RV UNIVERSITY Go, change the world an initiative of RV EDUCATIONAL INSTITUTIONS	USN **School of Computer Science and Engineering** **B.Tech. (Hons.)** **Continuous Internal Evaluation (CIE-2)** **Evaluation Scheme** **Academic Year 2024-2025**		
Course: Introduction to Machine Learning		**Course Code: CS2213**	**Semester: IV**
Time: 9:30 AM to 11:00 AM	**Duration: 90 minutes**	**Date: 18-03-2025**	**Max Marks: 25**

Notes/ Instructions:

a) **Answer all questions**
b) **Write answers to all the questions in the answer booklet**

Sl. No.	PART A – (MCQs/Short answers) Max Marks (5)	Marks	L1-L6	CO
1.	Which of the following is a hyperparameter the affects the performance of decision tree classifier? a). learning rate b). maximum depth of decision tree c). batch size d). number of estimators **Answer b)**	1	L2	CO2
2.	Which of the following is true about principal components? i) They are linear combinations of the original features ii) They are not orthogonal to each other iii) The first principal component captures the most variance a) i & ii only b) ii only c) i & iii only d) i, ii, & iii **Answer c)**	1	L2	CO3
3.	Calculate the number of parameters present in a perceptron (step function) which takes input feature of dimension 2. **Answer 3**	1	L2	CO4
4.	An MLP has 3 input features, 1 hidden layer with 4 neurons, and 1 output neuron. Total number of trainable parameters **Answer 21**	1	L2	CO4
5.	Which of the following is true about MLPs and FNNs? a) All MLPs are FNNs, but not all FNNs are MLPs b) All FNNs are MLPs, but not all MLPs are FNNs c) MLPs and FNNs are completely different architectures d) MLPs are used only for unsupervised learning **Answer a)**	1	L2	CO4

L1- Remember, L2- Understand, L3 – Apply, L4- Analyse, L5 – Evaluate, L6 – Create

<table>
<tr><th>Sl. No.</th><th>PART B – Max Marks (20)</th><th>Marks</th><th>L1-L6</th><th>CO</th></tr>
<tr><td>2</td><td>
Given the below confusion matrix for 3-class classification problem

Actual / Predicted	Apple	Orange	Banana
Apple	254	74	96
Orange	28	317	33
Banana	12	46	280

Find **a.** False positives for Orange - 1 Mark
b. Recall for Banana - 2 Marks
c. Accuracy for Apple - 2 Marks
Answer:
False positives for orange = 28+33 = 61 --------- 1 Mark
Recall for Banana --------- 2 Marks
TPs for banana = 280
FNs for banana = 96+33 = 129
Recall = TP/(TP+FN) = 0.6845
Accuracy for Apple ------------ 2 Marks
TPs for Apple = 254
FPs for Apple = 74+96 = 170
FNs for Apple = 28+12 = 40
TNs for Apple = 317+33+46+280 = 676
Accuracy = (TP+TN)/(TP+FP+FN+TN) = 0.8157
</td><td>5</td><td>L3</td><td>CO2</td></tr>
<tr><td>3</td><td>
Consider the dataset the below dataset in which last variable is dependant variable. Find the best splitting attribute (feature) at level-0 (root node) using Gini Index criterion. You are required to specify Gini index of each attribute (feature) explicitly.

Weather	Temperature	Humidity	Wind	Play?
Rainy	Cool	Normal	Strong	No
Sunny	Mild	High	Weak	No
Sunny	Hot	High	Strong	Yes
Cloudy	Hot	High	Strong	Yes

Answer:
Gini index = $1 - \sum_{i=1}^{C} (P_i)^2$
Gini index for Weather (Students are required to provide the three splits) -------- 1 Mark
Gini(Weather=Rainy) = 0
Gini(Weather=Sunny) = 0.5
Gini(Weather=Cloudy) = 0
Gini(Weather) = 0.25

Gini index for Temperature (Students are required to provide the three splits) -------- 1 Mark
Gini(Temperature=Cool) = 0
</td><td>5</td><td>L3</td><td>CO2</td></tr>
</table>

L1- Remember, L2- Understand, L3 – Apply, L4- Analyse, L5 – Evaluate, L6 – Create

Gini(Temperature =Mild) = 0
Gini(Temperature =Hot) = 0
Gini(Temperature) = 0

Gini index for Humidity (Students are required to provide the two splits) -------- 1 Mark
Gini(Humidity=Normal) = 0
Gini(Humidity =High) = 0.4444
Gini(Humidity) = 0.333

