**Project Initialization and Planning Phase**

| Date | 28 January 2025 |
| --- | --- |
| Team ID | NIL |
| Project Title | Online Payments Fraud Detection Using ML |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project develops a machine learning system to detect online payment fraud more accurately than rule-based methods. Using Python and scikit-learn, we'll train three models (Logistic Regression, Decision Trees, and Naive Bayes) on a dataset of transactions. The solution will focus on key features like transaction amounts and balance changes while addressing class imbalance.

Targeting 99% accuracy with high recall to catch fraud and minimal false positives, the system will help financial institutions reduce fraud losses by an estimated 25%. The deliverables include trained models, performance analysis, and documentation, with architecture designed for potential future deployment as a real-time API.

Key Features:

* Processes real transaction data (amounts, types, balances)
* Compares multiple ML algorithms
* Optimizes for both accuracy and fraud detection rate
* Provides foundation for production deployment

| **Project Overview** | |
| --- | --- |
| Objective | Build a machine learning pipeline to detect online payment fraud with >99% accuracy. |
| Scope | Covers data preprocessing, model training, evaluation, and API deployment. |
| **Problem Statement** | |
| Description | Fraudulent transactions cause financial losses and erode trust; current systems are reactive. |
| Impact | Solving this will reduce losses by 30% and improve user experience. |
| **Proposed Solution** | |
| Approach | Supervised learning (Logistic Regression, Decision Tree, Naive Bayes) with feature engineering. |
| Key Features | Real-time scoring, interpretable models, and scalability for large datasets. |

#### Resource Requirements

| **Type** | **Details** |
| --- | --- |
| **Hardware** | Laptop/Cloud VM (4-core CPU, 16GB RAM, 100GB storage). |
| **Software** | Python, scikit-learn, pandas, matplotlib, Jupyter Notebook. |
| **Data** | Kaggle dataset: transactions, 10 features (e.g., amount, isFraud). |