# Kumar Gaurav 20122065 , NLP lab

# Lab_5 NLP

## 7 and 8 , 16 oct 2021

### Qn 7. Write a program for Lemmatizing using WordNet

In [3]:

```
import nltk
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
lemmatizer.lemmatize('books')
import nltk
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
lemmatizer.lemmatize('books')
```

Out[3]:

```
'book'
```

### Qn 8. Write a program to differentiate stemming and lemmatization of words

In [1]:

```
import nltk
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
lemmatizer.lemmatize(' believes ')
```

Out[1]:

```
' believes '
```

In [5]:

```
import nltk
from nltk.stem import SnowballStemmer
spanishStemmer=SnowballStemmer("spanish", ignore_stopwords=True)
spanishStemmer.stem("Corriendo")
```

Out[5]:

```
'corr'
```

In [6]:

```
from nltk.stem import PorterStemmer
from nltk.stem import LancasterStemmer
```

In [7]:

```
porter = PorterStemmer()
lancaster= LancasterStemmer()
```

In [8]:

```
print("Porter Stemmer")
print(porter.stem("cats"))
print(porter.stem("trouble"))
print(porter.stem("troubling"))
print(porter.stem("troubled"))
print(porter.stem("believe"))
print("Lancaster Stemmer")
print(lancaster.stem("cats"))
print(lancaster.stem("trouble"))
print(lancaster.stem("troubling"))
print(lancaster.stem("troubled"))
print(lancaster.stem("believe"))
```

```
Porter Stemmer
cat
troubl
troubl
troubl
believ
Lancaster Stemmer
cat
troubl
troubl
troubl
believ
```

## The root words from different methods are also different in few cases

In [9]:

```
sentence="Christ University is the best place to learn about the new trending technologies"
porter.stem(sentence)
```

Out[9]:

```
'christ university is the best place to learn about the new trending technol
og'
```

In [10]:

```
from nltk.tokenize import sent_tokenize, word_tokenize
def stemSentence(sentence):
    token_words=word_tokenize(sentence)
    token_words
    stem_sentence=[]
    for word in token_words:
        stem_sentence.append(porter.stem(word))
        stem_sentence.append(" ")
    return "".join(stem_sentence)
x=stemSentence(sentence)
print(x)
```

```
christ univers is the best place to learn about the new trend technolog
```

## We are trying to find root words of a text. Reading the text to file

In [16]:

```
file = open("K://New_NLP_text1.txt")
file.read()
```

Out[16]:

```
'I would love to try or hear the sample audio your app can produce. I do not
want to purchase, because I\'ve purchased so many apps that say they do some
thing and do not deliver. \\n\\nCan you please add audio samples with text y
ou\'ve converted? I\'d love to see the end results.\\n\\nThanks!I would l\no
ve to tri or hear the sampl audio your app can produc . I do not want tpurch
as , becaus I \'ve purchas so mani app that say they do someth and do not de
liv . can you pleas add audio sampl with text you \'ve convert ? I \'d love
to see the end result . thank ! "`` I would love to tri or hear the sampl au
dio your app can produc . I do n ot want to purchas , becaus I \'ve purchas
so mani app that say they do someth and do not deliv . \\n\\ncan you pleas a
dd audio sampl with t ext you \'ve convert ? I \'d love to see the end resul
ts.\\n\\nthank ! I would l ove to tri or hear the sampl audio your app can p
roduc . I do not want to purcha , becau I \'ve purcha so mani app that say t
hey do someth and do not deli'
```

## Reading the text data line wise

In [17]:

```
file=open("K://New_NLP_text1.txt")
my_lines_list=file.readlines()
my_lines_list
```

