

Spacy	NLTK
Spacy is Object Oriented	NLTK is mainly a string processing library
Spacy is user friendly	NLTK is also user friendly but probably less user friendly compared to Spacy
Provides most efficient NLP algorithm for a given task. Hence if you care about the end result, go with Spacy	Provides access to many algorithms. If you care about specific algo and customizations go with NLTK
Spacy is new library and has a very active user community	NLTK is old library. User community as active as Spacy

Difference Btw SPACY & NLTK

Installation instructions

pip install spacy

python -m spacy download en

pip install nltk

▾ Sentence & Word Tokenization In Spacy

```
import spacy
```

▾ For language code in SPACY

see below URI

<https://spacy.io/usage/models>

```
nlp = spacy.load("en_core_web_sm")
```

```
doc = nlp("Dr. abhijeet etc. Like SPACY. Hulk loves NLTK")
```

▾ Spacy Sentence tokenization

```
for sentence in doc.sents:
    print(sentence)
```

```
Dr. abhijeet etc.
Like SPACY.
Hulk loves NLTK
```

▾ Spacy Word tokenization

```
for sentence in doc.sents:
    for word in sentence:
        print(word)
```

```
Dr.
abhiyeet
etc
.
Like
SPACY
.
Hulk
loves
NLTK
```

▾ Sentence & Word Tokenization In NLTK

```
from nltk.tokenize import sent_tokenize
import nltk
nltk.download('punkt')
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
True
```

```
sent_tokenize("Mr. abhiyeet Like SPACY. Hulk loves NLTK")
```

```
['Dr. abhiyeet etc.', 'Like SPACY.', 'Hulk loves NLTK']
```

```
from nltk.tokenize import word_tokenize
```

```
word_tokenize("Mr. abhiyeet Like SPACY. Hulk loves NLTK")
```

```
['Mr.', 'abhiyeet', 'Like', 'SPACY', '.', 'Hulk', 'loves', 'NLTK']
```

For spacy document Click below link

<https://spacy.io/api/doc>

