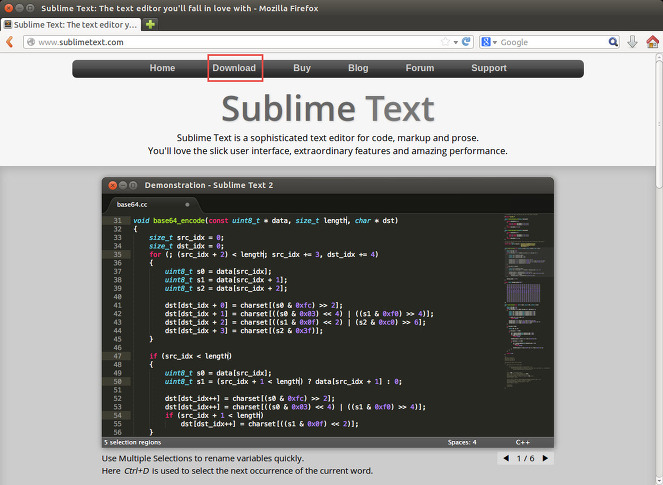
2016/07/04 ~ 2016/07/05

**sublime text 설치 방법**

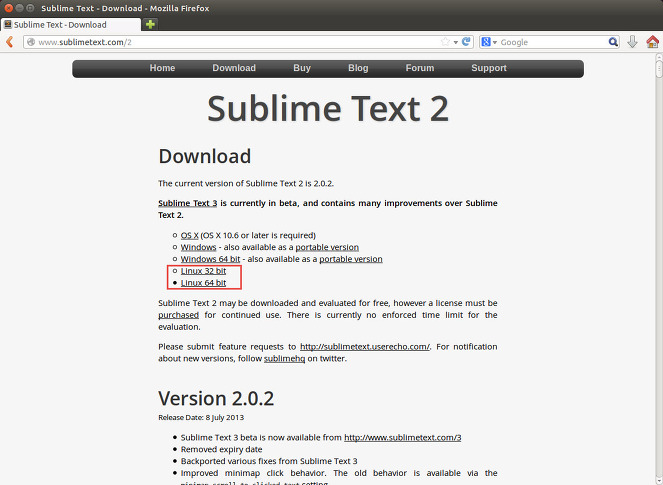
**1. Sublime Text 다운로드**

 Sublime Text 사이트로 접속 하도록 한다. 사이트 주소는 [**"www.sublimetext.com"**](http://www.sublimetext.com/)  이다.



[그림 1 - Sublime Text 사이트]

 [그림 1]과 같이 위 상단 다운로드(Download)메뉴를 클릭해서 다운로드 페이지로 이동한다.

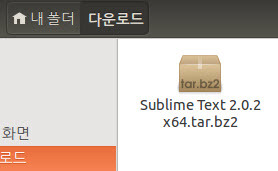


[그림 2 - Sublime Text 2 다운로드 화면]

 [그림 2]와 같이 다운로드 페이지에서 해당 PC에 맞는 파일을 선택해서 설치파일을 다운로드 받는다.

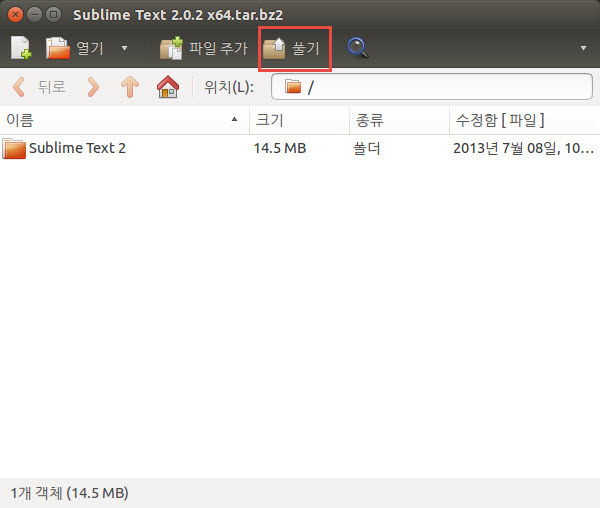
**2. Sublime Text 설치**

 파일을 다운로드 다 받았다며, 파일이 받은 폴더로 이동해 보자. 보통은 다운로드 폴더에 받아져 있을 것이다.



[그림 3 - Sublime Text 압축파일]

 현재 압축되어 있기때문에 압축 풀도록 하자. 해당 파일을 더블 클릭하도록 한다. 그럼 아래와 같이 "압축 관리자"창이 열릴 것이다.



[그림 4 - 압축 관리자 화면]

 [그림 4]와 같이 창이 열리면 "풀기"버튼을 클릭해서 압축을 풀도록 한다. 압축을 풀면 해당 압축파일에 위치한 경로에 압축이 풀린 폴더가 나타날 것이다. 폴더 이름이 "Sublime Text 2"라고 생성 될 텐데, 일단 공백을 다 붙어서 이름을 변경해준다. 다시말해 "SublimeText2"라고 이름을 변경한다. 그리고 해당 폴더를 "opt"폴더로 이동 시키도록 한다. 커맨드 창을 열고 아래와 같이 커맨드를 입력 한다.

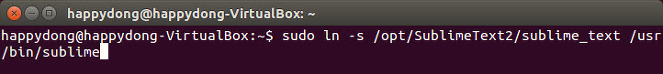
|  |
| --- |
| sudo mv ~/다운로드/SublimeText2 /opt/ |

http://cfile3.uf.tistory.com/image/2539524653A3D0362923E2

[그림 5 - 폴더 이동 커맨드]

해당 폴더에서 파일이 없어졌다면, opt폴더로 이동해서 잘 이동 되었는지 확인해 본다. 폴더가 잘 이동 되었다면, 이제는 "SublimeText2" 폴더 밑에 "sublime\_text"을 실행 시킬수 있도록, "/usr/bin"폴더에 심폴을 만들어 줄 것이다. 심폴 이름은 간단하게 "sublime"이라고 하자.

|  |
| --- |
| sudo ln -s /opt/SublimeText2/sublime\_text /usr/bin/sublime |



[그림 6 - sublime\_text 심볼 연결]

위와 같이 심볼까지 연결이 끝났다면, 이제는 우분투에 메뉴에 추가 하도록 하자.

**3. 우분투 메뉴 추가**

 우분투에서 Sublime를 검색하면 실행 아이콘이 나타나지 않을 것이다.



[그림 7 - Sublime 검색 화면]

 [그림 7]과 같이 실행파일이 나타나지 않는 것이 정상이다. 이제 실행파일이 검색에서 걸리수 있도록 등록을 해주자.

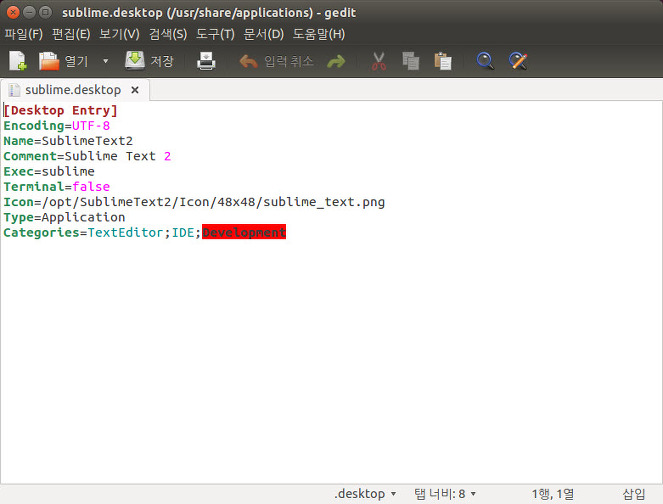
|  |
| --- |
| sudo gedit /usr/share/applications/sublime.desktop |

http://cfile23.uf.tistory.com/image/2732EF4653A3D03731B5E5

[그림 8 - gedit 열기]

[그림 8]과 같이 커맨드를 입력하면, 편집창이 열릴 것이다. 그럼 아래와 같이 입력 하도록 하자.

|  |
| --- |
| [Desktop Entry]   Encoding=UTF-8   Name=SublimeText2   Comment=Sublime Text 2   Exec=sublime   Terminal=false   Icon=/opt/SublimeText2/Icon/48x48/sublime\_text.png   Type=Application   Categories=TextEditor;IDE;Development |

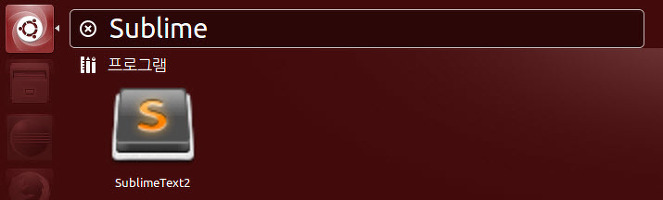


[그림 9 - 데스크탑 엔트리(Desktop Entry) 정의]

 [그림 9]와 같이 입력완료 되었다면, 저장하고 창을 닫도록 한다. 그리고 다시 검색창에서 "Sublime"를 검색해 보자.

경로를 설정할 때 직접 경로 폴더로 들어가서 파일이 존재하는지 확인한다.

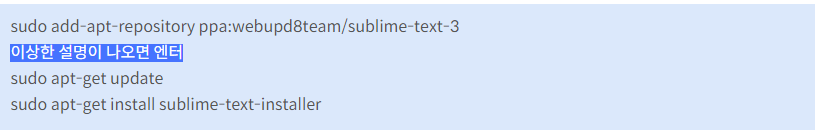
만약 저 경로가 틀리다면 맞는 경로로 수정해준다.



[그림 10 - Sublime 검색 화면]

[그림 10]과 같이 검색에 잘 나타난다면 설치및 설정이 완료 된 것이다.

만약 위의 방법이 안 된다면 아래 방법을 선택한다.



실행방법

터미널에서 subl & [Enter]

유빈언니가 해 놓은 파일에서 단어 검색하여 추출하기를 실행해봄

일단 세 개의 파일이 필요함

1. modified\_topic\_modeling\_and\_frequency.py

2. test\_pdf.py

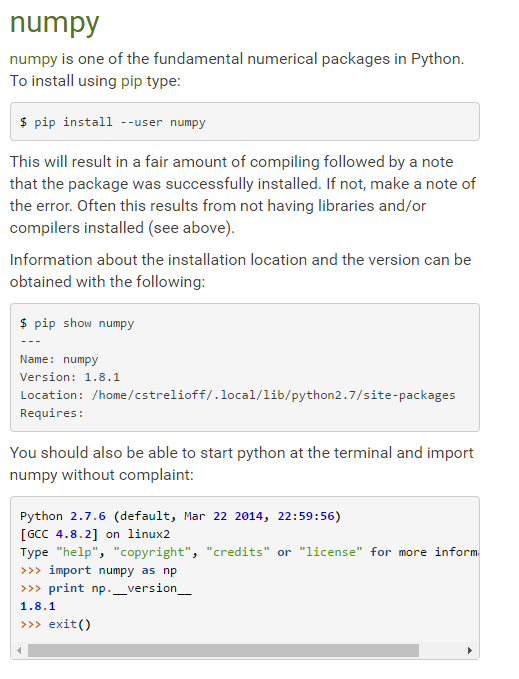
3. ui.py

ui.py 를 실행하면 됨

> python ui.py [Enter]

실행하기 전 설치해야 할 모듈들이 많음

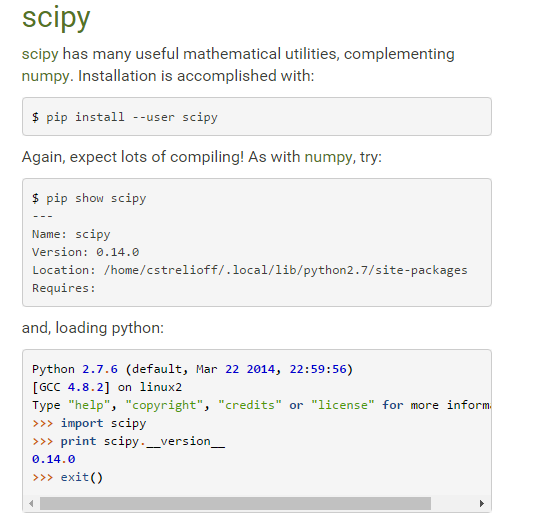
1. numpy



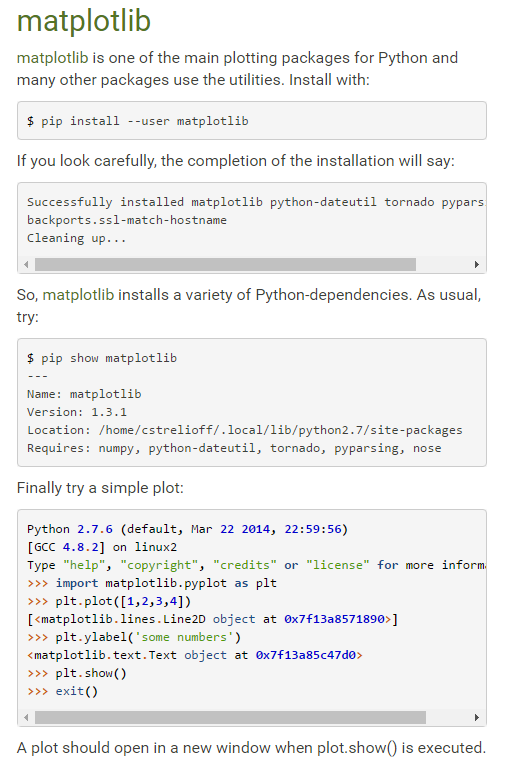
이런식으로 나오면 성공,

install 했을 때 또는 show numpy 했을 때 오류가 나면 다시 install 해보거나 인터넷 연결을 확인.

