Artificial Intelligence – Assignment 5

1. [25 pts.] Assume that you are given the set of labeled training examples below, where each of three features has four possible values: a, b, c, or d. You choose to apply the ID3 decision tree induction algorithm to this data.

	F1	F2	F3	Output
ex1	a	a	b	-
ex2	b	c	d	+
ex3	b	b	a	+
ex4	c	c	a	-
ex5	a	a	b	+
ex6	c	d	c	-
ex7	c	b	d	-

⇨ 1. $I(T) = \sum_{i=1}^{\{k\}} \left[\frac{|T_{C_i}|}{T}\right] X \log_2 \frac{|T_{C_i}|}{T}$

$I(T) = - 3/7\log_2 3/7 - 4/7\log_2 4/7 = 0.523 + 0.461 = 0.984$

2. Test F1:

1. $I(T_{\{F1 <-- a\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
2. $I(T_{\{F1 <-- b\}}) = -2/2\log_2 2/2 = 0$
3. $I(T_{\{F1 <-- c\}}) = -0/3\log_2 0/3 = 0$

$I(F1, T) = 2/7. I(T_{\{F1 <-- a\}}) + 2/7. I(T_{\{F1 <-- b\}}) + 3/7 I(T_{\{F1 <-- c\}})$
$= 2/7.1 + 0 + 0$
$= 0.285$

$Gain = I(T) – I(F1, T) = 0.984 – 0.285 = 0.699$

3. Test F2:

1. $I(T_{\{F2 <-- a\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
2. $I(T_{\{F2 <-- b\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
3. $I(T_{\{F2 <-- c\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
4. $I(T_{\{F2 <-- d\}}) = -0/1\log_2 0/1 = 0$

$I(F2, T) = 2/7. I(T_{\{F2 <-- a\}}) + 2/7. I(T_{\{F2 <-- b\}}) + 2/7 I(T_{\{F2 <-- c\}}) + 1/7 I(T_{\{F2 <-- d\}})$
$= 2/7.1 + 2/7.1 + 2/7.1 + 1/7.0$
$= 0.857$

$Gain = I(T) – I(F2, T) = 0.984 – 0.857 = 0.127$

4. Test F3:

1. $I(T_{\{F3 <-- a\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
2. $I(T_{\{F3 <-- b\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$
3. $I(T_{\{F3 <-- c\}}) = -0/1\log_2 0/1 = 0$
4. $I(T_{\{F3 <-- d\}}) = -1/2\log_2 1/2 - 1/2\log_2 1/2 = ½ + ½ = 1$

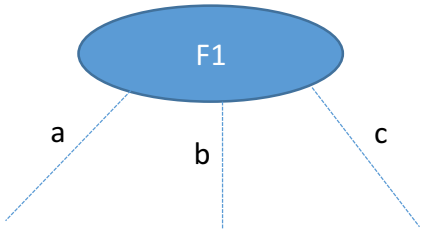
$I(F3, T) = 2/7.\ I(T_{\{F3 <-- a\}}) + 2/7.\ I(T_{\{F3 <-- b\}}) + 0/7\ I(T_{\{F3 <-- c\}}) + 2/7\ I(T_{\{F3 <-- d\}})$
$= 2/7.1 + 2/7.1 + 0 + 2/7.1 = 6/7 = 0.857$

Gain = I(T) – I(F3, T) = 0.984 – 0.857 = 0.127

Comparing Gains:

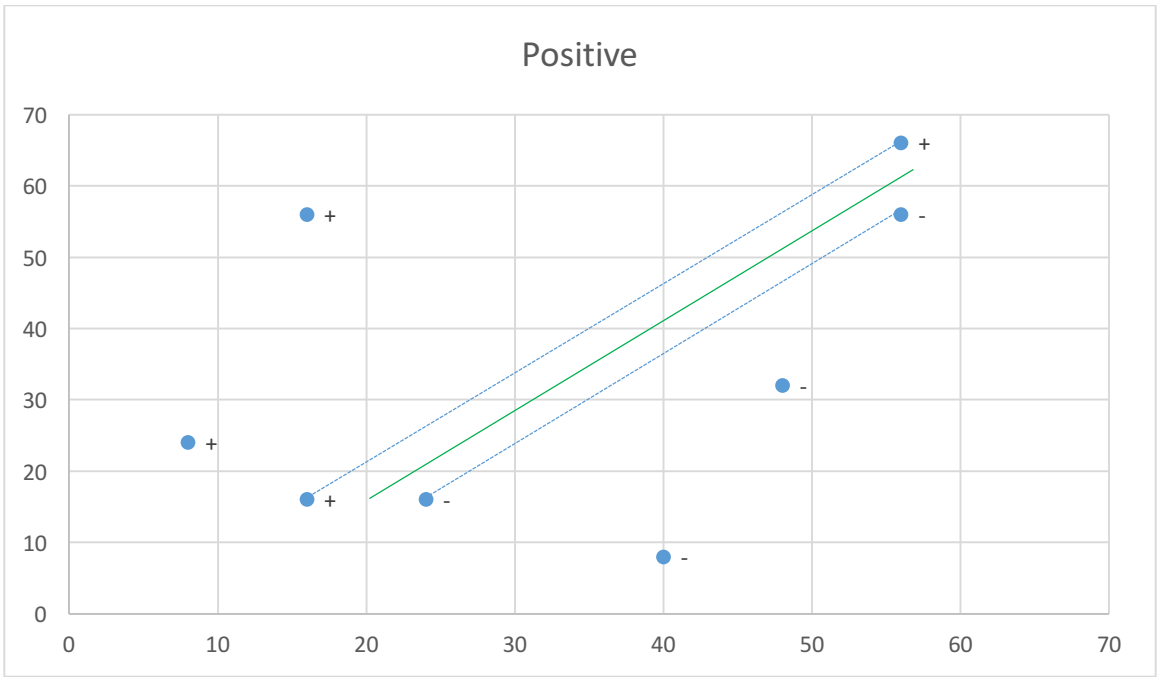
Gain(F1) > Gain(F2) = Gain(F3) :- 0.699 > 0.127 = 0.127

Hence the root attribute can be selected as F1 as it has the maximum gain.



2. [25 pts.] Derive the equation for the maximum margin separating hyperplane that a Support Vector Machine would find to classify the following set of points.

- **Positive:** (56,66), (16,16), (16,56), (8,24)
- **Negative:** (48,32), (40,8), (24,16), (56,56)


Positive
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
+
+
+
-
-
-
+
-