**Project: Real-Time Fraud Detection in Financial Transactions *(Big Data + AI + ML)***

**🚀 Goal:**

Develop a **real-time fraud detection system** to identify suspicious financial transactions using **Big Data, AI, and Machine Learning**, ensuring enhanced security for banking and fintech applications.

**🔹 Step-by-Step Execution Plan:**

**1️⃣ Data Ingestion & Streaming (Big Data Pipeline)**

✅ **Goal:** Collect real-time financial transaction data and enable continuous data flow.  
✅ **Tech Stack:** Apache Kafka (Open Source), PostgreSQL (Open Source), JSON/CSV datasets.  
✅ **Implementation:**

* Set up **Kafka Streams** to ingest real-time transaction data from banking APIs or simulated datasets.
* Store raw transactions in **PostgreSQL/Snowflake** for further analysis.

**2️⃣ Data Processing & Anomaly Detection (Stream Processing & ML)**

✅ **Goal:** Process financial transactions in real time and detect suspicious activities.  
✅ **Tech Stack:** Apache Flink (Open Source), Scikit-learn, PyOD (Open Source Anomaly Detection Library).  
✅ **Implementation:**

* Use **Apache Flink** to apply **stream processing** and detect anomalies dynamically.
* Implement **unsupervised learning models (Isolation Forest, One-Class SVM, LOF)** for anomaly detection.
* Train **Random Forest/XGBoost** models on historical fraud datasets to classify transactions as fraudulent or legitimate.

**3️⃣ Model Deployment & Real-Time Fraud Classification**

✅ **Goal:** Deploy the trained fraud detection model for real-time classification of transactions.  
✅ **Tech Stack:** Flask/FastAPI (Open Source), Docker (Containerization), Scikit-learn.  
✅ **Implementation:**

* Develop an **API-based fraud detection service** using **Flask or FastAPI**.
* Deploy the model on a local server or **cloud (GCP/AWS Free Tier)** using **Docker**.

**4️⃣ Visualization & Dashboard for Fraud Analytics**

✅ **Goal:** Build an interactive dashboard for financial analysts to monitor fraud trends.  
✅ **Tech Stack:** Tableau (Public Version for Free), Power BI, Streamlit (Open Source).  
✅ **Implementation:**

* Design an **interactive dashboard** with fraud detection insights, transaction patterns, and alert logs.
* Highlight **fraud probability, transaction locations, and anomaly scores** for better decision-making.

**💡 Value Add & Impact:**

🚀 **Real-time fraud detection**, reducing financial risks and improving transaction security.  
🚀 **Big Data + AI + ML** combination strengthens your profile for fintech and data-driven roles.  
🚀 Uses **open-source & free tools** where possible, minimizing infrastructure costs.

**🔥 Final Output:**

* **Fraud Detection Pipeline** (Kafka + Flink + PostgreSQL).
* **AI Model for Anomaly Detection** (Scikit-learn, PyOD, XGBoost).
* **API Deployment** for real-time fraud classification (Flask/FastAPI).
* **Interactive Fraud Monitoring Dashboard** (Tableau/Power BI/Streamlit).