Bank Campaign Project Description

Data Glacier Internship

Team Member Detail

Group Name: Zeru Zhou's Group

Name: Zeru Zhou

Email: zeruzhou9@gmail.com

Country: United States

College: University of Southern California

Specialization: Data Science

Problem Description

This project aims to fix the problem that there are too many potential clients associate with a Portuguese banking institution. The bank marketing campaigns are primarily based on phone calls. However, there are too many potential clients, and it is impossible to call all of them. Hence, some machine learning techniques are needed to classify the customers and predict if they are going to subscribe the term deposit or not. By implementing ML models, we are able to limit the number of 'potential clients' to an acceptable number that make sure the marketing strategies can be carried on.

Business Understanding

To evaluate this problem in the business scale, we should look at the profit we are going to have. The cost is relatively fixed, which is the salary of those agents that making the phone call to potential clients. The cost depends on the number of agents that the bank hired. What we need to do first is to evaluate the profit that could be made if one person subscribe the term deposit successfully. If that is a large amount profit, our model should put more weight on 'recall', that is, to make sure we capture as much potential clients as possible to maximize the profit. Conversely, if that profit is only a tiny amount, our model should put more weight on 'precision', that is, to make sure we precisely capture the potential clients even at cost of missing some potential customers. This action will reduce our cost by hiring less agents if our filtering algorithm is aggressive enough. In a word, the business scale is the tradeoff between cost and

revenue. All strategies need to depend on the cost of hiring and the profit made by convincing the clients.

Project lifecycle along with deadline

Data Understanding, Cleaning, Imputing, Encoding, Scaling. Finish all the feature engineering except for feature selection during the model construction step.	**Deadline: Feb 2nd, 2023**
Provide EDA analysis and insights. Visualize the data to present meaning that is reasonable to layperson.	**Deadline: Feb 9th, 2023**
Finish EDA presentation with detailed data visualization and presentation slides.	**Deadline: Feb 16th, 2023**
Build Machine Learning pipelines and validate, optimize them. Select the best model to deploy.	**Deadline: Feb 23rd, 2023**
Finish the final report with the result from last week. Explain the technical part with professional aspect.	**Deadline: Feb 28th, 2023**