2. scipy



3. matplotlib



matplotlib를 설치하는데 이것만 되지 않았다.

sudo pip install --upgrade setuptools

를 한 다음 다시 시도를 해보았지만 또 실패.

pip install -U pip

를 한 다음 다시 시도. -> 실패

sudo apt-get build-dep matplotlib

시도 -> 안됨

pip를 지우고 다시 깔았음

apt-get remove python-pip

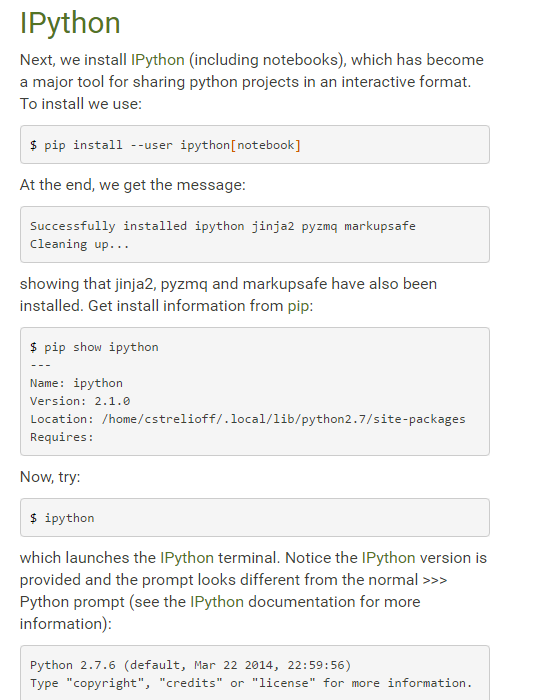
apt-get install python-pip

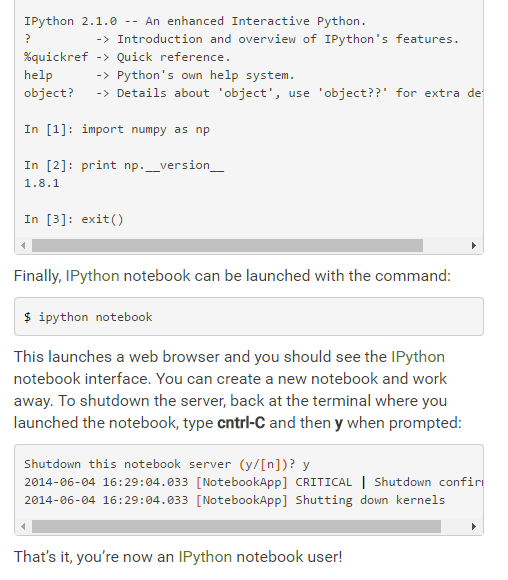
pip install --user matplotlib

4. sympy

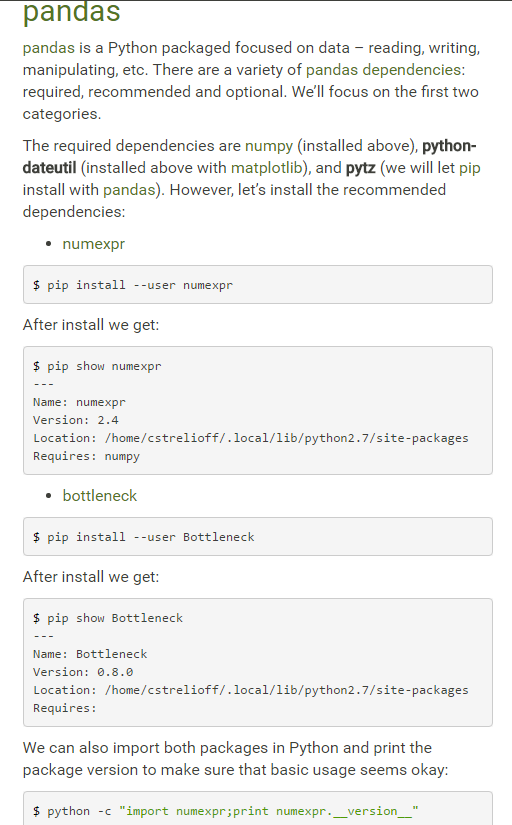


5. ipython

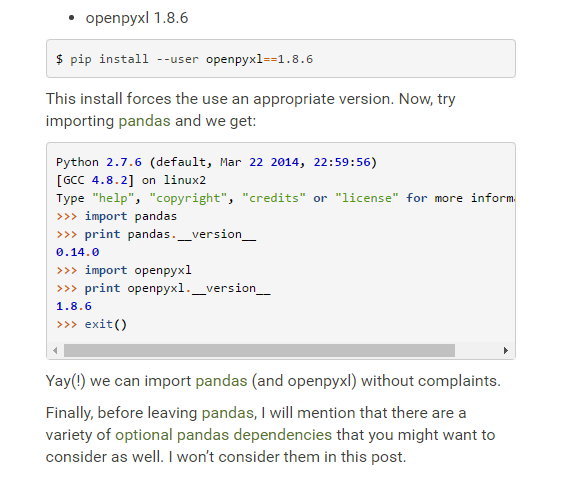




6. pandas







출처 : http://chrisstrelioff.ws/sandbox/2014/06/04/install\_and\_setup\_python\_and\_packages\_on\_ubuntu\_14\_04.html

(그 외 다른 모듈 다운 받는 건 사이트에 있음)

7. pdfminer

> python에서 pdf를 html / text 파일로 변환

sudo pip install pdfminer

8. scikit

sudo pip install -U scikit-learn

9. mallet

sudo apt-get update

sudo apt-get install ant mercurial openjdk-7-jdk -y

hg clone http://hg-iesl.cs.umass.edu/hg/mallet

cd mallet

ant

했는데 만약 안되거나 파일 경로에 파일이 없을 때

curl --remote-name http://mallet.cs.umass.edu/dist/mallet-2.0.7.tar.gz

tar zxf mallet-2.0.7.tar.gz

다시 해주면 됨

아마 밑에 방법이 확실한 듯

10. ipython

sudo apt-get install ipython-notebook

파이썬 패키지 설치 유틸(pip)도 설치하는 것이 좋음.

pip 설치

sudo apt-get install python-pip

notebook 실행

ipython notebook

[IPy Notebook 사용]

IPy Notebook 은 아래 주소로 접속해서 사용

http://127.0.0.1:8888/

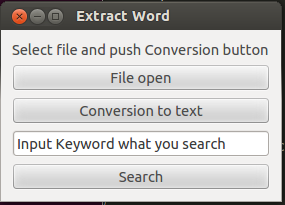
솔직히 이건 지금 상황에서 잘 모르겠음

나중에 유용하게 쓰인다던데..아직.. 잘...

실행 방법

1. python ui.py

아래 창이 뜬다.



2. File Open 을 하면



이런 창이 뜬다. 현재는 pdf파일이 한 개 존재한다.

pdf 파일을 open한다.

(파일 이름에 띄어쓰기가 있으면 conversion to text에서 에러가 난다. 되도록이면 띄어쓰기를 없애거나 \_를 써주는 게 좋다.

3. conversion to text

버튼을 클릭하면 text파일로 바꿔준다.

처음에는 문서 전체를 text 파일로 변경한 다음에 keyword를 추출한다. 아마 이게 csv 파일로 저장되는 것 같다.

topic-keys-report.txt 파일도 단어를 추출하긴 하는 것 같은데 csv 파일이랑은 또 다르다.

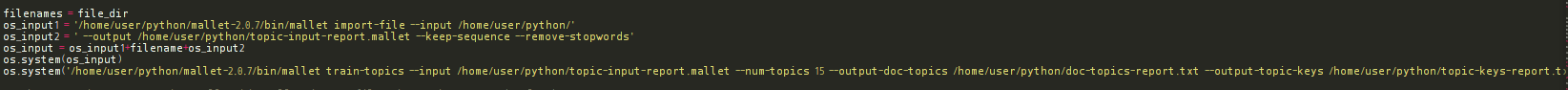
어쨋든 csv 파일이 주요 keyword를 추출하는 파일이라고 한다.

modified\_topic\_modeling\_and\_frequency.py 거의 마지막 쪽에 보면 new.csv 라 저장하라고 되어있는 것 같다. 그래서 이 과정이 완료되면 /home/user/python에 new.csv 파일이 만들어 진다. 또한 터미널 창에는 

이런 식으로 출력이 되면 성공적으로 conversion 한 것이다.

만약 오류가 뜬다면 path를 잘 확인하길 바란다.

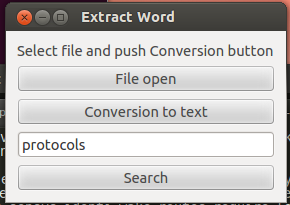
path를 변경하는 것에서 이틀 정도 소비 한 것 같다. 뭐가 어떤 path인지 몰라서...흑흑...역시 googling과 오류문을 분석한 결과 해결했다.



애 먹었던 path..

그것도 아니면 오류문을 분석하면 될 것 같다.

4. input word



찾고 싶은 단어를 입력하고 Search를 누른다.



그러면 위와 같은 창이 뜨는데 입력한 단어가 있는 파일들의 list를 보여준다고 한다.

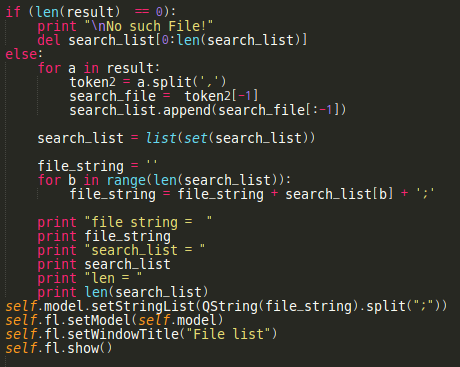
언니 말로는 만약 pdf 파일이 여러 개 있을 때 파일을 한 번이라도 open한 파일들 중에서 입력한 단어가 있는 파일들의 목록을 위의 창에 보여준다고 한다.

new.csv 파일을 들어가보았더니 단어 추출된 것이 append로 추가가 되는 것이었다. 그래서 한 번이라도 열었던 적이 있으면 기록이 남아있어서 계속 탐색이 가능하다.

**코드 수정 사항**

search를 하는데 만약 입력한 단어가 있는 파일이 검색되지 않을 때 file list창에 그 전 기록이 남아있었다.

터미널 창에는 no such file!이 출력이 되지만 terminal을 보지 않으면 해당 단어가 검색된 걸로 인식할 수 있어 수정하였다.



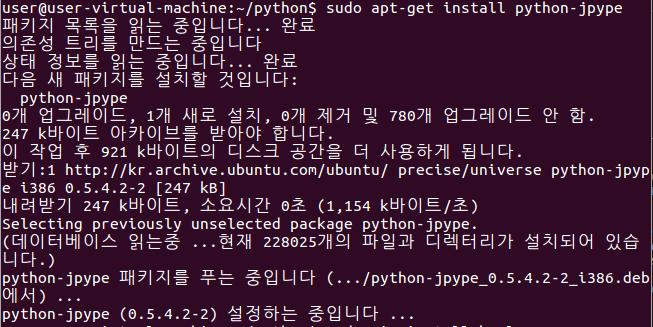
1. del search\_list[0:len(search\_list)]

이 명령문은 리스트를 초기화 하는 것이다. "del"이 핵심인 듯

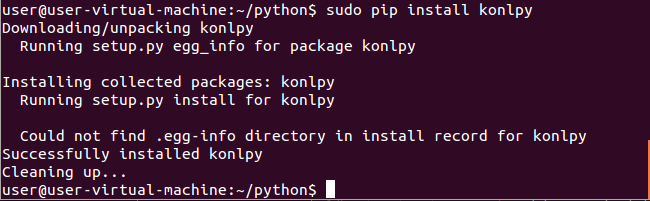
2. 두 번째 박스는 원래 else문에만 해당하는 것 이였는데 if 문에도 해당되도록 앞으로 back tab을 해주었다.

**한국어 처리**

한국어 처리를 위해 jpype를 설치한다.



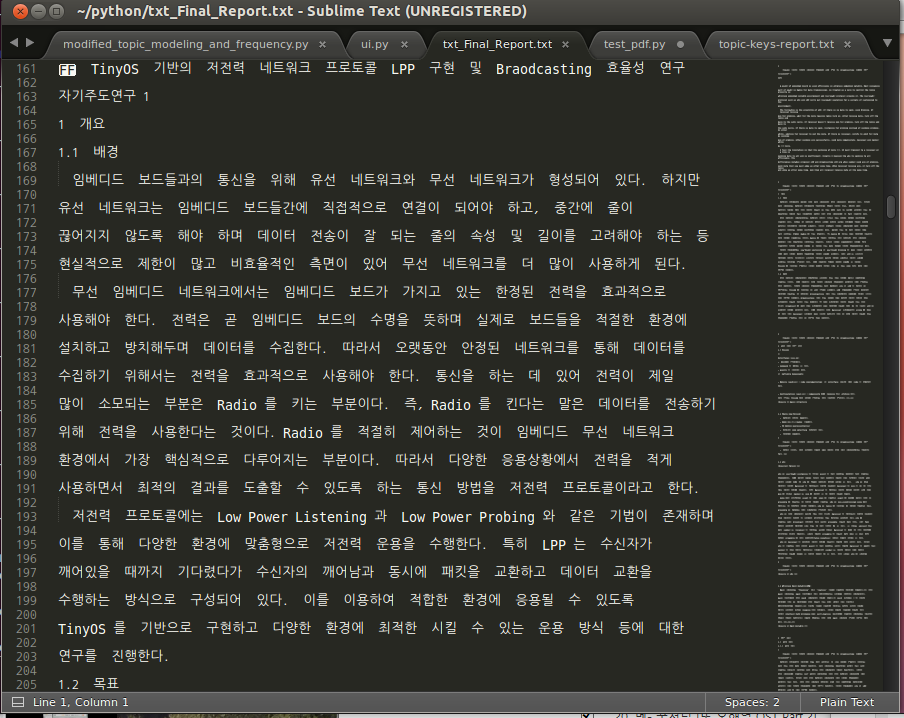
설치가 제대로 되었다면 konlpy를 설치한다.

