



USN

--	--	--	--	--	--	--	--	--	--

18CS743

## Seventh Semester B.E. Degree Examination, Feb./Mar. 2022 Natural Language Processing

Time: 3 hrs.

Max. Marks: 100

*Note: Answer any FIVE full questions, choosing ONE full question from each module.*

### Module-1

- 1 a. Define NLP. What makes NLP difficult. (10 Marks)  
 b. Explain Karaka theory of Paninian grammar. Identify different Karaka's in the following sentence in Hindi language. "Maan bachchi ko aangan mein haath se rotii khilaatii hei". (10 Marks)

OR

- 2 a. Explain transformational grammar with example. (10 Marks)  
 b. Write the C-structure and F-structure for the following sentence "she saw stars". Consider the CFG rule.  
 $S \rightarrow NPVP$   
 $VP \rightarrow V \{NP\} \{NP\} PP^* \{S'\}$   
 $PP \rightarrow P NP$   
 $NP \rightarrow Det N \{PP\}$   
 $S' \rightarrow comp S$  (10 Marks)

### Module-2

- 3 a. What is morphological passing? Explain 2-level morphological model with an example. (10 Marks)  
 b. Write and explain an algorithm for minimum edit distance spelling correction. Apply the same to find the minimum edit distance between words 'PEACEFUL' and 'PAECFLU'. (10 Marks)

OR

- 4 a. Explain Levensthein minimum edit distance algorithm. (10 Marks)  
 b. Write a note on different phrase level constructs with suitable example for each phrase. (10 Marks)

### Module-3

- 5 a. With neat diagram, explain the infact framework architecture over view. (10 Marks)  
 b. Write a neat diagram, explain the architecture used in the task of learning to annotale cases with knowledge roles. (10 Marks)

OR

- 6 a. Explain the strategies used in active learning approach for acquiring labels using committee based classification scheme. (10 Marks)  
 b. Write a short note on:  
 i) The shortest path hypothesis  
 ii) Learning with dependency path. (10 Marks)

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.  
 2. Any revealing of identification, appeal to evaluator and /or equations written eg. 42+8 = 50, will be treated as malpractice.

**Module-4**

- 7 a. Explain SVM (Support Vector Machine) learning method in sequence model estimation. (10 Marks)  
b. Explain in detail the high-level representation approaches in text mining. (10 Marks)

**OR**

- 8 a. Explain the functioning of word matching feedback system used in ISTART. (10 Marks)  
b. Write a note on various approaches to analyzing texts. (10 Marks)

**Module-5**

- 9 a. Explain design feature of IR with a neat diagram. (10 Marks)  
b. Explain classical information retrieval models. (10 Marks)

**OR**

- 10 a. With a suitable example explain cluster based Information Retrieval (IR) modeling. (10 Marks)  
b. Write short note on: i) Word Net ii) Frame Net. (10 Marks)

\* \* \* \* \*