K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2025-26
(B. Tech.) Program: Scheme IIB
Regular Examination: LY Semester/VIII
Course Code: _EXDLC8031 and Course Name: Natural language Processing
Date of Exam: 253/05/2026 Duration: 02.5 Hours Max. Marks: 60

Instructions:
(1)All questions are compulsory.
(2)Draw neat diagrams wherever applicable.
(3)Assume suitable data, if necessary.

Q. No.	Question	Max. Marks	CO	BT level
Q 1	Solve any **two** questions out of three: (05 marks each)	10		
a)	Apply the different stages of Natural Language Processing (NLP) to the sentence: 'Riya booked the train ticket using a mobile app,' and explain how each stage contributes to understanding it		6	A
b)	Why is it important to preprocess text data in Natural language? Explain in detail the steps of pre-processing		2	U
c)	What is Pragmatics?		5	U
Q 2	Solve any **two** questions out of three: (05 marks each)	10		
a)	What is phrase structure? Show the phrase structure (by a dendrogram or parse tree) for "Colorless green ideas sleep furiously".		3	U
b)	Differentiate Stemming and Lemmatization		2	U
c)	What is Word Sense Disambiguation?		4	U
Q.3	Solve any **two** questions out of three. (10 marks each)	20		
a)	Analyze the concept of measure *m* used in the Porter Stemmer algorithm. Break down each of the following words into consonant–vowel sequences, determine their measure values i) playing ii) trouble c) cats d) meeting		2	An
b)	Write a short note on N-Gram model and illustrate it for "Either my way or no way".		2	U
c)	List and explain syntactic and semantic constraints on co-reference.		5	U

Seat No.:

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2025-26
(B. Tech.) Program: Scheme IIB
Regular Examination: LY Semester/VIII
Course Code: _EXDLC8031 and Course Name: Natural language Processing
Date of Exam: 253/05/2026 Duration: 02.5 Hours Max. Marks: 60

Q.4	Solve any **two** questions out of three. (10 marks each)	20		
a)	Illustrate the following terms with suitable examples: Hyponymy, Hypernym, Meronym, Holonym.		4	U
b)	A Hidden Markov Model (HMM) based POS tagger is used to tag the sentence: "Dogs bark" Possible tags: N = Noun, V = Verb The following probabilities are given: Initial Probabilities: P(N)=0.6, P(V)=0.4 Transitional probabilities: P(N --> N) = 0.3, P(N --> V) = 0.7 P(V--> N) = 0.5, P(V --> V) = 0.5 Emission Probabilities: P(Dogs/N) =0.8, P(Dogs/V) = 0.2 P(bark/N) =0.4, P(bark/V) = 0.6 Analyze the sentence using the HMM model and determine the most probable POS tag sequence. Show the calculations.		3	A
c)	Write short notes on (i) Sentiment analysis (ii) Text Summarization		6	U

****

Seat No.: