

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

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (2025 - 2026)

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:-

- | | | |
|-------|--|---|
| (1.1) | This data is created using a fixed schema and is maintained in tabular format (CO1,K2) | 1 |
| (a) | Structured data | |
| (b) | Unstructured data | |
| (c) | Semi structured data | |
| (d) | None of the above. | |
| (1.2) | The main components of big data (CO1, K2) | 1 |
| (a) | HDFS | |
| (b) | Map Reduce | |
| (c) | Yarn | |
| (d) | All of the above | |
| (1.3) | Name of C++ interface to hadoop map reduce. (CO2,K2) | 1 |
| (a) | Hadoop streaming | |
| (b) | Hadoop Pipes | |
| (c) | Mahaout | |
| (d) | Oozie | |
| (1.4) | Used for collection and ingestion of semi-structured data. (CO2,K2) | 1 |
| (a) | Hbase | |
| (b) | PIG | |
| (c) | Flume | |
| (d) | Sqoop | |

- (1.5) Hadoop streaming allows running map reduce program in which of the languages are used (CO3, K2) 1
- (a) R
 - (b) Python
 - (c) Pearl
 - (d) All of the above
- (1.6) This is a failure in YARN.(CO3,K2) 1
- (a) Failure of running task
 - (b) Failure of Node manager
 - (c) Faiure of application manager
 - (d) All of the above
- (1.7) Which of these an example of SaaS? (CO4, K3) 1
- (a) Google Workspace
 - (b) Dropbox
 - (c) Salesforce
 - (d) All of the above
- (1.8) The future of federation in cloud computing is (CO4,K3) 1
- (a) It is expected to become less important as more organizations move to centralized cloud platforms.
 - (b) It is likely to remain an important concept as more organizations adopt multi-cloud environments.
 - (c) It is likely to become irrelevant as cloud providers merge and consolidate.
 - (d) It is difficult to predict the future of federation in cloud computing.
- (1.9) The uses of subnetting. (CO5, K4) 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
- (1.10) The combination of _____ and _____ is often termed the local address of the local portion of the IP address. (CO5, K4) 1
- (a) Network number and host number
 - (b) Network number and subnet number
 - (c) Subnet number and host number
 - (d) Host number

2. Attempt all parts:-

- (2.1) List the three primary types of digital data?. (CO1, K2) 2
- (2.2) Define the two primary daemons of HDFS. (CO2, K2) 2
- (2.3) Define Cloud computing. (CO3, K2) 2
- (2.4) List and define the three primary Cloud Service Models. (CO4, K2) 2
- (2.5) Describe the function of a Network Interface Card (NIC) (CO5, K2) 2

SECTION-B

30

Answer any one of the following:-

- (3.1) Can you explain the benefits and major challenges of big data architecture. (CO1, K2) 6
- (3.2) Explain cloud computing. Also write its benefits and challenges in detail. (CO1, K2) 6

Answer any one of the following:-

- (3.3) Write internal procedure of Map reduce work. (CO2, K4) 6
- (3.4) Write all the features of Hadoop 2.0. (CO2, K3) 6

Answer any one of the following:-

- (3.5) Draw and explain the components of HDFS architecture in detail. (CO3, K2) 6
- (3.6) Explain brief of the following. i) PIG ii) HIVE (CO3, K2) 6

Answer any one of the following:-

- (3.7) Define open stack. Also define its components. (CO4, K3) 6
- (3.8) Contrast Public Cloud and Private Cloud in terms of cost, security, and scalability. (CO4, K3) 6

Answer any one of the following:-

- (3.9) Explain in detail the concept of Subnetting. Explain. (CO5, K4) 6
- (3.10) Draw diagram and explain about the VPC architecture also elaborate its types? (CO5, K4) 6

SECTION-C 50

4. Answer any one of the following:-

- (4.1) Draw the big data architecture diagram and explain its components. (CO1, K2) 10
- (4.2) Define the working of Cloud Computing in detail. Also explain technology components for Cloud Computing. (CO1, K2) 10

5. Answer any one of the following:-

- (5.1) Explain the application workflow of YARN with diagram. (CO2, K4) 10
- (5.2) Define input and Output formats of Map Reduce. (CO2, K4) 10

6. Answer any one of the following:-

- (6.1) Can you explain how Apache Spark is used? (CO3, K2) 10
- (6.2) Explain the role of following nodes in detail: (CO3, K2) 10
- (i) Datanode
 - (ii) NodeManager
 - (iii) Job Tracker
 - (iv) Task Tracker
 - (v) NameNode

7. Answer any one of the following:-

- (7.1) Provide a comprehensive definition of Cloud Federation and explain the specific 'Four Levels of Federation' used to categorize it (CO4, K3) 10
- (7.2) List out features and challenges of GCP and explain in detail the features of any four services provided by GCP? (CO4, K3) 10

8. Answer any one of the following:-

- | | | |
|-------|---|----|
| (8.1) | Define what is Datastore? Explain Cloud Datastore concept.(CO5, K4) | 10 |
| (8.2) | Discuss the BigTable concept in cloud with an example? (CO5, K4) | 10 |

REG_JAN_JUNE_2026