

CSCI 4342 Natural Language Processing Semester 2 2023/24

Assignment 2: Text Analytics with Telegram (Group Submission) Due: Sunday, 23 June 2024 (11.59 pm)

Text Analytics on Telegram Data

- i. Register for a **new telegram account** or use an **existing telegram account** of one of your group members. Make sure you have “**pip install telethon**” or other similar packages.
- ii. Apply for **telegram API** (one member is enough). Complete the application form and wait for your approval. *Application may take a few days to be approved, apply early.
- iii. Refer to the links given in Slide #35 (Topic 10 (Part 1): Text Analytics & Sentiment Analysis)
- iv. Try to implement the “**chat-analytics**”, “**Telegrammetry**”, “**Automating Data Collection and Analysis**” or “**Collecting Messages from Telegram**” project. Find groups that discusses about the **current situation in Palestine such as #ceasefire or #freepalestine** possibly from a particular geolocation (e.g Malaysia) and get your telegram data. Limit your data to 200 messages at maximum.
- v. Use the uploaded **data files in italeem** containing the list of **positive, negative, neutral and booster words** to check if each tweet contains any of those terms. Assign and accumulate scores according to Slide #26 for each occurrence of the terms. You may also define your own scoring method that make sense/applicable (e.g., scores between the range of 0 – 1.0 as discussed in class).