Example

Independent/Condition attributes					Dependent/ Decision attributes
Animal	**Warm-blooded**	**Feathers**	**Fur**	**Swims**	**Lays Eggs**
Ostrich	Yes	Yes	No	No	Yes
Crocodile	No	No	No	Yes	Yes
Raven	Yes	Yes	No	No	Yes
Albatross	Yes	Yes	No	No	Yes
Dolphin	Yes	No	No	Yes	No
Koala	Yes	No	Yes	No	No

$$Entropy(S) = \sum_{i=1}^{c} - p_i \log_2 p_i$$

Entropy(4Y,2N): $-(4/6)\log_2(4/6) - (2/6)\log_2(2/6)$

$= 0.91829$

Now, we have to find the IG for all four attributes

Warm-blooded, Feathers, Fur, Swims

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

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

For attribute ‘Warm-blooded’:

Values(Warm-blooded) : [Yes,No]

S = [4Y,2N]

S_{Yes} = [3Y,2N] $E(S_{Yes})$ = 0.97095

S_{No} = [1Y,0N] $E(S_{No})$ = 0 (all members belong to same class)

Gain(S,Warm-blooded) = 0.91829 – [(5/6)*0.97095 + (1/6)*0]

= 0.10916

For attribute ‘Feathers’:

Values(Feathers) : [Yes,No]

S = [4Y,2N]

S_{Yes} = [3Y,0N] $E(S_{Yes})$ = 0

S_{No} = [1Y,2N] $E(S_{No})$ = 0.91829

Gain(S,Feathers) = 0.91829 – [(3/6)*0 + (3/6)*0.91829]

= 0.45914

For attribute ‘Fur’:

Values(Fur) : [Yes,No]

S = [4Y,2N]

S_{Yes} = [0Y,1N] $E(S_{Yes}) = 0$

S_{No} = [4Y,1N] $E(S_{No}) = 0.7219$

Gain(S,Fur) = 0.91829 – [(1/6)*0 + (5/6)*0.7219]= 0.3167

For attribute ‘Swims’:

Values(Swims) : [Yes,No]

S = [4Y,2N]

S_{Yes} = [1Y,1N] $E(S_{Yes}) = 1$ (equal members in both classes)

S_{No} = [3Y,1N] $E(S_{No}) = 0.81127$

Gain(S,Swims) = 0.91829 – [(2/6)*1 + (4/6)*0.81127]= 0.04411

Gain(S,Warm-blooded) = 0.10916

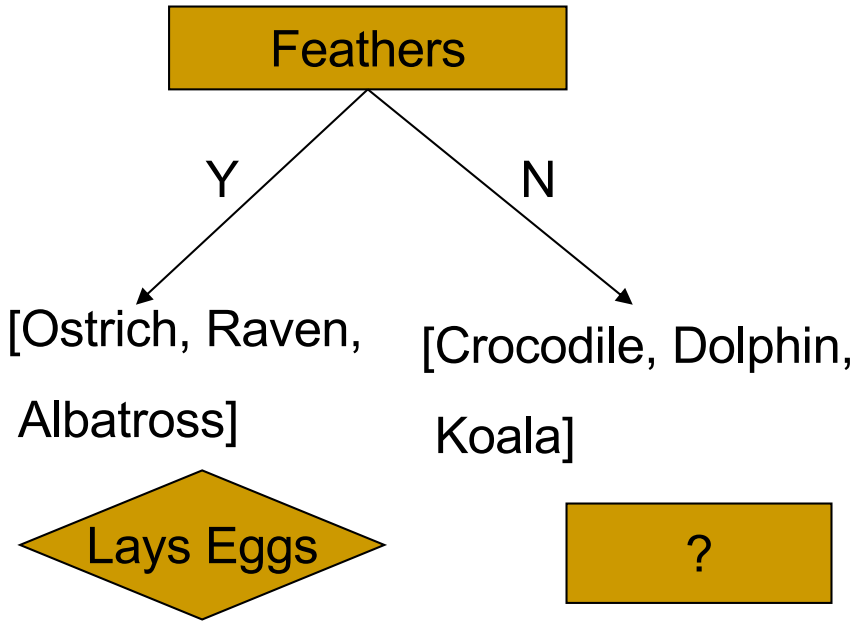
Gain(S,Feathers) = 0.45914

Gain(S,Fur) = 0.31670

Gain(S,Swims) = 0.04411

Gain(S,Feathers) is maximum, so it is considered as the root node

Animal	Warm-blooded	Feathers	Fur	Swims	Lays Eggs
Ostrich	Yes	Yes	No	No	Yes
Crocodile	No	No	No	Yes	Yes
Raven	Yes	Yes	No	No	Yes
Albatross	Yes	Yes	No	No	Yes
Dolphin	Yes	No	No	Yes	No
Koala	Yes	No	Yes	No	No



Animal	Warm-blooded	Feathers	Fur	Swims	Lays Eggs
Crocodile	No	No	No	Yes	Yes
Dolphin	Yes	No	No	Yes	No
Koala	Yes	No	Yes	No	No

We now repeat the procedure,

S: [Crocodile, Dolphin, Koala]

S: [1+,2-]

$$Entropy(S) = \sum_{i=1}^{c} - p_i \log_2 p_i$$

Entropy(S) = $-(1/3)\log_2(1/3) - (2/3)\log_2(2/3)$

= 0.91829

- For attribute ‘Warm-blooded’:

Values(Warm-blooded) : [Yes,No]

S = [1Y,2N] S_{Yes} = [0Y,2N] $E(S_{Yes}) = 0$

S_{No} = [1Y,0N] $E(S_{No}) = 0$

Gain(S,Warm-blooded) = 0.91829 – [(2/3)*0 + (1/3)*0] = **0.91829**

- For attribute ‘Fur’:

Values(Fur) : [Yes,No]

S = [1Y,2N] S_{Yes} = [0Y,1N] $E(S_{Yes}) = 0$

S_{No} = [1Y,1N] $E(S_{No}) = 1$

Gain(S,Fur) = 0.91829 – [(1/3)*0 + (2/3)*1] = **0.25162**

- For attribute ‘Swims’:

Values(Swims) : [Yes,No]

S = [1Y,2N] S_{Yes} = [1Y,1N] $E(S_{Yes}) = 1$

S_{No} = [0Y,1N] $E(S_{No}) = 0$

Gain(S,Swims) = 0.91829 – [(2/3)*1 + (1/3)*0] = **0.25162**

Gain(S,Warm-blooded) is maximum

The final decision tree will be: