Paul Mack, Alden Golab, Sam Sun
Machine Learning for Public Policy Project Proposal

seismos

Predicting Not-for-Profit Fiscal Health

In January 2012, one of Chicago's largest and oldest not-for-profit (NFP) agencies closed suddenly. Jane Addams Hull House Association, a 122 year old charity serving 60,000 residents annually, closed over 40 sites in Chicago within one week of announcing its plan to file for bankruptcy. The Department of Family and Social Services, which funded Head Start programs through Hull House, had less than one week to react to Jane Addam's Hull House center's closing of childcare facilities.

Many local governments rely heavily on NFP entities like Hull House to provide social services to their most disadvantaged constituents. Problematically, it can be difficult to predict those agencies' financial issues; indeed, there are so many agencies, reviewing records for all social service NFPs in order to assess risk is a herculean task. Foundations, too, face this dilemma: the larger the foundation, the more organizations to monitor. As such, when social service NFPs have financial difficulties---putting them at risk for closure---policy makers and foundation managers may not be able to intervene in time to prevent gaps in essential services.

The effects of these gaps can be extreme. In the case of Hull House, there was simply insufficient capacity among other Chicago NFPs to take-in such a large influx of cases without warning. Lacking services like child education, job training, and housing assistance, thousands of individuals served by Hull House simply lost a lifeline toward a brighter future.

Seismos seeks to develop risk scores for declining non-for-profit financial performance: that is, to predict the stability of NFP firms for the year ahead. There are several use cases for this type of risk factor:

- Assist Foundations in monitoring grant recipients' financial health, allowing them to provide guidance and support when appropriate.
- Give government agencies an early warning system for service gaps so that they can either intervene or find alternative service agencies.
- Allow NFP service agencies to investigate local NFP risks so that they can better prepare themselves in the event of a closure to a related NFP or to find NFPs that may be willing to merge operations.

Through this work, we can prioritize oversight and direct needed managerial attention or additional funds to those NFPs most at risk---regardless of whether they're aware or have asked for help themselves--preventing the human toll that would result from service providers' poor health or closure.

The Data

Currently, we have Guidestar 990 tax information for the universe of US NFPs in 2013 with over $200,000 in revenue or $500,000 in assets---about 988,000 organizations in total. These files include all 990 schedules and almost all columns, including general items like gross revenues, all the way down to CEO pay and Board Members. Out of these, we plan to develop a wide variety of features, including variables like percent of revenue from government grants and Board Member stability.

To best predict non-profit performance, we need additional years of longitudinal data: preferably at least 2012 and 2014. There is a more limited dataset available publically from the IRS, however, these files lack any character fields--like addresses and salaries--which make them not ideal for this application.

The National Center for Charitable Statistics has longitudinal data going back to 1989 for 60 financial variables for nonprofits that have filed 990 forms.

Regardless of which dataset we are able to secure, there are worries with tax data, primarily around its accuracy: many small organizations do not fill out their IRS forms correctly, lacking professional staffing to do so. We don't anticipate this being an issue, however, we will be looking through our dataset to diagnose errors like gaps in filing (e.g. an organization is missing but did not go under) or reporting the same revenue numbers year over year.

As such, we are also exploring use of the DataArts dataset, which has very clean, reliable, in-depth profiles of approximately 10,000 arts-based NFP organizations specifically including information like program participation.

The Analysis

We view NFPs as suppliers of collective goods (services). Donors (individuals, foundations, government) contribute money in return for an implicitly agreed-upon level of provision. The price of the collective goods is the amount of contribution needed by the NFP per unit of output. The value of the collective goods is determined by the valuation of the donor as well as the quality of the output the NFP provides. Donors are willing to contribute to the extent where the value equals the cost. With this view of NFPs' contribution/revenue generation mechanism, we first define relevant features:

Phase One: Feature Engineering

The label: our main variable of interest is the revenue change of NFPs, which can take several forms:

- Continuous: revenue change (%)
- Discrete: revenue change (up-stable-down or other categorization system)
- Source-specific: change in private contributions v.s. government grants

Eventually we will test on all the features available in the IRS tax document, which can be broadly classified into two group:

- Organization-level features
  - Lagged revenue change (control variable)
  - Price of NFP goods (various efficiency measures such as contribution/expenditure ratio, administration expense (%), CEO compensations (%), etc.)
  - Value of NFP goods (reputation--age, advertising, etc.; societal valuation--NTEE sector, mission statements. etc.; governance--transparency, organization structure, etc.)
- Market-level features
  - Competition (number of similar NFPs within MSA, etc.)
  - Macroeconomic environment (GDP per capita of the MSA, industry-wide performance, etc.)

Phase Two: Modelling

We begin by modeling a simple discrete model to understand the potential for predicting any kind of revenue change. Our primary worry is that revenue change is best predicted by last year's YOY change: we may very well find this to be correct. Using this as a baseline comparison, our goal is to improve our precision over a simple naive predictor based on last year's change.

1) Baseline Model:
   - Previous year's revenue change
2) Discrete Model:
   - Binary label of revenue change
   - Multi-class label of revenue change
3) Continuous Model:
   - Percentage change of revenue compared to the previous year

We will test on a group of classifiers: Logit, RFC, SVM, DT, NN, GB, Bayes, Ensembles, etc. As we close around those methods which most precisely estimate financial risk, we will employ ensemble methods to see if additional precision can be gained.

Phase Three: Evaluation

Our purpose is to develop an early warning system of financial risks of NFPs. For this purpose, we will evaluate our classifiers based on the following beliefs:

1) The precision of our prediction is more important to its recall capacity
2) We need to compare short-term prediction precision (1 year) vs long-term prediction precision (3-5 year)
3) We will use a temporal validation method to test real world applicability

Outcomes and Policy Impact

Based on the evaluation results we will construct a set of classifiers that can best identify the most financially risky NFPs. And the corresponding predicted probability/confidence statistics will be used to construct a Financial Risk Index (FRI) as our final deliverable. This index will likely be useful for 1) individuals/foundations to make more informed contributions; 2) government agencies to allocate resources and provide interventions; 3) NFPs to understand their financial health.

In this way, Seismos can reduce the risk of gaps in social service coverage due to poor NFP fiscal health, helping foundations and governments to target their limited resources to intervene on behalf of those organizations most at risk, limiting the human impacts of NFP turnover.