Spot The Scam

Fraudulent Job Posting Detector

Alice Karakachian, David Pu, David Nguyen,
Sai Ananya Channamsetti, Lakshmi Gayathri Divakaruni

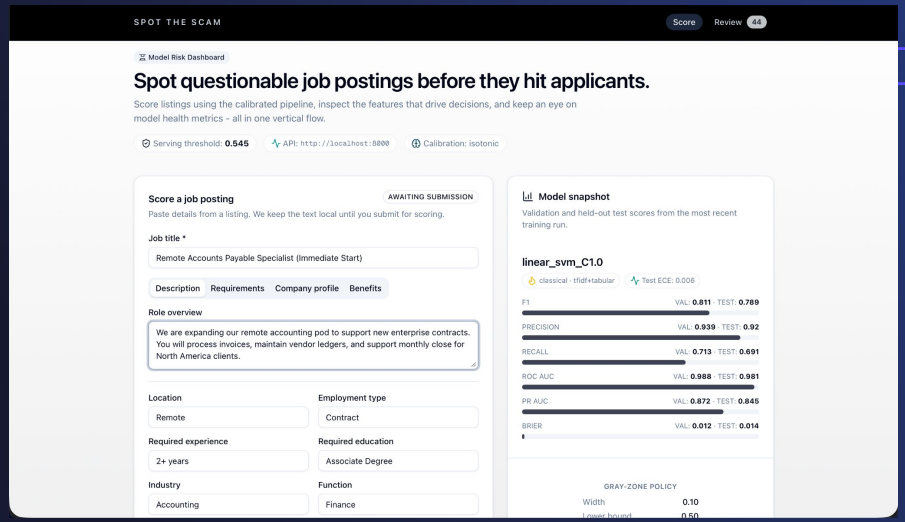
Introduction

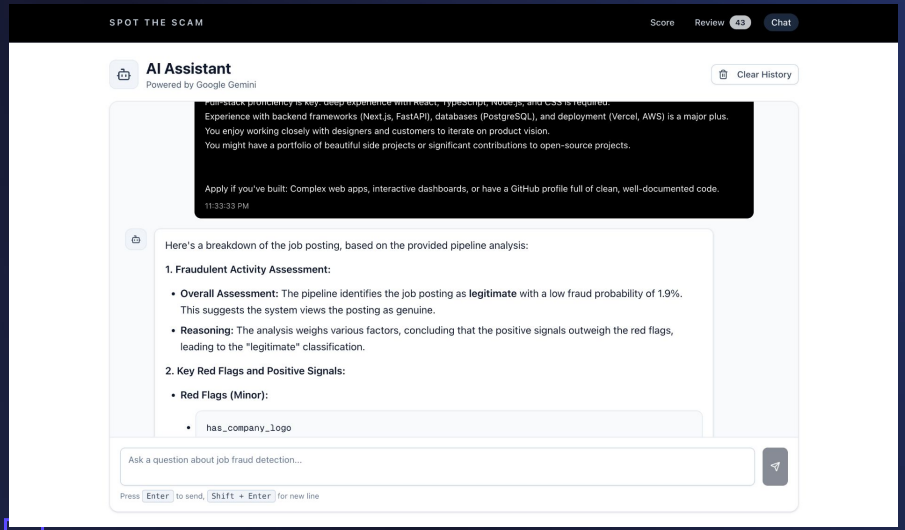
The Problem:

- Technology has increased accessibility to information about employment opportunities
- However, with this, there has also been an increase in fraudulent job postings
  - Collecting personal information, trying to take job seekers' money through upfront fees

Our project:

