

Time: 3-Hour

Max. Marks: 80

N.B

1. **Q.1** is compulsory
2. Attempt any three from the remaining five questions.
3. Assume suitable data, if required and state it clearly.

Q1. Attempt any 4

20M

- a. Explain percentiles and Boxplots with example.
- b. Illustrate central limit theorem with a neat diagram.
- c. Explain Null and alternative Hypothesis with example.
- d. What is Histogram. Give its applications.
- e. How to detect outliers?
- f. What is the F-Test used for?

Q 2.

10M

- a. What is Chi-Square Test? A study is conducted to examine the relationship between gender and whether a person prefers coffee or tea. The data collected is as follows:

Gender \ Preference Coffee Tea Total

Male	30	10	40
Female	20	30	50
Total	50	40	90

Use chi-square test to find association between gender and beverage preference? (Use $\alpha=0.05$).

- b. Explain Normal and Poisson Distribution. Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council's GMAT examination. Scores on the GMAT are roughly normally distributed with a mean of 527 and a standard deviation of 112. What is the probability of an individual scoring above 500 on the GMAT? How high must an individual score on the GMAT in order to score in the highest 5%?

10M

Q3.

10M

- a. Explain single and Multiple linear regression with example and show with suitable plot.
- b. Explain t-Distribution in detail. The CEO of light bulbs manufacturing company claims that a light bulb lasts 300 days. A researcher randomly selects 15 bulbs testing. The sampled bulbs last an average than 290 days? deviation of 50 days. If the CEO's claim were true, what is the probability that 15 randomly selected bulbs would have an average life of no more than 290 days?

10M

Q 4.

10M

- a. Explain briefly why use ANOVA? Give difference between one-way and two-way ANOVA test. Solve the following using one way annova

Col1	Col 2	Col 3
82	71	64
93	62	73
61	85	87
74	94	91
69	78	56
70	66	78
53	71	87

- b. Explain Mean, standard deviation, standard error and Confidence Interval. Following table shows values of 10 data points of a sample. Find mean, standard deviation, standard error, and 95% confidence interval for the given sample. **10M**

Data	Value
1	4
2	2
3	2
4	3
5	5
6	2
7	6
8	2
9	7
10	6

- Q.5. a.** Explain how Nonparametric Method are different from parametric methods. **10M**
 Department of Public Health and safety monitors the measures taken to cleanup drinking water were effective. Trihalomethanes (THMs) at 12 counties drinking water compared before cleanup, 1 week later, and 2 weeks after cleanup. solve the following using the Friedman Test ?

County	Trihalomethanes (THMs)		
	Before Cleanup	Week1	Week2
1	21.1	19.2	18.4
2	24.1	22.3	21.2
3	14.1	12.9	12.9
4	18.1	17.8	17.3
5	15.4	15.1	14.9
6	16.2	15.1	15.1
7	7.4	7.2	6.8
8	7.5	6.7	6.1
9	14.2	13.6	13.1
10	21.3	20.9	20.4
11	9.5	9.8	9.2
12	11.9	10.5	10.1

- b. Define Continuous Probability distribution and Probability Distribution Function (PDF). Consider a random variable X that is uniformly distributed between 0 and 20. This means that the probability density function (PDF) of X is given by: $f(x)=1/20$, for $0 \leq x \leq 20$.
- Calculate the probability that X lies between 5 and 15, i.e., $P(5 \leq X \leq 15)$.
 - Calculate the probability that X is less than 10, i.e., $P(X < 10)$
 - Find the probability that X is greater than 18, i.e., $P(X > 18)$

10M

Q 6. Write short note **20M**

- Type I and Type II error in detail.
 - Histogram and Scatterplot with example.
 - Boot strapping Vs Re-sampling.
 - Weibull Distribution.
-

(3 hours)

Total Marks: 80

- N.B. 1. Question **No. 1** is compulsory
 2. Attempt any **three** questions from remaining five questions
 3. Assume suitable data if **necessary** and justify the assumptions
 4. Figures to the **right** indicate full marks

Q1 Answer the following questions.

- A** Explain the role of DNS and TLS in web communication. 05
 - B** Describe the differences between ES5 and ES6 with examples. 05
 - C** Discuss the lifecycle of a React component and its significance in single-page applications. 05
 - D** Given a scenario where a Node.js application has to handle multiple client requests asynchronously, outline the best practices to manage the callback functions effectively. 05
- Q2** A Write a JavaScript program to validate an email input form. Include error handling for invalid formats. 10
 B Explain the purpose of Express Router and how it aids in organizing route handling in a Node.js application. 10
- Q3** A Illustrate the use of **State** and **Props** in React with a practical example demonstrating their interaction. 10
 B Develop a simple program in Node.js to read data from a file and send it as a response to an HTTP request. Explain the key steps in the code. 10
- Q4** A You are tasked to design a form validation system in JavaScript for a signup form with username, password, and email fields. Explain how you would implement this with error handling and ensure data security. 10
 B Create a React component using functional components and Hooks to display a list of items that a user can add to or remove from. Describe how the **useState** and **useEffect** hooks are used in your solution. 10
- Q5** A Describe how REST APIs function in web development and provide an example of an HTTP GET request. 10
 B Write a React component to render a form and handle its submit event. The form should capture user inputs and display a confirmation message upon submission. 10
- Q6** A Develop a Node.js application that uses streams to read and write data efficiently. Describe the scenarios where using streams is preferable over other data-handling techniques. 10
 B Given a complex web application, propose a strategy using advanced React features like Refs and Hooks to manage component states across multiple layers. Explain the benefits of this approach. 10

