seismos *Predicting NFP Stability*

Alden Golab, Paul Mack, Sam Sun
University of Chicago, Machine Learning for Public Policy

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.[1] The Department of Family and Social Services, which funded Head Start programs through Hull House, had less than one week to react to the Hull House center's closing of childcare facilities. Other area nonprofits such as the Metropolitan Family Services and Uhlich Children's Advantage Network hurried to take over serving some of Hull House's clients. 300 of Hull House's employees needed to seek new work.

Clients, government, foundations, individual donors, other area non-profits, and the agency's employees all are stakeholders in the fiscal health of a non-profit. Monitoring financial trouble in this sector is a challenge as there are about 1.5 million nonprofits in the country. Stephen Saunders, the leader of the Hull House board in its final year, described the difficulty of understanding the gravity of the financial situation even internally as the board came to realize, "previous financial reports, often arriving late, had sugar-coated the situation."[2]

When non-profits delay communicating their financial trouble either because they themselves do not realize its seriousness or are reluctant to communicate it, their options become limited. Terry Mazany, the CEO of the Chicago Community Trust, a community foundation, makes it clear this is not an issue unique to Hull House, describing how many nonprofits are "receiving less government support and are struggling to survive. Some waited too long before reaching out for help."[3]

If non-profits most in need of help can be identified earlier, stakeholders may be able to coordinate to improve outcomes and prevent closure. Realizing the gravity of the situation, individual donors who support the mission of the non-profit may be convinced to give more. Organizations such as the Chicago Community Trust may be able to work with the non-profit to make operational changes to reduce costs. Nonprofits could use findings to communicate to government funders the importance of their support, demonstrating their risk in the absence of government intervention. In some cases non-profits may be able to merge with another

---

[1] Thayer, Kate. "Hull House Closing Friday". *The Chicago Tribune.* January 25, 2012. http://articles.chicagotribune.com/2012-01-25/news/ct-met-hull-house-20120126_1_child-care-union-contract-employees Accessed 6/4/2016

[2] West, Maureen. "Some Fear Hull House Closure Is an Omen for Struggling Charities." *Chronicle of Philanthropy*, February 02, 2012. https://philanthropy.com/article/Collapse-of-Famous-Hull-House/157181. Accessed: 6/4/2016.

[3] Ibid.

non-profit. Finally, as a worst case scenario, government and other area non-profits may be able to make contingency plans to minimize the service gap if the non-profit closes programs.

To contribute to this important issue, our project seeks to identify non-profits who will experience a significant decline in revenue in the year ahead. We define a significant decline as a 20% decrease in revenue. We assume organizations that experience this much decline in revenue either must reduce the services they provide or will be at risk of closure at some point in the future. By identifying at risk non-profits a year ahead, stakeholders can prioritize oversight and direct needed managerial attention or additional funds to those NFPs most at risk regardless of whether the nonprofit has asked for help themselves.

Our model is able to achieve a precision of 44% in the top 10% of the population identified as most at-risk, which is substantially more than our naive model's 17% precision only using previous year's performance. While we find that this is a laudable accomplishment, we believe the model could be improved with additional data: when validating on 2015 data, we are only able to achieve a 17% precision. Fundamentally, this means that the patterns we identified from 2012-2013 to predict 2014 are not consistent year-over-year, something which can be rectified using additional years and expanding the scope of data we use for prediction.

Existing Literature on Sustainability

There is a large body of established literature on the topic of NFP sustainability, which encompasses both financial capacity--the ability to respond to opportunities and challenges--as well as financial stability--the ability to remain financially sound over time. Sontag-Padilla, et. al. (2012) outline five areas of focus for improving sustainability of NFP organizations common in the current literature:

1. Limit reliance on external funding sources and streams.
2. Invest in branding and marketing.
3. Exploration of external partnerships.
4. Demonstrate value and accountability to funders.
5. Promote community engagement and leadership.

For many NFPs, there is an inherent struggle to both adhere to these foci but at the same time achieve the organization's mission; as such, organizations vary in their ability to carry through on many of these items.

