

# CSE6060

## Statistical Natural Language Processing

### Activity 1

Name : Kavianand G

Reg. No. : 19MAI0050

Date : 13 – June – 2020

### Explore - Stemming, Tokenizing, POS Tagging

In [1]:

```
1 from nltk.stem import PorterStemmer
2 from nltk.stem import LancasterStemmer
3 from nltk.stem import RegexpStemmer
4 from nltk.stem import SnowballStemmer
5 from nltk.stem import WordNetLemmatizer
6 from nltk.tokenize import sent_tokenize, word_tokenize
7 from nltk import pos_tag
```

In [2]:

```
1 porter = PorterStemmer()
2 lancaster=LancasterStemmer()
3 #A list of words to be stemmed
4 word_list = ["friend", "friendship", "friends", "friendships", "stabil", "destabilize", "misunderstanding", "railroad", "moonlight", "football"]
5 print("{0:20}{1:20}{2:20}".format("Word", "Porter Stemmer", "lancaster Stemmer"))
6 print("-" * 60)
7 for word in word_list:
8     print("{0:20}{1:20}{2:20}".format(word, porter.stem(word), lancaster.stem(word)))
```

Word	Porter Stemmer	lancaster Stemmer
friend	friend	friend
friendship	friendship	friend
friends	friend	friend
friendships	friendship	friend
stabil	stabil	stabl
destabilize	destabil	dest
misunderstanding	misunderstand	misunderstand
railroad	railroad	railroad
moonlight	moonlight	moonlight
football	footbal	footbal

Using Shakespeare ceasar text file for stemming, Tokenizing

In [3]:

```
1 from nltk.corpus import gutenberg
2 abc = gutenberg.raw('shakespeare-caesar.txt')
3 print(abc)
```

[The Tragedie of Julius Caesar by William Shakespeare 1599]

Actus Primus. Scoena Prima.

Enter Flavius, Murellus, and certaine Commoners ouer the Stage.

Flavius. Hence: home you idle Creatures, get you home:  
Is this a Holiday? What, know you not  
(Being Mechanicall) you ought not walke  
Vpon a labouring day, without the signe  
Of your Profession? Speake, what Trade art thou?  
Car. Why Sir, a Carpenter

Mur. Where is thy Leather Apron, and thy Rule?  
What dost thou with thy best Apparrell on?  
You sir, what Trade are you?

Cobl. Truely Sir, in respect of a fine Workman, I am  
but as you would say, a Cobler

## Using Shakespeare ceasar text file for stemming, Tokenizing

### Sentence Tokenization using #sent\_tokenize option

In [4]:

```
1 sentence = sent_tokenize(abc)
2 print(sentence)
```

['[The Tragedie of Julius Caesar by William Shakespeare 1599]\n\n\nActus P  
rimus.', 'Scoena Prima.', 'Enter Flavius, Murellus, and certaine Commoners  
ouer the Stage.', 'Flavius.', 'Hence: home you idle Creatures, get you hom  
e:\nIs this a Holiday?', 'What, know you not\n(Being Mechanicall) you ough  
t not walke\nVpon a labouring day, without the signe\nOf your Professio  
n?', 'Speake, what Trade art thou?', 'Car.', 'Why Sir, a Carpenter\n\n M  
ur.', 'Where is thy Leather Apron, and thy Rule?', 'What dost thou with th  
y best Apparrell on?', 'You sir, what Trade are you?', 'Cobl.', 'Truely Si  
r, in respect of a fine Workman, I am\nbut as you would say, a Cobler\n\n  
Mur.', 'But what Trade art thou?', 'Answer me directly\n\n Cob.', 'A Tra  
de Sir, that I hope I may vse, with a safe\nConscience, which is indeed Si  
r, a Mender of bad soules\n\n Fla. What Trade thou knaue?', 'Thou naught  
y knaue,\nwhat Trade?', 'Cobl.', 'Nay I beseech you Sir, be not out with m  
e: yet\nif you be out Sir, I can mend you\n\n Mur.', 'What mean'st thou  
by that?', 'Mend mee, thou\sawcy Fellow?', 'Cob.', 'Why sir, Cobble you\n  
Fla. Thou art a Cobler, art thou?', 'Cob.', 'Truly sir, all that I li  
ue by, is with the Aule: I\nmeddle with no Tradesmans matters, nor womens  
matters;\nbut withal I am indeed Sir, a Surgeon to old shooes:\nwhen they  
are in great danger, I recouer them.', 'As proper\nmen as euer trod vpon N  
. . .']

