Paper ID : 250558

Subject Code: KDS601

Roll No: | | | | | | | | | | | | | |

BTECH

(SEM VI) THEORY EXAMINATION 2024-25

BIG DATA AND ANALYTICS

TIME: 3 HRS **M.MARKS: 100**

Note: Attempt all Sections. In case of any missing data; choose suitably.

SECTION A

1. Attempt *all* questions in brief. **2 x 10 = 20**

Q No.	Question	CO	Level
a.	Name the Different Sources of Big Data.	1	K1
b.	Elaborate different technology components in Big Data.	1	K2
c.	Define the purpose of MapReduce in Hadoop.	2	K1
d.	Explain the data replication in Hadoop Distributed File System (HDFS).	2	K2
e.	Describe the primary benefit of using NoSQL databases.	3	K6
f.	Explain the steps of data stored in MongoDB.	3	K6
g.	Discuss the purpose of the Grunt shell in Pig Latin.	4	K2
h.	Describe the concepts of file sizes, block sizes, and block abstraction in HDFS.	4	K5
i.	Illustrate the execution mode of Pig in Apache Pig.	5	K2
j.	Discover the significance of Spark in Big Data analytics.	5	K5

SECTION B

2. Attempt any *three* of the following: **10 x 3 = 30**

Q No.	Question	CO	Level
a.	Illustrate the working of MapReduce and the steps involved in the MapReduce framework.	1	K2
b.	Explore the architecture and components of the Hadoop Distributed File System (HDFS).	2	K2
c.	Discover the benefits and challenges associated with NoSQL databases and explain why they are suitable for Big Data applications.	3	K6
d.	Illustrate the features and capabilities of HBase in the context of Big Data applications.	4	K2
e.	Inspect the evolution of NOSQL databases, and explain CAP theorem.	5	K5

SECTION C

3. Attempt any *one* part of the following: **10 x 1 = 10**

Q No.	Question	CO	Level
a.	Explore the history of Big Data innovation and its impact on various industries. Discuss the drivers that have led to the growth of Big Data.	1	K2
b.	Compare Big Data Analytics from Traditional Data Analytics with its advantages and disadvantage along with Big Data privacy and ethics. Give a suitable example.	1	K2

4. Attempt any *one* part of the following: **10 x 1 = 10**

Q No.	Question	CO	Level
a.	Illustrate the design principles and key features of the Hadoop Distributed File System (HDFS). Explain how HDFS handles large-scale data storage and processing.	2	K2
b.	Illustrate the following: a) Security in Hadoop b) Administering Hadoop	2	K2

Paper ID : 250558

Roll No: | | | | | | | | | | | | | |

BTECH
(SEM VI) THEORY EXAMINATION 2024-25
BIG DATA AND ANALYTICS

TIME: 3 HRS **M.MARKS: 100**

5. Attempt any *one* part of the following: 10 x 1 = 10

Q No.	Question	CO	Level
a.	Explore the history of Big Data innovation and its impact on various industries. Discuss the drivers that have led to the growth of Big Data.	3	K6
b.	Explore the concept of the Hadoop ecosystem and its key components. Discuss the role of each component in processing and analyzing Big Data.	3	K6

6. Attempt any *one* part of the following: 10 x 1 = 10

Q No.	Question	CO	Level
a.	Examine role of DAG in Apache Spark. Classify the types of Deploy Modes in Spark.	4	K2
b.	Examine the setup and installation process of a Hadoop cluster. Explain the steps involved in configuring and administering Hadoop, including security measures and monitoring.	4	K2

7. Attempt any *one* part of the following: 10 x 1 = 10

Q No.	Question	CO	Level
a.	Illustrate the key features and capabilities of the Hadoop ecosystem frameworks, such as Pig, Hive, and HBase. Explain their role in data processing and analytics.	5	K5
b.	Illustrate the advantages of NoSQL databases and discuss the features and functionalities of MongoDB. Explain the data manipulation and querying operations in MongoDB.	5	K5