



**School of Computer Science Engineering and Information Systems**

**Winter Semester 2024-2025**

**Continuous Assessment Test – II**

**Programme Name & Branch: MCA**

**Course Name & code: PMCA507L – Machine Learning**

**Class Number (s): 3328 and 3334**

**Faculty Name (s): Dr. Parimala M & Dr. Anitha A**

**Exam Duration: 90 Min.**

**Maximum Marks: 50**

**General instruction(s): Answer All Questions**

**CO – Course Outcome; BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyze, 5 – Evaluate, 6 – Create)**

**Course Outcomes:**

**CO2: Provide solutions for classification, regression, and clustering approaches in real world applications**

**CO3: Gain knowledge to combine machine learning models to achieve better results**

Q.No	Question	Max Marks	CO	BL																																												
1.	<table><tr><th>Age</th><th>Likes height</th><th>Likes goats</th><th>Go rock climbing</th></tr><tr><td>23</td><td>0</td><td>0</td><td>0</td></tr><tr><td>31</td><td>1</td><td>1</td><td>1</td></tr><tr><td>35</td><td>1</td><td>0</td><td>1</td></tr><tr><td>35</td><td>0</td><td>0</td><td>0</td></tr><tr><td>42</td><td>0</td><td>0</td><td>0</td></tr><tr><td>43</td><td>1</td><td>1</td><td>1</td></tr><tr><td>45</td><td>0</td><td>1</td><td>0</td></tr><tr><td>46</td><td>1</td><td>1</td><td>1</td></tr><tr><td>46</td><td>1</td><td>0</td><td>0</td></tr><tr><td>51</td><td>1</td><td>1</td><td>1</td></tr></table> <p>Construct a Decision Tree for the above dataset. Use Information gain to split the tree for one iteration. Note: Convert the values of “Age” attribute into two class labels such as “&gt;=40.5” and “&lt;40.5”</p>	Age	Likes height	Likes goats	Go rock climbing	23	0	0	0	31	1	1	1	35	1	0	1	35	0	0	0	42	0	0	0	43	1	1	1	45	0	1	0	46	1	1	1	46	1	0	0	51	1	1	1	10	CO 2	BL 3
Age	Likes height	Likes goats	Go rock climbing																																													
23	0	0	0																																													
31	1	1	1																																													
35	1	0	1																																													
35	0	0	0																																													
42	0	0	0																																													
43	1	1	1																																													
45	0	1	0																																													
46	1	1	1																																													
46	1	0	0																																													
51	1	1	1																																													

$$\textcircled{1} \text{Info}(D) = I(5, 5) = -\frac{5}{10} \log_2 \left( \frac{5}{10} \right) - \frac{5}{10} \log_2 \left( \frac{5}{10} \right) \\ = \boxed{1}$$

<u>Age</u>		<sup>(0)</sup>	<sup>(1)</sup>	
age	$c_1$	$c_2$		$I(c_1, c_2)$

$> 40.5$	3	3	$I(3, 3)$
----------	---	---	-----------

$< 40.5$	2	2	$I(2, 2)$
----------	---	---	-----------

$$\text{Info}(\text{Age}) = \frac{6}{10} \times I(3, 3) + \frac{4}{10} \times I(2, 2)$$

$$= \frac{6}{10} \left[ -\frac{3}{6} \log_2 \left( \frac{3}{6} \right) - \frac{3}{6} \log_2 \left( \frac{3}{6} \right) \right] +$$

$$\frac{4}{10} \left[ -\frac{2}{4} \log_2 \left( \frac{2}{4} \right) - \frac{2}{4} \log_2 \left( \frac{2}{4} \right) \right]$$

$$= \boxed{1}$$

Likes height.

