

PROCESSING PYTHON ASSIGNMENT



Python Text Processing

"For the third time this monsoon, pounding rains paralysed Mumbai on Wednesday with total disruption in road and rail traffic, though flights were not hit severely. Mumbai turned into an extension of the sea as unprecedented rains deluged the country's financial capital, delaying flights, jamming traffic, flooding homes and having rescue teams out on the roads since morning. In several places, the local trains grinded to a halt, and long distance trains were cancelled. In the evening office hours, the waterlogging pushed people to look for the nearest shelter, abandoning thoughts of reaching home. The met office declared a red alert for the next 24 hours, indicating that the situation might worsen. Schools and junior colleges in Mumbai, Thane and Konkan will remain shut on Thursday as a precautionary measure."

1. Write a program to split the above paragraph into different sentences.

ALGORITHM

Step 1: Loading the library

Step 2: Split the paragraph into sentences

Step 3: Print the different sentences

2. Stop words are commonly used words (such as "the", "a", "an", "in") that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. NLTK has a list of English stop words. Write down a program to make a list of words from above paragraph other than stop words.

ALGORITHM

Step 1: Loading the new library

Step 2: Getting the stop words

Step 3: Tokenize the words

Step 4: Listing of words from other than stop words.

Step 5: Print without stop words

ALGORITHM

Step 5: Printing the nouns

ALGORITHM

Step 2: Printing the verbs

End