CMU Giving
Unlocking Generosity with Analytics

Group Members: Roopa Bharadwaj, Yang Liu, Mark Denton
Sponsors: Julie Knight, Dawn DiBartolo
Advisors: Joy Lu, Angela Xiao

Carnegie Mellon University
Tepper School of Business

Introduction

Growing a sustainable giving program is a key strategy for CMU. But finding and converting these donors to be regular or upgrade their donations isn't easy.

CMU has been growing the donation base of alumni that participate in annual giving, volunteering or other events. Appeals are a cornerstone for fundraising programs. But running them inefficiently by targeting wrong alumni with wrong ask amounts is often a pain point.

Research Goal

Predictive modelling to drive increased donations –

- Optimizing ask ladder for guiding a donor's next gift
- Better segmentation of donors to enhance communicate with them
- Identifying new potential donors
- Predicting donors Vs. non-donors
- Predicting who would donate in 365 days from their last donation.

Data and Methods

Our analysis covered two datasets: a biographic dataset and a donation dataset.

- Biographic dataset: 125,659 anonymized alumni records with demographic data, macro historical donation information, industry standard donor classification, event attendance history, and solicitation preferences
- Donation dataset: data for 109,078 donations made over the past 5 years including date, amount, method of donation, and appeal code

Both datasets included contact IDs for the involved alumnus so that donation data could be cross referenced against biographical data.

This data was then analyzed using methods including k-means classification, logistic regression, XGBoost analysis, and random forest classification in order to gain a quantitative understanding of the questions laid out in our research goals.

Additionally, an A/B test was conducted wherein potential donors were shown differing ask ladders. The data obtained from this test was then analyzed to better understand the impact of ask ladders on donation behavior.

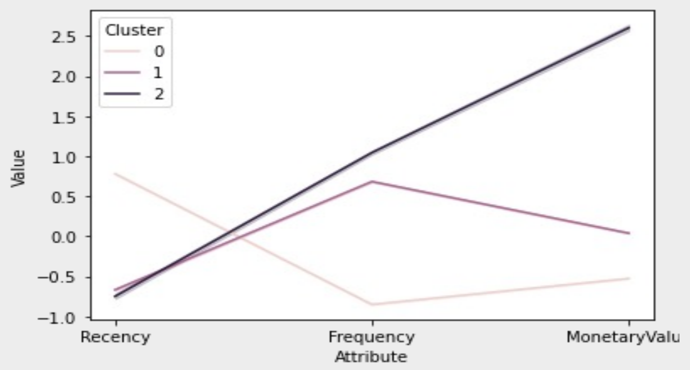
Models

Segmentation

Donors who give regularly have different characteristics and have their own profile. Because we can't target or reach out to every donor the same way, we implemented a segmentation method based on -

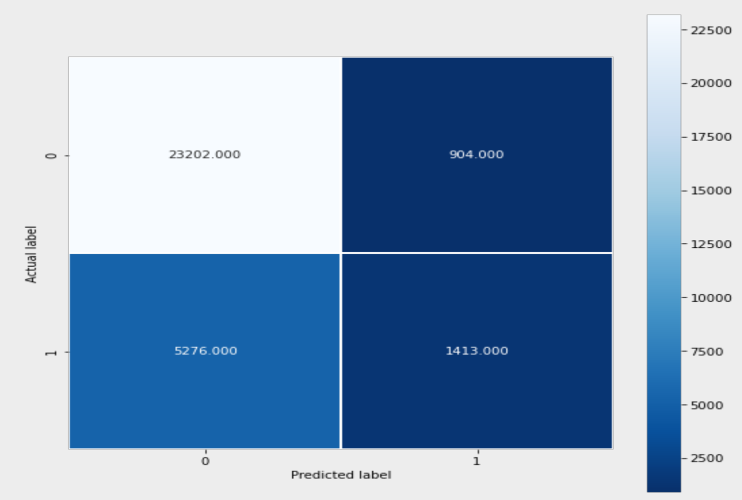
- Recency of Donation – based on most recent donation date and how many days they were inactive for.
- Frequency of Donation – total number of donations for each alumni
- Monetary value – total revenue from each alumni

Using K-means clustering, we segmented donors between 3 clusters.



