***POS Tagging & Chunking***

***What does Chunking mean?***

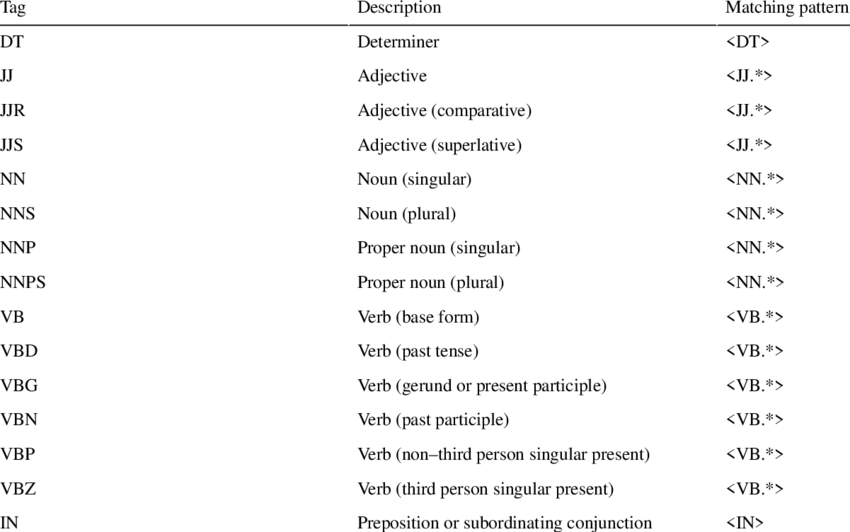
* Chunking, AKA: Partial/Shallow Parsing
* A process of meaningful extraction of short phrases from the sentence that are tagged with Part of Speech (POS)

***What does POS tagging mean:***

* A process to categorize words in a text (corpus) in correspondence with a particular POS, depending on the definition of the word and its context
* Why not tell someone ?

Why: adverb not: adverb tell: verb someone: noun ?: punctuation mark

* However, having to constantly write out these full terms when we perform text analysis can be very overwhelming, especially as the size of corpus grows. So, we use a short representation referred as “tags” to represent categories.



***Difference between POS tagging and Chunking:***

* Chunking is used to get the required phrases from a given sentence. However, POS tagging can be used only to spot the POS that every word of the sentence belongs to.

***Types of Chunking:***

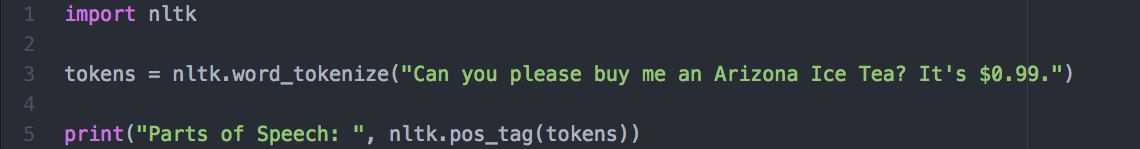
* Chunking up: Takes the important and brief stuff in the sentence
* Chunking down: The opposite; Gets the detailed information

For example, Chunking can be done to identify and thus group noun phrases or nouns alone, adjectives or adjective phrases and so on. Consider this:

“I had burgers and pastries for breakfast.”

In this case, if we wish to group or chunk noun phrases, we will get “burgers”, “pastries” and “lunch”, which are nouns

***Implementation of POS tagging in Python:***

* <https://medium.com/@gianpaul.r/tokenization-and-parts-of-speech-pos-tagging-in-pythons-nltk-library-2d30f70af13b>
* 

***Implementation of Chunking in Python:***

* If we want to extract to all verb phrases from the given piece of text, chunking is done with the help of Regular Expression and using an NLTK Library called RegexpParser, used to make a group of “Noun Phrases”. This helps us get the ENTITY