

PROGRAM [13]:

```
import nltk

nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')

from nltk.corpus import stopwords

from nltk.tokenize import word_tokenize

from nltk.stem import WordNetLemmatizer

# Define sample text

text = "Natural Language Processing (NLP) is a subfield of linguistics, computer science, and
artificial intelligence concerned with the interactions between computers and human (natural)
languages."

# Tokenize text

tokens = word_tokenize(text)

# Remove stop words

stop_words = set(stopwords.words('english'))

filtered_tokens = [word for word in tokens if word.lower() not in stop_words]

# Lemmatize tokens

lemmatizer = WordNetLemmatizer()

lemmatized_tokens = [lemmatizer.lemmatize(token) for token in filtered_tokens]

# Print lemmatized tokens

print(lemmatized_tokens)
```

OUTPUT [13]:

[nltk_data] Downloading package punkt to /root/nltk_data...

[nltk_data] Unzipping tokenizers/punkt.zip.

[nltk_data] Downloading package stopwords to /root/nltk_data...

[nltk_data] Unzipping corpora/stopwords.zip.

[nltk_data] Downloading package wordnet to /root/nltk_data...

['Natural', 'Language', 'Processing', '(', 'NLP', ')', 'subfield', 'linguistics', ',', 'computer', 'science', ',', 'artificial', 'intelligence', 'concerned', 'interaction', 'computer', 'human', '(', 'natural', ')', 'language', '.']