Course Project Report (Part I)

ERG3020 Web Analytics and Intelligence Project

Author: Piao Chuxin (115010058), Ye Xiaoxing (115010270)

In this part, we implemented the Naive Bayes Classifier using tweets data, using the code from sailor2017 (https://github.com/abisee/sailors2017). Copies of exported PDFs are attached to the report, and you can also view the files online. Due to some technique problems, the PDF is not wrapping lines and some characters will be missing.

What is this project?

The original project is a workshop in sailors2017, focus on the introduction to Natural Language Processing. The task was to automatically classify real tweets from Sandy using a Naive Bayes model.

(Introduction to the project?)

What we have also done?

Besides implementing the codes, we also adopted the original codes to Python 3. The original one is in Python 2 and it is going to be deprecated.

For the bonus part, we have finished a tweet sentiment analysis project, which is attached as well.

Jupyter Notebook Viewers

- lesson1_rulebased.ipynb: https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson1_rulebased.ipynb
- lesson2_evaluation.ipynb: https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson2_evaluation.ipynb
- lesson3_naivebayes_exercises.ipynb: https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson3_naivebayes_exercises.ipynb
- lesson4_languagemodel.ipynb: https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson4_languagemodel.ipynb
- lesson5_naivebayes.ipynb: https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson5_naivebayes.ipynb
- lesson6_visualization.ipynb:

https://github.com/Yexiaoxing/sailors2017/blob/Miley/lesson6_visualization.ipynb