Sessional 3

Big Data Analytics

Rohan Joshi

1. What is RCFile in hive? [2]
2. What is SerDe in hive?[2]
3. What is Machine Learning? list down its applications.[2]
4. When should we use SORT BY instead of ORDER BY in hive? [2]
5. Discuss Data units in the hive.[2]
6. What are the different components of a Hive architecture?[3]
7. Compare and contrast between Hive and Pig [3]
8. What is a partition in Hive? Explain with the help of an example.[3]
9. What are the different ways to import and export data in Cassandra? Explain with the help of an example. [4]
10. Write a short note on Mahout – Recommendation. [4]
11. What do you understand by bucketing in Hive? Explain with the help of an example. [4]
12. Explain the types of Collections available in Cassandra with the help of appropriate examples.[6]
13. Discuss different types of joins in Hive.[6]