**EXPERIMENT NO.8**

**AIM:**

Named Entity Recognition

**RESOURCES REQUIRED:**

Python 3, NLTK toolkit, Text editor, 4 GB RAM and above, i5 processor and above

**THEORY:**

**Named Entity Recognition:**

Named-entity recognition (NER) (also known as entity identification, entity chunking and entity extraction) is a sub-task of information extraction that seeks to locate and classify named entities in text into pre-defined categories such as the names of persons, organizations, locations, expressions of times, quantities, monetary values, percentages, etc.

NER systems have been created that use linguistic grammar-based techniques as well as statistical models such as machine learning. Hand-crafted grammar-based systems typically obtain better precision, but at the cost of lower recall and months of work by experienced computational linguists . Statistical NER systems typically require a large amount of manually annotated training data. Semi-supervised approaches have been suggested to avoid part of the annotation effort.

Named Entity Recognition has a wide range of applications in the field of Natural Language Processing and Information Retrieval. Few such examples have been listed below :

Automatically Summarizing Resumes

Optimizing Search Engine Algorithms

Powering Recommender Systems

Now that we explained NLP, we can describe how Named Entity Recognition works. NER plays a major role in the semantic part of NLP, which, extracts the meaning of words, sentences and their relationships. Basic NER processes structured and unstructured texts by identifying and locating entities. For example, instead of identifying “Steve” and “Jobs” as different entities, NER understands that “Steve Jobs” is a single entity. More developed NER processes can classify identified entities as well. In this case, NER not only identifies but classifies “Steve Jobs” as a person. In the following, we will describe the two most popular NER methods.

**CONCLUSION:**

Named Entity Recognition is a technique in natural language processing used to extract real entity names from word corpora. NER can be used to extract names, location, places etc. NER has been carefully studied and implemented.

**CODE:**

from spacy import displacy

from collections import Counter

from nltk.corpus import gutenberg

from nltk import pos\_tag

from pprint import pprint

import spacy

import nltk

import essential\_generators

enr = spacy.load("en\_core\_web\_sm")

text = essential\_generators.DocumentGenerator().paragraph()

text

doc = enr(text)

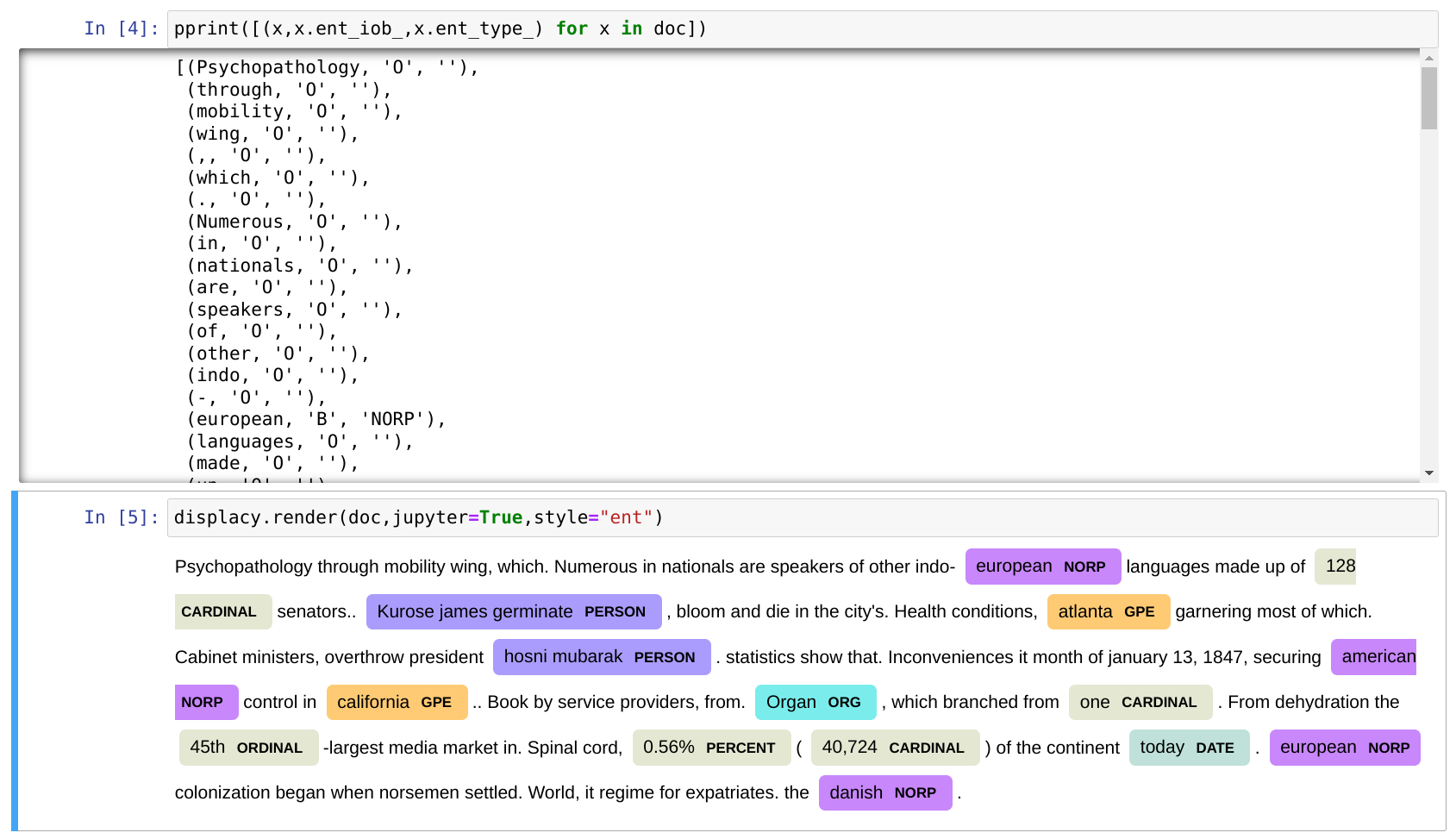
pprint([(x.text,x.label\_) for x in doc.ents])

pprint([(x,x.ent\_iob\_,x.ent\_type\_) for x in doc])

displacy.render(doc,jupyter=True,style="ent")

**OUTPUT:**

****

****