

**About code:**

The version of python is python3.6 and should have numpy and glob installed

There are nb.py, a file name 'stopwords' and a file named '20\_newsgroups'

How to run:

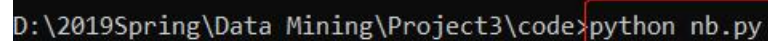
You should first enter the current folder in cmd.

That is, you should first unzip the 'code' folder and go to the directory of 'code' in cmd.

Make sure nb.py, 'stopwords' text and that data file are both in 'code' folder.

Then you can use following command:

>python nb.py



```
D:\2019Spring\Data Mining\Project3\code>python nb.py
```

This is a long data processing process, it will run about 10 hours. Sorry about that.