

Direct Marketing Campaign

(Phone Calls Only)

May 2008 —
November 2010

Project Goals

- Be able to predict, given a set of features, whether a customer will invest in a term deposit.
- Know which features are most important in the prediction.
- Maximize total available revenue.
- Minimize cost of future campaigns.

Our Data

Collected between May 2008 & November 2010

36,548 did **not** invest + 4,640 did invest = 41,188 total data points

Available features included (among 11 others):

- Presence of a personal loan
- Last contact date
- Consumer price index *
- Consumer confidence index *
- Euribor 3 month rate *
- Age
- Job
- Marital status
- Education
- Presence of a housing loan

* At time of contact

What is most important to measure?

Costs

**Did not
invest**

Mean call length:
3.681 mins

Mean cost per call:
\$0.92

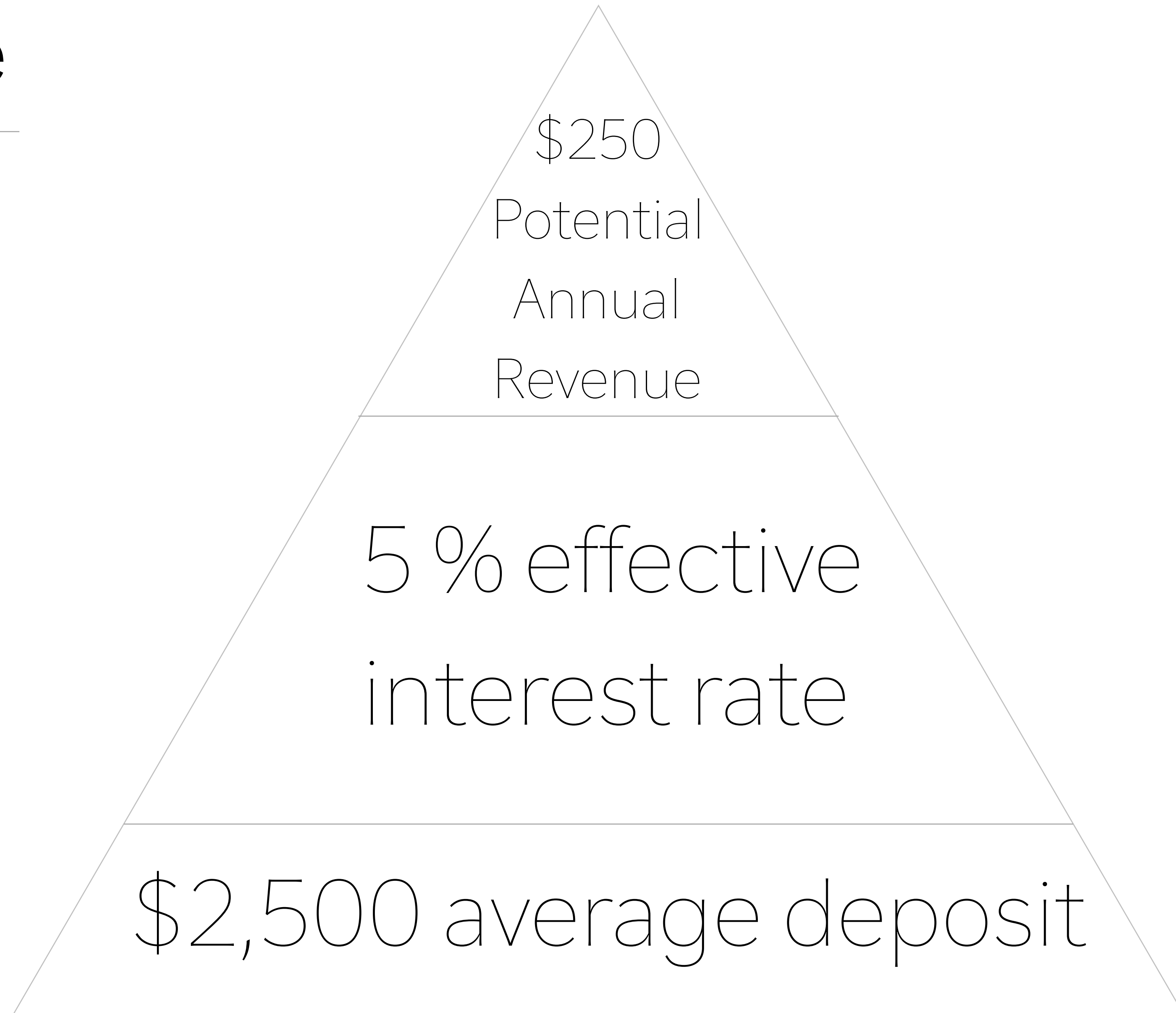
Cost per
minute:
\$0.25

**Did
invest**

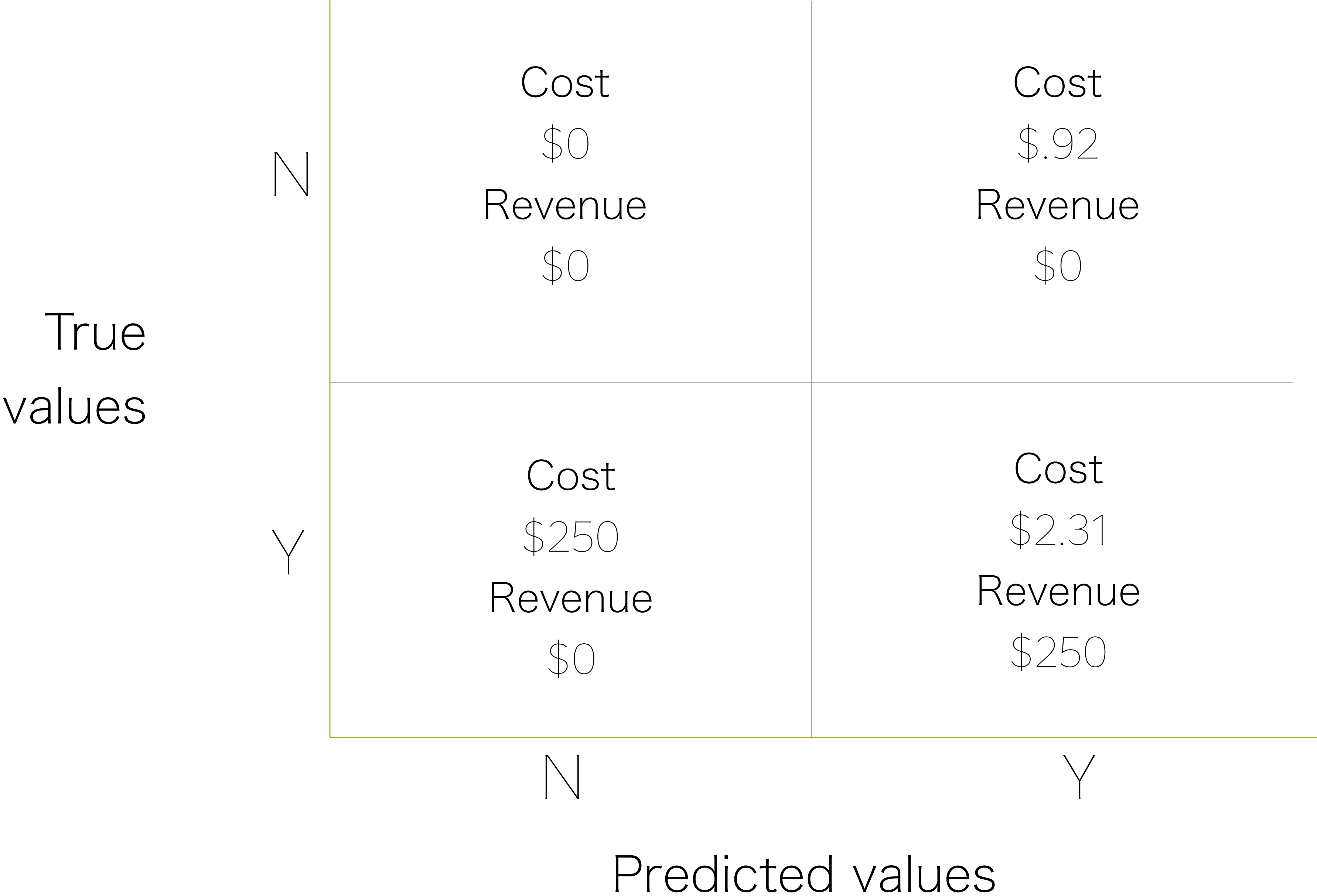
Mean call length:
9.220 mins

Mean cost per call:
\$2.31

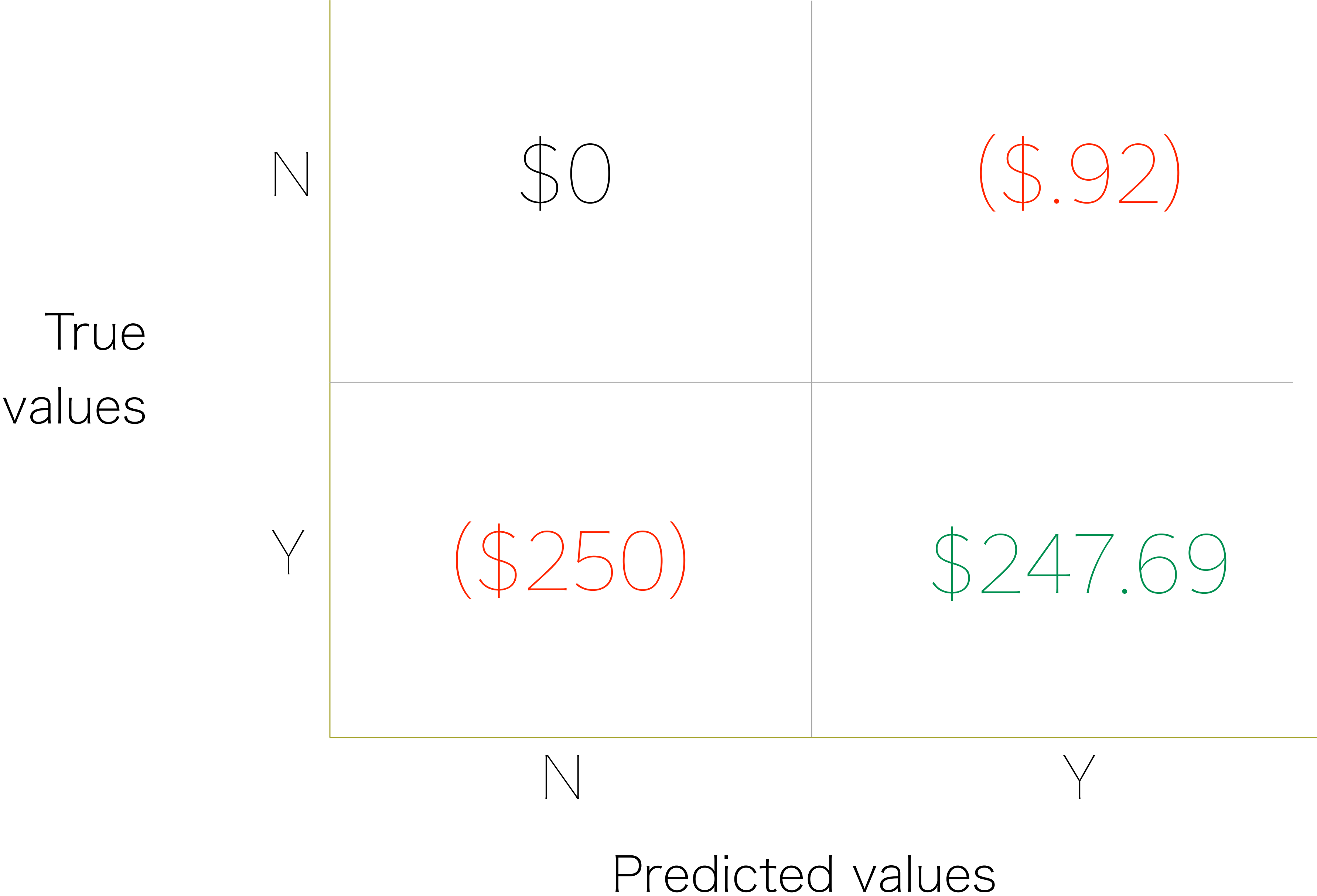
Revenue



