NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
B.Tech
SEM:VI CARRY OVER THEORY EXAMINATION-AUGUST 2023
Subject: Big Data

Time: 3 Hours　　　　　　　　　　　　　　　　　　　Max. Marks: 100

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*

1. *This Question paper comprises of* **three Sections -A, B, & C.** *It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.*

2. *Maximum marks for each question are indicated on right -hand side of each question.*

3. *Illustrate your answers with neat sketches wherever necessary.*

4. *Assume suitable data if necessary.*

5. *Preferably, write the answers in sequential order.*

6. *No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.*

SECTION A　　　　　　　　　　　　　　　　　　　20

1. Attempt all parts:-

1-a.　　　Which is not a benefit of big data architecture (CO1)　　　　　　　1

　　　　　(a) Technology choices

　　　　　(b) Performance through Parallelism

　　　　　(c) Non Elastic

　　　　　(d) Interperability with existing solution

1-b.　　　This should not be followed while dealing with Big data privacy and ethics (CO1)　　1

　　　　　(a) Private customer data should remain private

　　　　　(b) Transparent view

　　　　　(c) Big data should interfere with human will

　　　　　(d) Big data should not institutionalize unfair biases.

1-e.　　　Which is known as horizontal scaling? (CO2)　　　　　　　1

　　　　　(a) Scaling up

　　　　　(b) Scaling out

　　　　　(c) Scaling

.

(d) Scaling down

1-c.   HDFS files cannot be (CO2)   1

        (a) read

        (b) deleted

        (c) executed

        (d) archived

1-f.   Hadoop streaming allows running map reduce program in which of the   1
languages are used (CO3)

        (a) R

        (b) Python

        (c) Pearl

        (d) All of the above

1-d.   Hadoop is shared-nothing architecture (CO3)   1

        (a) TRUE

        (b) FALSE

        (c) Can't Say

        (d) May be

1-g.   How many phases are available in cloud computing planning? (CO4)   1

        (a) 1

        (b) 5

        (c) 3

        (d) 6

1-h.   Which is not a cloud stakeholder? (CO4)   1

        (a) Cloud providers

        (b) Clients

        (c) End users

        (d) Cloud users

1-i.   Which is the broadcast address for a Class B network ID using the default   1
subnetmask? (CO5)

        (a) 172.16.10. 255

        (b) 255.255.255. 255

        (c) 172.16.255. 255

        (d) 172.255.255. 255

1-j.   Default network mask for CLASS A is (CO5)   1

        (a) 255. 0.0.0

        (b) 255. 255.0.0

        (c) 255. 255.255.0

        (d) 255. 255.255.255

2. Attempt all parts:-

2.a.   Write any two real time examples of big data. (CO1)   2

2.b.   Why hadoop? (CO2)   2

2.c   Can you explain the components of HDFS architecture. (CO3)   2

2.d.   Explain in detail about the resource pooling in cloud computing? (CO4)   2

2.e.   Define subnet. (CO5)   2

SECTION B   **30**

3. Answer any <u>five</u> of the following:-

    6

3-a.   Explain in detail about Apache hadoop daemons, snowflake and AWS. (CO1)   6

3-b.   Explain 5 V's of big data in detail. (CO1)   6

3-c.   Write about history of hadoop. (CO2)   6

3-d.   Explain all the data formats available in hadoop? (CO2)   6

3.e.   Explain the difference between NameNode, Backup Node and Checkpoint NameNode? (CO3)   6

3.f.   Describe federation and levels of federation in detail (CO4)

3.g   Define virtual private cloud and explain its types in detail (CO5)   6

SECTION C   **50**

4. Answer any <u>one</u> of the following:-

4-a.   Draw the big data architecture diagram and explain its components. (CO1)   10

4-b.   Explain in detail about the cloud computing evolution and draw its proper diagram? (CO1)   10

5. Answer any <u>one</u> of the following:-

5-a.   Explain block and block size and why is it needed in hadoop? (CO2)   10

5-b.   Define heartbeat and load balancing in HDFS and List out the features of HDFS. (CO2)   10

6. Answer any <u>one</u> of the following:-

6-a.   Can you explain how to import relational data to sqoop with using hadoop?   10

(CO3)

6-b.   Explain the roles of Namenode, Datanode, Secodary Namenode and   10
ResourceManager? (CO3)

7. Answer any <u>one</u> of the following:-

7-a.   Differentiate between public, private and hybrid cloud and explain its both   10
properties? (CO4)

7-b.   Can you explain how the MapReduce job run in hadoop with an example? (CO4)   10

8. Answer any <u>one</u> of the following:-

8-a.   Explain the concept of Google cloud storage? (CO5)   10

8-b.   Explain how machine learning is related to Bigquery. (CO5)   10