MODULE - 1

1. Consider the Enjoy Sport concept and instance given below, identify the specific hypothesis using Find-S algorithm.

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

* First initialize 'h' to the most specific hypothesis in H

$$\boxed{h = \langle \phi, \phi, \phi, \phi, \phi, \phi \rangle}$$

* Consider the first training instance.

$$\boxed{x_1 = \langle \text{Sunny, warm, Normal, Strong, warm, Same} \rangle, +}$$

Observing instance x_1, 'h' is too specific so replace by the next more general constraint that fits the example.

$$\boxed{h_1 = \langle \text{Sunny, warm, Normal, Strong, warm, Same} \rangle}$$

* Consider second training instance.

$$\boxed{x_2 = \langle \text{Sunny, warm, High, Strong, warm, Same} \rangle, +}$$

Observe 'x_2' with 'h_1' and replace by more general constraint

$$\boxed{h_2 = \langle \text{Sunny, warm, ?, Strong, warm, Same} \rangle}$$

* Consider third training instance

x_3 = < Rainy, Cold, High, Strong, warm, Change >, −

Find-S algorithm ignores negative instances, so. $h_3 = h_2$

h_3 = < Sunny, warm, ?, Strong, warm, Same >

* consider fourth training instance.

x_4 = < Sunny, warm, High, Strong, warm, change >, +

compare x_4 with h_3 and replace by more general constraint

h_4 = < Sunny, warm, ?, Strong, ?, ? >

The final specific hypothesis for given instances is

h_4 = < Sunny, warm, ?, Strong, ?, ? >

2. Consider the EnjoySport concept and instance given below, identify the general and specific hypotheses using Candidate - Elimination learning algorithm

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

* The boundary sets are first initialized to G_0 & S_0 the most general & most specific hypotheses in H

S_0 <Ø . Ø . Ø . Ø . Ø . Ø>

G_0 <? , ? , ? , ? , ? , ?>

* Consider the first training instance.

x_1 = <Sunny, warm, Normal, Strong, warm, Same>

S_0 = <Ø Ø Ø Ø Ø Ø>

↓

S_1 = <Sunny, warm, Normal, Strong, warm, Same>

G_0 G_1 = <? ? ? ? ? ?>

* Consider second instance

x_2 = ⟨Sunny, warm, High, Strong, warm, Same⟩ +

S_1 = ⟨Sunny, warm, Normal, Strong, warm, Same⟩

↓

S_2 = ⟨Sunny, warm, ?, Strong, warm, Same⟩

G_1, G_2 = ⟨?, ?, ?, ?, ?, ?⟩

* Consider third instance which is <u>negative</u>.

x_3 = ⟨Rainy, Cold, High, Strong, warm, Change⟩ −

S_2 S_3 = ⟨Sunny, warm, ?, Strong, warm, Same⟩

G_3 = ⟨Sunny, ? ? ? ? ?⟩ ⟨? warm ? ? ? ?⟩
⟨? ? ? ? ? Same⟩

↑

G_2 ⟨? ? ? ? ? ?⟩

* Consider the fourth instance.

x_4 = < Sunny, warm, High. Strong, cool. Change >

S_3 = $\boxed{\text{< Sunny, warm, ?, Strong, warm, Same >}}$

↓

S_4 = $\boxed{\text{< Sunny, warm, ?, Strong, ? ? >}}$

G_4 = $\boxed{\text{< Sunny ? ? ? ? ? > < ? warm ? ? ? ? >}}$

↑

G_3 = $\boxed{\text{< Sunny, ? ? ? ? ? > < ? warm ? ? ? ? > < ? ? ? ? ? warm >}}$

S_4 and G_4 are the final set of hypotheses which are consistent to training instances

3. Consider the "Japanese Economy Car" concept and instance given below, identify the hypotheses using Candidate - Elimination learning algorithm.

Origin	Manufacturer	Color	Decade	Type	Target Value
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chryler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive

* Initialize G_0 & S_0.

S_0 ⟨Japan ∅ ∅ ∅ ∅ ∅⟩

G_0 ⟨? ? ? ? ?⟩

* Consider the first training instance

x_1 = ⟨Japan, Honda, Blue, 1980, Economy⟩ +

S_0 = ⟨∅ ∅ ∅ ∅ ∅⟩

S_1 = ⟨Japan, Honda, Blue, 1980, Economy⟩

~~G_0~~ G_1 = ⟨? ? ? ? ?⟩

* Consider second ~~training~~ instance.

x_2 = < Japan, Toyota, Green, 1970, Sports > —

Specialize G to exclude the negative example.

S_1, S_2 = < Japan, Honda, Blue, 1980, Economy >

G_2 = < ? Honda ? ? ? > < ? ? Blue ? ? >
< ? ? ? 1980 ? > < ? ? ? ? Economy >

↑

G_1 = < ? ? ? ? ? >

* Consider third traing instance.

x_3 = < Japan, Toyota, Blue, 1990, Economy >

Prune G to exclude inconsistent hypotheses with the positive example and generalize S to include with positive example.

S_3 = < Japan ? Blue ? Economy >

G_3 = < ? ? Blue ? ? > < ? ? ? ? Economy >

* Consider fourth instance

x_4 = < USA, Chrysler, Red, 1980, Economy > –

Specialize G to include the negative example but stay consistent with S.

G_4 = < Japan ? Blue, ? Economy >

G_4 = < ? ? Blue ? ? > < Japan ? ? ? Economy >

* Consider fifth instance

x_5 = < Japan, Honda, white, 1980, Economy >

Prune G to exclude inconsistent hypotheses with positive example and generalize S.

S_5 = < Japan, ? ? ? Economy >

G_5 = < Japan, ? ? ? Economy >

These are the final set of hypotheses which consistent with the training instance.

MODULE 2

1. Give Decision trees for the following set of training examples

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Solution :-

* $$\text{Entropy}(S) \equiv -P_{\oplus}\log_2 P_{\oplus} - P_{\ominus}\log_2 P_{\ominus}$$

$$\text{Gain}(S, A) \equiv \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|}\,\text{Entropy}(S_v)$$

* Note,

① when all members of S belong to the same class then, $\text{Entropy}(S) = 0$

② If S contains an equal number of positive and negative examples then,

$\text{Entropy}(S) = 1$

* The first step is to find the topmost node of the decision tree. ID3 determines the information gain for each attribute, then selects the one with highest information gain.