### Word Tokenization using #word\_tokenize option

In [5]:

```
1 words=word_tokenize(abc)
2 print(words)
```

```
[['', 'The', 'Tragedie', 'of', 'Julius', 'Caesar', 'by', 'William', 'Shake  
speare', '1599', ''], 'Actus', 'Primus', '.', 'Scoena', 'Prima', '.', 'Ent  
er', 'Flauius', ',', 'Murellus', ',', 'and', 'certaine', 'Commoners', 'oue  
r', 'the', 'Stage', '.', 'Flauius', '.', 'Hence', ':', 'home', 'you', 'idl  
e', 'Creatures', ',', 'get', 'you', 'home', ':', 'Is', 'this', 'a', 'Holid  
ay', '?', 'What', ',', 'know', 'you', 'not', '(', 'Being', 'Mechanicall',  
)', 'you', 'ought', 'not', 'walke', 'Vpon', 'a', 'labouring', 'day', ',',  
'without', 'the', 'signe', 'Of', 'your', 'Profession', '?', 'Speake', ',',  
'what', 'Trade', 'art', 'thou', '?', 'Car', '.', 'Why', 'Sir', ',', 'a',  
'Carpenter', 'Mur', '.', 'Where', 'is', 'thy', 'Leather', 'Apron', ',', 'a  
nd', 'thy', 'Rule', '?', 'What', 'dost', 'thou', 'with', 'thy', 'best', 'A  
pparrell', 'on', '?', 'You', 'sir', ',', 'what', 'Trade', 'are', 'you',  
'?', 'Cobl', '.', 'Truely', 'Sir', ',', 'in', 'respect', 'of', 'a', 'fin  
e', 'Workman', ',', 'I', 'am', 'but', 'as', 'you', 'would', 'say', ',',  
'a', 'Cobler', 'Mur', '.', 'But', 'what', 'Trade', 'art', 'thou', '?', 'An  
swer', 'me', 'directly', 'Cob', '.', 'A', 'Trade', 'Sir', ',', 'that',  
'I', 'hope', 'I', 'may', 'vse', ',', 'with', 'a', 'safe', 'Conscience',  
,', 'which', 'is', 'indeed', 'Sir', ',', 'a', 'Mender', 'of', 'bad', 'sou  
les', 'Fla.', 'What', 'Trade', 'thou', 'knaue', '?', 'Thou', 'naughty', 'k
```

