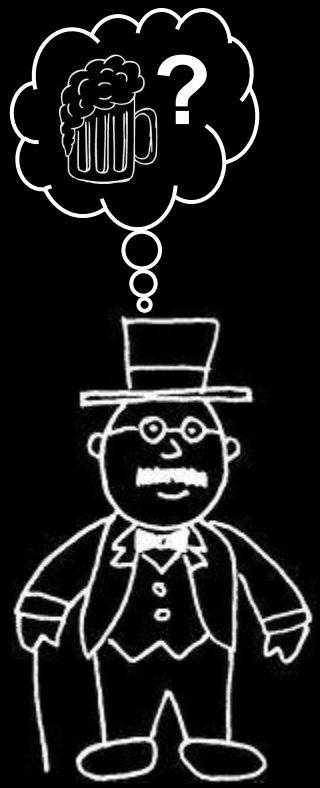
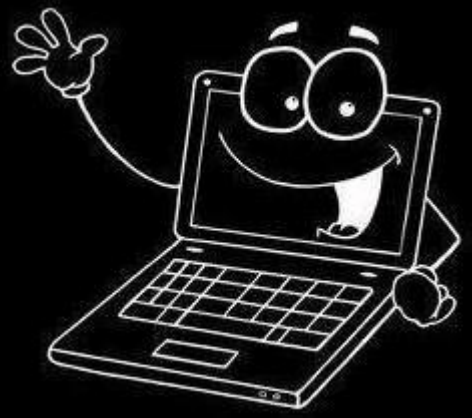
I want to drink new obscure craft beer that I will like, but it can be hard to find...

BeerInsight

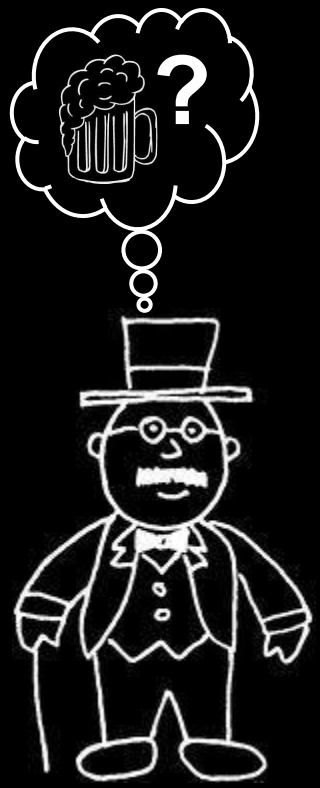
Alex Tronchin-James

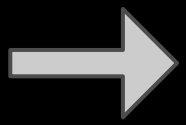


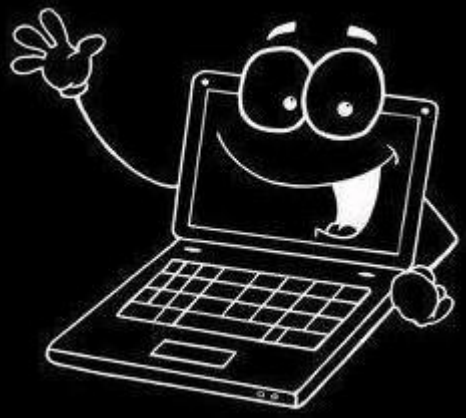
INSIGHT



user reviews





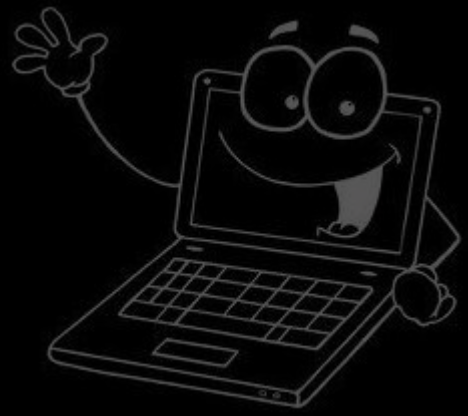


other user reviews

Beeradvocate™

SQLite

user reviews

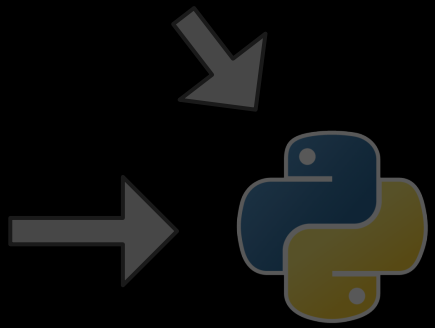


other user reviews

Beeradvocate™

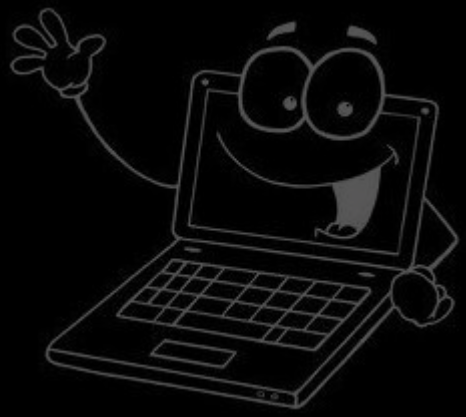
SQLite

user reviews



Cython

collaborative filtering recommender system



The recommender is trained using the BeerAdvocate.com database

- 1,586,614 reviews, text, 0-5 star ratings
- 56,857 unique beers
- 33,388 unique users
- Use these to build the **review matrices** Y and R

