Reference: Analyzing Messy Data Sentiment with Python and nltk:

<https://www.twilio.com/blog/2017/09/sentiment-analysis-python-messy-data-nltk.html>

Install the **nltk** library ( <http://www.nltk.org/)> as shown in the Environment Setup section

Download the tweets data files (neg\_tweets.txt, pos\_tweets.txt)

Use the **Jupyter** notebook if you like ( twilio-nltk-scikit-learn.ipynb **)**

Follow the examples in sections:

Preparing the Data

Building a Classifier

Classification

Accuracy

**Select 10 random tweets from Twitter:**

**Classify the tweets as 'pos' or 'neg’**

**Discuss the accuracy of the classification (i.e., is the classification "correct"?)**

**How can the accuracy be improved?**