

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

Nov – Dec 2025 B. Tech Program: AI-DS Scheme IIB		
Regular Examination: LY		Semester: VII
Course Code: AIC 702	Course Name: Natural Language Processing	
Date of Exam: 29/11/2025	Duration: 2.5 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)	List out and explain any 2 applications of NLP in health care industry		6	U
b)	What is the difference between syntactic and semantic information?		5	U
c)	What is the role of regular expression and automata in the development of NLP system?		2	U
Q 2	Solve any two questions out of three: (05 marks each)	10		
a)	Draw the block diagram to explain the NLP steps		1	U
b)	What are the problems encountered in PoS tagging in NLP?		3	U
c)	Explain Semantic analysis? Analyze various approaches for semantic analysis.		4	U
Q 3	Solve any two questions out of three. (10 marks each)	20		
a)	With examples, explain the grammatical ambiguities in designing NLP solutions.		1	U
b)	How do you identify and segregate the email messages based on urgency of tasks?		6	U
c)	What are orthographic rules used in FST in Porter stemmer? Explain with suitable examples		2	U
Q.4	Solve any two questions out of three. (10 marks each)	20		
a)	Explain Rule-Based POS tagging and Stochastic POS tagging in detail and apply Rule-Based POS tagging in the given sentences 1. "She promised to stop smoking." 2. "Mohan ran fast like a cheetah to catch the cunning thief."		3	Ap
b)	Consider the following corpus:		2	Ap

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

Nov – Dec 2025 B. Tech Program: AI-DS Scheme IIB		
Regular Examination: LY		Semester: VII
Course Code: AIC 702	Course Name: Natural Language Processing	
Date of Exam: 29/11/2025	Duration: 2.5 Hours	Max. Marks: 60

	<i>“natural language processing is fun . language models are useful . natural language models learn patterns .”</i> Calculate the probability of the sentence using Bigram and Trigram language models: <i>“natural language models are useful”</i>		
c)	Consider translation from your native language to English. Illustrate various challenges involved at Semantic level.	4	Ap
