Eisha Tir Raazia
17k-3730
Sec C.

Date: # Decision Tree Class Activity

$$Gain(S, A) = Entropy(S) - \sum [p(S|A) \cdot Entropy(S|A)]$$

(A) Attribute → Travel Cost
- a) Cheap
- b) Standard
- c) Expensive.

∴ we have :

$\boxed{E(S) = 1.5704}$ Entropy (Transportation)

Bus: $P(B) = 4/10 = 0.4$

Car: $P(C) = 3/10 = 0.3$

Train: $P(T) = 3/10 = 0.3$.

| Travelcost (A) | Transportation (B) | |
|---|---|---|
| Cheap | Bus | |
| Cheap | Bus | |
| Cheap | Train | |
| Cheap | Bus | $P(S|cheap) = 5/10 = 0.5$ |
| Cheap | Bus | $P(S|std) = 2/10 = 0.2$. |
| Standard | Train | $P(S|exp) = 3/10 = 0.3$. |
| Standard | Train | |
| Expensive | Car | |
| Expensive | Car | |
| Expensive | Car | |

For "cheap" class:
- Bus = 4/5 = 0.8
- Train = 1/5 = 0.2
- Car = 0

$$E(A_1) = -P(B|Cheap) \cdot \log_2(P(B|Cheap)) - P(T|Cheap) \cdot \log_2 P(T|Cheap) - P(C|cheap) \cdot \log_2(P(C|Cheap))$$

$$= -0.8 \cdot \log_2(0.8) - 0.2 \cdot \log_2(0.2)$$

$$= 0.257 + 0.464$$

Date:

$E(S|A=\text{cheap}) = 0.721$

for "standard" :
- Bus = 0/2 = 0
- Train = 2/2 = 1
- Car = 0/2 = 0

$$E(S|A=\text{std}) = -P(T|A=\text{std}) \cdot \log_2(T|A=\text{std}) - P(B|A=\text{std}) \cdot \log_2(P(B|A)) - P(C|A=\text{std})(\log_2(P(C|A=\text{std}))$$
$$= -1(\log_2(1)) - 0$$

$P(S|A=\text{std}) = 0$

for "Expensive" :
- Bus = 0/3 = 0
- Train = 0/3 = 0
- Car = 3/3 = 1

$$E(S|A=\text{Exp.}) = -P(T|A=\text{Exp.})(\log_2(T|A=\text{exp.})) - P(B|\text{exp}).\log_2(B|\text{exp}) - P(C|A=\text{exp}).(\log_2(C|A=\text{exp}))$$
$$= -0-0-1(\log_2(1))$$

$E(S|A=\text{exp.}) = 0.$

Now,

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum[P(S|A) \times \text{Entropy}(S|A)]$$
$$= 1.5704 - [0.5 \times 0.721] - [0.2 \times 0] - [0.3 \times 0]$$
$$= 1.5704 - 0.3605$$
$$= 1.2099$$

$\text{Gain}(S, A) = 1.21$ Ans.