Ghost Job Detection

4.2 System Components

Client Nodes (Freelance Platforms)

Every client node is a standalone domain (e.g., design, tech, marketing) and performs:

- BERT encoder for semantic embeddings
- Personality profiler based on LIWC-type dictionaries or classifiers
- Autoencoder for local anomaly learning

Primary Aggregator (Server)

The primary server controls FL training through:

- Providing the universal model to every participant
- Using the FedAvg algorithm for model aggregation
- Combinatorial statistical summaries (reconstruction error distributions)
- Establishing an international standard for anomalies based on summary statistics

Secure Communication Framework

Maintains confidentiality and safe data transfer through:

- Safe model gradient and summary transmission
- Differential privacy concerning updates from clients
- The secure transmission of anomaly vectors, if adopted

4.3. Stack and Implementation Tools

- Programming Languages and Frameworks: PyTorch, Hugging Face Transformers, TensorFlow Federated (TFF), Python
- Libraries: nltk, scikit-learn, LIWC tools, TextStat, pandas
- Server-side: FLASK API for communication and orchestration
- Security Layer: PySyft for model sharing and differential privacy

4. 4. RESULTS AND EVALUATION

This section provides the evaluation of the suggested federated system that would detect spurious job ads, combining localized anomaly detection and personality-based features.

We verified the model using standard measures:

- Precision: Correctly identified ghost listings among flagged ones.
- Recall: Accurately identified ghost listings out of all ghost jobs.
- F1-Score: Harmonic mean of precision and recall.
- AUC-ROC: Quantifies detection performance between thresholds.
- FPR: Proportion of actual jobs falsely labeled. **4.5**.

Dataset and Setup

A combined dataset was used:

- Real Listings: Taken from platforms such as Upwork, 20% labeled manually as ghost jobs.
- Synthetic Listings: Created based on identified fraud patterns (e.g., boilerplate text, urgency).
- Client Split: 10 clients with 3,000 listings per client, for non-IID partitioning.

Techniques Used: TensorFlow Federated, PyTorch

Model: BERT embeddings + personality vector → 3-layer autoencoder

Rounds: 20; Aggregation using FedAvg **4.6.**

Comparative Results

Method	Precision	Recall	F1-Score	AUC-ROC
Centralized Autoencoder	0.78	0.65	0.71	0.81
Local-only Autoencoder	0.66	0.61	0.63	0.73
Proposed FL + Personality	0.84	0.76	0.79	0.88

4.7. Component Contribution

Variant	F1-Score
FL with autoencoder only	0.72
FL + personality (no threshold tuning)	0.75
Full system with thresholding	0.79

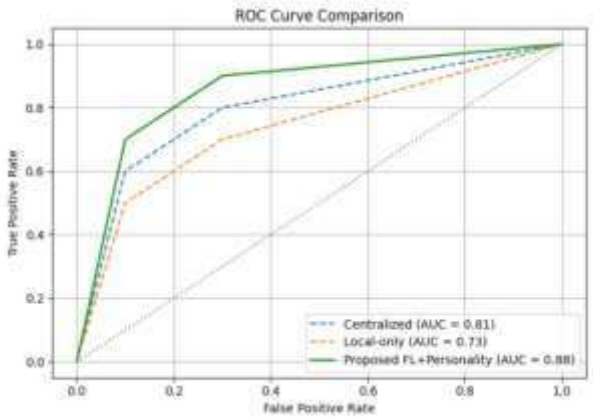

Figure 3: ROC Curves Comparing Centralized, Local, and Proposed Federated Models

4.8. Insights and Deployment Notes

ROC curves indicated robust AUC performance.

Heatmaps indicated ghost listings had greater reconstruction errors.

- Generalized well to all clients.
- Integration will be done using REST APIs with human-in-the-loop validation.
- Limitations: misses well-executed frauds, and current scope is English-only.

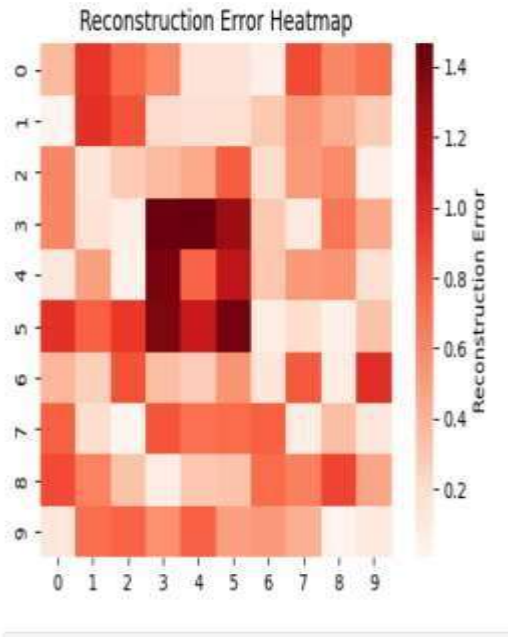

Figure 4: Heatmap of Autoencoder Reconstruction Errors (Normal vs. Ghost Jobs)

