University of Connecticut - School of Business

OPIM 5604 - Predictive Modeling

Project 2 Proposal

Team 8b - Rohit Kamineni, Ryan Nelson, Supraja Bala

* What data you have chosen, including a link to where you found the data.
  + We have chosen a data set that contains the information of Credit Card Customers
  + Link: <https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers>
* What business value there might be from your findings and prediction. What business questions are you hoping to answer in your modeling work?
  + Based on this dataset, we will look into what type of customer has more likelihood to stop being a customer.
  + This data could potentially help enable the company to predict what customers will churn, and offer incentives to these customers in order to get them to stay, before they decide to leave.
* What you plan to predict from the data. Be specific about your target variable.
  + We will look at the categorical and continuous variables, in order to predict whether that customer profile is likely to attrit (leave).
* How many rows and how many columns are in the dataset. Give a brief description of what each row represents and what type of information is included in the dataset.
  + **CLIENTNUM**: Client Identifier
  + **Attrition\_Flag**: Client status, whether or not they have left this company
  + **Customer\_Age:** The customer's age
  + **Gender:** Customers Gender (M/F)
  + **Dependent\_count**: How many dependents the customer has
  + **Education\_Level**: What is the education level of the customer (Unknown, uneducated, Graduated, etc)
  + **Marital\_Status:** What is the Marital status of the customer (Single, Married, Divorced)
  + **Income\_Category**: What is the income amount (already in categories) of the customer
  + **Card\_Category**: What is the card type of the customer (customer card level)
  + **Months\_on\_book**: How many months the customer has been a customer
  + **Total\_Relationship\_Count**: Unclear what this column signifies
  + **Months\_Inactive\_12\_mon**: Months the customer would be inactive (ie no purchases
  + **Contacts\_Count\_12\_mon**: Unclear what this column signifies
  + **Credit\_Limit**: What the customer's credit limit for their card is
  + **Total\_Revolving\_Bal:** What the customer's current balance is
  + **Avg\_Open\_To\_Buy**: What the available amount of credit the customer has (Credit Limit - Total Revolving Balance)
  + **Total\_Amt\_Chng\_Q4\_Q1**: Unclear what this column signifies
  + **Total\_Trans\_Amt:** Total amount of all transactions this customer has had
  + **Total\_Trans\_Ct**: Total count of all transactions this customer has had
  + **Total\_Ct\_Chng\_Q4\_Q1**: Unclear what this column signifies
  + **Avg\_Utilization\_Ratio**: What the average utilization this customer has (Most likely Average Revolving Balance / Average Credit Limit)