→ Entropy of S : positive examples = 09
negative examples = 05

$$\text{Entropy}([9+, 5-]) = -\left(\frac{9}{14}\right)\log_2\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right)\log_2\left(\frac{5}{14}\right)$$

$$= \underline{\underline{0.940}}$$

⇒ Information gain of the attribute <u>Outlook</u> is calculated as

Value (Outlook) = Sunny, Overcast, Rain

$S_{sunny} \leftarrow [2+, 3-]$

$S_{overcast} \leftarrow [4+, 0]$

$S_{Rain} \leftarrow [3+, 2-]$

$$\text{Gain}(S, \text{Outlook}) = \text{Entropy}(S) - \left[\left(\frac{5}{14}\right)\text{Entropy}(S_{sunny}) + \left(\frac{4}{14}\right)\text{Entropy}(S_{overcast}) + \left(\frac{5}{14}\right)\text{Entropy}(S_{Rain})\right]$$

* $$\text{Entropy}(S_{sunny}) = -\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right) - \left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right)$$
$$= -(0.4 * (-1.3219)) - (0.6 * (-0.7369))$$
$$= 0.52876 + 0.44214$$
$$= 0.970$$

* $\text{Entropy}(S_{overcast}) = 0$ (because of all members belong to same class)

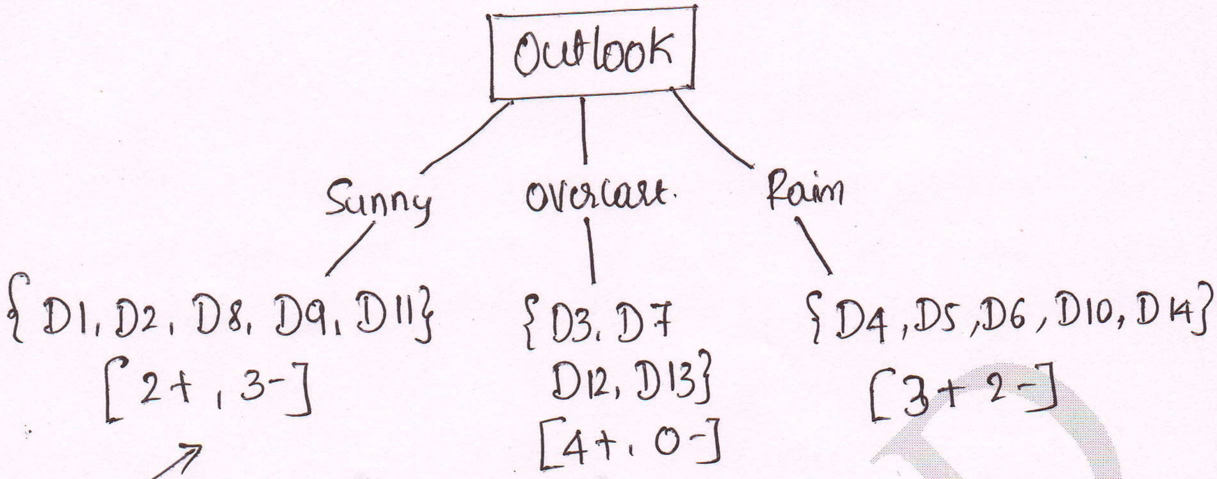
* $$\text{Entropy}(S_{Rain}) = -\left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right) - \left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right)$$
$$= -(0.6 * (-0.7369) - (0.4 * (-1.3219))$$
$$= 0.44214 + 0.52876$$
$$= 0.970$$

$\Rightarrow$
$$= (0.940) - \left[\left(\frac{5}{14}\right) * 0.9709 + 0 + \left(\frac{5}{14}\right) 0.9709\right]$$
$$= 0.9409 - [0.3467 + 0.3467]$$
$$= \underline{0.246}$$

So Gain (S, outlook) = 0.246

Similarly
Gain (S, Temperature) = 0.029
Gain (S, Humidity) = 0.151
Gain (S, wind) = 0.029

So Root node will be <u>Outlook</u>.



which attribute should be tested here?

$$S_{sunny} = \{D1, D_2, D_8, D_9, D_{11}\}$$

$$\text{Gain}(\text{Sunny}, \text{Humidity}) = 0.970 - \left[\left(\tfrac{3}{5}\right) * 0.0 + \left(\tfrac{2}{5}\right) \cdot 0.0\right]$$

$$= \underline{0.970}.$$

$$\text{Gain}(\text{Sunny}, \text{Temperature}) = 0.970 - \left[\left(\tfrac{2}{5}\right) * 0 + \left(\tfrac{2}{5}\right) * 1 + \left(\tfrac{1}{5}\right) \cdot 0\right]$$

$$= 0.970 - 0.4$$

$$= \underline{0.570}$$

$$\text{Gain}(\text{Sunny}, \text{Wind}) = 0.970 - \left[\left(\tfrac{3}{5}\right) * (0.918) + \left(\tfrac{2}{5}\right) \cdot 1\right]$$

$$= 0.970 - 0.9508$$

$$= \underline{0.0192}$$

So, attribute Humidity will be descendant node.

⇒ $S_{Rain} = \{D4, D5, D6, D10, D14\}$

$$Entropy(S_{Rain}) = -\left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right) - \left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right)$$
$$= \underline{\underline{0.970}}$$

