Sentiment Analysis Without Machine Learning

Putting Everything Together

ACTIVITY SLIDE

Write a program that:

1. Does the usual setup
    a. Loads the newspaper3k library
    b. Adds Article and nltk to your program
    c. Downloads, parses, and provides nlp processing to your article
    d. Saves the text of the article in a variable
2. Starts a count of both positive and negative words.
3. Splits the text of the article into a list of all the words in the article
4. Look through each word of the text
    a. If the word is positive add to the number of positive words seen
    b. If the word is negative add to the number of negative words seen
5. Check if there are more positive words or negative words.

Extension

Right now, you load one article into your program at a time. How about looking at several articles in one program!

1. Create a list of URLs of articles.
2. For each article in that list, download it, parse, it, and look through each word, deciding if the review is positive or negative.
3. Count the number of positive reviews.
4. Count the number of negative reviews.
5. Are there more positive reviews or more negative reviews? Print your answer to the screen.

DISCUSSION

- Where are some places your program might still have problems when it comes to sentiment analysis? How might you fix these problems?
- How about 2-grams (aka bigrams)? Our program currently view “not good” as just the words “not” and “good” rather than “not good” as a whole.

Unit Learning Goals

You will be able to:

- Write and run a program in Google Colab Notebook
- Write a program using conditionals
- Write a program that iterates through a list
- Write a program that stores input in a variable to be used later
- Write a program that sends information to the screen
- Write a program that changes numbers stored in variables.
- Debug a program that does all of the above