## Tagged using #pos\_tag option

In [6]:

```
1 tagged = pos_tag(words)
2 print(tagged)
```

```
rt', 'NN'), ('thou', 'NN'), ('?', '.'), ('Car', 'NNP'), ('.', '.'), ('Wh  
y', 'WRB'), ('Sir', 'NNP'), (',', ','), ('a', 'DT'), ('Carpenter', 'NNP'),  
( 'Mur', 'NNP'), ('.', '.'), ('Where', 'NNP'), ('is', 'VBZ'), ('thy', 'J  
J'), ('Leather', 'NNP'), ('Apron', 'NNP'), (',', ','), ('and', 'CC'), ('th  
y', 'JJ'), ('Rule', 'NNP'), ('?', '.'), ('What', 'WP'), ('dost', 'VBD'),  
( 'thou', 'NN'), ('with', 'IN'), ('thy', 'JJ'), ('best', 'JJ'), ('Apparrel  
l', 'NN'), ('on', 'IN'), ('?', '.'), ('You', 'PRP'), ('sir', 'VBP'), (',',  
,'), ('what', 'WP'), ('Trade', 'NNP'), ('are', 'VBP'), ('you', 'PRP'),  
( '?', '.'), ('Cobl', 'NNP'), ('.', '.'), ('Truely', 'RB'), ('Sir', 'NNP'),  
(',', ','), ('in', 'IN'), ('respect', 'NN'), ('of', 'IN'), ('a', 'DT'),  
( 'fine', 'JJ'), ('Workman', 'NNP'), (',', ','), ('I', 'PRP'), ('am', 'VB  
P'), ('but', 'CC'), ('as', 'IN'), ('you', 'PRP'), ('would', 'MD'), ('say',  
'VB'), (',', ','), ('a', 'DT'), ('Cobler', 'NNP'), ('Mur', 'NNP'), ('.',  
,'), ('But', 'CC'), ('what', 'WP'), ('Trade', 'NNP'), ('art', 'NN'), ('th  
ou', 'NN'), ('?', '.'), ('Answer', 'NNP'), ('me', 'PRP'), ('directly', 'R  
B'), ('Cob', 'NNP'), ('.', '.'), ('A', 'NNP'), ('Trade', 'NNP'), ('Sir',  
'NNP'), (',', ','), ('that', 'IN'), ('I', 'PRP'), ('hope', 'VBP'), ('I',  
'PRP'), ('may', 'MD'), ('vse', 'VB'), (',', ','), ('with', 'IN'), ('a', 'D  
T'), ('safe', 'JJ'), ('Conscience', 'NNP'), (',', ','), ('which', 'WDT'),  
( 'is', 'VBP'), ('indeed', 'RB'), ('Sir', 'NNP'), ('.', '.'), ('a', 'DT'),
```

## Apply Stemming on shakespeare work

In [7]:

```
1 porter = PorterStemmer()
2 lancaster=LancasterStemmer()
3 print("{0:20}{1:20}{2:20}".format("Word", "Porter Stemmer", "lancaster Stemmer"))
4 print("-" * 60)
5 for word in words:
6     print("{0:20}{1:20}{2:20}".format(word,porter.stem(word),lancaster.stem(word)))
```

Word	Porter Stemmer	lancaster Stemmer
[	[	[
The	the	the
Tragedie	tragedi	tragedy
of	of	of
Julius	juliu	juli
Caesar	caesar	caes
by	by	by
William	william	william
Shakespeare	shakespear	shakespear
1599	1599	1599
]	]	]
Actus	actu	act
Primus	primu	prim
.	.	.
Scoena	scoena	scoen
Prima	prima	prim
.	.	.
Enter	enter	ent

In [8]:

```
1 token_words=word_tokenize(abc)
2 stem_sentence=[]
3 for word in token_words:
4     stem_sentence.append(porter.stem(word))
5     stem_sentence.append(" ")
6 #print(stem_sentence)
7 print( "".join(stem_sentence))
```

[ the tragedi of juliu caesar by william shakespear 1599 ] actu primu . sc  
oena prima . enter flauiu , murellu , and certain common over the stage .  
flauiu . henc : home you idl creatur , get you home : Is thi a holiday ? w  
hat , know you not ( be mechanical ) you ought not walk vpon a labour day  
, without the sign Of your profess ? speak , what trade art thou ? car . w  
hi sir , a carpent mur . where is thi leather apron , and thi rule ? what  
dost thou with thi best apparel on ? you sir , what trade are you ? cobl  
. trueli sir , in respect of a fine workman , I am but as you would say ,  
a cobbler mur . but what trade art thou ? answer me directli cob . A trade  
sir , that I hope I may vse , with a safe conscienc , which is inde sir ,  
a mender of bad soul fla. what trade thou knaue ? thou naughti knaue , wha  
t trade ? cobl . nay I beseech you sir , be not out with me : yet if you b  
e out sir , I can mend you mur . what mean'st thou by that ? mend mee , th  
ou sawci fellow ? cob . whi sir , cobbl you fla. thou art a cobbler , art t  
hou ? cob . truli sir , all that I liue by , is with the aul : I meddl wit  
h no tradesman matter , nor women matter ; but withal I am inde sir , a su  
rgeon to old shooe : when they are in great danger , I recouer them . As p  
roper men as euer trod vpon neat leather , haue gone vpon my handy-work fl  
a . but wherefor art not in thi shop to day ? whi do'st thou lead these me

