PHASE II – DOCUMENT

STUDENT NAME : **G.PREMSUNDARI**

REGISTER NUMBER : **623323205022**

INSTITUTION : **VETRI VINAYAHA COLLEGE OF ENGINEERING AND TECHNOLOGY**

DEPARTMENT : **B.TECH INFORMATION TECHNOLOGY**

DATE OF SUBMISSION : **07/05/2025**

GITHUB REPOSITORY LINK : **https://github.com/Prem20066/Guarding-transaction-with-AI-powered-fraud-credit-card-detection-and-prevention-.git**

1.PROBLEM STATEMENT

Credit card fraud continues to challenge financial institutions, with increasing losses Each year. Traditional fraud detection systems based on static rules are limited in Handling complex, evolving fraud techniques. This project aims to build an AI-driven System capable of detecting and preventing fraudulent credit card transactions in Real time using advanced machine learning models.

2.PROJECT OBJECTIVES

- **Build an intelligent system that can detect and prevent fraudulent credit card Transactions.**
- **Reduce both false positives (legitimate transactions flagged as fraud) and false Negatives (missed frauds).**
- **Analyze patterns in transaction data to understand fraudulent behavior.**
- **Develop a scalable, efficient model suitable for real-time deployment.**

- **Implement a feedback mechanism to improve accuracy over time.**
- **Visualize results to support financial decision-making.**

3.FLOWCHART OF THE WORKFLOW

Start

|

V

Data Collection → Preprocessing → Feature Engineering

|

V

Model Selection → Model Training & Evaluation

|

V

Real-Time Detection → Fraud Prevention Mechanism

|

V

Feedback Loop → Visualization & Insights

|V

End

4.DATA DESCRIPTION

- **Sources:**

- **- Kaggle: Credit Card Fraud Detection dataset**

- **- Simulated transactions for real-time testing**
- **- Optionally anonymized partner institution data**

- **Features:**

- **- PCA-transformed inputs (V1–V28)**
- **- Time, Amount, and Class (0 = legitimate, 1 = fraud)**

5.DATA PROCESSING

- **Remove missing or duplicate entries**
- **Normalize numerical features (e.g., Time, Amount)**
- **Address class imbalance using SMOTE or under-sampling**
- **Split into training, validation, and testing datasets**

6.EXPLORATORY DATA ANALYSIS (EDA)

- **Analyze distribution of fraud vs. legitimate transactions**
- **Correlation heatmap between features and target class**
- **Visualize key variables (box plots, histograms)**
- **Analyze time-based transaction trends**

7.FEATURE ENGINEERING

- **Log transform for skewed features (e.g., Amount)**
- **Create temporal features (hour, day of week)**

- Use PCA/autoencoders for dimensionality reduction
- Encode categorical features (if applicable)

8.MODAL BUILDING

Compare multiple algorithms:

- - Logistic Regression
- - Decision Tree, Random Forest
- -XGBoost, Neural Networks

Evaluation metrics:

- Precision, Recall, F1-score, ROC-AUC
- - Confusion matrix
- - Cross-validation for model robustness

9.REAL TIME DETECTION & PREVENTION

- Integrate streaming API for live transaction analysis
- Auto-blocking mechanism for flagged transactions
- Trigger alert notifications to stakeholders
- Adaptive learning system to retrain with new cases

10.VISULAIZATION OF RESULTS AND MODEL INSIGHTS

- Confusion matrix, ROC and PR curves
- Feature importance charts

- **t-SNE or PCA plots to visualize class clusters**
- **Dashboard for fraud trends and alerts**

11. TOOLS AND TECHNOLOGIES

Language	**Python**
Libraries /Framework	**NumPy**
Visualization	**Matplotlib**
Notebook /IDE	**Google colob**
Deployment	**FastAPI**
Database	**MySQL**
Version Control	**Git**