Gini index for Wind (Students are required to provide the two splits) -------- 1 Mark
Gini(Wind=Strong) = 0.4444
Gini(Wind =Weak) =0
Gini(Wind) = 0.333

Choose Temperature as a best splitting attribute a level-0 since it has lowest Gini index value. -------- 1 Mark

4

Cluster the data points (with (x, y) representing locations) into two clusters {(2, 10), (2, 5), (8, 4), (5, 8), (7, 5))}. Employ the k-means clustering method by taking the seed mean points (i.e., the initial centroids) as the first two points, i.e., (2, 10) and (5, 8). At each stage (iteration) of the method give the centroids. Find the final centroids of the clusters and cluster members. Euclidian distance can be used for finding distance.

Iteration 1: 2 Marks

Samples	Distance from C1 (2,10)	Distance from C2 (5,8)	Cluster Assignment
(2,10)	0	√13	C1
(2,5)	5	√18	C2
(8,4)	√64	√25	C2
(5,8)	√13	0	C2
(7,5)	√50	√13	C2

Updated cluster centres:
C1 = (2,10)
C2 = (5.5,5.5)

Iteration 2: 2 Marks

Samples	Distance from C1 (2,10)	Distance from C2 (5.5,5.5)	Cluster Assignment
(2,10)	0	√32.5	C1
(2,5)	5	√12.5	C2
(8,4)	√64	√8.5	C2
(5,8)	√13	√6.5	C2
(7,5)	√50	√2.5	C2

5

L3

CO3

We can stop K-means clustering algorithm since there is no change in the cluster reassignment.
Final Cluster centroids C1 = (2,10) and C2 =(5.5,5.5) ----- 1 Mark

5 | Marks: 5 | L3 | CO4

Consider two class classification problem. Let the training set be $D = \{(X_1, t_1), (X_2, t_2), (X_3, t_3), (X_4, t_4)\}$ where each element (X, t) consists of the feature vector $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix}$, along with its given class-label t. Let $(X_1, X_2, X_3, X_4) = \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}, \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}\right)$, and $(t_1, t_2, t_3, t_4) = (+1, +1, +1, -1)$.
We want to learn the Perceptron classifier. Let the solution we are trying to get be f(X) such that $f(X) = w_0 + w_1x_1 + w_2x_2$. If $f(X) \geq 0$, we classify X to the +1 class, otherwise we classify X to the −1 class.
Consider the solution in parametric form be $W = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}$ and let the initial solution be $W_{init} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$. Let the objective function that we are trying to minimize be $J(W) = \frac{1}{2}\sum_{i=1}^{4}(t_i - o_i)^2$ where t_i is the given class label of X_i and $o_i = f(X_i)$.

a. Find $J(W_{init})$

b. Find $\nabla J(W_{init})$

Note: o_i represents predicted value not predicted label.

Answer Scheme:

Find predicted value for each sample

Sample	Actual Label (t_i)	Predicted value (o_i)
$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$	+1	-2
$\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$	+1	-2
$\begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$	+1	-2
$\begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$	-1	-2

a. $J(W_{init})$ = 14 **------ 2 Marks**

b. $$\nabla J(W) = \begin{bmatrix} -\sum_{i=1}^{4}\{t_i - (w_0 + w_1x_1 + w_2x_2)\} \\ -\sum_{i=1}^{4}\{t_i - (w_0 + w_1x_1 + w_2x_2)\}x_1 \\ -\sum_{i=1}^{4}\{t_i - (w_0 + w_1x_1 + w_2x_2)\}x_2 \end{bmatrix}$$

$$\nabla J(W\_init) = \begin{bmatrix} -10 \\ -4 \\ -4 \end{bmatrix}$$ -------- 3 Marks

Course Outcomes

1. Understand various types of machine learning algorithms and the role of data preprocessing in machine learning
2. Evaluate regression and classification model's performance on real-time datasets
3. Apply unsupervised learning algorithms for pattern discovery and structural analysis in datasets
4. Build Multilayer Perceptron to perform classification
5. Perform image classification using Convolutional Neural Networks
6. Design and implement machine learning solutions to solve a real-world problem through a guided or open-ended project.

Marks Distribution										
L1	L2	L3	L4	L5	L6	CO1	CO2	CO3	CO4	CO5
	5	20					11	6	8	

L1- Remember, L2- Understand, L3 – Apply, L4- Analyse, L5 – Evaluate, L6 – Create