Christine Miller

April 8,2020

CPS3320

Write-up for Natural Language Toolkit(NLTK)

Natural Language Toolkit(NLTK) is a platform for building python programs that work with human language data that is applied to natural language processing. Its central focus is human language that can be used in the aspects of learning English phrases or words, or the use or chat bots for customer service purposes, or even simple things like statistics of the top words written in a particular series of books or on a social media.

There may be many practical applications for NLTK one such example could be from publishing companies to use this in generating a tally of a particular writer’s use of certain words in an exceeding amount that they deem is unnecessary. Another practical example would be in the use of chatbots or automated messaging systems in customer service whether it is for checking the status of an account or order, or having issues with making a payment, or also a way of lessening the unnecessary line of people waiting for human counterpart to guide through fixing an issue.

Some functions of NLTK are tokenization, lemmatization, and part of speech (pos) tagging. The tokenization function is the breaking down of certain items into subitems if possible given a set of rules. The lemmatization function only removes the affixes if the end word result is in the dictionary. Part of speech (POS) tagging function tags the word by creating a tuple where the first is the word and the second part is the tag, which would indicate if it is a noun(NN), noun plural(NNS), noun proper,(NNP), adjective(JJ), .etc. Those are only a few functions that is in NLTK, the link is <https://www.nltk.org/api/nltk.html> .