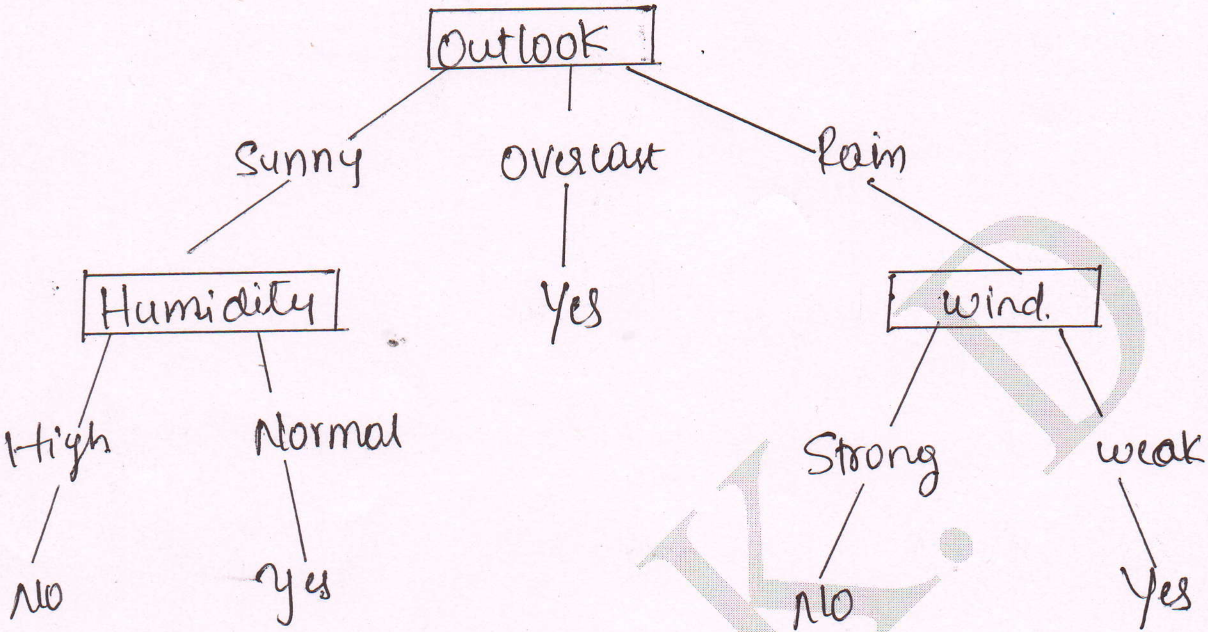
$$Gain(Rain, Temperature) = 0.970 - \left[\left(\frac{0}{5}\right)*0 + \left(\frac{3}{5}\right)*0.918 + \left(\frac{2}{5}\right)*1\right]$$
$$= \underline{\underline{0.0198}}$$

$$Gain(Rain, wind) = 0.970 - \left[\left(\frac{3}{5}\right)*0 + \left(\frac{2}{5}\right)*0\right]$$
$$= \underline{\underline{0.970}}$$

$$Gain(Rain, Humidity) = 0.97 - \left[\left(\frac{2}{5}\right)*1 + \left(\frac{3}{5}\right)*0.917\right]$$
$$= \underline{\underline{0.0198}}$$

So, highest information gain is ~~node~~ attribute ~~the~~ Wind.

So, the final tree is.



2. Consider the following set of training examples.
   a) What is the entropy of this collection of training example with respect to the target function classification?
   b) What is the information gain of a_2 relative to these training examples?

Instance	Classification	a_1	a_2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

⇒

$$Entropy(S) = -P_{\oplus}\log_2 P_{\oplus} - P_{\ominus}\log_2 P_{\ominus}$$

Information gain

$$Gain(S,A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

a) positive instances = 03
negative instances = 03

$$Entropy(S) = -\left(\frac{3}{6}\right)\log_2\left(\frac{3}{6}\right) - \left(\frac{3}{6}\right)\log_2\left(\frac{3}{6}\right)$$

$$\boxed{Entropy(S) = 1}$$

when there are equal number of +ve & -ve instances, then Entropy (S) = 1

b) $Gain(S, a_2) = Entropy(S) - \left[\frac{4}{6} Entropy(S_T) + \frac{2}{6} Entropy(S_F)\right]$

find

1. $Entropy(S_T) = -\left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right) - \left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right)$

$= \underline{\underline{1}}$

2. $Entropy(S_F) = -\left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right)$

$= \underline{\underline{1}}$

$\Rightarrow Gain(S, a_2) = 1 - \left[\frac{4}{6}(1) + \frac{2}{6}(1)\right]$

$= 1 - 1$

$\boxed{Gain(S, a_2) = \underline{\underline{0}}}$

3. Give decision trees to represent the following Boolean functions.

i) A && ¬B

ii) A ∨ [B && C]

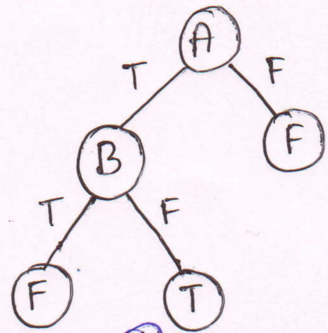
iii) A XOR B

iv) [A && B] ∨ [C && D]

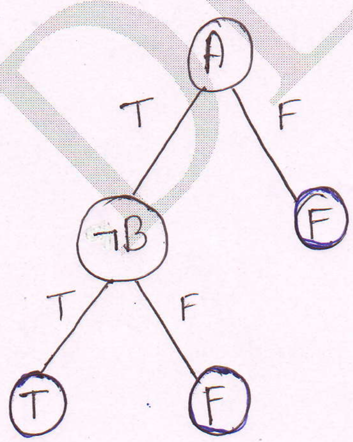
Solution

i) A && ¬B

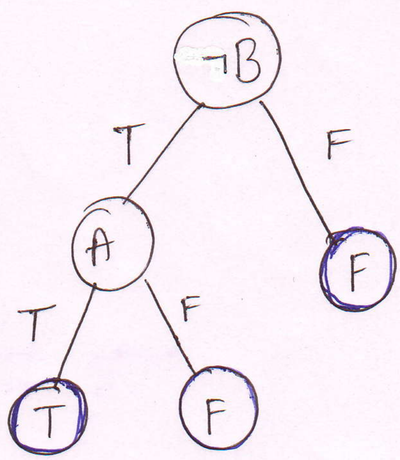
A	B	¬B	A && ¬B
T	T	F	F (−)
T	F	T	T (+)
F	T	F	F (−)
F	F	T	F (−)



OR

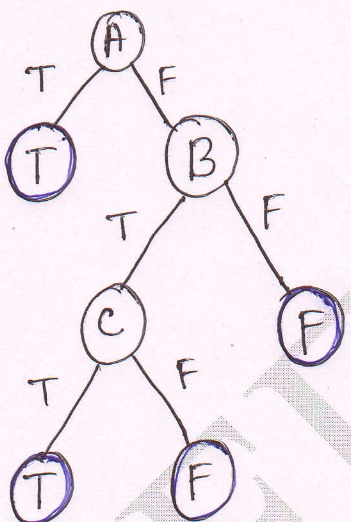


OR

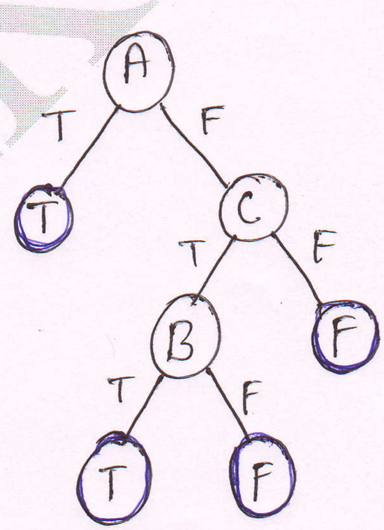


ii) A V [B && C]

