

28. a. What are the ways of representing the knowledge graph and semantic networks? Discuss with typical example for their needs. 10 4 3 4

(OR)

b. Explain in detail about discourse and its relevance to test coherence. 10 3 3 4

29. a. Justify the role of lingual based approach for machine translation. Elaborately your answer with relevant example. 10 4 4 5

(OR)

b. Distinguish extractive and abstractive summarization with an example. 10 3 4 4

30. a. Write short notes on: 10 2 5 1

- (i) Word embeddings
- (ii) Sentence embeddings

(OR)

b. Explain in detail about probabilistic and statistical approach. 10 3 5 4

\*\*\*\*\*

Reg. No.

**B.Tech. DEGREE EXAMINATION, NOVEMBER 2022**

Sixth/ Seventh Semester

**18CSE359T – NATURAL LANGUAGE PROCESSING**

(For the candidates admitted from the academic year 2020-2021 and 2021 -2022)

**Note:**

- (i) **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40<sup>th</sup> minute.
- (ii) **Part - B** should be answered in answer booklet.

Time: 2½ Hours

Max. Marks: 75

**PART – A (25 × 1 = 25 Marks)**

Answer **ALL** Questions

- |   | Marks | BL | CO | PO |
|---|-------|----|----|----|
| 1. What is the main challenges of NLP?<br>(A) Handling ambiguity of sentences<br>(B) Handling tokenization<br>(C) Handling pos-tagging<br>(D) Handling semantic   | 1     | 1  | 1  | 1  |
| 2. Which is the process of turning different morphologies of a word into its base form?<br>(A) Ngrams<br>(B) Tokenization<br>(C) Lemmatization<br>(D) Stopwords   | 1     | 1  | 1  | 2  |
| 3. The joint probability can be calculated using the _____.<br>(A) Conditional probability<br>(B) Probability<br>(C) Sentence<br>(D) Word   | 1     | 1  | 1  | 1  |
| 4. _____ is a process of assigning corresponding part of speech like noun, verb, adverb, adjective to each work in a<br>(A) Part of speech tagging<br>(B) Name entity tagging<br>(C) Parsing<br>(D) Disambiguation      | 1     | 1  | 1  | 2  |
| 5. How many bi-grams can be generated from the given sentence: "Gandhiji is the father of our nation"<br>(A) 7<br>(B) 6<br>(C) 8<br>(D) 9   | 1     | 2  | 1  | 2  |
| 6. Which of the following tasks would help us in choose the right meaning as per the context in which the word is used?<br>(A) Lemmatization<br>(B) Word sense disambiguation<br>(C) Stemming<br>(D) Discourse analysis | 1     | 1  | 2  | 1  |
| 7. Words may have multiple meanings. This leads to what type of ambiguity in NLP?<br>(A) Syntactic ambiguity<br>(B) Anaphoric ambiguity<br>(C) Semantic ambiguity<br>(D) Lexical ambiguity                              | 1     | 1  | 2  | 2  |

8. Which of the following is an example of indirect personal pronoun, reflexive  
(A) They (B) To-myself  
(C) myself (D) One – another
9. “I went to the school and they told me come on next day”. What type of ambiguity present in the given sentence?  
(A) Anaphoric ambiguity (B) Syntactic ambiguity  
(C) Semantic ambiguity (D) Lexical ambiguity
10. Which is a model of measuring the incidence of known words?  
(A) A low weight in TF – IDF (B) A high weight in TF – IDF  
(C) A bag of words (D) A corpus
11. What does the phenomena that operates at discourse level include?  
(A) Cohesion and coherence (B) Corrosion and erosion  
(C) Connection and resolution (D) Cooperation and coordination
12. Which of the following is an NLP task that involves determining all referring expressions that point to the same real world entity.  
(A) Named entity recognition (B) Coreference resolution  
(C) Information resolution (D) Discourse analysis
13. Which of the following is the graph used to represent semantic network?  
(A) Undirected graph (B) Directed complete graph  
(C) Directed acyclic graph (D) Directed graph
14. Select the extension of the semantic network?  
(A) Partitioned networks (B) Rule based expert systems  
(C) Decision tree based networks (D) Expert systems
15. Which of the following is the frame?  
(A) Data structure (B) A way of representing knowledge  
(C) Data type (D) Procedures and default values
16. Which of the following is the local method for improving recall of an information retrieval system?  
(A) Query expansion (B) Ontology based model  
(C) Relevance feedback (D) Language model
17. The vocabulary size (unique words) of a text can be estimated using  
(A) Zipf’s law (B) Inverted index rule  
(C) Scientific law (D) Heaps law
18. A metric used to measure the importance of a term in a text document collection is called  
(A) Inverse document frequency (B) Term frequency  
(C) Inverse term frequency (D) Document frequency

19. Machine translation converts \_\_\_\_\_.  
(A) Human language to machine language (B) One human language to another language  
(C) Any human language to English (D) Machine language to human language
20. In NLG \_\_\_\_\_ includes retrieving the relevant content from knowledge base1  
(A) Sentence planning (B) Text realization  
(C) Text planning (D) Text mapping
21. Same word can have multiple words embeddings possible with \_\_\_\_\_?  
(A) ELMO (B) Glove  
(C) wordzVec (D) Nltk
22. Which will be a suitable NLP method for Covid’19 news analysis from the online newspapers?  
(A) NER (B) Machine translation  
(C) Sentiment analysis (D) Text summarization
23. In RNN each unit has an internal state which is called the \_\_\_\_\_.  
(A) Visible state of the unit (B) Hidden state of the circle  
(C) Visible function (D) Hidden function
24. Spam email detection comes under which domain?  
(A) Text categorization (B) NER  
(C) Text classification (D) Sentiment analysis
25. When training a language model, if we use an overaly narrow corpus, the probabilities?  
(A) Don’t generalize (B) Don’t reflex the task  
(C) Reflect all possible working (D) Reflect intuition

**PART – B (5 × 10 = 50 Marks)**

Answer ALL Questions

Marks BL CO PO

26. a. Define vector representation of words. Explain in detail about language models. 10 3 1 4
- (OR)
- b. How will you represent the regular expression? Explain in detail about the steps of NLP. 10 3 1 4
27. a. Discuss in detail about 10 2 2 4  
(i) Pronoun resolution  
(ii) Dependency parsing
- (OR)
- b. Explain in detail word sense disambiguation techniques. 10 3 2 4