Donations raising campaign 2023

Sebastián Cabrera

Content

1. Business Problem

- A Charity Company currently has **1.000.000 members** subscribed to their annual newsletter.

- In 2022, the management team decided to call 100.000 members and ask them to donate for a new campaign for the children. Each call had a cost of $13.

- Results were positive but not as expected as only 14.719 members donated a total amount of $1,6 million. The net amount collected was $390k
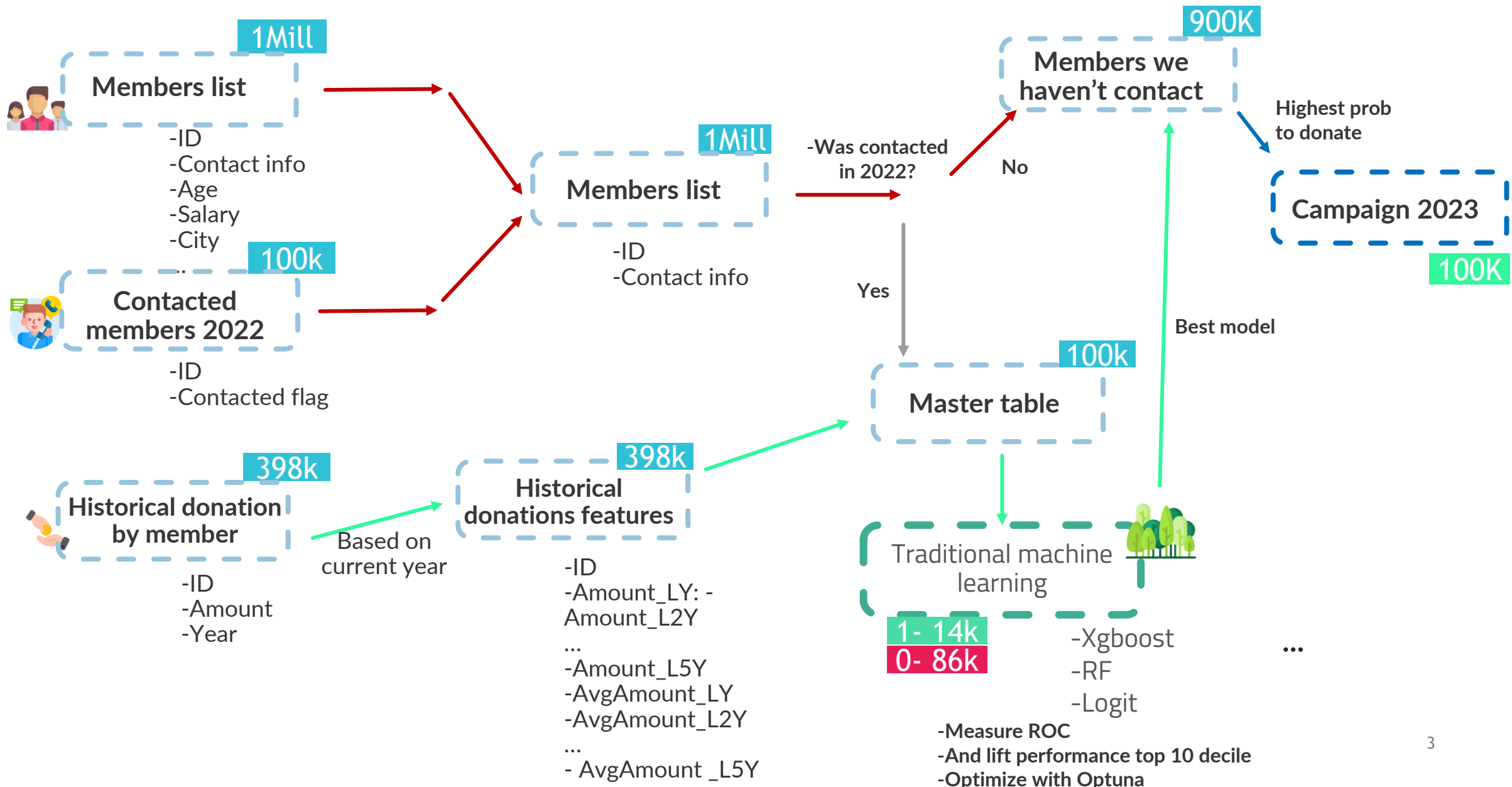
Contacted_2022	Potential_Donors	Real_Donors	Donations	Avg_Donations	Cost	Net_Collected
No	900.000	109.965	5.735.855	52,16	0	5.735.855
Yes	100.000	14.719	1.690.630	114,86	1.300.000	390.630
Total	1.000.000	124.684	7.426.485	59,56	1.300.000	6.126.485

- $5.7 million of the total collected, came from donors that we didn't call.

…

How can we better select the members to call for the upcoming 2023 campaign?

2. Methodology

Members list — 1Mill
- ID
- Contact info
- Age
- Salary
- City

Contacted members 2022 — 100k
- ID
- Contacted flag

Historical donation by member — 398k
- ID
- Amount
- Year

Based on current year

Historical donations features — 398k
- ID
- Amount_LY: - Amount_L2Y
...
- Amount_L5Y
- AvgAmount_LY
- AvgAmount_L2Y
...
- AvgAmount _L5Y

Members list — 1Mill
- ID
- Contact info

-Was contacted in 2022?