While philanthropic giving has steadily risen since the recession (National Philanthropic Trust, 2015), the recession serves as a warning regarding the uncertainty of relying on external funding. In the state of Illinois, because of political deadlock in the General Assembly, over $168 Million dollars has gone unpaid on contracts by the Illinois Department of Human Services for FY 2015 and, without a budget for 2016, it seems this trend of not paying NFPs for services provided will continue for Illinois,[4] and given the current political climate, is increasingly possible elsewhere. Diverse money flows for organizations is thus an important factor in understanding their stability. We take this into account in the model, including features that gauge funding reliance on government income and philanthropic grants rather than more stable sources, like income from programs. We are also attentive to reduction of costs of programs year over year and increased ROI from fundraising.

Other areas are difficult for us to grasp through IRS data. While we would have liked to access mission statements for processing branding concerns, these are not publicly released by the IRS and were outside of the scope of this project to obtain. We do measure some costs related to marketing and communications, which may serve as a proxy to determine whether any effort is going into branding or communicating success to donors, however it is likely we only are able to capture a portion of the determinates of NFP sustainability found .

Our Data

We have gathered data from the IRS on Non-Profits Tax Information for Calendar Years (CY) 2012, 2013, 2014, and 2015. The downloaded data represents the 990 forms which the Non-Profits have filed to claim their tax exemptions. Non-Profits which file 990's include tax exempt organizations which are not private foundations and have more than $200,000 in gross receipts or total assets more than $500,000. Private foundations and smaller nonprofits file separate tax documents not included in this data set.

While we have downloaded four years worth of data, we have kept Calendar Year 2015 in reserve for validation. Our goal is to predict Non-Profits which will experience the most drastic negative percent change in revenue for the next year since we assume these organizations will be more likely to have difficulty in providing the same level of service they have done in prior years.

Combining the Datasets

The organizations which file their taxes vary each year, so the same Employer Identification Number (EIN) is not present in every file. In addition, the datasets for each year contain duplicated values on the EIN. These are likely data entry errors of some kind. Because we link years with EIN number, and our dependent variable is the percentage change in revenue between years, all duplicated values must be dropped since the label for these observations would be ambiguous: organizations may be compared to different organizations between years and we would have no idea.

[4] Schoen, John W. “Illinois' epic budget fail sets a dubious record.” 3/30/2016. http://www.cnbc.com/2016/03/30/illinois-epic-budget-fail-sets-a-dubious-record.html. Accessed: 6/4/2016.

Table 1. *Record Counts for Calendar Years 2012-2014*

	2012	2013	2014
Total Records	294,019	289,603	299,405
Unique Records	265,473	266,089	274,263
Percent Dropped	9.7%	8.1%	8.4%

We combined these records on EIN to create longitudinal training dataset. After linking records and dropping records which are missing a total revenue or who have non-positive values for either 2012, 2013, or 2014 the remaining observations were 205,578 or about 77.4% of the original unique observations from 2012.

Absolute Variation in Total Revenue Changes

For the years we have observed, the median change in revenue is slightly positive. However, the standard deviation for the change in the two year time period from 2012 to 2014 is more than $20 million dollars which is more than 30 times the average total revenue of an organization. Although, there are large outliers in positive and negative changes in revenue, overall the skew is high in the positive direction.

Table 2. *Absolute Revenue Change YOY for Calendar Years 2012, 2013, and 2012-2014*

	2012 to 2013	2013 to 2014	2012 to 2014
Average	$318,766	$397,126	$715,892
Median	$6,256	$8,005	$14,878
Standard Dev.	$15,473,584	$18,096,780	$21,178,470
Min	-$--2,059,816,682	-$-3,534,121,809	-$-2,060,119,850
Max	$3,297,389,000	$2,197,928,696	$4,560,475,236

Percentile Variation in Total Revenue Changes

The average percentile change in revenue for each year is slightly positive 2.1% on average which is a little more than inflation. Percentile changes in revenue are greater in the two year time period on average 4.4%, which indicates that revenue trends between years are correlated.

Revenue changes in the bottom 40% are negative. For 2014, the year on which we are training our model, approximately 14% of our data fits into our category of a 20% or more decrease in revenue.

