TextPy

Introduction

TextPy is a Python package which can be used in extracting contents such as words, sentences, dates, numbers, telephone, URLs, and misspelled words from a given text.

System Requirement

You need a Python 2.7+ interpreter to install and run TextPy.

Platform Support

Currently, TextPy is supported only in Linux and Mac.

Installation Guidelines

Download TextPy

Open your terminal and clone the repository using the command below.

```
git clone https://github.com/shivapbhusal/textpy.git
```

Alternatively, you can also download the project from the link below.

```
https://github.com/shivapbhusal/textpy
```

Install TextPy

Navigate to the root of the project directory, and enter the command below.

```
sudo python setup.py install
```

Test and validate the installation using your Python terminal.

```
import textpy
```

```
print(textpy.words("Hello World !"))
```

Output

```
['Hello','World']
```

Function References

textpy.words(text)

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A list of all the words in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_words = textpy.words("Hello World !") # Returns ['Hello','World']
```

textpy.sentences(text):

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A list of all the sentences in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_sentences = textpy.sentences("Hello World.I am using TextPy.") # Returns ['Hello
World.','I am using TextPy.']
```

textpy.dates(text):

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A list of all the dates in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_dates = textpy.dates("My name is John Doe. Today is 05/20/2019") # Returns
['05/20/2019']
```

textpy.numbers(text):

- *Argument*: A string representing the text to be analyzed.

- *Returns*: A list of all the numbers in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_numbers = textpy.numbers('Hello World 123!') # Returns ['123']
```

textpy.telephone(text):

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A list of all the telephone numbers in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_words = textpy.telephone('My number is 319-378-8183') # Returns ['319-378-8183']
```

textpy.urls(text):

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A list of all the URLs in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_words = textpy.urls('Explore https://github.com/.') # Returns
'https://github.com/'
```

textpy.misspelled_words(text):

- *Argument*: A string representing the text to be analyzed.
- *Returns*: A List of all the misspelled words in the text.
- *Exception*: Throws *TypeError* exception if the argument is not a string.

```
import textpy
all_words = textpy.misspelled_words('Hello Worrlld') # Returns ['Worrlld']
```

Example Usage

Calculating the average length of sentences in a text.

First, get all the sentences from the text using TextPy.

```
import textpy

text = "Hello World.I am using TextPy."
sentences = textpy.sentences(text)
```

Then, compute the length of all the sentences and take the average.

```
total = 0
for sen in sentences:
        total += len(sen)
avg_length = total/len(sentences)
```

Complete Code

```
"""
This program computes the average length of sentences using TextPy.
"""
import textpy

text = "Hello World.I am using TextPy."
sentences = textpy.sentences(text)
total = 0
for sen in sentences:
        total += len(sen)
avg_length = total/len(sentences) # avg_length = 15
```

Computing Spelling Accuracy

First, get all the words from the text in a list.

```
import textpy
text = 'It was the best of the times, it was the worrrst of the times.'
sentences = textpy.words(text)
```

Then, get the list of misspelled words.

```
misspelled = textpy.misspelled_words(text)
```

Finally, compute the the spelling accuracy.

```
spelling_accuracy = float(len(sentences)-len(misspelled))/float(len(sentences))
```

Complete Code

```
"""
This program computes the spelling accuracy of a text using TextPy.
"""
import textpy

text = 'It was the best of the times, it was the worrrrrst of the times.'
sentences = textpy.words(text)
misspelled = textpy.misspelled_words(text)
spelling_accuracy = float(len(sentences)-len(misspelled))/float(len(sentences))
# spelling_accuracy = 0.928
```