./Sem-VII/COMP + CSE/C-Scheme/NOV-2025/18-11-25

(3 hours) **[Marks: 80]**

KONKAN GYANPEETH COLLEGE OF ENGINEERING KARJAT Dist. Raigad

NOTE:

1. Question No 1 is compulsory
2. Attempt any three questions from remaining.
3. Assume suitable data if necessary and state the same.

Q.1 Solve all questions below 20

a) What are the limitations of Hidden Markov Model (HMM) and MaxEnt Model for POS Tagging.

b) What is word level analysis? Explain the following terms with examples: Tokenization, Stemming, Lemmatization.

c) Explain the term Coreference resolution in NLP. Explain agreement constraints with examples.

d) Explain rules-based machine translational system.

Q.2

a) Design FST for regular and plural nouns. 10

b) Explain use of Probabilistic Context Free Grammar (PCFG) in NLP 10

Q.3

a) Explain n-gram language model with formulas. What is the role of perplexity in evaluating n-gram language model? 10

b) Describe lexical semantic analysis using examples. 10

Q.4

a) What is Word Sense Disambiguation (WDS) ? Explain Semi supervised algorithm for WSD. 10

b) Define POS Tagging. Explain why it is a necessary pre-processing step for higher-level NLP tasks like Named Entity Recognition (NER). Using the common Penn Treebank tagset, assign the appropriate POS tag to each word in the following sentence. "The fast **train** from **Dadar will** arrive at platform **seven** shortly." Briefly explain the reasoning for the tag assigned to the highlighted words. 10

Q.5

a) Explain Hobbs algorithm and Cantering algorithm for anaphora resolution. 10

b) Compare the structure, contents, and utility of two major language dictionaries: WordNet and Babelnet in NLP 10

Q.6

a) What is text summarization? 10

b) How conditional random fields (CRF) are used for sequence labelling task 10

****

muquestionpapers.com