Table 3. *Decile Revenue Changes for Calendar Years 2012, 2013, and 2014*

	2012 to 2013	2013 to 2014	2012 to 2014
0.1	-0.281	-0..266	-0.327
0.2	-0.130	-0.119	-0.150
0.3	-0.058	-0.051	-0.062
0.4	-0.014	-0.009	-0.004
0.5	0.019	0.023	0.044
0.6	0.057	0.060	0.100
0.7	0.109	0.114	0.177
0.8	0.205	0.209	0.311
0.9	0.456	0.459	0.657

Model Development Flow

To predict organizations that will experience a 20% decline in revenue or more, our model uses two years of historical data. CY 2012 and 2013 were used to train on 2014 test data using 90/10 splits while 2015 was held in reserve for out-of-sample validation.

Holding 2015 data out-of-sample helps us demonstrate (or disprove) the legitimacy of our model, but there was significant trade-off in the amount of features available for us to use: our model uses two years of historical data instead of three. Additionally, the IRS has published more features on on the nonprofits since 2012, increasing from 62 columns in 2012 to 245 in subsequent years.

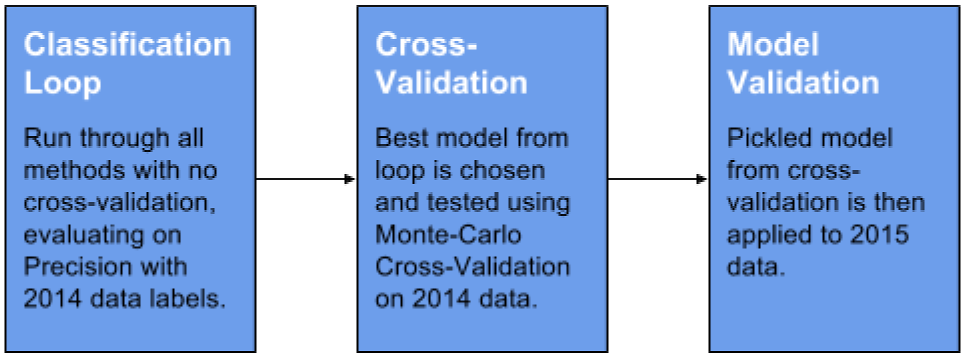
Figure 1 below illustrates the flow of our model evaluation and validation. We first run the set of features outlined above through a series of classification models to find the optimal method and set of parameters, evaluating each model based on its Precision at a threshold k = .10.[5] Given that the hypothetical application of the model is to focus a Foundation or Government Agency's resources and that approximately 14% of the dataset are at-risk in 2013, we find it pertinent to restrict Precision to only the 10% of the population we are most confident will see revenue reductions.

Second, to ensure generalizability, we cross-validate using a Monte-Carlo process on 90/10 splits of the data. While we do not execute this on a high n number of iterations, given a relatively close range of values, we are comfortable with asserting that the average of these simulations produces a good approximation of our expected generalizability. Here, we maintain the 10% population precision as our primary evaluation metric, however, understanding that

[5] That is, the percent of actual 2014 at-risk organizations we are able to predict using the model, if we only predict positive for the top 10% of cases. We use Decision Trees, Random Forests, Adaptive Boosting, Logistic Regression, Support Vector Machines (Stochastic ßGradient Descent), K-Nearest Neighbors, and Naive Bayes in this loop.

organizations may want fewer than ~30,000 organizations to inspect, we also do an analysis at 5% of the population. The best model from cross-validation is then preserved and used to predict on 2015 labels using 2013/2014 data for final model validation.

Figure 1*: Validation flow for seismos model development*



Feature Generation

We develop 154 features for our dataset, which are listed in the Appendix to this report (Table A1). These are extensive, based on CY2012 IRS 990 variable availability. Basing our feature development on 2012 variable availability was somewhat limiting: releases from CY2013 and after contain more detailed reporting for each organization. We pull in some CY2013 specific variables for our training on 2013/2014 data, however, since we are predicting YOY performance, most of our variables are YOY changes in variables and are thus limited. We explore later the extent to which this impacted our model's potential performance.

We decided classify organizations as 'At Risk' for instability using their percent change in total revenue YOY, labeling organizations that have a decrease of 20% or greater as 'At Risk'. This targets about 14% of the organizations within our dataset and represents those most at risk for fiscal instability. We do not chain this to any kind of local market average to avoid implicitly predicting macro-level performance for the year-ahead; instead, we include macro-level performance by MSA as an input within the dataset as a macro factor in these types of issues.