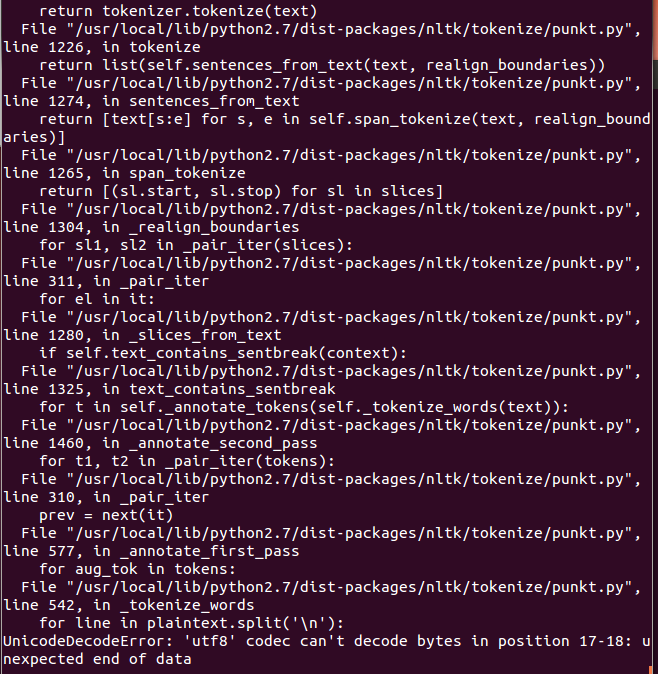


설치 완료

설치만 하고 한글로 된 pdf파일을 text파일로 변환해 보았다.

text 파일로 변환이 되긴 하는데 text파일로만 변환이 되고 단어 추출 등의 행위는 일어나지 않았다.





코덱 문제가 발생하는 듯 하다.

**mecab-ko-dic설치하기**

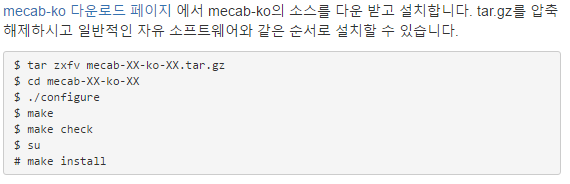
출처 : https://bitbucket.org/eunjeon/mecab-ko-dic

mecab-ko-dic 은 오픈 소스 형태소 분석 엔진은 mecab을 사용하여, 한국어 형태소 분석을 하기 위한 프로젝트이다. 말뭉치 학습과 사전 목록 일부는 21세기 세종계획의 성과물을 사용하였다.

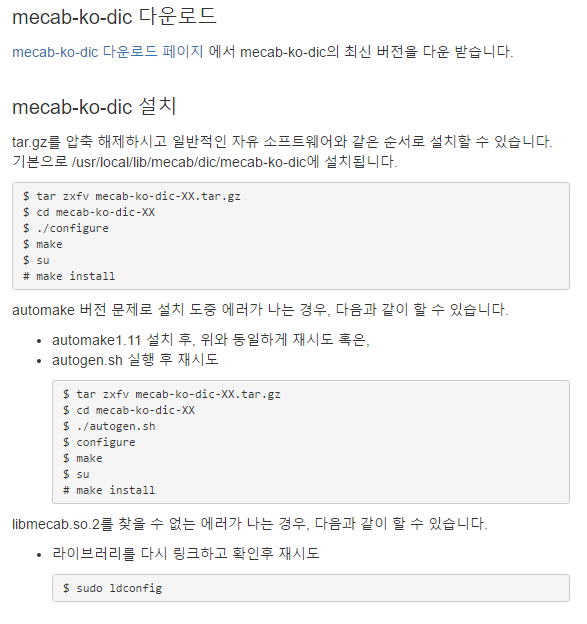
mecab-ko 설치

1. https://bitbucket.org/eunjeon/mecab-ko/downloads

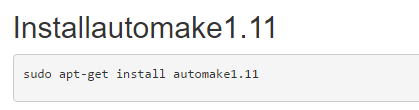
에서 최신 버전을 다운 받아 압축을 푼다.



mecab-ko-dic 다운로드 페이지 : https://bitbucket.org/eunjeon/mecab-ko-dic/downloads



automake1.11 설치방법은 다음과 같다.



(libmecab.so.2를 찾을 수 없는 에러가 나서 sudo ldfonfig 실행 후 다시 make 해주었음)

mecab 실행

-> mecab -d /usr/local/lib/mecab/dic/mecab-ko-dic 을 치면



위와 같이 실행 되어야 하는데 나는 실행이 안된다....

왜죠...ㅠㅠ

**20160706**

TensorFlow 설치하기

TensorFlow : 기계 학습과 딥러닝을 위해 구글에서 만든 오픈소스 라이브러리이다. 데이터 플로우 그래프방식을 사용하였다.

1. CUDA7.5 설치

<https://developer.nvidia.com/cuda-downloads> 여기서 다운로드



난 mac도 아니고 우분투 32bit라 이걸 설치할 수가 없다.. 어찌해야 하는가..ㅠ;;

일단 회사 서버로 접속해서 서버상에서 하는 걸로 코멘트 받았다.

그 전에 어제 한 거를 이어서 해야 하나 싶어서 그냥 해볼까 한다.

강의를 들어야 하나..

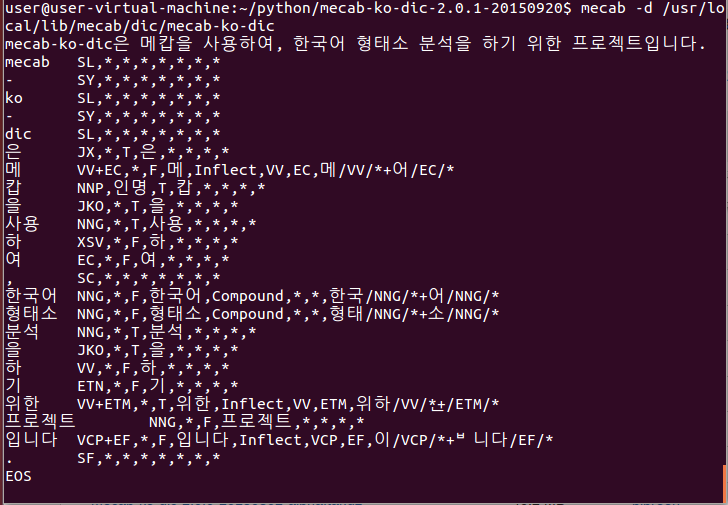
일단 한국어 모듈 mecab-ko-dic을 설치했다.

어제 실행한 결과 안 된다고 했는데 안 된 게 아니었다. ㅋㅋㅋ

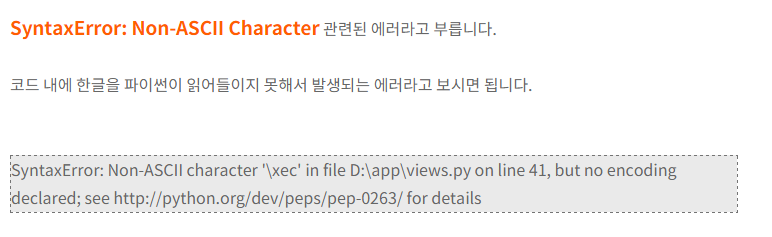
그냥 명령어 치고 그 다음에 한국어를 바로 치면 되는거였음

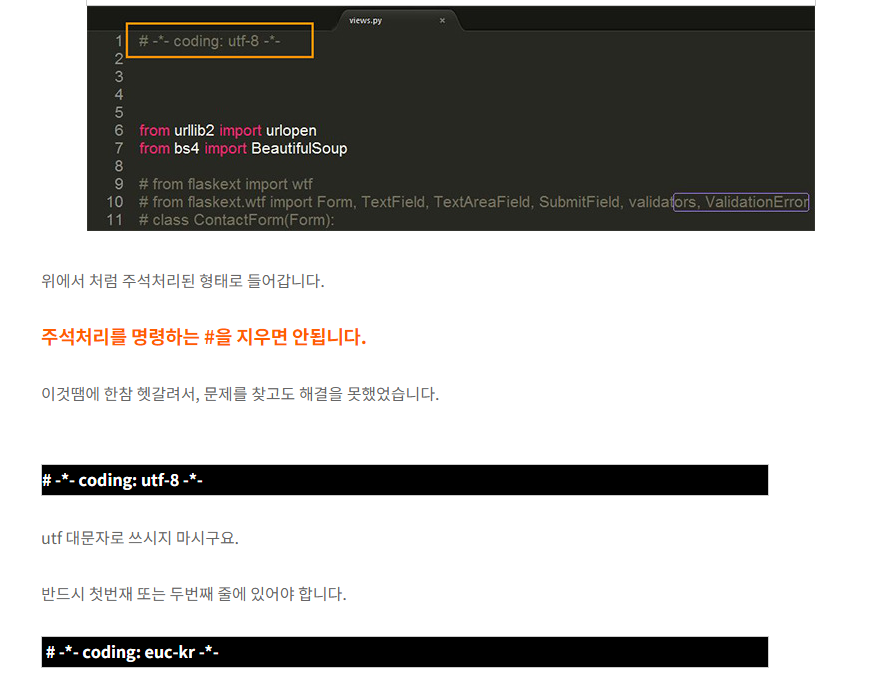
나는 무슨 치라는 명령어 나올 때까지 기다리고 있었는데..;;;;휴..

암튼 아래와 같이 성공

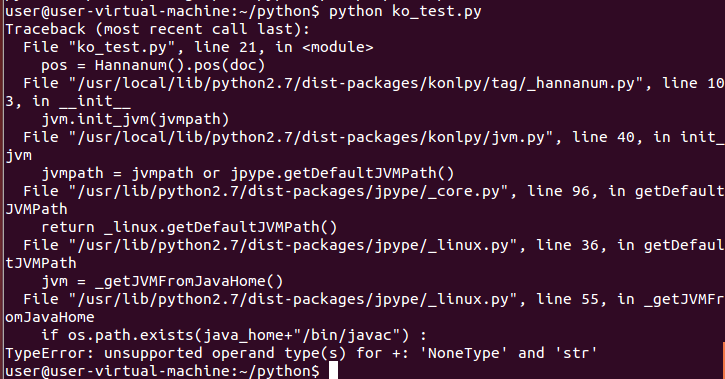


python ko\_test.py 실행->오류





두 번째 오류



참고자료 :

https://www.lucypark.kr/courses/2015-dm/text-mining.html

해결 못 하겠다..ㅠㅠ

서버에 대해 알려주심!!

**서버**([영어](https://ko.wikipedia.org/wiki/%EC%98%81%EC%96%B4): server)는 [클라이언트](https://ko.wikipedia.org/wiki/%ED%81%B4%EB%9D%BC%EC%9D%B4%EC%96%B8%ED%8A%B8)에게 [네트워크](https://ko.wikipedia.org/wiki/%EB%84%A4%ED%8A%B8%EC%9B%8C%ED%81%AC)를 통해 서비스를 제공하는 [컴퓨터](https://ko.wikipedia.org/wiki/%EC%BB%B4%ED%93%A8%ED%84%B0)(server computer) 또는**프로그램**(server program)을 말한다. 특히, 서버에서 동작하는 소프트웨어를 **서버 소프트웨어**(server software)라 한다. 주로 [리눅스](https://ko.wikipedia.org/wiki/%EB%A6%AC%EB%88%85%EC%8A%A4) 등의 [운영 체제](https://ko.wikipedia.org/wiki/%EC%9A%B4%EC%98%81_%EC%B2%B4%EC%A0%9C)를 설치한 대형 컴퓨터를 쓰지만, 그렇지 않은 경우도 있다.

서버는 프린터 제어나 파일 관리 등 네트워크 전체를 감시·제어하거나, 메인프레임이나 공중망을 통한 다른 네트워크과의 연결, 데이터·프로그램·파일 같은 소프트웨어 자원이나 모뎀·팩스·프린터 공유, 기타 장비 등 하드웨어 자원을 공유할 수 있도록 도와주는 역할을 한다.

서버는 사용자(클라이언트)의 요청에 의하여 서비스를 하는데 이와 같이 구성된 시스템을 [클라이언트-서버](https://ko.wikipedia.org/wiki/%ED%81%B4%EB%9D%BC%EC%9D%B4%EC%96%B8%ED%8A%B8-%EC%84%9C%EB%B2%84) 시스템이라고 하며, 이는 하나 이상의 응용 프로그램을 상호 협력적인 환경에서 운용하는 분산처리 형태를 의미한다. 즉, 서비스를 요청하는 클라이언트와 클라이언트의 요청을 처리하는 서버와의 협동작업을 통해서 사용자가 원하는 결과를 얻는 처리방식이 클라이언트-서버 시스템이다.

(출처 : 위키백과)

GPU : Graphics Processing Unit)는 메모리를 빠르게 처리하고 바꾸어 화면으로 출력할 프레임 버퍼 안의 영상 생성을 가속하도록 설계된, 전문화된 전자 회로이다.

wreck

유닛? 단위가 뭐엿지

하나에 4000만원 짜리

CPU는 그렇게 고성능은 아님

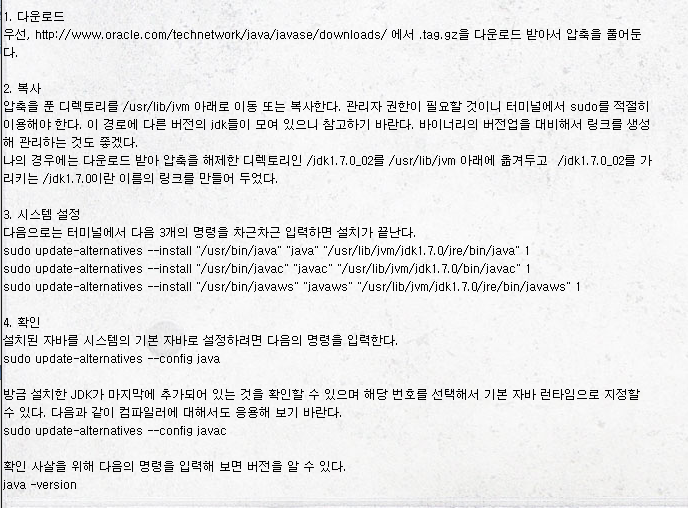
전력이 갑자기 나가거나 정전 일어났을 때 서버에 전력 공급해 주는 것

선 색깔마다 데이터 등 제공해주는 게 다름

gpu가 4개 들어가있는 서버도 있음

**20160707**

어제의 오류를 해결하기 위해 java와 jdk의 버전을 맞추게 하였다.



