May-June 2024				
(B.Tech / M.Tech.) Program: _M.Tech._ Scheme I/II/IIB/III: _II_				
Regular Examination: FY Semester: II				
Course Code: _PCEC202_ and Course Name: _Big Data Analytics_				
Date of Exam: 17/05/2024	Duration: 02.5 Hours		Max. Marks: 60	

Instructions:
(1) All questions are compulsory.
(2) Draw neat diagrams wherever applicable.
(3) Assume suitable data if required, and state it clearly.

		Max. Marks	CO	BT Level
Q-1	**Solve any Six questions out of Eight.**	12		
i)	What is Big Data? Give its importance.	2	CO2	U
ii)	Explain drivers of NoSQL system.	2	CO4	U
iii)	List limitations of Hadoop.	2	CO3	U
iv)	Differentiate between Hive and Pig	2	CO4	U
v)	Explain business intelligence , data science and data analytics.	2	CO1	U
vi)	Why Shuffle and sort phase is important in map reduce?	2	CO4	U
vii)	List Principles of Global Privacy.	2	CO5	U
viii)	Explain challenges in Stream Processing	2	CO6	U
Q-2	**Solve any Four questions out of Six.**	16		
i)	Describe decision support system and its components.	4	CO1	U
ii)	Find communities in the following graph using CPM algorithm	4	CO2	Ap
iii)	Explain Mahout- Hadoop ecosystem component and give its importance.	4	CO3	U
iv)	Explain Spark and SparkQL with example.	4	CO4	U
v)	Discuss anonymization Techniques and its solution	4	CO5	U
vi)	Differentiate between static data, incremental data and data streams.	4	CO6	U
Q-3	**Solve any Two questions out of Three.**	16		
i)	Describe Business Pressures–Responses–Support Model in detail.	8	CO1	U
ii)	Apply Matrix – Matrix Multiplication using MapReduce model to solve the following example 1 2 5 6 3 4 * 7 8	8	CO4	Ap
iii)	Explain collaborative filtering based recommendation system with example.	8	CO2	U

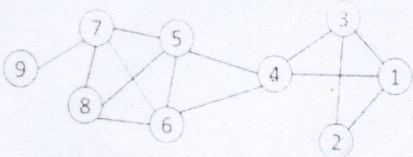| May-June 2024			
(B.Tech / M.Tech.) Program: ____M.Tech._____ Scheme I/II/IIB/III:___II_			
Regular Examination: FY Semester: II			
Course Code: _PCEC202 and Course Name: Big Data Analytics			
Date of Exam:	Duration: 02.5 Hours		Max. Marks: 60

Q-4	Solve any Two questions out of Three.	16		
i)	Describe HDFS architecture in detail with its advantages and limitations	8	CO3	Ap
ii)	What is Data Privacy? Explain different threats and privacy control techniques.	8	CO5	U
iii)	Explain Blooms filter for Stream mining. Let n =12, k=3, S= {17,23}. Build a bloom's filter and check for membership of 13 & 29 • $h_1(x) = x \% 12$ • $h_2(x) = (x+1) \% 12$ • $h_3(x) = (x+3) \% 12$	8	CO6	Ap

****