

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF ENGINEERING SCIENCES & TECHNOLOGY
JAMIA HAMDARD

ODD SEMESTER EXAM 2022

PAPER: BIG DATA ANALYTICS
TIME: 3 HRS

PAPER CODE: BTCSE DED 42
MAXIMUM MARKS: 75

Note: Do not write anything on paper except Roll No. & Enrolment No.
Use of mobile phone is not allowed
Please follow proper numbering while answering the questions
Proper steps are required to be specified for all questions

COURSE OUTCOME (CO)

CO1: The student will learn about the basics of big data, its need and significance.

CO2: An overview of the concepts, processes, and best practices needed for big data analytics and Hadoop.

CO3: Students will also learn about the HDFS.

CO4: Students will learn the basics of Map Reduce.

CO5: Students will also develop an understanding of the Hadoop Ecosystem.

CO6: Students will also develop an understanding of the basics of Machine Learning.

SECTION A

*All questions of Section A are compulsory (7*5)*

S.N O	QUESTIONS	MARK S	CO	BL
Q1.	If you have 3 input files of size 128K, 152MB and 256MB and the block size of cluster is 64 MB. How many input splits will be made by Hadoop framework?	5	CO1	L3
Q2.	"HDFS block is larger compared to disk block." State if the statement is true or False. Give a valid reason for your answer.	5	CO5	L4
Q3.	What is Rack Awareness in Hadoop? Discuss its main purpose.	5	CO2	L2
Q4.	What are the advantages of using a distributed storage system?	5	CO2	L2
Q5.	What will happen if a Namenode crashes in a Hadoop environment?	5	CO3	L3
Q6.	What is the small files problem in HDFS how to resolve it?	5	CO2	L1
Q7.	Discuss the various categories into which Machine Learning Algorithms are classified.	5	CO2	L1

SECTION B

*Answer any 4 questions in Section B (4*10)*

S . N O	QUESTIONS	MA RK S	CO	BL
Q1.	With suitable example explain the various phases of Map-Reduce in Hadoop.	10	C O3	L2
Q2.	Discuss in detail the architecture and operations of HDFS.	10	C O4	L1
Q3.	a. Discuss in brief the history of Big Data. b. Illustrate the various phases involved in Big Data Analytics with neat diagram.	10	C O3	L4
Q4.	Discuss in brief the basic features and purpose of the following w.r.t Hadoop: a. Apache Hive b. YARN c. Sqoop d. Flume	10	C O2	L2
Q5.	Discuss how can a music streaming startup like Spotify use Machine Learning to Create Value and Competitive Advantage.	10	C O5	L5
Q6.	You just moved into your dream apartment with a vision your interior decor. You've been to many shops but you haven't brought yourself up to buy anything because you are afraid it might not look great in your new space. Until one day a good friend of yours told you about IKEA app. There you can browse for your favorite furniture style and virtually place the objects in your house to see how they would look in person! And not just that, you can also choose from a range of wall colors to see what matches your style better. You finally can see how that beautiful sofa or coffee table will look without the need to buy and then return a bunch of items, all from the comfort of your own home. Considering the above scenario explain how Big Data Analytics is being leveraged by Ikea for its business.	10	C O3	L3

BLOOMS TAXANOMY LEVELS (BL)

1-REMEMBERING, 2-UNDERSTANDING, 3-APPLYING, 4-ANALYSING, 5-EVALUATING, 6-CREATING