-first we tokenize using nltk.word\_tokenize(data) into single words and giving the result in a list contain single word.

-secondly we use nltk.ngrams(words\_lower, 2) to contain 2 words together in list

-thirdly we calculate probability of single word and store it in dictionary that having

key is word and value is number of times that appear in file

-fourthly we calculate probability of two words and store it in dictionary that having

key is two words and value is number of times that appear in file

And this is an equation we used : P(B|A)=C(A|B)/C(A)

we calculate probability of upcoming word based on the rule and store the result in dictionary then we sort this as descending order and the word that have higer probability will come first as it appear most

finally we make GUI for our results as we show it by screen shots

Screensho



















