

Time: 3 Hours

Max. Marks: 80

Instructions:

- Figures to the right indicate max marks.
- Draw appropriate diagram whenever applicable.
- Assume suitable data wherever applicable. State your assumptions clearly.
- **Question number 1 is compulsory.**
- Attempt **any Three** questions from remaining questions

Q.1 Solve any Four

- A. Explain SVD and its applications? [05]
- B. Differentiate between supervised and unsupervised learning. [05]
- C. Explain Hebbian Learning rule [05]
- D. Explain Perceptron model with Bias. [05]
- E. Differentiate between Ridge and Lasso Regression [05]

Q.2 Solve the following

- A. Draw a block diagram of the Error Back Propagation Algorithm and explain with the flow chart the Error Back Propagation Concept. [10]
- B. Find a linear regression equation for the following two sets of data: [10]

X	Y
3	12
5	18
7	24
9	30

Q.3 Solve the following

- A. Diagonalize the matrix A [05]

$$\begin{bmatrix} 1 & 5 \\ 4 & 2 \end{bmatrix}$$

- B. Write short note on overfitting and underfitting of model [05]
- C. What are activation functions? Explain Binary, Bipolar, Continuous, and Ramp activation functions. [10]

4. Solve the following

- A. Explain Least-Squares Regression for classification. [10]
- B. What is the curse of dimensionality? Explain PCA dimensionality reduction technique in detail. [10]

Q. 5 Solve the following

- A. How to calculate Performance Measures by Measuring Quality of model. [10]
- B. Explain the Perceptron Neural Network [10]

Q. 6.

- A. Discuss the various steps of developing a Machine Learning Application. [10]
- B. Write a short note on LMS-Widrow Hoff [05]
- C. Explain the Maximization algorithm for clustering. [05]

Time: 3 hours

Max. Marks: 80

Instructions:

- 1) Only **Four question** need to be solved.
- 2) All question carries equal marks.
- 3) Illustrate your answers with neat sketches wherever necessary.
- 4) Figures to the right indicate full marks.
- 5) Assume suitable additional data, if necessary and clearly state it.
- 6) All sub-questions of the same question should be grouped together.

- Q.1** Solve any **four**
- | | | | |
|--|-----|---|-----------|
| | (a) | What are various issues of distributed system? | 05 |
| | (b) | Justify how Ricart-Agrawala's algorithm optimized the Message overhead in achieving mutual exclusion. | 05 |
| | (c) | Explain the election algorithm? | 05 |
| | (d) | Explain Suzuki-Kasami algorithm? | 05 |
| | (e) | Difference between RMI and RPC? | 05 |
- Q.2**
- | | | | |
|--|-----|--|-----------|
| | (a) | What is distributed computing? Explain various system models of distributed computing? | 10 |
| | (b) | Define Remote Procedure Call (RPC). Explain the working of RPC in Detail. | 10 |
- Q.3**
- | | | | |
|--|-----|--|-----------|
| | (a) | What is a logical clock? Why are logical clocks needed in a distributed system? Explain Lamport algorithm. | 10 |
| | (b) | Describe code migration issues in detail? | 10 |
- Q.4**
- | | | | |
|--|-----|---|-----------|
| | (a) | Explain Hadoop Distributed File System (HDFS). | 10 |
| | (b) | Differentiate between message-oriented communication and stream-oriented communication. | 10 |
- Q.5**
- | | | | |
|--|-----|---|-----------|
| | (a) | Compare Load sharing to Task Assignment and Load balancing strategies for scheduling processes in a distributed system. | 10 |
| | (b) | Discuss various client centric consistency models. | 10 |
- Q.6** Write Short note (Any 2)
- | | | | |
|--|-----|--------------------------------|-----------|
| | (a) | Physical Clock Synchronization | 10 |
| | (b) | Load balancing techniques | 10 |
| | (c) | Andrew File System (AFS) | 10 |
| | (d) | Fault tolerance | 10 |

Time: 3 Hrs

Marks: 80

Note: 1. Q.1 Compulsory

2. Solve any 3 question from remaining five questions

Q1. Each sub-question carries 05 marks.

- a) Define software engineering and explain umbrella activities 5 M
- b) Explain the 4'PS of project Management 5 M
- c) Explain functional and non-functional requirements 5 M
- d) Explain the Agile process model 5 M

- Q2 a) Elaborate COCOMO model for Cost estimation 10 M
b) Illustrate the SCM process of Software quality management. 10 M

- Q3. a) Describe the waterfall model and incremental process model 10 M
b) What is Risk management? Discuss RMMM plan for risk management 10 M

- Q4.a) What are the different phases in project life cycle explain with suitable example 10M
b) Explain the user interface design in details with example 10M

- Q5. a) Develop the SRS of Hospital Management system 10M
b) Describe the details of FTR and Walkthrough 10M

- Q6. a) Explain project scheduling and describe CPM and PERT 10 M
b) Differentiate between white box and black box testing 10 M

39291

(3 Hours)

(Total Marks: 80)

- N.B.:**
1. Question No. 1 is compulsory.
 2. Answer any three out of the remaining questions.
 3. Assume suitable data if necessary.
 4. Figures to the right indicate full marks.

Q1. Attempt the following (any 4):

(20)

- a. Why is data analytics lifecycle essential?
- b. The regression lines of a sample are $x + 6y = 6$ and $3x + 2y = 10$.
Find (i) sample means \bar{x} and \bar{y} .
(ii) coefficient of correlation between x and y
- c. Differentiate between linear regression and logistic regression.
- d. What is Pandas? State and explain key features of Pandas.
- e. Explain term frequency (TF), document frequency (DF), and inverse document frequency (IDF).

Q2. Attempt the following:

- a. Explain the data analytics lifecycle. (10)
- b. Find two lines of regression from the following data: (10)

Age of husband (x)	25	22	28	26	35	20	22	40	20	18
Age of wife (y)	18	15	20	17	22	14	16	21	15	14

Estimate (i) the age of husband when the age of wife is 19 and (ii) the age of wife when the age of the husband is 30.

Attempt the following:

- a. Explain Box-Jenkins intervention analysis. (10)
- b. What is text mining? Enlist and explain the seven practice areas of text analytics. (10)

Q4. Attempt the following:

- a. Explain different types of data visualizations in R programming language. (10)
- b. Fit a regression equation to estimate β_0 , β_1 , and β_2 to the following data of a transport company on the weights of 6 shipments, the distances they were moved and the damage of the goods that was incurred. (10)

Weight X_1 (1000 kg)	4.0	3.0	1.6	1.2	3.4	4.8
Distance X_2 (100 km)	1.5	2.2	1.0	2.0	0.8	1.6
Damage Y (Rs.)	160	112	69	90	123	186

Estimate the damage when a shipment of 3700 kg is moved to a distance of 260 km.

Q5. Attempt the following:

- a. From the following results, obtain two regression equations and estimate the yield when the rainfall is 29 cm and the rainfall when the yield is 600 kg. (10)

	Yield in Kg.	Rainfall in cm
Mean	508.4	26.7
SD	36.8	4.6
Coefficient of Correlation	0.52	

- b. What is stepwise regression? State and explain different types of stepwise regression. (10)

Q6. Write short notes on (any 2):

- a. Time series analysis
b. Exploratory data analysis
c. Regression plot
d. Generalized linear model (GLM)

(20)

(3 Hours)

(Total Marks: 80)

- N.B.:** 1. Question No. 1 is compulsory.
 2. Answer any three out of the remaining questions.
 3. Assume suitable data if necessary.
 4. Figures to the right indicate full marks.

