

AIM:

1. Basic concepts of Text Analytics
2. Text Analysis Operations using natural language toolkit
3. Text Analysis Model using TF-IDF.
4. Bag of Words (BoW)

[1]: `import nltk`

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

```
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\SSOS19\AppData\Roam
[nltk_data] Package punkt is already up-to
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\SSOS19\AppData\Roam
[nltk_data] Unzipping corpora\stopwords.zi
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\SSOS19\AppData\Roam
[nltk_data] Package wordnet is already up-
[nltk_data] Downloading package averaged_per
[nltk_data] C:\Users\SSOS19\AppData\Roam
[nltk_data] Package averaged_perceptron_ta
[nltk_data] date!
```

[2]: True

[3]: `text="Padmashree Dr. Dnyandeo Yashwantrao Patil"`

[4]: `from nltk.tokenize import sent_tokenize`

[5]: `x=sent_tokenize(text)`
`print("TOKENIZED TEXT IN SENTENCE\n",x)`

```
TOKENIZED TEXT IN SENTENCE
['Padmashree Dr. Dnyandeo Yashwantrao Patil
ions to education and social service, includ
1.']
```

[6]: `from nltk.tokenize import word_tokenize`