## Using Lemmatizer

In [9]:

```
1 from nltk.stem import WordNetLemmatizer
2 lemmatizer = WordNetLemmatizer()
```

In [10]:

```
1 for word in words:
2     print("{0:20}{1:20}".format(word, lemmatizer.lemmatize(word)))
```

(	(
Being	Being
Mechanicall	Mechanicall
)	)
you	you
ought	ought
not	not
walke	walke
Vpon	Vpon
a	a
labouring	labouring
day	day
,	,
without	without
the	the
signe	signe
Of	Of
your	your
Profession	Profession
?	?

## Inference on using Lemmatizer on shakespeare work

Not that of significance on fictional role play work. Even the Stemmer (Both porter and lancaster) are not of significant use, as the work has lot of poetic sentence formation. Let us explore it in normal webtext.

## Using Stemming, Tokenizing on Wikipedia context of Julius Ceasar. (Normal English - No poetic touch)

In [11]:

```
1 text = ""Gaius Julius Caesar (12 July 100 BC - 15 March 44 BC), known simply as Julius
2 ""
```

In [12]:

```
1 sentence = sent_tokenize(text)
2 for sent in sentence:
3     print(sent+ "\n")
```

Gaius Julius Caesar (12 July 100 BC – 15 March 44 BC), known simply as Julius Caesar, was a Roman general and statesman who played a critical role in the events that led to the demise of the Roman Republic and the rise of the Roman Empire.

He was also a historian and author of Latin prose.

In 60 BC, Caesar, Crassus and Pompey formed the First Triumvirate, a political alliance that dominated Roman politics for several years.

Their attempts to amass power as Populares were opposed by the Optimates within the Roman Senate, among them Cato the Younger with the frequent support of Cicero.

Caesar rose to become one of the most powerful politicians in the Roman Republic through a number of his accomplishments, notably his victories in the Gallic Wars, completed by 51 BC.

During this time, Caesar became the first Roman general to cross both the English Channel and the Rhine River, when he built a bridge across the Rhine and crossed the Channel to invade Britain.

Caesar's wars extended Rome's territory to Britain and past Gaul.

These achievements granted him unmatched military power and threatened to eclipse the standing of Pompey, who had realigned himself with the Senate after the death of Crassus in 53 BC.

With the Gallic Wars concluded, the Senate ordered Caesar to step down from his military command and return to Rome.

Leaving his command in Gaul meant losing his immunity from being charged as a criminal for waging unsanctioned wars.

As a result, Caesar found himself with no other options but to cross the Rubicon with the 13th Legion in 49 BC, leaving his province and illegally entering Roman Italy under arms.

This began Caesar's civil war, and his victory in the war by 45 BC put him in an unrivaled position of power and influence.