likes height	<sup>(0)</sup>	<sup>(1)</sup>	
	$c_1$	$c_2$	$I(c_1, c_2)$

0	4	0	$I(4, 0)$
---	---	---	-----------

1	1	5	$I(1, 5)$
---	---	---	-----------

$$\text{Info}(\text{likes height}) = \frac{4}{10} \times I(4, 0) + \frac{6}{10} \times I(1, 5)$$

$$= \frac{4}{10} \left[ -\frac{4}{4} \log_2 \left( \frac{4}{4} \right) - \frac{0}{4} \log_2 \left( \frac{0}{4} \right) \right] +$$

$$\frac{6}{10} \left[ -\frac{1}{6} \log_2 \left( \frac{1}{6} \right) - \frac{5}{6} \log_2 \left( \frac{5}{6} \right) \right] = \boxed{0.918}$$

Likes goats

likes goats	<sup>(0)</sup> c <sub>1</sub>	<sup>(1)</sup> c <sub>2</sub>	I(c <sub>1</sub> , c <sub>2</sub> )
0	4	1	I(4, 1)
1	1	4	I(1, 4)

$$\begin{aligned}
 \text{Info}(\text{likes goats}) &= \frac{5}{10} \times I(4, 1) + \frac{5}{10} \times I(1, 4) \\
 &= \frac{5}{10} \left[ -\frac{4}{5} \log_2 \left( \frac{4}{5} \right) - \frac{1}{5} \log_2 \left( \frac{1}{5} \right) \right] + \\
 &\quad \frac{5}{10} \left[ -\frac{1}{5} \log_2 \left( \frac{1}{5} \right) - \frac{4}{5} \log_2 \left( \frac{4}{5} \right) \right] \\
 &= \boxed{0.721}
 \end{aligned}$$

$$\text{Gain}(\text{age}) = 1 - 1 = \boxed{0}$$

$$\text{Gain}(\text{likes height}) = 1 - 0.39 = \boxed{0.61}$$

$$\text{Gain}(\text{likes goats}) = 1 - 0.721 = \boxed{0.279}$$

As Gain(likes height) has the maximum gain is selected as root node.

2.

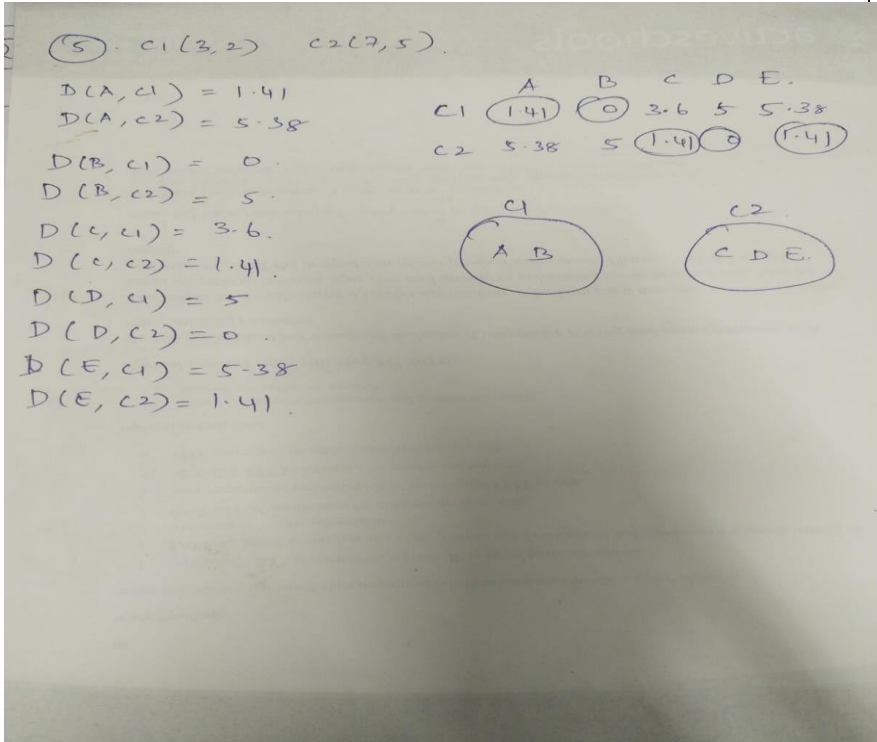
Consider the first five samples in the above dataset and apply kNN algorithm with k=3. Predict the class label for the input feature "Age=35, Likes height=0, Likes goats=1, Go rock climbing=??"

