(3 Hours)   Max. Marks: **80**

N. B: 1. Question **ONE** is compulsory.
     2. Attempt any **THREE** out of remaining.
     3. **Figures** to the **right** indicate **full** marks.
     4. Assume suitable data if **necessary**.

Qu-1	a)	What are the characteristics of big Data?	5
	b)	Explain Hadoop Architectural Model.	5
	c)	List the different issues and challenges in data stream query processing.	5
	d)	Explain NoSQL data Architecture patterns.	5
Qu-2	a)	Explain DGIM algorithm for counting ones in a stream with example.	10
	b)	Explain Social Network graph clustering algorithm with example.	10
Qu-3	a)	Explain Model for Recommendation System in detail	10
	b)	Explain Matrix - Matrix Multiplication using TWO step MapReduce model.	10
Qu-4	a)	Explain PageRank algorithm with suitable example.	10
	b)	Explain Bloom's filter for stream data mining with example.	10
Qu-5	a)	Explain PCY algorithm with suitable example.	10
	b)	i) Find Jaccard distance {1, 2, 3, 4} & {2, 3, 5, 7} and {a, a, a, b} & {a, a, b, b, c} ii) Find Hamming Distance between 110011 & 010101 and 11001 & 01011 iii) Compute the cosines of the angles between $(3, -1, 2)$ and $(-2, 3, 1)$.	10
Qu-6		Write a note on	20
	a)	Hadoop Ecosystem	
	b)	CURE Algorithm	
	c)	HITS	
	d)	MapReduce programming model	