해결이 안 되었음.

..

konlpy랑 pip랑 지우지 않고 다시 깔았던 것 같은데 갑자기 됨...;;;;;;;;;;;;

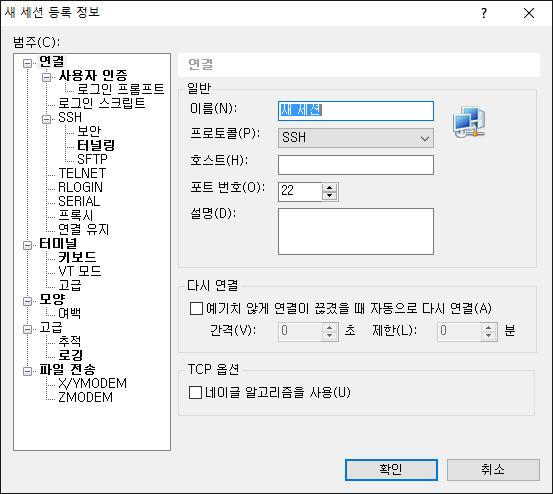
개빡..ㅋㅋ

**20160708**

xshell 을 이용하여 서버 접속

ssh starmichelle@cogcom.asuscomm.com 6622

아니면 아예 등록하기

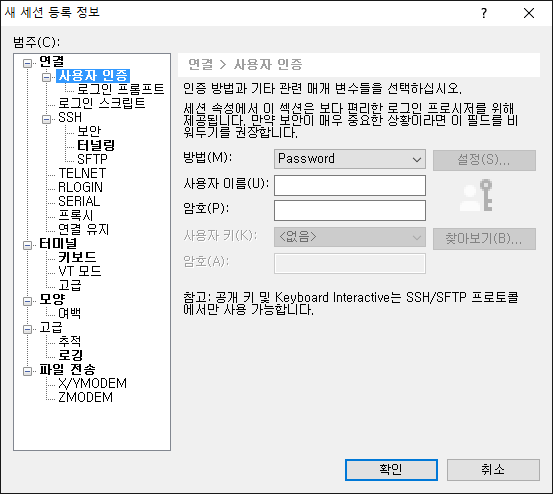
****

**이름 :** KETI

호스트 : cogcom.asuscomm.com

포트 번호 : 6622

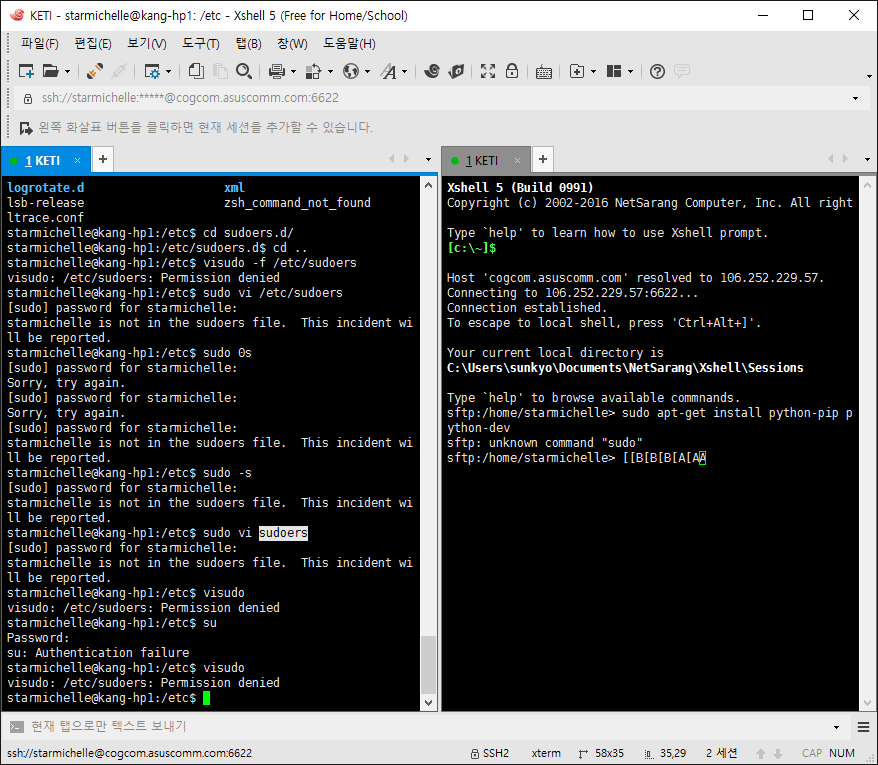
사용자 인증



사용자 이름 : starmichelle

암호 : 1234

확인 후 로그인

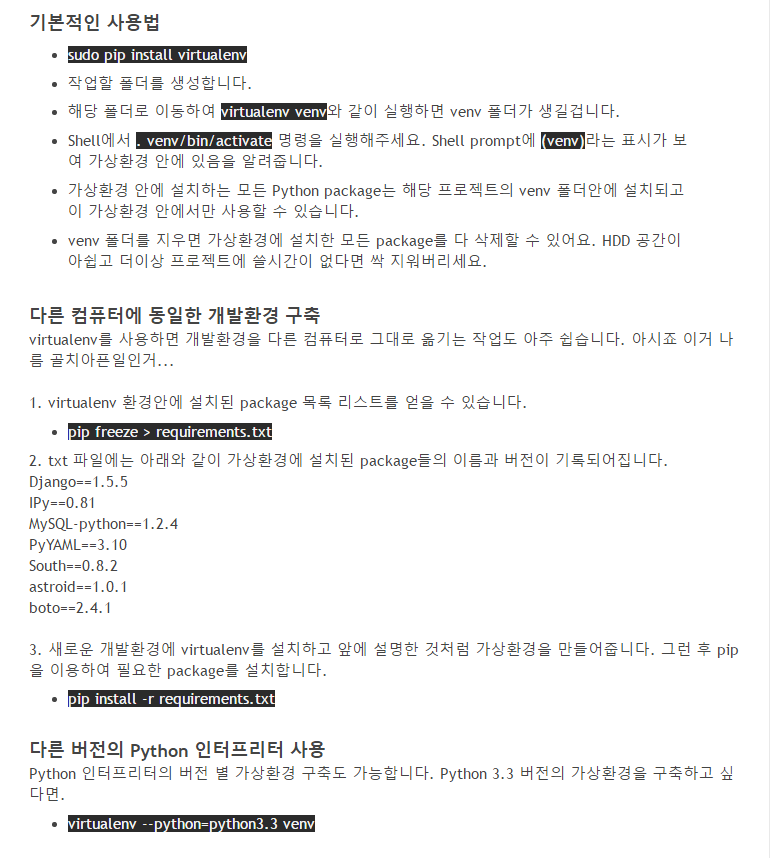


왼쪽 창에서 명령어 입력하고 하면 됨

tensorflow설치하려고 sudo계정으로 들어갔더니 위의 그림처럼 거부됨

root 권한을 얻으려면 강책임님한테 물어봐서 나를 추가해야 하는데 그러면 내 컴퓨터가 아니라 모든 사용자의 개발환경이 바뀔 수 있으므로

virtual environment 를 구축하여 사용하라는 코멘트 받음

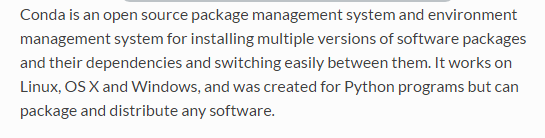


virtual environment를 위의 과정으로 구축

pip install -r requirements.txt

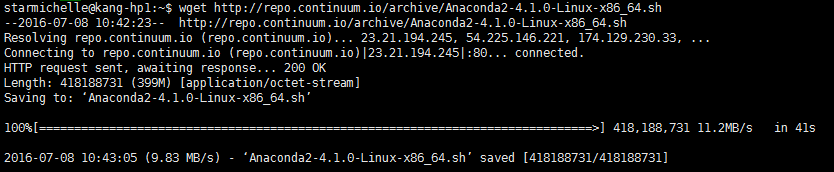
이 명령어는 원래 컴퓨터에 있는 개발환경을 똑같이 복사해서 끌어옴

Conda 라는 걸 설치할 것임.

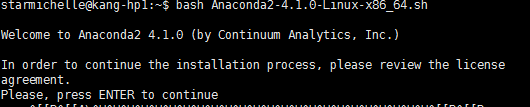


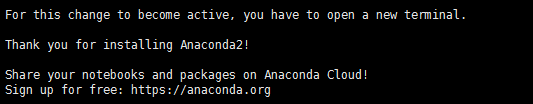


1. shell 에 anaconda를 설치.

wget 명령어 뒤에 링크를 복사하면 설치 됨

3. python 2.7 버전으로 명령어 침





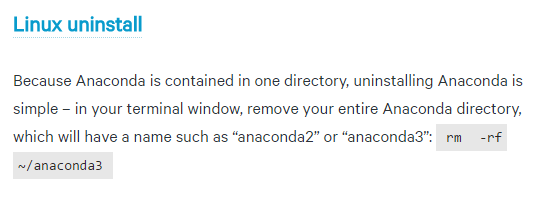
이게 나오면 된거임

이제 xshell을 껐다가 다시 키고

conda --version 을 사용해서 conda가 잘 깔렸는지 확인

(간단하지만 나는 막~ 하다가 안 되서 헤매다가 결국 지우고 다시 했더니 됨)

uninstall 방법





http://conda.pydata.org/docs/test-drive.html#managing-conda 참조

1. Managing Conda

update conda to the current version

> conda update conda

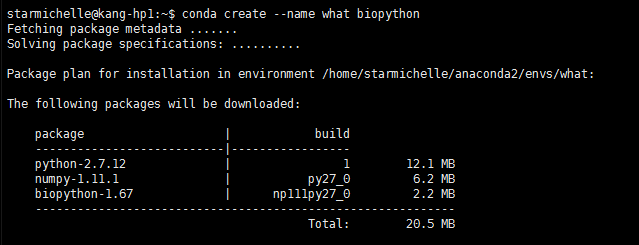
> proceed ([y/n]) ? y

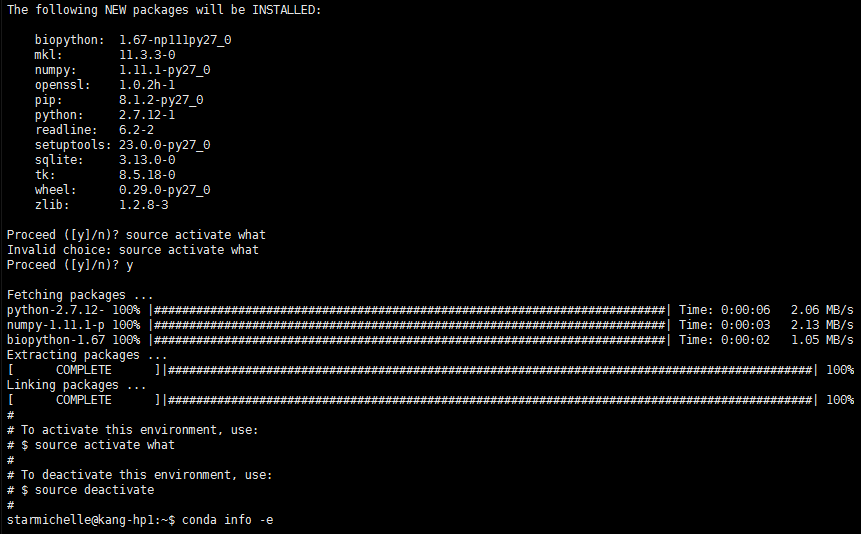
2. Managing environments

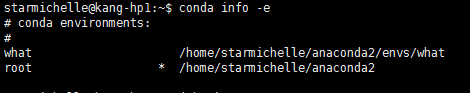
* create and activate an environment

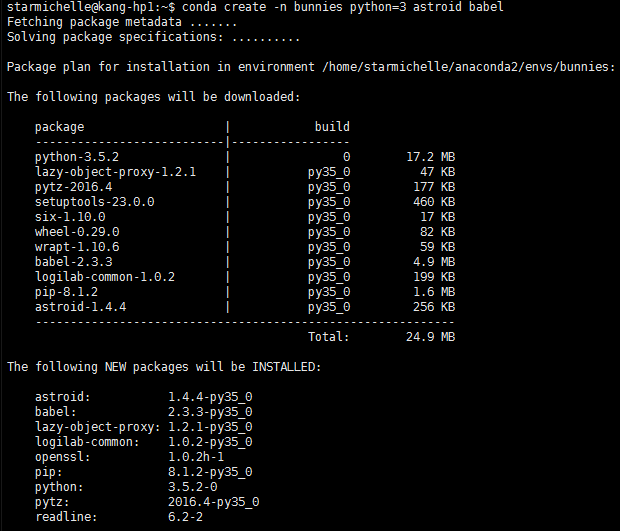
> conda create -n what biopython

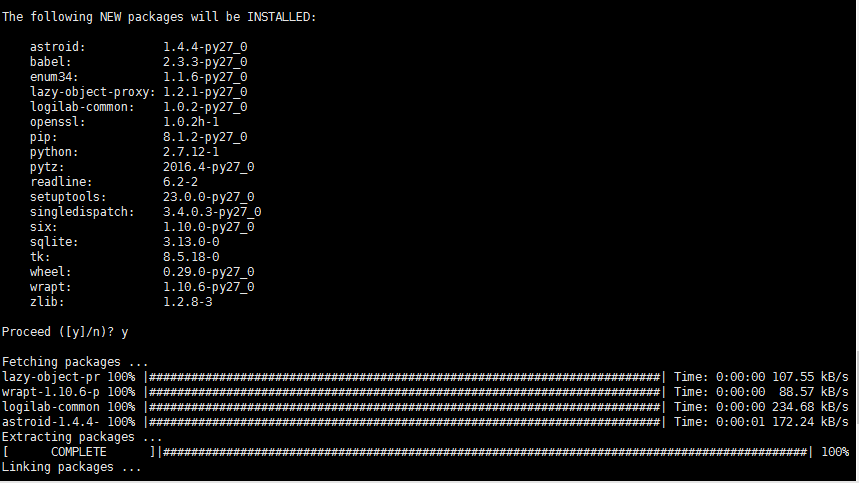
> source activate whatdf

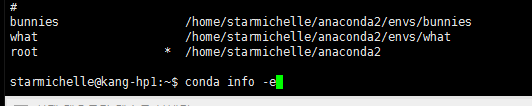












3. managing python



tensorflow 설치

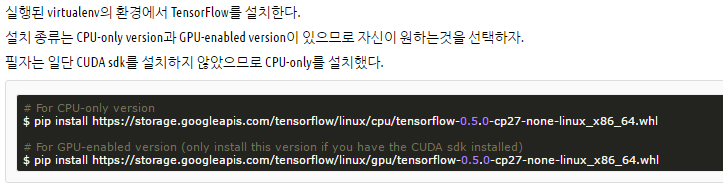
http://goodtogreate.tistory.com/entry/TensorFlow-%EC%84%A4%EC%B9%98-%EB%B0%8F-%EA%B8%B0%EB%B3%B8%EB%8F%99%EC%9E%91 참조



