Churn Down For What?

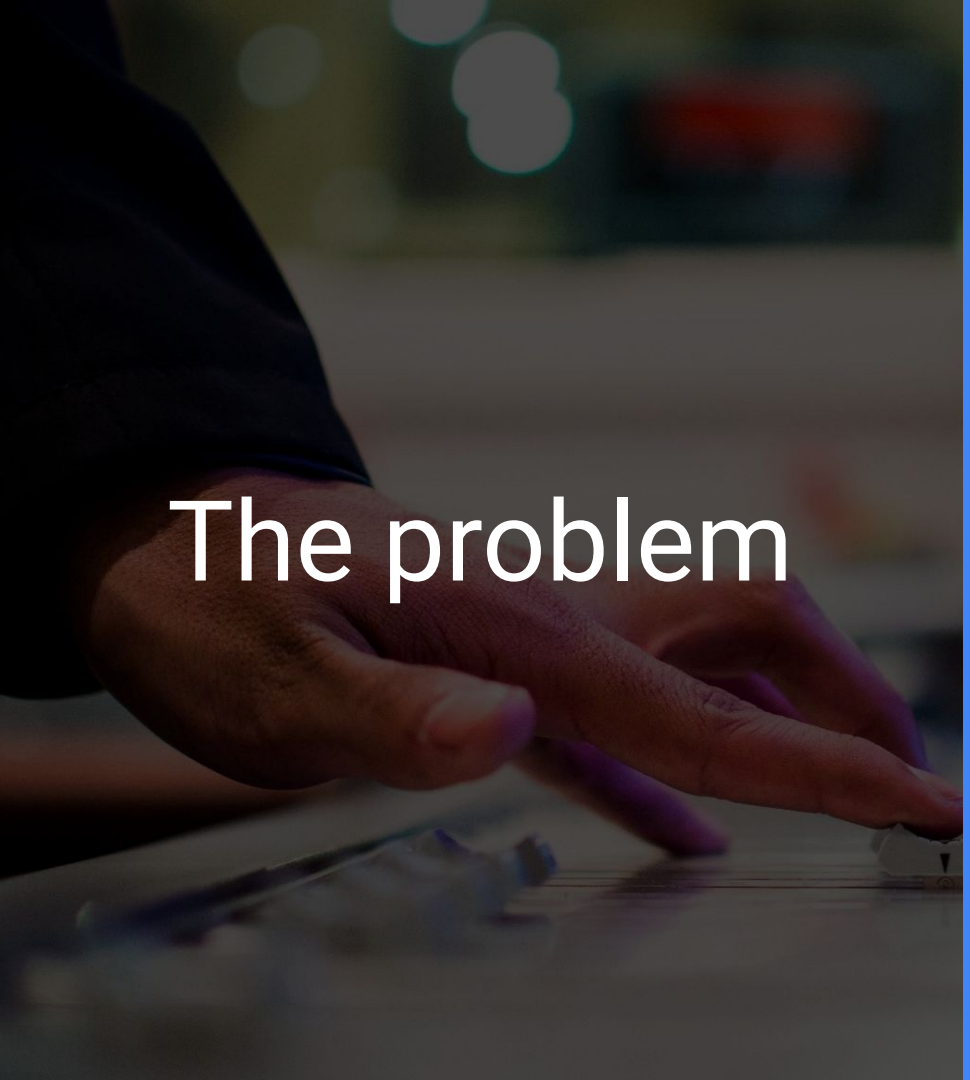
Your Own Good, Actually

By - Whitlee Pearson

Overview

- Customer churn is an ever-growing issue in telecommunications and other industries.

- It is more expensive to lure in new customers than to keep existing ones.



The problem

The Classification Station was asked to create a model able to predict whether or not a customer would “soon” stop doing business with (Client Redacted).

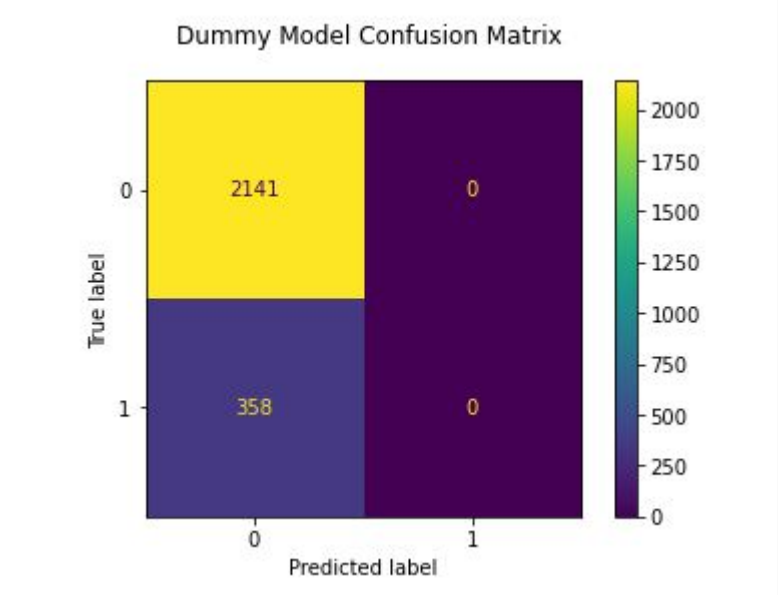
Data Understanding & Preparation

- Data from CrowdAnalytix
- 20 predictors & 3333 datapoints
- Of these, only 14% were recorded as having “churned”, so our data was imbalanced
- Quite common with churn data

