Credit Card Customer Segmentation

By Bruce Hou (Change-Hung Hou)

Banks provide numerous services and one of its services, credit card, accounts for a great portion of its total profit. Credit card service generates revenue by charging interest on the loan amount while affording losses due to default. To gain a higher profit, banks start to tweak credit limits, interest rates, promotions, and other services based on customer characteristics. In this project, we aim to discover customized strategies with an unsupervised machine learning approach. Our research could help banks optimize their revenue while mitigating potential loss.

Methods

The Kaggle dataset we acquired captures credit card information from different customers. It contains 8950 rows and 18 columns. There is 1 id column, 14 float columns, and 3 integer columns. There are no duplicate values, but there are some missing values in the credit limit and minimum payment columns. We filled all missing values with the median value of their respective column. Now, the data is prepared for further analysis.

To start our analysis, we implement two clustering methods, Hierarchical and K-means, that group customers into several categories by analyzing the variables. We also utilize PCA, a dimensionality reduction approach, to reduce the numbers of columns. We then check if the reduction improves the quality of the clustering. Next, we select the best model and adjust its hyper parameters based on certain criteria, including silhouette score, inertia, and numbers of observations in each cluster. Lastly, we provide customized strategies by analyzing the median value of each variable in each cluster. Median is chosen for analysis as it represents the major trend of data.

Results and Business Recommendations

With the help of a unsupervised machine learning model, we successfully group customers into 7 categories, define their characteristics, and provide tailored marketing plans. Our findings are as follows:

Cluster	Characteristics	Strategies
2	Highest purchasing power	• Encourage them to use more credit cards • Offer more cash backs and discounts for credit card payment
3	Second highest purchasing power	
6	Most frequent cash advance user	• Look for what motivates them to use cash advance • If they use cash advance due to cash shortage, consider reducing credit limit • If tight budget is not the case, advertise more cash advance service
5	Cash advance user, lower repayment capacity	
1	Cash advance user	
0	Lower purchasing power	• Remain usual marketing strategy • Notify cluster 4 to pay for credit cards once limit is nearly full
4	Lower purchasing power, more reluctant to update balance	

Though cluster 2 and 3 share the same type of strategies, cluster 2 requires stronger promotion initiatives to match its higher purchasing power. A similar concept applies to clusters 1, 4, and 5, as well as cluster 0, and 4. Using the table above, a bank is now able to leverage their resources to their customers. This eliminates unnecessary costs, and thus, increases overall profit.