[3 hrs]

[80 Marks]

- Note : 1. Question 1 is compulsory
2. Answer any three out of remaining questions
3. Assume suitable data where required

Q1 Solve any 4

- a) Describe the PEAS descriptor for AI agent-based Movie Ticket Booking System 5
b) Write the Environment properties of the Pacman Game 5
c) Describe an Intelligent Agent with a neat diagram. 5
d) Differentiate between supervised and unsupervised learning 5
e) Convert in to FOPL
 - EVERYONE LIKE EVERYONE
 - ALL GRADUATES ARE UNEMPLOYED

Q2

- a) Give the comparative analysis of BFS, DFS, Iterative Deepening, and Bidirectional Search Strategies with respect to Time Complexity, Space Complexity, Optimality, and Completeness 10
b) Describe the Hill Climbing algorithm with an example. Discuss its inherent limitations, and propose effective solutions to address those limitations 10

Q3

- a) Consider the following statements. 10
 (a) Ravi likes all kind of food.
 (b) Apple and Chicken are food.
 (c) Anything anyone eats and is not killed is food.
 (d) Ajay eats peanuts and still alive.
 (e) Rita eats everything that Ajay eats.

Prove that Ravi likes Peanuts using **Resolution**.

- b) Explain Total Order Planning and Partial Order Planning in detail. 10

Q4

- a) Explain Bayesian Belief Network with example 10
b) Define the initial and goal state of three missionaries and cannibals problem. Describe the set of operators using if-then rules.

Draw the entire state space graph (include only legal states, that is, states in which cannibals do not outnumber missionaries on either side of the river). State best searching algorithm for it

Q5

- a) Explain Genetic Algorithm in detail with suitable example. 10
b) Explain a heuristic function for an 8-puzzle problem and solve it using A* algorithm? 10

Q6

- a) Epidemiologists claim that the probability of breast cancer among Caucasian women in their mid-50s is 0.005. An established test identified people who had breast cancer and those that were healthy. A new mammography test in clinical trials has a probability of 0.85 for detecting cancer correctly. In women without breast cancer, it has a chance of 0.925 for a negative result. If a 55-year-old Caucasian woman tests positive for breast cancer, what is the probability that she, in fact, has breast cancer? 10
b) Explain reinforcement learning with example. 10
-

Time: 3 hours

Max. Marks: 80

N.B. (1) Question no one is Compulsory.

(2) Attempt any 3 questions out of the remaining.

(3) Assume suitable data if required.

Time: 3 hours

Max. Marks: 80

N.B. (1) Question one is Compulsory.

(2) Attempt any 3 questions out of the remaining.

(3) Assume suitable data if required.

- Q. 1 a) Explain CLARANS extension in web mining 05
b) Explain in detail the extract/transform/load (ETL) design of an automated warehouse. 05
c) What is prediction? Explain about Linear regression method. 05
d) Suppose data for clustering is {6,14,18,22,1,40,50,11,25}. Consider K=2 , Cluster the given data using K-means algorithm 05
- Q. 2 a) Briefly outline with example, how to compute dissimilarity between the objects describe following
i) Nominal attributes
ii)Asymmetric binary attributes 10
b) Discuss about a three-tier data warehouse architecture. 10
- Q. 3 a) Describe the various phases in knowledge discovery process with a neat diagram 10
b) Explain Decision tree induction algorithm for classification. Discuss the usage of information gain in this.
- Q. 4 a) Following table gives fat & protein content of the items. Apply single linkage clustering & dendogram 10

Food Item	Protein	Fat
1	1.1	60
2	8.2	20
3	4.2	35
4	1.5	21
5	7.6	15
6	2.0	55
7	3.9	39

b) Consider the following transactions

10

TID	ITEMS
01	1,3,4,6
02	2,3,5,7
03	1,2,3,5,8
04	2,5,9,10
05	1,4

Apply the Apriori algorithm with minimum support of 30% and minimum confidence of 75% and find large item set.

Q. 5 a) Explain Hyperlink Induced Topic Search Algorithm (HITS) Algorithm with example 10

b) What is market basket analysis? Explain with an example. State and explain with formula the meaning of the following term:

i) Support

ii) Confidence

10

Q. 6 a) Describe the working of K-medoid clustering with the help of sample dataset. 10

b) Define multidimensional and multilevel association mining 10