- Removed unnecessary columns
- Standardized data
- Prepped for modeling with recall/accuracy in mind
- Why?

Dummy Model

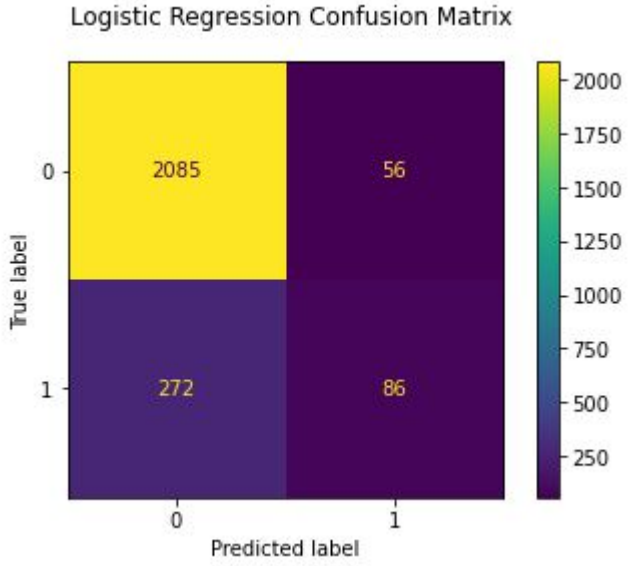
The first attempt



- Always predicts most common response
- 0% recall
- 85% accuracy

Logistic Regression

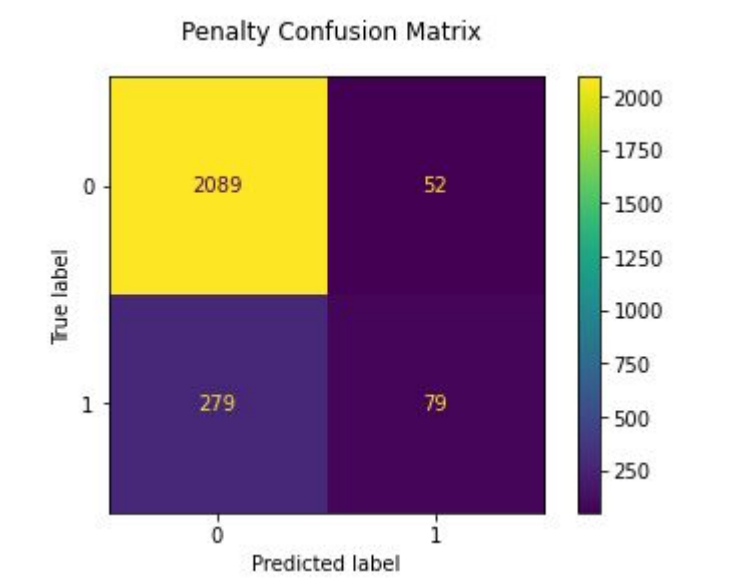
More attempts



- 21% recall
- 86% accuracy

Logistic Regression

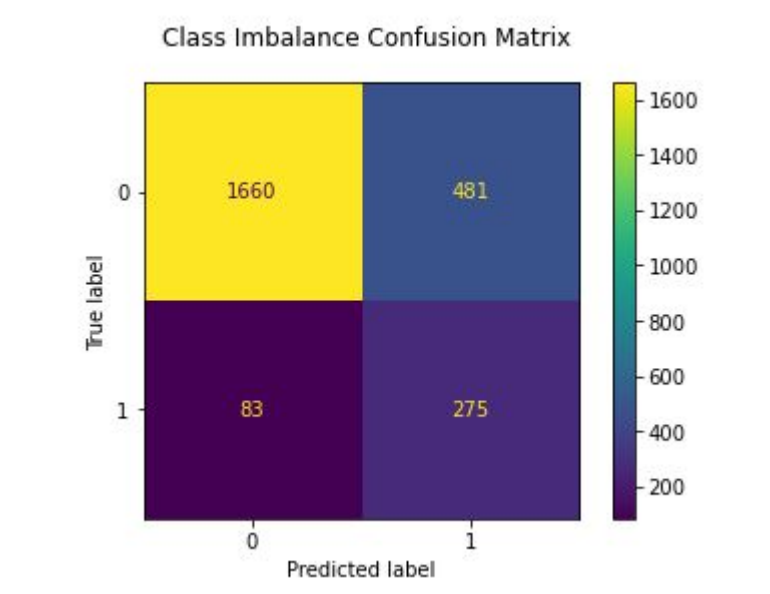
With penalty



- 20% recall
- 86% accuracy

Logistic Regression

With penalty and class imbalance

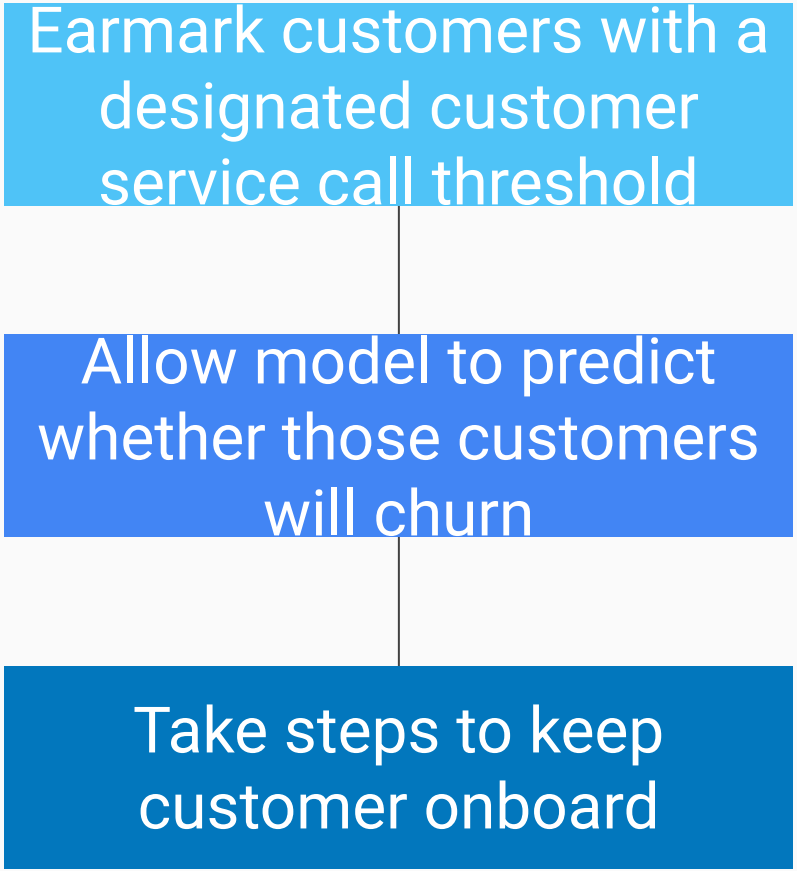


