Gary Chen

Charles Cheung

Matt Maloney

IST 664 – final report

On January 1st, 2013, Syracuse University’s Division of Advancement & External Affairs officially launched its *Forever Orange Campaign*. The goals of the campaign were broad and included a lofty aspiration of $1.5 billion in newly raised commitments over the ten-year period. These dollars have been raised through numerous channels: direct marketing efforts like print appeals, telephone calls from current students to alumni, and digital fundraisings strategies. Perhaps most significantly, the division relies heavily on the work of its fundraisers, also called gift officers. These staff members travel the country and the world meeting with alumni, parents, and other affiliated constituents to discuss their interests and their philanthropic priorities.

Our analysis in this project centers on the contact reports (written narratives) filed by gift officers after meeting with constituents. The data frame contains 32,950 of these contact reports, along with categorical and numerical attributes from the CITRUS fundraising database. These fields include the method of the contact (on-campus, off-campus, online), the category (the phase of the relationship), campaign commitment (funds pledged to the University by the constituent since January 2013), and the best rating (an internal calculation that estimates a constituent’s maximum philanthropic commitment). All interactions have occurred since January 2013 and summarize conversations with about 11,700 constituents. They have been authored by about 260 fundraisers.

To properly begin analyzing this data set, it was important to complete several pre-processing steps. After importing the data, we established several categories to group the giving and rating attributes. Additionally, we removed the stop words to better evaluate the content of the contact reports. Because some of the older contact reports included categorical strings like “Research Needs:,” “Professional Experience:,” and “Wealth/Assets:” within the summary, we also needed to remove instances of these word and punctuation combinations.

Once our data was in a suitable condition, we were able to embark on our natural language processing analysis. Our first investigation was to simply look at the frequency distribution for the entire data set, which we quickly found to feature many words that ought to have been stopped. After applying the additional stop words (as detail in the previous paragraph), we saw frequent words such as ‘su’, ‘gift’, ’time’, ‘met’, ‘like’, ‘students’, ‘alumni’, ‘program’, ‘dean’, and ‘interested’. These key word distributions highlight the types of conversations that fundraisers were most commonly having with donors.

Next, we utilized the Penn Treebank to identify the parts of speech employed within the corpora. Among adjective phrases, the most frequently employed by fundraisers included “very interested,” “not sure,” “very happy,” “very excited,” and “very pleased,” among others. The list of most frequently utilized adverb phrases contained “as well,” “very well,” “right now,” “very much,” “not yet,” and “still very.” Among the most common adjectives were words like “new,” “next,” “interested,” “other,” and “annual.” Words like “not,” “also,” “very,” “well,” “back,” and “now” were the most frequently written adverbs. Among verbs, “is,” “was,” “be,” “has,” “had,” and “have” topped the charts. The most used nouns were “gift,” “time,” “SU,” “students,” “year,” “school,” “years,” “program,” and “meeting.” Most of these entries are reasonable results to see, with so many relating to fundraising. In this section we also calculated the average length of a sentence, which was just over 159 characters.

Another investigative route for our group was to study the polarity of the comments. Drawing on the categorical buckets we had established in the pre-processing phase, we were able to study differences between the sentiments across several cross-sections of the data. Our real word data set did not have any tags rating the sentiment or scoring the comments from each donor interaction. We set up a process to analyze the sentiment, labeling the comments as positive, negative, or neutral. We imported the VADER lexicon which is a model used for text sentiment analysis and measures the positive and negative polarity and intensity of emotions within the text. The VADER lexicon has 7,500 sentiment features and words to help us calculate the polarity scores on our comments. Depending on the words in each comment, positive and negative words are scored according to the VADER lexicon and any word not listed in the lexicon will be scored as 0 (neutral). We used the SentimentIntensityAnalyzer() function to calculate the polarity scores on the lemmatized comments for each individual donor interaction. This output generated a negativity value, a neutrality value, a positivity value, and a compound value of the previous 3 fields combined. Utilizing that compound value, we tagged the interaction as positive if the compound score was greater than 0, negative if the compound score is less than 0, or neutral if the compound score is equal to 0.

Our results scored 88% of the comments as positive, 10% neutral, and 3% as negative. With such an overwhelming amount of positive tagged comments, we wanted to do further analysis of the average scores of the negativity, positivity, neutrality values group by our 3 different (‘Campaign Bin’, ‘Rating Bin’, & ‘Major Giving Bin’) bins. For the Campaign Giving bins, we saw an increase in the average positivity scores from giving amounts of Low to 100k. At the 1M bin, we saw a drop in the average positivity score. It is hard to explain why there was a drop off in average positivity scores at the 1M giving bin when it was trending up for the smaller bins. When we analyzed the Major Giving bins of low and high, we saw that average positivity score for high was much higher than low. This Major Giving bin was broken up by high for donations of more than $100k and low for donations of less than $100k. For the Ratings Bin analysis, the higher the rating bin, the higher the average positivity scores. The Rating bins are based on an internal metric to predict how much money a donor could give to the university. The interaction comments from the gift officers were probably more positive given that they expected donors with higher wealth indicators to give more money to the university.

To see which words were most impactful in determining the positive and negative scores, we used the Naïve Bayes classifier to train the data in finding the keyword features. It is not surprising that most of the informative features are those that detect a positive sentiment. Our comments are tagged as mostly positive, but we can see feature words like ‘thanks’, ‘amount’, ‘director’, ‘connections’, and ‘enjoyed’ had the most impact in rating comments as more positive. Predictably, gift officers choose to focus on those aspects of a conversation that serve their goals – including amounts of potential future gifts and a donor’s overall happiness with their alma mater. The negative feature words are ‘rob’ and ‘dick,’ which were likely utilized as nicknames for Robert and Richard, and may have been mis-classified by the algorithm. This shows some gaps in how our comments were interpreted in our sentiment polarity analysis. The VADER lexicon did provide some great insights, but more needs to be done to determine the full sentiment behind the interaction comments.

Lastly, our group completed several Naïve Bayes models to study the most informative text features of the data set. During this process, we created a model that predicted the overall sentiment of the contact report (positive or negative) as well as a model predicting whether a constituent had contributed at least $100,000 during the campaign period. ***Findings:***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Feature Function** | **Feature Set** | **Data Size (items)** | **Classification** | **Accuracy** |
| BOW/unigram | Comment, MajorGiving | 32950 | Low, High | 0.745 |
| POS tagger (+cardinal tag) | Comment, MajorGiving | 32950 | Low, High | 0.728 |
| BOW/unigram | Comment, Sentiment | 23756 | Positive, Negative | 0.663 |

Through our analysis, we have yet to find evidence of a relationship between certain text features of a fundraiser’s contact reports and a constituent’s giving. If we were to continue our analytical efforts, we might make some changes to our data set to improve the results. Could we separate the contact reports based on the category to glean more about the likelihood that a constituent will contribute? Or perhaps even more compelling, could we create separate data sets that restrict contact reports to those that come before a constituent’s largest gift? This might allow us to say something more definitively about the differences between a relationship that will bear fruit (an eventual major gift) and those that will not. In any case, our group’s efforts provided some interesting results, demonstrating the functionality of the frequency distributions, part of speech tagging, polarity scoring, and predictive modeling.

**What we worked on:**

Matt – Curated and imported the dataset. Created bins and categorical buckets. Completed part of speech tagging.

Gary – Tokenized & Lemmatized comment attributes from original data set. Created the word frequency distributions and calculated the sentiment polarity scores.

Charles – Created features with Naïve Bayes Algorithm. Ran different experiments and compared results.