CBCS SCHEME

USN | | | | | | | | | | |

15CS741

Seventh Semester B.E. Degree Examination, Dec.2019/Jan.2020
Natural Language Processing

Time: 3 hrs. Max. Marks: 80

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

Module-1

1 a. Define NLP. What makes NLP difficult? **(08 Marks)**

b. Explain transformational grammar with example. **(08 Marks)**

OR

2 a. Write a note on the following with example:
(i) $\overline{X}$ theory (ii) Theta theory **(04 Marks)**

b. Explain Lexical Functional Grammar [LFG]. **(06 Marks)**

c. Write the c-structure and f-structure for the following sentence "she saw stars". Consider the CFG rules.
$S \rightarrow NP\ VP$
$VP \rightarrow V\ \{NP\}\ \{NP\}\ PP^*\ \{S'\}$
$PP \rightarrow P\ NP$
$NP \rightarrow Det\ N\ \{PP\}$
$S' \rightarrow comp\ S$ **(06 Marks)**

Module-2

3 a. What is morphological parsing? Explain 2-level morphological model with an example. **(08 Marks)**

b. Comment on the validity of the following statements:
(i) Rule-based taggers are non-deterministic.
(ii) Stochastic taggers are language dependent **(04 Marks)**

c. Construct the parse tree for the sentence:"The girl plucked the flower with a long stick" Discuss the ambiguity arises from the parse tree. **(04 Marks)**

OR

4 a. Explain Levensthein minimum edit distance algorithm. **(08 Marks)**

b. Compute minimum edit distance between paceful and peaceful. **(08 Marks)**

Module-3

5 a. Explain shortest dependency path hypothesis. Show various shortest dependency path among the relations in the "Jellisc created an atmosphere of terror in the camp by killing abusing and threatening the detainees". **(08 Marks)**

b. Explain how the relation patterns can be captured using string kernel. **(08 Marks)**

OR

6 a. Compute the common features between x and y, where x = "his actions in Brocko" and y = "his arrival in Beijing" **(06 Marks)**

b. Explain the strategies used in active learning approach for acquiring labels using committee based classification scheme. **(10 Marks)**

Module-4

7 a. Explain the semantically guided model for effective text mining. **(08 Marks)**

b. Explain in detail the high-level representation approaches in text mining. **(08 Marks)**

OR

8 a. Define: (i) cohesion (ii) coh-metrix (iii) LSI **(06 Marks)**

b. Write a note on various approaches to analyzing texts. **(10 Marks)**

Module-5

9 a. Explain design features of IR with a neat diagram. **(08 Marks)**

b. How stemming affects the performance of IR systems? **(04 Marks)**

c. "Stop words elimination may be harmful". Justify. **(04 Marks)**

OR

10 a. A user submitted a query to an IR system. Out of the 1st 15 documents returned by the system, those ranked 1, 2, 5, 8, 12 were relevant. Compute non-interpolated average precision for this retrieval. Assume there are six relevant documents. **(06 Marks)**

b. Explain word net. List the applications of word net. **(06 Marks)**

c. Write the hypernym chain for "RIVER" extracted from wordnet 2.0. **(04 Marks)**

* * * * *