

CBCS SCHEME

USN

21AI643

Sixth Semester B.E./B.Tech. Degree Examination, Dec.2024/Jan.2025

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. List the various components of Government and Binding. Describe diagrammatically with examples the different types of representation of Government and Binding. (12 Marks)
- b. What does Karaka Theory deal with? Give a brief description of various Karakas. (08 Marks)

OR

- 2 a. Explain with relevant formulae how Add-one smoothing is carried out. Illustrate with an example how good turing discounting overcomes the drawback of Add one smoothing. (12 Marks)
- b. List any two issues in paninian Grammar processing differences encountered while processing Indian languages as opposed to European languages. (02 Marks)
- c. List the differences encountered while processing Indian languages as opposed to European languages. (06 Marks)

Module-2

- 3 a. Describe the information sources needed by a Morphological parser. (06 Marks)
- b. Write the algorithm of CYK parser. Parse the following sentence 'We buy jam with bread' using the following rules :

$S \rightarrow NP VP$ $VP \rightarrow VP PP$

$VP \rightarrow V NP$ $V \rightarrow \text{buy}$

$NP \rightarrow NP PP$ $PP \rightarrow PNP$

$P \rightarrow \text{with}$ $NP \rightarrow \text{we}$

$NP \rightarrow \text{bread}$ $NP \rightarrow \text{Jam.}$

Describe the manner in which this sentence is passed using CYK parser. (14 Marks)

OR

- 4 a. Describe in detail with examples the three methods of part of speech Tagging. (08 Marks)
- b. Mention the purpose of computing minimum Edit Distance. Compute the minimum edit distance between the words STRANGE and STORAGE. (06 Marks)
- c. Write short notes on feature structures. Give an example for unification of features structures. (06 Marks)

Module-3

- 5 a. Describe the manner in which relations are extracted from text. (08 Marks)
- b. Demonstrate the manner in which a dependency path kernel can be constructed for extracting relations. (12 Marks)

OR

- 6 a. Demonstrate the manner in which text reposts are mined through annotation of knowledge roles. (08 Marks)
- b. How is tagging done for the document and describe the manner in which a parser can be chosen. (12 Marks)

Module-4

- 7 Illustrate the manner in which a combination of probabilistic classification and finite state sequence modeling can be used to separate a document. Describe the process step by step. (20 Marks)

OR

- 8 a. Describe the manner in which cohesion of text structures can be measured using Latent semantic analysis. (10 Marks)
b. Describe the role of cohesion matrix in identifying textual signatures. (10 Marks)

Module-5

- 9 a. Describe in detail the design features of an informational retrieval system. How does Zip's law play a role in this retrieval? (12 Marks)
b. Consider a document represented by 3 terms namely tree, leaf and flower having raw and frequencies 6, 3 and 2 respectively. In a collection of 700 documents, 150 documents contain the term tree, 230 documents contain the term lead and 280 documents contain the term flower. Compute the weight (tf – idf product). (08 Marks)

CMRIT LIBRARY
BANGALORE - 560 037 **OR**

- 10 a. Write short notes on the three different non-classical models of information retrieval. Which of these models use a structure called infon? (10 Marks)
b. Illustrate with an example how precision and Recall are computed. (07 Marks)
c. Give the formulae for the terms : F- measure, E-measure and Normalized Recall. (03 Marks)

* * * * *