



# Vidyavardhini's College of Engineering & Technology

## Department of Computer Engineering

---

**Aim:** Perform morphological analysis and word generation for any given text.

**Objective:**

To study morphological analysis.

**Theory:**

Morphological analysis is a field of linguistics that studies the structure of words. It identifies how a word is produced through the use of morphemes. A morpheme is a basic unit of the English language. The morpheme is the smallest element of a word that has grammatical function and meaning. In inflected languages, words are formed through morphological processes such as affixation. For example, by adding the suffix '-s' to the verb 'to dance', we form the third person singular 'dances'.

**Parsing:**

It is the process of determining the morphemes from which a given word is constructed. Morphemes are the smallest meaningful words which cannot be divided further. Morphemes can be stem or affix. Stem are the root word whereas affix can be prefix, suffix or infix. For example-

Unsuccessfull → un success ful

(prefix) (stem) (suffix)

**Program:**

```
import nltk

from nltk.stem import WordNetLemmatizer

nltk.download('wordnet')

lemmatizer = WordNetLemmatizer()

text = "The quick brown foxes are jumping over the lazy dogs"

words = nltk.word_tokenize(text)

lemmatized_words = [lemmatizer.lemmatize(word) for word in words]
```



```
generated_words = []  
for word in lemmatized_words:  
    plural_form = word + 's'  
    generated_words.append(plural_form)  
  
print("Original words:", words)  
print("Lemmatized words:", lemmatized_words)  
print("Generated words:", generated_words)
```

Output:

Original words: ['The', 'quick', 'brown', 'foxes', 'are', 'jumping', 'over', 'the', 'lazy', 'dogs']

Lemmatized words: ['The', 'quick', 'brown', 'fox', 'are', 'jumping', 'over', 'the', 'lazy', 'dog']

Generated words: ['Thes', 'quicks', 'browns', 'foxs', 'ares', 'jumpings', 'overs', 'thes', 'lazys', 'dogs']

**Conclusion:** In computational linguistics and natural language processing (NLP), morphological analysis is of utmost importance. It acts as the first step in comprehending the word structure of a language, enabling NLP models to handle word variants like tense, plurals, and derivational forms successfully. Text normalisation, information retrieval, machine translation, and sentiment analysis are all made easier with accurate morphological analysis. Additionally, it helps to improve language understanding by enabling programmes to precisely grasp and produce human language, which is essential in today's data-driven and language-centric technological landscape.