Predicting Full Funding for Donors Choose Projects

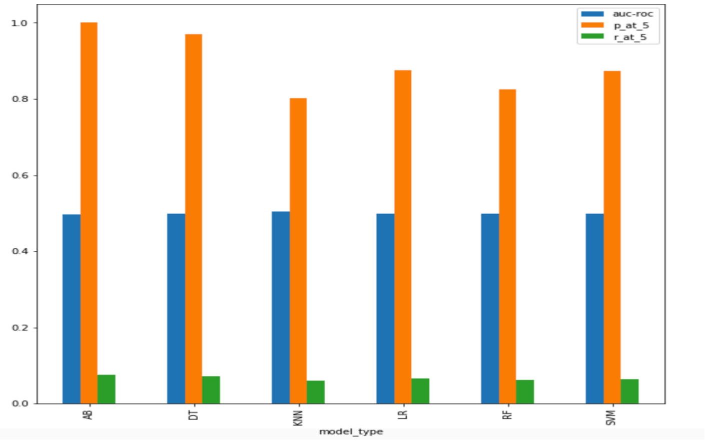
Given the high cost of school supplies, and decreasing funding in public schools across the US, many teachers struggle with the day-to-day challenge of providing basic resources to their students. An organization known as DonorsChoose connects teachers in high-need communities with donors who want to help. In order for this organization to be effective, it is helpful to know what factors increase the likelihood that a project receives funding. In order to investigate this issue, we ran various models to predict the likelihood that a project was funded within 60 days of it being posted. The purpose of the analysis was to classify observations as either being fully funded in 60 days, or not being fully funded in 60 days. Running 182 models over 7 different model types, we compared the model predictions to the actual outcome. Using various evaluation metrics, we compared the performance of the models. The metric we used to find the best model was the AUC-ROC metric, which is the likelihood that a positive result is correctly identified as being positive, and likelihood that a negative result is correctly identified as being negative. Based on our results, we recommend that DonorsChoose select the Random Forest classifier model to predict the likelihood of a project being fully funded in 60 days.

Model and Evaluation Details

In order to suggest an appropriate model to predict successfully funded projects, we had to run and compare the results of various models. For this particular analysis, we ran the following models: K-Nearest Neighbors, RandomForest, Boosting, Support Vector Machines, Logistic Regression, and Decision Trees. We trained the models using a rolling 6-month window, which resulted in 3 different timeframes, and 3 training sets and 3 testing sets, respectively. We varied the parameters to find the model with the optimal parameters for prediction. Overall, none of the models performed particularly well it came to their AUC-ROC measure. All the models had an AUC that hovered at around 50%, and even the best model (measured by having the highest AUC) only had an AUC of 50.8%. This is far less than the 69%-74% accuracy of the baseline model, which measures the accuracy of a random prediction. Basically, this suggests that the models trained do a worse job predicting whether a project would receive full funding within 60 days than a random prediction would do.

The figure below shows the relative performance of the models across a few metrics, including AUC, precision, and recall. The graph shows metrics measured for the bottom 5% of projects (those least likely to get funded). Precision tells us the percent of projects classified as having been fully funded that were actually fully funded. Recall tells us the percent of projects that were actually fully funded in 60 days that the model correctly classified. As the figure shows, all the models do significantly better when it comes to precision when compared to recall. Typically, whether the organization is more interested in optimizing precision or recall is dependent on the resources that an organization has to serve people, and the target group that they are hoping to serve. In this particular case, if we are interested in identifying the 5% of projects that are at highest risk of not getting fully funded, we would be interested in precision at 5%, because with limited resources, the organization would likely want to identify and target the projects in most need of help, and divert time away from projects in less need of help to achieve funding.

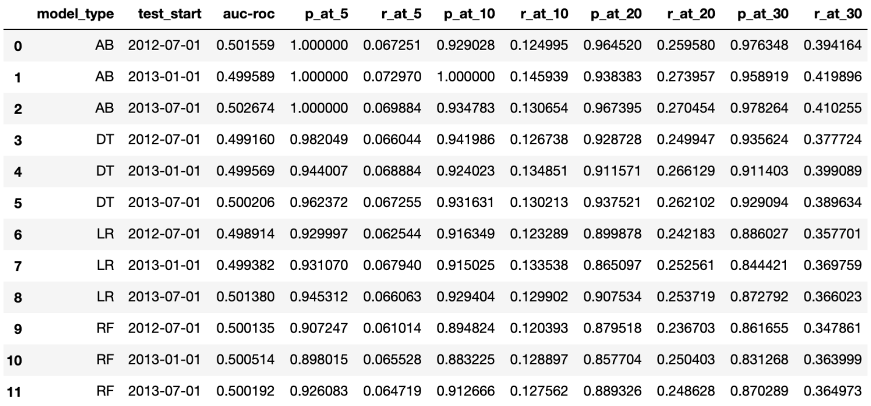
Figure 1: AUC, Precision, Recall at Bottom 5%



The boosting model appears to perform the best when it comes to precision, and the K nearest neighbor model appears to perform the worst when it comes to precision (and the other two metrics as well). The relative slump in performance for the K-Nearest Neighbor model could be due to the large number of features included in the model. The KNN model is typically sensitive to including inconsequential features. It is possible that pruning features could result in better performance of the KNN model. Given the poor overall AUC, it is likely that the organization would need to further analyze the data and tweak the models to more accurately predict the success/failure of funding.