Out[17]:

```
["I would love to try or hear the sample audio your app can produce. I do no
t want to purchase, because I've purchased so many apps that say they do som
ething and do not deliver. \\n\\nCan you please add audio samples with text
you've converted? I'd love to see the end results.\\n\\nThanks!I would l\n",
 'ove to tri or hear the sampl audio your app can produc . I do not want tpu
rchas , becaus I \'ve purchas so mani app that say they do someth and do not
deliv . can you pleas add audio sampl with text you \'ve convert ? I \'d lov
e to see the end result . thank ! "`` I would love to tri or hear the sampl
audio your app can produc . I do n ot want to purchas , becaus I \'ve purcha
s so mani app that say they do someth and do not deliv . \\n\\ncan you pleas
add audio sampl with t ext you \'ve convert ? I \'d love to see the end resu
lts.\\n\\nthank ! I would l ove to tri or hear the sampl audio your app can
produc . I do not want to purcha , becau I \'ve purcha so mani app that say
they do someth and do not deli']
```

## To display the original text data and the stemmed sentence

In [18]:

```
file=open("K://New_NLP_text1.txt")
my_lines_list=file.readlines()
my_lines_list
```

Out[18]:

```
["I would love to try or hear the sample audio your app can produce. I do no
t want to purchase, because I've purchased so many apps that say they do som
ething and do not deliver. \\n\\nCan you please add audio samples with text
you've converted? I'd love to see the end results.\\n\\nThanks!I would l\n",
 'ove to tri or hear the sampl audio your app can produc . I do not want tpu
rchas , becaus I \'ve purchas so mani app that say they do someth and do not
deliv . can you pleas add audio sampl with text you \'ve convert ? I \'d lov
e to see the end result . thank ! "`` I would love to tri or hear the sampl
audio your app can produc . I do n ot want to purchas , becaus I \'ve purcha
s so mani app that say they do someth and do not deliv . \\n\\ncan you pleas
add audio sampl with t ext you \'ve convert ? I \'d love to see the end resu
lts.\\n\\nthank ! I would l ove to tri or hear the sampl audio your app can
produc . I do not want to purcha , becau I \'ve purcha so mani app that say
they do someth and do not deli']
```

In [19]:

```
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.stem import PorterStemmer
def stemSentence(sentence):
    token_words=word_tokenize(sentence)
    token_words
    stem_sentence=[]
    for word in token_words:
        stem_sentence.append(porter.stem(word))
        stem_sentence.append(" ")
    return "".join(stem_sentence)
print(my_lines_list[0])
print("Stemmed Sentence")
x=stemSentence(my_lines_list[0])
print(x)
```

```
I would love to try or hear the sample audio your app can produce. I do not
want to purchase, because I've purchased so many apps that say they do somet
hing and do not deliver. \n\nCan you please add audio samples with text yo
u've converted? I'd love to see the end results.\n\nThanks!I would l

Stemmed Sentence
i would love to tri or hear the sampl audio your app can produc . i do not w
ant to purchas , becaus i 've purchas so mani app that say they do someth an
d do not deliv . \n\ncan you pleas add audio sampl with text you 've convert
? i 'd love to see the end results.\n\nthank ! i would l
```

In [20]:

```
stem_file=open("K://New_NLP_text1.txt",mode="a+",encoding="utf-8")
for line in my_lines_list:
 stem_sentence=stemSentence(line)
 stem_file.write(stem_sentence)
stem_file.close()
```

## The words are the root word using Porter Stemmer and Lancaster Stemmer for the input words

In [21]:

```
word_list = {"enemy", "fellowship", "friends", "friendshop","stabilise", "destablize",
"misunderstanding"}
print("{0:20}{1:20}{2:20}".format("word","Porter Stemmer", "Lancaster Stemmer"))
for word in word_list:
 print("{0:20}{1:20}{2:20}".format(word,porter.stem(word),lancaster.stem(word)))
```

```
word                Porter Stemmer      Lancaster Stemmer
stabilise           stabilis            stabl
fellowship          fellowship          fellow
misunderstanding    misunderstand       misunderstand
enemy               enemi               enemy
friends             friend              friend
friendshop          friendshop          friendshop
destablize          destabl             dest
```

In [ ]:

```
```