Q1. Attempt the following (any 4): (20)

- a. Compare descriptive and inferential statistics.
 b. A bag contains four balls. Two balls are drawn at random (without replacement) and are found to be white. What is the probability that all balls in the bag are white?
 c. Calculate the coefficient of correlation (r) for the following data:

x	9	8	7	6	5	4	3	2	1
y	15	16	14	13	11	12	10	8	9

- d. What are type I and type II errors?
 e. Construct 3 yearly moving averages from the following data:

Year	2017	2018	2019	2020	2021	2022	2023
Annual sale (in lakhs)	39	44	40	45	38	43	39

Q2. Attempt the following:

- a. The Dow Jones Travel Index reported what business travelers pay for hotel rooms per night in major U.S. cities (The Wall Street Journal, January 16, 2004). The average hotel room rates for 20 cities are as follows: (10)

Atlanta	310	Minneapolis	125
Boston	277	New Orleans	161
Chicago	166	New York	201
Cleveland	126	Orlando	146
Dallas	123	Phoenix	131
Denver	120	Pittsburgh	134
Detroit	144	San Francisco	167
Houston	173	Seattle	161
Los Angeles	160	St. Louis	143
Miami	192	Washington, D.C.	207

- (a) Calculate the mean hotel room rate (b) Calculate the median hotel room rate
 (c) Calculate the mode (d) Calculate the first quartile
 (e) Calculate the third quartile
- b. Explain Poisson probability distribution. A car distributor in a city Y experiences on an average 2.5 car sales per day. Find the probability that on a randomly selected day, (10)
 (a) they will sell 5 cars (b) they will sell no cars (c) they will sell at most 2 cars

Q3. Attempt the following:

- a. What is sampling? State and explain different sampling methods. (10)
 b. A random sample of size 64 is taken from a normal population with $\mu = 51.4$ and $\sigma = 6.8$. What is the probability that the mean of the sample will (10)
 (a) exceed 52.9 (b) fall between 50.5 and 52.3 (c) be less than 50.6

Q4. Attempt the following:

a. A company manufacturing automobile tyres finds that tyre life is normally distributed with a mean of 40,000 km and standard deviation of 3,000 km. It is believed that a change in the production process will result in a better product and the company has developed a new tyre. A sample of 100 new tyres has been selected. The company has found that the mean life of these tyres is 40,900 km. Can it be concluded that the new tyre is significantly better than the old one? Use 1% LOS. (10)

b. The population proportion (p) is 0.30. What is the probability that a sample proportion will be within ± 0.04 of the population proportion for each of the following sample sizes:

- (a) $n = 100$ (b) $n = 200$
 (c) $n = 500$ (d) $n = 1000$
 (e) What is the advantage of a larger sample size? (10)

Q5. Attempt the following:

a. The following data gives the experience of machine operators and their performance rating as given by the number of good parts turned out per 100 pieces. (10)

Operators	1	2	3	4	5	6
Performance rating (x)	23	43	53	63	73	83
Experience (y)	5	6	7	8	9	10

Calculate the regression line of performance rating (x) on experience (y) and also estimate the probable performance if an operator has 11 years of experience.

b. Consider a sample with a mean of 30 and a standard deviation of 5. Use Chebyshev's theorem to determine the percentage of the data within each of the following ranges. (10)

- (a) 20 to 40 (b) 15 to 45 (c) 22 to 38 (d) 18 to 42 (e) 12 to 48

Q6. Attempt the following:

a. From the following data, obtain two regression equations and estimate the value of sales when purchases were 75. Also, estimate the value of purchases when sales were 100. (10)

Sales	91	97	108	121	67	124	51	73	111	57
Purchases	71	75	69	97	70	91	39	61	80	47

b. What are non-parametric tests? Calculate Spearman's rank correlation coefficient between expenditure on advertising and sales from the data given below: (10)

Advertising expenses ('000 Rs.)	39	65	62	90	82	75	25	98	36	78
Sales (lakh Rs.)	47	53	58	86	62	68	60	91	51	84

(Time: 3 Hours)

(Total Marks: 80)

INSTRUCTIONS:

- (1) Question No. 1 is compulsory.
- (2) Attempt any (3) from remaining (5) questions.
- (3) Assume suitable data if required.
- (4) Figures to the right indicate full marks.

Q1 Attempt any four (4) questions.

