<https://www.anaconda.com/distribution/>

python 3.6

<https://blog.naver.com/vega2k>

pip.exe → npm과 비슷

ex) pip install requests

Ananconda 포함된 open source

web scraping  (수집)

: requests  → http통신

  beautifulsoup4 → http로 가져온 데이터 파싱하여 데이터 추출

  selenium  → 위의 두가지를 합쳐진듯 단점:느리다

(분석)

: numpy  → 난수발생이나 데이터 생성

  pandas  → data를 sql처럼 데이터를 처리할때 사용

 matplotlib  → 그래프로 표시

 seaborn  → 다양한 그래프 표시

 scikit-lean

개발할때 사용된 tool

: jupyter notebook

<https://www.kaggle.com/>

데이터 분석용 샘플 데이터 .. 학습용 데이터 제공해줌

tensorFlow

pytorch

http테스트 site <http://httpbin.org/>

<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

BeautifulSoup 설명

select() 🡪 다건

select\_one() 🡪 다건중 한 개

: tag명, attribute사용 가능

: CSS Selector 규칙을 적용해서 tag를 가져올 수 있다.

find\_all()

find()

: tag명, attribute

<https://regexr.com/> 정규표현식 연습

train

데이터(X\_train) /레이블(y\_train)

Test

데이터(X\_test) / 레이블(y\_test) (에측할 값)

F(x) = y