

# Online Payment Fraud Detection

## Milestone 1: Project Initialization and Planning Phase

The Project Initialization and Planning Phase is the first stage of the project management lifecycle. It sets the foundation for a successful project by defining the project's objectives, scope, timelines, budget, and resource allocation.

The output of this phase is a comprehensive project plan that serves as a guide for the entire project team. It's an essential phase that lays the groundwork for a well-executed project.

### Activity 1: Define Problem Statement

---

#### Problem Statement:

Online payment fraud poses a significant risk to our financial transactions, threatening both our business integrity and customer trust. Despite current preventive measures, the evolving tactics of fraudsters necessitate a robust and adaptive fraud detection system.

The challenge lies in developing an efficient online payment fraud detection system that accurately identifies fraudulent transactions while minimizing false positives. This system should integrate seamlessly into our existing payment processing infrastructure, ensuring minimal disruption to legitimate transactions.

Ref. template: [Click Here](#)

Online Payment Fraud Detection Problem Statement Report: [Click Here](#)

### Activity 2: Project Proposal (Proposed Solution)

Online payment fraud continues to be a pressing issue, affecting businesses and customers alike. To mitigate these risks and enhance transaction security, we propose the development of a comprehensive Online Payment Fraud Detection System. This system will leverage advanced machine learning techniques and real-time transaction monitoring to detect and prevent fraudulent activities effectively.

Ref. template: [Click Here](#)

Online Payment Fraud Detection Project Proposal Report: [Click Here](#)

### **Activity 3: Initial Project Planning**

Initial project planning for online payment fraud detection involves several key steps: 1. Define Project Scope

2. Set Goals and Objectives
3. Conduct Stakeholder Analysis
4. Perform Initial Risk Assessment
5. Determine Project Timeline
6. Estimate Resource Requirements
7. Define Project Deliverables
8. Establish Communication Plan
9. Identify Potential Roadblocks
10. Develop Preliminary Budget

**Ref. template:** [Click Here](#)

**Online Payment Fraud Detection Project Planning Report:** [Click Here](#)

## **Milestone 2: Data Collection and Preprocessing Phase**

The Data Collection and Preprocessing Phase for online payment fraud detection involves gathering and preparing data for use in modeling and analysis. Steps to be followed are : Data Sources, Data Collection, Data Cleaning, Data Transformation, Data Reduction, Data Anonymization, Data Quality Check, Feature Engineering, Data Split.

The output of this phase is a preprocessed dataset, ready for modeling and analysis. This phase is crucial in fraud detection, as high-quality data is essential for accurate model performance

### **Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report**

- Data Sources:
- Payment Gateway API
- Website Logs
- Credit Card Company API

- Internal Database    - Customer Feedback Form - Data Collection Methods:
- Kaggle website
- Web scraping for website logs
- Database queries for internal database    - Manual data entry for customer feedback - Data Format:
- CSV format for web scraping and database queries    - Excel format for customer feedback

**Ref. template:** [Click Here](#)

**Online Payment Fraud Detection Data Collection Report:** [Click Here](#)

### Activity 2: Data Quality Report

This report evaluates the quality of the data used for online payment fraud detection. The data was collected from various sources, including payment gateways, website logs, credit card companies, and internal databases. The data quality report highlights areas for improvement in data completeness, accuracy, consistency, timeliness, and validity. By addressing these issues, we can enhance the effectiveness of our online payment fraud detection system.

**Ref. template:** [Click Here](#)

**Online Payment Fraud Detection Data Quality Report:** [Click Here](#)

### Activity 3: Data Exploration and Preprocessing

Data exploration and preprocessing are crucial steps in building an online payment fraud detection model. Here's an outline of the tasks involved:

1. Data Review: Examine the data collection process, data sources, and data formats.
2. Descriptive Statistics: Calculate summary statistics (mean, median, mode, count, etc.) for each feature.
3. Data Visualization: Use plots and charts to understand the distribution of data, relationships between features, and outliers.
4. Correlation Analysis: Calculate correlation coefficients to identify highly correlated features.

**Ref. template:** [Click Here](#)

**Online Payment Fraud Detection Data Exploration and Preprocessing Report:** [Click Here](#)

## Milestone 3: Model Development Phase

The Model Development Phase for online payment fraud detection involves building and training a machine learning model to predict fraudulent transactions. The Model Development Phase entails initiating training with code, and rigorously validating and assessing model performance for informed decisionmaking in the lending process.

### Activity 1: Feature Selection Report

A feature selection report for online payment fraud detection would typically identify the most relevant factors that contribute to predicting fraudulent transactions. These features can help machine learning models identify patterns and anomalies that may indicate online payment fraud. However, it's important to note that feature selection may vary depending on the specific dataset and problem you're trying to solve.

Ref. template: [Click Here](#)

Online Payment Fraud Detection Feature Selection Report: [Click Here](#)

### Activity 2: Model Selection Report

The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, SVM and XGB models for transaction prediction. This Model Selection Report provides a structured approach to evaluating and recommending machine learning models for online payment fraud detection, ensuring alignment with project objectives and operational requirements. Adjustments can be made based on specific organizational needs and data characteristics encountered during the evaluation process

Ref. template: [Click Here](#)

Online Payment Fraud Detection Model Selection Report: [Click Here](#)

### Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

Initial Model Training Code:

1. Import libraries
  2. Load and preprocess dataset
  3. Split data into training and testing sets
  4. Define model architecture
  5. Train model on training data
  6. Save trained model for later use
- Model Validation:
1. Split testing data into validation and evaluation sets

2. Use validation set to tune hyperparameters and optimize model performance
3. Evaluate model on validation set using metrics Evaluation Report:
  1. Model performance metrics
  2. Confusion matrix
  4. Precision-Recall curve

Ref. template: [Click Here](#)

Online Payment Fraud Detection Model Development Phase Template: [Click Here](#)

## Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

The goal of model optimization and tuning is to improve the model's performance, robustness, and generalization to new, unseen data. By carefully tuning the model, you can increase the accuracy of fraud detection and reduce the number of false positives and false negatives.

### Activity 1: Hyperparameter Tuning Documentation

The Decision Tree model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex frauds, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

### Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Decision Tree Model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

### **Activity 3: Final Model Selection Justification**

The Final Model Selection Justification articulates the rationale for choosing Decision Tree as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal transaction approval predictions.

Ref. template: [Click Here](#)

Online Payment Fraud Detection Model Optimization and Tuning Phase Report: [Click Here](#)

### **Milestone 5: Project Files Submission and Documentation**

For project file submission in Github, Kindly click the link and refer to the flow: [Click Here](#)

For the documentation, Kindly refer to the link: [Click Here](#)

### **Milestone 6: Project Demonstration**

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.