Assignment 1

Due on 2019-06-16, 23:59 IST.

1. What data type is represented by an image in Python/DS?
   - String
   - Long
   - Real
   - Real
   Accepted Answers:
   - String

2. Which is the correct statement regarding similarities between text mining and data mining?
   - Both refer to natural language data analysis
   - Both refer to learning from variations of present experience or examples
   - Both refer to finding valuable patterns in data
   Accepted Answers:
   - Both refer to finding valuable patterns in data

3. Which of the following techniques is not a part of flexible text matching?
   - Soundex
   - Metaphone
   - City Distance
   Accepted Answers:
   - City Distance

4. What are the possible features of a text corpus?
   - 1. Count of word in a document
   - 2. Blocker feature – presence of word in a document
   - 3. Vector instead of word
   - 4. Part of Speech Tag
   - 5. Basic Dependency Grammar
   - 6. Entire document as a feature
   Accepted Answers:
   - 1, 2, 3, 4, 5, 6

5. Which of the following is a database for text mining?
   - TREC
   - TREC
   - TREC
   Accepted Answers:
   - TREC

6. While working with test data obtained from news articles, which of the grammar-based text processing techniques can be used for text pre-processing, detection, and sentence detection, subject detection, and object detection?
   - Part-of-speech tagging
   - Dependency Parsing and Constituency Parsing
   - Stop Gram and D-Gram variations
   - Common Bag of Words
   Accepted Answers:
   - Dependency Parsing and Constituency Parsing

7. Text mining is a field of:
   - Computer Science
   - Artificial Intelligence
   - Natural language processing (NLP)
   - All of the mentioned
   Accepted Answers:
   - All of the mentioned

8. Text mining can be used for:
   - Detecting spam
   - Predicting stock Movements
   - Finding trends in stock
   - All of the above
   Accepted Answers:
   - All of the above

9. The text mining process involves:
   - Data collection
   - Data preprocessing
   - Data analysis
   - All of the above
   Accepted Answers:
   - All of the above

10. The text mining project is a part of:
    - Data mining
    - Decision tree
    - Dataset mining
    - All of the above
    Accepted Answers:
    - All of the above