Data Science Home Assignment: Scam Detection and Intervention

Objective:

Develop a system capable of identifying and disrupting scam scenarios through brief interactions with potential victims.

Context:

Scams are evolving in sophistication, making it increasingly difficult for individuals to recognize fraudulent schemes. Your task is to create a model or system that can engage in a short conversation (text-wise) with a potential scam victim, assess the situation, and, if needed, intervene effectively to break the "scam spell."

Assignment:

You are provided with a dataset containing 1,000 examples of scam procedures. Your task is to build a solution that can:

1. **Detect** whether an individual is involved in a scam based on a short textual conversation.
2. **Intervene** effectively by communicating in a way that makes the victim aware of the scam and encourages them to disengage from it.

Key Considerations:

- You are free to choose your methodology, whether it is a machine learning model, a retrieval-augmented generation (RAG) approach, or other AI/heuristic-driven techniques.
- The system should be capable of handling a variety of scam types, including impersonation, investment scams, and social engineering attacks.
- Focus on real-world applicability: responses should be convincing, empathetic, and actionable.
- Consider adversarial strategies that scammers use to keep victims engaged, and develop techniques to counteract them.
- Your solution should go beyond simple keyword detection and demonstrate an understanding of scam tactics and human behavior, i.e., if the model can’t ‘break the spell,’ then it did not achieve its goal.

Deliverables:

- A functioning system or prototype that demonstrates the ability to detect and intervene in scam scenarios.
- A report detailing:
    - Your approach and rationale.

    - The methodology used (e.g., ML model, RAG, heuristics, etc.).
    - Any key challenges faced and how you addressed them.
    - Performance metrics and qualitative examples.
- A brief demonstration of how your system interacts with a potential scam victim and breaks the scam spell.

Evaluation Criteria:

- **Effectiveness:** How well does your system detect and intervene in scam scenarios? Or comparing to out-of-the-box vanilla models?
- **Innovation:** How novel is your approach compared to traditional scam detection methods?
- **Robustness:** Can your system adapt to different scam variations?
- **Clarity & Justification:** Is your approach well-reasoned and explained in the report?

This assignment is intentionally open-ended to allow creativity in problem-solving. The focus is not just on detection but on actionable intervention.

Timeframe:

You have **1 week** to complete this assignment. Please manage your time accordingly to ensure a comprehensive and well-structured submission.

Confidentiality Notice:

This document and the provided dataset are strictly confidential. Any use, sharing, or distribution of this material outside of the intended purpose is prohibited. Candidates are expected to handle all information responsibly and ensure that no data is exposed or misused.

Good luck!