Assignment-8

The due date for submitting this assignment has passed. As per our records, you have not submitted this assignment.

1. Which is not a component on the top of Spark Core?  
   - Spark RDD  
   - Spark Streaming  
   - VLSI

   Accepted Answers:
   - VLSI

   True or False?  
   - True

   Notes: (Rdd: Resilient Distributed Dataset) are fault-tolerant and immutable.

   2. In Kafka, _________ is the node responsible for all reads and writes for the given partition.

   Accepted Answers:
   - Leader
   - Follower
   - Other
   - None of the mentioned

   Notes: the partition is incremeneted.

3. What is type of processing in Apache Spark can handle?  
   - Batch Processing  
   - Stream Processing  
   - Graph Processing  
   - All of the mentioned

   Accepted Answers:
   - All of the mentioned

4. Kafka is a real-time collection of objects partitioned across a set of machines that can be rebuilt if a partition is lost.

   Accepted Answers:
   - True

5. Identify the correct Kafka commands for the following:

   P or topic in connect, subscribe, and consume large-scale data applications.

   Accepted Answers:
   - P (kafka)  
   - P (kafka-streams)  
   - P (kafka-consumers)  
   - P (kafka-connections)  

6. Kafka in run as a cluster comprised of one or more servers each of which is called

   Accepted Answers:
   - Brokers
   - Contexts
   - Connects
   - Workers

7. Kafka maintains feeds of messages in categories called

   Accepted Answers:
   - Topics
   - Topics
   - Domains
   - Workers

8. __________ is a distributed graph processing framework on top of Spark.

   Accepted Answers:
   - GraphX
   - Graph
   - All of the mentioned

9. __________ is a distributed machine learning framework on top of Spark.

   Accepted Answers:
   - MLlib
   - Spark Streaming
   - GraphX
   - REO