Experiment No.7					
Implement Named Entity	Recognizer	for	the	given	Text
input					
Date of Performance:					
Date of Submission:					

Vidyavardhini's College of Engineering and Technology

Department of Artificial Intelligence & Data Science

Aim: Implement Named Entity Recognizer for the given Text input

Objective: Understand the importance of NER in NLP and Implement NER.

Theory:

The named entity recognition (NER) is one of the most data preprocessing task. It involves

the identification of key information in the text and classification into a set of predefined

categories. An entity is basically the thing that is consistently talked about or refer to in the

text.

NER is the form of NLP.

At its core, NLP is just a two-step process, below are the two steps that are involved:

• Detecting the entities from the text

• Classifying them into different categories

Some of the categories that are the most important architecture in NER such that:

Person

Organization

Place/ location

Other common tasks include classifying of the following:

• date/time.

expression

• Numeral measurement (money, percent, weight, etc)

• E-mail address

Ambiguity in NE

For a person, the category definition is intuitively quite clear, but for computers, there is

some ambiguity in classification. Let's look at some ambiguous example:

England (Organisation) won the 2019 world cup vs The 2019 world cup happened in

England(Location).

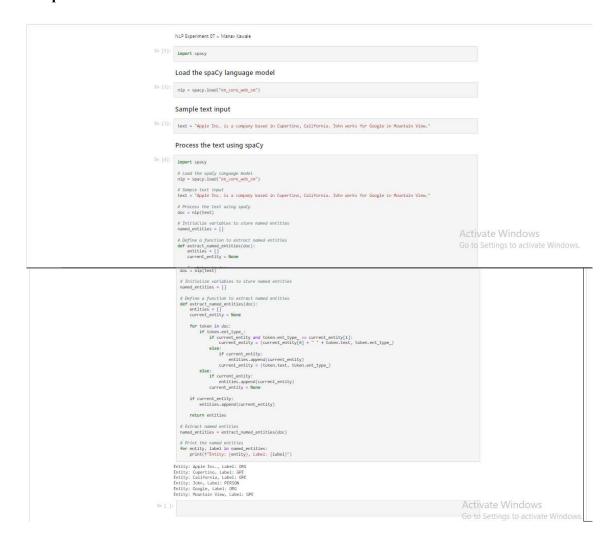
CSDL7013: Natural Language Processing Lab



Vidyavardhini's College of Engineering and Technology Department of Artificial Intelligence & Data Science

Washington(Location) is the capital of the US vs The first president of the US was Washington(Person).

Output:



Conclusion:

The results of the spaCy named entity recognition (NER) on the sample text input are correct. spaCy correctly identified all of the named entities in the text, including the following: Apple Inc. (organization) Cupertino (city) California (state) John (person) Google (organization) Mountain View (city) There are no words in the text that ought to be recognized as named entities but weren't.