

C

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code:U2E04

B.E./B.Tech. DEGREE EXAMINATION, NOV 2024

Second Semester

Artificial Intelligence & Data Science

21UAD204- FOUNDATION OF DATA SCIENCE

(Regulations 2021)

Duration: Three hours

Maximum: 100 Marks

Answer All Questions

PART A - (5x 1 = 5 Marks)

- Which of the following would be more appropriate to be replaced with question mark in the following figure? CO1-U
(a) Data Analysis (b) Data Science
(c) Descriptive Analytics (d) None of these
- Data in _____ bytes size is called Big Data. CO2- U
(a) Tera (b) Giga (c) Peta (d) meta
- According to analysts, for what can traditional IT systems provide a foundation when they're integrated with big data technologies like Hadoop? CO3- U
(a) Big data management and data mining
(b) Data warehousing and business intelligence
(c) Management of Hadoop clusters
(d) Collecting and storing unstructured data
- _____ is general-purpose computing model and runtime system for distributed data analytics. CO1- U
(a) MapReduce (b) Drill (c) Oozie (d) None of the mentioned

5. The daemons associated with the MapReduce phase are _____ and task-trackers. CO1- U

- (a) job-tracker (b) map-tracker (c) reduce-tracker (d) all of the mentioned

PART – B (5 x 3= 15Marks)

6. Write a structure of Data frame using 5 students Mark list. CO1- U

7. What are the advantages of clustering? CO2- U

8. Write the advantage and disadvantage of clustering CO3- U

9. How many name nodes are there in HDFS? CO4- U

10. What is Job Tracker? CO5- U

PART – C (5 x 16= 80Marks)

11. (a) Write Data Analytics Process Steps? Write types of Data Analytics and its applications. CO1-U (16)

Or

(b) Write Python decision making statements with example ? CO1-U (16)
Write Python Looping statements with example.

12. (a) How to construct a decision tree in R using classification techniques With suitable example. CO2-U (16)

Or

(b) What is Bayesian Classifier? Write the applications of naïve bayes Classifier? CO2-U (16)

13. (a) What is the function of Secondary Name node? Write the advantage and disadvantage of Hadoop? CO3-U (16)

Or

(b) What is a checkpoint? How is HDFS fault tolerant? CO3-U (16)

14. (a) What is meant by HDFS? What is Hadoop HDFS architecture with diagram? CO4- U (16)

Or

(b) Write the anatomy of read and write a file in HDFS? CO4- U (16)

15. (a) What is mapper phase and Reducer phase in Hadoop? CO5-U (16)

Or

(b) What is map reduce with Example? Write the advantage and disadvantage of map reduction? CO5-U (16)