Esther Edith Spurlock (12196692)
CAPP 30254
Assignment 3: Report

Identifying the Problem:

It is not novel to say that teachers often go above and beyond the call of duty. As the daughter of a middle school math teacher, and as someone who worked in a school for a year, I have seen how educators spend their evenings and weekends grading assignments, how they give up their summer holidays to prepare for the upcoming school year, how they take the time to chaperone school events, and how they spend their own money on classroom supplies they will not get from their district.

Teachers already give so much, but they always find a way to give even more. Nothing exemplifies this more than the amount of teachers fund classroom projects using donorschoose. Instead of looking to their district to fund projects that will help their students learn, teachers feel the need to crowdsource these projects.

As policymakers, I implore you to look at the data I have provided here and see that there is a way to fund the most important projects listed here.

Assumptions / Data Modeling:

Many of the projects posted on donorschoose get funded within 60 days of posting. I am identifying these as the most important projects. While this metric might not always accurately identify which projects are the most worthy of your resources, it is the metric I have used here.

You have pledged to support 5% of projects listed on this website, and I am here presenting you with models that will identify which 5% of projects are the most important ones to fund. While these projects would get funded on donorschoose, teachers should not have to go to the public to ask them for money for schools. Instead, teachers should look to the government first for project funding before they resort to crowdsourcing.

While analyzing this data, I created thousands of models that identify which projects are the most important to fund soon. Before I detail which models do better or worse, I would like to take a moment to discuss the evaluation metrics I used to identify how well my models performed on the data.

Evaluation:

When evaluating the models, I used 5 metrics

- Accuracy: looks at how many projects are correctly predicted
- Precision: the number of true positives divided by the number of predicted positives. Said another way, this is the percentage of people my model predicted as gaining funding within 60 days who actually received funding in that time period.
- Recall: the number of true positives divided by the number of true positives plus the number of false negatives. Or, the percentage of people who received funding within 60 days that we predicted as receiving funding in that time.
- F1: the average of precision and recall

- ROC AUC: Plots the fraction of true positives vs the fraction of false positives and takes the area under the curve

Which Models Do Best for Which Evaluations?

Unfortunately, not all of my models did equally as well on each of the metrics I measured.

The best accuracy score occurred when looking at the 10 nearest neighbors.

Precision worked best with SVM models.

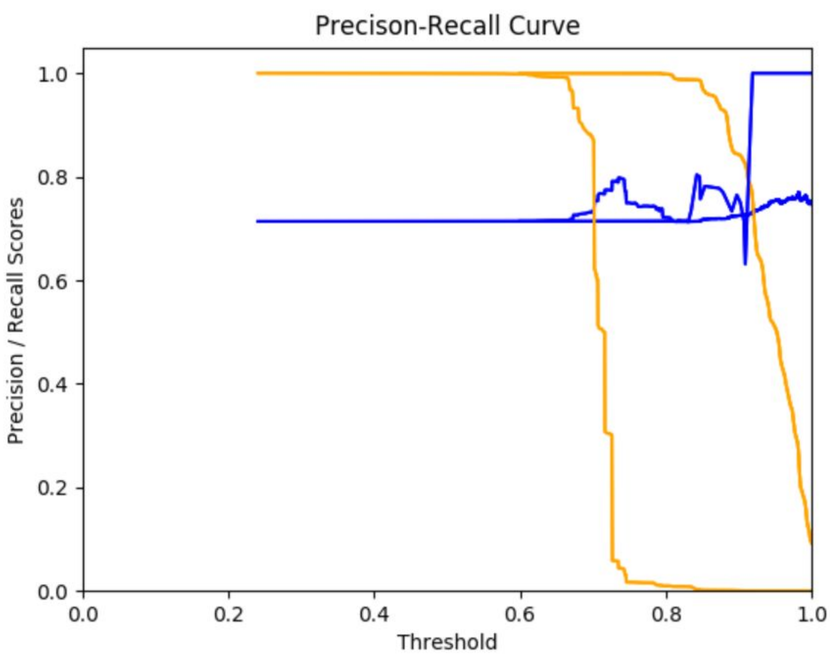
Recall did not have a clear best model because all model types returned a recall of 1 using different parameters and at different threshold levels.

F1 tended to do better with bagging models

ROC AUC worked best with Ada Boosting and SVM models.

Since none of these models overlapped a whole lot, I looked for the model that consistently worked well across all evaluation metrics (although perhaps not the best on all of them). The models that I chose as the one I liked the most were SVM with a C value of 1.5 and Extra Trees with 5 trees and a depth of 5.

I have here graphed their precision and recall scores at different thresholds.



Changing Results Over Time

Because I have data form 2012-2013, I looked to see how different models performed at different points in time. Generally speaking, I did not see much of a change in the evaluation metrics as time progressed. Different models did better or worse depending on the timeframe, but all in all, everything remained fairly consistent.

Choosing the Top 5%

Because you have pledged to support the top 5% of projects, I looked at models under a threshold of 0.05. This means that we are only looking at the top 5% of the population.

Using the Extra Trees model described above under this threshold, we can expect a recall of 100%, an average precision of 71.5%, an average F1 statistic of 83.3%, and an average accuracy of 71.5%.

Using the SVM model described above under this threshold, we can expect an average recall of 51.6%, an average precision of 82.2%, an average F1 statistic of 55.2% and an average accuracy of 53.7%.

Given these results, I would suggest you sue the Extra Trees model.

Conclusions:

It is high time that teachers start relying on the government more for their funding. I am grateful that you have begun thinking more seriously about funding teacher's extra projects. Please consider using my models to help you decide what projects are more important to fund.