# NLP

1. Which of the following reduce a word to its base form by cutting off the suffix?
   A) Lancaster Stemmer B) Porter Stemmer C) Snowball Stemmer D) WordNetLemmatizer
   **D) WordNetLemmatizer**

2. We need to perform stemming and lemmatization sothat:
   A) All the words can be reduced to their base form
   B) so that we do not end up with too many words in the vocabulary which are not adding information to the model.
   C) so that lengths of words are reduced.
   D) None of the above
   **A) All the words can be reduced to their base form**
   **B) so that we do not end up with too many words in the vocabulary which are not adding information to the model.**
   **C) so that lengths of words are reduced.**

3. Stemming and Lemmatization belongs to which of the following step in NLP?
   A) Semantic Processing B) Syntactic Processing C) Lexical Processing D) All of the above
   **C) Lexical Processing**

4. Which of the following is/are example of shallow parsing?
   A) POS tagging B) Chunking C) tokenization D) None of the above
   **A) POS tagging**
   **B) Chunking**

5. Which of the following are true regarding Lexicon Based taggers?
   A) These taggers assign that POS tag to the word whose frequency is maximum for that word in the training Corpus.
   B) These taggers also use tag of the previous word to find the tag of the word.
   C) These taggers take in to account the context in which the word is used to assign a tag.
   D) All of the above
   **D) All of the above**

6. Which of the following taggers uses predefined rules to assign tags?
   A) HMM B) Stochastic Rule taggers C) Rule Based Taggers D) None of the above
   **C) Rule Based Taggers**

7. Which of the following is /are true regarding HMM based POS tagger?
   A) It is used for tokenization.
   B) It uses tag of only the previous word to determine the tag of the current word.
   C) It assigns tag by finding the most frequent tag occurring for that word in the training corpora
   D) None of the above
   **B) It uses tag of only the previous word to determine the tag of the current word.**
   **C) It assigns tag by finding the most frequent tag occurring for that word in the training corpora**

8. What does the transition probability refer in to HMM bases POS tagging algorithm?
   A) The transition probabilities refer to probabilities of transitioning from one tag to another tag.
   B) Transition probabilities refer to the probability of emitting a given word from a tag.
   C) Transition probabilities are the probabilities of most occurring tag.
   D) HMM does not have the concept of transition probabilities.

9. Which of the following are terminal symbols in the following Context-Free Grammar? S -> NP VP NP -> DT N| N| N PP VP -> V| V NP N -> 'man'| 'bear' V -> 'ate' DT -> 'the'| 'a'
   A) 'a' B) 'ate' C) VP D) NP
   **A) 'a'**

10. In which of the cases Hidden Markov Model can be used?
    A) Modeling a Sequential process
    B) POS tagging
    C) Word Tokenization
    D) None of the above
    **A) Modeling a Sequential process**
    **B) POS tagging**

11. Which of the following is/are used to get the grammatical construction of the sentence?
    A) POS tagging B) Constituency Parsing C) Top-Down Parsing D) HMM based POS tagging
    **A) POS tagging**
    **D) HMM based POS tagging**

12. Which of the following are the approaches of constituency parsing?
    A) Top-Down Parsing B) Bottom-up Parsing C) Dependency Parsing D) None of these
    **B) Bottom-up Parsing**

13. Which of the following is true regarding Top-Down parsing?
    A) It starts with start symbol S.
    B) we use the CFG production rule to generate the sentence from the S start symbol.
    C) It starts with sentence and then we reduce it to the S symbol.
    D) All of the above
    **A) It starts with start symbol S.**

14. Which of the following statements are true regarding shift reduce parser algorithm?
    A) We start with start symbol S, then we use production rules of CFG and reach the sentence in the end.
    B) It's an algorithm of Bottom up parsing.
    C) In this algorithm we start from the sentence, take one word at a time from the sentence shift it to the stack or reduce the words present in the stack by using CFG rules, until we reach the S startsymbol.
    D) All of the above
    **B) It's an algorithm of Bottom up parsing.**

15. Which of the following are true regarding Chomsky Normal Form?
    A) It is normalized form of a CFG.
    B) The production rules can be written only in a particular way as defined by a set of rules.
    C) A CFG with no terminal symbol is called Chomksy Normal Form.
    D) It is used for POS tagging.
    **A) It is normalized form of a CFG.**

16. In Which of the following text processing technique we will remove stopwords as a preprocessing?
    A) Top-Down Parsing
    B) Bottom-Up parsing
    C) Count-vectorization to create BOW for lexical level analysis.
    D) All of the above
    **C) Count-vectorization to create BOW for lexical level analysis.**