```
import nltk
```

```
nltk.download('stopwords')
nltk.download('words')
nltk.download('wordnet')
nltk.download('averged_perception_tagger')
nltk.download('punkt')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package words to /root/nltk_data...
[nltk_data]   Unzipping corpora/words.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data] Error loading averged_perception_tagger: Package
[nltk_data]     'averged_perception_tagger' not found in index
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
True
```

```
import pandas as pd
import numpy as np
sent="they told that thier eges are 20 23 and 27 respectively"
```

```
add=[]
```

```
for word in sent.split():
  if word.isdigit():
    add.append(int(word))
```

```
print(add)
```

```
[20, 23, 27]
```

```
print("avg:",sum(add)/len(add))
```

```
avg: 23.333333333333332
```

```
from nltk.tokenize import word_tokenize,sent_tokenize
sent="hello all! how are you? welcome to pun"
sent_tokenize(sent)
```

```
['hello all!', 'how are you?', 'welcome to pun']
```

```
word_tokenize(sent)
```

```
['hello', 'all', '!', 'how', 'are', 'you', '?', 'welcome', 'to', 'pun']
```

```
from nltk.tokenize import SpaceTokenizer
tk=SpaceTokenizer()
tk.tokenize(sent)
```

```
['hello', 'all!', 'how', 'are', 'you?', 'welcome', 'to', 'pun']
```

```
#lemmatization
word='playing'
from nltk.stem import WordNetLemmatizer


wn1=WordNetLemmatizer()
print(wn1.lemmatize(word,'n'))
print(wn1.lemmatize(word,'v'))
print(wn1.lemmatize(word,'a'))
print(wn1.lemmatize(word,'r'))
```

```
playing
play
playing
playing
```

```
#pos tagging

from nltk import pos_tag



import nltk
nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
True
```

```
sents='Rajgad (literal meaning ruling fort) is a hill fort situated in the pune district
words=word_tokenize(sents)


nltk.download('omw-1.4')
```

```
[nltk_data] Downloading package omw-1.4 to /root/nltk_data...
True
```

```
pos_tag(words)
```

```
[('Rajgad', 'NNP'),
 ('(', '('),
 ('literal', 'JJ'),
 ('meaning', 'NN'),
 ('ruling', 'VBG'),
 ('fort', 'NN'),
 (')', ')'),
```

```
       ('is', 'VBZ'),
       ('a', 'DT'),
       ('hill', 'NN'),
       ('fort', 'NN'),
       ('situated', 'VBN'),
       ('in', 'IN'),
       ('the', 'DT'),
       ('pune', 'JJ'),
       ('district', 'NN'),
       ('of', 'IN'),
       ('maharastra', 'NN')]


tags=pos_tag(words)



for word in tags:
  if word[1].startswith('V'):
    print(word[0])

      ruling
      is
      situated


from textblob import TextBlob

t=TextBlob('computoor')
print(t.correct())

      computer
```