Reg No.:_____   Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester MCA Degree Regular and Supplementary Examination July 2021

Course Code: RLMCA204
Course Name: BIG DATA TECHNOLOGIES

Max. Marks: 60   Duration: 3 Hours

PART A
Answer all questions, each carries 3 marks.   Marks

1	Name and explain the four V's of big data.	(3)
2	Compare and contrast parallel data processing and distributed data processing.	(3)
3	Illustrate architecture of distributed file systems.	(3)
4	What are the features used by HDFS to ensure data integrity.	(3)
5	Explain the working of MapReduce algorithm.	(3)
6	Compare and contrast NoSQL and RDBMS database.	(3)
7	"In-Memory storage devices are faster at transferring data than on disk storage devices". Justify the statement.	(3)
8	List out the roles of data mining technique in big data analysis.	(3)

PART B
Answer any one question from each module. Each question carries 6 marks.

MODULE-I

9	List out different categories of big data analytic methodologies.	(6)

OR

10	Illustrate and explain various cloud deployment models used in big data.	(6)

MODULE-II

11	Write short notes on the following.	(6)

i) CAP

ii) ACID

iii) BASE

OR

12	Explain big data processing method in batch mode using map reduce algorithm.	(6)

MODULE-III

13   Illustrate and explain HDFS architecture.   (6)

OR

14   Write short notes on   (6)

   i) Hive

   ii) Sqoop

   iii) Flume

MODULE-IV

15   List out the prime features of MapReduce.   (6)

OR

16   Explain the role of map and reduce function with the help of an example.   (6)

MODULE-V

17   List out and explain types of NoSQL storage devices based on the way of   (6)
      storage.

OR

18   Write short notes on   (6)

   i) In-Memory datagrid

   ii) In-Memory database

MODULE-VI

19   Describe how machine learning supports data mining process.   (6)

OR

20   Explain the types of semantic analysis methods.   (6)

****