- 73% recall
- 77% accuracy

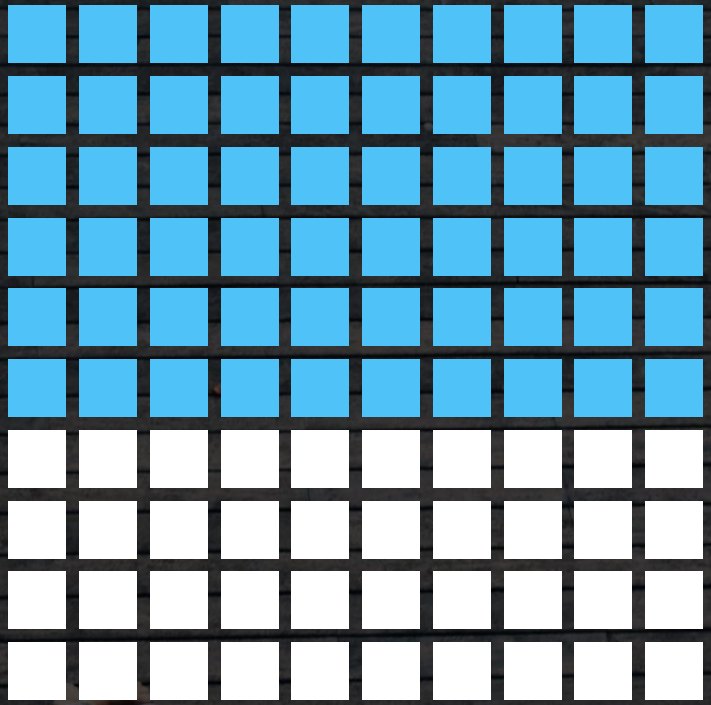
Conclusion

- Though our final model was not our most accurate, it did have the highest recall we could facilitate.
- Recall was our most important metric because it measures, in this context, which customers actually churned versus predicted churns.

Next Steps


Earmark customers with a designated customer service call threshold
Allow model to predict whether those customers will churn
Take steps to keep customer onboard

Thank you!



Whitlee Pearson
PWhitlee.2013@gmail.com
GitHub - falloutbtch