**NLP**

## **WORKSHEET – 3**

## **Solutions**

1. A) Lancaster Stemmer B) Porter Stemmer C) Snowball Stemmer D) WordNetLemmatizer
2. 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. A) Semantic Processing
4. 6
5. 7
6. 8
7. B) It uses tag of only the previous word to determine the tag of the current word
8. A) The transition probabilities refer to probabilities of transitioning from one tag to another tag
9. C) VP D) NP
10. B) POS tagging
11. A) POS tagging C) Top-Down Parsing D) HMM based POS tagging
12. A) Top-Down Parsing B) Bottom-up Parsing C) Dependency Parsing
13. D) All of the above
14. 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.
15. C) A CFG with no terminal symbol is called Chomksy Normal Form.

D) It is used for POS tagging.

**16)** C) Count-vectorization to create BOW for lexical level analysis.