Final Year B.Tech. Sem-II,Computer Science & Engg
Elective-IV Big Data Analytics
Question Bank

1. What are the issues with the unstructured data?
2. How to deal with the unstructured data?
3. What is Big data ? What are the sources of Big Data?
4. State the characteristics of Big Data.
5. What are differences between traditional BI & Big Data?
6. Explain with diagram Big Data Technology Stack.
7. What is Big Data Analytics? Explain
8. Compare Analytics 1.0,2.0 and 3.0
9. What is Data Science ?
10. Explain the process of Data Science
11. What are the responsibilities of Data Scientist?
12. What are the advantages of In-memory data analytics and In-database processing?
13. What is the difference between parallel systems and distributed systems?
14. Distinguish between SMP & MPP.
15. What are NoSQL databases? Explain different types of NoSQL databases.
16. What are the advantages of NoSQL Databases
17. Explain different components of Hadoop ecosystem.
18. What are functions of HDFE daemons: Namenode, datanode and secondary name node?
19. With diagram. Illustrate the anatomy of File write operation in hadoop
20. What are the features of MongoDB?
21. Explain how to import data from CSV file to Mongo DB collection with example.
22. What are the replication strategies used in Cassandra?
23. Demonstrate with example CRUD operations in Cassandra.
24. Illustrate Import & Export operations with example in Cassandra.
25. Explain the components of HIVE architecture.
26. With Diagram, illustrate the working of HIVE.
27. What is partitioning & bucketing in HIVE? What are its advantages?
28. How to create partitions and buckets using Hive query language?
29. What is Static and Dynamic partitioning in HIVE?
30. Write a HIVE function to convert the values of a field to uppercase using UDFs?
31. List four important features of Pig.
32. Explain the anatomy of Pig
33. List the advantages and limitations of Pig.
34. Illustrate LOAD & STORE operators in Pig with example.
35. Illustrate Grouping & Joining operations in Pig with example.
36. Illustrate with example, how MapReduce works.
37. Distinguish between RBBMS and NoSQL.
38. What are the different types of Digital data? Mention various sources of digital data.
39. With diagram, explain the application workflow in hadoop yarn.
40. What are the functions of Application and Global resource manager in Hadoop