Case Study: Direct Marketing Optimization

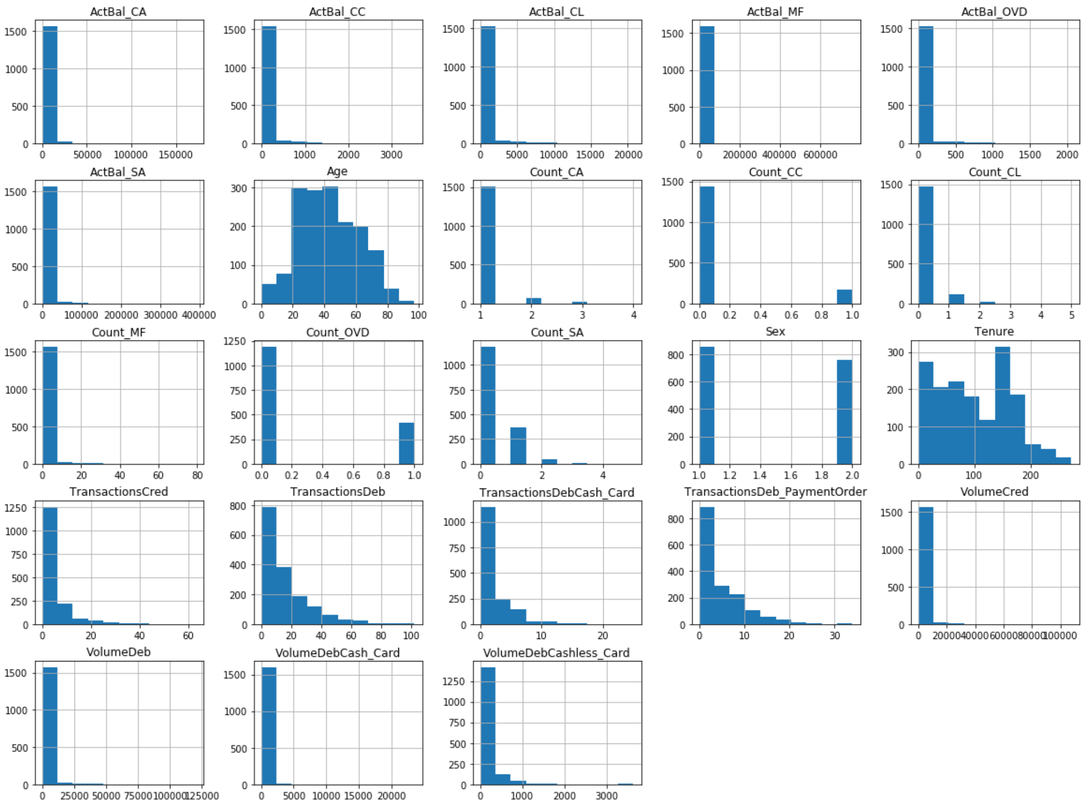
Ronald Wihal Oei

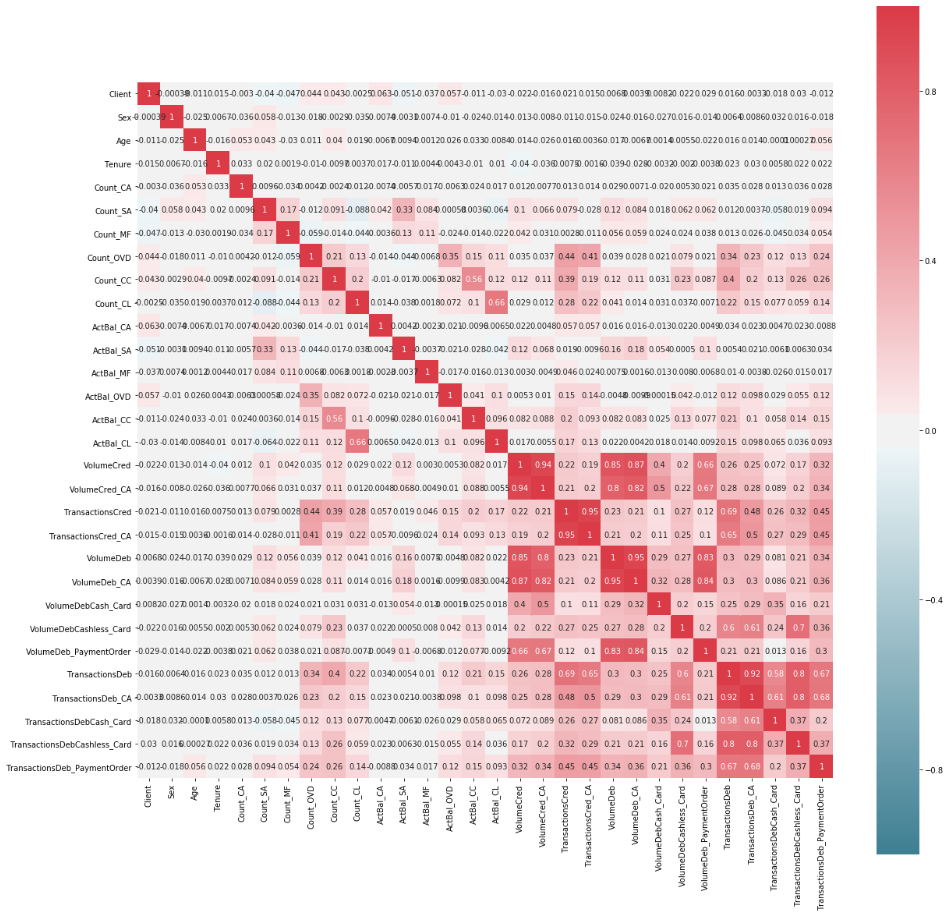
28th January 2025

Objective and Workflow

- Objective:
  - Select the most potential 100 clients to target with offers for mutual funds, credit cards and consumer loans to maximize revenue from direct marketing campaign.
- Workflow:
  - Data preparation: data loading and inspection, feature engineering, handling of missing data, creation of train and test sets.
  - Development of propensity score model and evaluation.