A brief overview of key features:

- CY 2013 GDP per Capita for the organization's MSA
- CY 2012-2013 revenue change
- Whether revenue fell by over 20% between CY 2012 and 2013
- Whether the organization is missing from the CY 2012 or 2013 datasets
- Dichotomous variables for each of the Major Groups of NTEE codes (which give the type of non-profit organization)[6]

[6] For more information, please see: http://nccs.urban.org/classification/ntee.cfm. Accessed 5-25-2016.

- Percent of income from various sources, including government and fundraising
- Fundraising ROI, Debt-Asset Ratios, and other revenue performance metrics

In exploring these features, we find a significant number of organizations are missing data. We remove organizations that are not present in our CY 2014 dataset, since there is no way for us to test the predictive power of the model on these organizations; we carry this policy over to 2015 when validating the final model. In order to reduce noise in the dataset caused by extensive missing values in our features, we run models both removing those features with over 40% missing and keeping them while simply labeling organizations as missing those features using a dichotomous variable. An analysis on the benefits this exercise is also included here.

The Naïve Model

Because we are trying to predict non-profits that have an unusually high fall in revenue, we could obtain a high level of accuracy simply by predicting that all nonprofits will not have a decrease in revenue of 20% or more. If we were to apply that model to the 2014 data set, we would obtain 86.63% accuracy. However, since we are creating an early warning detection system, we care more about precision so that stakeholders can confidently intervene with nonprofits that are predicted as having financial issues for the following year.

As a baseline, we applied the same label for 2013 to 2014 under the assumption that financial performance for the year prior may in many instances be the best predictor of the following year's financial performance. If a nonprofit had a 20% or more decline in revenue from 2012 to 2013, we predicted it would again have a 20% ore more decrease in 2014. We ran this model for all nonprofits in our dataset for which we had both 2013 and 2014 data (n = 177,443):

Table 4. *Confusion Matrix for Naïve Model*

True Positive: 2.41%	**False Negative: 10.95%**
False Positive: 11.74%	**True Negative: 74.89%**

Table 5. *Naïve Model Metric Evaluation*

Metric	Performance
Precision:	0.170
Recall:	0.180
Accuracy:	0.773

Final Model Development

After establishing a baseline in the naive model, we ran eight types of models with different set of parameters: K-Nearest Neighbors (KNN), Random Forests (RF), Logistic Regression (LR), Naive Bayes (NB), Decision Trees (DT), Adaptive Boosting with Decision Trees (AdaBoost), Extra Trees (ET) and Support Vector Machines with a Stochastic Gradient Descent (SVM/SGD). To find the "best" model, we trained and validated 1000+ models on 2012-2014 data with a 90/10 split. The classifier that achieved *highest* precision at 10% threshold is a Random Forest that takes in all the 98 features with 1000 trees and 100 max depth.[7] This model is also the one that has *highest* AUC and the *second highest* recall rate at 10% threshold. Meanwhile, none of the other types of models have achieved an accuracy of 0.35 or above at this 10% threshold. Based on all this information and under the assumption that 10% threshold is roughly the scale of our model application, we believe that the Random Forest Classifier is the ideal candidate.

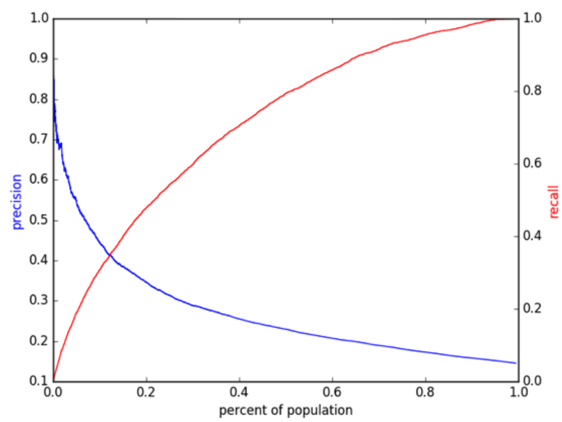
Table 6: *Tuned Model Performance*

Metric	RF (n=1000, features=98)
Precision (10%):	0.443
AUC (ROC):	0.752
Recall (10%):	0.306
Accuracy (10%):	0.844
10% Top Prob Threshold:	*0.323*

