USN

Internal Assessment Test II- March 2025

Sub:	**Machine learning-1**					Sub Code:	**BCS602**	Branch:	CSE- AIML
Date:	**23/05/2025**	Duration:	90min's	Max Marks:	50	Sem/Sec:	VI CSE-AIML		

Answer any FIVE FULL Questions — MARKS

1a) Differentiate between the local weighted regression with linear regression — 5

Scheme:
5 differences 1 marks each

Solution

Linear Regression	Local Weighted Regression (LWR)
Global model	Local model (varies across input space)
Fits a single line (or hyperplane) to the entire dataset	Fits a model around a specific query point using nearby data
All data points treated equally	Nearby points get higher weights; distant points are down-weighted
Fast and efficient (once trained)	More expensive, as it recalculates weights per prediction
Best when the relationship is approximately linear globally	Best when data shows local patterns or non-linearity

1.(b) For the given dataset given below compute the entropy and information gain — 5

Row	Age	BuysComputer
1	<=30	No
2	<=30	No
3	31–40	Yes
4	>40	Yes
5	>40	Yes

Scheme:-

Entropy Calculation-2.5Marks

Information gain-2.5Marks

Solution:

Entropy=.971

	$Entropy(S) = -\left(\frac{3}{5}\log_2\frac{3}{5} + \frac{2}{5}\log_2\frac{2}{5}\right)$ $\approx -(0.6 \cdot -0.737 + 0.4 \cdot -1.322) = 0.971$ Information Gain =.971 $Gain(S, Age) = Entropy(S) - Entropy_{Age} = 0.971 - 0 = \boxed{0.971}$	
2(a)	Explain how the continuous attributes are discretized. Scheme: 4 steps=4*1=4 Example=1Mark Solution: 1.Sort the elements 2.Remove the duplicates 3.Compute Gain 4.Based on the maximum Gain <=,> is categorised Example	5
2(b)	Explain the advantages and disadvantages of decision trees. Scheme: Advantages 2.5 Marks Disadvantages 2.5Marks	5
3(a)	What is the need of an activation function?Explain non linear activation functions. Scheme : Need of activation function-2Marks Types of Non linear activation function-6MArks Solution: To add non linearity To activate neurons Types:Sigmoid, Tanh,RELU and Softmax **Sigmoid Function** $\sigma(x) = \frac{1}{1+e^{-x}}$ Range: (0, 1) Shape: S-shaped curve Use case: Binary classification (e.g., logistic regression output) **Tanh (Hyperbolic Tangent)** $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ Range: (-1, 1) Shape: S-shaped but centered at 0 Use case: When negative outputs are useful (e.g., hidden layers)	8

ReLU (Rectified Linear Unit)

$$f(x) = \max(0, x)$$

Range: [0, ∞)

Shape: Linear for positive inputs, zero for negative

Softmax Function

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

Range: (0, 1), sum of outputs = 1

Use case: Output layer for multi-class classification

3(b)	What is the XOR problem in perceptron and how to overcome it? Scheme: Xor problem 1 Mark Solution for XOR 1 MArk Solution: XOR is not linearly seperable and single perceptron cannot solve this.We need multilayer perceptron.	2
4a)	Differentiate between clustering and classification. Scheme: Any 4 differences *1=4 Solution:	4

Clustering	Classification
Unsupervised Learning	Supervised Learning
No predefined labels; the algorithm groups data by similarity	Predefined labels are provided; model learns from them
Groups or clusters of similar items	Specific class labels for each input
Discover inherent structure or patterns in data	Predict the correct class for new, unseen data
K-Means, DBSCAN, Hierarchical clustering	Decision Trees, SVM, Random Forest, Neural Networks

4b)	Explain the components of Reinforcement Learning. Scheme: 6 components*1=6 Solution: Write about each component 1.Environment 2.State and actions 3.Episode 4.Policy 5.Reward 6.Reinforcement Problems	6

5a)	Analyze the student performance using Naive Bayes algorithm for continuous attribute. Predict whether a student will get a job offer or not in the final year by taking CGPA=8.9, Interactiveness='Yes'.	

