P19 - Uplift Modelling

TO SMS, OR NOT TO SMS: THAT IS THE QUESTION!
Team

Polina Shvechenko
CS Major
Marketing Data Analyst
at Bolt

Sander Sats
CS Major
Senior Backend Developer at Stebby

Tejas Anil Shah
CS Major
Full time student
Colab Notebook:
https://colab.research.google.com/drive/16lm2oULTMWL3Ezstxy-VWOD3ICFU_EU3

Slides:
https://slides.com/tejasanilshah/ml-p19/
Problem Statement

![Four-quadrant diagram showing decision outcomes: Persuadables, Sure things, Lost causes, Sleeping dogs.]

- **Persuadables**: Buy if treated Yes
- **Sure things**: Buy if not treated Yes
- **Lost causes**: Buy if not treated No
- **Sleeping dogs**: Buy if treated No
Problem Statement

Classification
### Data

#### X5 Retail Hero

<table>
<thead>
<tr>
<th>Filename</th>
<th>Description</th>
<th>Shape</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>products: All products that the chain sells</td>
<td>(43038, 10)</td>
<td>3.17 MB</td>
</tr>
<tr>
<td>2</td>
<td>clients: Attributes of the clients (id, age, gender etc)</td>
<td>(400162, 4)</td>
<td>20.73 MB</td>
</tr>
<tr>
<td>3</td>
<td>purchases: All purchases identified by the client ID</td>
<td>(45786568, 12)</td>
<td>4.16 GB</td>
</tr>
</tbody>
</table>
Explore  
Understand the data

Filter  
Clean it

Engineer  
Construct new features

Split  
The famous train, test split

Train  
model.fit()

Evaluate  
Change parameters and repeat
Metric

\[ \text{uplift} = P^T - P^C \]

- \( P^T = P(Y = 1|W = 1) \)
- is the probability of purchasing when we target the client.

- \( P^C = P(Y = 1|W = 0) \)
- is the probability of purchasing when we don’t target the client.
Model Parameters

xgb_params = {
    'max_depth': 2,
    'learning_rate': 0.05,
    'n_estimators': 200,
    'nthread': 50,
    'n_gpus': 0,
    'seed': 42,
    'use_label_encoder': False,
    'booster': 'gbtree',
    'reg_alpha': 0.9,
    'reg_lambda': 0.9
}

## Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Uplift@30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoModels</td>
<td>0.0516</td>
</tr>
<tr>
<td>ClassTransformation</td>
<td>0.0617</td>
</tr>
<tr>
<td>Baseline from competition</td>
<td>0.0622</td>
</tr>
<tr>
<td>Our model</td>
<td>0.0854</td>
</tr>
</tbody>
</table>
AUC

Baseline

Our model
Profit?
Lessons Learnt

- Understand the data
- Math is hard and that's okay
- Pair programming = awesome
Profit!
Questions