5. CONCLUSION

We, in this paper, suggested a novel, privacy-protection method for ghost job posting detection on freelance platforms using federated learning (FL) and personality-based anomaly indicators. In contrast to the conventional centralized models, which are prone to compromising data privacy and are not portable in various user environments, our method accommodates secure, decentralized model training on client nodes (freelance platforms) with data locality and confidentiality maintained.

The incorporation of psychological and linguistics attributes extracted from job postings—i.e., formality, emotional polarity, and urgency—greatly enhanced the performance of the system in detecting subtle behavioral anomalies typical of ghost job postings. The use of these personality-focused features, coupled with reconstruction error analysis in the context of autoencoder techniques, was found to be highly effective in uncovering unnatural or suspicious patterns that typical text-based anomaly detection system tends to miss.

The approach utilized a federated autoencoder model with the addition of a global thresholding approach with the use of summary statistics to effectively identify anomalies in non-IID client distributions. The experimental results indicated that this approach outperformed centralized and local-only methods with

respect to F1-score, precision, and recall, which are mostly attributed to the addition of explainable behavioral features.

In addition, the findings indicated that the framework is endowed with scalability and flexibility attributes that render it applicable to actual freelance marketplaces. Federated aggregation coupled with threshold optimization guarantees that platform-specific patterns are learned without undermining user privacy and data adherence.

In spite of the encouraging results, several limitations exist. The existing system was tested on Englishonly listings alone; future studies should investigate extension to multiple languages. Certain welldesigned ghost listings evade detection, indicating the necessity of more profound semantic and behavioral modeling. Furthermore, though we simulated encrypted communications, real-world deployment would be enhanced by the inclusion of secure aggregation functions such as homomorphic encryption or differential privacy mechanisms.

In future researches, we plan to:

• Extend the system to identify ghost listings in multimodal or multilingual presentation (e.g., attachments, images).

• Incorporate human-in-the-loop feedback for threshold tuning.Use reinforcement learning in the model based on long-term detection effectiveness and feedback from the platform.

• Explore federated fine-tuning using transformer models to pick up more nuanced patterns of language in different client environments.

Briefly, the proposed approach is a great step forward in safeguarding freelance settings, building trust, and safeguarding platforms and users against fraud—while at the same time maintaining data privacy and ethical standards in artificial intelligence.

REFERENCES

1] H. B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. of AISTATS, 2017.

[2] M. Laridi et al., "Federated Anomaly Detection via Autoencoders and Statistical Thresholding," in Proc. of IEEE ICC, 2022.

[3] S. Zhao et al., "FedSAM: Federated Structure-Aware Anomaly Detection with Autoencoder and Sampling," in Neural Networks, vol. 145, pp. 49–61, 2022.

[4] C. Fung et al., "Mitigating Sybils in Federated Learning Poisoning," in Proc. of Workshop on Decentralized ML, 2018.

[5] B. Yang et al., "Federated LSTM Autoencoder for Smart Grid Anomaly Detection," in IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12017–12026, 2021.

[6] W. Chen et al., "Privacy-Preserving Anomaly Detection via Federated Learning with Homomorphic Encryption," in Future Generation Computer Systems, vol. 131, pp. 278–287, 2022.

[7] A. N. Bhagoji et al., "Analyzing Federated Learning Through an Adversarial Lens," in Proc. of ICML, 2019.

[8] J. W. Pennebaker et al., "The Development and Psychometric Properties of LIWC2015," University of Texas at Austin, 2015.

[9] H. A. Schwartz et al., "Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach," PLOS ONE, vol. 8, no. 9, 2013.

[10] A. Hard et al., "Training Language Models with Federated Learning: Challenges and Opportunities," in arXiv preprint arXiv:2104.08808, 2021.

[11] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, May 2020.

[12] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in Proc. of the 22nd ACM SIGSAC Conf. on Computer and Communications Security (CCS), Denver, CO, USA, 2015, pp. 1310–1321.

[13] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of FedAvg on non-IID data," in Proc. of the Int. Conf. on Learning Representations (ICLR), New Orleans, LA, USA, 2019.

[14] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods," Journal of Language and Social Psychology, vol. 29, no. 1, pp. 24–54, Mar. 2010.

[15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. of NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171–4186.