Week 12 Deliverables

Group Name: Repeat Sales
Group Member #1: Griffin Palfrey
Email: griffinpalfrey@gmail.com
Country: Canada
College: Columbia University
Specialization: Data Analyst
Group Member #2: Phanindra Kumar
Email: phanindrakumar1998@gmail.com
Country: United States of America
College/Company: University of Maryland, Baltimore County
Specialization: Data Analyst

Problem Description

XYZ Credit Union in Latin America are selling business products such as Credit Cards, Deposit Accounts, and Retirement Accounts, but they are not performing well in cross selling: persuading customers to acquire more than one of their banking products. This is resulting in reduced customer satisfaction, loyalty, and opportunities for revenue growth.

Impact: Reduced revenue from business products, missed opportunities for enhanced client engagement, and potential loss of market share to competitors who offer more comprehensive solutions.

Objectives: To enhance the effectiveness of cross-selling business products, thereby increasing client uptake and revenue.

Model Rational

When deciding on the right model for our dataset, we had to take multiple factors into consideration. We eventually landed on XGBoost. This is our rationale.

XGBoost is highly suitable for our dataset because of its efficiency in handling large files and its ability to model complex, non-linear relationships. Given that we are measuring a non-linear relationship between 'Age', 'Seniority_Months', and 'Number of Accounts', XGBoost's decision-tree-based method is ideal for capturing intricate patterns that linear models might miss. Moreover, XGBoost is resilient to missing values, outliers, and multicollinearity, making it reliable for diverse real-world data scenarios, even though our data should be properly cleaned. On top of that, the other models we tried inflated or conflated at least one of our output measurements, making XGBoost an even more desirable model.

While XGBoost is sometimes regarded as a "black box" model, interpretability tools like SHAP can provide valuable insights into how individual features contribute to predictions. Its strong performance, scalability, and flexibility in managing non-linear relationships make XGBoost an excellent choice for our task.