

NLP assignment-2.4

Roll Number: [2403A52009]

Batch Number: [02]

Python Implementation

```
1 import nltk
2 from nltk.stem import PorterStemmer
3 from nltk.stem import WordNetLemmatizer
4
5 # Ensure required resources are downloaded
6 nltk.download('punkt')
7 nltk.download('punkt_tab')
8 nltk.download('wordnet')
9
10 medical_text = """
11 Diabetes is a chronic disease that affects how the body processes blood sugar.
12 If untreated, diabetes may cause heart disease, kidney failure, nerve damage and
13 vision problems.
14 Early diagnosis and proper treatment help improve patient outcomes.
15 """
16
17 # Tokenization
18 tokens = nltk.word_tokenize(medical_text)
19
20 # Apply stemming
21 stemmer = PorterStemmer()
22 stemmed_tokens = [stemmer.stem(token) for token in tokens]
23
24 # Apply lemmatization
25 lemmatizer = WordNetLemmatizer()
26 lemmatized_tokens = [lemmatizer.lemmatize(token) for token in tokens]
27
28 # Output results
29 print("Stemmed Tokens:", stemmed_tokens)
30 print("Lemmatized Tokens:", lemmatized_tokens)
```

Output

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data] Package punkt_tab is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
```

```
Stemmed Tokens: ['diabet', 'is', 'a', 'chronic', 'diseas', 'that', 'affect', 'how', 'the',
, 'bodi', 'process', 'blood', 'sugar', '.', 'if', 'untreat', ',', 'diabet', 'may', '
caus', 'heart', 'diseas', ',', 'kidney', 'failur', ',', 'nerv', 'damag', 'and', '
vision', 'problem', '.', 'earli', 'diagnosi', 'and', 'proper', 'treatment', 'help', '
improv', 'patient', 'outcom', '.']
```

```
Lemmatized Tokens: ['Diabetes', 'is', 'a', 'chronic', 'disease', 'that', 'affect', 'how',  
    'the', 'body', 'process', 'blood', 'sugar', '.', 'If', 'untreated', ',', 'diabetes',  
    'may', 'cause', 'heart', 'disease', ',', 'kidney', 'failure', ',', 'nerve', 'damage',  
    , 'and', 'vision', 'problem', '.', 'Early', 'diagnosis', 'and', 'proper', 'treatment',  
    , 'help', 'improve', 'patient', 'outcome', '.']
```