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

Reg. Number:	

Continuous Assessment Test (CAT) – II October 2025

Programme	:	**B.Tech (ECM/CSE)**	Semester	:	**FS 2025-26**
Course Code & Course Title	:	**BCSE409L & Natural Language Processing**	Class Number	:	**CH2025260101740 CH2025260100534**
Faculty	:	**Dr.G.Sudhakaran Dr.R.Krithiga**	Slot	:	**D1+TD1**
Duration	:	**90 Mins**	Max. Mark		**50**

General Instructions: < Use this space to provide additional information such as graph sheet, data book etc.>

- Write only your registration number on the question paper in the box provided and do not write other information
- Use statistical tables supplied from the exam cell as necessary
- Use graph sheets supplied from the exam cell as necessary
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks	CO	BT Level
1.		Consider the following sentences: a)**Sentence 1**: The bright star and the dim moon. [5 Marks] b) **Sentence 2**: A dim moon under the sky. [5 Marks] Use the Probabilistic Context-Free Grammar (PCFG) algorithm to calculate the probability of generating each sentence based on the following grammar and probabilities. **PCFG Rules and Probabilities:** S → NP VP | NP NP → Det (Adj)* N (PP)* | Det (Adj)* N (and NP)* | Quantifier NP VP → V NP Det → the | a | an Adj → bright | dim | large | small N → star | moon | sky | cloud | planet PP → P NP P → under | above | near V → shines | glows Quantifier → each | any | all Probabilities: S → NP VP [0.4]; S → NP [0.6] NP → Det (Adj)* N (PP)* [0.55]; NP → Det (Adj)* N (and NP)* [0.35]; NP → Quantifier NP [0.10] VP → V NP [1.0] Det → the [0.5]; Det → a [0.3]; Det → an [0.2] Adj → bright [0.3]; Adj → dim [0.25]; Adj → large [0.25]; Adj → small [0.2] N → star [0.25]; N → moon [0.2]; N → sky [0.2]; N → cloud [0.2]; N → planet [0.15] PP → P NP [1.0] P → under [0.4]; P → above [0.3]; P → near [0.3] V → shines [0.6]; V → glows [0.4] Quantifier → each [0.3]; Quantifier → any [0.4]; Quantifier → all [0.3]	10	3	2
2.		Illustrate how the Arc-Standard Dependency Parser constructs a dependency tree for the sentence: **"Rising inflation worried financial analysts"**, following the oracle	10	3	3

<table>
<tr><td></td><td></td><td>stack, buffer, and arcs, and provide the final dependency tree.</td><td></td><td></td><td></td></tr>
<tr><td>3.</td><td>a.

b.</td><td>Using the Term Frequency–Inverse Document Frequency (TF-IDF) technique, calculate the significance of terms in the following corpus. Determine the highest-ranked terms from each document based on their TF-IDF scores. [10 Marks]

Document 1: Data science drives modern innovation
Document 2: Machine learning powers data analysis
Document 3: Innovation is transforming technology rapidly.

Consider the following two sentences with 4-dimensional word vectors, denoted as X1 and X2: [5 Marks]
X1: "The book is old", vector values = [0.2, 0.3, 0.4, 0.1]
X2: "The table is new", vector values = [0.1, 0.4, 0.2, 0.3]
Evaluate the cosine similarity between X1 and X2.
Determine whether the sentences X1 and X2 are similar or dissimilar.</td><td>15</td><td>4</td><td>2</td></tr>
<tr><td>4.</td><td></td><td>Using the training data below, describe how a Naïve Bayes classifier can be trained for Word Sense Disambiguation (WSD). What role do context words play in classifying the sense of "crane"? [5 Marks]
<table><tr><th>Sentences</th><th>Sense</th><th>Key Context Words</th></tr><tr><td>"The crane lifted heavy steel beams at the site."</td><td>Machine</td><td>Lifted, steel, site</td></tr><tr><td>"The construction workers operated a crane."</td><td>Machine</td><td>Construction, operated, machine</td></tr><tr><td>"The tall crane was visible from a distance."</td><td>Machine</td><td>Tall, construction, heavy</td></tr><tr><td>"The crane spread its wings and flew gracefully."</td><td>Bird</td><td>Wings, flew, gracefully</td></tr><tr><td>"We saw a crane standing near the pond."</td><td>Bird</td><td>Saw, pond, bird</td></tr><tr><td>"The crane migrated south during winter."</td><td>Bird</td><td>Migrated, south, winter</td></tr></table>Classify the following sentences using the trained model:
a) "The crane perched near the riverbank at dawn." [5 Marks]
b) "The workers used a crane to move the concrete blocks." [5 Marks]</td><td>15</td><td>4</td><td>3</td></tr>
<tr><td colspan="6">**********All the best **********</td></tr>
</table>