Cochin University of Science and Technology
Department of Computer Science
21-805-0602 Machine Learning Algorithms
I Term, Date: Feb 11, 2025

| Timing: 09:00 AM to 11:00 AM | VI Semester | Max Marks: 20 |

Attempt ALL Questions

1. (a) Derive the Least Mean Squares (LMS) rule for a linear regression problem. Write the algorithm for training a dataset. **(Marks 3, CO2, DL 2, BLT 1,2)**

   (b) Use the dataset shown in the left to fit a linear regression model.

Weight (kg)	Height (ft)
63.5	5.0
70.3	5.2
72.1	5.6
87.1	5.9
90.7	6.0
96.2	6.3

   Show your calculations. **(Marks 2, CO2, DL 2, BLT 1,2)**

2. (a) When is linear regression inappropriate for a prediction task, and why is logistic regression preferred in those cases? Give examples to explain. **(Marks 2, CO1, DL 1, BLT 1,2)**

   (b) Derive the stochastic ascent rule for Logistic regression. **(Marks 3, CO2, DL 2, BLT 1,2)**

3. (a) Arrive to Bayes' theorem from first principles of Probability. How is it applied in Bayesian learning? Provide a simple example to illustrate its application. **(Marks 2, CO1, DL 1, BLT 1,2)**

   (b) You are given two coins. Coin A has a probability of heads of 0.6, and coin B has a probability of heads of 0.4. You randomly select a coin and flip it three times, observing two heads and one tail. Using Bayesian inference, determine the probability that you chose coin A. Assume a prior probability of 0.5 for each coin. **(Marks 3, CO1, CO2 DL 2, BLT 1,2)**

4. (a) Considering a set of sequences generated through tossing a coin, illustrate the idea of parametric learning. **(Marks 3, CO2, DL 1, BLT 1,2)**

   (b) Explain the concepts of bias and variance in machine learning. How do they relate to model complexity? **(Marks 2, CO1, DL 1, BLT 1,2)**

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: M.Sc. (Five Year Integrated) in Computer Science (Artificial Intelligence & Data Science)	
Course Code & Title: 21-805-0405 Optimization Techniques	

Name of Examination: Series I		Max. Marks: 20	Semester: VI
Batch: 2023-28	Duration: 2 Hours	Date: 14.02.2025	Time: 09:00 AM – 11 AM

Answer all questions

1. A company produces two types of products, A and B. Each product requires processing on two machines, M1 and M2. Product A requires 1 hour on M1 and 2 hours on M2, while product B requires 2 hours on M1 and 1 hour on M2. Machine M1 is available for a maximum of 16 hours per day. Machine M2 is available for a maximum of 12 hours per day. The profit per unit of product A is Rs 5, and the profit per unit of product B is Rs. 4. Formulate the linear programming problem to determine the daily production levels of products A and B that will maximize the total profit. [3 marks][CO2, DL-3, BTL-3]

2. Maximize
   $p=2u1+3u2$
   Subject to
   $u1+u2 \leq 4$
   $u1+2u2 \geq 2$
   $u1, u2 >= 0$ Solve using graphical model   [3 marks][CO2, DL-3, BTL-3]

3. Maximize $z=2x1+4x2+3x3$ subject to $x1+2x2+x3 \leq 12$, $2x1+x2+x3 \leq 9$, $8x1+4x2+3x3 \leq 33$ $x1, x2, x3 \geq 0$ using simplex method.   [6 marks][CO2, DL-3, BTL-3]

4. A company needs to transport goods from three factories to four destinations. The transportation costs are given in the table below: Determine the basic feasible solution using Northwest corner cell method and optimize the solution.   [5 marks][CO2, DL-3, BTL-3]

	D1	D2	D3	D4	SUPPLY
O1	3	1	7	4	250
O2	2	6	5	9	350
O3	8	3	3	2	400
DEMAND	200	300	350	150	

5. A software company has four expert programmers and needs to develop four application programmes. The head of the computer centre, estimates the programming time (in minutes) required by the respective experts to develop the application programmes as follows. Find the assignment pattern that minimises the time required to develop the application programmes   [3 marks][[CO2, DL-3, BTL-4]

	A	B	C	D
1	120	100	80	90
2	80	90	110	70
3	110	140	120	100
4	90	90	80	90

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: M.SC. (Five Year Integrated) In Computer Science (Artificial Intelligence & Data Science)/M.TECH. Computer Science And Engineering (Artificial Intelligence And Software Engineering)

Course Code & Title: 21-805-0604 Soft Computing Techniques/24-502-0208: Soft Computing

Name of Examination: Series II		Max. Marks: 20	Semester: VI/II
Batch: 2022-27/ 2024-26	Duration: 2 Hours	Date: 20.03.2025	Time: 9:00 AM – 11:00 AM

Course Outcome

Sl No	Description
CO1	Learn Fuzzy logic and its applications.
CO2	Understand the basic concepts of Artificial neural networks and its applications.
CO3	Solve single-objective optimization problems using GAs.
CO4	Solve multi-objective optimization problems using Evolutionary algorithms.
CO5	Apply Soft computing techniques to solve problems in various application domains.

