Customer Lifetime Value (LTV) Prediction Report

Introduction:

Customer Lifetime Value (LTV) is a crucial business metric used to estimate the total worth of a customer over the course of their relationship with a company. This project aims to develop a regression-based model to predict LTV using synthetic transactional data.

Abstract:

Using simulated data for 500 customers, we extracted features such as Recency (days since last purchase), Frequency (total purchases), and AOV (Average Order Value). These features were used to train a machine learning model that estimates LTV, helping businesses segment customers and prioritize marketing strategies.

Tools Used:

Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), Jupyter Notebook

Steps Involved:

1. Data generation using synthetic methods for Recency, Frequency, and AOV.
2. Feature engineering to calculate predicted LTV and customer segments.
3. Trained a Random Forest Regressor on 80% of the dataset.
4. Evaluated model using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
5. Exported predictions and segmented customers into High, Medium, Low, and Very Low.

Model Evaluation:

MAE (Mean Absolute Error): 192.13

RMSE (Root Mean Squared Error): 320.55

Conclusion:

This model provides a reliable prediction of LTV based on behavioral indicators. It enables marketing and business teams to target high-value customers effectively, minimize churn, and increase ROI. Future enhancements could include incorporating real-world transactional data, time-based cohort analysis, and advanced feature selection.