Cluster	Category	Count	Recency	Frequency	Monetary Value
0	Risk to Churn	46%	995	1	175
1	Loyal Donor	45%	242	5	629
2	Premium Donor	9%	246	8	13294

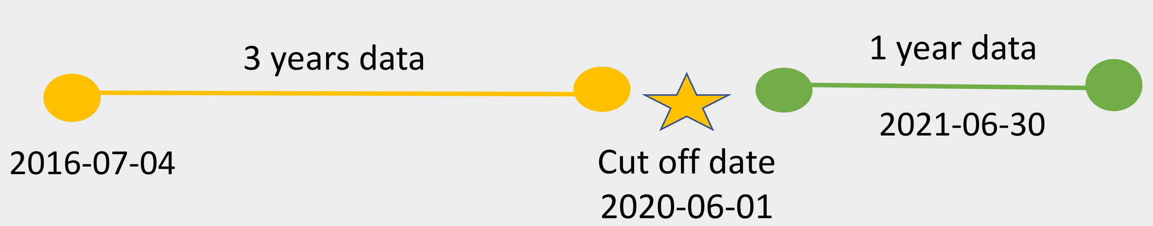
Who is a donor?



- Predicting whether alumni will appear in the previous 5 years' donation data based on demographic information from biographic dataset
- Individual alumni put into test/train datasets with a 75/25 split
- 79.9% accuracy
- Event attendance, class year, school within university, and whether receiving solicitations were key prediction factors

Predicting donations in the next year

To build our model, we split the data into two parts:

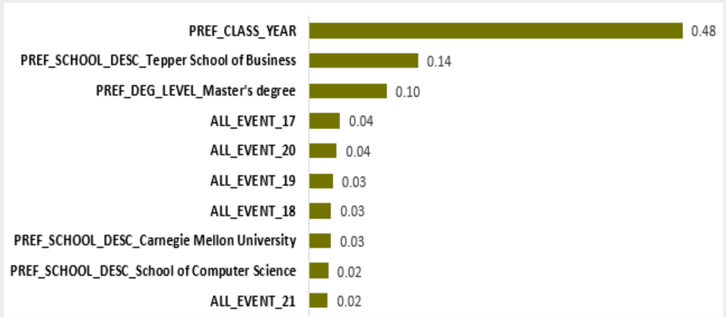


Model	Accuracy	F1 score	Recall	Precision
Random Forest	84.7%	69.05	67.01	75.9
Logistic Regression	84.2%	40.4	29.7	63.1
XG Boost	**87.2%**	**63.2**	**60.93**	**65.7**
SVC	83.9%	63.0	62.7	67.1
Decision Tree	83.8%	64.9	64.7	68.2
KNN	83.5%	65.8	65.3	68.8

- We used the first 3 year's historical data to predict donor's first donation in the next one year.
- Recency of donation, Frequency and Monetary value of donations from each donor were used as features. Among all the prediction models, XG Boost Classifier gave the highest accuracy in predicting who would donate next year.

Identifying potential donors

	Logistic Regression	Random Forest
Train Accuracy	60%	71%
Test Accuracy	62%	69%
Recall	25%	63%
Precision	71%	68%
AUC	67%	68%

