

Question 1

1 / 1 pts

We have an attribute set made of two boolean features (A, B) where $A, B \in \{0, 1\}$ with boolean labels $y \in \{0, 1\}$. In our dataset we have 4 categories of data points with counts:

- (i) $A = 0, B = 1, y = 0$: 30 examples
- (ii) $A = 0, B = 0, y = 0$: 60 examples
- (iii) $A = 1, B = 0, y = 1$: 15 examples
- (iv) $A = 1, B = 1, y = 1$: 15 examples

We want to train a decision tree with this data. Look at the data, and determine which of A or B should be at the root of the tree. Then verify your answer by computing the information gain.

A, IG = 0.8112

B, IG = 0.8112

A, IG = 0.562

B, IG = 0.562

Question 2

1 / 1 pts

The Boolean function, $\neg x_1 \wedge x_2 \wedge \neg x_3 \wedge x_4$, is a linear function over the boolean variables, x_1, x_2, x_3, x_4 . Which of the following is a correct "linear" representation for it?

$x_1 + x_2 + x_3 + x_4 \geq 2$

$x_1 + x_2 + x_3 + x_4 \geq 4$

$-x_1 + x_2 - x_3 + x_4 \geq 4$

$-x_1 + x_2 - x_3 + x_4 \geq 2$

Question 3

1 / 1 pts

We want to show that the Boolean function

$$y = 1 \text{ if and only if at least 3 out of 5 variables are 1}$$

can be written as a linear threshold function $w^T \cdot x \geq \theta$, where $x \in \{0, 1\}^5$ and $y \in \{0, 1\}$. What values of w^T and θ will show this?

$w^T = [1, 1, 1, 1, 1], \theta = 3$

$w^T = [1, 1, 0, 0, 0], \theta = 5$

$w^T = [0, 1, 1, 0, 1], \theta = 1$

$w^T = [1, 1, 1, 0, 0], \theta = 3$

Question 4

1 / 1 pts

Consider the following dataset:

Example ID	Feature A	Feature B	Label
0	0	0	0
1	0	1	1
2	1	0	0

3	1	1	0
4	2	0	0
5	2	1	0
6	3	0	0
7	3	1	0

Each instance has two features, A (with 4 possible values), B (with 2 possible values), and a binary label. If you build a decision tree using ID3 and split on Feature A first, for the node representing (A=0), what is the next step?

- Do nothing, the algorithm stops
- Split on B, IG=0.5
- Split on B, IG=0
- Split on B, IG=1

Question 5

1 / 1 pts

We want to learn a classifier over input that consists of a pair of character strings composed of lower case letters {a,b,c,...,z}. For example 'four seasons' or 'bill clinton' are possible input strings. We define two feature types:

- (i) "Whether or not the first character in the first string is a vowel" (This feature will be the x_1 in our feature vector.)
- (ii) "The first character in the second string is _" (These features will be the x_2, x_3, \dots, x_{27} in our feature vector, corresponding to whether the first character in the second string is 'a', 'b', ..., 'z' respectively.)

Which of the following examples is the feature-based representation of 'aditi rao'?

1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0

0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1

1, 1

0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0

Quiz Score: **5** out of 5