



Name : Yash Sarang	Class/Roll No. : D16AD / 47	Grade :
---------------------------	------------------------------------	----------------

Title of Experiment :

Part-of-Speech Tagging in Natural Language Processing (NLP)

Problem Statement :

The problem is to explore and understand part-of-speech (POS) tagging in natural language processing (NLP) and its significance in various NLP applications. The objective is to develop an efficient POS tagging algorithm and apply it to analyze text.

Description / Theory :

Description/Theory: Part-of-speech tagging involves assigning a grammatical part of speech (e.g., noun, verb, adjective, etc.) to each word in a sentence. This process helps in understanding the syntactic and semantic structure of the text, which is crucial for various NLP tasks such as named entity recognition, sentiment analysis, and machine translation. POS tagging can be performed using rule-based approaches, statistical methods, or machine learning algorithms.

Flowchart :

1. Input a text sentence.
2. Preprocess the text (tokenization, lowercasing, etc.).
3. Implement a POS tagging algorithm (rule-based, statistical, or ML-based).
4. Apply the algorithm to tag each word with its respective part of speech.
5. Analyze and validate the POS tags.
6. Use the POS-tagged data for NLP applications.

Program:

```
import spacy

# Load the pre-trained English POS tagging model
nlp = spacy.load("en_core_web_sm")

def pos_tagging(text):
```



Vivekanand Education Society's Institute Of Technology

Approved by AICTE & Affiliation to University of Mumbai

Artificial Intelligence and Data Science Department Natural Language Processing / Odd Sem 2023-24 / Experiment 6

```
# Process the input text using the spaCy NLP pipeline
doc = nlp(text)

# Iterate over tokens in the processed document and print POS tags
for token in doc:
    print(f"Token: {token.text}, POS Tag: {token.pos_}")

# Sample text for POS tagging
sample_text = "Part-of-speech tagging is an essential task in natural
language processing."

# Perform POS tagging on the sample text
print("POS Tagging results:")
pos_tagging(sample_text)
```

Output:

POS Tagging results:

Token: Part, POS Tag: NOUN
Token: -, POS Tag: PUNCT
Token: of, POS Tag: ADP
Token: speech, POS Tag: NOUN
Token: tagging, POS Tag: NOUN
Token: is, POS Tag: AUX
Token: an, POS Tag: DET
Token: essential, POS Tag: ADJ
Token: task, POS Tag: NOUN
Token: in, POS Tag: ADP
Token: natural, POS Tag: ADJ
Token: language, POS Tag: NOUN
Token: processing, POS Tag: NOUN
Token: ., POS Tag: PUNCT

Results and Discussions :

The results will include the POS tags assigned to each word in the input text, highlighting the accuracy and correctness of the tagging process. The discussion will



Artificial Intelligence and Data Science Department Natural Language Processing / Odd Sem 2023-24 / Experiment 6

focus on the strengths and weaknesses of the chosen POS tagging approach and its impact on downstream NLP tasks.

Conclusion:

POS tagging is a fundamental step in NLP, providing essential grammatical information about words in a sentence. The accurate identification of parts of speech allows for a deeper understanding of language structure and context. Various POS tagging algorithms and approaches can be tailored based on the specific application, and their accurate implementation is crucial for enhancing the accuracy and efficiency of NLP applications.
