# Customer shopping segmengtation 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
Develop a machine learning model to segment customers based on their shopping behavior. The objective is to identify distinct customer groups to enable personalized marketing strategies and improve customer satisfaction. The model should analyze various factors, such as purchase history, frequency, and spending patterns, to provide actionable insights for targeted promotions and optimized inventory management"
**Ref. template: [Click Here](#)**
**SmartLender Problem Statement Report: [Click Here](#)**

### Activity 2: Project Proposal (Proposed Solution)
The proposed project of customer shopping segmentation is To segment customers based on shopping behavior for personalized marketing and improved satisfaction.

Collect and preprocess customer data, then use clustering algorithms to identify distinct segments. Evaluate model performance and integrate it into business processes. Visualize segments with dashboards.

Enable targeted marketing, enhance customer satisfaction, and optimize inventory management.

**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 loan approval 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 loan

application 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

Collecting high-quality data is crucial for effective customer shopping segmentation using machine learning. The data collection process begins by gathering diverse sources of customer information, such as transaction records, website behavior, social media interactions, and customer feedback. Ensuring the data is comprehensive and up-to-date is essential, as outdated or incomplete data can lead to inaccurate segmentation. Data quality is maintained through rigorous cleaning processes that involve removing duplicates, correcting errors, and standardizing **formats.Ref. template: [Click Here](#)**
**SmartLender Data Collection Report: [Click Here](#)**

### Activity 2: Data Quality Report

Data Quality RA data quality report for customer shopping segmentation using machine learning typically covers several key aspects to ensure reliable results and insights. The report begins with an overview of data sources, detailing where the data comes from and its completeness. It assesses data accuracy by examining outliers and inconsistencies, ensuring that all data points are valid and correctly formatted

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

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

### Activity 3: Data Exploration and Preprocessing

Exploring and preprocessing data for customer shopping segmentation using machine learning involves several critical steps. Initially, gather relevant datasets from various sources such as transaction records, customer demographics, and behavioral data. Conduct exploratory data analysis (EDA) to gain insights into the data's distribution, correlations, and potential outliers. Visualizations such as histograms, scatter plots, and heatmaps can help uncover patterns and relationships among variables.
**Ref. template: [Click Here](#)**

# Milestone 3: Model Development Phase

During the model development phase for customer shopping segmentation using machine learning, several key steps are crucial. First, select appropriate algorithms suited to the nature of your data and segmentation goals. Common choices include clustering algorithms like K-means, hierarchical clustering, or density-based clustering methods such as DBSCAN.

**Activity 1: Feature Selection Report**

Feature selection is crucial for effectively segmenting customers based on shopping behavior using machine learning. In this context, selecting the right features involves identifying variables that most significantly influence customer segmentation outcomes. Techniques such as correlation analysis, feature importance from tree-based models like Random Forests or Gradient Boosting Machines, and domain knowledge can guide feature selectionm

**Ref. template:** [Click Here](#)
**SmartLender Feature Selection Report:** [Click Here](#)

## Activity 2: Model Selection Report

The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, KNN, and XGB models for customer shopping segmentation. 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

To begin customer shopping segmentation using machine learning, we first import essential libraries such as pandas for data manipulation and sklearn for model implementation. The dataset, typically stored as a CSV file named 'customer_data.csv', is loaded and relevant features essential for segmentation, such as purchase history, demographics, and browsing behavior, are selected. Data preprocessing steps, including scaling using StandardScaler to normalize feature values, ensure optimal model performance**.**

**Ref. template:** [Click Here](#)
**SmartLender Model Development Phase Template:** [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. Kindly refer to the link. Click Here