Phase-1

Student Name: Inbarasu I

Register Number: 410723106012

Institution: Dhanalakshmi college of engineering

Department: Electronics and Communication Engineering

Date of Submission: 28/04/2025

1.Problem Statement

AI-Based Credit Card Fraud Detection and Prevention

1. Credit card fraud is rising rapidly with the growth of digital transactions.
2. Traditional rule-based systems fail to detect evolving fraud patterns.
3. This project uses AI and Machine Learning to build a smart, adaptive detection system.
4. It learns from transaction behavior to identify fraud in real time. the goal is to reduce financial loss and enhance payment security.

2.Objectives of the Project

- Develop an AI-based fraud detection system capable of accurately identifying suspicious credit card transactions.
- Reduce false positives while maintaining high recall for fraudulent cases.
- Analyze transaction patterns to gain insights into how fraud is perpetrated.
- Evaluate and compare multiple ML models to find the most effective algorithm for fraud detection.

- Present findings through clear visualizations and potentially build a prototype dashboard for real-time detection.

3.Scope of the Project

- Use transaction datasets to train and test fraud detection algorithms.
- Apply preprocessing and handle imbalanced classes to improve model performance.
- Train, evaluate, and compare several ML models including tree-based classifiers and ensemble methods.
- Constraints:
  - Limited to publicly available anonymized datasets.
  - No real-time transaction access or integration with banking systems. • Deployment limited to simulation (dashboard or web app).

4.Data Sources

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

- **Dataset**: Credit Card Fraud Detection
- **Source**: Kaggle - Credit Card Fraud Detection Dataset
- **Type**: Public
- **Nature**: Static (downloaded once, not updated in real-time)
- **Description**: Contains anonymized credit card transactions over two days, with labeled legitimate and fraudulent transactions

5.High-Level Methodology

- **Data Collection**

Download the dataset from Kaggle.

- **Data Cleaning**

  Check and handle missing values, drop duplicates, and ensure uniform data formatting

- **Exploratory Data Analysis (EDA)**

  Bar charts, histograms, pair plots, and correlation matrices will be used to explore data relationships.

  Class distribution graphs will help identify imbalances between fraudulent and non-fraudulent transactions. These visualizations will guide feature selection and model improvement.

- **Feature Engineering**

  Data will be standardized and scaled to improve model performance, dimensionality reduction (e.g., PCA) will be applied to simplify features. New features will be created if they enhance predictive accuracy.

- **Model Building** –

Experiment with:

- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost
- LightGBM
- Possibly ANN (Artificial Neural Networks)

- **Model Evaluation** –

Use metrics such as:

- Confusion Matrix
- Precision, Recall, F1-Score
- ROC-AUC
- Cross-validation

- **Visualization & Interpretation**

Seaborn and Matplotlib will be used for creating visualizations. Graphs will highlight fraud trends, transaction patterns, and class distributions.

Feature importance and model performance will be visually compared. Confusion matrices will help evaluate prediction accuracy and error types.

- **Deployment** –

The project may include optional deployment as a Streamlit web app. It will simulate real-time fraud detection using user-input transactions or test cases.

This allows interactive testing and demonstration of the model's effectiveness.

6.Tools and Technologies

- **Programming Language**: Python
- **Notebook/IDE**: Google Colab or Jupyter Notebook
- **Libraries**:
  - Data Processing: pandas, numpy

- Visualization: matplotlib, seaborn, plotly
- Modeling: scikit-learn, xgboost, lightgbm, imbalanced-learn
- Deployment: Streamlit or Flask

7.Team Members and Roles

NAME	ROLE	RESPONSIBLE
Inbarasu I	Member	Data Collection, Data Cleaning
Deepak kumar S	Member	Visualization & Interpretation
Hemanth D	Member	Exploratory Data Analysis (EDA), Feature Engineering
Gokul S	Leader	Model Building, Model Evaluation