Time: 3 Hours **Marks: 80**

Note: 1. Question 1 is compulsory

2. Answer any three out of remaining questions.
3. Assume suitable data wherever required and justify the same.

Q1 a) Explain time series mining with an appropriate example. [5]

b) What is sharding? Explain the advantages of sharding. [5]

c) What is data leakage with respect to big data? [5]

d) What are the basic differences between relational database and HDFS? [5]

Q2 a) Use PCA to transform 2D data space to 1D data space for the given matrix A. [10]

$$A = \begin{matrix} 0 & 1 \\ -2 & -3 \end{matrix}$$

b) Explain singular value decomposition (SVD) with an example. [10]

Q3 a) Explain Gaussian (normal) distribution with respect to pdf and cdf and its use in statistics. [10]

b) Explain Independent Component Analysis (ICA) in descriptive modeling. [10]

Q4 a) What type of problem were you looking to solve with text mining? How did you know how to text mine? What could be the challenges when text mining? [10]

b) What is Recommendation System (RS)? Explain types of RS and various issues and challenges in RS. [10]

Q5 a) Draw and describe the information visualization process. [10]

b) How would you validate a model you created to generate a predictive model of a quantitative outcome variable using multiple regressions? [10]

Q6 a) Describe the working of the Map-Reduce with an example. [10]

b) What is No SQL? Compare No SQL with SQL. [10]

Q. P. Code: 38873

(3 Hours)

Total Marks: 80

N.B.: (1) Question No.1 is compulsory.

(2) Attempt any three questions from the remaining five questions.

(3) Make suitable assumptions wherever necessary but justify your assumptions.

1.	(a) What is hacking? Who are the different types of hackers?	05
	(b) Explain qualified forensic duplicate, restored image and mirror image.	05
	(c) What volatile data can be obtained from investigation of routers?	05
	(d) What are the different ways to recover deleted files from a Unix system?	05
2.	(a) What do you mean by incident response methodology? Explain the different phases.	10
	(b) Briefly explain the process of collecting the volatile data in Windows system.	10
3.	(a) Briefly explain the role of Windows registry in collecting forensic evidence.	10
	(b) Explain the method for performing the mobile forensic.	10
4.	(a) Discuss the steps for investigating routers.	10
	(b) Explain the steps for e-mail forensic investigation.	10
5.	(a) What are the requirements of forensic duplication tools? Elaborate different ways of creating a forensic duplicate of a hard-disk.	10
	(b) What is digital evidence? What are the different types of digital evidence?	10
6.	Write a short note on**:(any two)**	20

(1) Challenges of performing mobile forensic.

(2) Layout of report writing

(3) Chain of custody

****

Q.P Code : 39526

Total Marks:80 **Duration:3Hrs.**

N.B.:1) Question No.1 is **compulsory**.
2) Attempt **any three** from remaining 5 questions.
3) Draw the **relevant** diagram neatly.

1.a) Explain benefits of storage virtualization. (05)
b) Which components constitute the disk service time? Which one contributes largest percentage? (05)
c) Compare DAS, NAS, SAN technologies. (05)
d) Benefits of Content Address Storage? (05)

2.a) Current configuration is RAID 0 with 6 disks of 200GB capacity each and total file size is 900GB. Amount of data likely to increase by 30% over next six month with 40% writes and a new 200GB disk drive cost is Rs.3000.What is cost of new solution? Justify your choice based on cost, performance & data availability. (10)

b) State the issues of manual and automatic indexing. (10)

3.a Differentiate restoration process in incremental and cumulative backups. (10)

b) Explain copy on write technology. (10)

4.a) List and explain the two primary protocols available for transmitting storage data traffic over TECP/IP. (10)
b) What are the advantages and disadvantages of controlled vocabulary? (10)

5.a) Explain benefits of NAS and factors affecting on NAS performance. (10)

b) Explain probabilistic matching process in detail. (10)

6 . Write short note on (any two) (20)
a) Single point failure.
b) Ectosystem & Endosystem
c) Fibre Channel ports

[Total Marks 80]

(3 Hours)

N. B:

1. Question **No. 1 is Compulsory**.
2. Solve any **THREE** from Question No. 2 to 6.
3. Draw neat well labeled diagram wherever necessary

Q. 1	a)	What are the different network topologies? Explain in detail.	**(10)**
	b)	Compare and contrast loosely coupled and tightly coupled multiprocessors.	**(10)**
Q. 2	a)	Write a MPI program to find factorial of given number.	**(10)**
	b)	What is a Data Race? Why Data-Races are Undesired? How Data-Races Can be Prevented?	**(10)**
Q. 3	a)	Explain Flynn's classification in detail.	**(10)**
	b)	Explain Granularity, Concurrency and Dependency Path.	**(10)**
Q. 4	a)	Explain various levels of parallel processing.	**(10)**
	b)	What is OpenMP? Explain OpenMP compiler directives? What are the Pros and Cons of OpenMP.	**(10)**
Q. 5	a)	Draw and explain NVIDIA GPU architecture.	**(10)**
	b)	What are the different Performance metrics?	**(10)**
Q. 6		Write short notes on any **FOUR**: 1. Quantum Computers 2. Data flow computers 3. Memory organization. 4. Message passing interface 5. Non-uniform memory access model.	**(20)**