

```
!pip install nltk
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.3.2)
Requirement already satisfied: regex<=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2023.6.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.1)
```

```
text = 'Janvi is right of me, Vedantika is left of me.'
```

```
text
```

```
'Janvi is right of me, Vedantika is left of me.'
```

```
from nltk.corpus import stopwords
```

```
import nltk
```

```
nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
True
```

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

```
nltk.download('punkt')
```

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

```
from nltk.tokenize import word_tokenize
words = word_tokenize(text)
```

```
holder = list()
for w in words:
    if w not in set(stop_words):
        holder.append(w)
```

```
holder
```

```
['Janvi', 'right', ',', 'Vedantika', 'left', '.']
```

```
holder = [w for w in words if w not in set(stop_words)]
print(holder)
```

```
['Janvi', 'right', ',', 'Vedantika', 'left', '.']
```

```
from nltk.stem import PorterStemmer, SnowballStemmer, LancasterStemmer
```

```
porter = PorterStemmer()
snow = SnowballStemmer(language = 'english')
lancaster = LancasterStemmer()
```

```
words = ['play', 'plays', 'played', 'playing', 'player']
```

```
porter_stemmed = list()
for w in words:
    stemmed_words = porter.stem(w)
    porter_stemmed.append(stemmed_words)
porter_stemmed
```

```
['play', 'plays', 'played', 'playing', 'player']
```

```
['play', 'play', 'play', 'play', 'player']
```

```
porter_stemmed = [porter.stem(x) for x in words]
print (porter_stemmed)
```