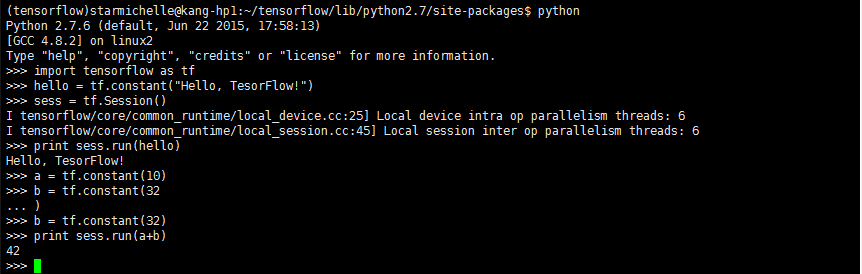
나는 source /home/starmichelle/tensorflow/bin/activate 라고 했음







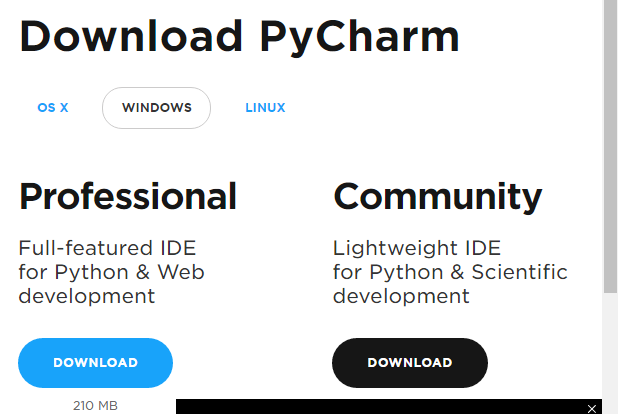
간단한 예제 코드 실행

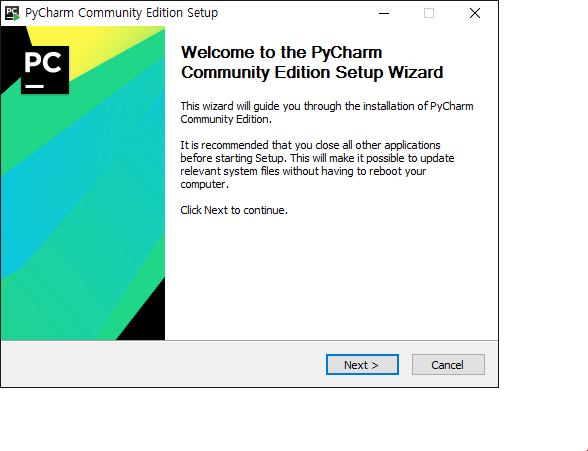


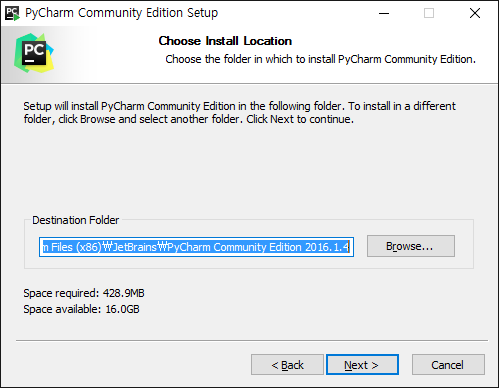
sublimeText 설치하고 싶었는데 찾다보니 python 환경은 pycharm이라는 환경을 많이 사용하고 좋다고 하고 있어서 pycharm을 깔기로 결정

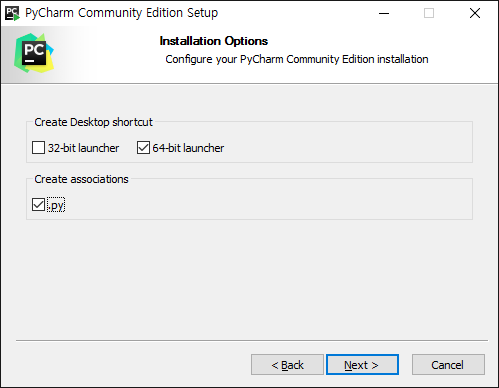
아래는 window에서 설치하는 것이고 그 밑에 linux 에서 설치하는 거 있음

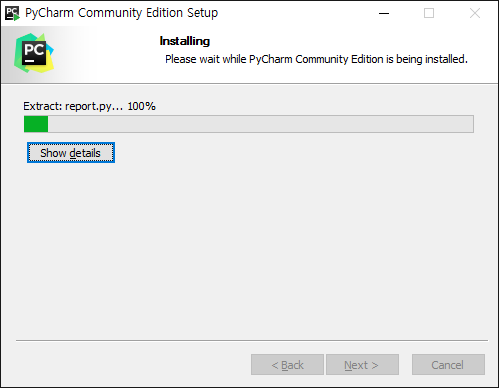
https://www.jetbrains.com/pycharm/download/#section=windows 여기서 window 버전 다운









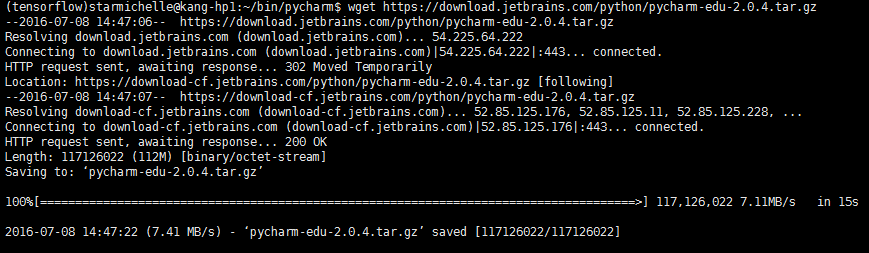


일단 자바 버전을 확인



자바가 깔려있는걸 확인 한 다음 설치

wget https://download.jetbrains.com/python/pycharm-edu-2.0.4.tar.gz

 압축 풀기

> tar -xvf pycharm-edu-2.0.4.tar.gz

깔았는데 실행이 안됨..

shell 창에서는 안 된다고 함..

window에서 pycharm실행.

github에 올려서 가져오는 식으로 함.

vi로 edit.

pycharm은 edu 버전으로 깔면 안되고 professional 버전으로 깔아야 한다.

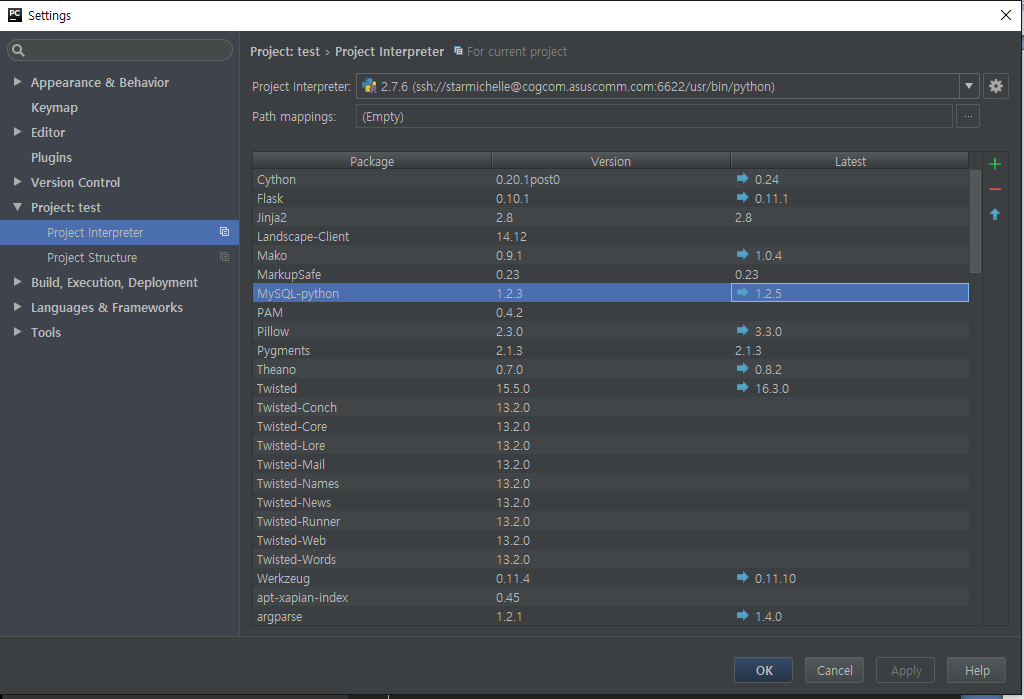
학교 메일이나 국제학생증이 있으면 신청이 되고 1년간 무료로 사용이 가능하다.

[**https://www.jetbrains.com/estore/students/**](https://www.jetbrains.com/estore/students/)

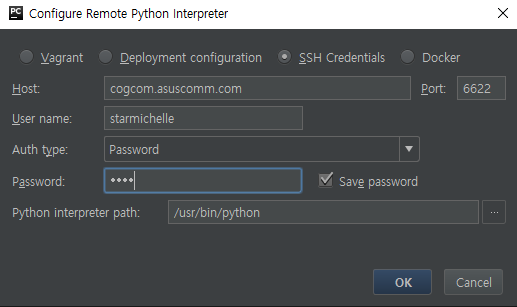
**이 사이트에서 신청하면 되고**

**licence 를 accept한 다음 code를 다운받아 복붙해서 실행하면 된다.**

**setting > project > project interpreter >**

톱니바퀴 모양 누르면 add remote 가 나온다.

거기서 설정해 주면 됨



pycharm 설정 끝

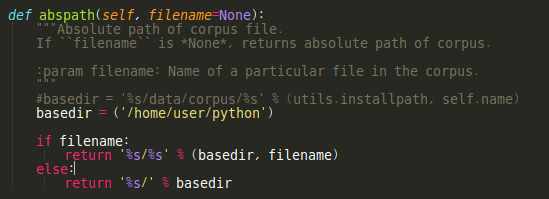
**20160711**

python 한국어 추출 시도해보기

1. txt파일을 기본 제공 파일이 아닌 내 파일로 해서 열어보고 형태소 분석하는 것 시도

① /usr/local/lib/python2.7/dist-packages/konlpy/corpus.py

에 있는 basedir 를 변경해주어 현재 python 폴더에 있는 문서를 찾도록 바꿈



② txt 파일을 바꾸어 시도



아래 사진처럼 start 다음에 실행이 안됨 (안되는건지 시간이 오래걸리는건지..)



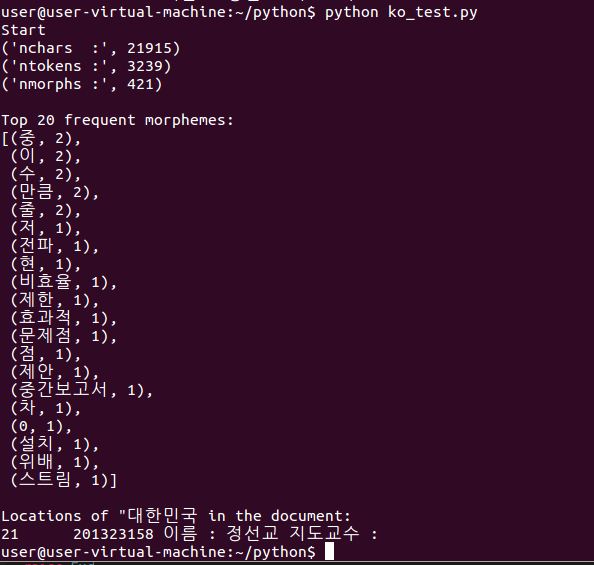
③ hannanum을 다른 걸로 바꿈



->



④ 실행결과



실행이 되는 걸 확인

⑤ 다시 hannanum 시도

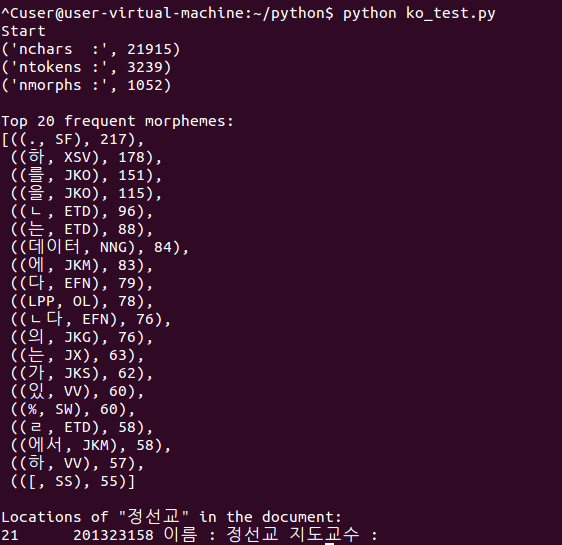




실행 안됨.

.....?? 왜쥬..

