

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF ENGINEERING SCIENCES & TECHNOLOGY
JAMIA HAMDARD
SUPPLEMENTARY EXAM 2023

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)*

SNo	Question	Marks	CO	BL
Q1	What are the characteristics of big data?	5	CO1	1
Q2	Discuss the conventional challenges in big data	5	CO1	2
Q3	Discuss Rack Awareness in Hadoop?	5	CO4	1
Q4	Discuss the various categories into which Machine Learning Algorithms are classified.	5	CO6	2
Q5	If you have 3 input files of size 238K, 132MB and 255MB and the block size of the cluster is 64 MB. How many input splits will be made by the Hadoop framework?	5	CO3	3
Q6	What is data preparation? What are the Steps for Data preparation.	5	CO2	2

Q7	What is the difference between InputSplit and Blocks in Hadoop	5	CO4	3
----	--	---	-----	---

SECTION B

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

SNo	Question	Marks	CO	BL
Q1	In the various stages of business analytics large volumes of data are processed at various steps. Depending on the stage of the workflow and the requirement there are different types of analytics. These different types of analytics together answer everything a company needs to know. Consider a scenario of a business enterprise. Name the different types of analytics that can be used and at which stage. Also, explain how these types of analytics can help the enterprise grow.	10	CO1	4
Q2	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	CO4	1
Q3	Discuss how a health and fitness application like Healthify uses Machine Learning to Create Value and Competitive Advantage.	10	CO2	5
Q4	Describe the MapReduce execution steps with neat diagram	10	CO4	2
Q5	Discuss in detail the architecture and operations of HDFS.	10	CO4	2
Q6	Define Big Data. Explain the Evolution of Big Data and their characteristics	10	CO1	2

BLOOM'S TAXONOMY LEVELS (BL): 1-REMEMBERING, 2-UNDERSTANDING, 3-APPLYING, 4-ANALYZING, 5-EVALUATING, 6-CREATING