

Experiment No. 6

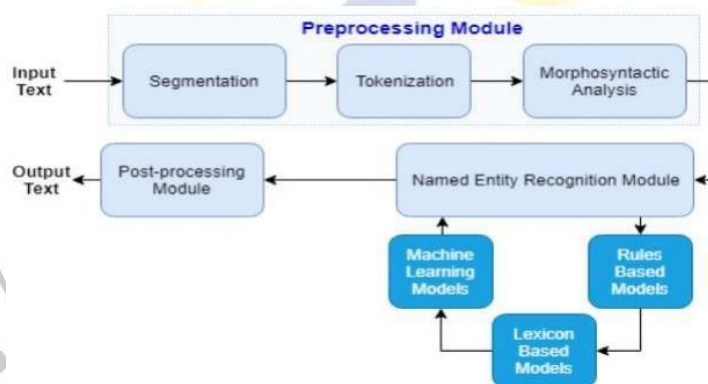
Named Entity Recognition

Aim: To study Named Entity Recognition

Theory: Named Entity Recognition (NER) is a standard NLP problem, which involves spotting named entities (people, places, organizations etc.) from a chunk of text, and classifying them into a predefined set of categories. Some of the practical applications of NER include:

1. Scanning news articles for the people, organizations and locations reported.
2. Providing concise features for search optimization: instead of searching the entire content, one may simply search for the major entities involved.
3. Quickly retrieving geographical locations talked about in Twitter posts.

In any text document, there are particular terms that represent specific entities that are more informative and have a unique context. These entities are known as named entities, which more specifically refer to terms that represent real-world objects like people, places, organizations, and so on, which are often denoted by proper names. A naive approach could be to find these by looking at the noun phrases in text documents. Named entity recognition (NER), also known as entity chunking/extraction, is a popular technique used in information extraction to identify and segment the named entities and classify or categorize them under various predefined classes.



How NER works

At the heart of any NER model is a two step process: -

1. Detect a named entity
2. Categorize the entity

Beneath this lie a couple of things. Step one involves detecting a word or string of words that form an entity. Each word represents a token: "The Great Lakes" is a string of three tokens that

represents one entity. Inside-outside beginning tagging is a common way of indicating where entities begin and end. We'll explore this further in a future blog post. The second step requires the creation of entity categories.

How is NER used?

NER is suited to any situation in which a high-level overview of a large quantity of text is helpful. With NER, you can, at a glance, understand the subject or theme of a body of text and quickly group texts based on their relevancy or similarity. Some notable NER use cases include:

Human resource: Speed up the hiring process by summarizing applicants' CVs; improve internal workflows by categorizing employee complaints and questions.

Customer support: Improve response times by categorizing user requests, complaints and questions and filtering by priority keywords.

Code:

Named Entity Recognition

```
locs = [('Omnicom', 'IN', 'New York'),
('DDB Needham', 'IN', 'New York'), ('Kaplan Thaler
Group', 'IN', 'New York'), ('BBDO South', 'IN',
'Atlanta'),
('Georgia-Pacific', 'IN', 'Atlanta')]
query = [e1 for (e1, rel, e2) in locs if e2=='Atlanta'] print(query)
```

Output:

['BBDO South', 'Georgia-Pacific']

Conclusion:

Thus, in the above experiment we have studied regarding named entity recognition, working of named entity recognition, how named entity recognition can be used and then implemented the code for the same and successfully executed it.

For Faculty Use

Correction Parameters	Formative Assessment [40%]	Timely completion of Practical [40%]	Attendance / Learning Attitude [20%]	
Marks Obtained				