

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

May-2026		
B.Tech. Program: AIML: HONOURS Scheme: IIB		
Regular Examination: LY- Honours Semester: VIII		
Course Code: HAIMLC801	Course Name: Text, Web and Social Media Analytics	
Date of Exam: 29/05/26	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)	Explain topic analysis in text classification.		CO1	U
b)	Explain L1 Regularization (LASSO).		CO2	U
c)	Discuss Jaccard coefficient?		CO3	U
Q 2	Solve any two questions out of three: (05 marks each)	10		
a)	Explain feature transformation in WUM?		CO4	U
b)	Explain User Behaviour Analysis (UBA).		CO5	U
c)	Discuss lexicon-based sentiment analysis?		CO6	U
Q.3	Solve any two questions out of three. (10 marks each)	20		
a)	Explain N-gram modelling with probability calculations.		CO1	U
b)	Discuss LDA, PLSA, and GMM in document clustering.		CO2	U
c)	Explain Latent Semantic Indexing with SVD process.		CO3	U
Q.4	Solve any two questions out of three. (10 marks each)	20		
a)	Explain association rule mining with Apriori and FP-Growth algorithms.		CO4	U

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

May-2026

B.Tech. Program: AIML: HONOURS Scheme: IIB

Regular Examination: **LY- Honours** Semester: **VIII**

Course Code: **HAIMLC801**

Course Name: **Text, Web and Social Media Analytics**

Date of Exam: 29/05/26

Duration: 2.5 Hours

Max. Marks: 60

b)	Explain the cold-start problem and data sparsity with solutions.	CO5	U
c)	Discuss VADER, SentiWordNet, and AFINN lexicons.	CO6	U
