BUSINESS ANALYST PROJECT 2

CUSTOMER LIFETIME VALUE

Instructor: Mr. Pham Tuan Minh
Group 4:
Hoang Thi Tu Anh - 11219255
Nguyen Kim Chi - 11211095
Tran Thi Ngan Ha - 11211959
Vo Thi Yen Nhi - 11219282
Kieu Thanh Tam - 11219286

Ha Noi, 2024

1. INTRODUCTION

a. Problem definition

Customer lifetime value (CLV or CLTV) is a metric that represents the total net profit a company can expect to generate from a customer throughout their entire relationship. It takes into account the customer's initial purchase, repeat purchases, and the average duration of their relationship with the company.

Customer lifetime value helps you understand and gauge current customer loyalty. If customers continue to purchase from you time and time again, that's usually a good sign you're doing the right things in your business. Furthermore, the larger a customer lifetime value, the less you need to spend on your customer acquisition costs.

b. Motivation

Understanding what your customer lifetime value is on a broader perspective is important for calculating your costs and return on investment in general. However, on a more granular level, the importance of customer lifetime value becomes even more apparent. Once you calculate CLV on a more granular and individual basis, you can use it for optimizing customer journeys, customer segmentation, return on investment calculation and more.

Measuring Marketing Efforts

Calculating customer lifetime value and changes within it allows marketers to gauge the impact of their marketing campaigns and other marketing efforts. Simply comparing prospective customer lifetime values - on an aggregate basis - from before and after the campaign will give insights into the efficacy of the campaign and the impact it had on customer lifetime value.

Not only is this a big help in determining what works and what doesn't, but it also helps with planning. Provided you have enough customer data, you can calculate predictive customer lifetime value, which will allow you to gauge whether or not you should pursue a certain campaign.

Understanding Customer Behavior

Using your customer lifetime value data can be a significant help in understanding customer behavior. Comparing the CLV of different customers at different times can expose insights into customer behavior. Not only this, but it can also help to uncover reasons for certain increases or decreases in revenue and sales within a given period.

This relates to measuring marketing efforts, as these can also be a reason for revenue fluctuations - profitable ones, we hope.

Understanding Marketing Objectives

For many companies, measuring and understanding customer lifetime value will bring a new perspective to marketing objectives and strategy. The importance of customer lifetime value is exposed within decisions on when to invest in certain customers, when to drop investments in other customers, and also follow how successful respective campaigns and efforts have been in increasing CLV.

Segmentation

The importance of customer lifetime value in segmentation is quite significant. Predictive customer lifetime value calculation is essential when it comes to customer acquisition campaigns, while either historic or predictive should be employed when it comes to deciding the investment in campaigns targeting existing customers.

ROI and Budget Allocation

This is the greatest benefit of knowing your CLV, but seems to be largely overlooked. Spending some time on understanding your CLV on an organization-wide basis, as well as on a segmental basis, will drastically improve your return on investment, which leads to more profit that can then be reinvested in CRM efforts and better customer service to maximize lifetime value across your customer base.

c. Problem Setting: Dataset/Goal

The first dataset we use is ***Online Retail II*** data set which contains all the transactions occurring for a UK-based and registered, non-store online retail between 01/12/2009 and 09/12/2011.The company mainly sells unique all-occasion gift-ware. Many customers of the company are wholesalers.

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

The second dataset is the ***CD NOW*** dataset. which contains the entire purchase history up to the end of June 1998 of the cohort of 23,570 individuals who made their first-ever purchase at CDNOW in the first quarter of 1997.

The main goal of this project is to develop a robust Customer Lifetime Value (CLV) system capable of estimating the value of customers throughout their interaction with the business. The system will be designed to analyze customer data from various sources, such as transaction history, behavioral information, and other relevant factors.

To achieve this goal, the system will utilize advanced methods and techniques in data analytics and machine learning. It will predict the future value of customers based on factors such as purchase frequency, average order value, and the likelihood of continued engagement with the business.

An important aspect of system development is ensuring the accuracy and reliability of CLV estimates. It should be able to handle variations in customer behavior and interactions, such as changes in consumer patterns, marketing campaigns, and other factors that may impact customer value. Additionally, the system will be optimized to provide CLV insights efficiently and rapidly, enabling businesses to make strategic decisions regarding marketing, customer care, and business development based on long-term customer value.

2. LITERATURE REVIEW

a. Research on Customer Life Cycle Value

- At present, there are two perspectives to define customer value.
    - ❖ Based on the customer perspective: It is calculated as the difference between total customer value and total customer cost. This perspective emphasizes how customers perceive the value gained from products or services compared to the cost incurred.
    - ❖ Based on the enterprise perspective: Customer value is the customer lifetime value (CLV) within the customer life cycle. Customer lifetime value refers to the present value of all profits created by customers for the enterprise in the whole process of maintaining a relationship with the enterprise.

→ The mainstream view currently focuses on studying customer value from the perspective of enterprises.

- Depending on the relationship between the sellers and the buyers, a business can either be a contractual business or a non-contractual business:
    - ❖ Contractual relationship, customers sign long-term contracts with a single business. If the contract is not renewed or transactions cease, the customer is considered lost permanently. Resuming transactions is viewed as acquiring a new customer. Examples include insurance and lending.
    - ❖ Non-contractual relationships, customers engage with multiple businesses concurrently with minimal switching costs. Pausing transactions doesn't necessarily mean customer loss; customers may return, and businesses still consider them as existing customers. An example is the retail industry.

b. Customer Life Cycle Model

- The Customer Life Cycle Value (CLV) model faces challenges in non-contractual relationships where customer churn is less visible. In such cases, CLV estimation relies on predicting future purchase probabilities or frequencies using recent transaction data and the RFM (Recency, Frequency, Monetary) model introduced by Arthur Hughes in the 1990s.

- However, scholars have identified shortcomings in CLV models:
    - ❖ They are often overly conceptual and idealized, making practical application difficult.
    - ❖ CLV typically focuses only on customer profitability, neglecting the broader value customers bring beyond revenue.
    - ❖ In the information age, the abundance of product options leads to intense competition among businesses, making it challenging to retain customers and gather comprehensive data for CLV calculation.

c. Probabilistic Model

- Probability models like the Pareto/NBD model and its variations are widely used to predict Customer Lifetime Value (CLV) by analyzing past purchase behavior. These models estimate the probability of customer retention and future purchase frequency based on historical data.
- Advantages:
    - ❖ Utilize historical data to estimate customer behavior.
    - ❖ Provide insights into customer retention and purchase frequency.
    - ❖ Adaptations enhance predictive accuracy by considering various factors.
- Disadvantages:
    - ❖ Reliance on assumptions that may not always hold true.
    - ❖ Limited use of purchase information in some models.
    - ❖ Complexity may require expertise in statistical analysis.

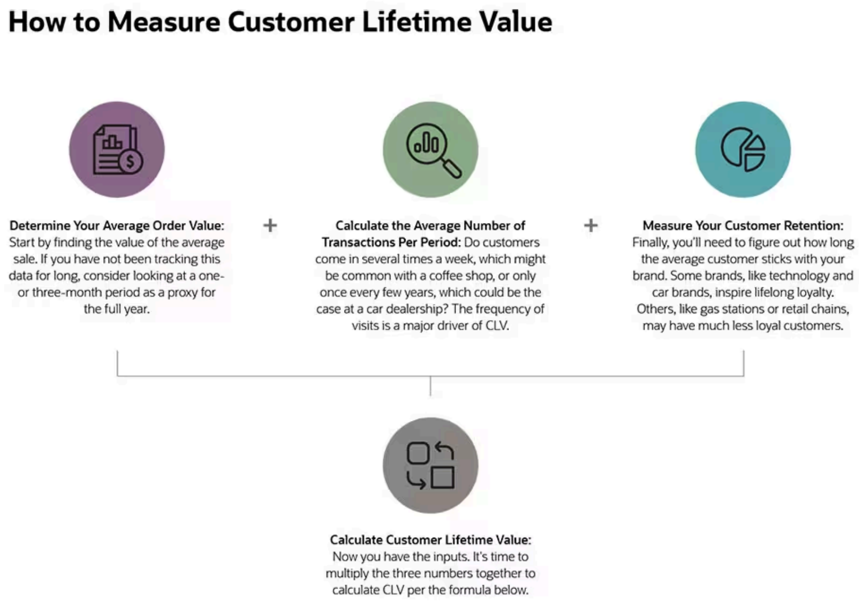
d. Machine learning model

- Machine learning techniques offer powerful predictive capabilities for CLV modeling. Unlike probability models, which rely on assumptions about underlying relationships, machine learning algorithms make no such assumptions, offering flexibility in modeling complex data. However, these techniques often require extensive parameter tuning and may suffer from overfitting, especially with large datasets. Despite these challenges, machine learning can effectively model customer acquisition, retention, and margin.
- Advantages:
    - ❖ Flexibility in modeling complex relationships in data.
    - ❖ Potential for high predictive performance.
    - ❖ Can incorporate various data types and features.
- Disadvantages:
    - ❖ Requires parameter tuning and careful model selection.
    - ❖ Prone to overfitting, especially with large datasets.
    - ❖ Computational complexity may be high, leading to longer training times.

3. SOLUTION

a. Traditional CLV

- **How to Measure Customer Lifetime Value**



Determine Average Order Value (AOV):
Begin by assessing the average value of each sale. Analyze data from a representative timeframe, such as a one- or three-month period, if yearly data is unavailable. AOV is calculated as the total revenue divided by the total number of orders.

Calculate Average Purchase Frequency (PF):
Evaluate the frequency at which customers engage with your business. This can vary widely depending on your industry, ranging from frequent visits (e.g., coffee shops) to more sporadic interactions (e.g., car dealerships). PF is derived from the total number of orders divided by the total number of customers.

Measure Customer Retention:
Assess the longevity of customer relationships with your brand. While some industries foster enduring loyalty, others may experience higher customer turnover. Determine the average duration of customer engagement to estimate churn rate, which is the percentage of customers who have not made repeat purchases. Churn rate can be calculated as 1 minus the repeat rate, where repeat rate is the proportion of customers with multiple orders compared to unique customers.

Calculate Customer Lifetime Value (CLV):
With the necessary inputs gathered, employ the CLV formula:
CLTV = ((Average Order Value x Purchase Frequency) / Churn Rate) x Profit Margin.
By multiplying AOV and PF and dividing by churn rate, you obtain the gross CLV. This figure can then be adjusted by factoring in the profit margin to yield the final CLV.

By systematically analyzing these metrics, you can gain actionable insights into customer behavior and optimize strategies to enhance CLV, thereby bolstering long-term profitability and sustainable growth

Average Order Value(AOV): The Average Order value is the ratio of your total revenue and the total number of orders. AOV represents the mean amount of revenue that the customer spends on an order.

Average Order Value = Total Revenue / Total Number of Orders

Purchase Frequency(PF): Purchase Frequency is the ratio of the total number of orders and the total number of customers. It represents the average number of orders placed by each customer.

Purchase Frequency = Total Number of Orders / Total Number of Customers

Churn Rate: Churn Rate is the percentage of customers who have not ordered again.
Customer Lifetime: Customer Lifetime is the period of time that the customer has been continuously ordering.

Customer Lifetime=1 / Churn Rate

Repeat Rate: Repeat rate can be defined as the ratio of the number of customers with more than one order to the number of unique customers. Example: If you have 10 customers in a month out of whom 4 come back, your repeat rate is 40%.

Churn Rate= 1 - Repeat Rate

- **Formula of CLV**

CLV for a customer (omitting customer subscript) is (Gupta, Lehmann, and Stuart 2004; Reinartz and Kumar 2003)

$$CLV = \sum_{t=0}^{T} \frac{(p_t - c_t)r_t}{(1+i)^t} - AC$$

where

- p_t= price paid by a consumer at time t,
- c_t = direct cost of servicing the customer at time t,
- i = discount rate or cost of capital for the firm,
- r_t = probability of customer repeat buying or being “alive” at time t,
- AC = acquisition cost, and
- T = time horizon for estimating CLV

Despite this simple formulation, researchers have used different variations in modeling and estimating CLV. Some researchers have used an arbitrary time horizon or expected customer lifetime (Reinartz and Kumar 2000; Thomas 2001),4 whereas others have used an infinite time horizon (e.g., Fader, Hardie, and Lee 2005; Gupta, Lehmann, and Stuart 2004). Gupta and Lehmann (2005) showed that using an expected customer lifetime generally overestimates CLV, sometimes quite substantially. Gupta and Lehmann (2003, 2005) also showed that if margins (p – c) and retention rates are constant over time and we use an infinite time horizon, then CLV simplifies to the following expression:

$$CLV = \sum_{t=0}^{\infty} \frac{(p-c)r^t}{(1+i)^t} = m \frac{r}{(1+i-r)}.$$

In other words, CLV simply becomes margin (m) times a margin multiple (r/1 + i – r). When retention rate is 90% and discount rate is 12%, the margin multiple is about four.6 Gupta and Lehmann (2005) showed that when margins grow at a constant rate "g," the margin multiple becomes r/[1 + i – r(1 + g)].

b. Probabilistic Models

Assumption of models

Customer Lifetime Value (CLV) estimation plays a crucial role in understanding the long-term value of customers for businesses. In this report, we explore the application of probabilistic models in CLV estimation, with a specific focus on the BG/NBD and Gamma-Gamma models. These models leverage probabilistic techniques to estimate the value of customers throughout their interaction with the business.

- **The BG/NBD model** combines the Beta Geometric and Negative Binomial distributions. The BG/NBD model is used to describe repeat purchase behavior in the context of non-contractual customer relationships. That is, users can buy products at any time without time constraints. The model can use historical user transaction data to predict the transaction times and turnover rate of each user in the future.

  The BG/NBD model is **based on the following assumptions**.
  (i). When the customer is active, the number of transactions follows a Poisson distribution with an action rate λ (while the customers are active).

  $$f(t_j \mid t_{j-1}; \lambda) = \lambda e^{-\lambda(t_j - t_{j-1})}, \qquad t_j > t_{j-1} \geq 0.$$

  (ii) The probability density function of nonuniformity of customer transaction rate λ obeys gamma distribution. The formula is:

$$f(\lambda \mid r, \alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)}, \qquad \lambda > 0.$$

(iii) After each transaction, the probability that the customer becomes silent is p, and the customer churn point follows the binomial distribution..

$$P(\text{inactive immediately after } j\text{th transaction}) = p(1-p)^{j-1}, \qquad j = 1, 2, 3, \ldots .$$

(iv) The nonuniformity probability density function of probability p follows a beta distribution. The formula is

$$f(p \mid a, b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a, b)}, \qquad 0 \leq p \leq 1,$$

where B(a,b) is a beta function, which can be expressed by the gamma function. The formula is as:

$$B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$$

- To estimate the Average Order Value (AOV), we employ **the Gamma-Gamma model**, which focuses on modeling the variation in purchase amounts. By accounting for the heterogeneity in customer spending patterns, this model provides a more accurate estimation of AOV.