- Obtained from a friendly blogger (thanks!)
- Previously available from Stanford

Hand-coded collaborative filtering

- regularized least squares linear regression with L-BFGS optimization
-  used to return a sparse component of an otherwise dense matrix multiplication

$$< X\Theta^T, R >= Y$$

- **pull request** in progress to include this to SciPy.org

other user reviews

location data

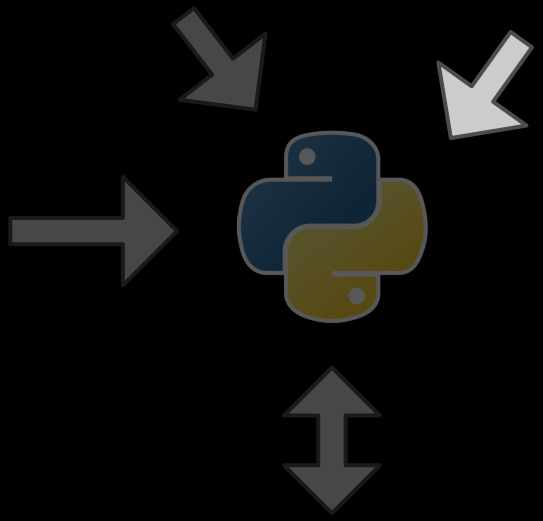
Beeradvocate

ratebeer

SQLite

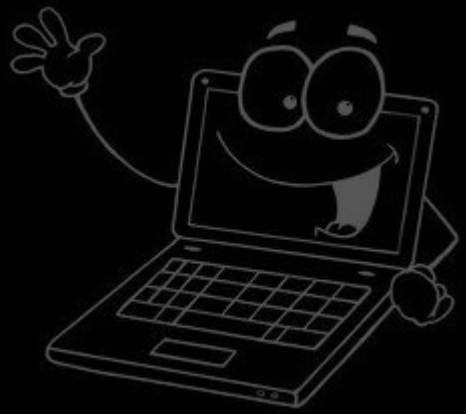
Requests

user reviews



Cython

collaborative filtering
recommender system



Filter results by location data scraped from RateBeer.com

- Names differ between databases

```
(48, u'Port Brewing/Lost Abbey|Port Brewing High Tide  Fresh Hop
IPA')
(46938, u'Port Brewing Company / Pizza Port|High Tide Fresh Hop IPA')
```

- … but can be matched by **n-gram similarity**
- mismatches are easy to **clean**

other user reviews

Beeradvocate

SQLite

location data

ratebeer

Requests

front end

HTML CSS

JavaScript

jQuery
write less, do more.

user reviews

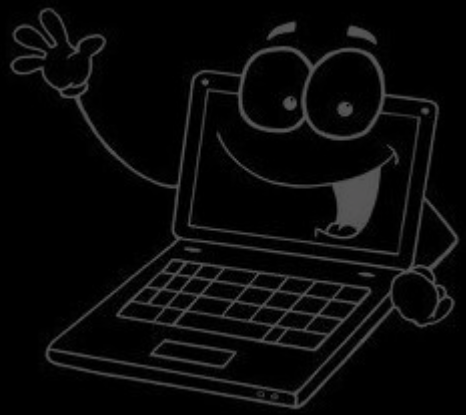
Flask
web development,
one drop at a time