Answer all questions

1. Design a neural network to perform the AND operation and demonstrate it as well.
   (8 marks) [CO2, DL-2, BTL-3]

2. Write the steps of order crossover in permutation representation. Given two parents: [1, 2, 3, 4, 5, 6, 7, 8, 9] and [5, 2, 7, 3, 6, 9, 1, 4, 8], perform order crossover and generate Child 1 and Child 2.
   (6 marks) [CO3, DL-2, BTL-2]

3. Explain three Arithmetic Recombination Operators for Floating-Point Representations with examples.
   (6 marks) [CO3, DL-2, BTL-2]

Cochin University of Science and Technology
Department of Computer Science
21-805-0602 Machine Learning Algorithms
II Term. Date: March 11, 2025

Timing: 09:00 AM to 11:00 AM	VI Semester	Max Marks: 20

Attempt ALL Questions

1. (a) What are the key differences between LDA and Principal Component Analysis (PCA) in terms of their objectives and methodologies?

    (Marks 2, CO3, DL 1, BLT 1,2)

    (b) Given the Fisher criterion, obtain the Fisher's linear discriminant as a solution to LDA.

    (Marks 3, CO3,CO4, DL 2, BLT 2)

2. (a) Consider a dataset with two features $(x1, x2)$ and three data points: Point A: (1, 2), Point B: (3, 4), Point C: (5, 3). Determine the Principal Component.

    (Marks 2, CO3,CO6, DL 1, BLT 1,2)

    (b) Describe the two main steps involved in LLE: (a) Constructing the neighborhood graph and computing the weights. (b) Computing the low-dimensional embedding.

    (Marks 3, CO3, DL 2, BLT 1,2)

3. (a) What is the concept of a hyperplane? How does it relate to maximizing the margin in SVMs?

    (Marks 2, CO3, DL 1, BLT 1,2)

    (b) What are kernels in SVM? List some of the kernel functions and explain.

    (Marks 3, CO4 DL 2, BLT 1,2)

4. (a) What are some potential advantages and disadvantages of using decision trees for classification?   (Marks 2, CO4, DL 1, BLT 1,2)

    (b) A small online retailer wants to predict whether a customer will make a purchase based on two features: Age: (Continuous, in years), Website visits: (Discrete, number of visits in the past month). Assuming a data set (Customer ID, Age, Number of Website Visits, Purchased (Yes/No)) with say 10 instances, illustrate how a CART tree is built.   (Marks 3, CO4, DL 2, BLT 1,2)

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: M.SC. (Five Year Integrated) In Computer Science (Artificial Intelligence & Data Science)/M.TECH. Computer Science And Engineering (Artificial Intelligence And Software Engineering)

Course Code & Title: 21-805-0604 Soft Computing Techniques/24-502-0208: Soft Computing

Name of Examination: Series I		Max. Marks: 20	Semester: VI/II
Batch: 2022-27/ 2024-26	Duration: 2 Hours	Date: 13.02.2025	Time: 9:00 AM – 11:00 AM

Course Outcome

Sl No	Description
CO1	Learn Fuzzy logic and its applications.
CO2	Understand the basic concepts of Artificial neural networks and its applications.
CO3	Solve single-objective optimization problems using GAs.
CO4	Solve multi-objective optimization problems using Evolutionary algorithms.
CO5	Apply Soft computing techniques to solve problems in various application domains.

Answer all questions

1. Consider a fuzzy set representing "temperatures close to 20°C". We define the membership function for this fuzzy set using a triangular membership function:

$$\mu_A(x) = \begin{cases} \frac{x-15}{5}, & \text{if } 15 \le x \le 20 \\ \frac{25-x}{5}, & \text{if } 20 < x \le 25 \\ 0, & \text{otherwise} \end{cases}$$

This function defines the degree of membership of a temperature x in the fuzzy set "close to 20°C".

   a. What is the membership value of the temperature 18°C in the fuzzy set A?
   b. What is the membership value of the temperature 22°C in the fuzzy set A?
   c. For the temperature 20°C, what is the membership value in the fuzzy set A?