In [13]:

```
1 words=word_tokenize(text)
2 print(words)
```

```
['Gaius', 'Julius', 'Caesar', '(', '12', 'July', '100', 'BC', '-', '15', 'Ma',
rch', '44', 'BC', ')', ',', ',', 'known', 'simply', 'as', 'Julius', 'Caesar',
',', ',', 'was', 'a', 'Roman', 'general', 'and', 'statesman', 'who', 'played',
'a', 'critical', 'role', 'in', 'the', 'events', 'that', 'led', 'to', 'the',
'demise', 'of', 'the', 'Roman', 'Republic', 'and', 'the', 'rise', 'of', 'th',
e', 'Roman', 'Empire', '.', 'He', 'was', 'also', 'a', 'historian', 'and', 'a
uthor', 'of', 'Latin', 'prose', '.', 'In', '60', 'BC', ',', ',', 'Caesar', ',',
'Crassus', 'and', 'Pompey', 'formed', 'the', 'First', 'Triumvirate', ',',
'a', 'political', 'alliance', 'that', 'dominated', 'Roman', 'politics', 'fo',
r', 'several', 'years', '.', 'Their', 'attempts', 'to', 'amass', 'power', 'a',
s', 'Populares', 'were', 'opposed', 'by', 'the', 'Optimates', 'within', 'th',
e', 'Roman', 'Senate', ',', ',', 'among', 'them', 'Cato', 'the', 'Younger', 'wit',
h', 'the', 'frequent', 'support', 'of', 'Cicero', '.', 'Caesar', 'rose', 't',
o', 'become', 'one', 'of', 'the', 'most', 'powerful', 'politicians', 'in',
'the', 'Roman', 'Republic', 'through', 'a', 'number', 'of', 'his', 'accompli',
shments', ',', ',', 'notably', 'his', 'victories', 'in', 'the', 'Gallic', 'Wars',
',', ',', 'completed', 'by', '51', 'BC', '.', 'During', 'this', 'time', ',', ',', 'Cae',
sar', 'became', 'the', 'first', 'Roman', 'general', 'to', 'cross', 'both',
'the', 'English', 'Channel', 'and', 'the', 'Rhine', 'River', ',', ',', 'when', 'h',
e', 'built', 'a', 'bridge', 'across', 'the', 'Rhine', 'and', 'crossed', 'th',
e', 'Channel', 'to', 'invade', 'Britain', '.', 'Caesar', '"s", 'wars', 'exte',
nded', 'Rome', '"s", 'territory', 'to', 'Britain', 'and', 'past', 'Gaul',
',', ',', 'These', 'achievements', 'granted', 'him', 'unmatched', 'military', 'po',
wer', 'and', 'threatened', 'to', 'eclipse', 'the', 'standing', 'of', 'Pompe',
y', ',', ',', 'who', 'had', 'realigned', 'himself', 'with', 'the', 'Senate', 'aft',
er', 'the', 'death', 'of', 'Crassus', 'in', '53', 'BC', '.', 'With', 'the',
'Gallic', 'Wars', 'concluded', ',', ',', 'the', 'Senate', 'ordered', 'Caesar', 't',
o', 'step', 'down', 'from', 'his', 'military', 'command', 'and', 'return',
'to', 'Rome', '.', 'Leaving', 'his', 'command', 'in', 'Gaul', 'meant', 'losi',
ng', 'his', 'immunity', 'from', 'being', 'charged', 'as', 'a', 'criminal',
'for', 'waging', 'unsanctioned', 'wars', '.', 'As', 'a', 'result', ',', ',', 'Cae',
sar', 'found', 'himself', 'with', 'no', 'other', 'options', 'but', 'to', 'cr',
oss', 'the', 'Rubicon', 'with', 'the', '13th', 'Legion', 'in', '49', 'BC',
',', ',', 'leaving', 'his', 'province', 'and', 'illegally', 'entering', 'Roman',
'Italy', 'under', 'arms', '.', 'This', 'began', 'Caesar', '"s", 'civil', 'wa',
r', ',', ',', 'and', 'his', 'victory', 'in', 'the', 'war', 'by', '45', 'BC', 'pu',
t', 'him', 'in', 'an', 'unrivaled', 'position', 'of', 'power', 'and', 'influ',
ence', '.']
```

