

# Yhelper: FAKE Review Detection & Applications

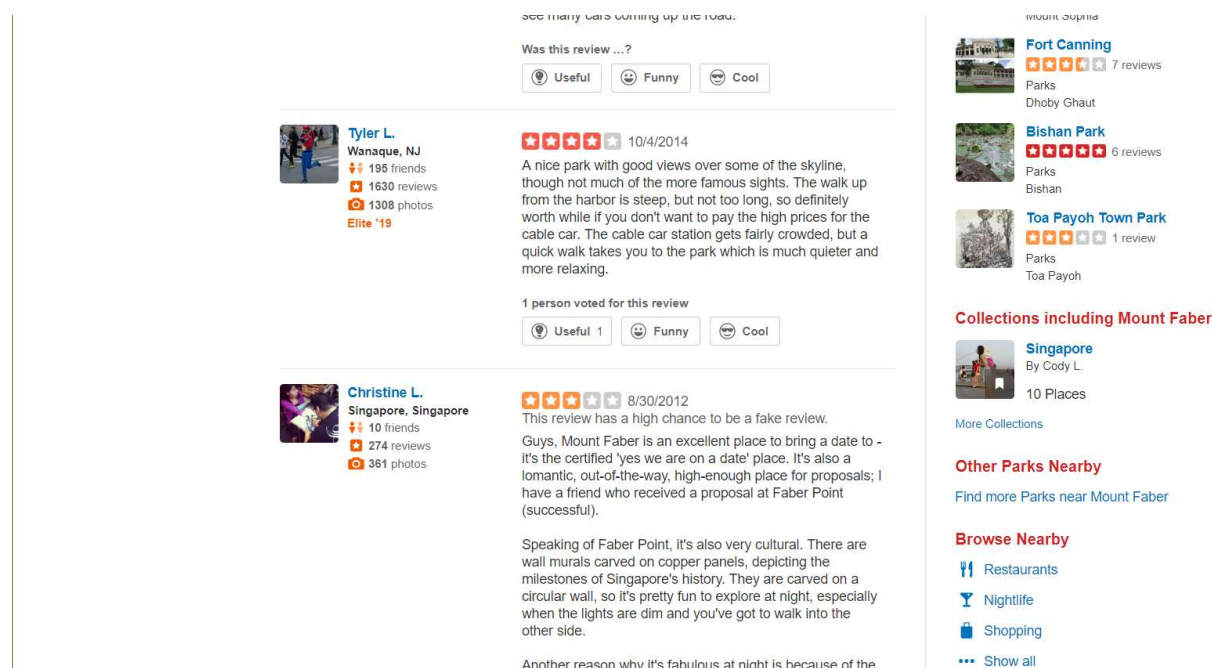
(C) 2019 Lee Sze Wei (A0180783X) & Wang Huaqiang (A0196556M)

- [1. Description](#)
- [2. Technology Used](#)
- [3. Targeted audience](#)
- [4. Program Structure](#)
- [5. Usage](#)
  - [5.1. Local version](#)
  - [Website](#)
  - [Browser Extension](#)

## 1. Description

This project aims to detect fake reviews on Yelp from sentiment features and user's social features. The form of this application will be a browser extension. When users browse reviews and grades on Yelp, the extension will point out potential fake reviews.

After enabling this extension on your web browser, the Yelp website will be like this:



