Sentiment Analysis for Good

Final Project

Project Introduction

In your project you will use an **API** to access data of some kind, similar to the way we used Genius to access song lyrics and newspaper3k for articles. You can reuse one of those two libraries or you can one of the new APIs listed here. In total, your choices are:

- Newspaper3k
- Genius
- Twitter
- New York Times
  - Movie reviews
  - Book reviews
  - Articles organized by subject matter or reader engagement

APIs

Newspaper3k

A library that allows you to find and analyze newspaper articles by using their hyperlinks.

Genius

Access to song lyrics searchable by artist

Twitter

Can be used to search for tweets by username, keyword, or location. You can download a series of tweets about a specific topic. In the tweets, you have access to the text of the tweet, hashtags, mentions, and more.

APIs

New York Times

The New York Times API offers access to multiple sources of information from their site. In our project, we can explore the API's access to:

- movie reviews
- book reviews
- top stories
- most viewed articles

PROJECT SLIDE

- First we have to decide which library your group will use for your project:
  - Assign each API to a group member who will be in charge of researching that particular API. Every API should be accounted for by at least one group member.
  - Spend 10 minutes researching your chosen API. While you research, answer the following questions:
    - What does this API do?
    - What might be difficult about implementing this API?
    - What kind of data does this API provide? (*For example, Genius provides song lyrics searchable by artist)*

PROJECT SLIDE

- Share the results of your research with your group members.
- Each member will have up to 5 minutes to explain what they know about their library.

PROJECT SLIDE

- Brainstorm project ideas
  - Without speaking to each other, every member spend 10 minutes brainstorming possible project ideas.
  - Think of ways you could use these APIs to create a project that could be helpful to your community or society as a whole, like the research we saw about toxicity in music.
- Each group member should submit ***at least 3 ideas,*** but aim for as many as possible. It's okay to have "bad" ideas, that's part of brainstorming!
- Record your ideas on your worksheet or a piece of paper
- With each idea, note which API you would use to build it.

PROJECT SLIDE

- Share your ideas with your group.
- Group similar ideas together
- Discuss each idea and decide on one to build as a group
  - It can be an idea exactly as someone had brainstormed, a combination of multiple ideas, or something new that the group comes up with through the discussion.

PROJECT SLIDE

- Come up with a problem statement or a question that you would like to answer
  - Note the library you plan to use.
- Example: *“Female authors often face discrimination in publishing. Do book reviewers seem to favor female authors, male authors, or neither? We will use the NY Times Book Reviews to attempt to answer this question.”*

PROJECT SLIDE

- We have to plan our projects on paper before we start coding
- Pseudocode your project.
  - Pseudocode is a way to write out the steps of an algorithm in plain English or a combination of English and code.
  - You can also draw or diagram the project if that helps you visualize it.

Pseudocode

```
avg_polarity = 0
for item in beatles_songs.songs:
    title = item.title
    lyrics = item.lyrics
    lyrics = TextBlob(lyrics)
    polarity = lyrics.sentiment.polarity
    print (title + " " + str(polarity))
    avg_polarity = avg_polarity + polarity
avg_polarity = avg_polarity/len(beatles_songs.songs)
print("Final average: " + str(average_polarity))
```

Pseudocode

```
create a variable for avg polarity
for songs in collection:
    get title
    get lyrics
    make TextBlob object from lyrics
    get polarity
    print song and polarity
    add to total polarity
calculate average polarity
print average
```

PROJECT SLIDE

- Code your project!
  - Open a new Google Collab notebook
  - Use your pseudocode to guide you

- If you need help, refer to:
  - Your old Google Collab notebooks
  - The API reference sheet
  - Your library's official documentation.

- Test your program as you go
  - Trial and error is an important and unavoidable part of the process

- To perform sentiment analysis, you will have to combine the new API code you're learning with what you've already learned in previous units.

PROJECT SLIDE

Once your code is finished and working, prepare a presentation to show the rest of the class.

PROJECT SLIDE

- In your presentation answer the following questions:
  - Explain your problem statement or question, and summarize your program. Feel free to share your pseudocode or actual code.
  - What did you learn from the data?
  - What were the limitations of your program and the library you used?
  - How could you imagine expanding or improving your program as you learn more computer science skills or get access to more APIs?
  - What other information would have been useful to you? What functionality would you like to see in an API?
  - Is there anything interesting about APIs or coding that you learned in the process?

Presentations

DISCUSSION

- What was your favorite project and why?

Unit Learning Goals

You will be able to:

- Describe how an API is useful
- Use an API
- Describe how AI can be used for censorship
- Analyze the pros and cons of using AI for censorship
- Showcase ways sentiment analysis can be used for good

What is NLP good at?
What isn't it good at?
Why?