Gitsome:
Github Recommender Project Proposal

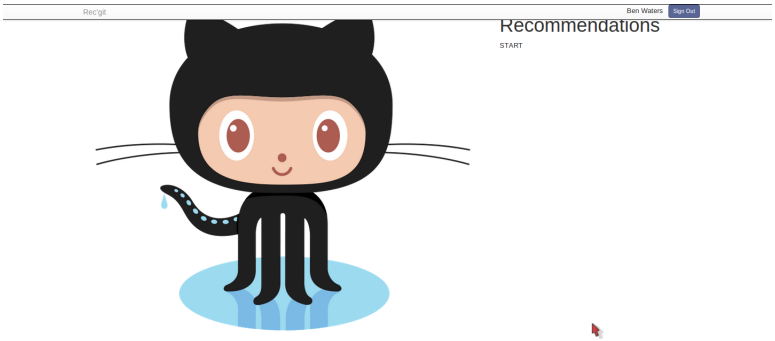
Alejandra Vigil
Benjamin Waters

1 Proposal

1.1 Summary

We will be finding repository recommendations for a user by using k-means algorithm and finding user-similarities using. Github does not currently have a recommendation system implemented. They just look amongst "trending" as highest values within a time frame. A recommendation system is easy to create because a user "stars" code repositories. A star relates to a 1 and not starring a repository relates to 0. The Jaccard coefficient is a useful measure for binary values. We chose to create a website that allows a user to log in with their Github credentials and repositories would be recommended. We chose to use MeteorJS because we are familiar with its technologies.

1.2 Mockup



1.3 Technologies

1. NodeJS
2. MeteorJS
3. MongoDB
4. Redis Server
5. Github v3.0 API `https://developer.github.com/v3/`
6. Recommendation Racoon `https://github.com/guymorita/recommendationRaccoon`

1.4 Process

1. Get user's starred values
2. Load into Raccoon to create binary vector
3. Get user's friend's stars and create binary vector
4. Load into Raccoon to create binary vector
5. Calculate Jaccard coefficient between friends
6. Use K-means to determine recommendations
7. Present recommendations to user through website medium

1.5 Evaluation

We want to look for Jaccard Index values closest to 1. Raccoon returns the top values, but does not have a x is $>$ function.