Proposal

Name:Haya Saud Alowias
Email:haya.alowias@gmail.com

Background:

For the dataset used in this paper, I have chosen an open-source online retail dataset that contains all the transactions occurring between 01/12/2010 and 09/12/2011 for many countries that registered non-store online retail. the companies mainly sell unique all-occasion giftware.

Motivation:

From this dataset, I am so interested to find the customer lifetime value (CLV) this is important for companies that have customers. It is one of the key stats to track as part of a customer experience program. because helping to determine customer segmentation, measurement of customer loyalty, aiding in the judgment of product quality, Increasing profitability overall.

Problem Statement:

The purpose of the model is to use regression models to predict the CLTV model and analysis the data that answer the following questions :

- What is the top 5 country Transaction?
- What is the number of transactions for all countries?
- What is the revenue by country?
- Average Price by Country?
- Transactions by Month and Year?
- Predicted Purchases for one month for each customer?
- which product has been purchased more?
- What is CLV for each customer?
- Predict CLTV using Linear Regression Model

Data Description:

- A real online retail transaction dataset for 12 months and 9 days . I got the data from Kaggle.
- Number of Features: 8
- Number of Instance: 541909

InvoiceNo	Nominal. A 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.
StockCode	Nominal. Product (item) code .A 5-digit integral number uniquely assigned to each distinct product.
Description	Nominal.Product (item) name.
Quantity	Numeric. The quantities of each product (item) per transaction
InvoiceDate	Numeric. Invice date and time. The day and time when a transaction was generated.
UnitPrice	Numeric. Product price per unit in sterling (Â£).
CustomerID	Nominal. Customer number. A 5-digit integral number uniquely assigned to each customer.
Country	Nominal. Country name. The name of the country where a customer resides

Tools:

- XSL
- Pandas
- Numpy
- Matplotlib
- Lifetimes
- Seaborn
- Datetime
- LinearRegression