  - Optimization of targeting strategy: allocate only one offer to each of 15% of clients to maximize revenue and estimate the expected revenue.

Data Description





Data Pre-Processing

- Remove those with unknown gender.
- Fill all other NA values with 0 (assuming missing indicates no holdings or activities or purchases).
- Remove features with correlation coefficients above 0.8.

Modelling Approach

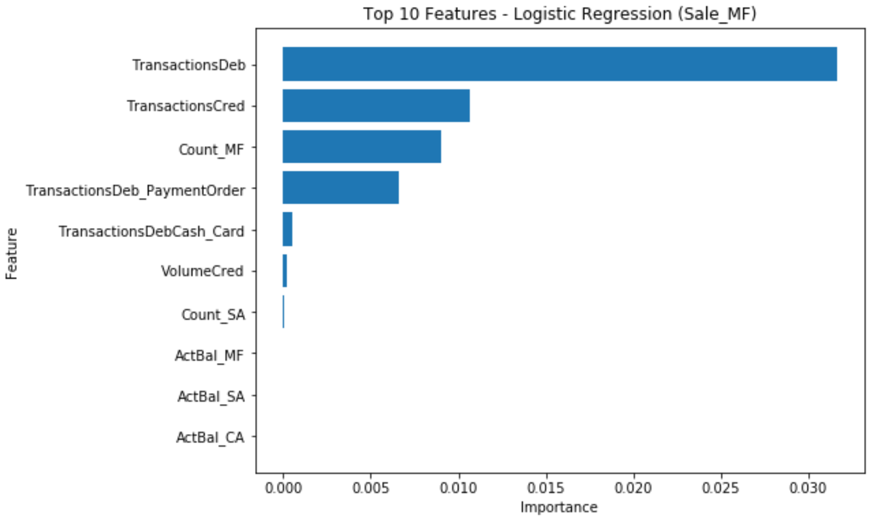
- Three models are compared:
  - Logistic regression
  - Random forest
  - Gradient boosting
- SMOTE and class weights are applied to address class imbalance.
- Evaluate models using area under thr ROC curve (AUC-ROC), Precision, Recall, and F1-score.

