

Project 1 for CS421 – University of Illinois at Chicago

Name 1: kwu3@uic.edu

Name 2: aserra23@uic.edu

-----Setup-----

Before running anything go to the following website and download the stanford parser:

<https://nlp.stanford.edu/software/stanford-parser-full-2018-02-27.zip>

We don't know if nltk differentiates between versions but if it does not then rename the file as stanford-parser-full-2018-02-27/ after it is unzip if you have a newer version. The same thing could be done with an older version.

Make sure you place the new unzipped folder in the execution/ folder.

Make sure to delete folder parameter files like .DS_Store (for Mac users) in the following paths:

execution/

input/testing/essays/

input/training/essays/

In the execution folder, give the following command on your command prompt terminal (Linux):

yourComputerName\$./run.sh

Alternatively, you can give the following command (in either Windows or Linux):

Please install required libraries using the following commands:

pip3 install nltk

pip3 install numpy

Navigate to the execution folder, give the following command on your command prompt(on linux systems you need to include the ./):

python3 ./cs421Project.py

-----Technique-----

Break the string into sentences. use nltk to get tags on each word and split the sentences into a word list and a tag list. Use the brown corpus to do spelling check on all words. Use the tag list to perform the other parts of grading. Make combinations of the verb-verb tags and subject-verb tags. If there is no verb in a sentence, it missed a verb. If the verbs tag combo is on the list created for improper combinations then, it is a mistake. Also, if the pronoun, proper noun and/or noun to verb singular/plural rule do not match then, it is

mistake. Also, the number of words and verbs of the essay are considered too because the less of those, the less chance to make mistake. On the other hand, the ratio of a mistake is higher if there is less word.