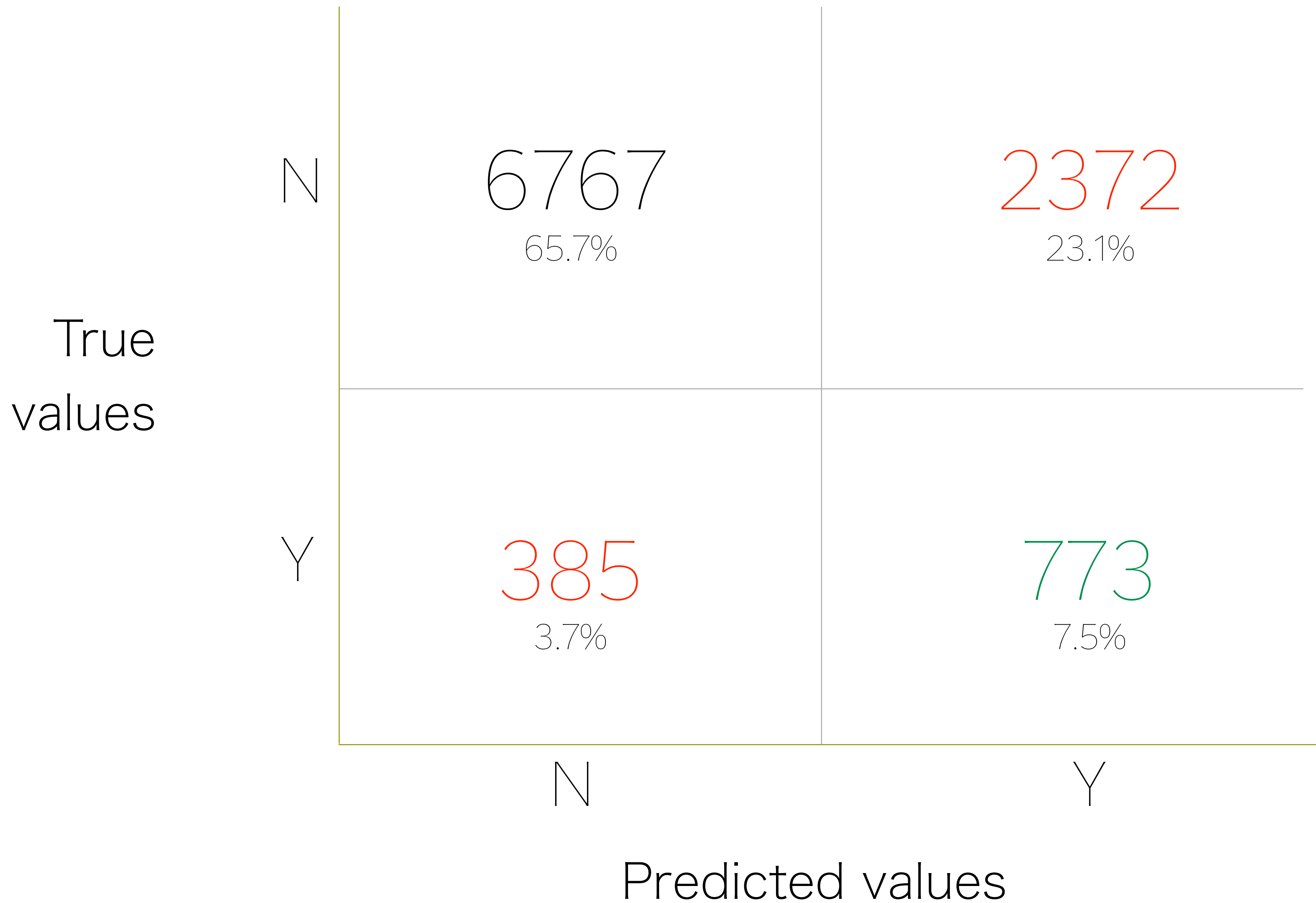
Y = subscribed
N = didn't subscribe



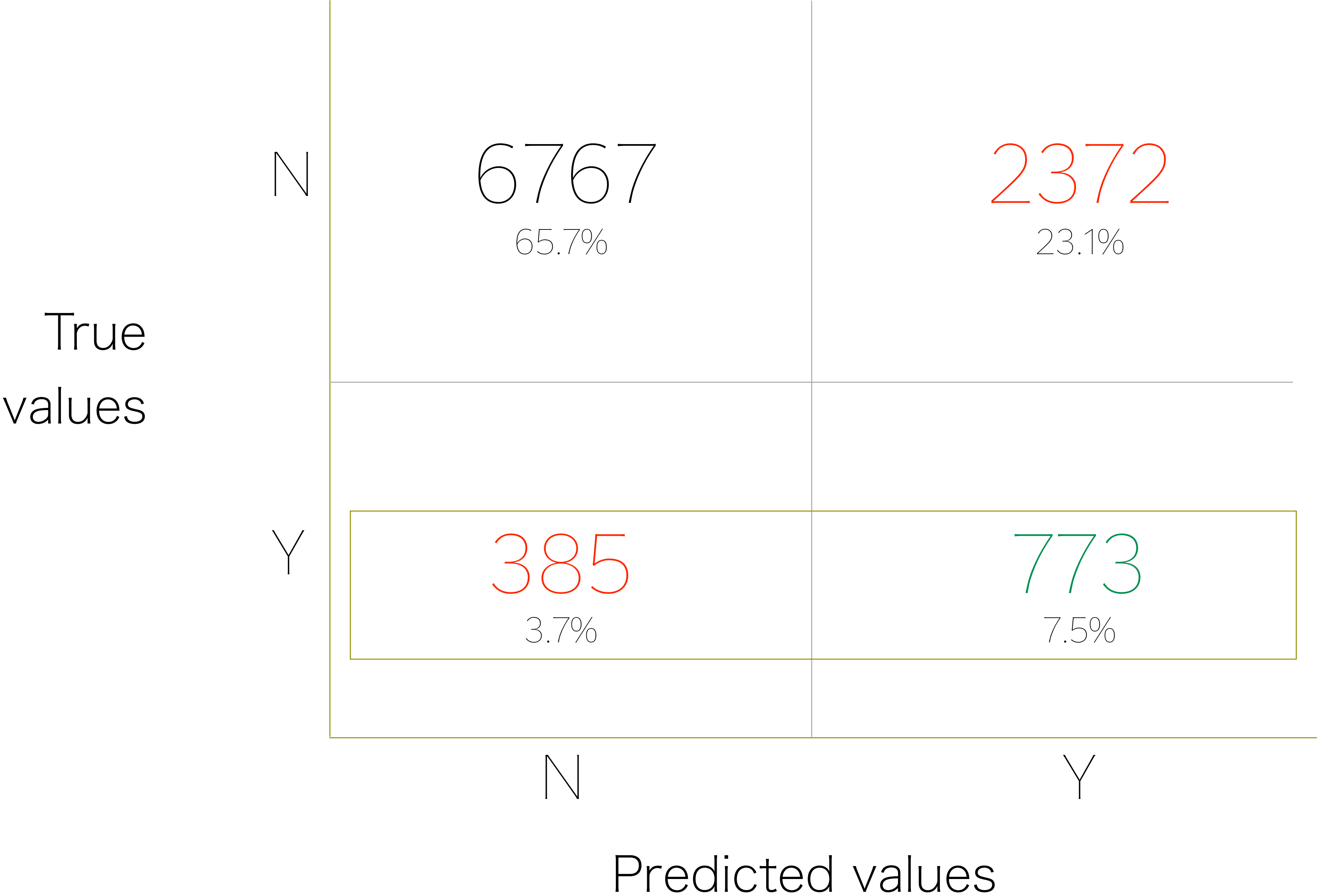
Y = subscribed
N = didn't subscribe



Y = subscribed
N = didn't subscribe



Y = subscribed
N = didn't subscribe

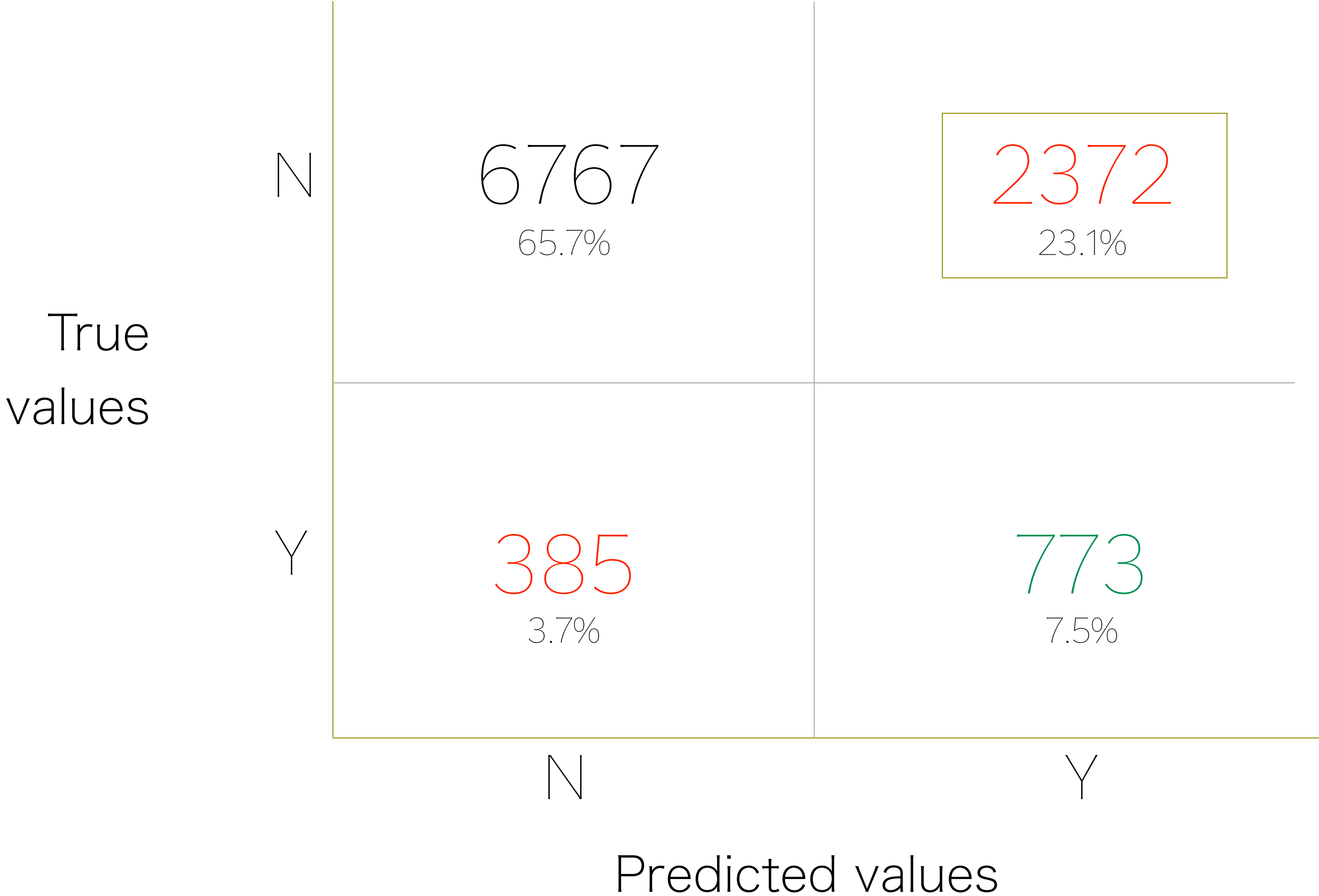


Designed to
maximize recall:

$$\frac{773}{385 + 773}$$

↓
66.8%

Y = subscribed
N = didn't subscribe



Okay with 'false positives' - costs are minimal.

What is the earning potential of a list with
100,000 new clients?

Y = subscribed
N = didn't subscribe