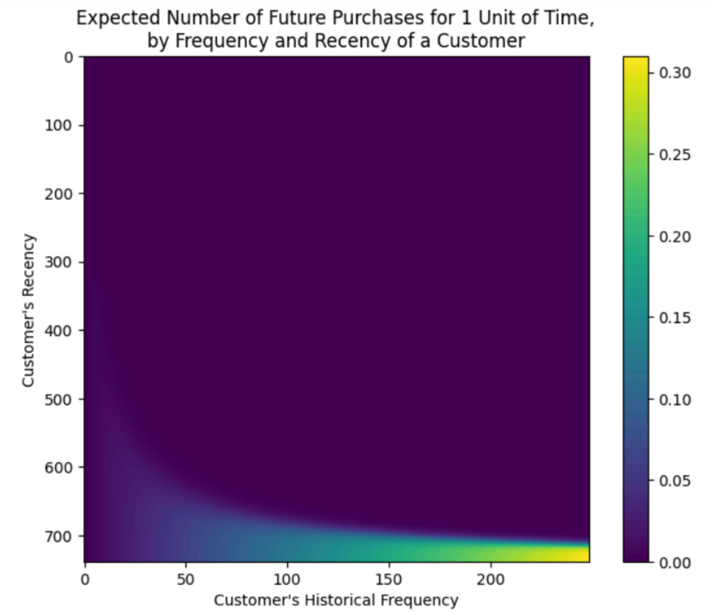
The application of these probabilistic models in CLV estimation enables businesses to make informed decisions regarding marketing strategies, customer segmentation, and enhancing customer interactions. The insights derived from these models help optimize marketing campaigns, identify high-value customers, and foster long-term customer relationships.

Analysis of the Online Retail II dataset using the BG/NBD model

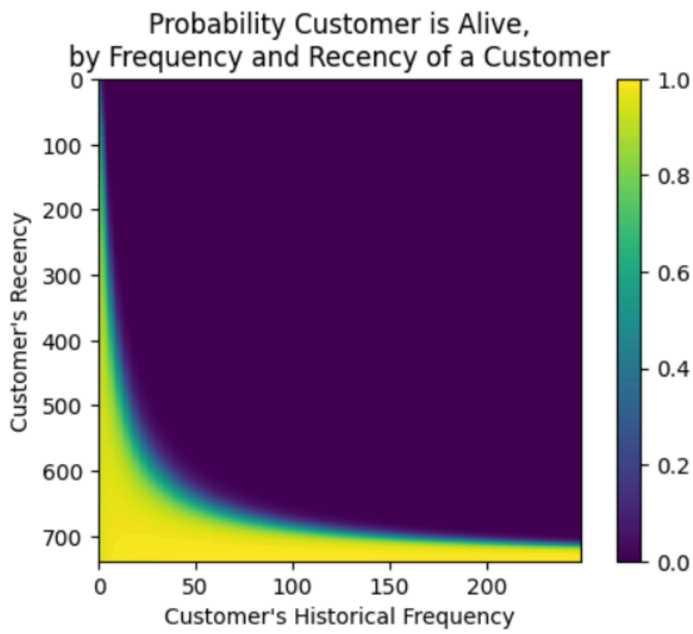
We can use the BG/NBD model to analyze customers' purchasing behavior and lifetime. The dataset provides the purchasing data of the customers; the BG/NBD model uses the frequency (the number of repeat purchases), T (the age of the customer in time unit), recency (the most recent purchase), and monetary (the average value of purchases) from this dataset.

	frequency	recency	T	monetary_value
Customer ID				
12346.0	2.0	322.0	647.0	38662.955000
12347.0	7.0	402.0	404.0	615.714286
12348.0	4.0	363.0	438.0	359.310000
12349.0	2.0	571.0	589.0	1305.085000
12350.0	0.0	0.0	310.0	0.000000

Using the lifetime library in python, we can predict the customers' buying behavior.

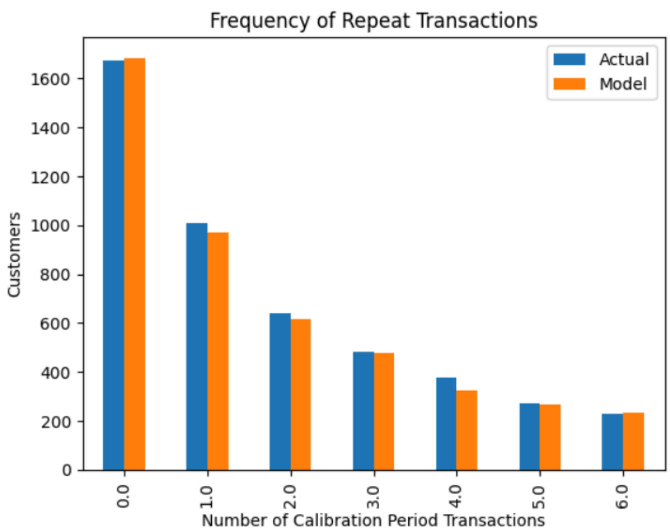


The above matrix shows that the expected number of future purchases starts low (top-left), increases around (100, 600) days, and rises towards the bottom-right. The top-right section doesn't reflect a high number of purchases and instead shows that those customers who bought more frequently at the beginning didn't return. The customers who made 200 purchases when they were 700 days old are the loyal customers the company should focus on.



We can see from the above matrix that customers with a high probability of being alive are in the bottom section of the graph. It's easy to conclude that the more recent customers make a purchase repeatedly, the higher the probability that they are active and loyal.

The following graph reveals if the model is accurate compared to the actual behavior of customers. We can notice that the value of the actual frequency of repeat transactions closely matches the predicted frequency of repeat transactions.