No

Yes

Members we haven't contact — 900K

Highest prob to donate

Campaign 2023 — 100K

Best model

Master table — 100k

Traditional machine learning

1- 14k
0- 86k

- Xgboost
- RF
- Logit
...

-Measure ROC
-And lift performance top 10 decile
-Optimize with Optuna

3. Key findings and insights

Group	Min	Max	Members	Donated in 2022	%	Lift
0	0.006017	0.030590	2000	29	1.45	0.098639
1	0.030622	0.046198	2000	74	3.70	0.251701
2	0.046234	0.061354	2000	91	4.55	0.309524
3	0.061357	0.079561	2000	134	6.70	0.455782
4	0.079572	0.100424	2000	204	10.20	0.693878
5	0.100426	0.127963	2000	236	11.80	0.802721
6	0.127968	0.162264	2000	284	14.20	0.965986
7	0.162283	0.224331	2000	374	18.70	1.272109
8	0.224402	0.332871	2000	560	28.00	1.904762
9	0.332983	0.936550	2000	954	47.70	3.244898
Total			20,000	2,940	14,7%	1

In the testing set, we found that the members with the highest probability had a **Conversion rate of 47,7%**

3.2 higher than the campaign from 2022

If we now **use the model** to select 100.000 members to call, and have a similar conversion, we **expect $5,3** collected for a **net of $4.0** after calling costs

*100.000 * 47% * 114.8 = 5,39 million; $114.8 was the avg collected by the Members that were called in 2022 and donated.

4. Model performance metrics

While training each model, we assessed two key metrics:

(1) the AUC, which is preferred over accuracy, as the dataset is imbalanced (% of donors is 14%)

It behaves reasonably well, outperforming the commonly used logistic regression as benchmark

(2) The Lift of the 10 decile in the testing set for each model.

For both Random Forest and Xgboost exceeded 3.0, which means the model in the individuals with identifies with highest probability expects a conversion rate 3 times higher than the average.

ROC Curve Comparison

Logistic Regression (AUC = 0.68)
Random Forest (AUC = 0.77)
XGBoost (AUC = 0.77)
Random Classifier
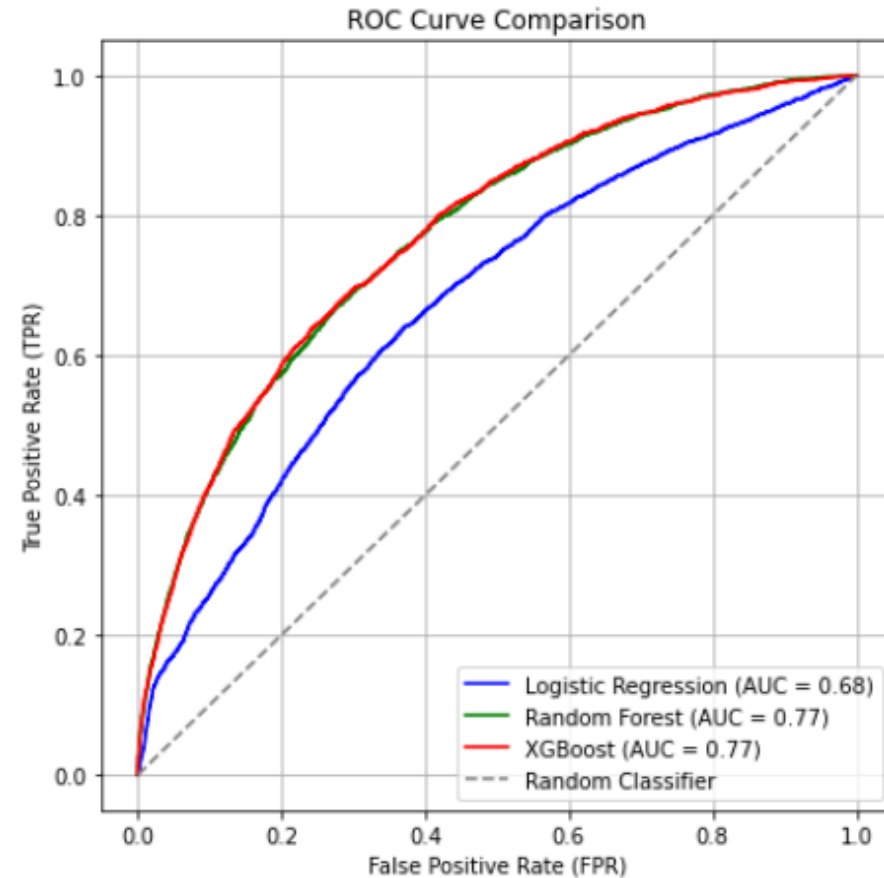
5. Instructions for running the code

Running main.py, will start the procedure to merge the different data sources, activates the training, and lastly generates the csv with the 100.000 donors to be used in the 2023 campaign.