In [14]:

```
1 tagged = pos_tag(words)
2 print(tagged)
```

```
[('Gaius', 'NNP'), ('Julius', 'NNP'), ('Caesar', 'NNP'), ('(', '('), ('12', 'CD'), ('July', 'NNP'), ('100', 'CD'), ('BC', 'NNP'), ('-', '$'), ('15', 'CD'), ('March', 'NNP'), ('44', 'CD'), ('BC', 'NNP'), (',', ','), (',', ','), ('known', 'VBN'), ('simply', 'RB'), ('as', 'IN'), ('Julius', 'NNP'), ('Caesar', 'NNP'), (',', ','), ('was', 'VBD'), ('a', 'DT'), ('Roman', 'NNP'), ('general', 'JJ'), ('and', 'CC'), ('statesman', 'NN'), ('who', 'WP'), ('played', 'VBD'), ('a', 'DT'), ('critical', 'JJ'), ('role', 'NN'), ('in', 'IN'), ('the', 'DT'), ('events', 'NNS'), ('that', 'WDT'), ('led', 'VBD'), ('to', 'TO'), ('the', 'DT'), ('demise', 'NN'), ('of', 'IN'), ('the', 'DT'), ('Roman', 'NNP'), ('Republic', 'NNP'), ('and', 'CC'), ('the', 'DT'), ('rise', 'NN'), ('of', 'IN'), ('the', 'DT'), ('Roman', 'NNP'), ('Empire', 'NNP'), ('.', '.'), ('He', 'PRP'), ('was', 'VBD'), ('also', 'RB'), ('a', 'DT'), ('historian', 'JJ'), ('and', 'CC'), ('author', 'NN'), ('of', 'IN'), ('Latin', 'NNP'), ('prose', 'NN'), ('.', '.'), ('In', 'IN'), ('60', 'CD'), ('BC', 'NNP'), (',', ','), ('Caesar', 'NNP'), (',', ','), ('Crassus', 'NNP'), ('and', 'CC'), ('Pompey', 'NNP'), ('formed', 'VBD'), ('the', 'DT'), ('First', 'NNP'), ('Triumvirate', 'NNP'), (',', ','), ('a', 'DT'), ('political', 'JJ'), ('alliance', 'NN'), ('that', 'WDT'), ('dominated', 'VBD'), ('Roman', 'NNP'), ('politics', 'NNS'), ('for', 'IN'), ('several', 'JJ'), ('years', 'NNS'), ('.', '.'), ('Their', 'PRP$'), ('attempts', 'NNS'), ('to', 'TO'), ('amass', 'VB'), ('power', 'NN'), ('as', 'IN'), ('Populares', 'NNS'), ('were', 'VBD'), ('opposed', 'VBN'), ('by', 'IN'), ('the', 'DT'), ('Optimates', 'NNP'), ('within', 'IN'), ('the', 'DT'), ('Roman', 'NNP'), ('Senate', 'NNP'), (',', ','), ('among', 'IN'), ('them', 'PRP'), ('Cato', 'NNP'), ('the', 'DT'), ('Younger', 'NNP'), ('with', 'IN'), ('the', 'DT'), ('frequent', 'JJ'), ('support', 'NN'), ('of', 'IN'), ('Cicero', 'NNP'), ('.', '.'), ('Caesar', 'NNP'), ('rose', 'VBD'), ('to', 'TO'), ('become', 'VB'), ('one', 'CD'), ('of', 'IN'), ('the', 'DT'), ('most', 'RBS'), ('powerful', 'JJ'), ('politicians', 'NNS'), ('in', 'IN'), ('the', 'DT'), ('Roman', 'NNP'), ('Republic', 'NNP'), ('through', 'IN'), ('a', 'DT'), ('number', 'NN'), ('of', 'IN'), ('his', 'PRP$'), ('accomplishments', 'NNS'), (',', ','), ('notably', 'RB'), ('his', 'PRP$'), ('victories', 'NNS'), ('in', 'IN'), ('the', 'DT'), ('Gallic', 'NNP'), ('Wars', 'NNP'), (',', ','), ('completed', 'VBN'), ('by', 'IN'), ('51', 'CD'), ('BC', 'NNP'), ('.', '.'), ('During', 'IN'), ('this', 'DT'), ('time', 'NN'), (',', ','), ('Caesar', 'NNP'), ('became', 'VBD'), ('the', 'DT'), ('first', 'JJ'), ('Roman', 'NNP'), ('general', 'NN'), ('to', 'TO'), ('cross', 'VB'), ('both', 'DT'), ('the', 'DT'), ('English', 'NNP'), ('Channel', 'NNP'), ('and', 'CC'), ('the', 'DT'), ('Rhine', 'NNP'), ('River', 'NNP'), (',', ','), ('when', 'WRB'), ('he', 'PRP'), ('built', 'VBD'), ('a', 'DT'), ('bridge', 'NN'), ('across', 'IN'), ('the', 'DT'), ('Rhine', 'NNP'), ('and', 'CC'), ('crossed', 'VBD'), ('the', 'DT'), ('Channel', 'NNP'), ('to', 'TO'), ('invade', 'VB'), ('Britain', 'NNP'), ('.', '.'), ('Caesar', 'NNP'), ('s', 'POS'), ('wars', 'NNS'), ('extended', 'VBD'), ('Rome', 'NNP'), ('s', 'POS'), ('territory', 'NN'), ('to', 'TO'), ('Britain', 'NNP'), ('and', 'CC'), ('past', 'JJ'), ('Gaul', 'NNP'), ('.', '.'), ('These', 'DT'), ('achievements', 'NNS'), ('granted', 'VBD'), ('him', 'PRP'), ('unmatched', 'JJ'), ('military', 'JJ'), ('power', 'NN'), ('and', 'CC'), ('threatened', 'VBD'), ('to', 'TO'), ('eclipse', 'VB'), ('the', 'DT'), ('standing', 'NN'), ('of', 'IN'), ('Pompey', 'NNP'), (',', ','), ('who', 'WP'), ('had', 'VBD'), ('realigned', 'VBN'), ('himself', 'PRP'), ('with', 'IN'), ('the', 'DT'), ('Senate', 'NNP'), ('after', 'IN'), ('the', 'DT'), ('death', 'NN'), ('of', 'IN'), ('Crassus', 'NNP'), ('in', 'IN'), ('53', 'CD'), ('BC', 'NNP'), ('.', '.'), ('With', 'IN'), ('the', 'DT'), ('Gallic', 'NNP'), ('Wars', 'NNP'), ('concluded', 'VBD'), (',', ','), ('the', 'DT'), ('Senate', 'NNP'), ('ordered', 'VBD'), ('Caesar', 'NNP'), ('to', 'TO'), ('step', 'VB'), ('down', 'RP'), ('from', 'IN'), ('his', 'PRP$'), ('military', 'JJ'), ('command', 'NN'), ('and', 'CC'), ('return', 'NN'), ('to',
```



