

Duration: 3hrs

[Max Marks:80]

- N.B.: (1) Question No 1 is Compulsory.
(2) Attempt any three questions out of the remaining five.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

- 1 Attempt any FOUR [20]**
- a Describe any five characteristics of Big data.
 - b Explain Hamming distance measure with an example.
 - c Define Social Networks and Social Network Mining.
 - d Explain Market – Basket Model with example
 - e Explain CAP Theorem? How it is different from ACID properties.
- 2 a Explain Hadoop Ecosystem with core components. Explain its Physical architecture. State Limitations of Hadoop. [10]**
- b What is MapReduce ? Explain How Map and Reduce Work? What is Shuffling in MapReduce? [10]**
- 3 a Draw the diagram showing the structure of the World Wide Web and Explain the different parts. [10]**
- b Explain the concept of Bloom Filter with the help of an example. [10]**
- 4 a What is NOSQL? What are the business drivers for NoSQL? Discuss any two architectural patterns of NoSQL. [10]**
- b What is a “Community “ in a social Network graoh? Explain any one algorithm for finding communities in a Social Graph. [10]**
- 5 a Why is finding similar items important in Big Data? Illustrate using two example applications. [10]**
- b Explain in detail any one Ranking algorithm used by Search Engines. [10]**
- 6 a Explain Collaborative Filtering based recommendation System. How it is different from content based recommendation systems ? [10]**
- b Explain with diagram the Park Chn Yu (PCY) algorithm for frequent itemset mining. [10]**