S.No.	CGPA	Interactiveness	Job Offer
1.	9.5	Yes	Yes
2.	8.2	No	Yes
3.	9.3	No	No
4.	7.6	No	No
5.	8.4	Yes	Yes
6.	9.1	Yes	Yes
7.	7.5	Yes	No
8.	9.6	No	Yes
9.	8.6	Yes	Yes
10.	8.3	Yes	Yes

Scheme:

Frequency matrix, mean and standard deviation-3Marks

Test instance probability-1 Mark

Posterior Probabillity-2Marks

Likelihood Interactiveness Yes Yes=5/7 No=⅓

No Yes=2/7 No=⅓

Mean and S.D for CGPA JF=Yes 8.814 .538

Mean and S.D for CGPA JF=No 8.133 .825

Test Instance Guassian Yes=.732,No=.313

Posterior Yes=.366 No=.0313

Job offer=Yes

5b)	Explain prior, posterior and likelihood probabilities with the help of an example Scheme: Prior, Posterior and Likelihood 1Mark each Example 1Mark Solution: Prior:- it is the general probability of an uncertain event before an observation is seen or some evidence is collected. Likelihood Probability:- Relative probability of the observation occurring for each class for the evidence given the hypothesis. Posterior-Probability of an event after considering new evidence/data	4
6a)	Find cosine similarity , SMC and Jaccard coefficient for the following data (1011) (1100) Scheme: Each method*1=3	3

Solution:

1. Cosine Similarity

$$\text{Cosine Similarity} = \frac{A \cdot B}{\|A\|\|B\|}$$

Dot Product (A · B):

$$(1 \cdot 1) + (0 \cdot 1) + (1 \cdot 0) + (1 \cdot 0) = 1 + 0 + 0 + 0 = 1$$

Magnitude of A:

$$\|A\| = \sqrt{1^2 + 0^2 + 1^2 + 1^2} = \sqrt{3}$$

Magnitude of B:

$$\|B\| = \sqrt{1^2 + 1^2 + 0^2 + 0^2} = \sqrt{2}$$

$$\text{Cosine Similarity} = \frac{1}{\sqrt{3} \cdot \sqrt{2}} = \frac{1}{\sqrt{6}} \approx 0.408$$

2.**Simple Matching Coefficient (SMC)**

SMC is the proportion of matching attributes (both 1s and 0s):

- **Matches:**
  - 1st bit: 1 == 1 → match
  - 2nd bit: 0 ≠ 1 → no match
  - 3rd bit: 1 ≠ 0 → no match
  - 4th bit: 1 ≠ 0 → no match

Only 1 match out of 4:

7.

$$\text{SMC} = \frac{\text{Number of Matches}}{\text{Total Number of Attributes}} = \frac{1}{4} = 0.25$$

$$\text{Jaccard} = \frac{1}{4} = 0.25$$

3.

6b)	Solve using Single linkage or MIN algorithm clustering algorithm, Object X Y A 3 5 B 7 8 C 12 5 D 16 9 Scheme Each iteration -3 MArks each 3*2=6 Final cluster -1 MArk Solution:	7

Compute Pairwise Euclidean Distances

The Euclidean distance between two points (x1,y1)(x_1, y_1)(x1,y1) and (x2,y2)(x_2, y_2)(x2,y2) is:

distance=(x2−x1)2+(y2−y1)2\text{distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}distance=(x2−x1)2+(y2−y1)2

Now calculate the distances:

- **AB** = √((7−3)² + (8−5)²) = √(16 + 9) = √25 = **5.0**
- **AC** = √((12−3)² + (5−5)²) = √(81 + 0) = √81 = **9.0**
- **AD** = √((16−3)² + (9−5)²) = √(169 + 16) = √185 ≈ **13.6**
- **BC** = √((12−7)² + (5−8)²) = √(25 + 9) = √34 ≈ **5.83**
- **BD** = √((16−7)² + (9−8)²) = √(81 + 1) = √82 ≈ **9.06**
- **CD** = √((16−12)² + (9−5)²) = √(16 + 16) = √32 ≈ **5.66**
- **Closest distance is 5.0 (A-B) → Merge {A} and {B}**

Clusters now: {AB}, {C}, {D}

Iteration 2: Find minimum distance between clusters

- **AB–C: Use minimum of distance(AC)=9.0 and distance(BC)=5.83 → MIN = 5.83**
- **AB–D: distance(AD)=13.6, distance(BD)=9.06 → MIN = 9.06**
- **C–D: 5.66**

Minimum is 5.66 (C–D) → Merge {C} and {D}

Clusters now: {AB}, {CD}

Final{A,B,C,D}

Faculty Signature CCI Signature HOD Signature