**NLP – WORKSHEET 4**

1. Which of the following are true regarding Context Free Grammars?

Answer: A) It consists of a set of production rules

B) The production rules are of the following form: A-> BC where A is non terminal while B, C can be either terminal or non-terminal

C) These grammars are free of context in which they are used, they will remain same regardless of the context in which they are used.

2. Advantages of using PCFG over CFG are:

Answer: A) All the production rules in PCFG has probability associated with them while in CFG we do not have Probability of a production rule.

B) With PCFG we can find the most probable parse tree of a sentence which we cannot find CFG.

3. The problem with constituent parsing is:

Answer: D) All of the above

4. In order to deal with free word order languages what type of parsing is suitable?

Answer: A) Free word parsing

5. Which of the following are true regarding Dependency Parsing?

Answer: A) It establish dependencies between words of a sentence B) The dependencies are established in terms of subject-object-verb and other dependencies.

6. Which of the following are techniques for Named Entity Recognition?

Answer: A) Chunking C) unigram chunker D) bigram chunker

7. Which of the following statements are true regarding Unigram chunker for NER?

Answer: A) It uses the POS tag of a word and find the most probable IOB label for that POS tag

8. Which of the following statements are true regarding Unigram chunker for NER?

Answer: C) It assigns that IOB label which has the maximum probability based on the POS tag

9. Which of the following statements are true regarding Bigram chunker for NER?

Answer: A) It is a sequential modeling process for assigning POS tags to the word

B) It uses a dictionary of IOB labels to assign a IOB entity label

C) It uses the POS tag of the word and its previous word to assign the most probable IOB label

10. Which of the following technique uses a dictionary to extract an entity?

Answer: B) Bigram chunker

11. Which of the following are preprocessing steps in Information Extraction Systems?

Answer: A) Word Tokenization B) Lemmatization D) POS tagging

12. Which of the following cases require NER (Named Entity Recognition)?

Answer: A) POS Tagging B) Chat Bot creation C) A flight booking system which books flights for a customer according to information given by him in query D) All of the above

13. Which of the following is true regarding Top-Down parsing?

Answer: A) It starts with start symbol S B) we use the CFG production rule to generate the sentence from the S start symbol

14. Consider the following string and tell what should be the pattern to extract all #tags String = “there should be justice for #sushant #singh @rindia” Import re Re.search(pattern, string)

Answer: A) pattern = “#\w\*”

15. Consider the following string and tell what should be the pattern to extract the mention (@rindia) only String = “there should be justice for #sushant #singh @rindia” Import re Re.search(pattern, string)

Answer: C) pattern = “@\w\*”