Customer Lifetime Value Prediction

Using Python, RFM Features, and XGBoost

Objective & Tools

Goal: Predict Customer Lifetime Value (LTV) to improve marketing strategies and target the right segments.

Tools Used:

- Python (Pandas, Scikit-learn, XGBoost)
- Excel (for raw data viewing)
- Matplotlib & Seaborn (for visualization)

Data & Feature Engineering

Dataset: Insurance company customer records.

Features Used:

- Recency: Days since last purchase

- Frequency: Total number of purchases

- AOV (Average Order Value): Mean total claim amount

Method: Group by Customer ID to calculate RFM values.

Model Training & Evaluation

Model: XGBoost Regressor

Target Variable: Customer Lifetime Value

Metrics:

- MAE: 4409.72 (avg prediction error)

- Visual Evaluation: Scatterplot of predicted vs actual values

Split: 80% train, 20% test

Insights & Segmentation

Segmentation:

- Low, Medium, High, VIP (quartiles based on predicted LTV)

Use Case:

- Target VIP customers with high-value campaigns
- Retain Medium/High segments with promotions
- Low segment: upselling or automated retention efforts