   (3 marks) [CO1, DL-1, BTL-2]

2. Construct a membership function that defines a person's height (cm) as medium. Calculate membership value of the heights (cm) 150, 160, 168, 172, 180 as per your constructed membership function. (7 marks) [CO1, DL-1, BTL-2]

PTO

3. Consider two fuzzy relations R_1 and R_2 defined on the set of values representing temperature and fan speed, respectively. Let the fuzzy relation R_1 (between temperature and heating requirement) and R_2 (between heating requirement and fan speed) be defined as follows:

$$R_1 = \begin{bmatrix} 0.8 & 0.6 & 0.4 \\ 0.7 & 0.9 & 0.3 \\ 0.6 & 0.5 & 0.7 \end{bmatrix} \qquad R_2 = \begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.3 & 0.8 & 0.7 \\ 0.6 & 0.9 & 0.2 \end{bmatrix}$$

a. Compute the Max-Min composition of the two fuzzy relations R1 and R2.
b. Interpret the resulting fuzzy relation in terms of heating requirement and fan speed.

(6 marks) [CO1, DL-2, BTL-3]

4. Prove that the following function is a valid fuzzy union function:

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

(4 marks) [CO1, DL-2, BTL-3]

Cochin University of Science And Technology
Department of Computer Science

VI Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence & Data Science)

End Semester Examination - April 2025

21-805-0601 Inferential Statistics

Duration: 3 Hours

Maximum Marks: 50

Answer all questions. From each question fully answer either (a) or (b)

1) (a) i) Define the Central Limit Theorem for statistics mean.

[1 mark][CO1, DL-1, BTL-1]

ii) The population standard deviation is 35.2. Determine the minimum sample size required to estimate the population mean with the margin error of 0.1 and a 99% confidence level.

[5 marks][CO1, DL-2, BTL-3]

iii) A researcher claims that the average salary of a private school teacher is greater than 40,000 with a standard deviation of 7500. A sample of 35 teachers has a mean salary of 42000. Test the claim of the researcher at 0.05 level of significance.   [6 marks][CO1, DL-2, BTL-3]

OR

(b) i) A survey found that women over the age of 55 consume an average of 1660 calories per day. The standard deviation was also found to be 56 calories. The researcher sampled 43 women over the age of 55 and found that the mean number of calories they consume was 1446. At α = 0.10, can it be concluded that there is no significant difference between the number of calories consumed by women over age 55?   [6 marks][CO1, DL-2, BTL-3]

ii) A study shows that the average number of hours of TV viewing per household per week in the Philippines is 50.4 hours and the standard deviation is 11.8 hours. If a random sample of 42 households is taken, what is the probability that the sample average is less than 52 hours?

[6 marks][CO1, DL-1, BTL-3]

2) (a) Three types of fertilizers are used on three groups of plants for 5 weeks. We want to check if there is a difference in the mean growth of each group. Using the data given below apply suitable test at 0.05 significant level.

Fertilizer 1	6	8	4	5	3	4
Fertilizer 2	8	12	9	11	6	8
Fertilizer 3	13	9	11	8	7	12

[9 marks][CO2, DL-1, BTL-3]

OR

(b) A group of 5 patients treated with medicines "A" weights 42, 39, 48, 60 and 41 kgs. Second group of 7 patients treated with medicine "B" weights 38, 42, 56, 64, 68, 69 and 62 kgs. Do you agree with the claim that the medicine "B" increases the weight significantly at 0.05 significant level?   [9 marks][CO2, DL-1, BTL-3]

3) (a) i) It has been found that 85.6% of all enrolled college and university students in the United States are undergraduates. A random sample of 500 enrolled college students in a particular state revealed that 420 of them were undergraduates. Is there sufficient evidence to conclude that the proportion differs from the national percentage with significance level 0.05?

[6 marks][CO3, DL-2, BTL-3]

ii) For a class project, a political science student at a large university wants to estimate the percent of students who are registered voters. He surveys 500 students and finds that 300 are registered voters. Construct a 90% confidence interval for the percent of students who are registered voters.

[6 marks][CO3, DL-2, BTL-3]

OR

(b) i) Two samples of executives, 120 in Finance and 180 in Marketing are polled for their views (Agree/Disagree) on the question "Will Brexit be good for business?" 54 Finance executives agree and 63 Marketing executives agree. Conduct a test of proportion at the significance level 0.1 to see if there is evidence of a difference of opinion between the two sets of executives.

