



## BRAINWARE UNIVERSITY

Term End Examination 2023-2024  
Programme – B.Tech.(CSE)-AIML-2021  
Course Name – Speech and Natural Language Processing  
Course Code - PEC-CSM602C  
( Semester VI )

Full Marks : 60

Time : 2:30 Hours

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

### Group-A

(Multiple Choice Type Question)

1 x 15=15

1. Choose the correct alternative from the following :

- (i) Select which of the following is NOT a common type of machine learning algorithm used in NLP?
- a) Naive Bayes  
b) Support Vector Machines (SVMs)  
c) Decision Trees  
d) K-Nearest Neighbors (KNN)
- (ii) Select the main challenge associated with using deep learning models in NLP
- a) They are not accurate enough for real-world applications.  
b) They are easy to understand and interpret  
c) They require large amounts of data for training  
d) They are not suitable for tasks like sentiment analysis
- (iii) Identify the potential impact of advances in machine learning on the future of NLP
- a) NLP tasks will become increasingly complex and require specialized hardware  
b) NLP models will be able to understand and respond to natural language perfectly  
c) NLP applications will become more accessible and user-friendly  
d) Human language understanding will be completely replaced by machines
- (iv) Which of the following best describes the use of forward chaining inference systems in cognitive modeling?
- a) They are primarily utilized for data storage.  
b) They are extensively employed in cognitive modeling tasks.  
c) They are mainly employed for backward reasoning.  
d) They are primarily used for mathematical calculations.
- (v) Select the correct characteristic of the resolution technique.
- a) Efficient  
b) Sound and Complete  
c) Biased  
d) Incomplete
- (vi) Which of the following is associated to a restricted class of pairings where one lexeme denotes a subclass of the other?
- a) Hyponymy  
b) Polysemy

- c) Monosemy  
 (vii) Choose the correct option from the following:  
 a) Discourse Analysis is beneficial for Information retrieval  
 b) Discourse Analysis is not beneficial for Question answering  
 c) Coherence in terms of Discourse in NLP means making sense of repetition of words  
 d) Discourse Analysis is partly beneficial for Information retrieval
- (viii) Determine the correct type of segmentation of unsupervised discourse segmentation:  
 a) Linear  
 b) Non-linear  
 c) Exponential  
 d) Probabilistic
- (ix) "The new restaurant downtown is excellent. They have a delicious menu with fresh, seasonal ingredients." Determine the type of discourse relation.  
 a) Result  
 b) Explanation  
 c) Parallel  
 d) Elaboration
- (x) Choose the correct statement regarding Treebank:  
 a) Represents the grammatical structure of words  
 b) Represents the grammatical structure of sentences  
 c) Represents the grammatical structure of paragraphs  
 d) Represents the grammatical structure of documents
- (xi) Choose the correct option from the following regarding Supervised discourse segmentation:  
 a) It needs to have generic training data  
 b) It needs to have unlabeled training data  
 c) It needs to have boundary-labeled training data  
 d) It needs to have boundary-labeled testing data
- (xii) Trace the word which would definitely be rejected by the English syntactic analyzer?  
 a) Rahul is going to school  
 b) Rahul is coming from school  
 c) Rahul is in the school  
 d) The Rahul the go to the school
- (xiii) In Context-Free Grammars (CFGs), non-terminal is used to interpret the  
 a) parts of speech (e.g., noun, verb)  
 b) categories of words or phrases (e.g., NP, VP)  
 c) specific words in a sentence  
 d) punctuation marks
- (xiv) The aim of Syntactic parsing is to examine the text and:  
 a) Correct spelling errors in text  
 b) Break down text into individual words (tokenization)  
 c) Reveal the grammatical structure of a sentence  
 d) Translate text from one language to another
- (xv) Predict the correct one which is NOT typically used in a rule-based parsing system:  
 a) Finite-state automata  
 b) Context-free grammars (CFGs)  
 c) Recurrent neural networks (RNNs)  
 d) Lexicons (dictionaries) containing word information

### Group-B

(Short Answer Type Questions)

3 x 5=15

2. Apply the CYK algorithm to parse a simple sentence with probabilities assigned to CFG rules. (3)
3. State the common evaluation metrics used to assess the performance of speech recognition systems. (3)
4. Explain the benefits of Discourse Analysis. (3)
5. Explain representation of meaning in Natural Language Processing (NLP). (3)
6. Compare Lemmatization and Stemming. (3)

OR

- Summarize the challenges in Coreference Resolution in terms of Ambiguity, Nested Coreferences and Long-Distance Dependencies. (3)

**Group-C**  
(Long Answer Type Questions)

5 x 6=30

7. Describe the following regarding requirements of representation: Neural Network Architectures, Frame Semantics (5)
8. Recognize the importance of several components that NLP is made up of. (5)
9. Differentiate between the following: WSD using Supervised learning, WSD using Dictionary, WSD using Thesaurus. (5)
10. How to apply context-free grammar rules to analyze the structure of simple English sentences. (5)
11. Summarize the main aspects of Reference Phenomena. (5)
12. Write down the steps of the Hobb's algorithm. (5)

**OR**

Write down the steps of Coreference Resolution. (5)

\*\*\*\*\*