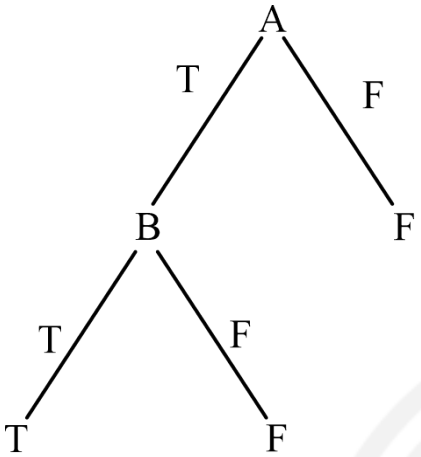
Machine Learning

Decision Tree

DPP: 1

Q1 What does the following figure represent?



(A) Decision tree for OR
(B) Decision tree for AND
(C) Decision tree for XOR
(D) Decision tree for XNOR

Q2 Consider the dataset given below where T and F represent True and False respectively. What is the entropy H (Rain)?

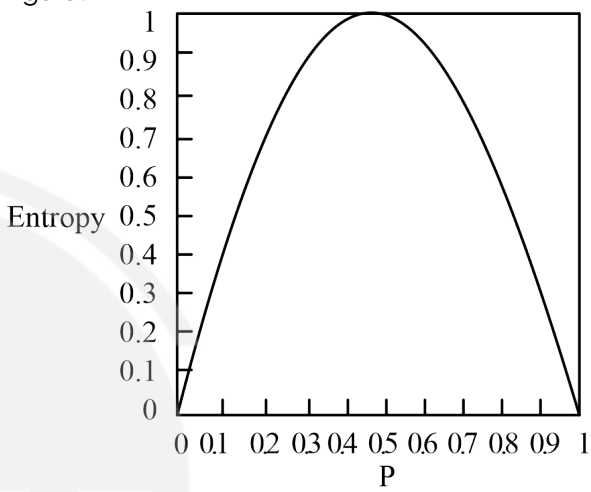
Temperature	Cloud	Rain
Low	T	T
Low	T	T
Medium	T	F
Medium	T	T
High	T	F
High	F	F

(A) 1 (B) 0.5
(C) 0.2 (D) 0.6

Q3 Given the entropy for a split, Esplit = 0.39 and the entropy before the split, Ebefore = 1. What is the Information Gain for the split?

(A) 1 (B) 0.39
(C) 0.61 (D) 2.56

Q4 What is the entropy at P = 0.5 from the given figure?



(A) 0.5 (B) -0.5
(C) 1 (D) -1

Q5 Given entropy of parent = 1, weights average = $\left(\frac{3}{4}, \frac{1}{4}\right)$ and entropy of children = (0.9, 0). What is the information gain?

(A) 0.675 (B) 0.75
(C) 0.325 (D) 0.1

Q6 Which of the following statements is not true about the Decision tree?

(A) A Decision tree is also known as a classification tree
(B) Each element of the domain of the classification in decision tree is called a class
(C) It is a tree in which each internal node is labeled with an input feature
(D) It cannot be used in data mining applications as it only classifies but not predicts anything

Answer Key

Q1	(B)	Q4	(C)
Q2	(A)	Q5	(C)
Q3	(C)	Q6	(D)

Hints & Solutions

Q1 Text Solution:

The given figure represents the decision tree implementation of Boolean AND as per the following truth table.

A	B	A AND B
F	F	F
F	T	F
T	T	T
T	F	F

Q2 Text Solution:

We know entropy = $\sum_{i=1}^{n}$ – Pi log2 Pi.

Entropy = – (3/6) * log2 (3/6) – (3/6) * log2 (3/6)

= – (1/2) * log2 (1/2) – (1/2) * log2 (1/2)

= – 0.5 * -1 – 0.5 * -1

= 0.5 + 0.5

= 1

Q3 Text Solution:

Information Gain is calculated for a split by subtracting the weighted entropies of each branch from the original entropy. We have Esplit = 0.39 and Ebefore = 1.

Then Information Gain, IG = Ebefore – Esplit

= 1 – 0.39

= 0.61

Q4 Text Solution:

We know the entropy E = -p log2p – q log2q. Here p = 0.5 and q = 1 – p = 1 – 0.5 = 0.5. So we have p = 0.5 and q = 0.5.

Entropy = (–0.5 * log2 0.5) – (0.5 * log2 0.5)

= (-0.5 * -1) – (0.5 * -1)

= 0.5 + 0.5

= 1

Q5 Text Solution:

We know information gain = Entropy (parent) – Σ (weights average × entorpy (child) information gain

= $1 - 1 - \left(\frac{3}{4} \times 0.9 + \frac{1}{4} \times 0\right)$

= 1 – (0.675 + 0)

= 1 – 0.675

= 0.325

Q6 Text Solution:

Decision trees can be widely used in data mining applications because it is able to classify and predict as well. It is also known as a classification tree. Each element of the domain of the classification in the decision tree is called a class and each internal node is labeled with an input feature.