

Bank Customer Prediction

Developer: Claudia Xu
QA: Ziying Wang

Motivation

- The Portuguese bank spends excessive budget on marketing but there is not a significant change in the number of bank clients who buy the financial product (term deposit).
- In order to handle this problem, we need to evaluate the area of improvement for the bank's marketing strategy that can help ultimately increase more bank clients purchase the produce. In this way, the bank can earn more revenue
- Drive subscription of the term deposit by using data of clients and marketing strategies to inform targeting decision





Demo

<http://18.219.212.25:3000/index>



Data Description

Kaggle Bank Market Dataset

The data is related with direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to predict if the client will subscribe a term deposit (variable y).

Features:

Customers: age, job, marital, education, balance, etc

Campaign: campaign(number of contact), previous(number of previous contact before this campaign),etc





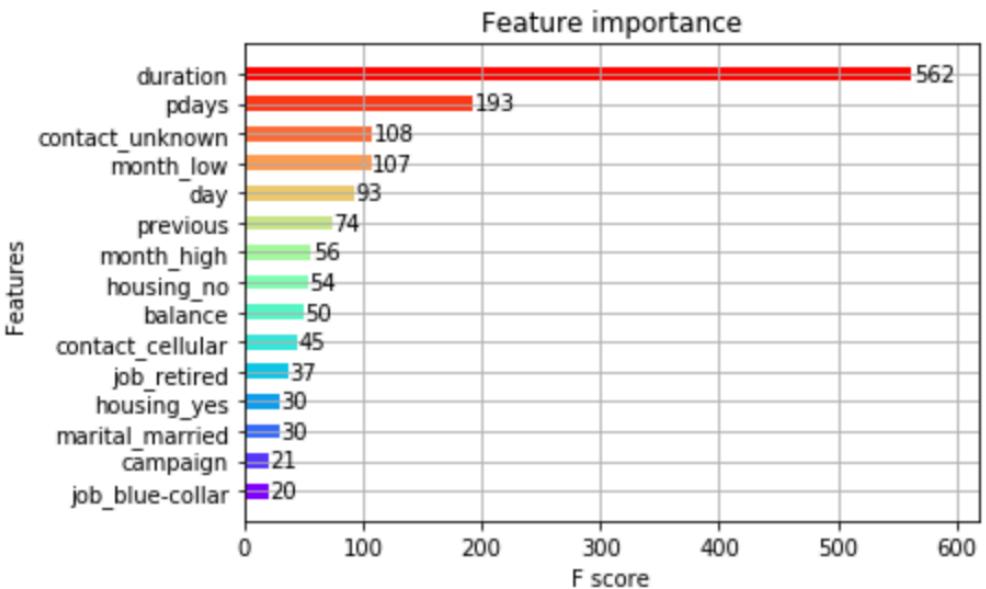
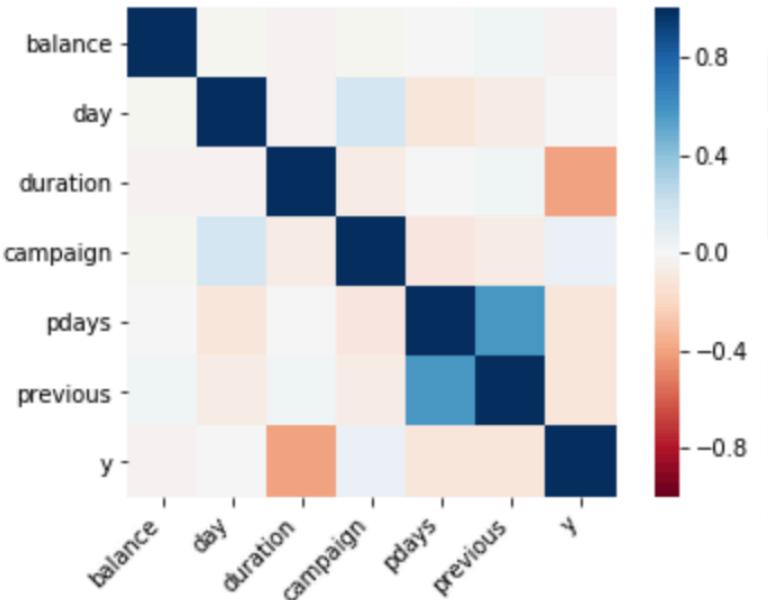
Model Description

