

M.Sc. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2023.

Third Semester

Computer Science with Artificial Intelligence–Core

NATURAL LANGUAGE PROCESSING

(For those who joined in July 2022 onwards)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer.

1. What is the field of Natural Language Processing (NLP)?
 - (a) Computer Science
 - (b) Artificial Intelligence
 - (c) Linguistics
 - (d) All of the mentioned

6. Which is a collection of documents?
 - (a) Corpus
 - (b) Ngrams
 - (c) Tokenization
 - (d) Stopwords
7. Natural language processing is divided into the two subfields of
 - (a) symbolic and numeric
 - (b) algorithmic and heuristic
 - (c) time and motion
 - (d) understanding and generation
8. The process of converting data to something a computer can understand is referred to as _____
 - (a) Post processing
 - (b) Pre processing
 - (c) Pre defined
 - (d) Post defined
9. Which of the following is an example of a natural language processing task?
 - (a) Creating a website layout
 - (b) Designing a logo
 - (c) Identifying named entities in a text
 - (d) Generating a musical composition

2. Choose form the following areas where NLP can be useful.
 - (a) Automatic Text Summarization
 - (b) Automatic Question-Answering Systems
 - (c) Information Retrieval
 - (d) All of the mentioned
3. Given a sound clip of a person or people speaking, determine the textual representation of the speech.
 - (a) Text-to-speech
 - (b) Speech-to-text
 - (c) Security-to-text
 - (d) None of the Above
4. Which of the following is the main challenge of NATURAL LANGUAGE PROCESSING?
 - (a) Handling Ambiguity of Sentences
 - (b) Handling Tokenization
 - (c) Handling POS-Tagging
 - (d) All of these
5. Which step is the process of breaking down documents into smaller units of analysis?
 - (a) Ngrams
 - (b) Tokenization
 - (c) Stopwords
 - (d) Corpus

10. Which of the following is an example of a language model that uses a probabilistic approach?
 - (a) Hidden Markov model (HMM)
 - (b) Rule-based model
 - (c) Decision tree
 - (d) Convolutional neural network

PART B — (5 × 5 = 25 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 250 words.

11. (a) Elaborate the applications of natural language processing.

Or

 (b) Describe the basic concepts of parts-of-speech.
12. (a) Explain the main purpose of n-gram.

Or

 (b) How will you evaluate language models? Explain.
13. (a) Write down the early approaches of parts-of-speech.

Or

 (b) What are the basic concepts of parts-of-speech? Describe.

14. (a) Summarize the methods of syntactic parsing.
Or
(b) Differentiate between the top down and bottom up parsing.

15. (a) Point out the advantages of vector semantics.
Or
(b) Assess the uses of WordNet.

PART C — (5 × 8 = 40 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 600 words.

16. (a) Enumerate the different types of phases of natural language processing.
Or

- (b) Outline the spelling error and noisy channel model.

17. (a) Discuss the training methods of neural network with example.

Or

- (b) Analysis the application of neural language model in NLP system development.

Page 5 Code No. : 7743

18. (a) Examine the importance of POS tagging using HMM.

Or

- (b) Determine the POS tagging using neural model with diagram.

19. (a) Evaluate the basic of statistical parsing with example.

Or

- (b) Explain the implementation of Probabilistic Context Free Grammar.

20. (a) What are the methods of measuring similarity in vector? Explain.

Or

- (b) Illustrate the difference between skip-gram and CBOW.

Page 6 Code No. : 7743

