Project proposals consist of the names of the team members, project topic, a brief summary of the project, and a plan for completing the project including specifications of tools/libraries that will be used. The proposals must be submitted in writing and presented to the class on Feb 7, 2019. Each team will get 5 minutes for proposal presentation.

**Team Members:**

Bill Fisher, Ron Richardson, Christina Eusanio

**Project Topic:**

Predict which Reddit posts (subreddits) will receive the highest score based on features extracted from the Reddit API.

**Project Plan:**

Reddit is the fifth most popular site on the internet according to the data captured by Amazon’s Alexa. Reddit is a social message board where users can create profiles, create topics, and interact with others in a variety of ways. When users create posts, they receive upvotes and downvotes. The aggregate of those votes constitutes as a post’s score. Our plan is to extract various features from the Reddit API to create a machine learning model that will predict the score of specified Reddit posts.

**Data we want to scrape for possible features:**

* Post title
* Post content
* User that made the post
* User data
* Sub Reddit
* Sub Reddit data
* Create date/time

**Possible Libraries and Packages:**

* Reddit API
* Scikit Learn

<https://www.alexa.com/topsites/countries/US>