

IERG3320/ESTR3306 Social Media and Human Information Interaction

Mini Project One Sentiment Analysis – Project Specification

Deadline: 15 November 2020 (Total: 20 scores)

Use any method of your own to collect the comments you received from your blog. Save them as a .txt files. Name the file by your student ID (e.g., 1155115511.txt). (In case you did not receive any comments from your blog, please use all of the comments you have ever provided to the other classmates).

Write a Python program to perform sentiment analysis on the comments using the dictionary-based approach. Two lists of sentiment words (positive-words.txt and negative-words.txt) are provided to you. Your program should be able to:

1. Read your own .txt file.
2. Perform dictionary-based sentiment analysis on your own .txt file; and
3. Output the overall sentiment score (a normalized value between -1 and 1) of your received comments on the screen when the program is run.

Remarks: You do not need to expend the given list (although we will learn how to generate the lists [i.e. positive-words.txt and negative-words.txt] during the lectures). You are allowed to handle exceptions such as negation and emoticons manually before running the program.

Submit the following items to the assignment dropbox (Project One) on eLearn:

1. The.txt file;
2. Source codes of your Python program;
3. Veriguide proof document of the program source codes.

Our tutors will run your program to see if it works properly, and verify your score independently using your .txt files and his own program. (Note: the sentiment score will not affect your project score).

Assessment scheme: 1. The dataset (i.e. the .txt file) (5 scores) 2. An executable Python program (15 scores)