A	B	C	B && C	A V [B && C]
T	T	T	T	T
T	T	F	F	T
T	F	T	F	T
T	F	F	F	T
F	T	T	T	T
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

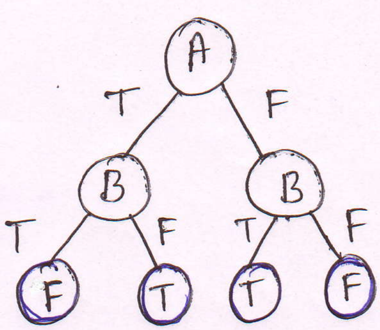


OR.



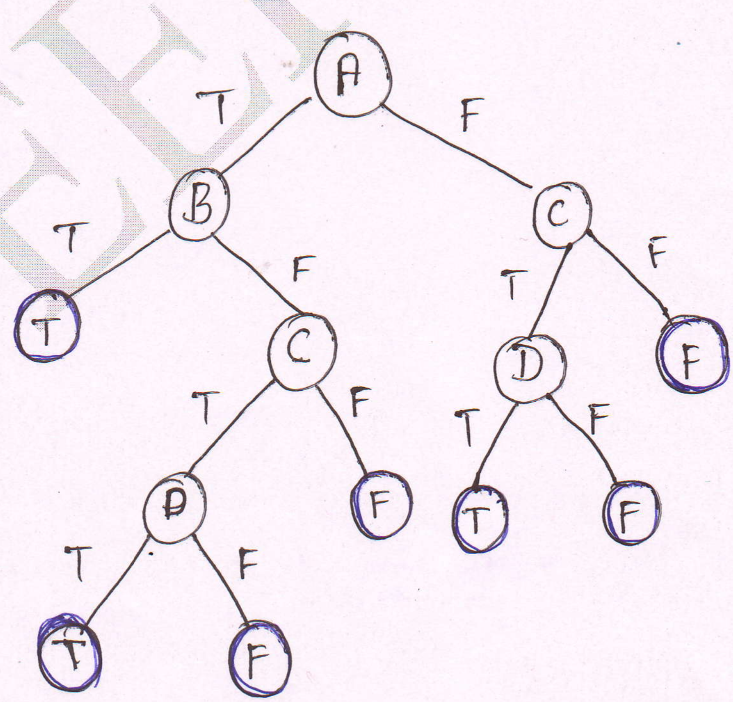
iii) A XOR B

A	B	A XOR B
T	T	F
T	F	T
F	T	T
F	F	F



iv) [A && B] v [C && D]

A	B	C	D	A && B	C && D	(A && B) v (C && D)
T	T	T	T	T	T	T
T	T	T	F	T	F	T
T	T	F	T	T	F	T
T	T	F	F	T	F	T
T	F	T	T	F	T	T
T	F	T	F	F	F	F
T	F	F	T	F	F	F
T	F	F	F	F	F	F
F	T	T	T	F	T	T
F	T	T	F	F	F	F
F	T	F	T	F	F	F
F	T	F	F	F	F	F
F	F	T	T	F	T	T
F	F	T	F	F	F	F
F	F	F	T	F	F	F
F	F	F	F	F	F	F



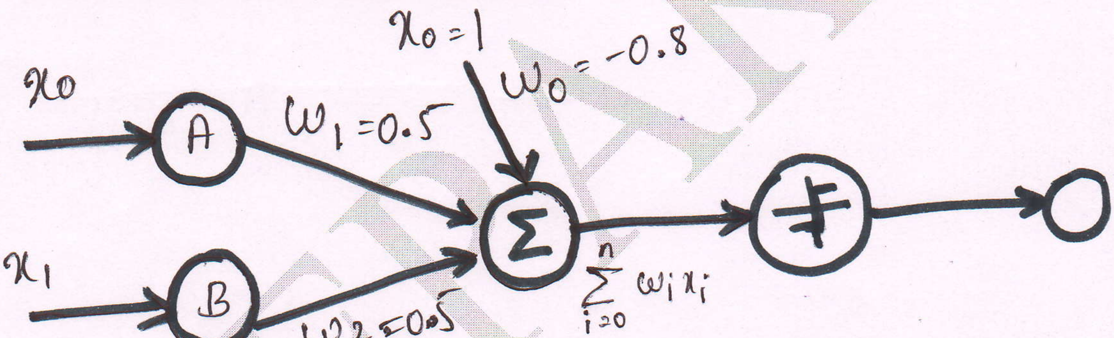
MODULE 3

1. How a single perceptron can be used to represent the Boolean functions such as AND, OR

⇒ Boolean function AND

A	B	A ∧ B
0	0	0
0	1	0
1	0	0
1	1	1

* Set $w_0 = -0.8$
$w_1 = 0.5$
$w_2 = 0.5$



$$O(x_1 \ldots x_n) = \begin{cases} 1 & \text{if } w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n > 0 \\ -1 & \text{otherwise.} \end{cases}$$

1) if $\underset{x_1}{A=0}$ & $\underset{x_2}{B=0}$ ⇒ $\cancel{0*0}$ $-0.8 + (0.5*0) + (0.5*0) =$
$= -0.8 < 0$ So, output $= 0$