In addition, we can use customers' past transaction history to predict whether the customers are still alive or not.

c. Machine Learning Model

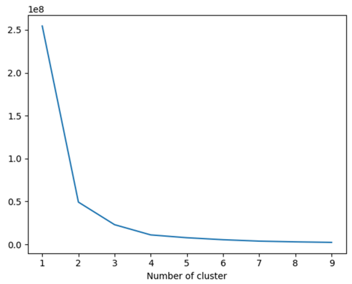
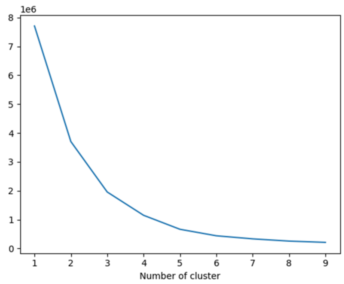
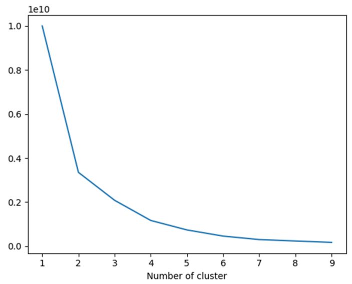
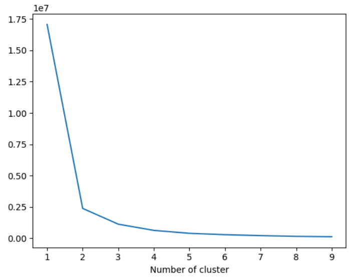
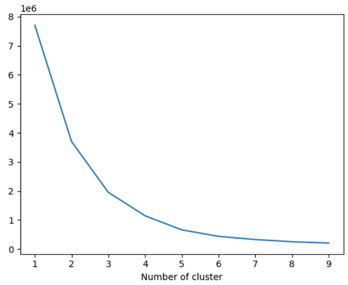
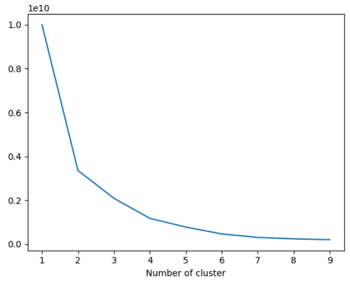
01. Customer classification based on the RFM model

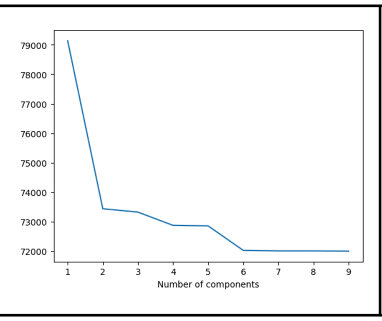
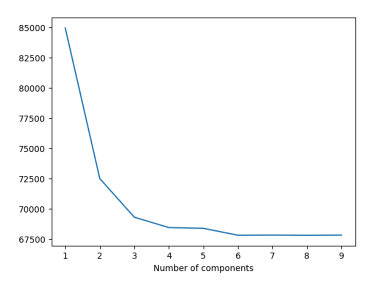
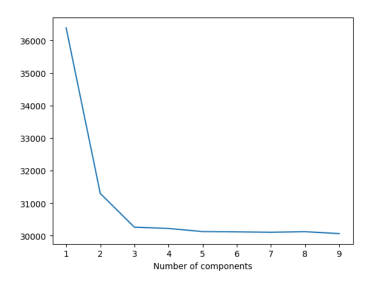
Customer segmentation using RFM (Recency, Frequency, Monetary) analysis is a powerful technique used by businesses to categorize customers into different groups based on their purchasing behavior. RFM analysis allows businesses to identify their most valuable customers, understand their preferences, and tailor marketing strategies accordingly.

RFM stands for Recency - Frequency - Monetary Value

- Recency: To calculate recency, we need to find out the most recent purchase date of each customer and see how many days they are inactive.
- Frequency: Frequency represents the total number orders for each customer
- Monetary (or Revenue): Monetary value represents the total amount of money a customer has spent over a given period

In order to partition each score into clusters based on similarity of characteristics, we use various unsupervised algorithms such as GMM, k-means, k-means++, and hierarchical clustering. Using the elbow method to specify the effective number of clusters, we decided to evaluation based on a 3-point scale for each criterion

	Recency	Frequency	Revenue
KMeans			
KMeans ++			

GMM			

After receiving each recency, frequency, revenue score, we combine them into a single RFM score. Theoretically, we will have 3 segments:

- Low value (having RFM score range from 0 to 2): customers who are less active than others, not very frequent buyer, and generates very low revenue
- Mid value (having RFM score range from 3 to 4):: in the middle of everything. Often uses the platform, fairly frequent and generates moderate revenue
- High value (having RFM score as 5):: the group you don't want to lose. High revenue and frequency, very active

Once customers are segmented into different clusters, we will use the first 3 months of these labeled data and predict customer segments for the next 6 months. The classification Machine Learning we choose are XGBClassifier, K-Nearest Neighbors, Logistic Regression, Random Forest and Gradient Boosting

02. Regression Model

One of other methods to predict customer lifetime value is the Regression Model. Our approach involves generating new features based on CLV customer behavior and incorporating demographic information to enhance prediction accuracy.

Feature engineering plays a pivotal role in CLV prediction. In our methodology, we generate new features based on CLV customer behavior. Specifically, we calculate CLV for each month throughout the dataset's time span and create features representing these monthly CLV values. This provides the model with a comprehensive view of customer spending patterns over time, enabling it to capture trends and seasonality.

We tested our approach on two datasets. The first, Online Retail II, uses only the newly engineered CLV features to train and test a regression model that predicts CLTV for the following month. We evaluate the model's performance using metrics like R-squared (R^2) and Mean Squared Error (MSE).

The second dataset, CDNow, takes a more comprehensive approach. We incorporate demographic information alongside the CLV features. This richer data allows the model to capture more complex patterns and potentially achieve better CLTV prediction accuracy. We then compare the performance of this model to the one trained solely on CLV features from Online Retail II.

The results reveal that adding demographic information significantly improves model performance. In Online Retail II (CLV features only), the model shows moderate performance with acceptable R^2 scores and reasonable MSE. However, in CDNow (CLV features + demographics), the inclusion of demographics leads to a substantial increase in R^2 and a decrease in MSE, indicating a more accurate CLTV prediction. This suggests that demographic characteristics provide valuable insights into customer behavior and spending patterns, significantly improving the model's predictive capabilities.

d. Deep Learning Model: Deep Neural Networks

Deep learning neural networks, a powerful subset of machine learning, offer significant advantages in CLV prediction compared to traditional methods. Their ability to handle complex data and model non-linear relationships can significantly improve the accuracy and insights businesses gain from CLV analysis. In this study, we use Deep Neural Networks in CLV prediction compared with above approaches.

Feature Engineering

After the data preprocessing process, we use feature engineering techniques such as recency, frequency, monetary (RFM) analysis and a naive benchmark is defined using the following parameters:

- Average basket value. This is calculated on all orders that are placed before the threshold date.
- Order count. This is calculated for the training interval on all orders that are placed before the threshold date.
- Count multiplier. This is calculated based on the ratio of the number of days before the threshold date and the number of days between the threshold date and now.

All features can be used to create meaningful features that capture customer behavior.

Training and target interval

Training a deep learning model for customer lifetime value (CLV) prediction involves strategically dividing your customer data. Here's how this process works:

1. Choosing a Threshold Date:

This critical step establishes a point in time that separates your customer order data into two distinct sets:

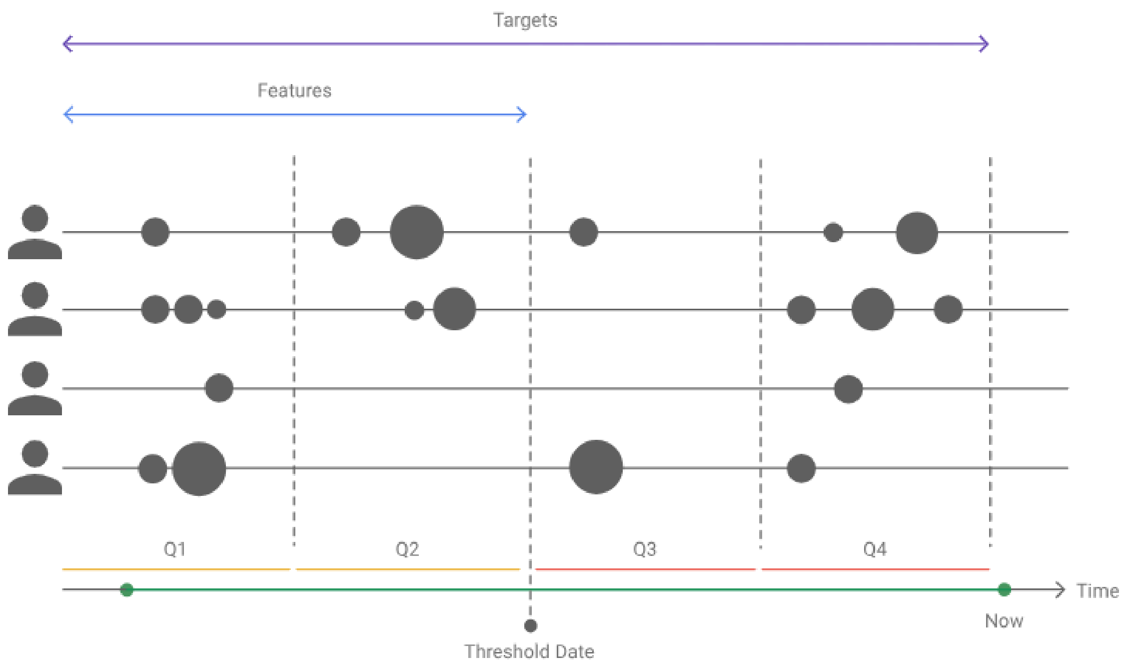
- Training Data**:** Orders placed before the threshold date are used to train the deep learning model. This data provides the model with historical patterns and customer behavior to learn from.
- Target Data: Orders placed after the threshold date are used to calculate the actual CLV (the target value). The model's performance will be evaluated based on its ability to predict CLV for these orders it hasn't seen before.

2. Rationale for Splitting Data:

Separating data this way ensures the model doesn't simply memorize historical data for prediction. By using unseen orders (post-threshold), we can assess the model's ability to generalize and accurately predict CLV for future customers.

