**Documentation for**

**ARI1102 – An Online NLTK Demo –**

**Assignment 2020/2021**

Nathan Bonavia Zammit – 0198402L

**Overview of which parts of the specification are complete:**

* The server is running correctly.
* The Web Interface is really basic and simple.
* The POS visualisation was not done as I could not figure out how to output it, however from the console one can see that it worked.
* The NER visualization was not done as I could not figure out how to output it, however from the console one can see that it worked.

**A list of the third party components and material used:**

* Filezilla is used to upload the html and python scripts to the server.
* Bootstrap 4 is used for the web interface.

**Overview of how the core components work:**

The web interface is really simple. It is asks the user to input text correctly into the textarea that is provided. The text that is inputted is then checked if it is English or not. If it is English a pop-up alert will not appear, if it is another language a pop-up alert appears telling the user that the input must be English. If the inputted text is inputted incorrectly, a message is outputted telling the user that an error occurred, if the user does not even input text a message is outputted telling the user that text must be inputted. After the user inputs text correctly, and in English, the text is split into sentences, and then tokenized using part-speech-tags. This can be seen in the console under POS1 where each token (word) is given a small description regarding whether they are adjectives, nouns, and so forth. I printed a few rows and I gave each tag a different colour. Under POS2 the categories are finer and with more detailed input, such as singular, plural and so forth. Under Names, one can find POS2 however this time any people, organizations, or locations can be identified.