2) if A=0 & B=1 ⇒ $-0.8 + (0.5*0) + (0.5*1)$
$= -0.3 < 0$ So, output $= 0$

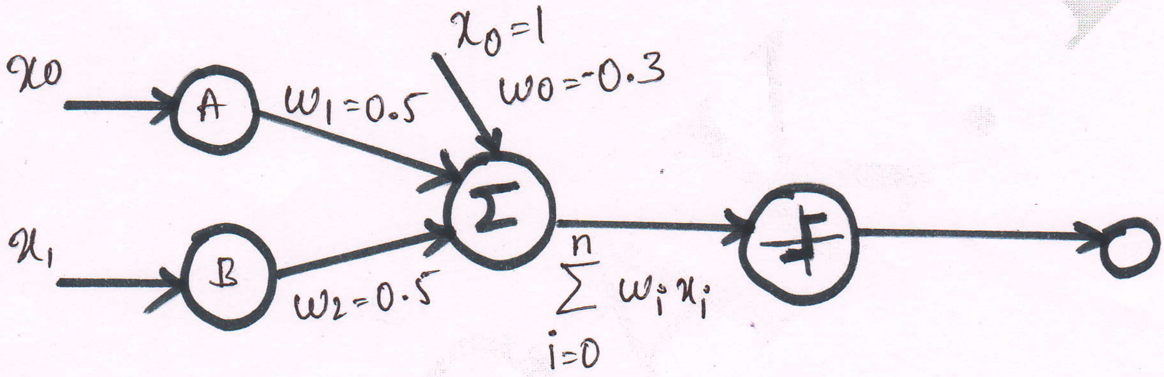
3) if A=1 & B=1 ⇒ $-0.8 + (0.5*1) + (0.5*0) = -0.3 < 0$
Output = 0

4) if A=1 & B=1 ⇒ $-0.8 + (0.5*1) + (0.5*1) = 0.2 > 0$
Output = 1

⇒ Boolean function OR

A	B	A V B
0	0	0
0	1	1
1	0	1
1	1	1

* Set $w_0 = -0.3$
$w_1 = 0.5$
$w_2 = 0.5$



i) A = 0 B = 0 ⇒ $-0.3 + (0.5 * 0) + (0.5 * 0)$
$= -0.3 < 0$ So output = 0

2) A = 0 B = 1 ⇒ $-0.3 + (0.5 * 0) + (0.5 * 1)$
$= 0.2 > 0$ So output = 1

3) A = 1 B = 0 ⇒ $-0.3 + (0.5 * 1) + (0.5 * 0)$
$= 0.2 > 0$ So output = 1

4) A = 1 B = 1 → $-0.3 + (0.5 * 1) + (0.5 * 0)$
$= 0.7 > 0$ So output = 1

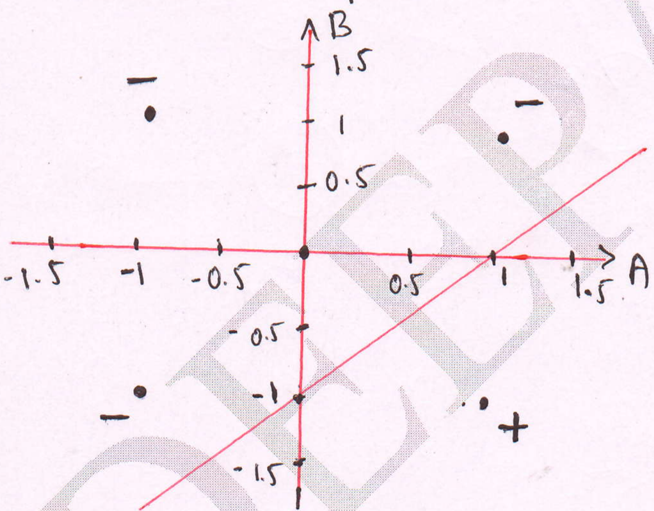
2. (a) Design a two-input perceptron that implements the boolean function $A \wedge \neg B$. (b) Design a two-layer network of perceptron's that implements A XOR B.

⟹ a) The perceptron has two input A, B and constant 1

A	B	¬B	A ∧ ¬B
0 (-1)	0 (-1)	1	0 (-1)
0 (-1)	1	0 (-1)	0 (-1)
1	0 (-1)	1	1
1	1	0	0 (-1)

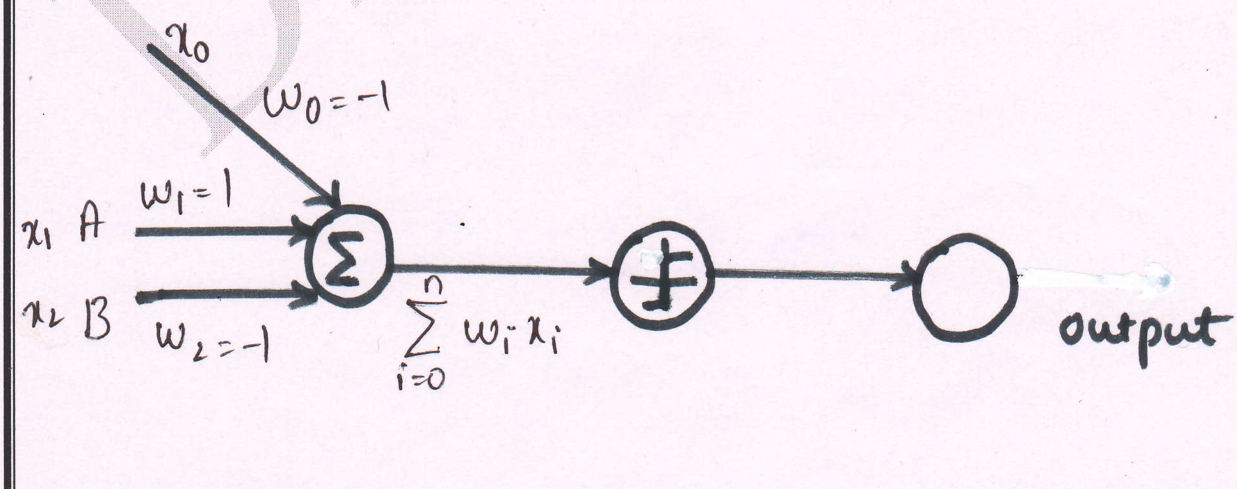
The values of A & B are 1 (true) or -1 or 0 for false.

Decision surfaces



* The line crosses the A axis at 1 and B axis at -1

* The weights are
$w_0 = -1$
$w_1 = 1$ $w_2 = -1$.



