



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

End Semester Exam

Nov - Dec 2021

Program: B.Tech Electronics and Telecommunication

Examination: LY Semester: VII

Course Code: IUEXDLC7034 and Course Name: Big Data Analytics

Duration: 03 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 six questions out of eight:	12		
i)	Summarize the data types for big data.	2	CO1	R
ii)	For each of the following write its NoSQL data architecture pattern.	2	CO3	An
iii)	List the types of queries fired on stream data.	2	CO5	U
iv)	Draw and explain the architectural diagram for physical organization of compute nodes.	2	CO4	U
v)	Discuss any four HDFS commands in brief.	2	CO2	U
vi)	Explain with example collaborative filtering.	2	CO6	U
vii)	What is NoSQL? Give example of NoSQL systems.	2	CO3	U
viii)	Discuss any two applications of near-neighbor search.	2	CO5	U
Q.2	Solve any four questions out of six.	16		
i)	Explain Column Family store and Document Store with example.	4	CO3	U
ii)	Explain Utility matrix and long tail.	4	CO6	U
iii)	Explain in detail about the challenges of conventional system.	4	CO1	U
iv)	Discuss any two limitations of Hadoop.	4	CO2	Ap
v)	How to compute Page Rank for any web graph.	4	CO5	An
vi)	Describe the operations of "shuffle" and "sort" in the Map reduce framework? Explain with the help of one example.	4	CO4	U
Q.3	Solve any two questions out of three.	16		
i)	Explain various phases of Map-Reduce execution pipeline and role of combiner with suitable example.	8	CO4	Ap
ii)	What do you mean by NoSQL databases? What is the alternative to ACID property in NoSQL databases?	8	CO3	U
iii)	What is a community in a Social Network Graph?	8	CO6	An



	How the Girvan Newman algorithm finds the different Communities in the graph.			
Q.4	Solve any two questions out of three.	16		
i)	Illustrate with an example the application of the Apriori algorithm to a relatively simple case that generalizes to those used in practice. Show how to use the Apriori algorithm to generate frequent item sets and rules and to evaluate and visualize the rules.	8	CO5	C
ii)	How does Hadoop work? Explain in detail with an example.	8	CO2	An
iii)	Show why traditional approach is not appropriate for Big Data solutions.	8	CO1	R