True values	N	65,700 people x \$0 = \$0	23,100 people x (\$.92) = (\$21,252)
	Y	3,700 people x (\$250) = (\$925,000)	7,500 people x \$247.69 = \$1,857,675
		N	Y
		Predicted values	

Y = subscribed
N = didn't subscribe

True
values

N	65,700 people x \$0 = \$0	23,100 people x (\$0.92) = (\$21,252)
	3,700 people x (\$250) = (\$925,000)	7,500 people x \$247.69 = \$1,857,675
Y	N	Y

Predicted values

\$1,857,675

-

\$925,000

-

\$21,252

=

\$911,423

But what if we just . . .
called all 100,000 people?

Y = subscribed
N = didn't subscribe

True values		N	<div>0 people x \$0 = \$0</div>	<div>88,800 people x (\$0.92) = (\$81,696)</div>
		Y	<div>0 people x (\$250) = (\$0)</div>	<div>11,200 people x \$247.69 = \$2,774,128</div>
			N	Y
		Predicted values		

Y = subscribed
N = didn't subscribe

True values	N	0 people x \$0 = \$0	88,800 people x (\$0.92) = (\$81,696)
	Y	0 people x (\$250) = (\$0)	11,200 people x \$247.69 = \$2,774,128
		N	Y
Predicted values			

\$2,774,128

-

\$81,696

=

\$2,692,432

Then why is this model useful?

In The Future

- Marketing campaigns are not always this inexpensive.
- Potential revenue will not always be this high.
- We might not have the time to call every single person.
- Generalizable - has the potential to be used to detect behaviors with a similar distribution.

Future Work

- Engineer new features to better identify likely investors.
- Continue to tune the model via hyperparameters.
- Calculate minimum cost:revenue ratio needed to run a viable campaign.
- Identify which financial products each customer is most likely to invest in.



Questions?

Thank You.

