N.B. Question No: 1 is Compulsory
Attempt any three from the remaining
Assume suitable data wherever necessary

1. a) Compare Big data analytics with traditional data mining and warehousing system.  
   b) With respect to data stream querying, give example of  
      a) one time Queries  
      b) continues queries  
      c) predefined Queries  
      d) ad-hoc queries  
   c) How Big Data Analytics can be useful in the development of Digital India  
   d) What are distance measures? Brief any two distance measures.

2. a) Explain K mean algorithms for large data set. Explain its significance as compared to other clustering algorithms.
   b) How Bloom filter is useful for big data analytics. Explain with one example.

3. a) Explain DGIM algorithm for counting ones in a window.  
   b) Elaborate collaborative filtering system. How is the system different from a content based system.

4. a) Explain Hadoop ecosystem with core components? Explain the Physical Architecture of Hadoop. State its limitations  
   b) What is the MapReduce? Explain the role of combiner with the help of an example.

5. a) Differentiate between a RDBMS and a No-Sql Database  
   b) Explain the Clique Percolation Method (CPM) used in direct discovery of communities in a Social Graph with example.

6. a) List down the steps in HITS Algorithm with one example.  
   b) Explain Park-Chen-Yu algorithm. How memory mapping is done in PCY.