Predicting Churned Credit Card Customers

Nouf Alotaibi

Email: noufmitla@gmail.com

1. Introduction

This is the third T5 Data Science Bootcamp project, which is about building classification models that address a useful prediction and/or interpretation problem using Python with Sklearn. Below is a detail about the company *Bank Unity* I am collaborating with to help them with their problem, the dataset description, and the tools I will be using for the project.

2. Problem Statement

A manager at *Bank Unity* is disturbed by more and more customers leaving their credit card services. They would appreciate it if one could predict who is going to get churned, so they can proactively go to the customer to provide them better services and turn customers' decisions in the opposite direction [1]. Moreover, in this project, I will get familiar with Machine Learning modeling methods such as Logistic Regression, KNN, Decision Trees, Random Forest, and more to predict the customers who will get churned and who will not.

3. Dataset

For this project, I gathered the data from Kaggle:

- **Credit Card Customers Dataset:** it was uploaded to Kaggle.com, see [1]. The dataset contains 10,127 customers records and 21 features about the customers such as their age, salary, marital status, credit card limit, credit card category, etc. The table below illustrates the dataset's features and their types.

Column Name	Column type
CLIENTNUM	int64

Attrition_Flag	Object
Customer_Age	int64
Gender	Object
Dependent_count	int64
Education_Level	Object
Marital_Status	Object
Income_Category	Object
Card_Category	Object
Months_on_book	int64
Total_Relationship_Count	int64
Months_Inactive_12_mon	int64
Contacts_Count_12_mon	int64
Credit_Limit	float64
Total_Revolving_Bal	int64
Avg_Open_To_Buy	float64
Total_Amt_Chng_Q4_Q1	float64
Total_Trans_Amt	int64
Total_Trans_Ct	int64
Total_Ct_Chng_Q4_Q1	float64
Avg_Utilization_Ratio	float64

4. Tools

These are the technologies and libraries that I will be using for this project:

- **Technologies:** Python, Jupyter Notebook.
- **Libraries:** NumPy, Pandas, Matplotlib, Seaborn, Statsmodels, Scikit-learn, plotly.

5. Resources

[1] Credit Card customers. (n.d.). Kaggle: Your Machine Learning and Data Science Community. https://www.kaggle.com/sakshigoyal7/credit-card-customers