

Smart Lender – Online Payment Fraud Detection Using ML

Milestone 1: Project Initialization and Planning Phase

To initiate and plan the online payments fraud detection project using machine learning, we first defined the project scope and identified the specific problem to be solved. We determined the goals and objectives of the project, established key performance indicators to measure success, and conducted a stakeholder analysis to identify key stakeholders and their roles in the project. Next, we performed a literature review to research existing solutions and approaches to online payments fraud detection, and identified machine learning algorithms and techniques relevant to the problem. We then gathered requirements by collecting data on online payments fraud cases, identifying features and variables relevant to fraud detection, and determining data quality and availability.

Activity 1: Define Problem Statement

Problem Statement: Design and develop a machine learning model that can detect and prevent online payments fraud in real-time, with a focus on accuracy, scalability, and interpretability. The model should classify transactions as fraudulent or legitimate with a high degree of precision, minimizing false positives and false negatives. It must adapt to evolving fraud patterns and trends, providing insights into the underlying factors contributing to fraud.

Ref. template: [Click Here](#)

SmartLender Problem Statement Report: [click here](#)

Activity 2: Project Proposal (Proposed Solution)

The rapid growth of e-commerce and online transactions has led to an increase in online payment fraud, resulting in significant financial losses for individuals and businesses. This project aims to develop a machine learning-based online payment fraud detection system to identify and prevent fraudulent transactions in real-time. The system will utilize advanced algorithms and techniques, such as anomaly detection, predictive modeling, and data analytics, to analyze transaction patterns and identify potential fraud indicators

Ref. template: [Click Here](#)

SmartLender Project Proposal Report: [Click Here](#)

Activity 3: Initial Project Planning

To initiate the project on online payment fraud detection using ML, we will first conduct a thorough literature review to understand the current state-of-the-art techniques and algorithms used in fraud detection. Next, we will collect and preprocess a dataset of historical online payment transactions, labeling them as fraudulent or legitimate. We will then explore feature engineering techniques to extract relevant information from the data, such as transaction amount, location, and velocity.

Ref. template: [Click Here](#)

SmartLender Project Planning Report: [Click Here](#)

Milestone 2: Data Collection and Preprocessing Phase

The project on online payment fraud detection using Machine Learning will commence with a literature review to understand current techniques and algorithms used in fraud detection.

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

The dataset for "SmartLender – online payments fraud detection using ML" is sourced from Kaggle. It includes applicant details and financial metrics. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [Click Here](#)

SmartLender Data Collection Report: [Click Here](#)

Activity 2: Data Quality Report

The data quality report evaluates the quality of the data used to train and test machine learning models in online payments fraud detection. The report assesses the data across several dimensions, including completeness, accuracy, consistency, and timeliness.

Ref. template: [Click Here](#)

SmartLender Data Quality Report: [Click Here](#)

Activity 3: Data Exploration and Preprocessing

Data exploration and preprocessing are crucial to ensure the quality and integrity of the data. This includes handling missing values, normalizing and scaling numerical data, encoding categorical data, and transforming the data into a suitable format for ML algorithms

Ref. template: [Click Here](#)

SmartLender Data Exploration and Preprocessing Report: [Click here](#)

Milestone 3: Model Development Phase

The model development phase in online payment fraud detection using Machine Learning (ML) involves training and validating ML algorithms to detect fraudulent transactions. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, XGB), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

Activity 1: Feature Selection Report

The Feature Selection Report for online payments fraud detection using Machine Learning (ML) highlights the most relevant features used to train the ML models. These features were chosen based on their correlation with fraudulent transactions, ability to distinguish between legitimate and fraudulent transactions, and robustness to noise and missing values.

Ref. template: [Click Here](#)

SmartLender Feature Selection Report: [Click Here](#)

Activity 2: Model Selection Report

The Model Selection Report for online payment fraud detection using Machine Learning (ML) summarizes the performance of various ML algorithms in detecting fraudulent transactions.

These models were selected based on their high accuracy, precision, recall, and F1-score, as well as their robustness to class imbalance and noise in the data.

Ref. template: [Click Here](#)

SmartLender Model Selection Report: [Click Here](#)

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

The initial model training code involves implementing the selected machine learning algorithms, Random Forest (RF) and Gradient Boosting (GB), using the scikit-learn library in Python. The code includes data preprocessing, feature scaling, and model training functions..

Ref. template: [Click Here](#)

SmartLender 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.

Activity 1: Hyperparameter Tuning Documentation

Hyperparameter tuning is a crucial step in optimizing the performance of machine learning models in online payments fraud detection. The goal is to find the optimal combination of hyperparameters that result in the best performance metrics. The hyperparameter tuning process involves using techniques such as Grid Search, Random Search, and Bayesian Optimization to iterate through a range of hyperparameter values.

Activity 2: Performance Metrics Comparison Report

This report compares the performance metrics of different machine learning models in online payments fraud detection. The models evaluated include Random Forest (RF), Gradient Boosting (GB), and Neural Network (NN). The performance metrics used for comparison are accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Gradient Boosting as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring online payments fraud detection using ML.

Ref. template: [Click Here](#)

SmartLender 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.