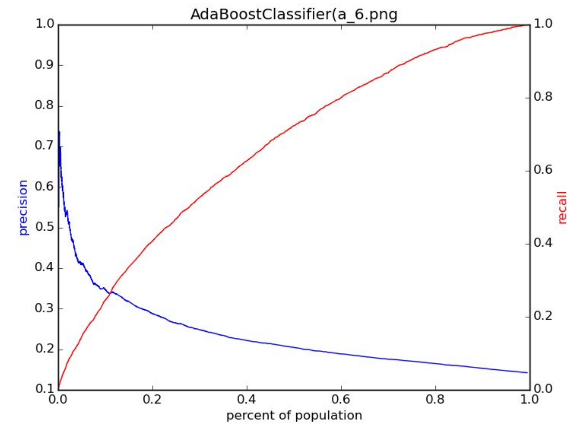
Below is a comparison of the precision-recall curve between the Random Forest Classifier we identified and an Adaptive Boosting Classifier that achieves the highest precision rate among all non-RF classifiers. And it is easy to see that the RF classifier out-performs the AB classifier at almost all thresholds k.

[7] Precise specifications: RandomForestClassifier; bootstrap = True; class_weight = None; criterion = 'gini'; max_depth = 100; max_features = 'sqrt'; max_leaf_nodes = None; min_samples_leaf = 1; min_samples_split = 5; min_weight_fraction_leaf = 0.0; n_estimators = 1000; n_jobs = -1; oob_score = False; verbose = 0; warm_start = False).

Chart 1: *Comparison of Precision Recall Curve for Tuned Models*



Random Forest



Adaptive Boosting

Table 7: *Feature importance for tuned RF Model*

Rank	Features:	Importance
1	Total Contributions Change %, 1 Year Prior	0.0504
2	Total Asset Change %, 1 Year Prior	0.0474
3	Total Functional Expense Change %, 1 Year Prior	0.0444
4	Program Revenue % of Total Revenue, 1 Year Prior	0.0422
5	Total Asset, 1 Year Prior	0.0404
6	Total Functional Expense, 1 Year Prior	0.0395
7	Program Revenue % of Total Revenue, 2 Year Prior	0.0286
8	Investment revenue % of Total Revenue, 1 Year Prior	0.0284
9	Total Functional Expense, 2 Year Prior	0.0278
10	Office Expense, 1 Year Prior	0.0275

We can see that: (1) the most relevant type of features are those measuring the trend last year (i.e. Change %); (2) the component of the revenue matters (reliance on external fundings); (3) some aggregate financial indicators are important (total asset and expense); (4) 1 year prior features are more relevant than 2 year prior features. These results confirm our intuition.

To refine the model, we have done several additional analyses:

Additional features with null dummies

We have generated 154 variables during the feature generation process. However, in the models above we have dropped 56 features that have missing values>40%. The rationale is that we might bring about additional noise by introducing variables with large proportion of missing values imputed. Yet on the other hand we may lose valuable information by dropping the features completely. To address this issue we impute all columns with missing values and indicate null/not-null with dummies. This expand our feature space to 277 (compare to 98). And we ran the RF Classifier on this new feature space for the 2012-2014 data. We saw a slight performance decrease, probably indicating that the noise outweighed new information during imputation.

Additional features with 2013-2015 data

As we have discussed above, the 2012-2014 data have much smaller feature space (154 vs 260) than the 2013-2015 data (and probably the data in the future as IRS is revealing more information to the public). Here our choice of starting year is limited by data availability validation requirement. To address the potential loss of explanatory power due to limited feature space in 2012-2014 data, we ran again the eight types of models on the 2013-2015 data. And the best one is still the Random Forest with 1000 trees, indicating stability in model types. We did not see an increase in precision with additional features, probably indicating that the newly available features after year 2013 are not that relevant for our purpose.

Alternative revenue change specification

Previously we labeled organizations that have a decrease of 20% or greater as 'At Risk'. However the choice of 20% is partly arbitrary. We here changed the threshold to 40% to see how our model performs for the even more risky organizations. Previously about 14% of the nonprofits in our training set met the criteria to be at risk, now with this adjustment, about 6.5% meet the criteria. Taking this into consideration, we also adjusted the k threshold from 10% to 5% since the 40% decrease is a rarer event and we wanted to send fewer warnings.

