

SPARK written assignment

1. What is Spark? Why do we require Apache Spark?
2. Differentiate between Spark and Hadoop MapReduce.
3. List the areas where Spark and Hadoop MapReduce are good.
4. How Spark manages the Cluster and Resources?
5. Write short note on Spark Driver and Executors.
6. Write the Spark Unified Stack and explain Spark Core.
7. What are RDDs? Explain the properties of RDD.
8. What are transformation and Actions in Spark? Give suitable examples.
9. What is Lazy evaluation in Spark? Explain with an example.
10. With Logical and Physical plans explain how structured APIs will be executed in Spark.
11. Explain some use cases of stream processing.
12. List out the challenges of stream processing.
13. Explain Continuous Stream Processing method. What are its advantages and disadvantages?
14. Explain Micro-Batch Stream Processing method. What are its advantages and disadvantages?
15. Write short notes on DStreaming and Structured Streaming.
16. Write Short notes on Data Sources and Data Sinks of Structured Streaming in Spark.