[6 marks][CO3, DL-2, BTL-3]

ii) A professor believes that a new online learning curriculum is increasing the median final exam score from the previous year, which was 75. A random sample of final exam scores were collected for students that went through the new curriculum is given below. Use the sign test to see if the new curriculum has any effect using α=0.05. Given that the critical value for n =11 at given level of two tailed test is 1.

78 100 75 64 87 80 72 91 89 70 82 76

[6 marks][CO4, DL-1, BTL-3]

4) (a) In the game rock-paper-scissors, Kenny expects to win, tie and lose with equal frequency. Kenny plays rock-paper-scissors often, but he suspected his own games were not following that pattern, so he took a random sample of 24 games and recorded their outcomes as 4, 13 and 7. Perform Chi-square test at 0.05 significant level.

[7 marks][CO4, DL-1, BTL-3]

OR

(b) 500 elementary school boys and girls are asked which is their favorite color: blue, green, or Pink as follows. Examine whether there is a relationship between gender and favourite color using chi-square test at significance level 0.05.

	Blue	Green	Pink
Boys	100	150	20
Girls	20	30	180

[7 marks][CO4, DL-1, BTL-3]

5) (a) i) There are 6 models of car A, B, C, D, E and F and the number of purchases who prefer particular models are 14, 12, 32, 32, 20 and 16 respectively. Determine whether the distribution of preferences same with the expected distribution under the assumption of equal preference, using a significance level of 0.05.

[7 marks][CO4, DL-2, BTL-3]

ii) Define Complete and Minimal Complete Class of Decision Rules. Give an example for each.

[3 marks][CO5, DL-1, BTL-1]

OR

(b) i) The lifetime of bulbs of two different brands Brand X is 80, 100, 90, 110, 125, 130 and 70 and brand Y is 100, 120, 80, 140, 130, 160, 115 and 120. Determine whether the two samples come from the same distribution at significance level 0.05.

[7 marks][CO4, DL-2, BTL-3]

ii) Explain how Bayes' rule and the Minimax rule are used in statistical decision-making using a suitable example.

[3 marks][CO5, DL-1, BTL-2]

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: VI Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence & Data Science)			
Course Code & Title: 21-805-0601 Inferential Statistics			
Name of Examination: Series II		Max. Marks: 20	Semester: VI
Batch: 2022-27	Duration: 2 Hours	Date: 10.03.2025	Time: 09:00 AM – 11:00 AM

Answer all questions

1.
  a) State Central limit theorem for statistics proportion. (1 mark) [CO3, DL-1, BTL-1]

  b) A recent report states that 48% of high school graduates are interested in a STEM (Science, Technology, Engineering, Math) discipline. What is the chance the proportion for a sample of 90 between 0.45 and 0.60? (2 marks) [CO3, DL-1, BTL-3]

  c) A study found that 73% of prekindergarten children ages 3 to 5 whose mothers had a bachelor's degree or higher were enrolled in early childhood care and education programs. How large a sample is needed to estimate the true proportion within 3% with 95% confidence? (2 marks) [CO3, DL-2, BTL-3]

2.
  a) Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes - they own cell phones. Using a 98% confidence level, compute a confidence interval estimate for the true proportion of adult residents of this city who have cell phones. (2 marks) [CO3, DL-2, BTL-3]

  b) A car manufacturer is trying to lower the amount of defects it has in its cars. Two distinct assembly procedures are used to build the cars. In a sample of 350 cars cossming off the line using the first procedure there were 28 with major defects. While in a sample of 500 cars coming off the second assembly line shows 32 cars with defects. Is the difference between the two assembly procedures significant at 95% confidence level? (3 marks) [CO3, DL-2, BTL-3]

3. Two sample polls of votes of two candidates A and B for a public office are taken, one from among the residents of rural areas. The results are given in the adjoining table. Examine whether the nature

of the area is related to voting preference in the elections using chi-square test at significance level 0.05.

Area	Votes for	
	A	B
Rural	620	380
Urban	550	450

(4 marks) [CO4, DL-2, BTL-3]

4. The following information summarized related to the classification of the men and women with respect to the age group. Determine whether the two samples come from the same distribution at significance level 0.05.