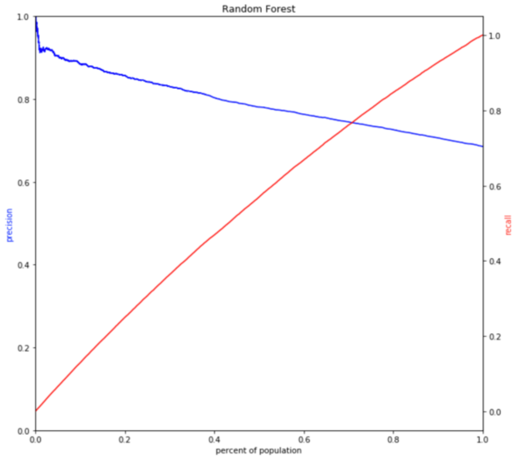
The full tables (in the appendix) show precision and recall for the bottom 1%, 2%, 5%, 10%, 20%, 30%, and 50% of the projects. An abridged table below shows the trends across models over time. Precision is particularly high for boosting, decision trees, and logistic regression. Recall is relatively low over all the models over time, but increases as the threshold increases.

Figure 2: AUC, Precision, Recall Over Time



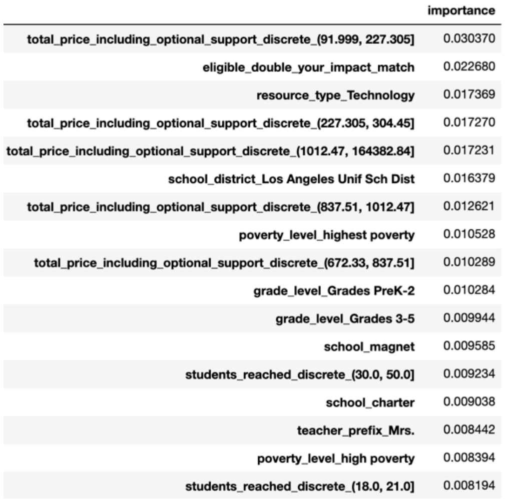
	model_type	test_start	auc-roc	p_at_5	r_at_5	p_at_10	r_at_10	p_at_20	r_at_20	p_at_30	r_at_30
0	AB	2012-07-01	0.501559	1.000000	0.067251	0.929028	0.124995	0.964520	0.259580	0.976348	0.394164
1	AB	2013-01-01	0.499589	1.000000	0.072970	1.000000	0.145939	0.938383	0.273957	0.958919	0.419896
2	AB	2013-07-01	0.502674	1.000000	0.069884	0.934783	0.130654	0.967395	0.270454	0.978264	0.410255
3	DT	2012-07-01	0.499160	0.982049	0.066044	0.941986	0.126738	0.928728	0.249947	0.935624	0.377724
4	DT	2013-01-01	0.499569	0.944007	0.068884	0.924023	0.134851	0.911571	0.266129	0.911403	0.399089
5	DT	2013-07-01	0.500206	0.962372	0.067255	0.931631	0.130213	0.937521	0.262102	0.929094	0.389634
6	LR	2012-07-01	0.498914	0.929997	0.062544	0.916349	0.123289	0.899878	0.242183	0.886027	0.357701
7	LR	2013-01-01	0.499382	0.931070	0.067940	0.915025	0.133538	0.865097	0.252561	0.844421	0.369759
8	LR	2013-07-01	0.501380	0.945312	0.066063	0.929404	0.129902	0.907534	0.253719	0.872792	0.366023
9	RF	2012-07-01	0.500135	0.907247	0.061014	0.894824	0.120393	0.879518	0.236703	0.861655	0.347861
10	RF	2013-01-01	0.500514	0.898015	0.065528	0.883225	0.128897	0.857704	0.250403	0.831268	0.363999
11	RF	2013-07-01	0.500192	0.926083	0.064719	0.912666	0.127562	0.889326	0.248628	0.870289	0.364973

In other words, recall is better for the bottom 50% of projects compared to the bottom 5% of projects. There do not appear to be any noticeable differences in performance over time. Overall, the high precision and low recall for all models over time shows that the models are doing well at identifying the projects that are fully funded within 60 days, but maybe over-classifying projects as not being fully funded within 60 days. The precision-recall curve for the best performing model, the Random Forest model with max depth of 100 is shown below.



Policy Implications

How can the results of these models help DonorsChoose more effectively connect donors to projects and ensure that projects are quickly funded? One approach is to look at the best performing model, and explore which features were most useful in prediction. The features importance table below shows these features. The top features are the total price of the project, the eligibility for matching, and the type of technology. The table suggests that in order to get fully funded, teachers should limit the total price of their posting. Of course, from a policy perspective, this is challenging because of the continued threat to public school funding. However, none of the features have particularly high importance which suggests that there are likely too many features included in the model. However, as the models continue to be improved, DonorsChoose could use feature importance to determine what areas to focus on when helping the bottom 5% of projects receive full funding.



	importance
total_price_including_optional_support_discrete_(91.999, 227.305]	0.030370
eligible_double_your_impact_match	0.022680
resource_type_Technology	0.017369
total_price_including_optional_support_discrete_(227.305, 304.45]	0.017270
total_price_including_optional_support_discrete_(1012.47, 164382.84]	0.017231
school_district_Los Angeles Unif Sch Dist	0.016379
total_price_including_optional_support_discrete_(837.51, 1012.47]	0.012621
poverty_level_highest poverty	0.010528
total_price_including_optional_support_discrete_(672.33, 837.51]	0.010289
grade_level_Grades PreK-2	0.010284
grade_level_Grades 3-5	0.009944
school_magnet	0.009585
students_reached_discrete_(30.0, 50.0]	0.009234
school_charter	0.009038
teacher_prefix_Mrs.	0.008442
poverty_level_high poverty	0.008394
students_reached_discrete_(18.0, 21.0]	0.008194