NLP Experiment 7

Aim: To study and implement Morphological Analysis of English sentences.

Theory:

Morphological analysis in Natural Language Processing (NLP) involves the study and decomposition of words into their morphemes, which are the smallest units of meaning within a language. Morphemes can be prefixes, suffixes, roots, and other linguistic units that contribute to the meaning and structure of a word.

To perform morphological analysis we do the following steps:
- Tokenize the sentences (for which we download the "punkt" module).
- Remove stopwords (for which we download the "stopwords" module).
- We perform Stemming to find the suffix or prefix of a word.

To perform stemming:
- We make a list of affixes (suffixes as well as prefixes)
- first 18 are prefixes, the next 15 are suffixes.
- If it starts with a prefix or ends with a suffix we extract that from the word and show the extracted suffix or prefix for every word.

Libraries and Tools Used:
- Pandas (for loading data)
- nltk (for tokenizing and removing stopwords)