Experiment No 07: Named Entity Recognition

Rebecca Dias

Roll no: 18 BE CMPN A2

Pid: 182027

Aim : To implement the NER with NLTK toolkit.

```
# Step One: Import nltk and download necessary packages
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
nltk.download('maxent_ne_chunker')
nltk.download('words')
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Package averaged_perceptron_tagger is already up-to-
[nltk_data]       date!
[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...
[nltk_data]   Package words is already up-to-date!
True
```

```
# Step Two: Load Data
sentence="""
 Men like Schiaparelli watched the red planet—it is odd, by-the-bye, that
 for countless centuries Mars has been the star of war—but failed to
 interpret the fluctuating appearances of the markings they mapped so well.
 All that time the Martians must have been getting ready.

 During the opposition of 1894 a great light was seen on the illuminated
 part of the disk, first at the Lick Observatory, then by Perrotin of Nice,
 and then by other observers. English readers heard of it first in the
 issue of Nature dated August 2."""
```

```
# Step Three: Tokenise, find parts of speech and chunk words

for sent in nltk.sent_tokenize(sentence):
  for chunk in nltk.ne_chunk(nltk.pos_tag(nltk.word_tokenize(sent))):
    #POS tagging using nltk: Identifies parts of speech of each word and returns an array
    #Chunking on POS: Lastly, we perform a chunking operation that returns a nested nltk.t
     if hasattr(chunk, 'label'):
        print(chunk.label(), ' '.join(c[0] for c in chunk))
```

```
PERSON Schiaparelli
```

```
PERSON Mars
ORGANIZATION Lick Observatory
PERSON Perrotin
GPE Nice
GPE English
PERSON Nature
```

Conclusion:

We implemented the NER with NLTK Toolkit , the desired output was obtained.