**Advanced Credit Card Fraud Detection using Machine Learning and Power BI**

**Introduction**

With the rise of digital payments, credit card fraud has become a significant challenge for financial institutions. Traditional rule-based systems fail to detect subtle patterns used by fraudsters. This project aims to develop a machine learning-based fraud detection system that can intelligently flag suspicious transactions and help reduce financial loss. We also integrate **Power BI** to visualize trends and improve decision-making for fraud analysts.

**Project Overview**

This project uses a real-world credit card transactions dataset that includes over 280,000 records with anonymized features. The goal is to build a **classification model** that can distinguish between genuine and fraudulent transactions with high accuracy and recall, especially since fraudulent cases represent a very small percentage of the data (imbalanced dataset problem).

The ML output is then visualized using **Power BI**, allowing analysts to interact with fraud trends, risk indicators, and transaction behavior.

**💻 3. Programming Language & Tools Used**

* **Language**: Python
* **Libraries**:
  + Pandas, NumPy – for data handling
  + Scikit-learn – for ML models and evaluation
  + XGBoost, Logistic Regression, Random Forest – models used
  + Matplotlib, Plotly – for visualization
  + imblearn – for handling class imbalance (SMOTE)
* **Power BI**:
  + Used to create dashboards with fraud trends, KPIs, and filters
  + Integrated using CSV exports from the Python model

**Training the Model (Workflow)**

**Data Preprocessing**:

* + Remove duplicates
  + Scale Amount and Time using StandardScaler
  + Drop irrelevant columns (if any)

**Handle Imbalance**:

* + Apply SMOTE to balance fraud/genuine classes in training data

**Model Training**:

* + Train models using train\_test\_split (80:20)
  + Evaluate using metrics: **Recall**, **Precision**, **F1-score**, **ROC-AUC**

**Prediction & Output**:

* + Predict fraud probability on test set
  + Save results with transaction IDs, amounts, and fraud scores

**Power BI Integration**:

* + Export cleaned + prediction data as .csv
  + Load into Power BI for dashboard creation

**Merits**

* 🎯 **High Precision & Recall**: ML models outperform manual rules
* 📊 **Power BI Dashboards**: Easy monitoring of fraud trends
* 🔁 **Retrainable**: Model can improve with more labeled data
* ⚡ **Real-time capable**: Can be deployed as an API for streaming detection

**⚠️ Demerits / Limitations**

* ❗ **False Positives**: May block genuine users occasionally
* 🔍 **Data Privacy**: Needs secure handling of sensitive data
* 💾 **Model Drift**: Patterns change over time, requiring regular retraining
* ⚖️ **Explainability**: Advanced models (like XGBoost) are harder to interpret without SHAP

**🏁 Conclusion**

This project demonstrates how **Machine Learning** can revolutionize the fight against credit card fraud when combined with **Power BI dashboards**. With intelligent prediction models and dynamic visualization tools, businesses can quickly detect suspicious activities, minimize financial loss, and stay ahead of evolving fraud tactics. Future extensions could include deploying the model in real-time systems, incorporating feedback loops, and improving explainability using tools like SHAP.