

# CBCS SCHEME

USN

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

18AI641

**Sixth Semester B.E. Degree Examination, June/July 2023**

## **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? Explain the applications of Natural Language processing. (06 Marks)
- b. Illustrate with suitable examples the different levels of NLP. (08 Marks)
- c. Explain the role of transformational rules in transformational grammar with the help of an example. (06 Marks)

**OR**

- 2 a. List and explain the challenges of Natural Language Processing. (06 Marks)
- b. Explain Statistical Language Model and find the probability of the test sentence P (they play in a big garden) in the following training set using bi-gram model.  
<S> There is a big garden  
Children play in the garden  
They play inside beautiful garden <S> (06 Marks)
- c. List the problems associated with n-gram model, explain how these problems are handled. (08 Marks)

### Module-2

- 3 a. Explain the working of two-steps morphological parser. Write a simple Finite State Transducer (FST) for mapping English nouns. (08 Marks)
- b. Illustrate parts of Speech Tagging and explain different categories of POS tagging. (06 Marks)
- c. Explain the Minimum Edit Distance algorithm and compute the minimum edit distance between EXECUTION and INTENTION. (06 Marks)

**OR**

- 4 a. Design CYK algorithm. Tabulated the sequence of states created by CYK algorithm while parsing "A pilot likes flying planes".  
Consider the following simplified grammar in CNF  
S → NP VP      NN → Pilot      VBG → flying  
NP → DT NN      NNS → planes  
NP → JJ NNS      JJ → flying  
VP → VBG NNS      DT → a  
VP → VBZ NP      VBZ → likes (08 Marks)
- b. Explain top-down parsing and bottom up parsing with an example. (08 Marks)
- c. List out the disadvantages of Probabilistic Context Free Grammar (PCFG). (04 Marks)

**Module-3**

- 5 a. Explain the four patterns used to extract relationship between two entries with an example for each. (08 Marks)
- b. Explain dependency path Kernel for Relation Extraction. (08 Marks)
- c. Discuss the knowledge rules for below sentences with the same domain concepts.
- i) The calculated insulating resistance values lay in the safe operating area
  - ii) Compared to the last examination, lower values for the insulating resistance were ascertained due to dirtiness at the surface. (04 Marks)

**OR**

- 6 a. With a neat diagram, explain the architecture used in the task of learning to annotate cases with knowledge roles. (10 Marks)
- b. Explain functional overview of Infact system with a neat diagram. (10 Marks)

**Module-4**

- 7 a. Explain the functioning of word Matching Feedback Systems. (08 Marks)
- b. Discuss iSTART System and their modules. (08 Marks)
- c. Illustrate Topic Models (TM) Feedback System. (04 Marks)

**OR**

- 8 a. Define :  
i) Cohesion ii) Coh – Metrix iii) Latent Semantic Analysis. (10 Marks)
- b. Write a note on various approaches to analyzing texts. (10 Marks)

**Module-5**

- 9 a. Explain design features of information retrieval systems with a neat diagram. (08 Marks)
- b. Define term weighing. Consider a document represented by the 3 terms {tornado, swirl, wind} with the raw tf 4, 1 and 1 respectively. In a collection of 100 documents 15 documents contains the terms tornado, 20 contains, swirl and 10 contains wind. Find the idf and term weight of a 3 terms. (06 Marks)
- c. Explain the benefits of eliminating stop words. Give examples in which stop word eliminating may be harmful. (06 Marks)

**OR**

- 10 a. List different IR models. Explain classical Information Retrieval models. (10 Marks)
- b. Explain :  
i) Wordnet ii) Framenet. (10 Marks)

\* \* \* \* \*