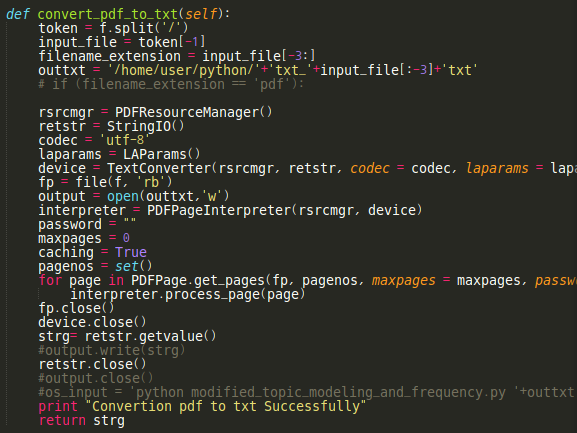
⑥ 다시 kkma로 사용



이제 언니가 한 것에서 한국어 추출로 바꾸기를 시도

1. pdf -> txt

ko\_ui.py ( 원본 : ui.py ) 에서 convert\_pdf\_to\_txt 함수



output을 일단 주석처리.

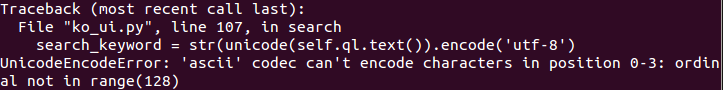
python\_modified\_topic\_modeling\_and\_frequency.py 에서 또 다른 기능을 수행했던 것을 주석 처리하였다.

일단은 한국어 처리하는데 필요 없을 것 같았다.

하지만 나중에 top 20개의 topic을 추출해서 new.csv에 append하는 것을 추가 해야 한다.

2. 입력한 단어 search 해서 추출

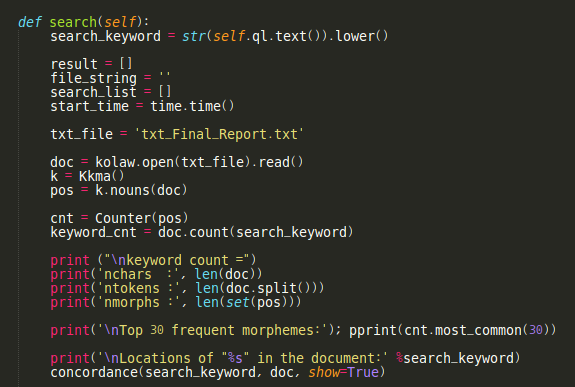
단어를 한글로 입력했을 때 그게 인식이 안되어서 변환해주어야 한다.

이런 식으로 인코딩 문제가 발생

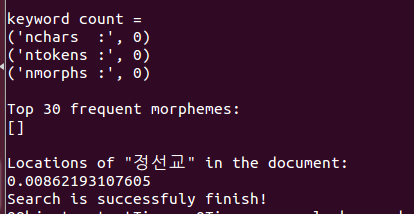


해결 완료

ko\_ui.py 에 아래처럼 수정해줌



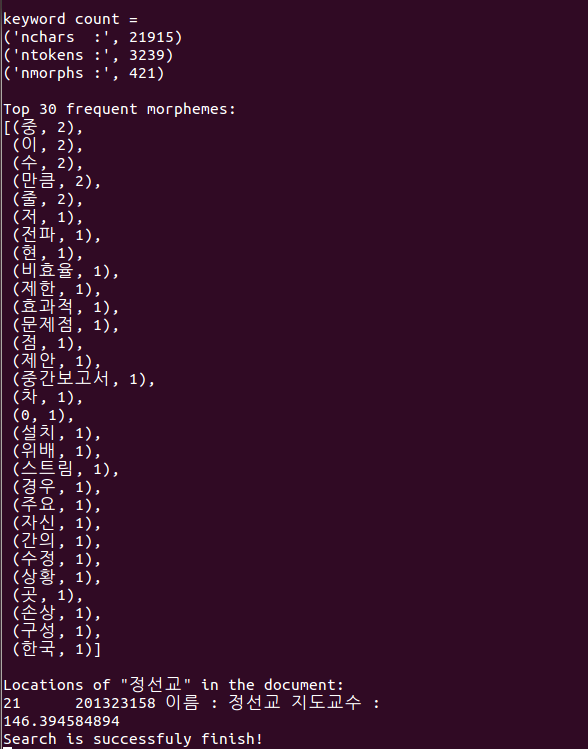
그랬더니 출력 결과가 아무것도 안나오고 keyword에 대한 걸 인식하지 못 한 것 같은 결과가 나옴



헤매다가 text파일을 열어보니 빈 파일...

1번을 주석처리 해주면 안 되는 거였다.

다시 원상복구하고 실행해보니



성공적인 결과가 나왔음

해결 해야 할 점 : 키워드 top #을 new.csv에 추가하기

더 의미 있는 키워드 뽑아내기

더 빠르게 탐색

참고하면 좋을 만한 페이지 : *https://www.lucypark.kr/courses/2015-dm/text-mining.html*

*http://davidmburke.com/2014/02/04/python-convert-documents-doc-docx-odt-pdf-to-plain-text-without-libreoffice/*

pycharm 실행 시 path가 잘못되어 자꾸 오류가 남.

해결 해야 한다.

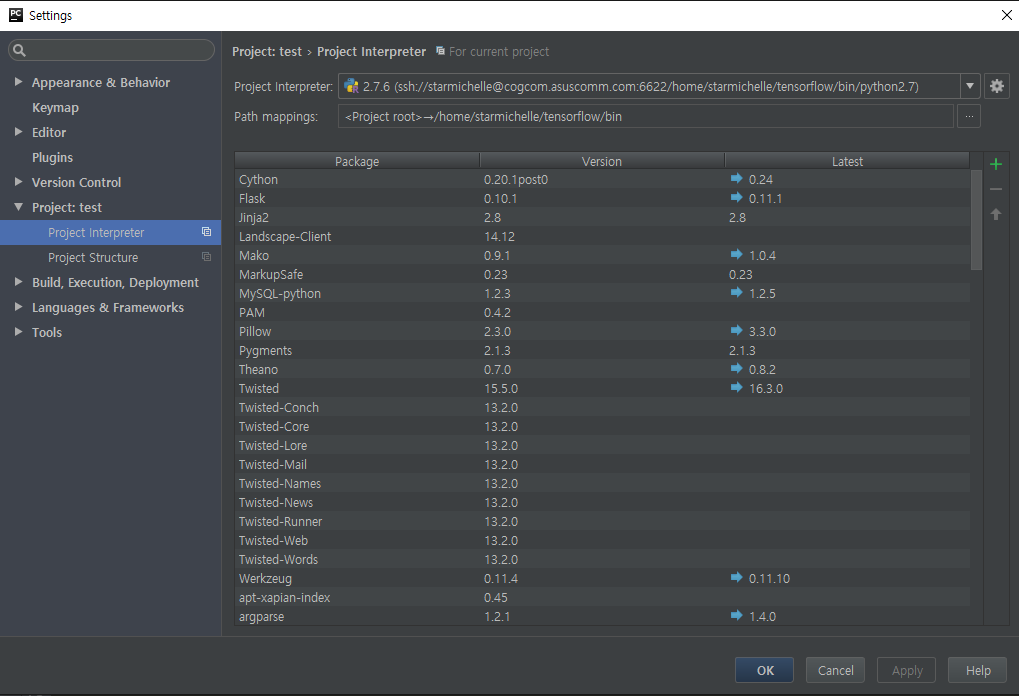
http://dejavuwing.tistory.com/category/Linux

**20160712**

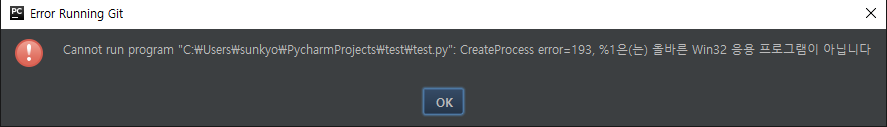
난 바보였다....

path가 잘못되었을 때 오류로 no such file or directory 라는 게 나왔는데 그게 내가 폴더가 아니라 그냥 python2.7이라는 파일을 선택했는데 그냥 directory를 선택해야 하는 정말 간단한 문제였다. 하핫

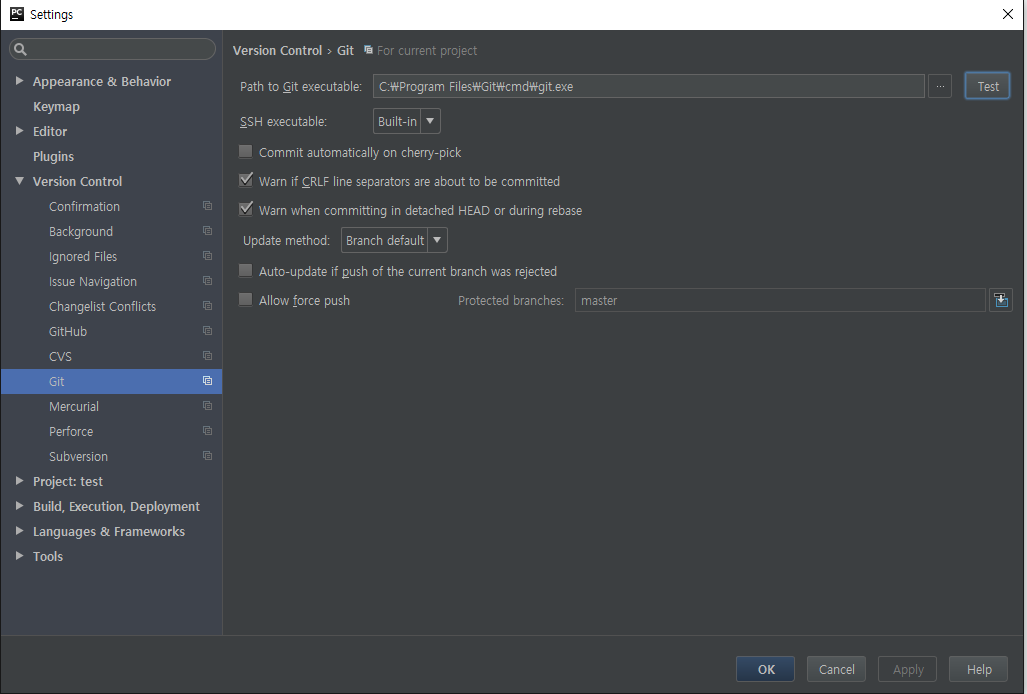
어쨌든 설정 완료



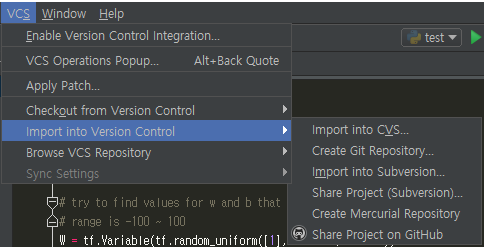
run은 되지만 github 연결이 되지 않는다.

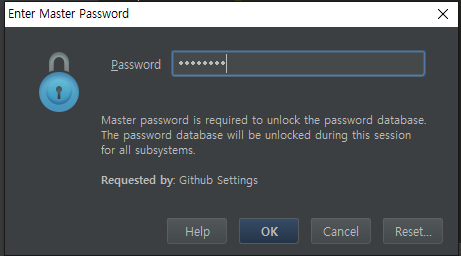


git.exe 를 찾아서 path 연결

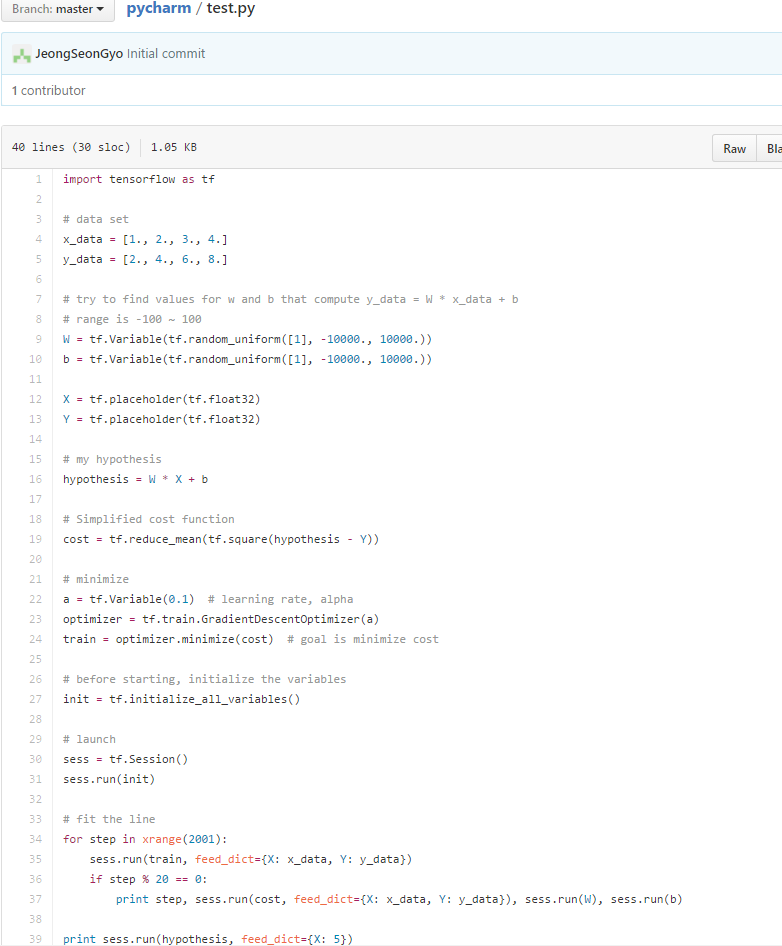


VCS -> import into Version Control -> Share Project On GitHub





공유 완료



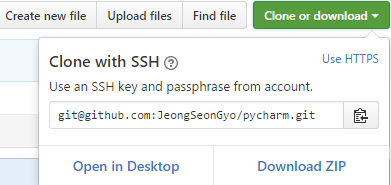
이제 서버에서 github와 연결하는 걸 해야 하는데 아직 잘 모르겠음

https://mylko72.gitbooks.io/git/content/remote/repository.html

참조

서버에서

git clone 주소(github 에서 clone을 누르면 주소가 나온다.)



하면 github에 있는 디렉토리가 그대로 복사가 되고

virtual environment 해서 conda로 해서 tensorflow 들어간 다음 python 으로 하면 됨.

'tensorflow' virtual environment 만든 거 폴더 옮김. anaconda2/envs로!

