

程式碼說明文件

一、實驗環境

CPU : Intel Core i5-9600K

RAM : 16 GB

OS : Windows 11

Program Language : Python

Python Dependencies:

1. pytorch == 1.3.1
2. torchtext == 0.4.0
3. scikit-learn == 0.22
4. spacy == 2.2.3
5. scipy == 1.4.1
6. gensim == 3.8.1
7. numpy == 1.17.4

二、資料集來源

<https://github.com/peiranli/ARYA>

三、運行方式

透過 commandline 的方式先運行 `nltk_download.py`，最後再 `train.py` 檔即可，參數調整方面可輸入以下指令來進行調整。

1. `--quantile [hnorm threshold]`
2. `--score_threshold [score threshold]`
3. `--seedword_limit [seedwords number]`
4. `--no_filtering`
5. `--no_tuning`