

```
In [25]: import string
        from nltk.corpus import stopwords
```

```
In [26]: data = "Welcome to AI and ML Module.!!!"
```

```
In [27]: data
```

```
Out[27]: 'Welcome to AI and ML Module.!!!'
```

```
In [28]: remove_pun = [ c for c in data if c not in string.punctuation ]
```

```
In [29]: remove_pun
```

```
Out[29]: ['W',
          'e',
          'l',
          'c',
          'o',
          'm',
          'e',
          ' ',
          't',
          'o',
          ' ',
          'A',
          'I',
          ' ',
          'a',
          'n',
          'd',
          ' ',
          'M',
          'L',
          ' ',
          'M',
          'o',
          'd',
          'u',
          'l',
          'e']
```

```
In [30]: sentences = ''.join(remove_pun)
```

```
In [31]: ## Removal of punctuation marks
        sentences
```

```
Out[31]: 'Welcome to AI and ML Module'
```

```
In [32]: ## Conversion of sentence into words
        words = sentences.split(" ")
```

```
In [33]: words
```

```
Out[33]: ['Welcome', 'to', 'AI', 'and', 'ML', 'Module']
```

```
In [34]: ## Removal of stopwords  
vocabulary = [ word for word in words if word not in stopwords.words('english') ]
```

```
In [35]: vocabulary
```

```
Out[35]: ['Welcome', 'AI', 'ML', 'Module']
```

```
In [23]: def textPreprocessing(data):  
         #Removal of Punctuations  
         remove_pun = [ c for c in data if c not in string.punctuation ]  
         sentences = ''.join(remove_pun)  
         #Converting Sentences to Words  
         words = sentences.split(" ")  
         #Removal of Stopwords  
         vocabulary = [ word for word in words if word not in stopwords.words('english') ]  
         #Return Vocabulary  
         return vocabulary
```

```
In [24]: ### Function Calling  
  
         textPreprocessing(data)
```

```
Out[24]: ['Welcome', 'AI', 'ML', 'Module']
```

```
In [ ]:
```