NLP – WORKSHEET 5

All the questions in this worksheet have one or more than one correct answers. Choose all the correct options to answer your questions.

1. Which of the following NLP tasks are done by sequential labelling technique?
   A) POS tagging   B) Named Entity Recognition
   C) Speech recognition   D) All of the above
   ANS D

2. Word ambiguity is major challenge in NLP. What type of ambiguity exists in the word sequence "Time flies"?
   A) Semantic   B) Syntactic
   C) Phonological   D) None of the above
   ANS B

3. In NLP many words have more than one meanings but we have to select the meaning which makes the most sense in the context in which it was used. This problem of finding the most likely sense can be resolved by:
   A) Shallow Semantic Analysis   B) Word Sense Disambiguation
   C) Fuzzy Logic   D) None of the above
   ANS A

4. Which of the following techniques are used to reduce inflected words to their base form?
   A) Lemmatization   B) Stemming
   C) Dependency Parsing   D) None of the above
   ANS B

5. Which of the following are challenges in NLP?
   A) POS tagging   B) Handling Tokenization
   C) Word Sense Disambiguation   D) None of the above
   ANS B

6. In which of the following areas NLP can be used?
   A) Information retrieval from text   B) Automatic summarization
   C) Making sales forecasting from past data   D) None of the above
   ANS B

7. What is Morphological Segmentation?
   A) Does Discourse Analysis   B) Is an extension of propositional logic
   C) Separate words into individual morphemes and identify the class of the morphemes
   D) All of the above
      ANS C

8. Which of the following are word embedding techniques which are used to capture the semantics of a word?
   A) Word2Vec   B) Top-Bottom parsing
   C) GloVe   D) None of the above
   ANS A

9. Advantages of Word Embeddings are:
   A) They capture the semantics of the word.
   B) The word embeddings provide dense vector for the word representation
   C) The word embeddings are sparse vectors   D) All of the above
      ANS D

10. Which of the following can be used to match similarity between two word vectors?
    A) Cosine similarity   B) POS tags similarity
    C) L2 normalization   D) None of the above
    ANS B

11. The term distributional semantics basically says that:
    A) The words similar to each other in meaning will have same POS tags
    B) The words which occur in similar contexts tend to have similar semantics
    C) The dependency parsing represents the meaning of a word
    D) The words with dissimilar contexts tend to have similar meanings

ANS B

12. Which of the following are used to represent words as vectors?
    A) Occurrence matrix   B) Dependency parsing
    C) Co-occurrence matrix   D) All of the above
    ANS D

13. The problem with occurrence and co -occurrence matrix is:
    A) The word vectors does not capture the semantics
    B) The word vectors are very complex
    C) The word vectors are very high dimensional
    D) None of the above

    ANS C

14. The advantages of using word embeddings over co-occurrence matrix are:
    A) The word embedding are very high dimensional
    B) The word embeddings are very sparse
    C) The word embeddings are very dense and low-dimensional vectors are compared to co-occurrence matrix word vectors.
    D) The word embeddings are much more complex
    ANS C
15. Which of the following techniques are used for creation of word embeddings?
    A) we try to generate a word given the context of the word by using a deep neural network
    B) we use naïve Bayes to create word embeddings
    C) The word embeddings are created by performing SVD on co-occurrence matrix
    D) None of the above
    ANS B