10

CO  
2

BL  
4

	<div><p><b>KNN</b></p><table><thead><tr><th></th><th>Age</th><th>Likes Height</th><th>likes goats</th><th>climbing</th><th>Distance</th><th>Rank</th></tr></thead><tbody><tr><td>1.</td><td>1</td><td>0</td><td>0</td><td>0</td><td><math>\sqrt{(1-1)^2 + (0-0)^2} = 0</math></td><td>1</td></tr><tr><td>2.</td><td>1</td><td>1</td><td>1</td><td>1</td><td><math>\sqrt{(1-1)^2 + (1-0)^2} = 1</math></td><td>2</td></tr><tr><td>3.</td><td>1</td><td>1</td><td>0</td><td>1</td><td><math>\sqrt{1+4} = 2.24</math></td><td>4</td></tr><tr><td>4.</td><td>1</td><td>0</td><td>0</td><td>0</td><td><math>\sqrt{1+1} = 1.41</math></td><td>3</td></tr><tr><td>5.</td><td>0</td><td>0</td><td>0</td><td>0</td><td><math>\sqrt{1+4} = 2.24</math></td><td>5</td></tr></tbody></table><p>New point [Age=1, Likes height=0, likes goats=1] then climbing</p><p>Majority target class with K=3 is 0, so for given x/p, climbing = 0</p></div>		Age	Likes Height	likes goats	climbing	Distance	Rank	1.	1	0	0	0	$\sqrt{(1-1)^2 + (0-0)^2} = 0$	1	2.	1	1	1	1	$\sqrt{(1-1)^2 + (1-0)^2} = 1$	2	3.	1	1	0	1	$\sqrt{1+4} = 2.24$	4	4.	1	0	0	0	$\sqrt{1+1} = 1.41$	3	5.	0	0	0	0	$\sqrt{1+4} = 2.24$	5																
	Age	Likes Height	likes goats	climbing	Distance	Rank																																																					
1.	1	0	0	0	$\sqrt{(1-1)^2 + (0-0)^2} = 0$	1																																																					
2.	1	1	1	1	$\sqrt{(1-1)^2 + (1-0)^2} = 1$	2																																																					
3.	1	1	0	1	$\sqrt{1+4} = 2.24$	4																																																					
4.	1	0	0	0	$\sqrt{1+1} = 1.41$	3																																																					
5.	0	0	0	0	$\sqrt{1+4} = 2.24$	5																																																					
3.	<div><p>Apply Linear SVM model to find the best hyperplane which splits the below given data points into two classes</p><p>Positively labelled data points (0,3)(0,4)(5,-1)</p><p>Negatively labelled data points (0,1)(0,2)(1,2)</p></div>	10	CO 2	BL 3																																																							
