# Craigslist Rank & Search

## Group

* Abhinav Vora
* Riley Pittman
* Rohith Pottabathni
* Henry Who

## Problem

Searching Craigslist is a tedious task if a user is looking for something specific since CL inventory is a moving target as it sells very fast. The person who spots the item first has an edge. Consequently a user needs to visit the site often and often scour through duplicate postings to find the item they want.

Another related problem is when users are looking for a particular item (e.g., camera) but may not care about the specific model s/he gets. In this scenario it is tedious to sort through various cameras for A. Their criteria (price, location, etc), B. How good the product is relative to others in the category.

We plan to solve both these related problems by allowing the user to provide a wish list of items they are interested in purchasing (including a preference for buying a specific model vs. finding the best listed product, how long do they want to complete the purchase in, etc). The system will monitor postings and email the user when an appropriate listing appears.

## Project Breakdown

### Acquire data from CraigsList

This stage will involve being able to crawl craigslist for postings.

### Extract product categories, price, location, etc. from postings

This stage will involve extracting information from CL postings that will help the system filter user criteria.

### Extract product reviews

Product reviews for various products are available from search engines and sites such as Amazon.com product pages. In this stage we will extract reviews and attemp to classify positive/negative reviews for a particular product and compare it against other products in that category so that we can use that in our ranking function.

### Rank based on users preferences and product reviews

This stage will involve buildling a ranking function that will use data learnt from the earlier stages to sort through multiple listings. If possible we will also attempt to predict how often popular products are listed so that the ranking function can recommand the user to wait for a better product that is not currently listed.

### Build a UI

We will build a basic search interface. We will also provide product reviews in the form of a mashup in the results page.

## Timeline

1. Oct 29 – Complete crawlers
2. Nov 17 – Complete all extraction
3. Dec 3 – Complete ranking algorithm, and UI.