TextBlob: Simplified Text Processing

Release v0.16.0. (Changelog)

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.

```
from textblob import TextBlob

text = '''
The titular threat of The Blob has always struck me as the ultimate movie
monster: an insatiably hungry, amoeba-like mass able to penetrate
virtually any safeguard, capable of--as a doomed doctor chillingly
describes it--"assimilating flesh on contact.
Snide comparisons to gelatin be damned, it's a concept with the most
devastating of potential consequences, not unlike the grey goo scenario
proposed by technological theorists fearful of
artificial intelligence run rampant.
'''

blob = TextBlob(text)
blob.tags           # [('The', 'DT'), ('titular', 'JJ'),
                    #  ('threat', 'NN'), ('of', 'IN'), ...]

blob.noun_phrases   # WordList(['titular threat', 'blob',
                    #           'ultimate movie monster',
                    #           'amoeba-like mass', ...])

for sentence in blob.sentences:
    print(sentence.sentiment.polarity)
# 0.060
# -0.341
```

TextBlob stands on the giant shoulders of NLTK and pattern, and plays nicely with both.

Features

- Noun phrase extraction
- Part-of-speech tagging
- Sentiment analysis
- Classification (Naive Bayes, Decision Tree)
- 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
```

Ready to dive in? Go on to the Quickstart guide.

Guide

- License
- Installation
  - Installing/Upgrading From the PyPI
  - With conda
  - From Source
  - Get the bleeding edge version
  - Migrating from older versions (<=0.7.1)
- Tutorial: Quickstart
  - Create a TextBlob
  - Part-of-speech Tagging
  - Noun Phrase Extraction
  - Sentiment Analysis
  - Tokenization
  - Words Inflection and Lemmatization
  - WordNet Integration
  - WordLists
  - Spelling Correction
  - Get Word and Noun Phrase Frequencies
  - Parsing
  - TextBlobs Are Like Python Strings!
  - **n**-grams
  - Get Start and End Indices of Sentences
- Tutorial: Building a Text Classification System
  - Loading Data and Creating a Classifier
  - Classifying Text
  - Classifying TextBlobs
  - Evaluating Classifiers
  - Updating Classifiers with New Data
  - Feature Extractors
  - Next Steps
- Advanced Usage: Overriding Models and the Blobber Class
  - Sentiment Analyzers
  - Tokenizers
  - Noun Phrase Chunkers
  - POS Taggers
  - Parsers
  - Blobber: A TextBlob Factory
- Extensions
  - Available extensions
- API Reference
  - Blob Classes
  - Base Classes
  - Tokenizers
  - POS Taggers
  - Noun Phrase Extractors
  - Sentiment Analyzers
  - Parsers
  - Classifiers
  - Blobber
  - File Formats
  - Wordnet
  - Exceptions

v: dev

Project info

- Changelog
- Authors
- Contributing guidelines