

- Q-5 Attempt all questions**
- (a) Explain Map reduce architecture and it's working in detail. (07)
 - (b) What is supervised and unsupervised learning? (03)
 - (c) Explain CAP theorem in brief. (04)

- Q-6 Attempt all questions**
- (a) Explain the architecture of Hive. List out features of Hive. (07)
 - (b) How RDBMS differs from HBase? (04)
 - (c) Write down the features of HDFS. (03)

- Q-7 Attempt all questions** (14)
- (a) Explain K-means algorithm in detail with suitable example.
 - (b) Explain CRUD operations in MongoDB.

- Q-8 Attempt all questions** (14)
- (a) You are given the transaction data shown in the Table below from a fast food restaurant. There are 9 distinct transactions (order:1–order:9) and each transaction involves between 2 and 4 meal items. There are a total of 5 meal items that are involved in the transactions. For simplicity we assign the meal items short names (M1 –M5) rather than the full descriptive names (e.g., Big Mac).

Meal Item	List of Item IDs	Meal Item	List of Item IDs
Order:1	M1, M2, M5	Order:6	M2, M3
Order:2	M2, M4	Order:7	M1, M3
Order:3	M2, M3	Order:8	M1, M2, M3, M5
Order:4	M1, M2, M4	Order:9	M1, M2, M3
Order:5	M1, M3		

For all of the parts below the minimum support is $2/9(.222)$ and the minimum confidence is $7/9(.777)$. Note that you only need to achieve this level, not exceeds it. Apply the Apriori algorithm to the dataset of transactions and identify all frequent k-item sets. Show all of your work. You must show candidates but can cross them off to show the ones that pass the minimum support threshold.

- (b) Write a short note on intellectual property challenges for big data.