With this new label, at the new threshold, the model has a precision of 0.387 and a recall of 0.26 with an overall AUC of 0.79. As there was no significant benefit, we stuck to our initial 20% threshold.

Additional randomness by model tuning

Within the different RF Classifiers we noticed that there is an increase in precision with the increase of trees (n_estimators). So we also explore higher number of trees. We did not see an increase in performance by increasing n from 1000 to 2000, indicating a decreasing marginal effect.

Table 8: *Increased RF classifiers number results*

Metric	RF (features=98, n=2000)
Precision (10%):	0.431
AUC (ROC):	0.749
Recall (10%):	0.301
Accuracy (10%):	0.843
10% Top Prob Threshold:	*0.293*

Reduction of population threshold to .05

After the analyses above, we then lowered our population threshold for labeling; that is to say, we used the model to label as 'At Risk' the top 5% of scores coming from the model rather than the top 10%. We find that this is able to achieve 5 point increase in Precision over the tuned model, above, though with substantially lower Recall.

Table 9. *Results for reduction of population threshold to .05*

Metric	RF (features=98, n=1000)
Precision (5%):	0.490
AUC (ROC):	0.737
Recall (5%):	0.174
Accuracy (5%):	0.858
5% Top Prob Threshold:	*0.385*

While a 5 point increase in Precision is valuable, it is not clear to us that identifying half the number of organizations actually at risk is justified in order to reduce incorrect identifications. If we assume a population of around 280,000 non-profits, a 5% threshold yields 6,860 True Positives, about half the correct identifications--13,200 organizations--with a 10% threshold while reducing False Positives by around 8,540.

This requires some consultation with actual non-profit professionals: there are ethical implications, explored in the Applications section, that we are not sufficiently familiar enough with non-profit governance and management to identify and think through. Because of this, we keep the .10 threshold for our model.

Cross-Validation

For the tuned Random Forests model, we ran Monte-Carlo cross-validation simulations to determine generalizability. Iterating over 10 90/10 train/test splits, we found our model approached an average of 42.6% precision in the top 10% of organizations. Given the small standard deviations across all metrics and computational strain of extensive cross-validation, we determined that doing additional simulations would be unnecessary to illustrate consistent generalizable expected performance.

Table 10. *Final Model Monte-Carlo Cross-Validation Metrics*

Model	Precision	AUC (ROC)	Recall	Accuracy
Average	0.426	0.746	0.302	0.844
Standard Dev.	0.011	0.005	0.006	0.002

Validation on 2015 Data

Upon validation with the 2015 data, we find that our model far underperforms our expectations. We see a reduction from 42.6% cross-validated Precision at the 10% population threshold down to 18.1% Precision. This is not significantly improved by lowering that threshold:

Table 11: *Validation results on 2015 data using trained model*

Metric	RF (n=1000, features=98)
Precision (10%):	0.181
AUC (ROC):	0.541
Recall (10%):	0.127
Accuracy (10%):	0.793
10% Top Prob Threshold:	*0.563*

Table 12. *Results for validation on 2015 with .05 threshold*

Metric	RF (features=98, n=1000)
Precision (5%):	0.188
AUC (ROC):	0.541
Recall (5%):	0.066
Accuracy (5%):	0.826
5% Top Prob Threshold:	*0.567*

These results are disappointing and indicate that the patterns we identified in 2012/2013 data to predict 2014 simply were not consistent enough over time to predict 2015 performance well. There are a variety of implications for this; most notably, it implies that we were far too restricted in terms of the number of years available for predictive work on this project. With more years, patterns that are strong for small windows are smoothed out, such that persistent patterns are more easily detected.

Additionally, this may also indicate that the fiscal factors which we are able to identify through IRS 990 returns are not sufficient to determine performance over time, rather only for specific years. It could very well be that coming out of the downturn fiscal aspects of NFP organizations' management were highly indicative of their ability to maintain revenue, whereas in 2015, as the market begins to recover, non-financial factors like non-profit management and branding and macro factors like local government effectiveness may become more important. Without additional data, we won't be able to measure these well.

