

```
In [1]: import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer
```

```
In [2]: nltk.download('punkt')
nltk.download('stopwords')
```

```
[nltk_data] Downloading package punkt to
[nltk_data]   C:\Users\dell\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]   C:\Users\dell\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

```
Out[2]: True
```

```
In [9]: def preprocess_text(text):
        ps = PorterStemmer()

        words = word_tokenize(text)

        stop_words = set(stopwords.words('english'))
        filtered_words = [word for word in words if word.lower() not in stop_words]

        stemmed_words = [ps.stem(word) for word in filtered_words]

        return ' '.join(stemmed_words)
```

```
In [10]: sample_text = """
Natural language processing (NLP) is a subfield of artificial intelligence (AI)
"""
```

```
In [11]: processed_text = preprocess_text(sample_text)
```

```
In [12]: print("Original Text:\n", sample_text)
print("\nProcessed Text:\n", processed_text)
```

Original Text:

Natural language processing (NLP) is a subfield of artificial intelligence (AI) that focuses on the interaction between computers and humans through language.

Processed Text:

natur languag process (nlp) subfield artifici intellig (ai) focus interact comput human languag .

```
In [ ]:
```