Age	21-22	23-24	25-26	27-28	29-30	31-32	33-34	35-36	37-38	39-40
Men	4	11	5	7	0	5	9	13	20	6
Women	7	4	1	11	12	4	2	4	8	9

(4 marks) [CO4, DL-2, BTL-3]

5. A travel agency which promotes a particular holiday resort, advertises that the median cost per day of a motel in the city is $50. A cautions traveler selects a random sample of 8 motels in that city and records the cost (in dollar) per day as follows:

52   51   49   50   53   52   48   47

Use the sign test, at 5% level of significance to test the travel agency's claim. Given that the critical value for n =8 at given level of two tailed test is 0.

(2 marks) [CO4, DL-2, BTL-3]

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: VI Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence & Data Science)

Course Code & Title: 21-805-0601 Inferential Statistics

Name of Examination: Series 1		Max. Marks: 20	Semester: VI
Batch: 2022-27	Duration: 2 Hours	Date: 10.02.2025	Time: 09:00 AM – 11:00 AM

Answer all questions

1. .

   a) State Central limit theorem. (1 mark) [CO1, DL-1, BTL-1]

   b) A survey found that the American family generates an average of 17.2 pounds of glass garbage each year. Assume the standard deviation of the distribution is 2.5 pounds. Find the probability that the mean of a sample of 55 families will be between 17 and 18 pounds.

   (2 marks) [CO1, DL-1, BTL-3]

   c) Suppose you want to determine the mean distance between to cells inside the human body. The margin of error (E) in this case is 0.01 micrometers. The population standard deviation is known to be 0.16 micrometers. Determine the minimum sample size required to estimate the population mean with the given error and a 95% confidence. (2 marks) [CO1, DL-2, BTL-3]

2. An auto company decided to introduce a new six cylinder car whose mean petrol consumption is claimed to be lower than that of the existing auto engine. It was found that the mean petrol consumption for the 50 cars was 10 km per litre. The standard deviation of population is 3.5 km per litre.

   a) Test at 5% level of significance, whether the claim of the new car petrol consumption is 9.5 km per litre on the average is acceptable. (3 marks) [CO2, DL-2, BTL-3]

   b) Construct 98% Confidence Interval for mean petrol consumption.

   (2 marks) [CO2, DL-2, BTL-3]

3. Suppose you want to see if patients with PTSD have higher than normal systolic blood pressure. You sample 25 patients and obtain a sample mean Systolic BP of 131.66 and a sample standard deviation of 22.5352 mm Hg. Using an alpha value of 0.05 is this observed mean greater than a 'normal' Systolic BP of 120 mm Hg? (3 marks) [CO2, DL-2, BTL-3]

4. Test the following new drug measure before and after. Are the means before and after are equal for alpha value 0.05.

Before	265	262	267	260	264	266
After	261	257	263	259	261	264

(3 marks) [CO2, DL-2, BTL-3]

5. Per-pupil costs (in thousands of dollars) for cyber charter school tuition for school districts in three areas are shown. Test the claim that there is a difference in means for the three areas using an appropriate test for alpha=0.05.

Area 1	6.2	9.3	6.8	6.1	6.7	7.5
Area 2	7.5	8.2	8.5	8.2	7	9.3
Area 3	5.8	6.4	5.6	7.1	3	3.5

(4 marks) [CO2, DL-2, BTL-3]

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: M.Sc. (Five Year Integrated) In Computer Science (Artificial Intelligence & Data Science)

Course Code & Title: 21-805-0604 Soft Computing Techniques

Name of Examination: End Semester	Max. Marks: 50	Semester: VI
	Date: 11.04.2025	Time: 9:30 AM – 12:30 PM
Batch: 2022-27	Duration: 3 Hours	

Course Outcome

Sl No	Description
CO1	Learn Fuzzy logic and its applications.
CO2	Understand the basic concepts of Artificial neural networks and its applications.
CO3	Solve single-objective optimization problems using GAs.
CO4	Solve multi-objective optimization problems using Evolutionary algorithms.
CO5	Apply Soft computing techniques to solve problems in various application domains.

