In [1]:
```
import string as st

def remove_pun (text):
    temp= []
    for word in text:
        if word not in st.punctuation:
            temp.append(''.join(word))
    return ''.join(temp)

remove_pun('my, name is srivatsa')
```

Out[1]:
```
'my name is srivatsa'
```

In [2]:
```
import string as st

def remove_pun (text):
    temp= ''
    for word in text:
        if word not in st.punctuation:
            temp = temp+word
    return temp

remove_pun('my, name is srivatsa')
```

Out[2]:
```
'my name is srivatsa'
```

In [20]:
```
import re

def tokenization(text):
    token = re.split('\W+',text)
    return token

nxt = tokenization('he is a good boy.him his they yourselves see cannot happy because'.lower())
nxt
```

Out[20]:
```
['he',
 'is',
 'a',
 'good',
 'boy',
 'him',
 'his',
 'they',
 'yourselves',
 'see',
 'cannot',
 'happy',
 'because']
```

In [15]:
```
import nltk
#nltk.download('stopwords')
stop_words = nltk.corpus.stopwords.words('english')
len(stop_words)
```

Out[15]:
```
179
```

In [17]:
```
def remove_stop_words(text):
    temp = []
    for word in text:
        if word not in stop_words:
            temp.append(word)
    return temp

remove_stop_words(nxt)
```

Out[17]:
```
['good', 'boy', 'see', 'cannot', 'happy']
```

In [29]:
```
ps = nltk.PorterStemmer()
ps.stem("good")
```

Out[29]:
```
'good'
```

In [ ]:
```

```