

DAT-203 Final Project Proposal

Predicting Customer Churn

Colin Bowers

June 30, 2023

Customer churn, or the loss of customers over a specific period, is a critical business problem with significant implications for companies. I intend to analyse existing data on customers that have churned in order to accomplish a few objectives. First, to learn more about customer behaviour and other insights that would be useful for the business in general. Second, to develop a classification model that would be able to identify any existing customers that would be at risk of leaving. Third, provide some interpretations and recommendations that would be useful for account managers and marketing teams to optimize their customer retention strategy.

This depends on a sufficient dataset in a domain that would be most relevant to my own line of work. So far I have discovered 3 that could be used (Ecommerce and a couple from Telecom)

- <https://www.kaggle.com/datasets/ankitverma2010/ecommerce-customer-churn-analysis-and-prediction>
- <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>
- <https://www.kaggle.com/datasets/mnassrib/telecom-churn-datasets>

Notes:

- The previous course I took (DAT 201) was all on linear regression so I wanted to focus more on classification or clustering for this project.
- I'm still considering other problems that are more relevant to my line of work (Telematics, GPS and trips data, engine status data, etc.). However, I haven't been able to find a usable dataset that I can also share publicly. (And advice here would be more welcome!)