```
#NLP
import nltk
import nltk.corpus



#tokenizaton - split sentence - then only bot will understand
from nltk.tokenize import word_tokenize
from nltk.tokenize import sent_tokenize


a="Ben 10 is a cartoon network"
b="Ben's girlfriend name is July.His cousin name s Gwen"




nltk.download('punkt')
print(word_tokenize(a))
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
['Ben', '10', 'is', 'a', 'cartoon', 'network']
```

```
print(sent_tokenize(b))
```

```
["Ben's girlfriend name is July.His cousin name s Gwen"]
```

```
print(len(word_tokenize(a)))
print(len(word_tokenize(b)))
```

```
6
10
```

```
#bi-grams and n-grams
c="Ben is a plmber"
c_token=(word_tokenize(c))

list(nltk.bigrams(c_token))
```

```
[('Ben', 'is'), ('is', 'a'), ('a', 'plmber')]
```

```
list(nltk.trigrams(c_token))
```

```
[('Ben', 'is', 'a'), ('is', 'a', 'plmber')]
```

```
list(nltk.ngrams(c_token,1))
```

```
[('Ben',), ('is',), ('a',), ('plmber',)]
```

```
list(nltk.ngrams(c_token,4))
```

```
[('Ben', 'is', 'a', 'plmber')]
```

```
#Stemming - root word


from nltk.stem import PorterStemmer
d=PorterStemmer()


print(d.stem("kissing"))
print(d.stem("dating"))
```

```
kiss
date
```

```
#pas-tagging
e="Ben is my favorite alien is Feedback"
```

```
f=word_tokenize(e)
```

```
nltk.download('averaged_perceptron_tagger')
nltk.pos_tag(f)
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
[('Ben', 'NNP'),
 ('favorite', 'JJ'),
 ('alien', 'NN'),
 ('is', 'VBZ'),
 ('Feedback', 'NNP')]
```

```
# name- entity recogniton
from nltk import ne_chunk
```

```
g="loki is creating his own cinematic universe"
h=word_tokenize(g)
```

```
i=nltk.pos_tag(h)
print(i)
```

```
[('loki', 'NN'), ('is', 'VBZ'), ('creating', 'VBG'), ('his', 'PRP$'), ('own', 'JJ'), ('cinematic', 'JJ'), ('universe', 'NN')]
```

```
nltk.download('maxent_ne_chunker')
nltk.download('words')
```

```
print(ne_chunk(i))
```

```
[nltk_data] Downloading package maxent_ne_chunker to
[nltk_data]     /root/nltk_data...
[nltk_data]   Package maxent_ne_chunker is already up-to-date!
[nltk_data] Downloading package words to /root/nltk_data...
(S
  loki/NN
  is/VBZ
  creating/VBG
  his/PRP$
  own/JJ
  cinematic/JJ
  universe/NN)
[nltk_data]   Unzipping corpora/words.zip.
```

```
#convert text to speech
!pip install gTTs
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting gTTs
  Downloading gTTS-2.2.4-py3-none-any.whl (26 kB)
Requirement already satisfied: six in /usr/local/lib/python3.7/dist-packages (from gTTs) (1.15.0)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from gTTs) (7.1.2)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from gTTs) (2.23.0)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->gTTs) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->gTTs) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->gTTs) (2022.6.15)
Installing collected packages: gTTs
Successfully installed gTTs-2.2.4
```

```
#making text into voice
from gtts import gTTS
from IPython.display import Audio
```

```
j = gTTS("Sundar Pitchai is the ceo of Google")
j.save('1.wav')
sound_file='1.wav'
Audio(sound_file,autoplay=True)
```