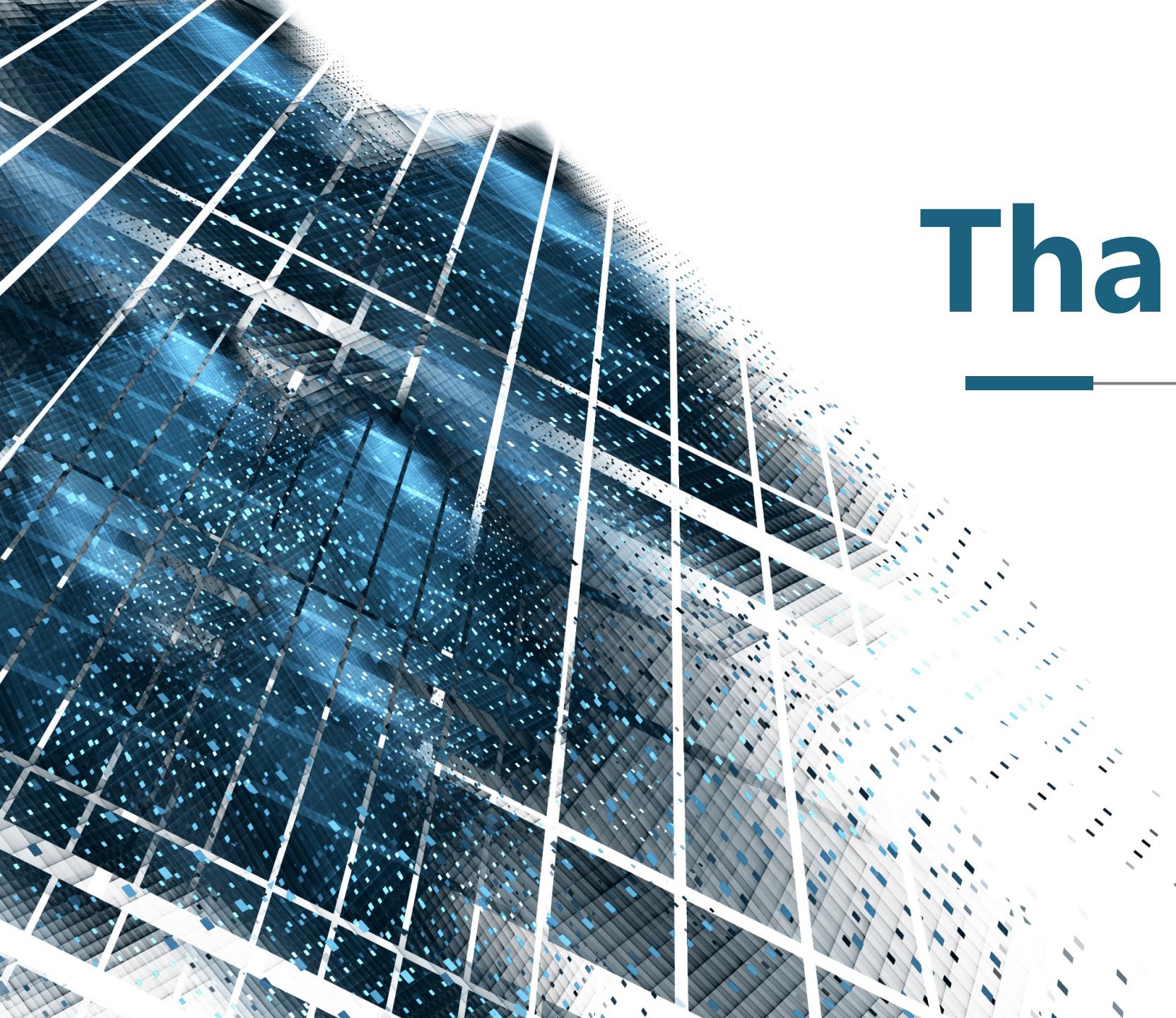
- Comparison between Logistic Regression, XGBoost, Decision Tree, and Random Forest models on the down-sampled data with classification cutoff point 0.5
- **XGBoost**
 - max_depth: 3
 - n_estimators: 300
 - learning_rate: 0.05
- **Best result**
 - AUC: 0.7226
 - Accuracy: 0.6741
 - F1 Score: 0.6577
- **Success Criteria**
 - Accuracy score > 0.65
 - An increase in clients' subscription rate of the term deposit



Interesting Insights

- Duration of the call is highly correlated with y , and is statistically important.
- However, we remove it in our final model
 - Since the duration is only known after a call is finished when y is clearly known
- Thus, this feature should be discarded if the intention is to have a realistic predictive model.



A large, abstract graphic on the left side of the slide features a complex pattern of blue and white geometric shapes, resembling a stylized globe or a network of data points. It has a high-resolution, pixelated appearance with a color gradient from dark blue to light blue and white.

Thank you

Contact Info:

Email:
lininxu2019@u.northwestern.edu

Github:
[https://github.com/claudia1462656918/MSI
A-Value-Chain-Project](https://github.com/claudia1462656918/MSI-A-Value-Chain-Project)