4.	<div><p>For the dataset given in Q.No.1 apply “AdaBoost” algorithm and find the following. Assume the 9<sup>th</sup> sample in the dataset is misclassified.</p><div><p><b>Build first Model - Data and Weights</b></p><table><thead><tr><th>age</th><th>likes height</th><th>likes goats</th><th>go rock climbing</th><th>weight</th></tr></thead><tbody><tr><td>23</td><td>0</td><td>0</td><td>0</td><td>0.1</td></tr><tr><td>31</td><td>1</td><td>1</td><td>1</td><td>0.1</td></tr><tr><td>35</td><td>1</td><td>0</td><td>1</td><td>0.1</td></tr><tr><td>35</td><td>0</td><td>0</td><td>0</td><td>0.1</td></tr><tr><td>42</td><td>0</td><td>0</td><td>0</td><td>0.1</td></tr><tr><td>43</td><td>1</td><td>1</td><td>1</td><td>0.1</td></tr><tr><td>45</td><td>0</td><td>1</td><td>0</td><td>0.1</td></tr><tr><td>46</td><td>1</td><td>1</td><td>1</td><td>0.1</td></tr><tr><td>46</td><td>1</td><td>0</td><td>0</td><td>0.1</td></tr><tr><td>51</td><td>1</td><td>1</td><td>1</td><td>0.1</td></tr></tbody></table><p>Initially, all data samples get the same weight, which is 1/N, with N the number of data points, in this case N=10. All weights sum up to 1.</p></div><div><p><b>AdaBoost - First Stump</b></p><div><div><p>likes height</p><p>false</p><p>going rock climbing</p><p>yes: 0</p><p>no: 4</p><p>misclassified: 0</p></div><div><p>true</p><p>going rock climbing</p><p>yes: 5</p><p>no: 1</p><p>misclassified: 1</p></div></div><div><p>Total Error: 1/10</p><p>Influence:</p><math display="block">\alpha = \frac{1}{2} \ln \left( \frac{1 - 1/10}{1/10} \right)</math><math display="block">\alpha = \frac{1}{2} \ln \left( \frac{9/10}{1/10} \right)</math><math display="block">\alpha = \frac{1}{2} \ln(9) = 1.099</math></div></div></div>	age	likes height	likes goats	go rock climbing	weight	23	0	0	0	0.1	31	1	1	1	0.1	35	1	0	1	0.1	35	0	0	0	0.1	42	0	0	0	0.1	43	1	1	1	0.1	45	0	1	0	0.1	46	1	1	1	0.1	46	1	0	0	0.1	51	1	1	1	0.1	10	CO 3	BL 4
age	likes height	likes goats	go rock climbing	weight																																																							
23	0	0	0	0.1																																																							
31	1	1	1	0.1																																																							
35	1	0	1	0.1																																																							
35	0	0	0	0.1																																																							
42	0	0	0	0.1																																																							
43	1	1	1	0.1																																																							
45	0	1	0	0.1																																																							
46	1	1	1	0.1																																																							
46	1	0	0	0.1																																																							
51	1	1	1	0.1																																																							

