

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI  
(END SEMESTER EXAMINATION)**

**CLASS: BCA  
BRANCH: BCA**

**SEMESTER : V  
SESSION : MO/2024**

**SUBJECT: CA337 NATURAL LANGUAGE PROCESSING**

**TIME: 3 Hours**

**FULL MARKS: 50**

**INSTRUCTIONS:**

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
  2. Attempt all questions.
  3. The missing data, if any, may be assumed suitably.
  4. Before attempting the question paper, be sure that you have got the correct question paper.
  5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
- 

		CO	BL
Q.1(a)	Define the terms Phonology and Morphology related to NLP with example.	[5] 1	2
Q.1(b)	Describe the types of spelling errors commonly encountered in natural language text. Provide examples of each type.	[5] 3	4
Q.2(a)	Explain the concept of N-gram language models. How does an N-gram model estimate the probability of a sequence of words?	[5] 2	3
Q.2(b)	Here are a few other useful probabilities: $P(\text{Children} \mid <s>) = 0.25$ , $P(\text{want} \mid \text{Children}) = 0.33$ , $P(\text{to} \mid \text{want}) = 0.0011$ , $P(\text{play} \mid \text{to}) = 0.6$ , $P(\text{footwal} \mid \text{play}) = 0.5$ , $P(</s> \mid \text{footwal}) = 0.68$ . Then compute the probability of sentences like "Children want to play footwal"	[5] 3	4
Q.3(a)	What is Perplexity in the context of language modeling? Provide an example of how it can be calculated for a simple sentence.	[5] 3	3
Q.3(b)	For the sentence "data science is fun", if the unigram probabilities are: $P(\text{data}) = 0.2$ , $P(\text{science}) = 0.25$ , $P(\text{is}) = 0.3$ , $P(\text{fun}) = 0.25$ , find the perplexity of the sentence.	[5] 5	5
Q.4(a)	Explain the importance of POS tagging in Natural Language Processing. Describe how POS tagging helps in understanding the structure of sentences in NLP applications.	[5] 3	3
Q.4(b)	Define Named Entity Recognition (NER) and explain its significance in Natural Language Processing. What are some common types of named entities that NER systems are designed to recognize?	[5] 4	4
Q.5(a)	Explain the concept of vector semantics in NLP. For Given Context Free Grammar: $S \rightarrow NP VP$ $NP \rightarrow \text{Pronoun} \mid \text{NOM} \mid \text{Det NOM}$ $\text{NOM} \rightarrow \text{Noun} \mid \text{Noun NOM}$ $VP \rightarrow \text{Verb NP}$	[5] 2	3
Q.5(b)	-- Lexicon -- (parts of speech) $\text{Pronoun} \rightarrow I \mid they$ $\text{Noun} \rightarrow \text{flight} \mid \text{morning} \mid \text{evening}$ $\text{Verb} \rightarrow \text{prefer}$ $\text{Det} \rightarrow a \mid the \mid that$ Derive derivation for sentence "I Prefer a morning flight" and also draw parse tree.	[5] 5	5

:::21/11/2024:::M