**Fraud Detection System**

This project detects fraudulent financial transactions using a Random Forest model, with real-time streaming via Kafka and visualization on a Streamlit dashboard.

**Features**

* **Data Processing**: Analyzes transaction data from Kaggle.
* Fraud Detection: Random Forest model predicts fraud.
* Streaming: Kafka streams transaction data.
* Dashboard: Streamlit displays real-time fraud predictions.

**Setup**

Download dataset: https://www.kaggle.com/datasets/sriharshaeedala/financial-fraud-detection-dataset

Install Kafka (https://kafka.apache.org/):

bin/zookeeper-server-start.sh config/zookeeper.properties

bin/kafka-server-start.sh config/server.properties

**Install dependencies:**

pip install streamlit altair pandas numpy scikit-learn kafka-python

**Run scripts:**

* python train\_model.py
* python producer.py
* python consumer.py
* streamlit run dashboard.py
* View dashboard at http://localhost:8501.

**Files**

* **train\_model.py:** Trains and saves the fraud detection model.
* **producer.py:** Streams transactions to Kafka.
* **consumer.py:** Applies model and saves predictions.
* **dashboard.py:** Displays results in Streamlit.