Analytics Lab

Customer Segmentation and Lifetime Value Prediction Using Machine Learning Approaches

ABSTRACT

- **CRISP-DM** framework was followed to solve specific problems for an Egyptian online retail company for more effective and efficient marketing activities.
- **Machine learning clustering algorithms** were applied to perform RFM customer segmentation for launching customized promotion campaigns that solve monthly orders fluctuations.
- **BG/NBD** and **Gamma-Gamma** models were used to predict 6-month customer lifetime values for launching feasible campaigns.
- A business intelligence **KPIs dashboard** was built to track the company's business performance.
- **Business recommendations** are provided to solve the company's business problems and enhance its performance.

PROBLEM STATEMENTS

1. Fluctuations in the number of **monthly orders** made by customers.
2. Low retention rates due to the **churned customers**.
3. The campaigns' impact on the **business performance is not measurable**.

OBJECTIVES

1. To determine the **best ML clustering** model for the company's RFM customer segmentation.
2. To **predict customers' lifetime value** for feasible retention and reactivation campaigns in solving churn problems.
3. To keep track of the company's key performance indicators **(KPIs)**.

Methodology

- The study follows the **CRISP-DM** framework:

✔ **Business Understanding:** where the problems and potential solutions were identified.

✔ **Data Understanding:** where the raw transaction data was collected using SQL and its description was documented as well.

✔ **Data Preparation:** where data went through cleaning and validation, extracting RFM and cohorts data, and finally preprocessed for machine learning models.

✔ **Modeling:** where K Means, Gaussian Mixture and HDBSCAN were built for segmentation. Also, BG/NBD and Gamma-Gamma models were used to predict customer lifetime values.

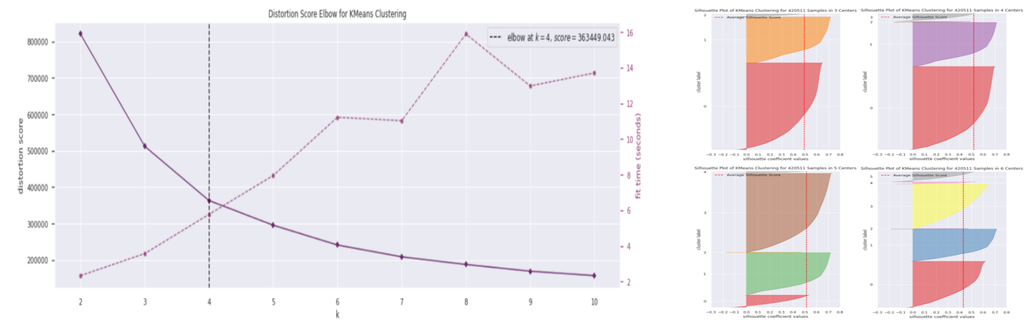
✔ **Evaluation:** Models were evaluated mainly using the Silhouette score analysis.

✔ **Deployment:** Finally, the models were deployed on the company's data to make recommendations based on the insights & results.

*(Coding was implemented in **Python** & Dashboard using **Tableau**)*

Discussion

- For the RFM segmentation, the **K Means** model with 4 clusters showed the **best** performance with a Silhouette score = 0.521



- **Gamma-Gamma** model was used to predict a **6-month** customer lifetime value putting a per-user threshold regarding the retention campaigns budgets.
- A BI KPIs **dashboard** was built to **monitor** active users, orders and revenues.

CONCLUSION

- **RFM customer segmentation** using **K Means** clustering algorithm is an excellent marketing tool for more **customized** and effective targeting.
- **Customer lifetime value prediction** using **Gamma-Gamma** model is a great way to get valuable insights into planning **feasible and efficient** campaigns.
- Business intelligence **dashboards** are indispensable tools for tracking business **KPIs performance**.

Master Candidate: Ayman Mohammed Altaweel

Supervisor: Assoc. Prof. Dr. Teh Sin Yin