Answer all questions

1.
   a. Explain the concept of computing and differentiate between hard computing and soft computing. Discuss the key characteristics of soft computing and how it overcomes the limitations of traditional computing approaches. Provide suitable examples to support your answer. (10 marks) [CO5, DL-1, BTL-2]

OR

   b. List any four applications of soft computing and briefly explain how soft computing techniques are used in each application. (10 marks) [CO5, DL-2, BTL-3]

2.
   a. Explain the concept of fuzzy composition in fuzzy set theory. Discuss its significance in fuzzy relational modeling, particularly in decision-making systems. Illustrate your explanation with an example, such as a student performance evaluation system, where fuzzy relations between students, subjects, and grades are used to assess overall performance. (10 marks)[CO1, DL-2, BTL-3]

OR

   b. Explain the concept of a Fuzzy Expert System (FES) and describe its key components. Provide an example to illustrate its application. (10 marks)[CO1, DL-2, BTL-3]

PTO

3.

a. Explain the working principles of Evolutionary Algorithms (EAs) and their importance solving real-world optimization problems. Discuss the roles of selection, crossover, mutation and fitness evaluation in the evolutionary process. Provide a suitable example to illustrate it application. (10 marks)[CO3, DL-2, BTL-3]

OR

b. Explain the Eight-Queens Problem and discuss how Evolutionary Algorithms (EAs) can used to find a solution. Describe the representation, fitness function, selection, crossover, and mutation processes involved in solving this problem using evolutionary computing. (10 marks)[CO3, DL-2, BTL-3]

4.

a. Differentiate between single-objective and multi-objective optimization problems. Discuss their significance in real-world applications and explain the key challenges associated with multi-objective optimization. Provide examples of each type to illustrate your answer. (10 marks)[CO4, DL-2, BTL-2]

OR

b. Write a short note on Pareto-based optimization and explain the concept of the Multi-Objective Genetic Algorithm (MOGA) with a flowchart. (10 marks)[CO4, DL-2, BTL-3]

5.

a. Explain the fundamental concepts of Artificial Neural Networks (ANNs) and their significance in solving complex computational problems. Discuss the key components of an ANN, including neurons, activation functions, layers, and learning algorithms. Provide an example to illustrate its application. (10 marks)[CO2, DL-2, BTL-2]

OR

b. Design a multi-layer neural network (MLP) to solve the XOR problem. Describe the network structure, including the number of layers and activation functions. Perform one iteration of forward and backward propagation using back propagation, showing the necessary calculations for weight updates. (10 marks)[CO2, DL-2, BTL-3]

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE

Programme: M.Sc. (Five Year Integrated) In Computer Science (Artificial Intelligence & Data Science)
Course Code & Title: 21-805-0604 Soft Computing Techniques

Name of Examination: End Semester	Max. Marks: 50	Semester: VI
Batch: 2022-27	Duration: 3 Hours	Date: 11.04.2025	Time: 9:30 AM – 12:30 PM

Course Outcome

Sl No	Description
CO1	Learn Fuzzy logic and its applications.
CO2	Understand the basic concepts of Artificial neural networks and its applications.
CO3	Solve single-objective optimization problems using GAs.
CO4	Solve multi-objective optimization problems using Evolutionary algorithms.
CO5	Apply Soft computing techniques to solve problems in various application domains.

Answer all questions

1.
   a. Explain the concept of computing and differentiate between hard computing and soft computing. Discuss the key characteristics of soft computing and how it overcomes the limitations of traditional computing approaches. Provide suitable examples to support your answer. (10 marks) [CO5, DL-1, BTL-2]

   OR

   b. List any four applications of soft computing and briefly explain how soft computing techniques are used in each application. (10 marks) [CO5, DL-2, BTL-3]

2.
   a. Explain the concept of fuzzy composition in fuzzy set theory. Discuss its significance in fuzzy relational modeling, particularly in decision-making systems. Illustrate your explanation with an example, such as a student performance evaluation system, where fuzzy relations between students, subjects, and grades are used to assess overall performance. (10 marks)[CO1, DL-2, BTL-3]

   OR

   b. Explain the concept of a Fuzzy Expert System (FES) and describe its key components. Provide an example to illustrate its application. (10 marks)[CO1, DL-2, BTL-3]

3.
a. Explain the working principles of Evolutionary Algorithms (EAs) and their importance in solving real-world optimization problems. Discuss the roles of selection, crossover, mutation, and fitness evaluation in the evolutionary process. Provide a suitable example to illustrate its application.
(10 marks)[CO3, DL-2, BTL-3]

OR

b. Explain the Eight-Queens Problem and discuss how Evolutionary Algorithms (EAs) can used to find a solution. Describe the representation, fitness function, selection, crossover, and mutation processes involved in solving this problem using evolutionary computing.
(10 marks)[CO3, DL-2, BTL-3]

4.
a. Differentiate between single-objective and multi-objective optimization problems. Discuss their significance in real-world applications and explain the key challenges associated with multi-objective optimization. Provide examples of each type to illustrate your answer.
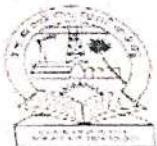(10 marks)[CO4, DL-2, BTL-2]

