

5. Attempt any *one* part of the following: 10*1 = 10

- (a) Describe briefly the architecture of Data Stream Management System (DSMS).
- (b) Explain any one algorithm for counting oneness in a window.

6. Attempt any *one* part of the following: 10*1 = 10

- (a) For the given data, find the association rule using apriori algorithm.

TID	ITEMSETS
T1	A, B
T2	B, D
T3	B, C
T4	A, B, D
T5	A, C
T6	B, C
T7	A, C
T8	A, B, C, E
T9	A, B, C

Given: Minimum Support= 2, Minimum Confidence= 50%

- (b) What do you mean by k-means clustering? How does the k-means algorithm work? Write k-means algorithm for partitioning.

7. Attempt any *one* part of the following: 10*1 = 10

- (a) Define HDFS. Discuss the HDFS architecture and HDFS commands in brief.
- (b) Differentiate between Map Reduce and Apache Pig.