- a) Explain various types of accounts used in ethereum. (05)
- b) Explain the need of smart contract. (05)
- c) Explain the concept of transaction hash in blockchain. (05)
- d) Compare private and public blockchain. (05)
- e) Explain UTXO model for bitcoin. (05)

Q2 a) Explain the concept of Gas and Gas fees used in ethereum transactions. (10)
b) Explain Architecture of Fabric. (10)

Q3 a) Compare various consensus used in private blockchain. (10)
b) Explain the role of ethereum virtual machine. Explain with example for state change. (10)

Q4 a) How does Blockchain used in e-voting application? (10)
b) Write a note on Blockchain on AWS or Azure cloud. (10)

Q5 a) Explain the phases of development in ETH 2.0 with suitable diagram. (10)
b) Explain different types of test networks used in ethereum. (10)

Q6 Write a short note on: (20)
a) Mining pool difficulty
b) Corda architecture
c) Tools and frameworks of Hyperledger.
d) Blockchain in Ecommerce website

Duration: 3 Hrs.

Marks: 80

Note:

1. Question 1 is compulsory.
2. Attempt any 3 questions out of the remaining questions.

Q1. Attempt any Four.

- a. Explain the different modes of block ciphers. **05**
- b. List with examples the different mechanisms to achieve security. **05**
- c. Differentiate MD5 and SHA-1 algorithms. **05**
- d. List and explain security requirements of database. **05**
- e. Explain phishing and list different types of phishing techniques. **05**

Q2.

- a. User A and B want to use RSA to communicate securely. A chooses public key as (7, 119) and B chooses public key as (13, 221). Calculate their private keys. A wishes to send message $m = 10$ to B. Produce the ciphertext. Formulate the key using which A encrypt the message "m" if A need to authenticate itself to B. **10**
- b. Explain memory and address protection in detail. Write a note on file protection. **10**

Q3.

- a. List the functions of the different protocols of SSL. Explain the handshake protocol. **10**
- b. List different poly-alphabetic substitution ciphers. Encrypt "The key is hidden under the door" using playfair cipher with keyword "domestic". **10**

Q4.

- a. Define digital signature. Explain any digital signature algorithm in detail. **10**
- b. Give the format of X.509 digital certificate and explain the use of a digital signature in it. **10**

Q5.

- a. Explain session hijacking and management. **10**
- b. What is need of Diffie-Hellman algorithm. User A and B decide to use Diffie-Hellman algorithm to share a key. They choose $p = 23$ and $g = 5$ as the public parameters. Their secret keys are 6 and 15 respectively. Compute the secret key that they share. **10**

Q6. Attempt any Four.

- a. Explain the different types of firewalls and mention the layer in which they operate. **05**
- b. List and explain vulnerabilities in windows operating system. **05**
- c. List and explain characteristics needed in secure hash function. **05**
- d. Explain Triple DES in short. **05**
- e. Explain with examples, keyed and keyless transposition ciphers. **05**

(3 Hours)

(Maximum Marks: 80)

- NB. 1. Question number one is compulsory**
2. Attempt any three out of remaining five questions
3. Assume suitable data
4. Figures to the right indicate the maximum marks

- Q1 Attempt any FOUR: (20)**
- a) Cyber Terrorism and Cyber Stalking
 - b) How to Present Digital Evidences
 - c) Comment on Windows Systems- FAT32 and NTFS
 - d) Sources of Network-Based Evidence
 - e) Describe Goals of Incident Response
- Q2 a) Explain Challenging Aspects of Digital Evidence (10)**
b) Describe DoS and Trojan Attacks (10)
- Q3 a) Describe Digital Investigation Staircase Model (10)**
b) Describe Forensic Investigation Report Writing in terms of Standards, Content, Style, Formatting and Organization. (10)
- Q4 a) Describe and Compare UNIX and MAC File System. (10)**
b) Memory Forensic and RAM Forensic Analysis (10)
- Q5 a) Explain Computer Forensic Tools in detail (10)**
b) Comment on Investigation of Routers and Firewalls (10)
- Q6 a) Describe Android OS Architecture and File Systems basics (10)**
b) Explain Image Acquisition over a Network and Removable Media (10)
