

#Name:-Vaibhav Shete

#Batch:-B3

Import spacy

Raw\_text="""The Board of Control for Cricket in India (BCCI) is the governing body for cricket in India and is under the jurisdiction of Ministry of Youth Affairs

And Sports, Government of India.[2] The board was formed in December 1928 as a society, registered under the Tamil Nadu Societies Registration Act.

It is a consortium of state cricket associations and the state associations select their representatives who in turn elect the BCCI Chief.

Its headquarters are in Wankhede Stadium, Mumbai. Grant Govan was its first president and Anthony De Mello its first secretary. """

NER = spacy.load("en\_core\_web\_sm", disable=["tok2vec", "tagger", "parser", "attribute\_ruler", "lemmatizer"])

Text= NER(raw\_text)

For w in text.ents:

Print(w.text,w.label\_)

The Board of Control for Cricket ORG

India GPE

BCCI ORG

India GPE

Ministry of Youth Affairs

ORG

December 1928 DATE

The Tamil Nadu Societies Registration Act ORG

BCCI ORG

Wankhede Stadium FAC

Mumbai GPE

Grant Govan PERSON

First ORDINAL

Anthony De Mello PERSON

First ORDINAL

```
Spacy.displacy.render(text, style="ent",jupyter=True)
```

The Board of Control for Cricket ORG in India GPE ( BCCI ORG ) is the governing body for cricket in India GPE and is under the jurisdiction of Ministry of Youth Affairs ORG and Sports, Government of India.[2] The board was formed in December 1928 DATE as a society, registered under the Tamil Nadu Societies Registration Act ORG .

It Is a consortium of state cricket associations and the state associations select their representatives who in turn elect the BCCI ORG Chief.

Its headquarters are in Wankhede Stadium FAC , Mumbai GPE . Grant Govan PERSON was its first ORDINAL president and Anthony De Mello PERSON its first ORDINAL secretary.

```
Spacy.explain(u"NORP")
```

‘Nationalities or religious or political groups’

Import nltk

```
From nltk.tokenize import word_tokenize
```

```
From nltk.tag import pos_tag
```

```
Nltk.download('treebank')
```

```
Nltk.download('words')
```

```
Sent = nltk.corpus.treebank.tagged_sents()
```

```
Print(nltk.ne_chunk(sent[0]))
```

[nltk\_data] Downloading package treebank to

[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...

[nltk\_data] Package treebank is already up-to-date!

[nltk\_data] Downloading package words to

[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...

(S

(PERSON Pierre/NNP)

(ORGANIZATION Vinken/NNP)

,/,

61/CD

Years/NNS

Old/JJ

./,

Will/MD

Join/VB

The/DT

Board/NN

As/IN

a/DT

nonexecutive/JJ

director/NN

Nov./NNP

29/CD

./.)

[nltk\_data] Unzipping corpora\words.zip.

Nltk.download('punkt')

Nltk.download('averaged\_perceptron\_tagger')

Raw\_words= word\_tokenize(raw\_text)

Tags=pos\_tag(raw\_words)

[nltk\_data] Downloading package punkt to

[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...

[nltk\_data] Unzipping tokenizers\punkt.zip.

[nltk\_data] Downloading package averaged\_perceptron\_tagger to

[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...

[nltk\_data] Unzipping taggers\averaged\_perceptron\_tagger.zip.

Nltk.download('maxent\_ne\_chunker')

Nltk.download('words')

Ne = nltk.ne\_chunk(tags,binary=True)

Print(ne)

[nltk\_data] Downloading package maxent\_ne\_chunker to  
[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...  
[nltk\_data] Package maxent\_ne\_chunker is already up-to-date!  
[nltk\_data] Downloading package words to  
[nltk\_data] C:\Users\Lenovo\AppData\Roaming\nltk\_data...  
[nltk\_data] Package words is already up-to-date!

(S

The/DT

(NE Board/NNP)

Of/IN

(NE Control/NNP)

For/IN

Cricket/NNP

In/IN

(NE India/NNP)

/(

(NE BCCI/NNP)

)/)

Is/VBZ

The/DT

Governing/VBG

Body/NN

For/IN

Cricket/NN

In/IN

(NE India/NNP)

And/CC

Is/VBZ

Under/IN

The/DT

Jurisdiction/NN

Of/IN

(NE Ministry/NNP)

Of/IN

(NE Youth/NNP Affairs/NNPS)

And/CC

(NE Sports/NNP)

,/,

Government/NNP

Of/IN

(NE India/NNP)

./.

[/CC

2/CD

]/VBP

The/DT

Board/NN

Was/VBD

Formed/VBN

In/IN

December/NNP

1928/CD

As/IN

a/DT

society/NN

,/,

Registered/VBN

Under/IN

The/DT

(NE Tamil/NNP Nadu/NNP Societies/NNP)

Registration/NNP

Act/NNP

./.

It/PRP

Is/VBZ

a/DT

consortium/NN

of/IN

state/NN

cricket/NN

associations/NNS

and/CC

the/DT

state/NN

associations/NNS

select/VBP

their/PRP\$

representatives/NNS

who/WP

in/IN

turn/NN

elect/VBP

the/DT

(NE BCCI/NNP Chief/NNP)

./.

Its/PRP\$

Headquarters/NNS

Are/VBP

In/IN

(NE Wankhede/NNP Stadium/NNP)

,/,

(NE Mumbai/NNP)

./.

Grant/NNP

Govan/NNP

Was/VBD

Its/PRP\$

First/JJ

President/NN

And/CC

(NE Anthony/NNP De/NNP Mello/NNP)

Its/PRP\$

First/JJ

Secretary/NN

./.)