Zach Shim

Capstone Proposal

CSS 490

**Description:**

Fake reviewers promote false advertisement, which negatively affects the business of sellers and the experience of customers.

**Requirements:**

* Improve code readability and documentation.
* Importing CSV/JSON files to a database
  + Currently, the project uses a flat file format for storing and manipulating data. Information will now be stored inside a relational database, using SQLite as the DBMS.
* Deploy project to a cloud server.
  + The website will include four screens:
    - The index page will allow users to search the static database.
    - Another page will allow users to paste links to amazon products, allowing for dynamic fake score generation.
    - An about page featuring more information about the project.
* Allow user to paste a link to an Amazon review for analysis.
  + Allows user to see real-time data instead of relying on a static database.
  + Scraping directly from an amazon link given a ASIN key.
* Add another machine learning methodology
  + Possible implementation: logical regression or Linear SVM

**Tech Stack**

* Backend:
  + Django
  + Python
  + SQLite3
* Frontend:
  + HTML5
  + CSS
  + Javascript

**Dependency Packages**

* + beautifulsoup4
  + django
  + matplotlib
  + mpld3
  + nltk
  + numpy
  + pandas
  + py2app
  + pyculiarity
  + rpy2
  + regex
  + requests
  + scipy
  + scikit-learn
  + sqlalchemy

to download all packages, simply run the ‘’dependencies.py’’ script in the base directory of the project.

***python3 dependencies.py***

**Static Dataset**: <http://jmcauley.ucsd.edu/data/amazon/>

Table

Description automatically generated

5-core .json files must be downloaded from the website and sanitized/inserted into the database. See the /API Documentation/Commands/ to see how to insert data.