

First of all, I am really sorry that I can only deliver .py files in my homework.
Anaconda doesn't support jieba for python 3.0 but my homework needs it.
All .py files can be run in python3 environment with current package installation like
sklearn, jieba..etc.

The program files structure is described like below.

Pttmobile/: It is the folder of scrapy spider project. It can crawl ptt mobile site to
fetch information from there. The main file of spider code is

pttmobile->pttmobile->pttmobile.py

The scrapy items configuration is at

pttmobile->items.py

The wordinganalysis folder has the below files.

1. Brand_hot_ratio.py: It will show a graph about the amount of hot topics for each brand in ptt mobile
2. Handledata.py: It uses jieba to extract key word and remove stop words. Then it display hot keywords for each brand by drawing graphs.
3. Prediction.py: It provides a example about how to classify an article by keyword prediction
4. User_preference.py: Analysis which group of users prefer for each brand. It will list top 10 lovers for each brand and also gives an example about a user's favorite level for each brand.
5. Train_data.json: It is the data set file which grabbed from ptt mobile.
6. Stop_word.txt: The stop words I use in my homework.