**Additional Features**

The following are the features that I have used

1. **right\_word.isalpha()**: This feature checks if the word to the right of the period is an alphabet or not. I have used this feature because, if the word is an alphabet than it might probably be an End of sentence (EOS)
2. **left\_word.isalpha()**: This feature checks if the word to the left of the period is an alphabet or not. I have used this feature because, if the word is an alphabet than it might probably be an End of sentence (EOS) and vice-versa
3. **right\_word.isdigit()**: This feature checks if the word to the right of the period is a number or not. I have used this feature because if the word is a number and not an alphabet, it might be Not an End of Sentence (NEOS)

**Program Flow**:

I have opened the file and started reading every line. I have appended every line to a list named “data” only when its length is greater than two else it would be an empty line(“\n”) and I have split them near spaces and have checked if its NEOS or EOS and applied all the features required including my additional features. At the end I have replaced EOS with ‘1’ and NEOS with ‘0’ to make it compatible with the data frame. I have created two dataframes namely, train\_df and test\_df and used LabelEncoder that assigns numbers to the labels for easy comparison. Finally, I was able to train the classifier and **achieve** the desired output.

**Accuracy:**

Accuracy of the system using all the eight features: 99.07%

Accuracy of the model with only the given features:98.94%

Accuracy of the model with only my additional features: 98.99%