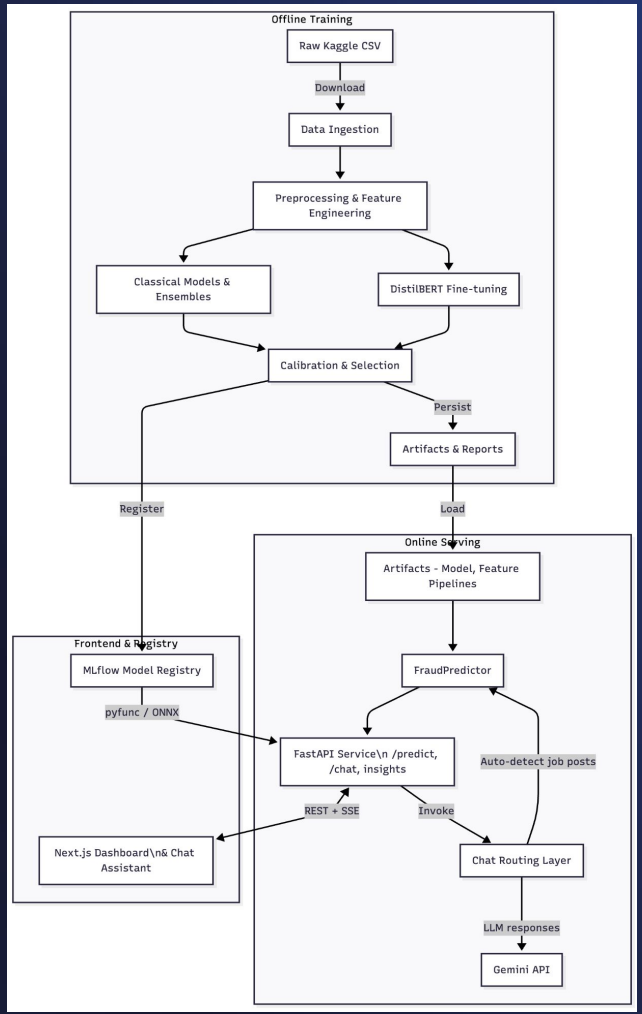
- **Aim:** Improve the accuracy of the identification of fraudulent job postings
- **Our solution:** full-stack application that applies both classical and transformer models to this modern problem





Methods

- Kaggle job-post datasets merged. Used 70/15/15 split for model training, validation, and testing.
- Classical models: Logistic Regression, Linear SVM, LightGBM, XGBoost
- Transformer model: fine-tuned DistilBERT classifier
- All candidates tuned and ranked by validation F1; best calibrated model evaluated on test
- Gemini API chatbot wraps the model using conversation history, job fields, and model outputs to produce natural language responses



Results

Table 1: Performance of `ensemble_top3` on validation and test splits.

Split	F1	Precision	Recall	ROC AUC	PR AUC	Brier
Validation	0.8561	0.9297	0.7933	0.9890	0.9053	0.0103
Test	0.7721	0.8537	0.7047	0.9863	0.8659	0.0143

- Validation winner: calibrated classical ensemble ensemble_top3, selected over transformer baselines by validation F1
- Generalizes well: strong F1 on validation and test with high ROC/PR AUC and low Brier (see table)

References

[1] Sultana Umme Habiba, Md Khairul Islam, and Farzana Tasnim. A comparative study on fake job post prediction using different data mining techniques. In 2021 *2nd international conference on robotics, electrical and signal processing techniques (ICREST)*, pages 543–546. IEEE, 2021.

[2] Vijay Madaan, Neha Sharma, Raghubeer Singh Bangari, and Srinivas Aluvala. Fraudulent job posting detection using logistic regression. In *2024 International Conference on Information Science and Communications Technologies (ICISCT)*, pages 1–6. IEEE, 2024.

[3] Gayathri Malaichamy, Cristina Hava Muntean, and Anderson Augusto Simiscuka. Online job posting authenticity prediction with machine and deep learning: Performance comparison between n-gram and tf-idf. In *International Conference on Deep Learning Theory and Applications*, pages 143–162. Springer, 2024.

[4] K Ramakrishna Reddy, G Indrani, N Pavan Kumar, and K Vamshi Krishna. Fake job posting detection using machine learning algorithms. In *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pages 822–827. IEEE, 2025.

[5] K. Suparna and R. Anil Kumar. A comparative study on fake job post prediction using different datamining techniques. *International Journal of Information Technology and Computer Engineering*, 12(3):317–323, 2024

GitHub Repo

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**

Thank you!