

Business Analytics Student Competition 2022



Team:
Data Squad

Overview

Agenda

- Identify a shortlist of B2B customers who have a likelihood to buy products after being subjected to marketing campaign.
- Find characteristics of potential customers.

Every company has a limited budget for marketing campaign. A better strategy to effectively use the budget is by targeting customers who have a higher probability of buying a product or service.

Data Overview

- The dataset consists of 46332 records and 34 features.
- Broadly, the features tell us about the:
 - Contact Attributes (Details about the person in contact)
 - Digital Interaction (Communication between client and companies)
 - Firmographics (Additional information about the companies)
 - IT Spend (Past deals)
 - Past Purchases (Whether the companies bought from us)

Data challenges

◆ Missing Values

- 469 missing values were found in **persona_tech** predictor.
- In the **number of employees** and **industry vertical information** features, nearly 80% of the records are empty

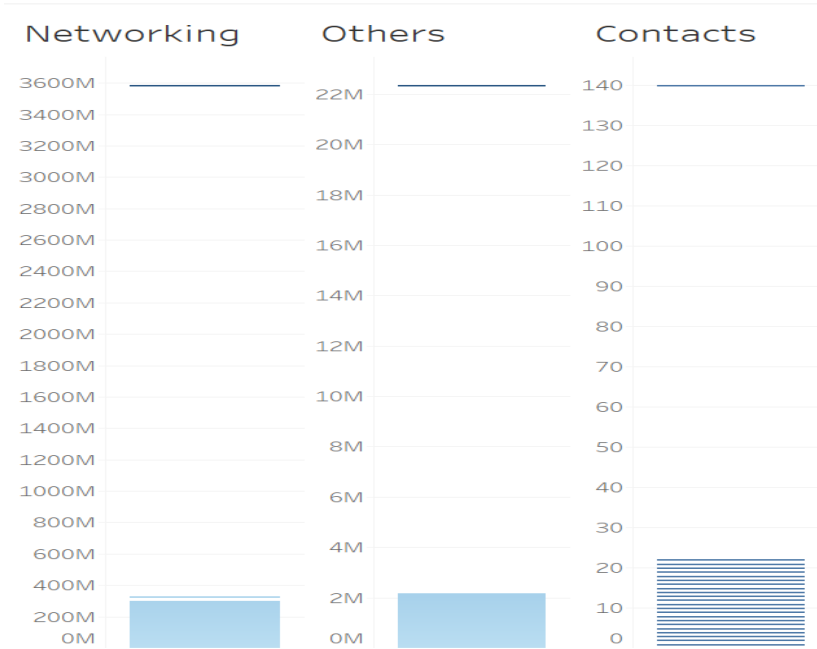
Handling Missing Values

- ◆ For the **persona_tech** feature the missing values account for only 1% of the data so they were dropped.
- ◆ For the other 2 features, the percentage of missing values are quite high and did not provide much information hence, these features were also dropped



Data challenges

◆ Outliers



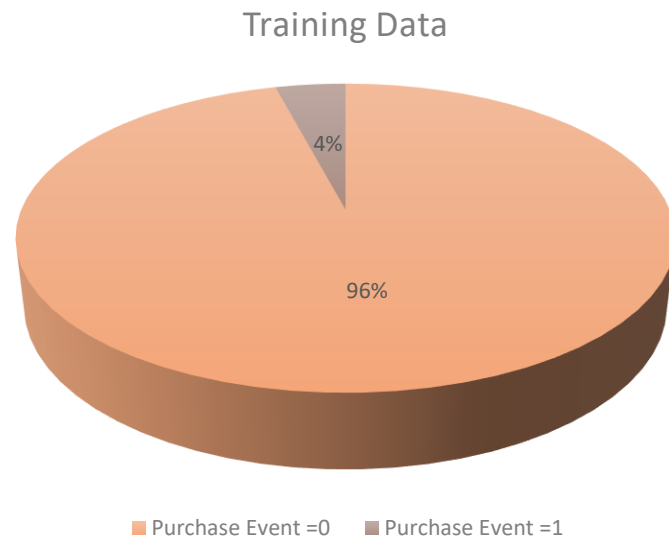
Handling Outliers

- ◆ Outliers comprise of a very minute percentage of data and we can see that these are extreme values hence it is better to get rid of them

