**DOCUMENTATION**

**KHUSHI TIWARI(BATCH-4)**

**Description: Problem Statement-1**

Develop an AI model capable of detecting fraudulent transactions in real-time. Use historical transaction data to train the model to identify anomalies and flag suspicious activities.

**Evaluation Matrix:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S.NO | Precision (30 POINTS): | Recall (Sensitivity) (30 POINTS): | F1 Score (5 POINTS) | AUC-ROC (5 POINTS): | Feature Importance (5 POINTS): | Insights from Data (10 POINTS): | Progress During Hackathon (5 POINTS): | Innovation & Documentation (10 POINTS): |
| 1 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

**Model details & Configurations**

**1. Model Choice:** XGBoost Classifier and Random Forest Classifier

**2. Key Configurations:**

* Objective: binary: logistic

This specifies that the model is used for binary classification problems, where the output is either 0 or 1 (non-fraudulent or fraudulent).

* Evaluation Metric: logloss

Logarithmic loss measures the performance of the classification model whose output is a probability value between 0 and 1. It is used to evaluate how well the predicted probabilities match the actual labels.

* Use Label Encoder: False

This avoids the deprecated warning related to the label encoder, ensuring compatibility with the latest versions of XGBoost.

* Random State: 42

This sets the random seed for reproducibility of the results. Using a fixed random state ensures that the same results can be achieved if the code is run multiple times.