TextBlob: Simplified Text Processing

TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.

TextBlob aims to provide access to common text-processing operations through a familiar interface. You can treat TextBlob objects as if they were Python strings that learned how to do Natural Language Processing.

Terminologies in Text mining

- Noun phrase extraction
- Part-of-speech tagging
- Sentiment analysis
- Classification (Naive Bayes, Decision Tree)
- Language translation and detection powered by Google Translate
- Tokenization (splitting text into words and sentences)
- Word and phrase frequencies
- Parsing
- **n**-grams
- Word inflection (pluralization and singularization) and lemmatization
- Spelling correction
- Add new models or languages through extensions
- WordNet integration

Get it now

```
$ pip install -U textblob
$ python -m textblob.download_corpora
```

One of the more powerful aspects of the NLTK module is the Part of Speech tagging that it can do for you. This means labeling words in a sentence as nouns, adjectives, verbs...etc. Even more impressive, it also labels by tense, and more. Here's a list of the tags, what they

mean, and some examples:

POS tag list:

CC	coordinating conjunction

CD	cardinal digit

DT	determiner

EX	existential there (like: "there is" ... think of it like "there exists")

FW	foreign word

IN	preposition/subordinating conjunction

JJ	adjective 'big'

JJR	adjective, comparative 'bigger'

JJS	adjective, superlative 'biggest'

LS	list marker 1)

MD	modal could, will

NN	noun, singular 'desk'

NNS	noun plural 'desks'

NNP	proper noun, singular 'Harrison'

NNPS	proper noun, plural 'Americans'

PDT	predeterminer 'all the kids'

POS	possessive ending parent\'s

PRP	personal pronoun I, he, she

PRP$	possessive pronoun my, his, hers

RB	adverb very, silently,

RBR	adverb, comparative better

RBS	adverb, superlative best

RP	particle give up

TO	to go 'to' the store.

UH	interjection errrrrrrm

VB	verb, base form take

VBD	verb, past tense took

VBG	verb, gerund/present participle taking

VBN	verb, past participle taken

VBP	verb, sing. present, non-3d take

VBZ	verb, 3rd person sing. present takes

WDT	wh-determiner which

WP	wh-pronoun who, what

WP$	possessive wh-pronoun whose

WRB	wh-abverb where, when

Sentiment Analysis

Sentiment analysis is used to determine the emotional tone behind words to gain understanding of

the attitudes, opinions and emotions expressed within an online mention. It is also known as opinion mining, deriving the opinion or attitude of a speaker

What is the difference between Polarity and Subjectivity?

Polarity : It simply means emotions expressed in a sentence.
Emotions are closely related to sentiments. The strength of a sentiment or opinion is typically linked to the intensity of certain emotions, e.g., joy and anger.
Opinions in sentiment analysis are mostly evaluations(although not always).

According to consumer behaviour research, evaluations can be broadly categorized into two types:
1. Rational evaluations
2. Emotional evaluations.

1. Rational evaluation: Such evaluations are from rational reasoning, tangible beliefs, and utilitarian attitudes. For example, the following sentences

Express rational evaluations: "The voice of this phone is clear," "This car is worth the price," and "I am happy with this car."

2. Emotional evaluation: Such evaluations are from non-tangible and emotional responses to entities which go deep into people's state of mind.

For example, the following sentences express emotional evaluations: "I love iPhone," "I am so angry with their service people" and "This is the best car ever built."

To make use of these two types of evaluations in practice, we can design 5 sentiment ratings, emotional negative (-2), rational negative (-1), neutral (0), rational positive (+1), and emotional positive (+2). In practice, neutral often means no opinion or sentiment expressed.

Subjectivity
Subjective sentence expresses some personal feelings, views, or beliefs.

An example
subjective sentence is "I like iPhone." Subjective expressions come in many forms, e.g., opinions, allegations, desires, beliefs, suspicions, and speculations.
A subjective sentence may not express any sentiment.

For example, "I think that he went home" and "I want a camera that can take good photos" are a subjective sentences, but does not express any sentiment.

The sentiment property returns a namedtuple of the form Sentiment(polarity,subjectivity).
The polarity score is a float within the range [-1.0, 1.0].
The subjectivity is a float within the range [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective.

TextBlob Sentiment: Calculating Polarity and Subjectivity

Sunday June 7, 2015

The TextBlob package for Python is a convenient way to do a lot of Natural Language Processing (NLP) tasks. For example:

```
from textblob import TextBlob

TextBlob("not a very great calculation").sentiment
## Sentiment(polarity=-0.3076923076923077, subjectivity=0.5769230769230769)
```

This tells us that the English phrase "not a very great calculation" has a *polarity* of about -0.3, meaning it is slightly negative, and a *subjectivity* of about 0.6, meaning it is fairly subjective.