

Customer Lifetime Value (LTV)

Prediction: *Project Report*

Y. BHAVANI SANKAR

Introduction :

Customer Lifetime Value (LTV) is a critical metric for businesses, especially in the e-commerce domain. It helps determine the total revenue a business can expect from a customer during their entire relationship. By accurately predicting LTV, businesses can optimize marketing strategies, personalize customer experiences, and increase retention rates. This project focuses on building a machine learning model to predict Customer LTV using historical transaction data from an e-commerce platform.

Abstract :

This project aims to predict the Customer Lifetime Value (LTV) for e-commerce transactions by applying machine learning techniques. Using historical transaction data, we calculate key features like Recency, Frequency, and Monetary values (RFM analysis). We train a Random Forest Regressor model to predict the total amount spent by each customer over their lifetime. The model's performance is evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The predicted LTV values can provide actionable insights for marketing, customer retention, and segmentation strategies.

Tools Used :

- **Python:** For data processing, model building, and evaluation.
- **Pandas:** For data manipulation and handling missing values.
- **NumPy:** For numerical operations.
- **Matplotlib & Seaborn:** For visualizations like feature importance and model evaluation.
- **Scikit-learn:** For machine learning algorithms and metrics (Random Forest Regressor, MAE, RMSE).
- **Google Colab:** For code execution and collaborative working.

Steps Involved in Building the Project :

Data Preprocessing:

- Loaded the dataset containing e-commerce transaction data.
- Cleaned the data by removing missing CustomerIDs and converting the InvoiceDate to a datetime format.
- Created a new column, **TotalPrice**, by multiplying Quantity and UnitPrice.

Feature Engineering:

- Performed RFM (Recency, Frequency, and Monetary) analysis to create meaningful features:
 - **Recency**: The number of days since the last purchase.
 - **Frequency**: The total number of transactions per customer.
 - **Monetary**: The total spend of each customer.

Data Splitting:

- Split the dataset into training and testing sets using an 80-20 ratio.

Model Training:

- Trained a Random Forest Regressor model to predict the monetary value (Customer LTV).
- Evaluated the model using MAE and RMSE metrics.

Results & Evaluation:

- Visualized feature importance to understand the contribution of Recency and Frequency in predicting LTV.
- Predicted LTV values and saved them into a CSV file for further analysis.

Conclusion

The Customer Lifetime Value prediction model built using Random Forest Regressor effectively predicts the total spend of each customer. The model provides valuable insights that can be leveraged for marketing strategies, customer retention programs, and targeted campaigns. The evaluation metrics, MAE and RMSE, indicate the model's reliable performance in predicting LTV. Further improvements can be made by exploring advanced algorithms, tuning hyperparameters, and adding more features like customer segmentation or seasonal trends. Ultimately, this project demonstrates how predictive analytics can enhance business decision-making in e-commerce.