Assignment 8

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

1) The process of transforming unstructured text into text involves following steps:
   - Cleaning Text
   - Stopword removal
   - Stemming or lemmatization
   - All of the above

   No, the answer is incorrect.
   Accepted Answers:
   All of the above

2) A large collection of text data in text mining is known as:
   - Text database
   - Cloud Collection
   - Cloud Data
   - All of the above

   No, the answer is incorrect.
   Accepted Answers:
   Corpus

3) The primary assumption that sentence tokenization has already performed is necessary in:
   - Textual Memory
   - Word Tokenization
   - N-gram Sentences Tokenizer

   None of the above

   No, the answer is incorrect.
   Accepted Answers:
   Textual Memory

4) Pattern "\=" in (a-zA-Z0-9) extract from text:
   - Characters
   - Words
   - Alphabet or Characters
   - None of the above

   No, the answer is incorrect.
   Accepted Answers:
   Characters

5) Process of converting the word to its base form or root form is called:
   - Inflection
   - Lemmatization
   - Stemming
   - Both a and c

   No, the answer is incorrect.
   Accepted Answers:
   Both a and c

6) Which of the following are stopwords in the English language?
   - n
   - no
   - the
   - All of the above

   No, the answer is incorrect.
   Accepted Answers:
   All of the above

7) In the text mining, we transform and represent text document as numeric vectors of specific terms. This form of presentation is called:
   - Vector space model
   - Term space model
   - Both a and b
   - None of the above

   No, the answer is incorrect.
   Accepted Answers:
   Both a and b

8) Which of the following features can be used for accuracy improvement of a classification model?
   - Frequency count of terms
   - Part of Speech Tag
   - Dependancy Grammar
   - All of these

   No, the answer is incorrect.
   Accepted Answers:
   All of these

9) What is the exact order for text classification model components?
   - Text cleaning
   - Text normalization
   - Stemming
   - Text predictions

   No, the answer is incorrect.
   Accepted Answers:
   1,2,3,4

10) A document term matrix for the tweets is given, where every tweet has been treated as one document. Which of the following is correct?
    - Removal of stopwords from the text will reduce the dimensionality of the matrix
    - Normalization of weights of the matrix will produce more relevant features

    No, the answer is incorrect.
    Accepted Answers:
    1 and 2