FIT9133: Programming foundations in python

Assignment2

Documentation

Yuhan Zhang
29400988

目录

Version

Python version 3.6
Numpy version 1.15.2
Matplotlib version 3.0.0

Version list

Date of previous version	Version	Comments
03/10/2018	V 1.0	Using **readlines()** and only process one file in each time
06/10/2018	V 2.0	Using **read()** and implement all function
08/10/2018	V 3.0	Bug fixed

Introduction

This project is a basic language analyser for analyse the children that have some form of language discords. The target is two group. SLI group and TD group, each have ten files.

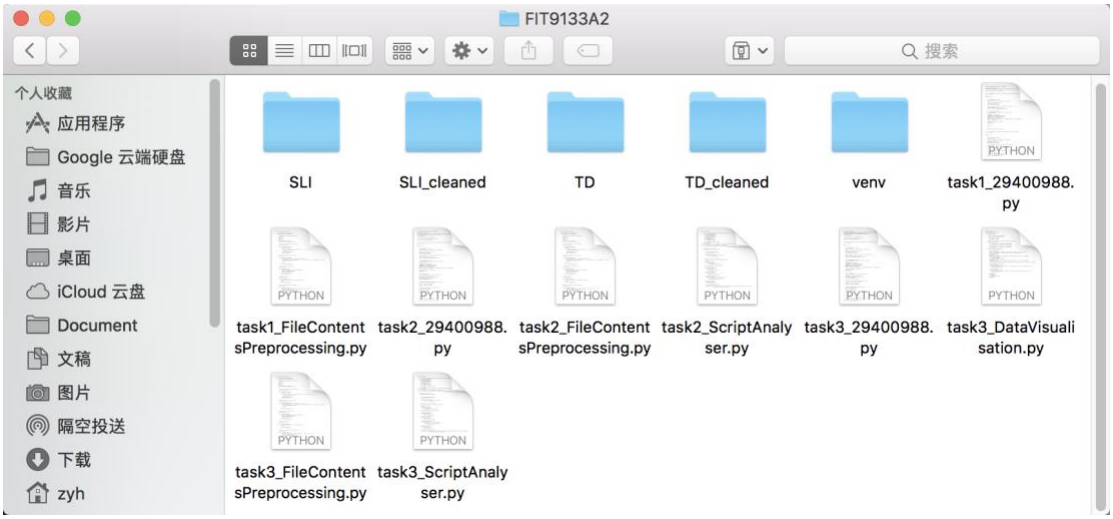
Step of working flow:

1. Pick up the necessary data
2. Store these contents
3. Counting the needful data
4. Data visualisation

Important! Before Run

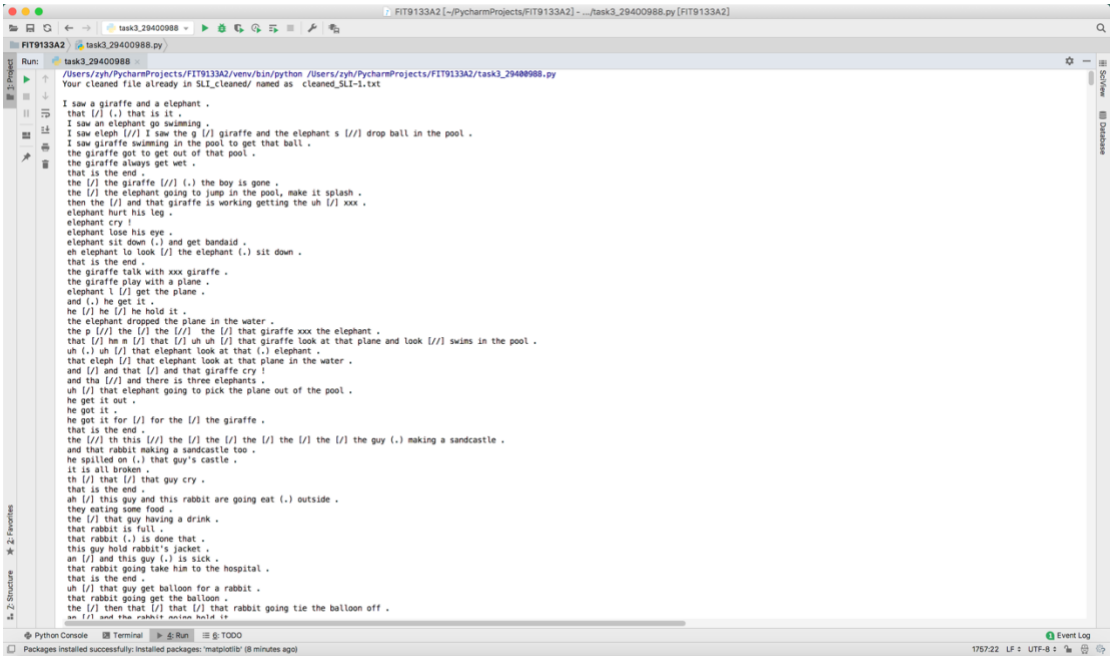
Please create the correct folder before run the program

1. Create *SLI_cleaned, TD_cleaned* folder.
2. Decompression the SLI and TD under same path with *.py* file.
3. Example for mac user.



Function

1. task_29400988.py
   a) Used to run all function and initialization variable.
   b) This project will process 10 files for each group just by one click.
   c) File contents will be preprocessing and save in new folder(*SLI_cleaned, TD_cleaned*) with a new name, then the ScriptAnalyser will do counting roles and output the valuable data. Finally, these data will be visualization and output a bar chart.
2. task_FileContentsPreprocessing.py
   a) File contents will be pick-up according to the requirements.
   b) Output



3. task_ScriptAnalyser.py
   a) This function will count six variables in fixed contents for each file
   b) The counted data will be build
   c) Output

```
Analyse Report of cleaned_SLI-1.txt
Number of statements:67
Number of vocabulary:123
Number of repetition:47
Number of retracing:10
Number of errors:1
Number of pauses:12
```

4. task_DataVisualisation.py
   a) The summation of each file for two group will be used to compute the average value.
   b) A bar chart created by this function and the average will be visualization.
   c) Output