Esther Edith Spurlock (12196692)
CAPP 30254
Assignment 5: 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 do not get funded within 60 days of posting. I am identifying these as the most important projects as they are projects teachers obviously feel deeply about but they will not get the funding they need in a timely manner.

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.

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 not 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 will not 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?

Changing Results Over Time

Choosing the Top 5%

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.