

▼ Parts of Speech

▼ Tag|Meaning|English Examples

ADJ adjective new, good, high, special, big, local
ADP adposition on, of, at, with, by, into, under
ADV adverb really, already, still, early, now
CONJ conjunction and, or, but, if, while, although
DET determiner, article the, a, some, most, every, no, which
NOUN noun year, home, costs, time, Africa
NUM numeral twenty-four, fourth, 1991, 14:24
PRT particle at, on, out, over per, that, up, with
PRON pronoun he, their, her, its, my, I, us
VERB verb is, say, told, given, playing, would
. punctuation marks . , ; !
X other ersatz, esprit, dunno, gr8, univeristy

```
text = "TON 618 (short for Tonantzintla 618) is a hyperluminous, broad-absorption-line, radio-loud quasar and Lyman-alpha blob located near the border of the constellation Coma Berenices"
```

▼ Importing necessary dependencies

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

▼ Word Tokenization

```
words = word_tokenize(text)
```

▼ Parts of Speech Tagging

```
tagged_words = nltk.pos_tag(words, tagset = 'universal')
```

```
tagged_words

[('TON', '.'),
 ('618', 'NUM'),
 ('(', '.'),
 ('short', 'ADJ'),
 ('for', 'ADP'),
 ('Tonantzintla', 'NOUN'),
 ('618', 'NUM'),
 (')', '.'),
 ('is', 'VERB'),
 ('a', 'DET'),
 ('hyperluminous', 'ADJ'),
 (',', '.'),
 ('broad-absorption-line', 'ADJ'),
 ('', '.'),
 ('radio-loud', 'ADJ'),
 ('quasar', 'NOUN'),
 ('and', 'CONJ'),
 ('Lyman-alpha', 'NOUN'),
 ('blob', 'NOUN'),
 ('located', 'VERB'),
 ('near', 'ADP'),
 ('the', 'DET'),
 ('border', 'NOUN'),
 ('of', 'ADP'),
 ('the', 'DET'),
 ('constellations', 'NOUN'),
 ('Canes', 'NOUN'),
 ('Venatici', 'NOUN'),
 ('and', 'CONJ'),
 ('Coma', 'NOUN'),
 ('Berenices', 'NOUN'),
```

```
(',', '.'),  
( 'with', 'ADP'),  
( 'the', 'DET'),  
( 'projected', 'VERB'),  
( 'comoving', 'NOUN'),  
( 'distance', 'NOUN'),  
( 'of', 'ADP'),  
( 'approximately', 'ADV'),  
( '18.2', 'NUM'),  
( 'billion', 'NUM'),  
( 'light-years', 'NOUN'),  
( 'from', 'ADP'),  
( 'Earth', 'NOUN'),  
( '.', '.')] ]
```

```
for t in tagged_words:  
    print(t)
```

```
( 'TON', '.')  
( '618', 'NUM')  
( '(', '.')  
( 'short', 'ADJ')  
( 'for', 'ADP')  
( 'Tonantzintla', 'NOUN')  
( '618', 'NUM')  
( ')', '.')  
( 'is', 'VERB')  
( 'a', 'DET')  
( 'hyperluminous', 'ADJ')  
( ',', '.')  
( 'broad-absorption-line', 'ADJ')  
( ',', '.')  
( 'radio-loud', 'ADJ')  
( 'quasar', 'NOUN')  
( 'and', 'CONJ')  
( 'Lyman-alpha', 'NOUN')  
( 'blob', 'NOUN')  
( 'located', 'VERB')  
( 'near', 'ADP')  
( 'the', 'DET')  
( 'border', 'NOUN')  
( 'of', 'ADP')  
( 'the', 'DET')  
( 'constellations', 'NOUN')  
( 'Canes', 'NOUN')  
( 'Venatici', 'NOUN')  
( 'and', 'CONJ')  
( 'Coma', 'NOUN')  
( 'Berenices', 'NOUN')  
( ',', '.')  
( 'with', 'ADP')  
( 'the', 'DET')  
( 'projected', 'VERB')  
( 'comoving', 'NOUN')  
( 'distance', 'NOUN')  
( 'of', 'ADP')  
( 'approximately', 'ADV')  
( '18.2', 'NUM')  
( 'billion', 'NUM')  
( 'light-years', 'NOUN')  
( 'from', 'ADP')  
( 'Earth', 'NOUN')  
( '.', '.')
```