소스는 ~/anaconda2/envs/tensorflow/lib/python2.7/ 로!

tensorflow 하려고 했는데 갑자기 tensorflow 모듈이 없다고 해서 다시 깔아야 함

pip install https://storage.googleapis.com/tensorflow/linux/cpu/tensorflow-0.5.0-cp27-none-linux\_x86\_64.whl

를 복붙해서 다시 깔아야 하는 것 같기도 하다.

test 해보기

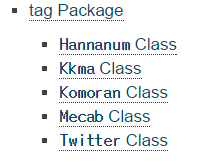
-> python

import tensorflow as tf

해결 해야 할 점 : 키워드 top #을 new.csv에 추가하기

더 의미 있는 키워드 뽑아내기

더 빠르게 탐색

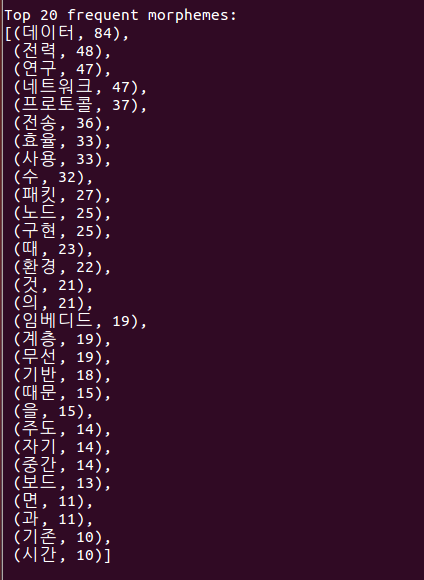


5개의 클래스가 있음

-> Kkma 는 별로 의미 없는 keyword가 많이 나옴

-> Hannanum 으로 실험 : 너무 오래걸려서 결과도 안나옴

from konlpy.tag import Komoran -> komoran 이라는 모듈을 사용



의미도 있고 cnt가 많이 나오는 것이 제대로 나옴.

mecab 이라는 모듈을 사용해 보겠음

->mecab은 따로 깔아야 함..그래서 안할꾸임..히히

======================

cf)

linux 창

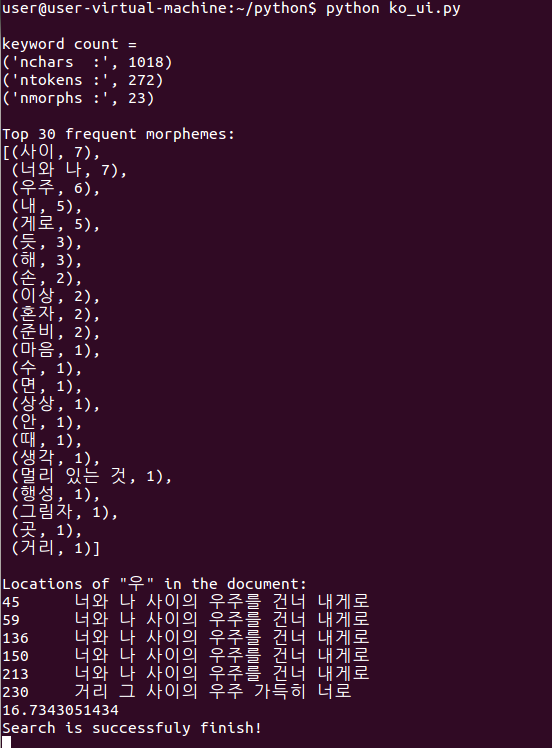
du . -hs ; 현재 사용량 확인

-ll : 모든 파일 출력

======================

ko\_ui.py 로 test해보았음

ko\_test.py와 결과가 같음



해결 해야 할 점 : 키워드 top #을 new.csv에 추가하기

더 의미 있는 키워드 뽑아내기

더 빠르게 탐색

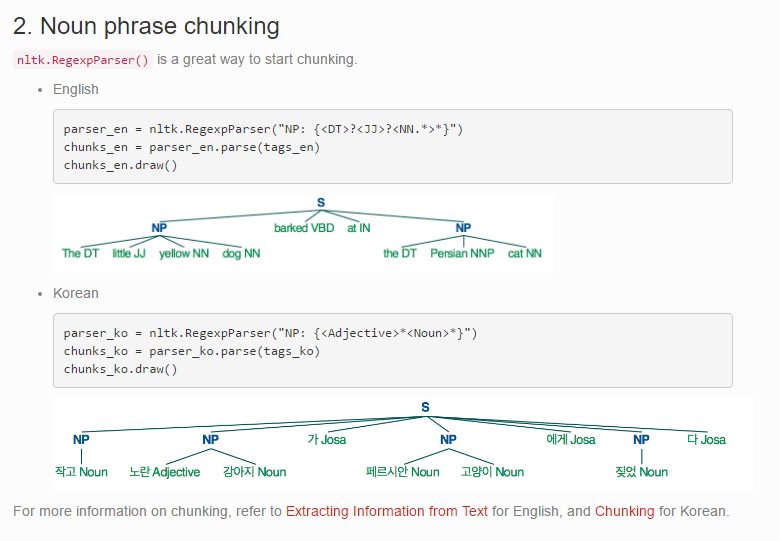
================

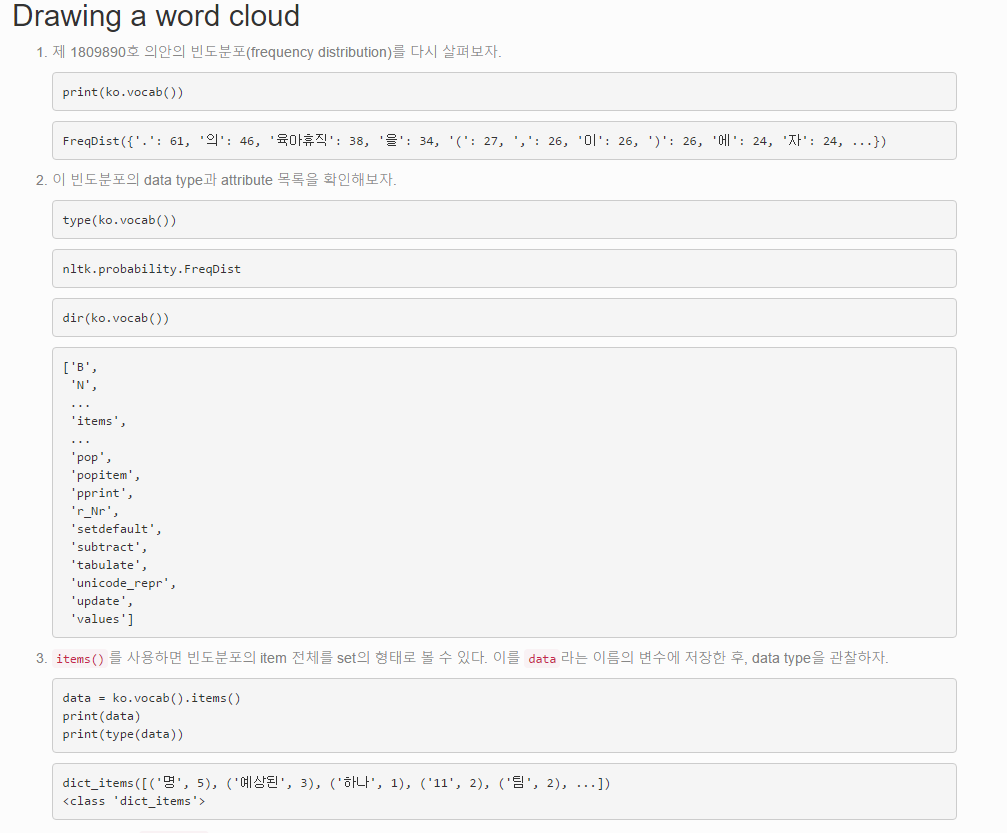
참고 : 유사 언어 찾기

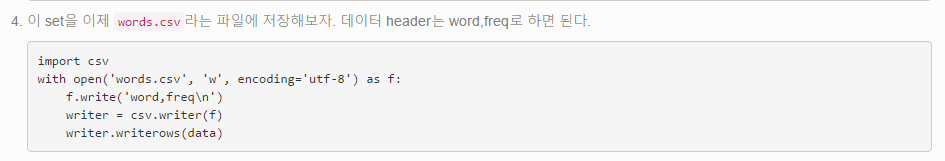
ko.similar(' 유사')

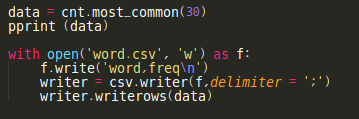
collocations

ko.collocations()

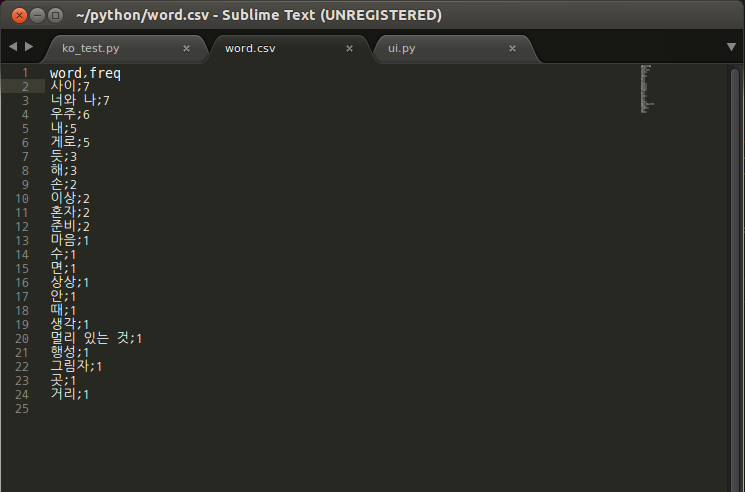








csv 파일 성공적 생성.



svn VS git

svn : version으로 관리 / 코드 전체를 snap shot

git : 변경된 부분만 따로 저장

================

csv 파일 읽고 쓰기

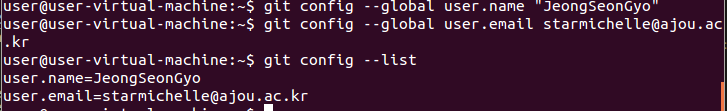
csv : 각 항목의 값들이 콤마로 분리되어 있고, 행은 개행 문자(new line)로 구분하는 파일 포맷

git 사용하기

-linux용 git 설치

sudo apt-get install git

사용자 설정



- 새로운 저장소 만들기

폴더를 하나 만들고 git init 명령 실행 : 새로운 git 저장소가 만들어짐

- 저장소 받아오기

1. 로컬 저장소를 복사 > git clone /로컬/저장소/경로

Linux에 있는걸 서버로 옮기는 것 ex : git clone ~/python

2. 원격 서버의 저장소 복사 > git clone starmichelle@ajou.ac.kr:/원격/저장소/경로

서버에 있는 걸 linux로 옮기는 것 git clone



<https://help.github.com/articles/generating-an-ssh-key/>

참조

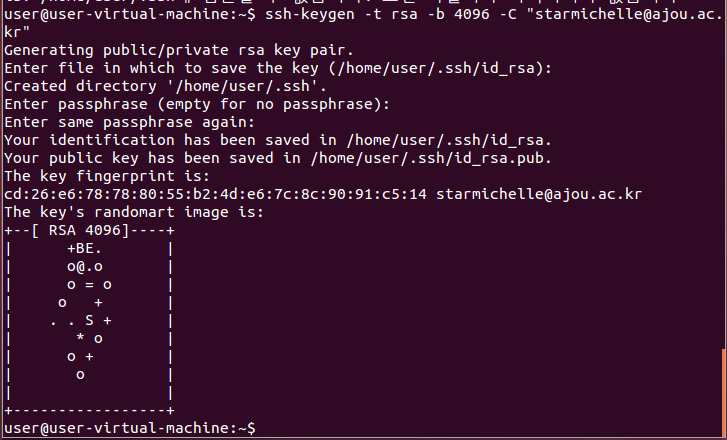
**20160713**

Generating an SSH key

1. ls –al ~/.ssh 로 이미 key가 있는지 확인



1. ssh-keygen –t rsa –b 4096 –C [starmichelle@ajou.ac.kr](mailto:starmichelle@ajou.ac.kr) 해서 rsakey 만들기



1. ensure ssh-agent is enabled



1. add your ssh key to the ssh-agent

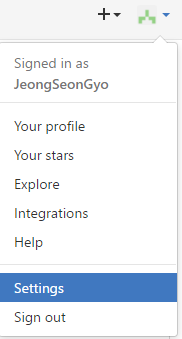


1. copy the SSH key to your clipboard

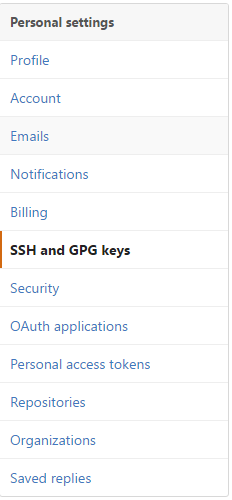
* sudo apt-get install xclip
* xclip –sel clip < ~/.ssh/id\_rsa.pub