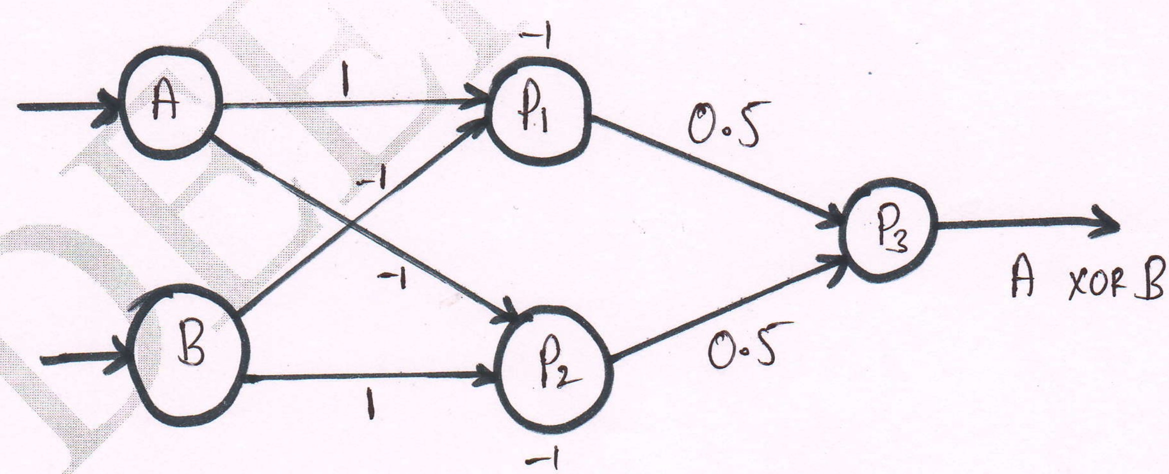
⟹ b) A XOR B cannot be calculated by a single perceptron, so build a two-layer network of perceptrons

* Express A XOR B in terms of other logical connectives

$$A \text{ XOR } B = (A \wedge \neg B) \vee (\neg A \wedge B)$$

* Define the perceptron P_1 and P_2 for $(A \wedge \neg B)$ & $(\neg A \wedge B)$

* Composing the outputs of P_1 & P_2 into a perceptron P_3 that implements $O(P_1) \vee O(P_2)$



3. Consider two perceptrons defined by the threshold expression $w_0 + w_1x_1 + w_2x_2 > 0$.

Perceptron A has weight values

$$w_0 = 1, w_1=2, w_2=1$$

and perceptron B has the weight values

$$w_0 = 0, w_1=2, w_2=1$$

True or false? Perceptron A is more-general than perceptron B.

Solution

True, Perceptron A is more-general than Perceptron B.

$\Rightarrow O(x_1 \ldots x_n) = w_0x_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$

$\Rightarrow B(\langle x_1, x_2\rangle) = 1$ & $w_0 = 0$, $w_1 = 2$, $w_2 = 1$

$0 + 2x_1 + x_2 > 0 \Rightarrow \underline{0 + 2 + 1 > 0}$

where, x_0 is constant which is equal to 1 i.e., $x_0 = 1$

$\Rightarrow A(\langle x_1, x_2\rangle) = 1$ & $w_0 = 1$, $w_1 = 2$, $w_2 = 1$

$1 + 2x_1 + x_2 > 0 \Rightarrow \underline{1 + 2 + 1 > 0}$

Here, Perceptron A is more general than perceptron B because every instance of x_1 & x_2 that satisfies perceptron B also satisfies perceptron A.

MODULE 4

1. Consider a medical diagnosis problem in which there are two alternative hypotheses: 1. that the patient has a particular form of cancer (+) and 2. That the patient does not (-). A patient takes a lab test and the result comes back positive. The test returns a correct positive result in only 98% of the cases in which the disease is actually present, and a correct negative result in only 97% of the cases in which the disease is not present. Furthermore, .008 of the entire population have this cancer. Determine whether the patient has Cancer or not using MAP hypothesis.

Solution :-

* From the above situation, probabilities can be summarized as follow.

$$P(cancer) = 0.008 \qquad P(\neg cancer) = 0.992$$
$$P(\oplus | cancer) = 0.98 \qquad P(\ominus | cancer) = 0.02$$
$$P(\oplus | \neg cancer) = 0.03 \qquad P(\ominus | \neg cancer) = 0.97$$

Should we diagnose the patient as having cancer or not?

$$P(\oplus | cancer)\, P(cancer) = (0.98) * (0.08) = 0.0078$$
$$P(\oplus | \neg cancer)\, P(\neg cancer) = 0.03 * 0.992 = 0.0298$$

So, $h_{MAP} = \neg cancer$

* The exact posterior probabilities can also be determined by normalizing the above quantities so that they sum to 1

$$P(cancer | \oplus) = \frac{0.0078}{0.0078 + 0.0298} = 0.21$$

$$P(\neg cancer | \oplus) = \frac{0.0298}{0.0078 + 0.0298} = 0.79$$

2. Apply Naïve Bayes classifier for ***PlayTennis*** concept learning problem to classify the following novel instance

< Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong >

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Table: Training examples for the target concept *PlayTennis*

⇒ Solution

* Task is to predict the target value (yes or no) of the target concept PlayTennis for this new instance.

* $$V_{NB} = \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \prod_i P(a_i | v_j) \longleftarrow \text{equ } (1)$$

$$V_{NB} = \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \cdot P(\text{Outlook} = \text{Sunny} | v_j) \cdot$$

$$P(\text{Temp} = \text{cool} | v_j) \cdot P(\text{Humidity} = \text{high} | v_j) \cdot P(\text{Wind} = \text{Strong} | v_j)$$

* First, the probabilities of the different target values is estimated based on their frequencies over 14 training examples

$$P(\text{playTennis} = \text{yes}) = \frac{9}{14} = 0.64$$

$$P(\text{playTennis} = \text{no}) = \frac{5}{14} = 0.36$$

* Next, estimate the conditional probabilities

$$P(\text{Outlook} = \text{sunny} \mid \text{playTennis} = \text{yes}) = \frac{2}{9} = 0.22$$

$$P(\text{Outlook} = \text{sunny} \mid \text{playTennis} = \text{no}) = \frac{3}{5} = 0.60$$

$$P(\text{Temp} = \text{cool} \mid \text{playTennis} = \text{yes}) = \frac{3}{9} = 0.33$$

$$P(\text{Temp} = \text{cool} \mid \text{playTennis} = \text{no}) = \frac{1}{5} = 0.20$$

$$P(\text{Humidity} = \text{high} \mid \text{playTennis} = \text{yes}) = 3/9 = 0.33$$

$$P(\text{Humidity} = \text{high} \mid \text{playTennis} = \text{no}) = 4/5 = 0.80$$

$$P(\text{wind} = \text{Strong} \mid \text{playTennis} = \text{yes}) = \frac{3}{9} = 0.33$$

$$P(\text{wind} = \text{Strong} \mid \text{playTennis} = \text{no}) = \frac{3}{5} = 0.60$$

