

3. Programs on stemming,Lemmatization stopwords

.Take a document with 200 words as the input and perform the following below,

1. Read the document line by line
2. Tokenize the line
3. Stem the words
4. Output the stemmed words (print on screen or write to a file)
5. For the below given sentence ,perform Lemmatization, :”for sentence containing words like visit,visitor,visiting,visited ”
6. Output the Lemma (print on screen or write to a file)
7. Use stop words for the given document(with a minimum of 10) and show the output