

Grading Criterion of Assignment 1

Total Score: 100 points

Part 1: Collecting user's profile information (25 points):

Task:

Given a list of user's IDs, please write a data crawler to collect the users' profile information.

What to Turn In:

- (1) A result file that contains the profile information of the Twitter users with the following IDs: 18165866, 26257166, 12579252.
- (2) The source code of your crawler to finish this task.

Grade Criterion:

1. Whether submitted codes can run successfully.
 - Successfully: 7 points
 - Unsuccessfully:
 - Authenticate function correctly: 2 points
 - Profile function correctly: 2 points
2. Whether result file contain the correct profile information.
 - Contain: 6 point for each given ID, $6 \times 3 = 18$ point in total.
 - Not Contain:
 - For each given ID, only contain correct but incomplete profile information: 4 points.
 - For each given ID, contain complete but incorrect profile information: 4 points.

Part 2: Collecting user's social network information: (25 points):

Task 2:

Given a list of user's IDs, please write a data crawler to collect the users' social network information (i.e., the lists of screen names of the user's friends and followers)

What to Turn In:

- (1) A result file that contains the social network information of the Twitter users

with the following IDs: 18165866, 26257166, 12579252. Note: you only need to collect the first 20 followers and friends of a specified user in the result file.

(2) The source code of your crawler to finish this task.

Grade Criterion:

1. Whether submitted codes can run successfully.
 - Successfully: 7 points
 - Unsuccessfully:
 - Authenticate function correctly: 2 points
 - Friend function correctly: 1 point
 - Follower function correctly: 1 point
2. Whether result file contain the correct social network information.
 - Contain: 6 points for each given ID, $6 \times 3 = 18$ points in total.
 - Not Contain:
 - For each given ID, contain correct but incomplete friend social network information: 2 points.
 - For each given ID, contain correct but incomplete (less than 20) friend social network information: 2 points.
 - For each given ID, contain complete but incorrect friend social network information: 2 points.
 - For each given ID, contain complete but incorrect follower social network information: 2 points.

Part 3.1: Collecting tweets in real-time filtered by specified keywords: (25 points):

Task 3.1:

Write a data crawler to collect tweets that contain one of the following two keywords: [Illinois, weather]

What to Turn In:

- (1) A result file of tweets that contain one of above two keywords (50 tweets will be enough and the result file size should be less than 1 M).
- (2) The source code of your crawler to finish this task.

Grade Criterion:

1. Whether submitted codes can run successfully.
 - Successfully: 10 points.
 - Unsuccessfully:

- Authenticate function correctly: 4 points.
 - Keywords filter function correctly: 4 points.
2. Whether result file contain correct tweets filtered by keywords.
 - Contain: 15 points.

Part 3.2: Collecting tweets in real-time filtered by specified geo-location information: (25 points):

Task 3.2:

Write a data crawler to collect tweets that originate from the geographic region around Champaign.

What to Turn In:

- (1) A result file of tweets that originate from the specified geographic region (50 tweets will be enough and the result file size should be less than 1 M).
- (2) The source code of your crawler to finish this task.

Grade Criterion:

1. Whether submitted codes can run successfully.
 - Successfully: 10 points.
 - Unsuccessfully:
 - Authenticate function correctly: 4 points.
 - Location filter function correctly: 4 points.
2. Whether result file contain correct tweets filtered by location information.
 - Contain: 15 points.