Data Scientist Case Study: Direct Marketing Optimization

Objective

Maximize revenue from direct marketing campaigns using the provided dummy data.

Business Context: This case study simulates a real-world marketing scenario aimed at optimizing resource allocation to maximize revenue.

Data Description

You will work with several datasets (see attached) containing the following information:

- Social-Demographic Information: Age, gender, and bank tenure.
- Product Holdings and Volumes: Details on current accounts, savings accounts, mutual funds, overdrafts, credit cards, and consumer loans.
- Financial Transactions: Aggregated inflow/outflow on current accounts and average monthly card turnover over the past three months.
- Sales and Revenue Data: Available for 60% of clients, serving as a training set.

Constraints

- Contact Limitation: The bank can reach out to only 15% of clients (approximately 100 individuals).
- Single Offer Per Client: Each client can receive only one marketing offer. For example, a client might receive an offer for a consumer loan, a new credit card, or an investment in mutual funds based on their profile and financial behavior.

Task Breakdown

1. Create Analytical Datasets: Prepare both train and test datasets using the provided data.
2. Develop Propensity Models: Build three models to estimate the likelihood of purchase for:
    a. Consumer Loan
    b. Credit Card
    c. Mutual Fund
3. Optimize Targeting Strategy: Select clients to target with marketing offers to maximize revenue, adhering to the constraints. Clearly document your approach.

Submission Requirements

1. **High-Propensity Client Lists and Targeting Strategy:** Submissions must directly answer the following questions with justifications:
    - Which clients have a higher propensity to buy a consumer loan?
    - Which clients have a higher propensity to buy a credit card?
    - Which clients have a higher propensity to buy a mutual fund?

    - Which clients are to be targeted with which offer?
    - What would be the expected revenue based on your strategy?
2. Deliverables
    a. A concise executive summary (up to 2 pages) OR a slide deck (5–8 slides) for a technical audience.
    b. A Targeted Client List specifying selected clients for each offer.
    c. A Codebase with all necessary files
3. Instructions on How to Submit
    - **Commit all deliverables to a public GitHub repository** in your personal account. *(We do not accept zipped submissions attached in emails.)*
    - **DO NOT include this document** or sensitive terms like "Singlife" in the repository.
    - **Email the GitHub repository link** to the sender of this test. Submission deadline is communicated by the sender of the test, please check your email.

Evaluation Criteria

Submissions will be assessed based on the following:

- **Correctness**: The correctness and soundness of the data science methodology, i.e. preprocessing, sampling, modelling, evaluation, optimization, inference, explainability, etc.
    - *Note: we do not grade submissions based on model performance.*
- **Code Quality**: The cleanliness and organization of your code
- **Communication**: Clarity and effectiveness in presenting findings through the executive summary or slide deck.
- **Business Impact**: Alignment of the targeting strategy with the objective of maximizing revenue.

Key Notes

- **Time Commitment:** This assessment is expected to take **1-2 days**. We understand candidates may have full-time roles & other commitments. Therefore, we welcome partial submissions.
- Assumptions: If certain details are unspecified, make reasonable assumptions and proceed accordingly.
- Tool Requirement: All submissions must use Python as the primary programming language.
- For Candidates with Production Code Experience: **While not required, candidates who demonstrate production-ready code will be recognized**. This includes writing tests, managing dependencies, containerizing workflows, modularizing Jupyter notebooks, utilizing experimentation frameworks, proper documentation, etc.