OR

b. Write a short note on Pareto-based optimization and explain the concept of the Multi-Objective Genetic Algorithm (MOGA) with a flowchart.
(10 marks)[CO4, DL-2, BTL-3]

5.
a. Explain the fundamental concepts of Artificial Neural Networks (ANNs) and their significance in solving complex computational problems. Discuss the key components of an ANN, including neurons, activation functions, layers, and learning algorithms. Provide an example to illustrate its application.
(10 marks)[CO2, DL-2, BTL-2]

OR

b. Design a multi-layer neural network (MLP) to solve the XOR problem. Describe the network structure, including the number of layers and activation functions. Perform one iteration of forward and backward propagation using back propagation, showing the necessary calculations for weight updates.
(10 marks)[CO2, DL-2, BTL-3]

Cochin University of Science And Technology

Department of Computer Science

VI Semester M.Sc(Five Year Integrated) in Computer Science (Artificial Intelligence & Data Science)

End Semester Examination - April 2025

21-805-0603: Feature Engineering

Duration: 3 Hours   Maximum Marks: 50

Answer all questions. From each question fully answer either (a) or (b)

1. (a) (i) A company trained a deep learning model for customer sentiment analysis, but the model performs well on training data but poorly on real-world reviews. What problem is the model facing? Suggest the techniques to improve its performance.

[5 marks][CO1, DL-2, BTL-3]

(ii) Consider the binary classification problem where the model predicts whether a patient has a disease (Positive) or not (Negative). The confusion matrix is:

Actual/Predicted	Positive	Negative
Positive	50	30
Negative	10	110

Compute accuracy, precision, recall, and F1-score. [5 marks][CO1, DL-2, BTL-3]

OR

(b)(i) Explain the key steps and techniques involved in feature engineering for machine learning models, including how to transform raw data into meaningful predictors to improve model performance.[4 marks][CO1, DL-1, BTL-1]

(ii) A regression model predicts house prices (in $1000s) based on features like size and location. Given:

Actual Prices (y): [250, 300, 400, 500, 350],
Predicted Prices (ŷ): [260, 290, 410, 490, 340]
Number of samples (n):5
Calculate MAE, MSE, RMSE, Relative MSE, and Coefficient of Variation
[6 marks][CO1, DL-3, BTL-3]

2. (a) (i) Given a 2-class dataset with the following samples:
Class 1 (C₁): (2, 3), (3, 4), (4, 5)
Class 2 (C₂): (6, 2), (7, 3), (8, 4)
Compute the LDA projection vector (direction of maximum class separation) Project the data onto this vector [6 marks][CO2, DL-2, BTL-3]

(ii) What is regularization in Deep Learning? Why is it important? Explain different regularization techniques [4 marks][CO2, DL-1, BTL-2]

OR

(b) (i)Consider the Data Values and Find the outlier using Modified Z-score Method
29, 46, 26, 26, 22, 24, 19, 16, 16, 15, 14, 12, 8, 7, 7, 6.

[7 marks][CO2, DL-2, BTL-2]

(ii) A dataset contains categorical variables with missing values. What Imputation techniques would you suggest? [3 marks][CO2,DL-1,BTL-3]

3. (a)(i) Outline the key phases involved in the machine learning lifecycle and explain what occurs at each stage? [4 marks][CO3, DL-2, BTL-2]

(ii) Given a dataset with correlated features, reduce its dimensionality using PCA. Consider a 2D dataset with 4 samples:

$$X = \begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \\ 5 & 6 \end{bmatrix}$$

[6 marks][CO3, DL-2, BTL-2]

OR

(b) (i) How do different feature selection methods impact model performance metrics? Explain the feature selection methods with examples [5 marks][CO3, DL-1, BTL-2]

(ii) How does dimensionality reduction mitigate the curse of dimensionality, and what are common techniques used for this purpose? [5 marks][CO3, DL-2, BTL-1]

4. (a) (i) What is meant by Normalization of Data. Discuss various normalization techniques with examples. [5 marks][CO4, DL-1, BTL-2]

(ii) Consider the following dataset representing the ages of people in a group: 22, 25, 27, 28, 29, 30, 31, 34, 36, 38, 40, 50, 60 . Find whether there is an outlier present in the given dataset using Inter-Quartile Range. [5 marks][CO4, DL-2, BTL-2]

OR

(b)(i) Define Local Outlier Factor Method. Briefly explain the steps to calculate LOF.
[4 marks][CO4, DL-2, BTL-3]

(ii) What is the role of Instance-Based Learning in recommendation systems like Netflix or Amazon? [6 marks][CO4, DL-2, BTL-3]

5. (a) (i) Describe Kernel-Induced Feature Expansion and Discuss Various Types of Kernels.
[6 marks][CO5, DL-2, BTL-3]

(ii) A social media platform wants to automatically detect hate speech in comments. Would traditional ML work, or is deep learning required? Why?