Limitations, Caveats, and Future Work

Ultimately, the precision of our model is not significantly higher than the naive model's precision of applying last year's label to this year's label. If being labeled at risk could be harmful to a nonprofit that is not actually at risk, our model should not be applied.

As the IRS continues to publish more data, more features will be available. Currently there are several sections of the Form 990 Tax Documents not accessible for the public, including nonprofit mission statement, programs, governing structures, and violations records. It would also be helpful if we could access tax data that contains less noise: as mentioned previously, more than 9% of the nonprofits were dropped each year due to duplicated values which decreases our confidence in the overall reliability of this dataset. For next year, we could train on three years of historical data and keep the 2016 data in reserve as a validation set. Adding in longer historical trends may increase performance some, but we should also consider that self-reported tax information may not be very predictive of future financial performance in general and be an inherent limitation to constructing a model of this type. We would like to consider accessing other data sources to gather more information on the quality of the donors of the nonprofits, such as how consistently do they donate, the structural information on the nonprofit (this is missing from the IRS published data), and board membership ties.

Before attempting to iterate and build another model, we should first discuss its efficacy with nonprofits, foundations, large individual donors, and government funders. Our model assumes risk scores would be used to intervene with non-profits earlier and help them - but in practice it may be used as a method of triage and could further accelerate declining non-profit financial performance. Perhaps, this is why some nonprofits are reluctant to be open about their financial troubles - and we should consider whether those fears are justified.

We also could use subject matter expertise on the relative costs of false positives and false negatives. If we miss a nonprofit which will experience a significant revenue decline how much more likely is it that the nonprofit will need to reduce programs or close than if we had identified it? Does incorrectly putting a non-profit which will not experience significant revenue decline into the at risk category affect donors confidence and perhaps cause decreasing revenue? Or could it cause a foundation to give it extra grants when it is not at risk and those grants could be better used elsewhere? More information on these risks would be needed to tune a model that could be applied.

Appendix

Table A1. *Features Generated*

Feature Generated	Feature Type
YOY Gross Revenue Percent decrease over 20 Percent	Revenue
GDP for 2013	Macro
Five-year GDP Trend	Macro
One Year Prior YOY Gross Revenue Percent increase or decrease	Revenue
NTEE Major Code (2016)	Demographic
Missing for 2013	Demographic
Missing for 2012	Demographic
One year prior has negative revenue	Revenue
One Year Prior number of employees	Demographic
One Year Prior YOY Change in Payroll Taxes	Expense
One year Prior YOY Change in net assets	Balance Sheet
One year prior positive income from sale of goods	Balance Sheet
One year YOY prior gross receipts change	Revenue
Two year prior gross receipts	Revenue
One year prior gross receipts	Revenue
Two year prior collected member dues	Revenue
One year prior collected member dues	Revenue
One year prior Percent of revenue from program services	Revenue
One year prior Percent of revenue from investment	Revenue
One year prior Percent of revenue from rental income	Revenue
One year prior Percent of revenue from net sale in assets	Revenue
One year prior Percent of revenue from fundraising income	Revenue
One year prior fundraising expenses ROI in income per dollar	Revenue
One year prior Percent of revenue from sales of inventory income	Revenue-Expense
Two year prior Percent of revenue from program services	Revenue
Two year prior Percent of revenue from investment	Revenue
Two year prior Percent of revenue from rental income	Revenue
Two year prior Percent of revenue from net sale in assets	Revenue
Two year prior Percent of revenue from fundraising income	Revenue
Two year prior fundraising expenses ROI in income per dollar	Revenue

