|  |  |
| --- | --- |
| No | **Week 1 - Lab Experiments** |
| 1 | **Lexical Analysis - Preprocessing Text** |
| a) | Write a program to remove punctuations, special symbols, numbers using regular expression |
| b) | Tokenize the given text (Tokenizing the text into sentences) |
| c) | Add Custom Stopwords and List Removed Stopwords |
| d) | Perform stemming and lemmatization on text |
| e) | Extract the usernames from the email addresses present |
| f) | Find the most common words in the text to exclude as stopwords |
| g) | Write a program to Correct the spelling errors using textblob |
| **Apply Level on Text Preprocessing** | |
| a) | Replace social media slangs from a text (create dictionary) |
| b) | Apply stemming or lemmatization only if the word doesn't have any meaning in a dictionary |