Netflix's $1,000,000 Model and Customer Retention at Grocery Consolidators

Lucas Wu, Dani Chu, Nikola Surjanovic - SFU Sports Analytics Club

Taking Action on our Predictions (What Next?)

- Our predictions are the 2nd best on the leaderboard
- Can predict a lost customer 84% of the time on new data
- We can leverage our accurate predictions to establish:
  - High-value customers we are at risk of losing
  - Expected revenue losses over a fixed time frame
    - Allows us to allocate resources wisely

Understanding our Predictions

- The high indicators of loss are: Email Open Rate, Email Click Through Rate, Order Frequency and Average Order Size
  - Describe the customer, tougher to take action on
- City CHO has a higher loss rate and BWI has a low rate
  - Adjust the setup of the program in this CHO, compare with BWI
- Encourage sign up for paperless communication for the environment
- Favourite day of Sunday has a high loss rate
  - Adjust Sunday pick up strategy
- Surprisingly, doorstep delivery and automatic refill were the least important for predicting a lost customer when all other variables are considered.
  - Still encourage doorstep delivery and automatic refill

Customer Retention Decisions

Below are the expected revenue losses over a fixed time frame of one month (**30 days**), for the five customers which have the highest expected loss.

Customer ID	Score	Average Order	Order Frequency	Value	Expected Loss
NQQTH5	0.298051	129.34	2	258.68	$2312.997
AVEDW8	0.791003	262.53	0.285714	73.51	$1779.959
V97YHB	0.57046	141.33	0.6	84.80	$1451.217
RLARBF	0.474267	85.39	1.171429	99.90	$1423.202
SSRWWF	0.19929	69.48	3.25	225.81	$1350.048

High Value Customer Surveillance

- Value of a customer depends on their order frequency and sizes of their orders, relative to our other customers
- Customers in the upper quartile of value are considered high value customers
- With our predictions we can decide how much to spend in retention for a particular customer in a given month.