3. Creating the training, evaluation, and test sets for DNN

- The *training* (70–80%) dataset is used to learn weights to reduce a loss function.
- The *evaluation* (10–15%) dataset is used during the training phase to prevent overfitting.
- The *test* (10–15%) dataset should be used only once, after all training and evaluation has been completed, to perform a final measure of model performance.



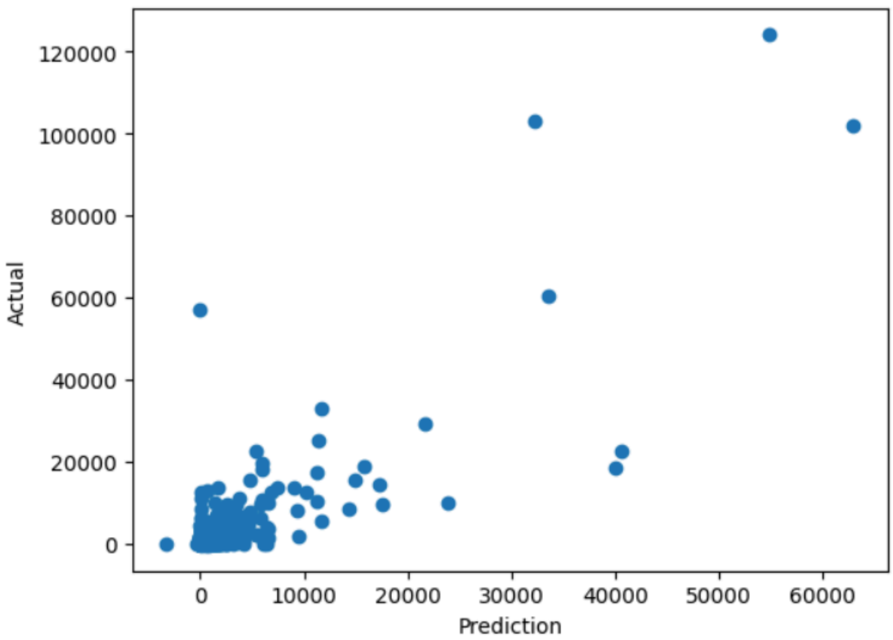
Model Development

Model development entails designing and training the deep learning architecture for CLV prediction. This process involves selecting an appropriate neural network architecture, defining the number of layers and neurons, choosing activation functions, and configuring optimization algorithms and regularization techniques.

Training the model involves feeding the preprocessed data to the neural network and adjusting the model parameters iteratively to minimize the prediction error. Hyperparameter tuning, using techniques like grid search or random search, helps optimize the model's performance on a validation dataset.

Evaluation Metric

For evaluating deep neural networks (DNNs) in Customer Lifetime Value (CLV) prediction, several key metrics are commonly used. These metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2), Adjusted R-squared, Mean Absolute Percentage Error (MAPE), Precision, and Recall. These metrics provide different insights into the performance of the model in predicting CLV values. Lower values of MSE, RMSE, MAE, and MAPE indicate better accuracy, while higher values of R^2 and Adjusted R^2 suggest a better fit of the model to the data. Precision and Recall are relevant if CLV prediction is treated as a classification problem.



4. **Results**

a. Probabilistic Model

	Online Retail II		CD Now	
	R2 Score	**RMSE**	**R2 Score**	**RMSE**
BG/NBD Model	**0.62**	**1.267**	**0.45**	**1.84**
Note	**Same duration hold out = 100 days**			

b. Machine Learning Model

- **Classification**

	XGBClassifier	K-Nearest Neighbors	Logistic Regression	Random Forest	Gradient Boosting
KMeans	**0.8**	**0.859**	**0.859**	**0.869**	**0.869**
KMeans ++	**0.85**	**0.788**	**0.838**	**0.879**	**0.848**
GMM	**0.84**	**0.869**	**0.889**	**0.848**	**0.838**

- **Regression**

	Online Retail II	
	R2 Score	**MSE**
LinearRegression	**1.00**	**3.453470e-21**
DecisionTreeRegressor	**0.96**	**1.186331e+07**
GradientBoostingRegressor	**0.86**	**4.262696e+07**
RandomForestRegressor	**0.65**	**1.031761e+08**
LGBMRegressor	**0.54**	**1.380273e+08**
XGBRegressor	**0.43**	**1.701170e+08**
CatBoostRegressor	**0.37**	**1.883615e+08**
Note	**Online Retail II do not use demographic features**	

	CD Now	
	R2 Score	**MSE**
LinearRegression	**1.00**	**5.420061e-23**
DecisionTreeRegressor	**0.999**	**1.058404e+04**
GradientBoostingRegressor	**0.999**	**3.669319e+03**
RandomForestRegressor	**0.999**	**1.438535e+04**
LGBMRegressor	**0.995**	**4.091690e+05**
XGBRegressor	**0.999**	**1.386258e+04**
CatBoostRegressor	**0.994**	**4.772642e+05**
Note	**CD Now use demographic features**	

More evaluate metric can found at Appendix

c. Deep Learning Model

	Online Retail II			
	R2 - Score	**MSE**	**MAPE**	**MAE**
Deep Neural Networks	**0.6052**	**3597224.67**	**3.13**	**424.97**

5. Conclusion

In summary, the implementation of Customer Lifetime Value (CLV) offers significant advantages for businesses, including strategic alignment with financial objectives, enhanced marketing strategies, and optimized resource allocation. However, potential drawbacks such as misalignment with company goals, incorrect customer segmentation, unrealistic lifetime expectations, and neglecting flexibility over time need to be carefully addressed to ensure effective utilization of CLV strategies.

