Customer Lifetime Value Prediction Model

Introduction

In today's competitive market, understanding and forecasting a customer's Lifetime Value (LTV) is crucial for optimizing marketing strategies and retaining high-value clients. This project aims to build a machine learning model to predict LTV using customer transactional behavior, enabling better customer segmentation and targeted marketing.

Abstract

The project involved analyzing a digital wallet transaction dataset to predict the LTV of customers. Feature engineering was performed to extract behavioral metrics like frequency, recency, and average order value (AOV). An XGBoost regression model was trained to predict customer LTV. The model's predictions were saved and segmented into customer tiers for strategic business applications.

Tools Used

- Language: Python

- Libraries: Pandas, NumPy, Seaborn, Matplotlib, Scikit-learn, XGBoost

- Output: CSV file of predicted LTV values

Steps Involved in Building the Project

1. Data Loading and Cleaning

   - Imported digital_wallet_ltv_dataset.csv

   - Checked for and handled null values

   - Converted invoice_date to datetime and ensured consistency in formats

2. Feature Engineering

   - Recency: Days since last transaction

   - Frequency: Number of transactions per user

   - Monetary (AOV): Average order value

   - Calculated LTV as the sum of purchase values for each user

3. Data Preprocessing

Customer Lifetime Value Prediction Model

- Normalized features

- Split dataset into training and testing sets (80-20 split)

4. Model Building

- Used XGBoost Regressor

- Evaluated model using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)

5. Customer Segmentation

- Based on predicted LTV values, customers were segmented into:

  - High Value

  - Medium Value

  - Low Value

- Saved output to ltv_predictions.csv

Conclusion

The LTV prediction model effectively segments customers based on their projected long-term value. Such segmentation allows marketing teams to allocate resources efficiently, improve retention strategies, and personalize campaigns. This project provides a scalable framework for LTV forecasting in digital commerce or fintech environments.