

---

# CS 4780 Final Project Proposal

---

## 1. Team

This project will be completed by a team of four students.

## 2. Motivation

DonorsChoose is a crowd funding website that helps public school teachers request and receive funding. 70 percent of campaigns on DonorsChoose are successfully funded. While this is a healthy margin, it could be vastly improved. It would be valuable for teachers to learn what factors may affect the success of their campaign.

Extensive research about Kickstarter and other similar crowd-funding sources has already been done to investigate causes of successful commercial campaigns. It would be interesting to compare and contrast techniques for creating a successful crowd-funding campaign in the commercial realm to the philanthropic realm.

## 3. Problem Statement

The main goal of this project is to determine what factors have the greatest influence on if a project will be fully funded. In particular, we are interested in investigating whether characteristics of a project such as the location of the school, the poverty level, the grade level, or area of study (such as English versus chemistry) affect the likelihood of funding.

We are also interested in looking at how characteristics of already pledged donations affect likelihood of future donations, and thus success of projects in a time-series framework.

DonorChoose enables various promotions such as having a corporation match donations. We would like to investigate if these promotions affect the number or size of donation as well as if they affect the likelihood of a project being funded.

We also have access to project description essays, written by the creating teachers. We will analyze text samples to determine if certain keywords or other text features increase likelihood of funding for a given project.

## 4. Approach

## 5. Resources

The full dataset for this project is provided publicly by Kaggle. It is available here: <https://www.kaggle.com/c/kdd-cup-2014-predicting-excitement-at-donors-choose/data>

-something about the linear classifier software we plan on using...

File reading and other custom code written for this project is written in Scala, a publicly accessible and OS-agnostic language.

## 6. Schedule

- 24th October: Peer reviews for project proposal due (via CMT).
- 11th November: Submit progress report (via CMT).
- 04th December: Poster presentation. Submit poster (via CMT).
- 05th December: Peer reviews for posters due (via CMT).
- 10th December: Final project report (and code) due (via CMT).
- 15th December: Peer reviews for final project reports due (via CMT).
- 16th December: Author Feedback on reviews for final project reports due (via CMT).