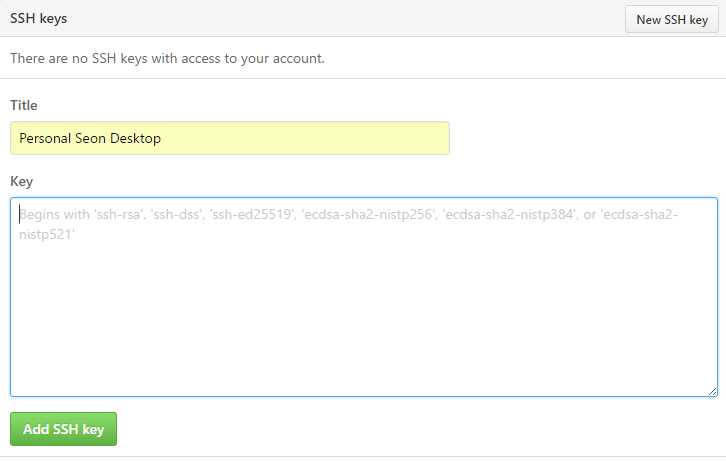
1. 내 github로 가서 setting 클릭

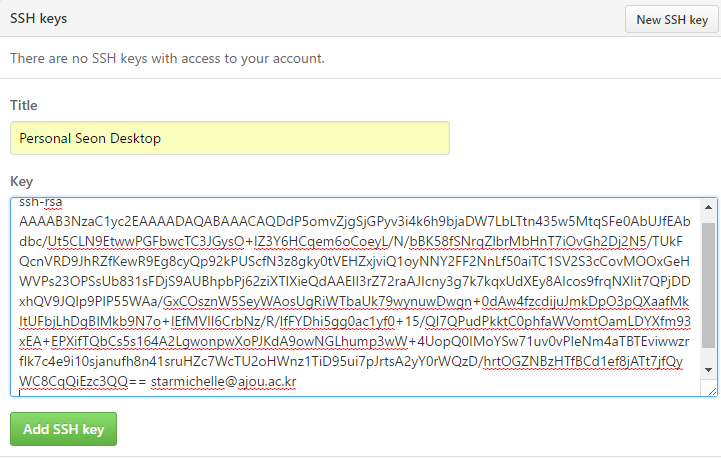


1. Ssh and GPG keys 클릭

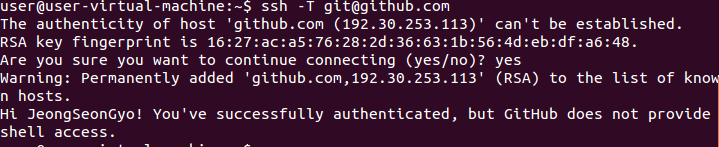


1. New SSH Key 를 누르고 입력한다.





1. Testing ssh connection

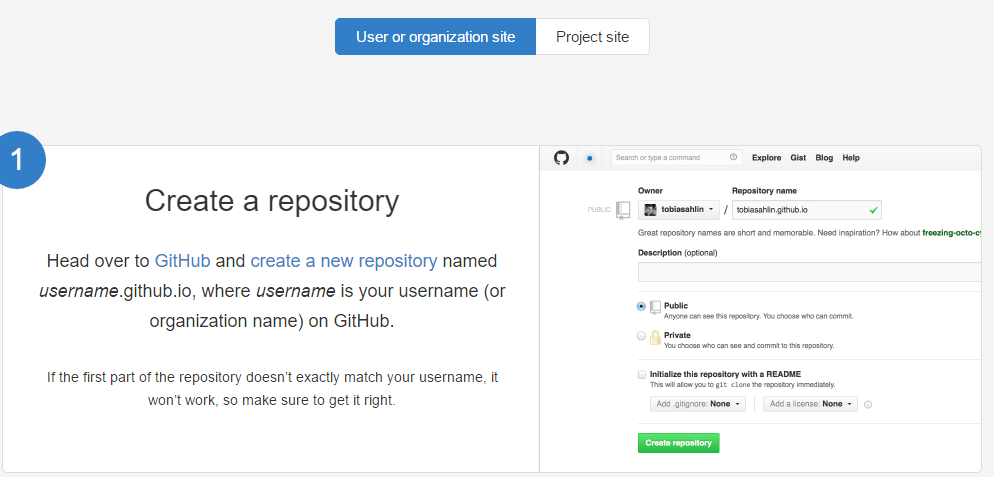


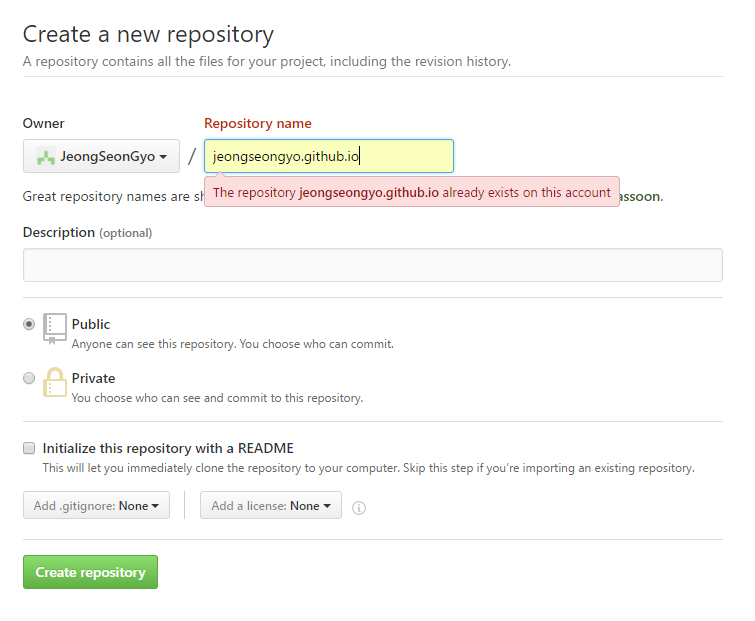
Successfully authenticated 가 되면 성공한거임!! 야호!

Webpage 만들기 using git

* <http://jeongseongyo.github.io/> 만들기

1. Repository 를 만든다.

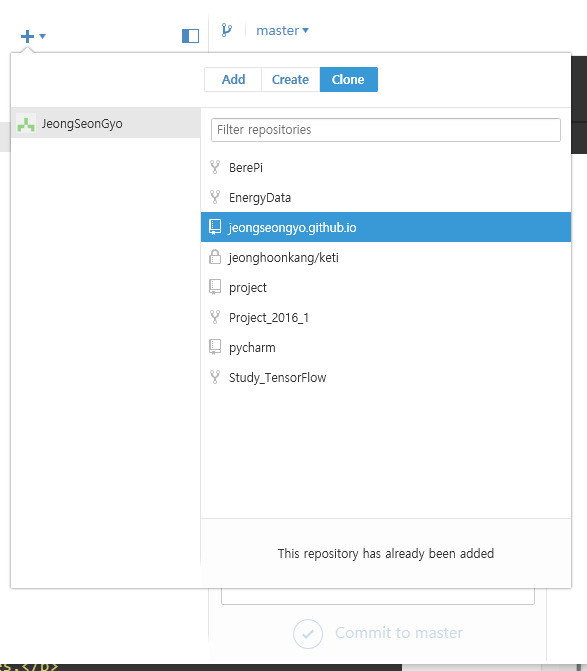




1. Set up in Desktop 을 누르면 git 앱이 실행된다.



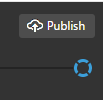
1. 만든 repository를 clone한다.



1. Index.html 파일을 올린다. (chrome에서)

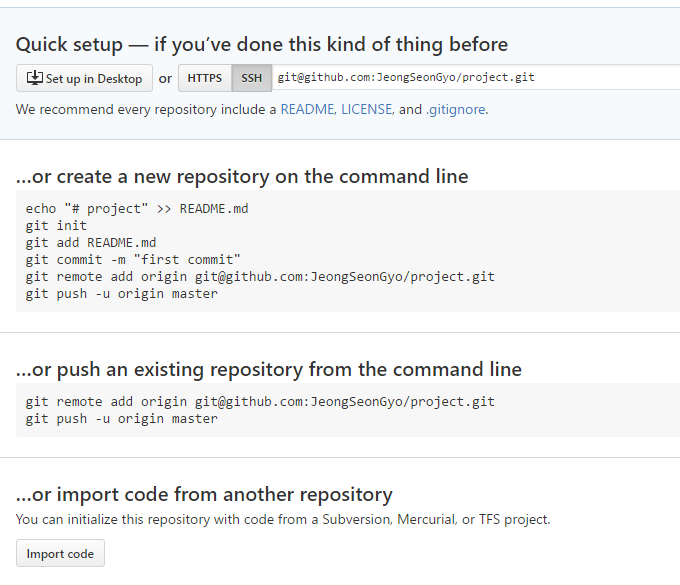


1. 다시 앱으로 돌아와서 publish를 누른다.



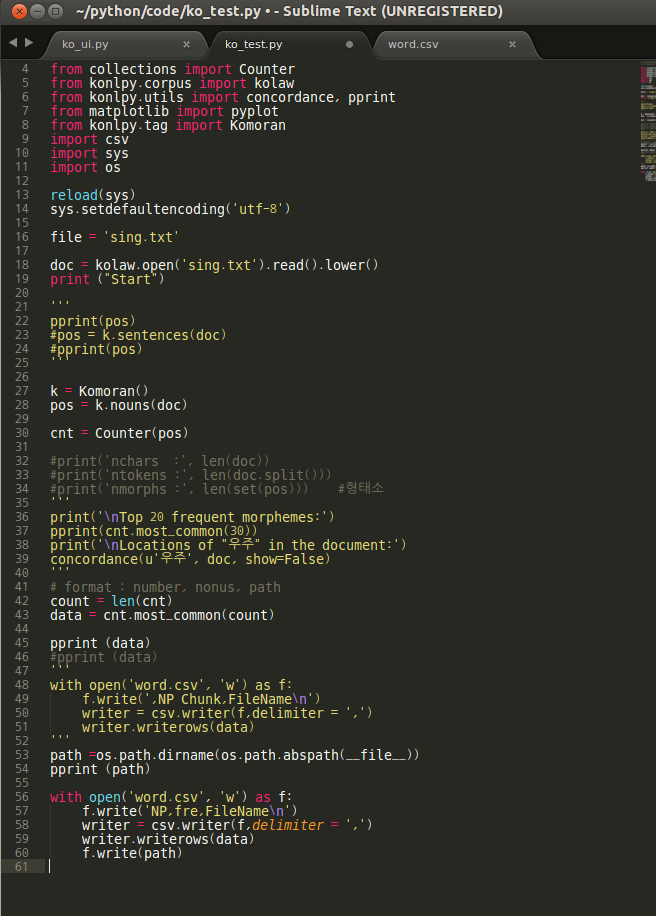
1. <http://jeongseongyo.github.io> 에 들어가본다.(끝)12345





D아… git 실패…

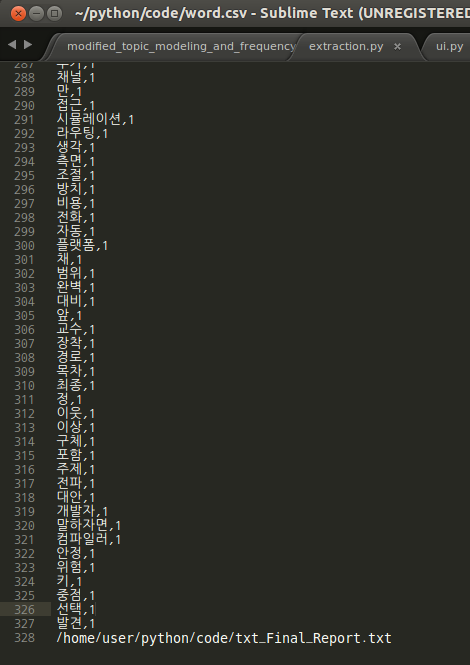
Csv 생성 방법 강구 중





일단 이런식으로 csv를 작성하고 유빈언니의 ui.py를 고치고 있음

현재 search전까지 완성



구체적으로 파일 이름까지 출력하도록 해줌

Ko\_ui.py에서 search 함수의



이 부분을 바꿔야 함.

왜냐면 내 csv의 파일 path와 이름은 맨 파일의 탐색 끝에 붙이게 했으므로.

파일 path/filename은 밑에 있는데….그러면 또 그걸 찾아가야 하는데..그게 정확할까..?

Csv에 저장해놓은 것을 parsing해서 문장 끝마다 path/filename을 붙일 순 없을까..?

Line을 읽어서 뒤에 붙이는 방법은..?

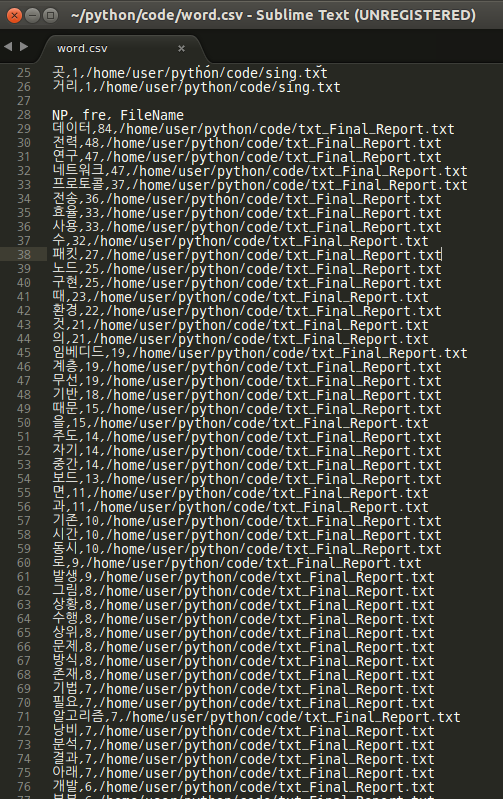
**20160714**

csv파일을 새로 만들어서 거기다가 일단 위의 사진에서 출력되는 것을 한 줄 한 줄 입력시킨다.

다시 그 파일을 열어서 줄마다 경로와 파일 이름을 쓴다.

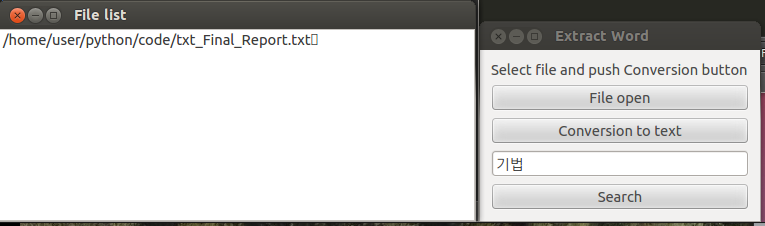
그리고 원래 저장해놓는 csv파일에 append mode로 쓴다.

그러면



원하는 것처럼 출력이 됨!!!!!!!

Search 결과!!



원하는 결과를 도출했음

해야 할 것

1. 영어 + 한글
2. Hwp 파일 변환
3. 오류 처리 -> file open 한 적이 없는데 conversion to text 버튼을 눌렀을 경우
4. 현재 디렉토리에 있지 않은 파일도 탐색 + 그 파일의 디렉토리 주소 가져오기. (지금은 절대경로임)