

(3 Hours)

Max. Marks: 80

- N.B.:** (1) Question No. 1 is compulsory.  
 (2) Attempt any **Three** questions out of remaining **Five** questions.  
 (3) **Figures** to the **right** indicate **full** marks.  
 (4) Assume suitable data if **necessary**.



Q.1 Answer the following: [20]

- A) Explain the role and effect of damping factor (teleportation) in PageRank computation.  
 B) Agility is a NoSQL business driver. Justify.  
 C) Give the updating buckets approach of DGIM algorithm.  
 D) Find Cosine Distance between the d1 and d2 vectors:

Index	1	2	3	4	5	6	7	8	9	10
d1	5	2	1	0	0	0	0	1	3	7
d2	5	2	1	0	0	1	2	2	0	2

Q.2 A) List the different NoSQL data stores. Explain any two with diagram. [10]

Q.2 B) Write steps of Girvan-Newman Algorithm. Explain clustering of Social-Network Graphs using GN algorithm with example? [10]

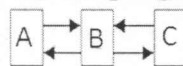
Q.3 A) Explain Flajolet Martin Algorithm with example. [10]

Q.3 B) Distinguish the following: [10]

- i) DBMS and DSMS
- ii) PCY, Multistage and Multihash

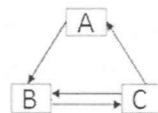
Q.4 A) List Relational-Algebra Operations. Explain any two using MapReduce. [10]

Q.4 B) Compute Efficient PageRank with the damping factor  $d = 0.8$  for web. [10]



Q.5 A) What are different recommender systems. Explain any one with example. [10]

Q.5 B) Define Hub and Authority. Compute Hub and Authority scores for web. [10]



Q.6 Answer the following: [20]

- A. Core Hadoop Components
- B. CURE Algorithm
- C. SON Algorithm and MapReduce
- D. Matrix-Vector Multiplication by MapReduce