

BACHELOR OF TECHNOLOGY (CBCS) (2021-COURSE)
B. Tech. Sem - VI Computer Science & Engineering : WINTER: 2025
SUBJECT: NATURAL LANGUAGE PROCESSING

Day : Monday
Date : 24/11/2025

W-25597-2025

Time : 02:30 PM-05:30 PM
Max. Marks : 60

N.B.

- 1) All questions are **COMPULSORY**.
- 2) Figures to the right indicate **FULL** marks.
- 3) Use of non-programmable calculator is allowed.

Q.1 How underflow problem can be solved using N-gram model. Explain with suitable example. (10)

OR

How zero probability problem can be solved using tri-gram model consider below example. Given corpus:

- <S> I am Henery </S>
<S> I like college </S>
<S> Do Henery like college </S>
<S> Henery I am </S>
<S> Do I like Henery </S>
<S> Do I like college </S>
<S> I do like Henery </S>

Identify which sentence is better:

- 1) <S> like college </S>
- 2) <S> Do I like Henery </S>

Q.2 Describe the various common techniques used for extracting information from the text. (10)

OR

Define concept of vector space model. Also solve below example to find cosine similarity of document.

- $D1 = (0.4, 0.7, 0.2, 0.6)$
 $D2 = (0.8, 0.4, 0.3, 0.7)$
 $Q = (1.5, 1.0, 0, 1.2)$

Q.3 Describe classification of POS tagger in detail. (10)

OR

Discuss in detail architecture of Speech Recognition.

P.T.O.

- Q.4 Use CKY algorithm for PCFG to find most probable parse tree.
Sentence: "The flight includes a meal."
CFG rules with probabilities. (10)
- S → NP VP [0.8]
NP → Det N [0.3]
VP → V NP [0.20]
VP → include [0.05]
Det → the [0.4]
Det → a [0.4]
N → meal [0.001]
N → flight [0.02]

OR

- Define concept of parsing. Also, explain the concept of Shift Reduce Parser.
Solve the below example using Shift Reduce Parser.
- | | |
|---------|--------------|
| Grammar | Input String |
| E → 2E2 | 32423 |
| E → 3E3 | |
| E → 4 | |

- Q.5 Discuss the concept of non-negative matrix factorization with suitable (10)
example.

OR

- Apply concept of singular value decomposition on below matrix.

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$$

- Q.6 Describe topics TreeBank, ProBank, WordNet, and VerbNet. (10)

OR

- Write in detail about NLTK library of Python.

* * * * *