

Printed Page:-

Subject Code:- ACSAI0621

Roll. No:

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

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (2024-2025)

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. This data is created using a fixed schema and is maintained in tabular format (CO1) 1

- (a) Structured data
- (b) Unstructured data
- (c) Semi structured data
- (d) None of the above.

1-b. The main components of big data (CO1) 1

- (a) HDFS
- (b) Map Reduce
- (c) Yarn
- (d) All of the above

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

- (a) read
- (b) deleted
- (c) executed
- (d) archived

1-d. Every TaskTracker is configured with a _____.(CO2) 1

- (a) Maximum memory available in the node
- (b) Not limited

- (c) set of slots
- (d) As decided by the job tracker
- 1-e. Which one is taking a set of data and converted into another set of data? (CO3) 1
- (a) Reduce stage
- (b) Shuffling stage
- (c) Map stage
- (d) Shuffling and sorting
- 1-f. Hadoop streaming allows running map reduce program in which of the languages are used (CO3) 1
- (a) R
- (b) Python
- (c) Pearl
- (d) All of the above
- 1-g. It is not a property of cloud computing (CO4) 1
- (a) virtualization
- (b) composability
- (c) scalability
- (d) all of the mentioned
- 1-h. Which of these an example of SaaS? (CO4) 1
- (a) Google Workspace
- (b) Dropbox
- (c) Salesforce
- (d) All of the above
- 1-i. The uses of subnetting. CO5 1
- (a) It divides one large network into several smaller ones
- (b) It divides network into network classes
- (c) It speeds up the speed of network
- (d) None of above
- 1-j. The combination of _____ and _____ is often termed the local address of the local portion of the IP address. (CO5) 1
- (a) Network number and host number
- (b) Network number and subnet number
- (c) Subnet number and host numbe
- (d) Host number

2. Attempt all parts:-

- 2.a. Explain types of Big data. (CO1) 2
- 2.b. Explain the limitation of traditional approach in big data process? (CO2) 2
- 2.c. Can you explain the components of HDFS architecture.(CO3) 2

- | | | |
|------|----------------------------------|---|
| 2.d. | Define Cloud. (CO4) | 2 |
| 2.e. | State Virtual Private cloud. CO5 | 2 |

SECTION-B

30

3. Answer any five of the following:-

- | | | |
|------|---|---|
| 3-a. | Can you explain the digital data types and describe the data classification features? (CO1) | 6 |
| 3-b. | Define principles of Big Data ethics. (CO2) | 6 |
| 3-c. | How would you present the apache hadoop functions? (CO2) | 6 |
| 3-d. | Define HDFS. Also explain components of hadoop. (CO2) | 6 |
| 3.e. | Draw and explain the components of HDFS architecture in detail.CO3 | 6 |
| 3.f. | Define open stack. Also define it's components. (CO4) | 6 |
| 3.g. | Explain in detail the concept of Subnetting. Explain. (CO5) | 6 |

SECTION-C

50

4. Answer any one 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 one of the following:-

- | | | |
|------|--|----|
| 5-a. | Rank the importance of the HDFS and Demonstrate the HDFS architecture. (CO2) | 10 |
| 5-b. | Can you explain in detail about the HDFS architecture. (CO2) | 10 |

6. Answer any one of the following:-

- | | | |
|------|--|----|
| 6-a. | Can you explain how Apache Spark is used? (CO3) | 10 |
| 6-b. | Explain the role of following nodes in detail:
(i) Datanode
(ii) NodeManager
(iii) Job Tracker
(iv) Task Tracker
(v) NameNode (CO3) | 10 |

7. Answer any one of the following:-

- | | | |
|------|--|----|
| 7-a. | Differentiate between public, private and hybrid cloud and explain its both properties? (CO4) | 10 |
| 7-b. | List out features and challenges of GCP and explain in detail the features of any four services provided by GCP? (CO4) | 10 |

8. Answer any one of the following:-

- | | | |
|------|---|----|
| 8-a. | Define what is Datastore? Explain Cloud Datastore concept.CO5 | 10 |
| 8-b. | Discuss the BigTable concept in cloud with an example? (CO5) | 10 |