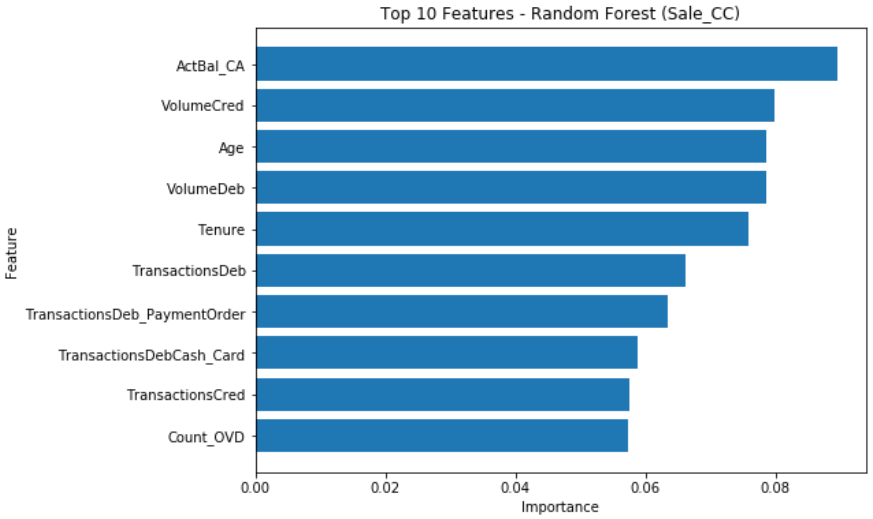
Results

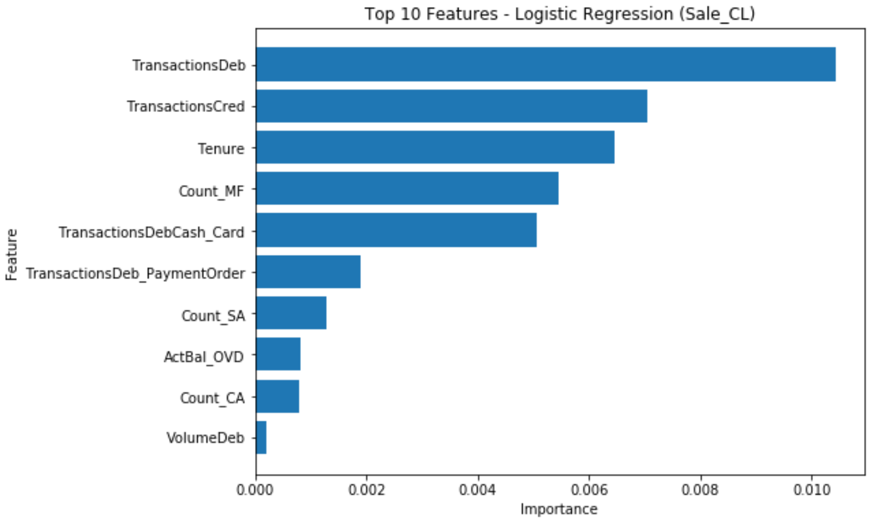
Model	Mutual fund		Credit card		Consumer loan	
	AUC-ROC	F1-score	AUC-ROC	F1-score	AUC-ROC	F1-score
Logistic Regression	**0.696**	**0.61**	0.567	0.53	**0.731**	**0.63**
Random Forest	0.652	0.53	**0.582**	**0.58**	0.597	0.58
Gradient Boosting	0.585	0.56	0.533	0.52	0.558	0.53

Based on the model performance, logistic regression model is chosen for mutual fund modelling and consumer loan modelling, while random forest model is chosen for credit card modelling.

Feature Importance Based on the Best Models







High-Propensity Client Lists and Targeting Strategy

- Use the best models to output the propensity of clients buying the three products.
- Rank clients by propensity and select the top 100 clients with the highest propensity.
- In case there is a duplicated client, remove duplicated clients and keep the offer with the highest propensity.
- Complete list can be seen on the ipynb file.
- Calculate the total expected revenue as follows:

$$ExpectedRevenue = propensity * \overline{revenue}$$

  where $\overline{revenue}$ is the mean of revenue for that specific product.
- The total expected revenue is calculated by summing the expected revenue from the top 100 clients, which is **869.452**.

Steps to Improve Model Performance

- Feature engineering: create new features, such as aggregated product ownership, ratios of credit/debit transactions or balances, age groups and tenure categories.
- Address class imbalance (low ratio of Sale_MF, Sale_CC and Sale_CL with ~20%, 25% and 30% respectively): SMOTE, ADASYN, oversampling or class weights in the model.
- Build a separate model to predict expected revenue, such as regression model, rather than using the mean revenue for a specific product.