* Next calculate V_{NB} according equation (1)

$$= P(\text{yes}) \cdot P(\text{sunny} \mid \text{yes}) \cdot P(\text{cool} \mid \text{yes}) \cdot P(\text{high} \mid \text{yes}) \cdot P(\text{strong} \mid \text{yes})$$

$$= 0.64 * 0.22 * 0.33 * 0.33 * 0.33$$

$$= \underline{\underline{0.0053}}$$

$= P(no) \cdot P(Sunny|no) \cdot P(cool|no) \cdot P(high|no) \cdot P(Strong|no)$

$= 0.036 * 0.60 * 0.20 * 0.80 * 0.60$

$= \underline{\underline{0.0206}}$

Thus, the naive Bayes classifier assigns the target value "no" to the new instance.

ie., $\underline{\underline{\text{PlayTennis} = \text{no}}}$

Outlook	Temp	Humidity	Wind	PlayTennis
Sunny	Cool	high	Strong	no

* Normalizing the quantities to sum to 1, calculate the conditional probability of target values.

$$\text{Yes} = \frac{0.0053}{0.0053 + 0.0206} = \underline{\underline{0.205}}$$

$$\text{No} = \frac{0.0206}{0.0053 + 0.0206} = \underline{\underline{0.795}}$$

CBCS SCHEME

USN | | | | | | | | | |

15CS73

Seventh Semester B.E. Degree Examination, Dec.2018/Jan.2019

Machine Learning

Time: 3 hrs. Max. Marks: 80

Note: ***Answer FIVE full questions, choosing ONE full question from each module.***

Module-1

1 a. Specify the learning task for 'A checkers learning problem'. **(03 Marks)**

b. Discuss the following with respect to the above,
   (i) Choosing the training experience.
   (ii) Choosing the target function and
   (iii) Choosing a function approximation algorithm. **(09 Marks)**

c. Comment on the issues in machine learning. **(04 Marks)**

OR

2 a. Write candidate elimination algorithm. Apply the algorithm to obtain the final version space for the training example. **(10 Marks)**

Sl. No.	Sky	Air temp	Humidity	Wind	Water	Forecast	Enjoy sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

b. Discuss about an unbiased Learner. **(06 Marks)**

Module-2

3 a. What is a decision tree & discuss the use of decision tree for classification purpose with an example. **(08 Marks)**

b. Write and explain decision tree for the following transactions: **(08 Marks)**

Tid	Refund	Martial status	Taxable Income	Cheat
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Married	120 K	No
5	No	Divorced	95 K	Yes
6	No	Married	60 K	No
7	Yes	Divorced	220 K	No
8	No	Single	85 K	Yes
9	No	Married	75 K	No
10	No	Single	90 K	Yes

OR

4 a. For the transactions shown in the table compute the following :
   (i) Entropy of the collection of transaction records of the table with respect to classification.
   (ii) What are the information gain of a_1 and a_2 relative to the transactions of the table? **(08 Marks)**

Instance	1	2	3	4	5	6	7	8	9
a_1	T	T	T	F	F	F	F	T	F
a_2	T	T	F	F	T	T	F	F	T
Target class	+	+	_	+	_	_	_	+	_

b. Discuss the decision learning algorithm. **(04 Marks)**

c. List the issues of decision tree learning. **(04 Marks)**

Module-3

5 a. Draw the perceptron network with the notation. Derive an equation of gradient descent rule to minimize the error. **(08 Marks)**

b. Explain the importance of the terms : (i) Hidden layer (ii) Generalization (iii) Overfitting (iv) Stopping criterion **(08 Marks)**

OR

6 a. Discuss the application of Neural network which is used for learning to steer an autonomous vehicle. **(06 Marks)**

b. Write an algorithm for back propagation algorithm which uses stochastic gradient descent method. Comment on the effect of adding momentum to the network. **(10 Marks)**

Module-4

7 a. What is Bayes theorem and maximum posterior hypothesis? **(04 Marks)**

b. Derive an equation for MAP hypothesis using Bayes theorem. **(04 Marks)**

c. Consider a football game between two rival teams: Team 0 and Team 1. Suppose Team 0 wins 95% of the time and Team 1 wins the remaining matches. Among the games won by team 0, only 30% of them come from playing on teams 1's football field. On the otherhand, 75% of the victories for team 1 are obtained while playing at home. If team 1 is to host the next match between the two teams, which team will most likely emerge as the winner? **(08 Marks)**

OR

8 a. Describe Brute-force MAP learning algorithm. **(04 Marks)**

b. Discuss the Naïve Bayees classifier. **(04 Marks)**

c. The following table gives data set about stolen vehicles. Using Naïve bayes classifier classify the new data (Red, SUV, Domestic) **(08 Marks)**

Table

Color	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

Module-5

9 a. Write short notes on the following:
   (i) Estimating Hypothesis accuracy.
   (ii) Binomial distribution. **(08 Marks)**

b. Discuss the method of comparing two algorithms. Justify with paired to tests method. **(08 Marks)**

OR

10 a. Discuss the K-nearest neighbor language. **(04 Marks)**

b. Discuss locally weighted Regression. **(04 Marks)**

c. Discuss the learning tasks and Q learning in the context of reinforcement learning. **(08 Marks)**

* * * * *

4a. For the transactions shown in the table compute the following.

(i) Entropy of the collection of transaction records of the table with respect to classification.

(ii) what are the information gain of a_1 and a_2 relative to the transactions of the table?

