



VIT

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.: *0629*

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
WINTER SEMESTER 2025-2026

SLOT:C1+TC1

Programme Name & Branch : B.Tech CSE
Course Code and Course Name : BCSE409L , Natural Language Processing
Faculty Name(s) : Dr. SARAVANAKUMAR K, Dr. KANNADASAN R,
 Dr. RAJESHKANNAN R, Dr. SATHYARAJ R,
 Dr. JABANJALIN HILDA J, Dr. ARIVOLI A
Class Number(s) : VL2025260502034, 2035, 2032, 2033, 2039, 2270
Date of Examination : 17/03/26
Exam Duration : 90 minutes **Maximum Marks: 50**

General instruction(s): Answer All Questions

- M - Max mark; CO – Course Outcome; BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Create)
- Course Outcomes

CO3 - Check the syntactic and semantic correctness of natural language.

CO4 – Select a suitable language modelling and feature representation to develop real-world applications.

Q. No	Question	M	CO	BL
1.	Consider the sentence: “The young scientist analyzed the complex data in the laboratory.” Using transition-based dependency parsing, derive the dependency structure for the given sentence. Present the step-by-step parsing process in a transition table, listing the stack, input word list (buffer), action taken, and dependency relation added at each step. Finally, draw the resulting dependency tree/graph.	10	3	3
2.	Consider the following sentences: 1. The idea slept quietly. 2. The table opened the door with a key. 3. The number two is red. 4. The car read the newspaper. For each sentence, perform Semantic Role Labelling (SRL) by identifying the appropriate semantic roles. Examine whether these roles satisfy the selectional restrictions of the predicate. Based on this analysis, determine whether the sentence is semantically valid or anomalous, and justify your answer.	10	3	4
3.	Consider the following sentences: Sentence 1: The young boy met Ravi near the park. Sentence 2: He greeted him with a friendly smile. The following Context-Free Grammar (CFG) is provided: S → NP VP VP → V NP V NP PP NP → Pronoun ProperNoun Det N Det Adj N PP → P NP Det → the a Adj → young friendly	10	3	3



VIT

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING CONTINUOUS ASSESSMENT TEST - II WINTER SEMESTER 2025-2026

SLOT:C1+TC1

	<p>V → met greeted N → boy park smile ProperNoun → Ravi Pronoun → he him P → near with</p> <p>Construct the parse trees for the sentences using the given grammar and apply Hobbs' Algorithm to resolve the anaphors "he" and "him." Identify the possible antecedents and demonstrate the step-by-step traversal of the parse tree used by the algorithm to determine the most likely antecedents. Finally, rewrite the sentences by replacing the pronouns with their resolved antecedents.</p>																											
4	<p>Consider the following sentences:</p> <p>Sentence A: "The tiger chased the deer." Sentence B: "The lion hunted the deer." Sentence C: "The tiger hunted the deer." The vocabulary is: {tiger, lion, deer, chased, hunted} The sentences are represented using Term frequency (TF) vectors:</p> <table border="1" data-bbox="287 918 1021 1153"> <thead> <tr> <th>Term</th> <th>Sentence A</th> <th>Sentence B</th> <th>Sentence C</th> </tr> </thead> <tbody> <tr> <td>tiger</td> <td>1</td> <td>0</td> <td>1</td> </tr> <tr> <td>lion</td> <td>0</td> <td>1</td> <td>0</td> </tr> <tr> <td>deer</td> <td>1</td> <td>1</td> <td>1</td> </tr> <tr> <td>chased</td> <td>1</td> <td>0</td> <td>0</td> </tr> <tr> <td>hunted</td> <td>0</td> <td>1</td> <td>1</td> </tr> </tbody> </table> <p>a) Compute the Inverse Document Frequency (IDF) for each Term. b) Compute the TF-IDF vectors for Sentence A, Sentence B, and Sentence C. c) Using the TF-IDF vectors, compute the cosine similarity between: i) Sentence A and Sentence B ii) Sentence A and Sentence C d) Interpret the results. Which pair of sentences is more similar and why?</p>	Term	Sentence A	Sentence B	Sentence C	tiger	1	0	1	lion	0	1	0	deer	1	1	1	chased	1	0	0	hunted	0	1	1	10	4	5
Term	Sentence A	Sentence B	Sentence C																									
tiger	1	0	1																									
lion	0	1	0																									
deer	1	1	1																									
chased	1	0	0																									
hunted	0	1	1																									
5.	<p>Consider the following training corpus:</p> <p><s> birds build small nests </s> <s> birds eat small insects </s> <s> cats chase small mice </s> <s> dogs chase small cats </s> <s> birds watch small insects </s></p> <p>Test sentence: <s> birds chase small mice </s></p> <p>(a) Compute the Bigram and Trigram probabilities of the test sentence using Laplace smoothing.</p> <p>(b) Compute the perplexity of the test sentence using both the Bigram and Trigram models after applying Laplace smoothing.</p> <p>(c) Compare the results and justify which model performs better for this case.</p> <p style="text-align: right; margin-right: 100px;"><i>3+4+1+4 = 12</i></p>	10	4	5																								
