BACHELOR OF TECHNOLOGY (CBCS) (2020 COURSE)
Computer Science & Engineering
B.Tech.Sem - VI :SUMMER : 2023
SUBJECT : NATURAL LANGUAGE PROCESSING

Day : Friday
Time : 02:30 PM-05:30 PM

Date : 26-05-2023
S-24330-2023
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 zero probability problem can be solved using trigram model. Consider below example. **(10)**

Given corpus:
<S> I am Henry</S>
<S> I like college </S>
<S> Do Henry like college </S>
<S> Henry I am </S>
<S> Do I like Henry </S>
<S> Do I like college </S>
<S> I do like Henry </S>
Identify which sentence is better using zero Probability Resolution.

1) <S> like college </S>
2) <S> Do I like Henry </S>

OR

Define NLP? Also explain NLP stages in detail with a neat diagram. **(10)**

Q.2 Define Tokenization. Also explain various methods of Tokenization. **(10)**

OR

Define concept of vector space model. Also solve below example to find cosine similarity of document **(10)**
D_1= (0.5, 0.8, 0.3, 0.6)
D_2= (0.9, 0.4, 0.2, 0.7)
Q= (1.5, 1.0, 0, 0)

Q.3 Define POS in NLP. Also explain categories of POS with suitable example. **(10)**

OR

How ambiguity in POS tag is resolved using HMM veterbi algorithm. Solve below example. **(10)**

Training Corpus with POS tags:

<S>	book	a	Car	</S>			
	V	Det	N				
<S>	dark	the	Car	</S>			
	V	Det	N				
<S>	the	Book	is	in	the	Car	</S>
	Det	N	V	prep	Det	N	
<S>	the	Car	is	in	a	Park	</S>
	Det	N	V	prep	Det	N	

Test Data

<S>	the	Park	is	a	Book	</S>
	Det	N	V	Det	N	
		V			V	

Q.4 Define context free grammar? Derive top down and bottom up parser for the given sentence. **(10)**

"The angry bear chased the frightened little squirrel".

CFG Rules-

S → NP VP
NP → Det Nom
VP → V NP
Nom → Adj Nom | N
Det → the
Adj → little | angry | frightened
M → squirrel | bear
V → chased

OR

Use cky algorithm for PCFG to find most probable parse tree. **(10)**

Sentence: "Book the dinner flight."

CFG Rules with probabilities are:-

S → VP [0.05]
VP → V NP [0.030]
NP → Det N [0.020]
N → NN [0.020]
V → book [0.30]
Det→the [0.60]
N → dinner [0.10]
N → flight [0.30]

Q.5 Apply concept of singular value decomposition on below matrix. **(10)**

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

OR

Discuss various topic modelling strategies in detail. **(10)**

Q.6 Write a short note on: **(10)**

a) TreeBank
b) PropBank
c) WordNet
d) VerbNet

OR

Write note on OWL, NLTK and ontologies. **(10)**

* * * * *