Paper / Subject Code: 42175 / NATURAL LANGUAGE PROCESSING (DLOC – III)

BE Sem VII (C Scheme) R-2019 "Computer" Dec '2022

Time: 3 Hours **Max. Marks: 80**

N.B. (1) Question No. 1 is compulsory
(2) Assume suitable data if necessary
(3)Attempt any three questions from remaining questions

Q.1	Any Four	**20[M]**
a	Differentiate between Syntactic ambiguity and Lexical Ambiguity.	**[5M]**
b	Define affixes. Explain the types of affixes.	**[5M]**
c	Describe open class words and closed class words in English with examples.	**[5M]**
d	What is rule base machine translation?	**[5M]**
e	Explain with suitable example following relationships between word meanings. Homonymy, Polysemy, Synonymy, Antonymy	**[5M]**
f	Explain perplexity of any language model.	**[5M]**
Q.2 a)	Explain the role of FSA in morphological analysis?	
Q.2 b)	Explain Different stage involved in NLP process with suitable example.	**[10M]**
Q.3 a)	Consider the following corpus <s> I tell you to sleep and rest </s> <s> I would like to sleep for an hour </s> <s> Sleep helps one to relax </s> List all possible bigrams. Compute conditional probabilities and predict the next ord for the word "to".	**[5M]**
Q.3 b)	Explain Yarowsky bootstrapping approach of semi supervised learning	**[5M]**
Q.3 c)	What is POS tagging? Discuss various challenges faced by POS tagging.	**[10M]**
Q.4 a)	What are the limitations of Hidden Markov Model?	**[5M]**
Q.4 b)	Explain the different steps in text processing for Information Retrieval	**[5M]**
Q.4 c)	Compare top-down and bottom-up approach of parsing with example.	**[10M]**
Q.5 a)	What do you mean by word sense disambiguation (WSD)? Discuss dictionary based approach for WSD.	**[10M]**
Q.5 b)	Explain Hobbs algorithm for pronoun resolution.	**[10M]**
Q.6 a)	Explain Text summarization in detail.	**[10M]**
Q.6 b)	Explain Porter Stemming algorithm in detail	**[10M]**

****