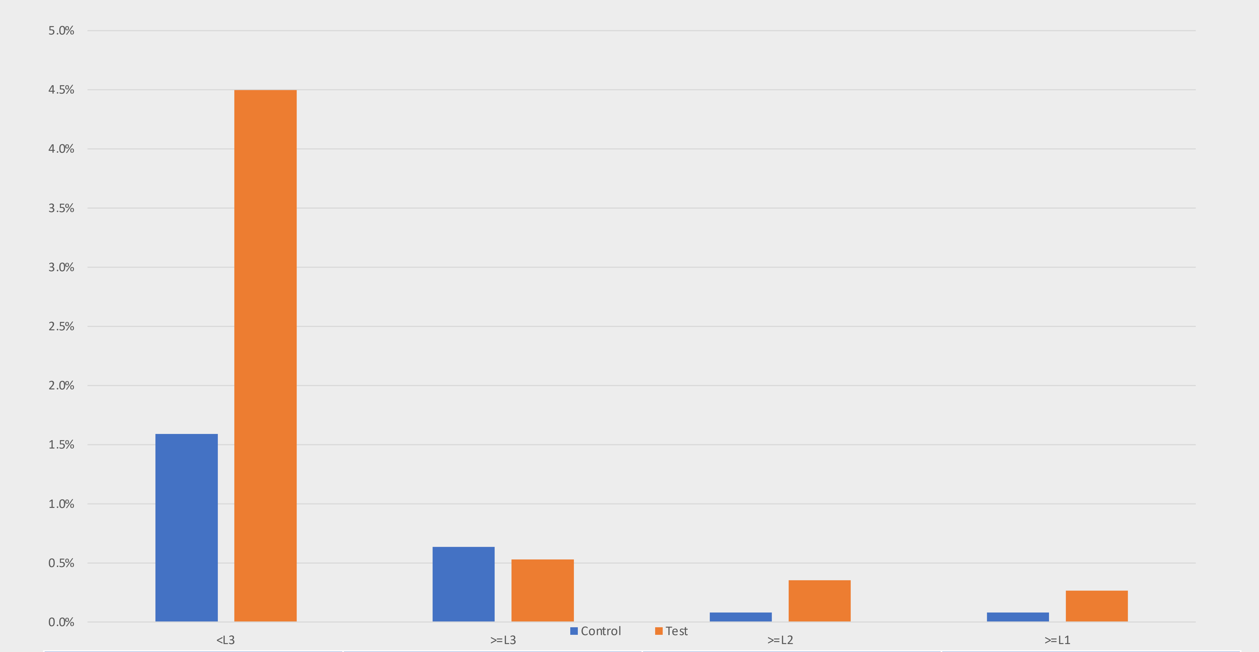


Predicting first-time donor and the probability of being a donor by using the biographic information provided. Random Forest model give better result on accuracy at 69%. Feature importance allows us to rank the predictors contributed the most.

- As expected, Alumnus class year dominates with over most of the overall importance.
- Following are event attendance, specific schools, gender are the also significant indicators for predicting donors.

By assigning a donation probability score to each Alumnus which would be helpful for sorting the best contact when running future campaigns.

A/B Test



Group	Participants	Donors	Conversion Rate
Control	1259	30	2.4%
Test	1134	64	5.6%

We conducted an A/B test to examine how being presented with different ask ladders would impact donation behavior.

Donors were shown ask ladders with three possible donation amounts based on past giving behavior.

- Donors in the control group were shown the existing ask ladder for their tier of giving.
- Donors in the test group were shown an *increased ask ladder* with the higher ask amounts than their tier of giving.

The results of the A/B test were significant with >95% confidence that the test group were more likely to donate and more likely to donate at the higher ask levels of their tier.

Results & Findings

The various models provide an outline for finding potential donors. Collectively, they identify segments for targeted solicitation, individuals who are likely to donate, non-donors who have the potential to become donors, and which donors are likely to donate in the near term.

Additionally, the A/B test produced interesting and somewhat counterintuitive results suggesting individuals may be more likely to donate when asked for larger amounts. This result may be worth further investigation and testing.

Acknowledgements

We would like to thank our sponsors – Julie Knight and Dawn DiBartolo for all their help in getting additional data we needed and for being so flexible with their schedules.
We are also thankful to our advisors – Prof. Joy Lu & Angela Xiao for their help and support throughout the project. Prof. Lu's ideas and directions helped us a lot and gave a proper direction to our project. We would also like to thank Angela Xiao for her constructive feedback.