

Relazione primo progetto di Social Computing

Andrea Gritti (152206)

Anna Sacchet (153005)

Giovanni Palma (152336)

Paola Sagliocca (151861)

Introduzione

In questa prima esperienza di laboratorio, il nostro gruppo ha avuto l'occasione di prendere dimestichezza con il linguaggio di programmazione Python e di accedere alle API di Twitter, oltre che di effettuare analisi su grafi e reti grazie, alle librerie: Tweepy, utilizzata per i primi 3 punti del progetto, e NetworkX, per i rimanenti 7 punti. Al centro della nostra analisi vi è il noto social network Twitter: un social finalizzato alla comunicazione e alla divulgazione di informazioni tramite lo scambio di messaggi di testo di al massimo 280 caratteri che pendono il nome di “tweet”, e per cui si contano oltre 500 milioni di iscritti e 330 milioni di utenti attivi in tutto il mondo, motivo per cui gli oltre 5780 tweet al secondo lo rendono uno dei più grandi database al mondo da cui poter estrapolare qualsiasi tipo di informazione.

Tweepy

Tweepy è una libreria Python per l'accesso alle API di Twitter. Consente agli sviluppatori di integrare facilmente le funzionalità di Twitter nelle proprie applicazioni Python. Con Tweepy è possibile eseguire operazioni quali la lettura e la scrittura di messaggi Twitter, la ricerca di tweet, ricavare i followers di un utente ed altro. Per utilizzare la libreria Tweepy è, dunque, necessario aver accesso agli strumenti da sviluppatore Twitter indispensabili per estrarre i dati pubblicati sulla piattaforma. Ogni componente del gruppo ha dovuto creare dapprima un account Twitter e in seguito un account Twitter Developers che fornisce all'utente le credenziali d'accesso alle API, tra cui il bearer token. Installata la libreria ed importata all'interno del progetto, è possibile utilizzare la funzione `tweepy.Client()`, autenticandoci tramite `BEARER_TOKEN`, creeremo un oggetto `Client` che fungerà da interfaccia per le API di Twitter

NetworkX

NetworkX è una libreria Python che fornisce strumenti per la creazione e la manipolazione di reti complesse. Può essere utilizzata per creare e analizzare grafi, che sono rappresentazioni matematiche di reti composte da nodi e archi. NetworkX può essere utilizzato per modellare e studiare un'ampia gamma di sistemi complessi, tra cui reti sociali, biologiche e tecnologiche. È uno strumento potente per comprendere la struttura e la dinamica dei sistemi complessi e può aiutare i ricercatori e gli analisti a capire meglio le relazioni tra i vari elementi di una rete. Nel nostro caso utilizzeremo NetworkX per analizzare i followers di un utente, @KevinRoitero, analizzeremo le relazioni di following tra essi e andremo a svolgere varie misurazioni sul grafo risultate.

Svolgimento

In ultimo abbiamo definito due funzioni, una per salvare dati JSON in locale e l'altra invece per leggere file JSON da locale. A questo punto abbiamo potuto finalmente iniziare il nostro progetto! Per il primo punto ci è stato richiesto di utilizzare la libreria Tweepy e scaricare tutti i follower del profilo Twitter con username @KevinRoitero, scaricando per ciascuno di essi : attributi di default, descrizione del profilo,

metriche pubbliche dell'account (`followers_count`, `following_count`, `tweet_count`, `listed_count`) e lo stato di protezione dell'account, e infine salvare il risultato in una unica serializzazione JSON. Per fare ciò abbiamo creato una funzione `get_parsed_users_followers(id)` che sfrutteremo anche nei punti successivi, e che, specificato l'id dell'utente di interesse, ci restituisce, grazie alla funzione `client.get_users_followers` di Tweepy, un dizionario di tutti i suoi followers con salvate tutte le informazioni richieste nella consegna del punto 1. In questo passaggio, inoltre, abbiamo dovuto usare anche il `Paginator` per paginare i risultati, dato che...??????. Abbiamo, quindi, fornito come parametro l'id di @kevinRoitero e salvato tutte le informazioni scaricate in un JSON utilizzando la funzione da noi definita nei primi passi.