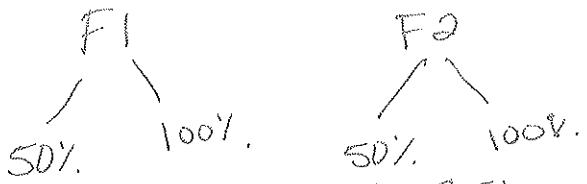
CS691 Quiz 2 Name Solution

1. Why would it be bad to use accuracy to generate a decision tree on the following data? (3 points)

Sample	F1	F2	Label
1	1	1	1
2	1	0	0
3	0	1	1
4	0	0	1



75% overall 75%.

If we always say "no" we also get 75%. So we wouldn't even build a tree.

2. What is the best accuracy a binary decision tree can achieve on the following data? How do you know? (3 points)

Sample	F1	F2	Label
1	0	0	1
2	0	1	0
3	1	0	1
4	1	1	1
5	1	0	0

3 1 0 1
5 1 0 0
> can only get one of these right.
So 4/5

3. What are the minimum and maximum values of entropy in a binary classification problem? Explain. (2 points) max = 1 min = 0.

Two extreme cases: ① one class has P(c) = 1 (so all others are 0)
② each class has equal P(c).

P(c) = 1 → H = 0.

all same P(c) → H = 1.

4. Given that we have 4 binary features, what is the minimum number of samples we would need to cover all possible combinations of features? Explain. (2 points)

$2^4 = 16$

every combination of 4 binary features.
1 binary feature is $2^1 \to 2$
2 binary features $2^2 \to 4$ ...

CS491 Quiz 2 Name Solution

1. Using the greedy algorithm learned in class (with Information Gain), generate the best depth-1 decision tree for the given training data. What accuracy does your model achieve on the training data? Show your work. (8 points)

Sample	F1	F2	F3	Label
1	0	0	0	1
2	0	0	1	0
3	1	0	1	1
4	1	1	1	0
5	1	0	0	0

$H = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.52 + 0.44$

$H(0) = 0.96$

F1

1,2 — H=1

3,4,5 — $H = 0.52 + 0.38 = 0.90$

$IG = 0.96 - \frac{2}{5}\cdot 1 - \frac{3}{5}\cdot 0.90$

$= 0.96 - 0.4 - 0.54 = 0.02$

F2

1,2,3,5 — $H = 1$

4 — $H = 0$

$IG = 0.96 - \frac{4}{5}\cdot 1 - \frac{1}{5}\cdot 0$

$= 0.96 - 0.8 = 0.16$

F3

1,5 — H=1

2,3,4 — $H = 0.90$

IG = Same as F3.

F2

0 → YES

1 → NO

accuracy = $\frac{3}{5}$

2. How would the model you generated in problem 1 classify the following sample? (2 points)

Sample	F1	F2	F3	Label
6	1	0	1	0

F2 = 0 → YES

or ∅