Moreover, exploring various CLV modeling techniques provides valuable insights into customer behavior and revenue forecasting. Traditional CLV methods offer a foundational framework, while probabilistic models like BG/NBD offer more sophisticated approaches. Machine learning models such as RFM and regression models provide enhanced predictive capabilities, with notable variations in performance across datasets. Additionally, deep learning models like Deep Neural Networks exhibit promising potential for improving accuracy, as evidenced by their higher R2-Score compared to other models.

Overall, the choice of CLV model should be guided by the specific needs and characteristics of the business, considering factors such as data availability, complexity of customer interactions, and desired level of predictive accuracy. Integrating these insights into strategic decision-making processes can significantly enhance the effectiveness of CLV initiatives and drive sustainable growth.

6. Future Work

Expanding on the future work, we consider this task:

- **Seeking more data for retail banking sector:** Continue researching and gathering suitable data for the retail banking sector, including financial transaction information, purchasing behavior, and customer demographics. This data will help improve the accuracy of CLV models and generate more realistic forecasts.

- **Advanced Probabilistic Modeling Techniques:** Explore advanced probabilistic modeling techniques beyond Pareto/NBD and BG/NBD models, such as Hierarchical Bayesian Models or Hidden Markov Models. These sophisticated approaches can better capture complex patterns in customer transactions and improve CLV estimation accuracy.
- **Continuous Model Monitoring and Updating:** Implement a system for continuous monitoring of CLV models in production, regularly updating them with new data and recalibrating parameters as needed. This proactive approach ensures that the models remain accurate and relevant over time, despite changing market conditions and customer behavior.
- **Utilizing multi-source data, evaluating and comparing models:** Combine data from various sources such as transaction systems, customer relationship management (CRM) systems, and online services to create a comprehensive picture of customers and enhance the predictive power of CLV models, and Implement various evaluation and comparison methods between models to identify which model performs best for the retail banking sector. This helps determine the optimal model to deploy in practice and ensure the highest performance for the business's CLV strategies.

7. Appendix

a. Evaluation For Machine Learning Regression

Online Retail

	Regressor	R2	MAE	MSE	RMSE	MEDAE
0	LinearRegression	1.000000	1.027341e-11	3.453470e-21	5.876623e-11	1.818989e-12
1	DecisionTreeRegressor	0.960332	9.223990e+02	1.186331e+07	3.444316e+03	2.060500e+02
4	GradientBoostingRegressor	0.857468	4.561246e+02	4.262696e+07	6.528932e+03	4.792705e+01
2	RandomForestRegressor	0.655009	7.318080e+02	1.031761e+08	1.015756e+04	1.053502e+02
5	LGBMRegressor	0.538476	8.941187e+02	1.380273e+08	1.174850e+04	1.070854e+02
3	XGBRegressor	0.431177	8.685348e+02	1.701170e+08	1.304289e+04	8.626444e+01
6	CatBoostRegressor	0.370173	1.380410e+03	1.883615e+08	1.372449e+04	7.011096e+02

CD Now

	Regressor	R2	MAE	MSE	RMSE	MEDAE
0	LinearRegression	1.000000	2.195761e-12	5.420061e-23	7.362106e-12	6.821210e-13
4	GradientBoostingRegressor	0.999959	3.146992e+01	3.669319e+03	6.057491e+01	1.582883e+01
1	DecisionTreeRegressor	0.999882	3.589418e+00	1.058404e+04	1.028788e+02	0.000000e+00
3	XGBRegressor	0.999845	3.019992e+01	1.386258e+04	1.177395e+02	1.487908e+01
2	RandomForestRegressor	0.999839	9.003824e+00	1.438535e+04	1.199389e+02	3.060000e-02
5	LGBMRegressor	0.995420	9.037253e+01	4.091690e+05	6.396632e+02	3.206026e+01
6	CatBoostRegressor	0.994658	2.790080e+02	4.772642e+05	6.908431e+02	1.766007e+02

b. **Source Code**

The source code of this study can be found here.

8. References

1. Customer segmentation based on recency frequency monetary model: A case study in E-retailing - Bilişim Teknolojileri Dergisi, 2020
2. An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation H Abbasimehr, A Bahraini - Expert Systems with Applications, 2022
3. Modeling customer lifetime value S Gupta, D Hanssens, B Hardie… - Journal of service …, 2006
4. Modeling and application of customer lifetime value in online retail P Jasek, L Vrana, L Sperkova, ZZ Smutny, M Kobulsky - Informatics, 2018
5. Baesens B, Viaene S, Van den Poel D, Vanthienen J, Dedenne G (2002) Bayesian neural network learning for repeat purchase modeling in direct marketing. Eur J Oper Res 138(1):191–211
6. McCarty JA, Hastak M (2007) Segmentation approaches in data-mining: a comparison of RFM, CHAID, and logistic regression. J Bus Res 60(6):656–662