

PAI LAB

BS in Artificial Intelligence



Department of Software Engineering
Faculty of Computer Science & Information Technology
The Superior University, Lahore

Sr.#	Reg.#	Student Name
1.	SU92-BSAIM-F23-107	Ayesha Faisal

Table of Contents

1	INTRODUCTION:	3
2	CODE IMPLEMENTATION:	4
3	CONCLUSION:	6

TABLE OF FIGURES

Figure a. IMPORT LIBRARIES	4
Figure b. APPLYING NER	5
Figure c. CREATING DF FOR ENTITIES	6

Named Entity Recognition (NER)

1 INTRODUCTION:

This project demonstrates a **basic but complete NER pipeline using SpaCy in Python**. The implementation covers:

- **Installing and loading** required models.
- **Processing real-world text** for named entity recognition.
- **Displaying** the entities.
- **Creating a structured dataset** from the recognized entities.

2 CODE IMPLEMENTATION:

```
NAMED ENTITY RECOGNITION USING NLP:

!pip install spacy

!pip install nltk

!python -m spacy download en_core_web_sm

!pip install pandas

pip install --upgrade pip
```

Figure a. IMPORT LIBRARIES

```
[10] !pip install BeautifulSoup4

import pandas as pd
import spacy
import requests

[14] from bs4 import BeautifulSoup

> ✓ [19] nlp = spacy.load("en_core_web_sm")
pd.set_option("display.max_rows", 200)

[20] content = '''Trinamool Congress leader Mahua Moitra has moved the Supreme Court
against her expulsion from the Lok Sabha over the cash-for-query allegations
against her. Moitra was ousted from the Parliament last week after the Ethics
Committee of the Lok Sabha found her guilty of jeopardising national security by
sharing her parliamentary portal's login credentials with businessman Darshan Hiranandani.'''
doc = nlp(content)
for ent in doc.ents:
    print(ent.text, ent.start_char, ent.end_char, ent.label_)
```

Figure b. APPLYING NER

```
for ent in doc.ents:
    print(ent.text, ent.start_char, ent.end_char, ent.label_)
```

```
Trinamool Congress 0 18 ORG
Mahua Moitra 26 38 PERSON
the Supreme Court 49 66 ORG
Moitra 157 163 NORP
Parliament 184 194 ORG
last week 195 204 DATE
the Ethics
Committee 211 231 ORG
the Lok Sabha 235 248 ORG
Darshan Hiranandani 373 392 PERSON
```

```
entities = [(ent.text, ent.label_, ent.lemma_) for ent in doc.ents]
df = pd.DataFrame(entities, columns=['text', 'type', 'lemma'])
print(df)
```

	text	type	lemma
0	Trinamool Congress	ORG	Trinamool Congress
1	Mahua Moitra	PERSON	Mahua Moitra
2	the Supreme Court	ORG	the Supreme Court
3	Moitra	NORP	Moitra
4	Parliament	ORG	Parliament
5	last week	DATE	last week
6	the Ethics\nCommittee	ORG	the Ethics\nCommittee
7	the Lok Sabha	ORG	the Lok Sabha

Figure c. CREATING DF FOR ENTITIES

3 CONCLUSION:

This code serves as an effective introduction to **NER using SpaCy**, showing how unstructured text can be converted into structured, labeled information. It can be extended further for real-world tasks like **news analysis, chatbot understanding, or information extraction from documents**