The function of files in this document is listed as below:

1. Project #2\_data\_reading.py

Shows the process of getting data frame for raw data.iTS reult already uploaded as .pkl and .txt file. **Please donot use it or run it.**

1. .pkl files

Stored raw data

1. File\_id.txt

Shows the order of reading files into all\_X\_test.plk

1. Project #2 Feature\_extraction\_and\_running.py

Shows the examples of extracting features and training NB model

1. **If you want to submit prediction results:**

Please (1) Create a excel file (.csv file). The heading of first row is id and category

(2) paste items in File\_id.txt as the first column of a excel file, and delete .txt (using replace in excel)

(3) paste items in new.csv (result of Project #2 Feature\_extraction\_and\_running.py) as the second column of that excel file.