Reg. No.																

Question Paper Code	13237

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2024

Fifth Semester

Artificial Intelligence and Data Science

(Common to Computer Science and Engineering (AIML))

20AIPC503 – NATURAL LANGUAGE PROCESSING AND CHATBOT

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (MCQ) (20 × 1 = 20 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	What does the term "tokenization" refer to in NLP? (a) The process of converting text into tokens or individual words and punctuation marks (b) The process of removing stop words (c) The process of summarizing text (d) The process of translating text	1	K1	CO1
2.	What is the purpose of a stop word list in NLP? (a) To identify the main subjects of a text (b) To remove common words that do not carry significant meaning (c) To convert text into lowercase (d) To highlight important keywords	1	K1	CO1
3.	Which algorithm is typically used for language modelling and text generation? (a) Random Forest (b) Recurrent Neural Networks (RNN) (c) Decision Tree (d) k-Means Clustering	1	K1	CO1
4.	What is the Bag of Words (BoW) model in NLP? (a) A model that captures the order of words in a text (b) A model that represents text as an unordered collection of words (c) A model that uses neural networks for text processing (d) A model that identifies named entities in text	1	K1	CO2
5.	What is the purpose of using term frequency (TF) in the BoW model? (a) To count the number of documents containing a word (b) To measure the importance of a word in a document (c) To count the number of times a word appears in a document (d) To normalize the text	1	K1	CO2
6.	Words may have multiple meanings. This leads to what type of ambiguity in NLP? (a) Syntactic ambiguity (b) Anaphoric ambiguity (c) Semantic ambiguity (d) Lexical ambiguity	1	K1	CO2
7.	What is the primary purpose of a Context-Free Grammar (CFG)? (a) To define the syntax of programming languages (b) To perform semantic analysis (c) To convert natural language to machine code (d) To analyze the meaning of sentences	1	K1	CO3
8.	What is a primary characteristic of lexicalized parsing? (a) It ignores the specific words in a sentence. (b) It incorporates specific words into the parsing process. (c) It relies solely on statistical models. (d) It uses fixed grammatical rules without variation.	1	K1	CO3
9.	Which of the following terms refers to the phenomenon where two words have the same spelling and pronunciation but different meanings? (a) Synonymy (b) Antonymy (c) Homonymy (d) Polysemy	1	K1	CO3

10. What is machine translation (MT)? *1 K1 CO4*
    (a) The process of translating human speech into text
    (b) The process of converting one programming language into another
    (c) The use of computers to translate text or speech from one language to another
    (d) The process of converting handwritten text into digital format
11. Which of the following types of machine translation relies heavily on statistical models and large corpora of bilingual text? *1 K1 CO4*
    (a) Rule-based MT (RBMT) (b) Statistical MT (SMT)
    (c) Neural MT (NMT) (d) Hybrid MT
12. What does BLEU stand for in the context of evaluating machine translation? *1 K1 CO4*
    (a) Bilingual Evaluation Understudy (b) Best Language Evaluation Unit
    (c) Basic Language Understanding Engine (d) Bilingual Language Estimation Utility
13. What is the primary goal of Natural Language Understanding (NLU)? *1 K1 CO5*
    (a) To convert text to speech
    (b) To understand and extract meaning from user input
    (c) To generate a response based on user input
    (d) To maintain conversation context
14. In dialog management (DM), which component is responsible for determining the next action based on the current state? *1 K1 CO5*
    (a) Intent Recognizer (b) Entity Extractor (c) Policy Manager (d) Response Generator
15. Which of the following best describes the role of entity extraction in NLU? *1 K1 CO5*
    (a) Converting user input to structured data (b) Generating natural language responses
    (c) Understanding user sentiment (d) Managing dialog state
16. What technique is commonly used in NLG to ensure that responses are coherent and contextually appropriate? *1 K1 CO5*
    (a) Rule-based generation (b) Template-based generation
    (c) Statistical generation (d) Neural network generation
17. Which feature of Microsoft Bot Framework allows developers to create bots that can adapt their responses based on user interactions? *1 K1 CO6*
    (a) QnA Maker (b) LUIS (Language Understanding Intelligent Service)
    (c) Bot Connector (d) Adaptive Cards
18. In Dialogflow, what is a "fulfillment" webhook used for? *1 K1 CO6*
    (a) Storing user session data (b) Providing dynamic responses based on backend logic
    (c) Translating user input (d) Generating static responses
19. What is the primary benefit of using a rule-based approach in NLG? *1 K1 CO6*
    (a) High flexibility and adaptability
    (b) Consistency and predictability in responses
    (c) Lower computational requirements
    (d) Easier integration with machine learning models
20. Which of the following is a common challenge in dialog management? *1 K1 CO6*
    (a) Text-to-speech conversion
    (b) Maintaining context over multiple turns
    (c) Identifying user intent
    (d) Generating creative responses

PART - B (10 × 2 = 20 Marks)

Answer ALL Questions

21. What is meant by Language Modelling? *2 K1 CO1*
22. What are the applications of NLP? *2 K1 CO1*
23. Define Bagging. *2 K1 CO2*
24. Define WSD. *2 K1 CO2*
25. Define meaning of a sentence in terms of lexical and semantic understanding. *2 K1 CO3*
26. Define Parsing. *2 K1 CO3*

27.	What are the various summarization methods in MT?	2	K1	CO4
28.	What are the different Language divergence methods?	2	K1	CO4
29.	What is meant by Vauquois Triangle?	2	K1	CO5
30.	What is meant by Constatives?	2	K1	CO6

PART - C (6 × 10 = 60 Marks)

Answer ALL Questions

31.	a) i)	Explain the architecture of NLP System.	7	K2	CO1
	ii)	Describe Language Modelling.	3	K2	CO1
		OR			
	b) i)	Compare and Contrast between Bigram and Trigram Model with evaluation pattern.	5	K2	CO1
	ii)	Explain the applications of NLP.	5	K2	CO1
32.	a)	Write about Bag-of-words. Elaborate on the method to calculate average word length and tf-idf.	10	K2	CO2
		OR			
	b)	Explain in detail about Embedding representation of Words.	10	K2	CO2
33.	a)	What is Chomsky Hierarchy and what is its use? Explain.	10	K2	CO3
		OR			
	b)	Explain Probabilistic CFGs with example.	10	K2	CO3
34.	a)	Classify the Machine Translation System in the NLP. Explain.	10	K2	CO4
		OR			
	b)	Discuss the objectives of Encoder- decoder model.	10	K2	CO4
35.	a)	Explain about Statistical and Knowledge Based Machine Translation System.	10	K2	CO5
		OR			
	b)	Explain Embedded based methods in detail.	10	K2	CO5
36.	a)	Explain in detail the Architecture in Chatbot system.	10	K2	CO6
		OR			
	b)	Explain in detail the dialog system in Amazon Alexa System.	10	K2	CO6