Two year prior Percent of revenue from sales income	Expense
One year prior total assets	Revenue
Two year prior total assets	Balance Sheet
One year prior absolute total liabilities	Balance Sheet
Two year prior absolute total liabilities	Balance Sheet
One year prior debt-asset ratio	Balance Sheet
Two year prior debt-asset ratio	Balance Sheet
One year prior Change in assets	Balance Sheet
One year prior Change in liabilities	Balance Sheet
One year prior total gifts, grants, membership fees	Balance Sheet
Two year prior total gifts, grants, membership fees	Revenue
One year prior total functional expenses	Revenue
Two year prior total functional expenses	Expense
One year prior YOY absolute change in total functional expenses	Expense
One year prior YOY Change in total functional expenses	Expense
One year prior total compensation of officers, directors etc.	Expense
Two year priot total compensation of officers, directors, etc.	Expense
One year prior YOY absolute change in total compensation of officers	Expense
One year prior YOY Change in total compensation of officers	Expense
One year prior Used a professional fundraising agency	Expense
Two year prior Used a professional fundraising agency	Revenue
One year prior YOY change in Percent receipts from members	Revenue
One year prior YOY change in Percent of revenue from program services	Revenue
One year prior YOY change in Percent of revenue from investment	Revenue
One year prior YOY change in Percent of revenue from rental income	Revenue
One year prior YOY change in Percent of revenue from net sale in assets	Revenue
One year prior YOY change in Percent of revenue from fundraising income	Revenue
One year prior YOY change in fundraising expenses ROI in income per dollar	Revenue
One year prior YOY change in Percent of revenue from sales income	Expense
Paid for professional fundraising after drop in Percent revenue from fundraising	Revenue
Paid for professional fundraising after drop in fundraising ROI	Revenue
Paid for professional fundraising one year prior and fundraising ROI increased the following year	Revenue
Paid for professional fundraising one year prior and Percent of revenue from fundraising increased the following year	Revenue

One year prior large (+25Percent) rental expense increase (have new location)	Revenue
One year prior Collected royalties	Revenue
Two year prior collected royalties	Expense
One year prior government facilities/services support as Percent of total support	Expense
Two year prior government facilities/services support as Percent of total support	Revenue
One year prior taxes levied on behalf of organization / total support	Revenue
Two year prior taxes levied on behalf of organization / total support	Revenue
YOY Change in Percent government facitiles/services of total support	Revenue
YOY Change in Percent taxes levied of total support	Revenue
One year prior Support-Revenue Ratio	Revenue
Two year prior Support-Revenue Ratio	Revenue
YOY Change in support-revenue ratio	Revenue
YOY absolute change in paid-in or capital surplus	Revenue
YOY Change in paid-in or capital surplus	Revenue
One year prior any paid-in or capital surplus	Revenue
Two year prior any paid-in or capital surplus	Revenue
One year prior ratio of unsecured loans to revenue	Revenue
Two year prior ratio of unsecured loans to revenue	Revenue
YOY Change in ratio of unsecured loans to revenue	Revenue
One year prior functional expenses as a Percent of revenue	Revenue
Two year prior functional expenses as a Percent of revenue	Revenue-Expense
YOY Change in expenses / revenue	Revenue-Expense
One year prior program expense	Expense
Two year prior program expense	Expense
One year prior administrative expense	Expense
Two year prior administrative expense	Expense
One year prior fundraising expense	Expense
Two year prior fundraising expense	Expense
One year prior expenses exceeded revenue	Expense
Two year prior expenses exceeded revenue	Expense
One year prior fundraising efficiency	Expense
Two year prior fundraising efficiency	Revenue-Expense
One year prior working capital ratio	Revenue-Expense

Two year prior working capital ratio	Revenue-Expense
One year prior number of Individuals over 100k	Revenue-Expense
Two year prior number of Individuals over 100k	Expense
One year prior number of Contractors over 100k	Expense
Two year prior number of Contractors over 100k	Expense
One year prior distribution equality	Expense
Two year prior distribution equality	Expense
One year prior interest expense	Expense
Two year prior interest expense	Expense

Sources

1. National Philanthropic Trust, 2015. *2015 Annual Report*. http://www.nptrust.org/annual-report/. Accessed: June 4, 2016.
2. Sontag-Padilla, Lisa M., Lynette Staplefoote, and Kristy Gonzalez Morganti, 2012. *Financial Stability of Non-Profit Organizations: A Review of the Literature*. RAND Corporation.
3. Thayer, Kate. "Hull House Closing Friday". *The Chicago Tribune.* January 25, 2012. http://articles.chicagotribune.com/2012-01-25/news/ct-met-hull-house-20120126_1_child-care-union-contract-employees Accessed 6/4/2016
4. West, Maureen. "Some Fear Hull House Closure Is an Omen for Struggling Charities." *Chronicle of Philanthropy*, February 02, 2012. https://philanthropy.com/article/Collapse-of-Famous-Hull-House/157181. Accessed: 6/4/2016.