('TO'), ('Rome', 'NNP'), ('.', '.'), ('Leaving', 'VBG'), ('his', 'PRP\$'), ('command', 'NN'), ('in', 'IN'), ('Gaul', 'NNP'), ('meant', 'NN'), ('losing', 'VBG'), ('his', 'PRP\$'), ('immunity', 'NN'), ('from', 'IN'), ('being', 'VBG'), ('charged', 'VBN'), ('as', 'IN'), ('a', 'DT'), ('criminal', 'NN'), ('for', 'IN'), ('waging', 'VBG'), ('unsanctioned', 'JJ'), ('wars', 'NNS'), ('.', '.'), ('As', 'IN'), ('a', 'DT'), ('result', 'NN'), ('.', '.'), ('Caesar', 'NNP'), ('found', 'VBD'), ('himself', 'PRP'), ('with', 'IN'), ('no', 'DT'), ('other', 'JJ'), ('options', 'NNS'), ('but', 'CC'), ('to', 'TO'), ('cross', 'VB'), ('the', 'DT'), ('Rubicon', 'NNP'), ('with', 'IN'), ('the', 'DT'), ('13th', 'CD'), ('Legion', 'NNP'), ('in', 'IN'), ('49', 'CD'), ('BC', 'NNP'), ('.', '.'), ('leaving', 'VBG'), ('his', 'PRP\$'), ('province', 'NN'), ('and', 'CC'), ('illegally', 'RB'), ('entering', 'VBG'), ('Roman', 'NNP'), ('Italy', 'NNP'), ('under', 'IN'), ('arms', 'NNS'), ('.', '.'), ('This', 'DT'), ('began', 'VBD'), ('Caesar', 'NNP'), ('s', 'POS'), ('civil', 'JJ'), ('war', 'NN'), ('.', '.'), ('and', 'CC'), ('his', 'PRP\$'), ('victory', 'NN'), ('in', 'IN'), ('the', 'DT'), ('war', 'NN'), ('by', 'IN'), ('45', 'CD'), ('BC', 'NNP'), ('put', 'VBD'), ('him', 'PRP'), ('in', 'IN'), ('an', 'DT'), ('unrivaled', 'JJ'), ('position', 'NN'), ('of', 'IN'), ('power', 'NN'), ('and', 'CC'), ('influence', 'NN'), ('.', '.'))]