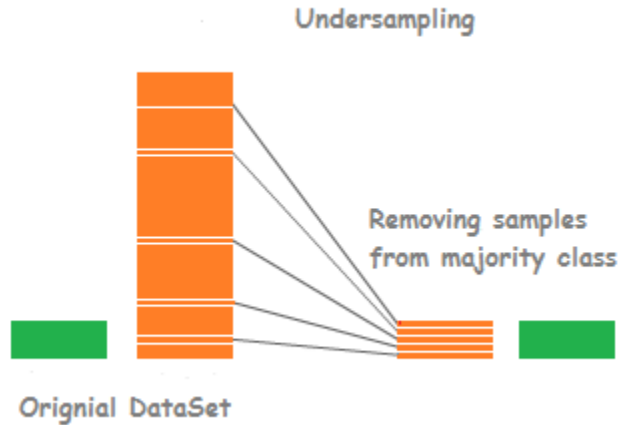


Data challenges

- ◆ Imbalanced dataset
- ◆ We observed that 96% of the records belong to non-responders whereas only 4 % of the records belong to the responders



Handling Imbalance



Results



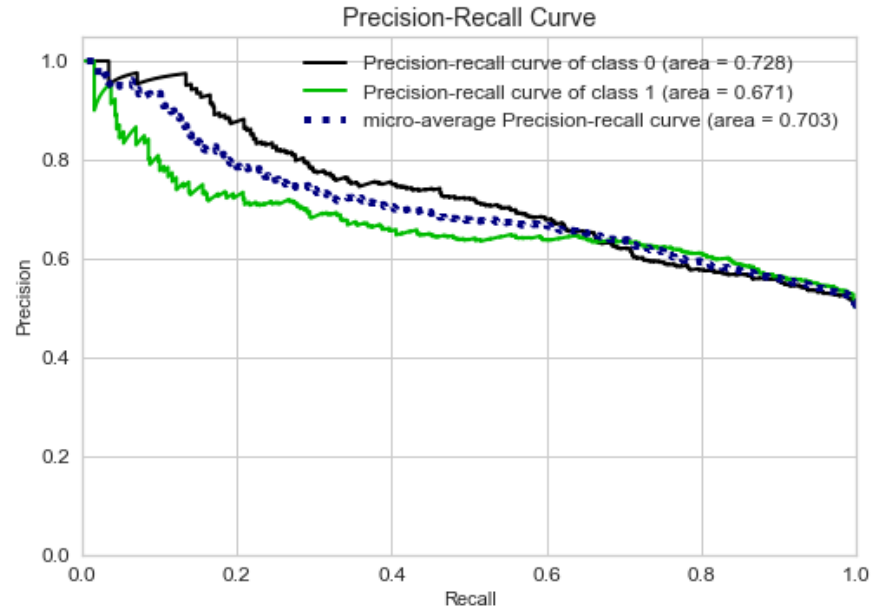
Buyer Characteristics

- 66% of the buyers have a technical contact associated with the account in the known contact database.
- 73% of buyers are decision headquarters.
- 91% of buyers didn't show up for any events



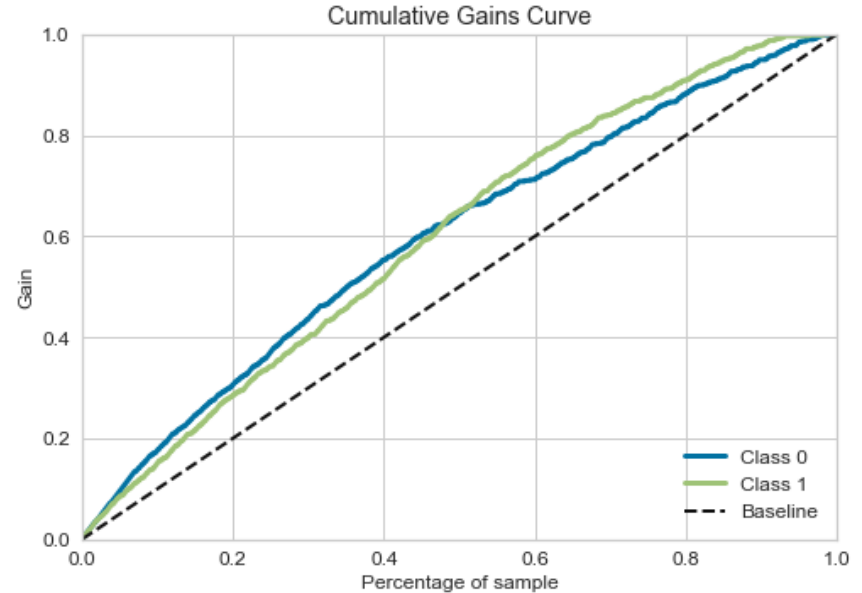
Results

As a result , we have achieved precision-recall score of 0.70



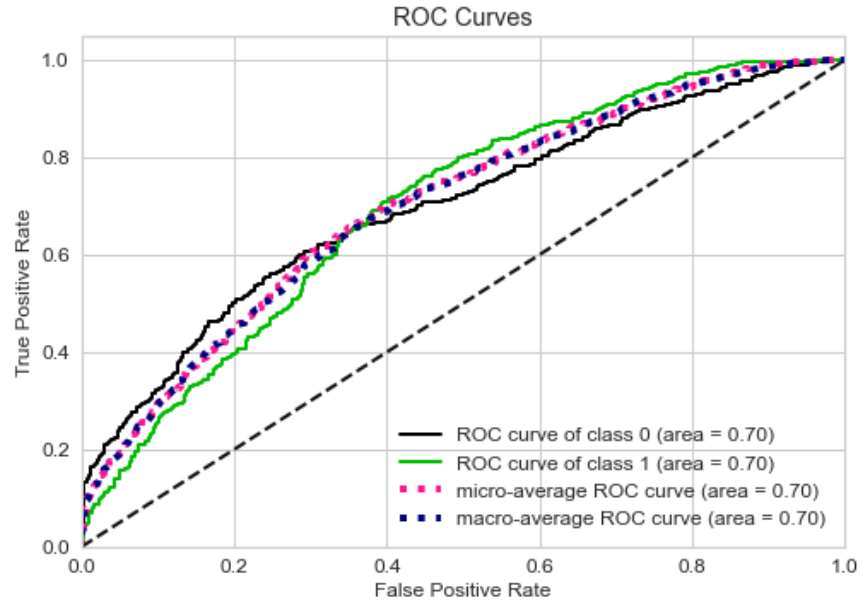
Results

For Top 20 % of the results we are able to target more than 20% of the responders



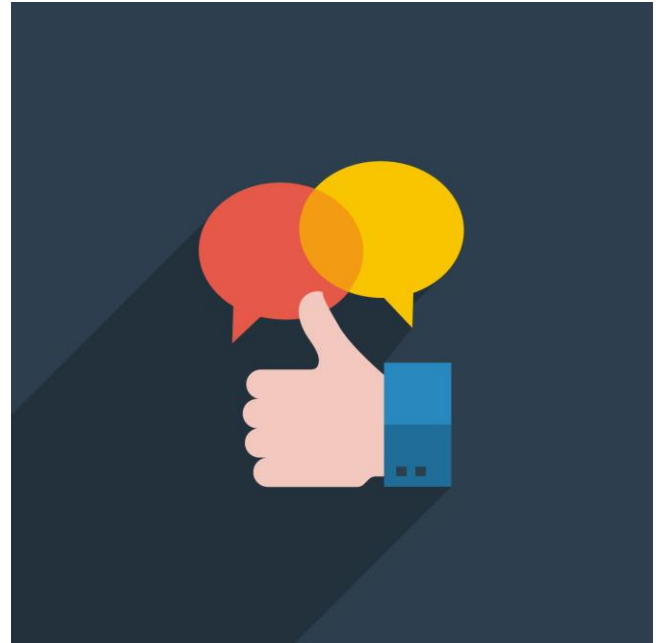
Results

As a result , we have achieved roc_auc score of around 0.70



Recommendations

- 3% of the total event attendance are customers
- 96% of buyers didn't engage in any chat activities with the company.



THANKS!

