**WORKSHEET-3 NLP**

1- Porter Stemmer, Lancaster Stemmer, Snowball Stemmer.

2- All the words can be reduced to their base form

so that we do not end up with too many words in the vocabulary which are not adding information to the model.

3- All of the above

4- Chunking

5- All of the above

6- Rule Based Taggers

7- It is used for tokenization

It uses tag of only the previous word to determine the tag of the current word.

8- The transition probabilities refer to probabilities of transitioning from one tag to another tag.

9- ‘a’

10- POS tagging

11- POS tagging

HMM based POS tagging

12- None of these

13- It starts with start symbol S.

we use the CFG production rule to generate the sentence from the S start symbol.

14- It’s an algorithm of Bottom up parsing.

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 start symbol.

15- It is normalized form of a CFG.

A CFG with no terminal symbol is called Chomksy Normal Form.

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