In [15]:

```
1 porter = PorterStemmer()
2 lancaster=LancasterStemmer()
3 print("{0:20}{1:20}{2:20}".format("Word", "Porter Stemmer", "lancaster Stemmer"))
4 print("-" * 60)
5 for word in words:
6     print("{0:20}{1:20}{2:20}".format(word, porter.stem(word), lancaster.stem(word)))
```

Word	Porter Stemmer	lancaster Stemmer
Gaius	gaiu	gai
Julius	juliu	juli
Caesar	caesar	caes
(	(	(
12	12	12
July	juli	july
100	100	100
BC	BC	bc
-	-	-
15	15	15
March	march	march
44	44	44
BC	BC	bc
)	)	)
,	,	,
known	known	known
simply	simpli	simply

In [16]:

```
1 token_words=word_tokenize(text)
2 stem_sentence=[]
3 for word in token_words:
4     stem_sentence.append(porter.stem(word))
5     stem_sentence.append(" ")
6 #print(stem_sentence)
7 print( "".join(stem_sentence))
```

gaius julius caesar ( 12 juli 100 BC - 15 march 44 BC ) , known simply as julius caesar , was a roman general and statesman who played a critical role in the events that led to the demise of the roman republic and the rise of the roman empire . He was also a historian and author of latin prose . In 60 BC , caesar , crassus and pompey formed the first triumvirate , a political alliance that dominated roman politics for several years . their attempt to amass power as popular leaders was opposed by the optimates within the roman senate , among them cato the younger with the frequent support of cicero . caesar rose to become one of the most powerful politicians in the roman republic through a number of his accomplishments , notably his victories in the gallic wars , completed by 51 BC . during this time , caesar became the first roman general to cross both the english channel and the rhine river , when he built a bridge across the rhine and crossed the channel to invade britain . caesar's wars extended rome's territory to britain and beyond gaul . these achievements granted him unmatched military power and threatened to eclipse the standing of pompey , who had realigned himself with the senate after the death of crassus in 53 BC . with the gallic wars concluded , the senate ordered caesar to step down from his military command and return to rome . leaving his command in gaul meant losing his immunity from being charged as a criminal for waging unauthorized wars . As a result , caesar found himself with no other option but to cross the rubicon with the 13th legion in 49 BC , leaving his province and illegally entering roman italy under arms . this began caesar's civil war , and his victories in the war by 45 BC put him in an unrivaled position of power and influence .

**---End of Documentation---**