Instance	1	2	3	4	5	6	7	8	9
a_1	T	T	T	F	F	F	F	T	F
a_2	T	T	F	F	T	T	F	F	T
Target class	+	+	−	+	−	−	−	+	−

Solution:

* $$\text{Entropy}(S) \equiv -p_{\oplus}\log_2 p_{\oplus} - p_{\ominus}\log_2 p_{\ominus}$$

* $$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|}\text{Entropy}(S_v)$$

(i) Here, there are 9 instances out of which

* 4 are positive instances
* 5 are Negative instances

$$\text{Entropy}(4+, 5-) = -\left(\frac{4}{9}\right)\log_2\left(\frac{4}{9}\right) - \left(\frac{5}{9}\right)\log_2\left(\frac{5}{9}\right)$$

$$= \underline{0.9910}$$

So, $\boxed{\text{Entropy}(S) = 0.9910}$

(ii) * For attribute a_1,

Target class

a_1	+	−
T	3	1
F	1	4

Values

find the entropy for a_1. i.e., $\frac{|S_v|}{|S|}$ Entropy(S_v)

$$= \frac{4}{9}\left[-\left(\frac{3}{4}\right)\log_2\left(\frac{3}{4}\right) - \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)\right] +$$

$$\frac{5}{9}\left[-\left(\frac{1}{5}\right)\log_2\left(\frac{1}{5}\right) - \left(\frac{4}{5}\right)\log_2\left(\frac{4}{5}\right)\right]$$

$= 0.3605 + 0.4010$

$= \underline{\underline{0.7615}}$ substitute in gain formula.

ie,

$$Gain(S, a_1) = Entropy(S) - \sum_{v \in Values(a_1)} \frac{|S_v|}{|S|} Entropy(S_v)$$

$$= 0.9910 - 0.7615$$

$$\boxed{Gain(S, a_1) = 0.2295}$$

* For attribute a_2

Target class

a_2	+	−
T	2	3
F	2	2

Values

find the entropy for a_2 i.e., $\frac{|S_v|}{|S|}$ Entropy(S_v)

$$= \frac{5}{9}\left[-\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right) - \left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right)\right] + \frac{4}{9}\left[\cdot 1\right]$$

$\because$ Equal no. of +ve & -ve intances

$= 0.5394 + 0.4444$

$= \underline{0.9838}$ Substitute in Information Gain formula

i.e., $Gain(S, a_2) = Entropy(S) - \sum_{v \in Values(a_2)} \frac{|S_v|}{|S|} Entropy(S_v)$

$= 0.9910 - 0.9838$

$$\boxed{Gain(S, a_2) = 0.0072}$$

7. C. Consider a football game between two rival teams: Team0 and Team1. Suppose Team0 wins 95% of the time and Team1 wins the remaining matches. Among the games won by Team0, only 30% of them come from playing on Team1's football field. On the otherhand, 75% of the victories for Team1 are obtained while playing at home. If Team1 is to host the next match between the two teams, which team will most likely emerge as the winner?

Solution:

* Probability that Team0 wins is $P(Y_0) = \underline{0.95}$
* Probability that Team1 wins is $P(Y_1) = 1 - P(Y_A)$
$$= 1 - 0.95$$
$$= \underline{0.05}$$
* Probability that Team1 hosted the match it had won is $P(X_1|Y_1) = \underline{0.75}$
* probability that Team1 hosted the match won by Team0 is $P(X_1|Y_0) = \underline{0.30}$

The problem can be solved by computing $P(Y_1|X_1)$, which is the conditional probability that Team1 wins the next match it hosts.

Using Bayes theorem,

$$P(Y_1|X_1) = \frac{P(X_1|Y_1) * P(Y_1)}{P(X_1)}$$

$$= \frac{P(X_1|Y_1) * P(Y_1)}{P(X_1|Y_1) * P(Y_1) + P(X_1|Y_0) P(Y_0)}$$

$$= \frac{0.75 * 0.05}{(0.75 * 0.05) + (0.30 * 0.95)}$$

$$P(Y_1|X_1) = \underline{0.1162}$$

$$P(Y_0|X_1) = 1 - P(Y_1|X_1) = 1 - 0.1162$$

$$= \underline{0.8838}$$

Since $P(Y_1|X_1) < P(Y_0|X_1)$

Team0 has a better probability of winning than Team1

Note: Here, there are two events - win & host

1. win ← denoted by Y
2. Host ← denoted by X
3. Team0 ← denoted by 0
4. Team1 ← denoted by 1

8. c The following table gives data set about stolen vehicles. Using Naïve Bayes classifier classifythe new data {Color = Red, Type = SUV, Origin =Domestic}

Color	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

$$y' = \underset{y \in Y}{\text{argmax}}\ P(y) \prod_{i=1}^{m} P(x_i | y)$$

(Attributes, Values) ⇒ (Color | Red, yellow)
(Type | Sports, SUV)
(Origin | Domestic, Imported)

Target class

Values

Color	Yes	No
Red	3	2
Yellow	2	3

$P(\text{Color} = \text{Red} | \text{Stolen} = \text{yes}) = \frac{3}{5} = 0.6$

$P(\text{color} = \text{Red} | \text{Stolen} = \text{No}) = \frac{2}{5} = 0.4$

$P(\text{color} = \text{yellow} | \text{Stolen} = \text{yes}) = \frac{2}{5} = 0.4$

$P(\text{Color} = \text{yellow} | \text{Stolen} = \text{No}) = \frac{3}{5} = 0.6$

Type	Yes	No
Sport	4	2
SUV	1	3

$P(\text{Type} = \text{Sport} | \text{Stolen} = \text{yes}) = 4/5 = 0.8$

$P(\text{Type} = \text{Sport} | \text{Stolen} = \text{No}) = 2/5 = 0.4$

$P(\text{Type} = \text{SUV} | \text{stolen} = \text{yes}) = 1/5 = 0.2$

$P(\text{Type} = \text{SUV} | \text{Stolen} = \text{No}) = 3/5 = 0.6$

Target

Origin	yes	No
Domestic	2	3
Imported	3	2

Value

P(origin = Domestic | Stolen = yes) = 2/5 = 0.4
P(origin = Domestic | stolen = No) = 3/5 = 0.6
P(origin = Imported | Stolen = yes) = 3/5 = 0.6
P(origin = Imported | Stolen = no) = 2/5 = 0.4

Classify the new data = (Red, SUV, Domestic)

* For Stolen = yes :

⇒ (color = Red | stolen = yes) * (Type = SUV | stolen = yes) * (origin = Domestic | Stolem = yes) * P(yes)

⇒ 0.6 * 0.2 * 0.4 * 0.5

⇒ 0.024

* For Stolen = No :

⇒ (color = Red | Stolen = No) * (Type = SUV | stolen = No) * (Origin = Domestic | stolen = No) * P(NO)

⇒ 0.4 * 0.6 * 0.6 * 0.5

⇒ 0.072

So, we would classify the new data as not Stolen