

## Practical No. 04

**AIM** - Write a program for Pre-processing of a Text Document: stop word Removal.

### Source Code -

```
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
example_sent="we are the students of ckt college"
stop_words=set(stopwords.words('english'))
word_tokens=word_tokenize(example_sent)
filtered_sentence=[w for w in word_tokens if not w in stop_words]
filtered_sentence=[]

for w in word_tokens:
    if w not in stop_words:
        filtered_sentence.append(w)
    print(word_tokens)
    print(filtered_sentence)
```

### OUTPUT -

```
In [16]: runfile('C:/Users/ckt/Documents/KUNAL-workspace/P4.py',
wdir='C:/Users/ckt/Documents/KUNAL-workspace')
['we', 'are', 'the', 'students', 'of', 'ckt', 'college']
['students']
['we', 'are', 'the', 'students', 'of', 'ckt', 'college']
['students', 'ckt']
['we', 'are', 'the', 'students', 'of', 'ckt', 'college']
['students', 'ckt', 'college']

In [17]:
```