CODE 4 GovTech
Building Digital Public Goods

C4GT DMP - Proposal Template

Name	BHUMIKA S
Email ID	bhumikasp.2004@gmail.com
Phone Number	8431465524
GitHub ID	https://github.com/Bhumika-SP
Discord ID	https://discord.com/users/bhumika_2608_34831
Current occupation *(Working Professionals - add current organization & years of exp)*	Student
Education Details *(College Name - Degree Name and branch of engineering or other course/specialization)*	College Name – Rajarajeshwari College of Engineering Branch – Computer Science and Engineering
Technical skills with level *(Mention tech skills/languages known/UI-UX and level - Novice/Intermediate/Expert)*	Languages: C, Java, JavaScript, Python |Web Development: HTML, CSS, Bootstrap, Tailwind, React, Node.js, Express.js| Database: MySQL, Mongo DB Machine Learning Libraries: Numpy, Pandas, Scikit-Learn, Matplotlib| Platforms: PyCharm, Jupiter Notebook, IntelliJ IDEA, AWS, GitHub

Title:

Designing an ML-Based Fraud Detection System for Real-Time Financial Risk Monitoring

---

Summary:

This proposal presents my plan to build a modular fraud detection system using machine learning techniques tailored for financial applications. The goal is to create a robust backend module that can detect anomalous transactions in real time by learning patterns from historic data. My approach focuses on data preprocessing, model training, integration with APIs, and visual feedback to support actionable insights.

Project Detail

1. Project Overview

a. Understanding of the project
The project aims to develop a backend service for fraud detection that can process financial transaction data and flag suspicious behavior. The module should support streaming or batch inputs and return a probability score or label for fraud. It should be designed to plug into a larger finance application or dashboard.

b. Issues that might come up and the support needed from org

- Availability of clean labeled transaction data.
- Real-time processing pipeline integration.
- Guidance on deployment best practices for production-readiness.

c. Solutions

- Use open datasets (e.g., Kaggle, IEEE-CIS) for prototyping.
- Build a modular design (data ingestion → model inference → alert).
- Consult with mentors for integration points, testing strategy, and model evaluation.

---

2. Macro Implementation Details with Timelines

Milestone 1 (Week 1–2):

- Research and select relevant datasets.
- Perform data cleaning, exploratory analysis, and feature engineering.
- Set up basic binary classification model (e.g., XGBoost or Logistic Regression).

Milestone 2 (Week 3–4):

- Integrate the trained model into a microservice (FastAPI/Flask).
- Design APIs for prediction and result logging.
- Set up performance dashboards (e.g., streamlit or simple frontend).

Milestone 3 (Week 5–6):

- Test against unseen data and fine-tune thresholds.
- Add logging, monitoring, and basic visualization of alerts.
- Prepare documentation, deployment script (Docker), and submit final PR.

Availability

- **Available hours/week:** 20–25 hours
- **Other engagements:** None currently planned
- **Notes:** I am highly available and committed to this project for the full duration. I can adapt to any sync schedule as needed by the mentors.

Personal Information

About Me:
I am a developer and learner passionate about building ML-based applications, especially in the finance and security domain. I enjoy solving real-world problems using data.

Motivation:
I am applying for this project to contribute to building a scalable fraud detection system and improve my real-world machine learning deployment skills. I've previously worked on personal projects in this space and am excited to take that further under guidance and through open-source collaboration. C4GT offers a great platform to learn, build, and give back — and I'm eager to be part of that journey.

Please mention if you have solved any issues/tickets for this or other C4GT projects: (Optional)

Link to to Issue	Resolution description in short	Link to pull request
https://github.com/AgentTorch/visualize/issues/3	Added example usage with inline comments for GeoPlot visualization to improve developer onboarding and open source clarity.	https://github.com/AgentTorch/visualize/pull/62

Add more rows if required