



CBCS SCHEME

USN

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

18CS743

Seventh Semester B.E. Degree Examination, July/August 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. What is NLP? List and explain applications of NLP. (10 Marks)
- b. Briefly explain various levels of natural language processing. (10 Marks)

OR

- 2 a. Describe C-structure and F-structure in LFG. Write C-structure and F-structure for the sentence 'She saw stars' using the CFG rules as below.
 $S \rightarrow NP VP$
 $VP \rightarrow V\{NP\}\{NP\}PP^*\{S'\}$
 $PP \rightarrow PNP$
 $NP \rightarrow Det N\{PP\}$
 $S' \rightarrow Comps$ (10 Marks)
- b. Describe Paninian Framework for Indian languages. Explain Layered representation of Paninian Grammar and Karaka theory. (10 Marks)

Module-2

- 3 a. Describe DFA and NFA. Mention the properties of Finite automation. (06 Marks)
- b. Define Morphology. Explain stem and affix classes of Morphemes with example. (04 Marks)
- c. Explain two step morphological parser with a neat diagram. (10 Marks)

OR

- 4 a. Write minimum Edit Distance Algorithm. (05 Marks)
- b. What is POS tagging? Explain Rule based Tagger and Hybrid Taggers. (10 Marks)
- c. Write CYK syntactic parsing algorithm. (05 Marks)

Module-3

- 5 a. Explain two major approaches of relation extraction. (05 Marks)
- b. Describe the patterns used to extract relationship between two entity with a given sentence. (05 Marks)
- c. Write short notes on the following:
(i) Globalsecurity.org
(ii) InFact system (10 Marks)

OR

- 6 a. Explain Active Learning Strategy steps for acquiring labels from a human annotator. (07 Marks)
- b. What are the different steps used to achieve the goal of case annotation? (07 Marks)
- c. Briefly describe the following :
(i) Domain knowledge.
(ii) Domain concept
(iii) Knowledge Roles. (06 Marks)

Module-4

- 7 a. What is iSTART? List the reading strategies used by iSTART. (06 Marks)
 b. Explain Literal word matching and soundex word matching approaches. (06 Marks)
 c. Briefly describe LSA feedback systems. Mention four benchmarks used by LSA to Assess the level of an explanation. (08 Marks)

OR

- 8 a. Briefly describe the evolutionary model for knowledge discovery from texts with a neat diagram. (10 Marks)
 b. Describe the following:
 (i) Topic models
 (ii) Cohesion and Cohesion matrix. (10 Marks)

Module-5

- 9 a. Explain basic information Retrieval process with a neat diagram. (06 Marks)
 b. Describe the following approaches used in I.R. :
 (i) Indexing.
 (ii) Stop words elimination.
 (iii) Stemming. (09 Marks)
 c. State and explain the importance of Zipf law related to words distribution in NLP. (05 Marks)

OR

- 10 a. Explain the following alternative-IR models:
 (i) Cluster model.
 (ii) Fuzzy model. (10 Marks)
 b. Describe cosine and Jaccard Similarity measures used for IR. (06 Marks)
 c. Describe tf-idf term weightage approach used in IR. (04 Marks)

CMRIT LIBRARY
 BANGALORE - 560 037
