

## Customer shopping segmentation using machine learning

### Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

#### Activity 1: Define Problem Statement

**Problem Statement:** A retail company aims to segment its customer based on their shopping behaviors to improve personalized marketing efforts, enhance product recommendations, and boost customer satisfaction. The challenge is to identify distinct customer groups using purchase history, spending patterns, and products preferences. By understanding these segments, the company can tailor its strategies to meet the unique needs and preferences of each group, ultimately increasing customer loyalty and sales.

Ref. template: [Click Here](#)

SmartLender Problem Statement Report: [Click Here](#)

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

The proposed solution involves leveraging machine learning to segment customers based on demographics, purchase history, and behavioral data. This segmentation will enable targeted marketing, personalized experiences, and optimized inventory management. By analyzing these data points, the model can identify distinct customer groups, leading to enhanced customer satisfaction and increased sales. This initiative aligns with the goal of improving operational efficiency and strategic decision-making.

Ref. template: [Click Here](#)

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

#### Activity 3: Initial Project Planning

Initial Project Planning involves outlining key objectives, defining scope, and identifying stakeholders for a customer shopping system. It encompasses setting timelines, allocating resources, and determining the overall project strategy. During this phase, the team establishes a clear understanding of the dataset, formulates goals for analysis, and plans the workflow for data processing. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.

Ref. template: [Click Here](#)

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

### Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant customers

shopping data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

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

The dataset for "customer shopping segmentation using machine learning" is sourced from Kaggle. It includes customer details and shopping 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 dataset for "customer shopping segmentation using machine learning" is sourced from Kaggle. It includes customers details and shopping 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 Quality Report: [Click Here](#)

### **Activity 3: Data Exploration and Preprocessing**

Data Exploration involves analyzing the customer shopping dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses in the customer shopping project.

Ref. template: [Click Here](#)

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

## **Milestone 3: Model Development Phase**

The Model Development Phase entails crafting a predictive model for customer shopping. It encompasses strategic feature selection, evaluating and selecting models (kmeans, 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 outlines the rationale behind choosing specific features (e.g., Gender, category, shoppingmalls, payment method, age) for the clustering model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to discern distinct customer shopping segments effectively.

Ref. template: [Click Here](#)

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

### **Activity 2: Model Selection Report**

The Model Selection Report details the rationale behind choosing kmeans, Decision Tree, KNN, and XGB models for customer shopping prediction. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

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 employs selected algorithms on the customer shopping dataset, setting the foundation for predictive modeling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, shopping metrics like accuracy and precision to ensure reliability and effectiveness in predicting cluster outcomes.

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**

The kmeans and knn model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, 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 kmeans and knn 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 kmeans and knn as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal customer shopping predictions.

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.