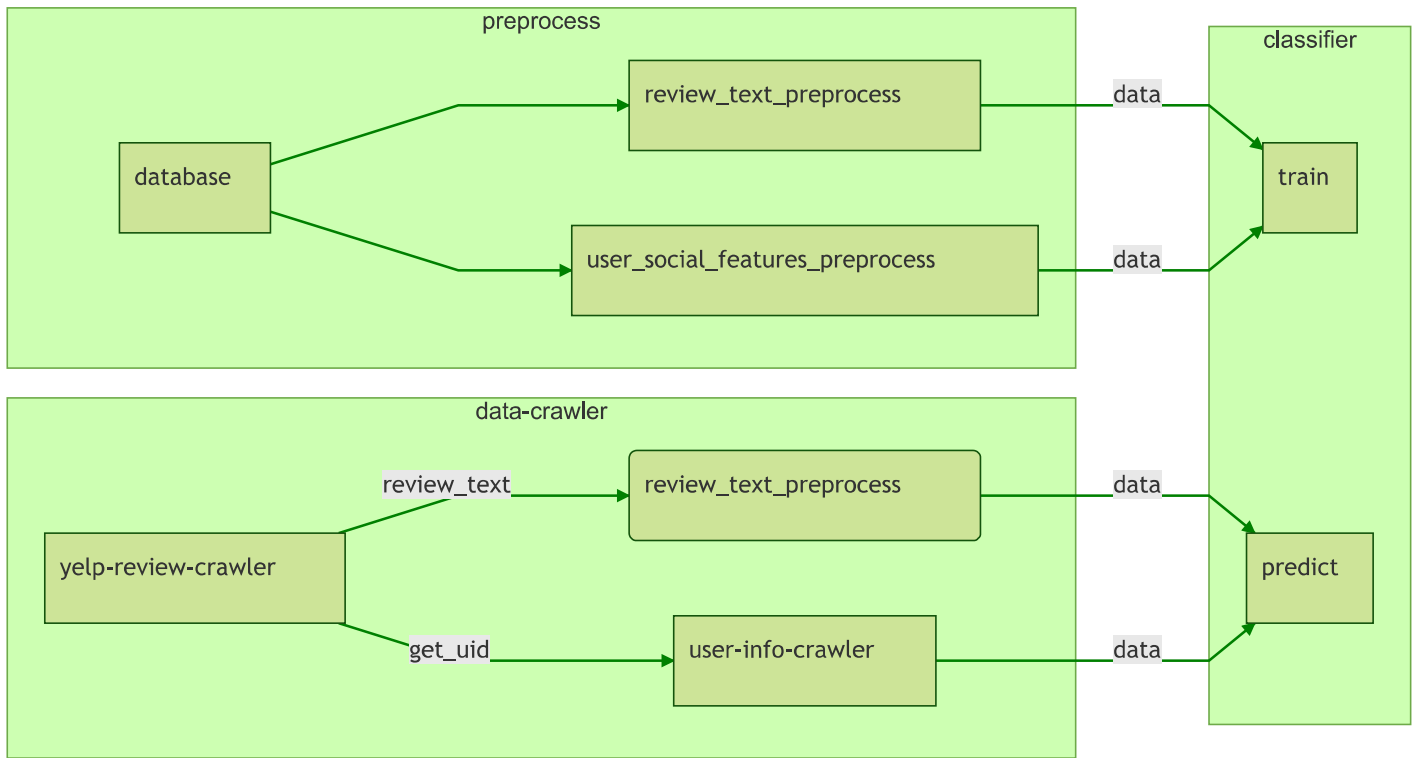
## 2. Technology Used

- Sentiment analysis.
- Fake user detection.
- Profile level behavioral analysis.

## 3. Targeted audience

People who wants to know about a business from its remarks and grades on Yelp.

## 4. Program Structure



TODO: File introduction and comments.

Preprocess:

- [yelpdbreader.py](#)
- `preprocess_db.py`
- `preprocess_review.py`
- `vader_awmod.py`

Runtime data crawler:

- [pageparser.py](#)
- `user_crawler.py`

Classifier:

- [classifier.py](#)

## 5. Usage

### 5.1. Local version

The local version requires `python2.7` with `sklearn`, `xgb`, and `nltk`. You can install them via `pip`.

To do preprocess, run:

```
python preprocess_db.py
python preprocess_review.py
```

Warning: `preprocess_review.py` will take long to finish.

To train classifier, run:

```
python classifier.py
```

The trained model will be saved in `model11.data` and `model12.data` .

To detect fake reviews from a yelp page, call (in python):

```
import pageparser
pageparser.detect_fake_review(url)
```

You can run:

```
python pageparser.py
```

To run a built-in fake reviews detection unit test.

For debug: the result will be saved in:  
./data/preprocessed/yelppage.json  
./data/preprocessed/new\_yelppage.htm

## Website

TBD

## Browser Extension

TBD

---

TODO: