Question 1: What is the entropy of this collection of training examples with respect to the target class? (3 points)

T = 6/10

F = 4/10

Entropy = $-(0.6\log2(0.6)+0.4\log2(0.4))$

$=-(0.6\times-0.73697+0.4\times-1.32193)$

$= 0.97095$

The entropy of this collection of training examples with respect to the target class is 0.97095

Question 2: What are the different options for the first split when constructing your decision tree? (3 points)

The options are:

- X1=0 and X1=1
- X2=0 and X2!=0 (x={1,2})
- X2=1 and X2!=1 (x={0,2})
- X2=2 and X2!=2 (x={0,1})

Question 3: For each potential first split option, compute the information gain. Only provide the results, there is no need to provide your calculations (3 points).

For X1=0:

(2T and 3F)

X1 != 0

(4T and 1F)

Information Gain = $0.971-0.846465 = 0.124535$

For X2=0:

(4T and 0F)

X2!=0:

(2T and 4F)

Information Gain = 0.971−0.55098 = 0.42002

For X2=1:

(1T and 2F)

X2 != 1:

(5T and 2F)

Information Gain = 0.971 – 0.87967 = 0.09133

For X2=2:

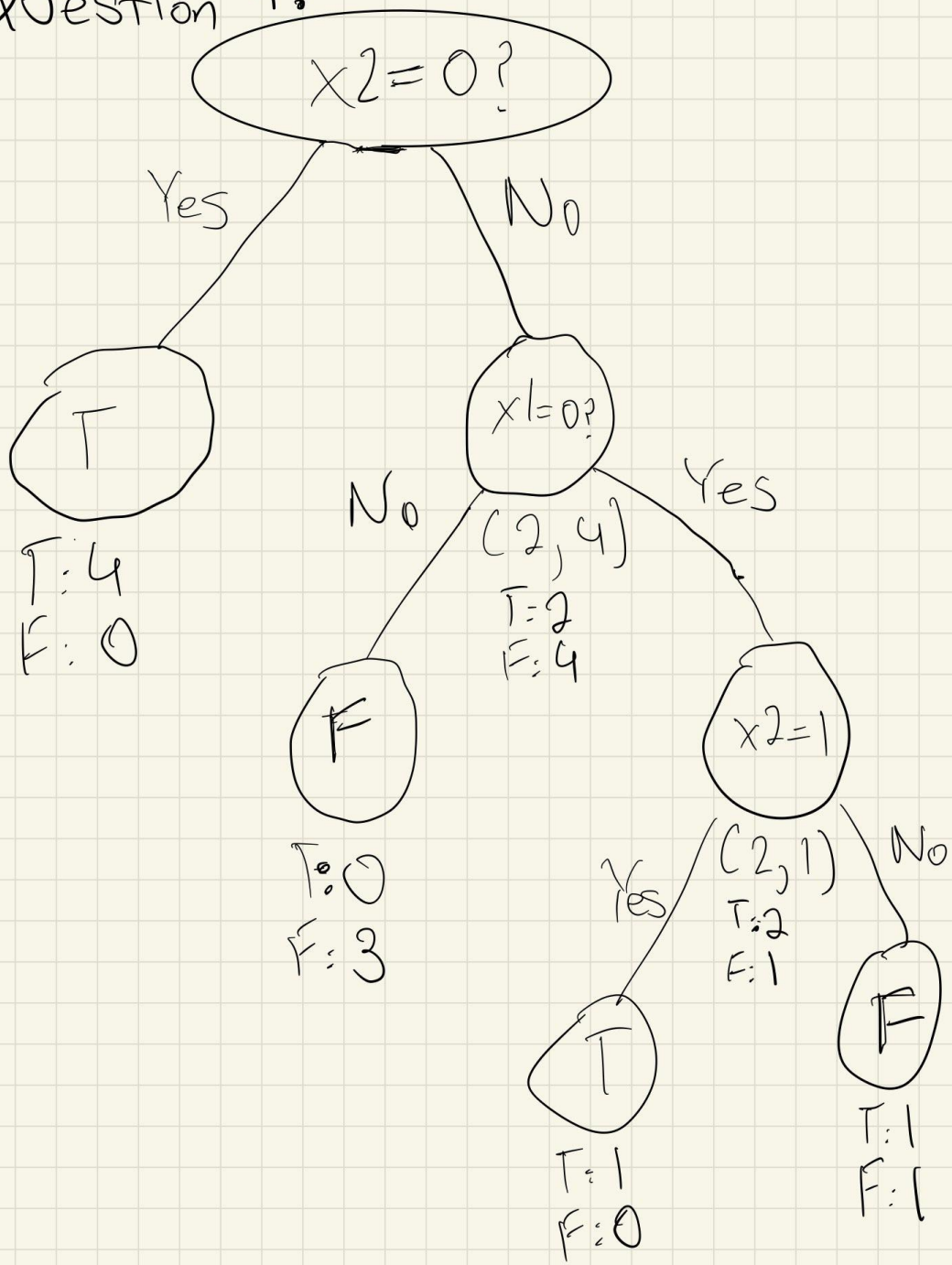
(1T and 2F)

X2!=2:

(5T and 2F)

Information Gain = 0.9704 – 0.87967 = 0.09073

Question 4:


x2=0?
Yes
No
T
T: 4
F: 0
x1=0?
No
(2, 4)
T: 2
F: 4
Yes
F
T: 0
F: 3
x2=1
(2, 1)
T: 2
F: 1
Yes
No
T
T: 1
F: 0
F
T: 1
F: 1