gunicorn

powered by
amazon
web services™

Cython

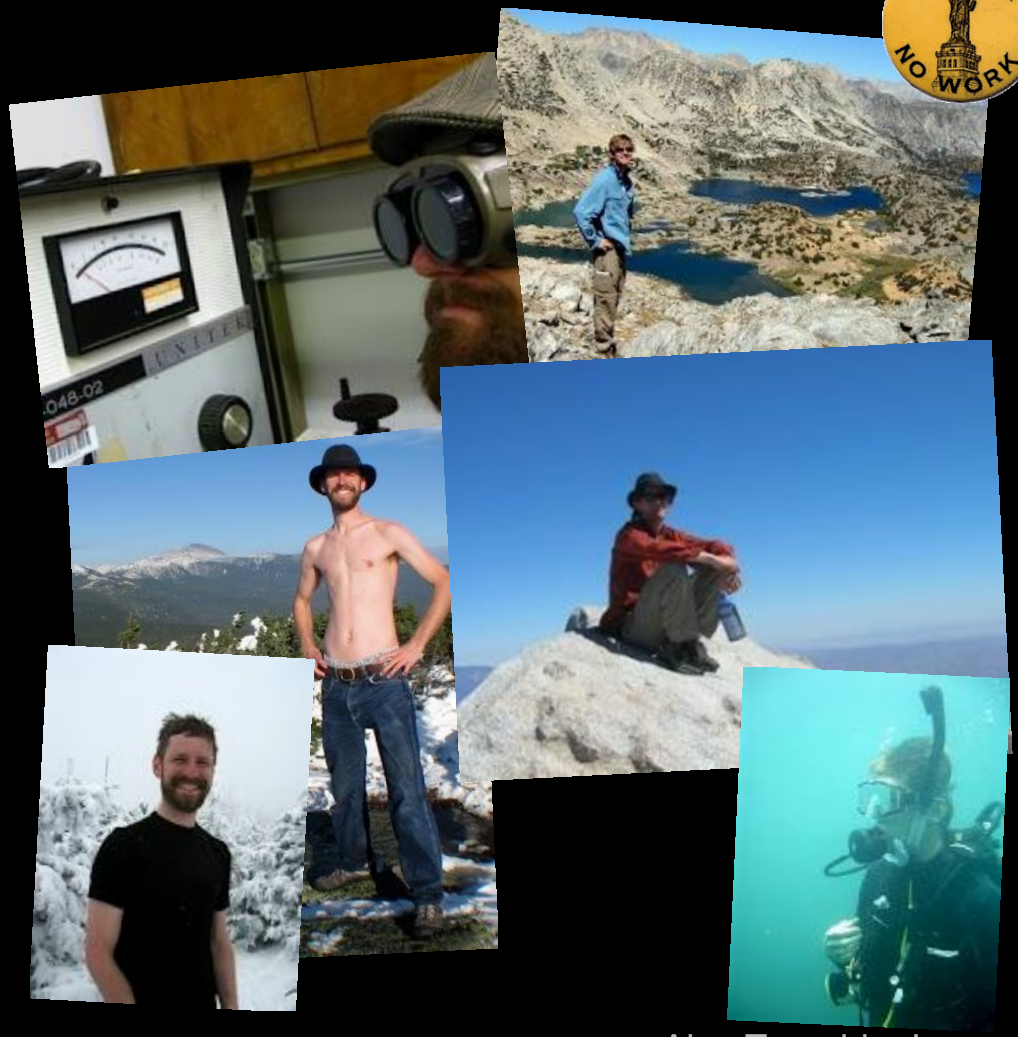
collaborative filtering
recommender system

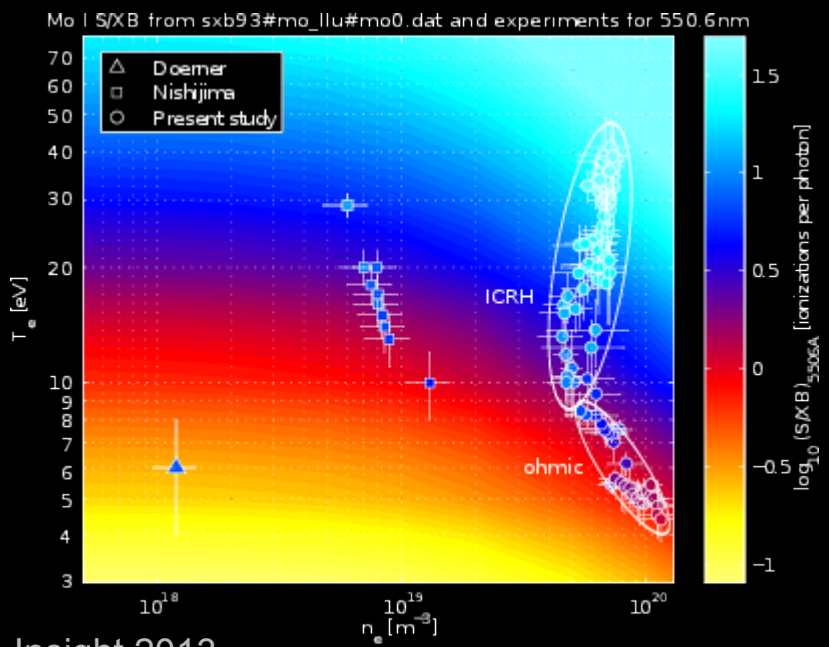


About me

Alex Tronchin-James

UCSD Jacobs





Tuning and validation

- Applying learning curves to tune number of features and regularization parameter

  … TODO