

(Time: 3 Hrs)

Marks: 80

N.B. : 1. Question no. 1 is compulsory.

2. Solve any **Three** questions out of remaining **Five** questions.**Q-1 Write short Note on:**

- a) Big Data & its characteristics 5
- b) Distance Measures for Big Data 5
- c) The Map & Reduce Tasks 5
- d) Multistage Frequent Itemset Mining Algorithm 5

- Q-2**
- a) Explain HDFS architecture with diagram. 10
  - b) Explain Column family store and Graph Store NoSQL architectural pattern with examples. 10

- Q-3**
- a) Explain Matrix-Vector multiplication algorithm by MapReduce? 10
  - b) Explain Issues in Data stream query processing? 10

- Q-4**
- a) Explain 10
    - 1) Bloom Filter with the help of an example
    - 2) Steps of HITS algorithm
  - b) Explain structure of web with suitable diagram? 10

- Q-5**
- a) Explain Park-Chen-Yu algorithm? How memory mapping is done in PCY? 10
  - b) Explain CURE algorithm with initialization & completion phase? 10

[Turn Over]

- Q-6** a) How recommendation is done based on properties of product? Explain with suitable example? 10
- b) For following graph, show how the clique percolation method (CPM) find cliques. 10  
Explain with steps?