[4 marks][CO5, DL-2, BTL-3]

OR

(b) (i) Define Automated Feature engineering and Elaborate on its tools.
[5 marks][CO5, DL-2, BTL-1]

(ii) Explain the Convolution Neural Network Architecture. Describe the layers in detail.
[5 marks][CO5, DL-2, BTL-2]

Cochin University of Science and Technology
Department of Computer Science

VI Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence & Data Science)
End Semester Examination – April 2025
Regular and Supplementary
21-805-0605: Parallel Computing

Duration: 3 Hours

Maximum Marks: 50

Answer all questions. From each question fully answer either (a) or (b)

1. (a) i) Explain the concept of hybrid architectures in parallel computing. Discuss the benefits of hybrid architectures. Provide examples of hybrid systems and analyze their suitability for different types of applications.

   [6 marks] [CO1, DL-1, BTL-1]

   ii) Compare and contrast SIMD and MIMD architecture giving examples of their application.

   [4 marks] [CO1, DL-1, BTL-2]

   OR

   (b)i)Describe the distributed memory architecture and advantages for large-scale parallel computing. Discuss the communication mechanisms used in distributed memory systems.

   [5 marks] [CO1, DL-1, BTL-2]

   ii) Analyze the limits of parallel programming, including Amdahl's Law and Gustafson's Law. Discuss the various costs associated with developing and running parallel applications.

   [5 marks] [CO1, DL-1, BTL-2]

2. (a) i) Explain the key principles of designing efficient parallel algorithms. Discuss techniques for partitioning problems, managing communication, and ensuring load balancing.

   [7 marks] [CO2, DL-1, BTL-2]

   ii) Explain the challenges that I/O operations pose to parallel programs. What are some solutions for handling I/O in parallel environments?

   [3 marks] [CO2, DL-1, BTL-2]

   OR

   (b)i) Explain the threads model of parallel programming. How does it differ from the traditional process model? What are the advantages of using threads?

   [6 marks] [CO2, DL-2, BTL-3]

ii) How does granularity affect the performance of parallel programs? Discuss the trade-offs between fine-grain and coarse-grain parallelism.

[4 marks] [CO2, DL-2, BTL-3]

3. (a)i) Explain the parallel algorithm for Odd-Even Transposition Sort. Analyze the time complexity of the algorithm.

[6 marks] [CO3, DL-3, BTL-3]

ii) Explain the purpose of synchronization in parallel programming. Describe Barrier and Lock synchronization types.

[4 marks] [CO3, DL-3, BTL-3]

OR

(b) i) Explain the parallel prefix computation problem, its applications, and how it can be solved using a binary tree. Analyze the time and processor complexity of this parallel prefix sum algorithm.

[6 marks] [CO3, DL-1, BTL-1]

ii)Discuss factors affecting communication efficiency in parallel programs. How does asynchronous communication improve performance?

[4 marks] [CO3, DL-3, BTL-3]

4. (a)i) Compare and contrast the OpenMP and Message-Passing (MPI/PVM) models for parallel programming. Discuss the concept of dynamic process creation in each model.

[6 marks] [CO4, DL-1, BTL-2]

ii) Compare and contrast the packaging approaches of compact and slack clusters, highlighting their distinguishing features.

[4 marks] [CO4, DL-1, BTL-2]

OR

(b)i) Discuss the concept of a "single-system image" in the context of computer clusters. Why is achieving a single-system image considered a challenging task, and what are the advantages of presenting a cluster as a unified resource?

[6 marks] [CO4, DL-3, BTL-3]

ii) Discuss the different connection models for nodes in a cluster, detailing their characteristics and use cases.

[4 marks] [CO4, DL-3, BTL-3]

5. (a) Describe the CUDA programming model, with a detailed explanation of its thread hierarchy. Discuss how the SIMT execution model impacts kernel design and performance

[10 marks] [CO4, DL-3, BTL-3]

OR

(b) Detail the different types of memory available in CUDA. Discuss the characteristics of each type in terms of speed, scope, and usage.

[10 marks] [CO4, DL-1, BTL-3]