	<table><thead><tr><th>age</th><th>likes height</th><th>likes goats</th><th>go rock climbing</th><th>weight</th><th>normalized weight</th></tr></thead><tbody><tr><td>23</td><td>0</td><td>0</td><td>0</td><td>0.033</td><td>0.056</td></tr><tr><td>31</td><td>1</td><td>1</td><td>1</td><td>0.033</td><td>0.056</td></tr><tr><td>35</td><td>1</td><td>0</td><td>1</td><td>0.033</td><td>0.056</td></tr><tr><td>35</td><td>0</td><td>0</td><td>0</td><td>0.033</td><td>0.056</td></tr><tr><td>42</td><td>0</td><td>0</td><td>0</td><td>0.033</td><td>0.056</td></tr><tr><td>43</td><td>1</td><td>1</td><td>1</td><td>0.033</td><td>0.056</td></tr><tr><td>45</td><td>0</td><td>1</td><td>0</td><td>0.033</td><td>0.056</td></tr><tr><td>46</td><td>1</td><td>1</td><td>1</td><td>0.033</td><td>0.056</td></tr><tr><td>46</td><td>1</td><td>0</td><td>0</td><td>0.3</td><td>0.5</td></tr><tr><td>51</td><td>1</td><td>1</td><td>1</td><td>0.033</td><td>0.056</td></tr></tbody></table> <div>Updated weights, based on the results of the first weak learner. All weights sum up to 1</div> <table><thead><tr><th>age</th><th>likes height</th><th>likes goats</th><th>go rock climbing</th><th>bins</th></tr></thead><tbody><tr><td>23</td><td>0</td><td>0</td><td>0</td><td>[0.0, 0.056]</td></tr><tr><td>31</td><td>1</td><td>1</td><td>1</td><td>[0.056, 0.111]</td></tr><tr><td>35</td><td>1</td><td>0</td><td>1</td><td>[0.111, 0.178]</td></tr><tr><td>35</td><td>0</td><td>0</td><td>0</td><td>[0.178, 0.222]</td></tr><tr><td>42</td><td>0</td><td>0</td><td>0</td><td>[0.222, 0.278]</td></tr><tr><td>43</td><td>1</td><td>1</td><td>1</td><td>[0.278, 0.333]</td></tr><tr><td>45</td><td>0</td><td>1</td><td>0</td><td>[0.333, 0.389]</td></tr><tr><td>46</td><td>1</td><td>1</td><td>1</td><td>[0.389, 0.444]</td></tr><tr><td>46</td><td>1</td><td>0</td><td>0</td><td>[0.444, 0.944]</td></tr><tr><td>51</td><td>1</td><td>1</td><td>1</td><td>[0.944, 1]</td></tr></tbody></table>	age	likes height	likes goats	go rock climbing	weight	normalized weight	23	0	0	0	0.033	0.056	31	1	1	1	0.033	0.056	35	1	0	1	0.033	0.056	35	0	0	0	0.033	0.056	42	0	0	0	0.033	0.056	43	1	1	1	0.033	0.056	45	0	1	0	0.033	0.056	46	1	1	1	0.033	0.056	46	1	0	0	0.3	0.5	51	1	1	1	0.033	0.056	age	likes height	likes goats	go rock climbing	bins	23	0	0	0	[0.0, 0.056]	31	1	1	1	[0.056, 0.111]	35	1	0	1	[0.111, 0.178]	35	0	0	0	[0.178, 0.222]	42	0	0	0	[0.222, 0.278]	43	1	1	1	[0.278, 0.333]	45	0	1	0	[0.333, 0.389]	46	1	1	1	[0.389, 0.444]	46	1	0	0	[0.444, 0.944]	51	1	1	1	[0.944, 1]				
age	likes height	likes goats	go rock climbing	weight	normalized weight																																																																																																																									
23	0	0	0	0.033	0.056																																																																																																																									
31	1	1	1	0.033	0.056																																																																																																																									
35	1	0	1	0.033	0.056																																																																																																																									
35	0	0	0	0.033	0.056																																																																																																																									
42	0	0	0	0.033	0.056																																																																																																																									
43	1	1	1	0.033	0.056																																																																																																																									
45	0	1	0	0.033	0.056																																																																																																																									
46	1	1	1	0.033	0.056																																																																																																																									
46	1	0	0	0.3	0.5																																																																																																																									
51	1	1	1	0.033	0.056																																																																																																																									
age	likes height	likes goats	go rock climbing	bins																																																																																																																										
23	0	0	0	[0.0, 0.056]																																																																																																																										
31	1	1	1	[0.056, 0.111]																																																																																																																										
35	1	0	1	[0.111, 0.178]																																																																																																																										
35	0	0	0	[0.178, 0.222]																																																																																																																										
42	0	0	0	[0.222, 0.278]																																																																																																																										
43	1	1	1	[0.278, 0.333]																																																																																																																										
45	0	1	0	[0.333, 0.389]																																																																																																																										
46	1	1	1	[0.389, 0.444]																																																																																																																										
46	1	0	0	[0.444, 0.944]																																																																																																																										
51	1	1	1	[0.944, 1]																																																																																																																										
5.	<p>Group the following points using k-means algorithm into two clusters. Assume 'B' and 'D' as initial centroid. Find the clusters after one iteration. A=(2,3) B=(3,2) C=(6,4) D=(7,5) E=(8,4)</p> 	10	CO 3	BL 3																																																																																																																										

