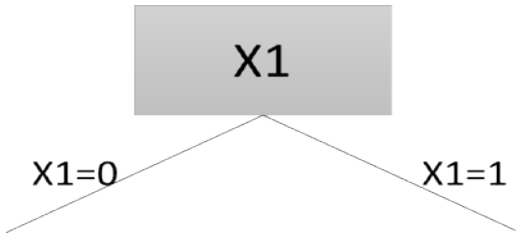
Boolean operation (OR)
Sepaseh Hakiminejad

	x1	x2	x3	OR (x1,x2,x3)
1	0	0	0	FALSE
2	0	0	1	TRUE
3	0	1	0	TRUE
4	0	1	1	TRUE
5	1	0	0	TRUE
6	1	0	1	TRUE
7	1	1	0	TRUE
8	1	1	1	TRUE

P(y=True)=7/8
P(y=False)=1/8
E(y)= - 7/8*log(7/8) - 1/8*log(1/8) = 0.544

X1		X2		X3	
P(x1=1,y=True)= 4/4=1		P(x2=1,y=True)= 4/4=1		P(x3=1,y=True)= 4/4=1	
P(x1=1,y=False)= 0/4=0		P(x2=1,y=False)= 0/4=0		P(x3=1,y=False)= 0/4=0	
E(x1=1) = - 1*log(1) - 0*log(0)	0	E(x2=1) = - 1*log(1) - 0*log(0)	0	E(x3=1) = - 1*log(1) - 0*log(0)	0
P(x1=0,y=True)= 3/4		P(x2=0,y=True)= 3/4		P(x3=0,y=True)= 3/4	
P(x1=0,y=False)= 1/4		P(x2=0,y=False)= 1/4		P(x3=0,y=False)= 1/4	
E(x1=1) = - 3/4*log(3/4) - 1/4*log(1/4)	0.811	E(x2=1) = - 3/4*log(3/4) - 1/4*log(1/4)	0.811	E(x3=1) = - 3/4*log(3/4) - 1/4*log(1/4)	0.811
E(y|x1) = 4/8*0 + 4/8*0.81 =	**0.405**	E(y|x2) = 4/8*0 + 4/8*0.81 =	**0.405**	E(y|x3) = 4/8*0 + 4/8*0.81 =	**0.405**
IG(x1) = 0.544 - 0.405 =	**0.139**	IG(x2) = 0.544 - 0.405 =	**0.139**	IG(x3) = 0.544 - 0.405 =	**0.139**

The Entropy and Information Gain is equal for all three nodes. So we do not have any other parameter to choose. I choose X1.

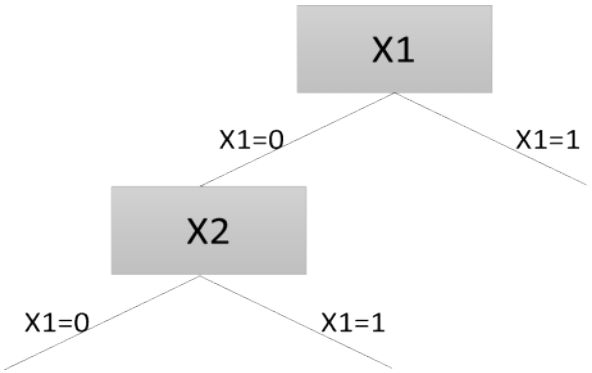


Now we shoud decide to choose x2 or x3 according to x1=0

	x1	x2	x3	OR (x1,x2,x3)
1	0	0	0	FALSE
2	0	0	1	TRUE
3	0	1	0	TRUE
4	0	1	1	TRUE

P(y=True)=3/4
P(y=False)=1/4
E(y)= - 3/4*log(3/4) - 1/4*log(1/4) = 0.811

X2		X3	
P(x2=1,y=True)= 2/2 =1		P(x3=1,y=True)= 2/2=1	
P(x2=1,y=False)= 0/2=0		P(x3=1,y=False)= 0/2=0	
E(x1=1) = - 1*log(1) - 0*log(0)	0	E(x1=1) = - 1*log(1) - 0*log(0)	0
P(x2=0,y=True)= 1/2		P(x3=0,y=True)= 1/2	
P(x2=0,y=False)= 1/2		P(x3=0,y=False)= 1/2	
E(x2=1) = - 1/2*log(1/2) - 1/2*log(1/2)	1	E(x3=1) = - 1/2*log(1/2) - 1/2*log(1/2)	1
E(y|x2) = 2/4*0 + 2/4*1 =	**0.500**	E(y|x3) = 2/4*0 + 2/4*1 =	**0.500**
IG(x2) = 0.811 - 0.500 =	**0.311**	IG(x3) = 0.811 - 0.500 =	**0.311**



The Entropy and Information Gain is equal for both nodes. I choose X2.

Now we shoud decide to choose x2 or x3 according to x1=1

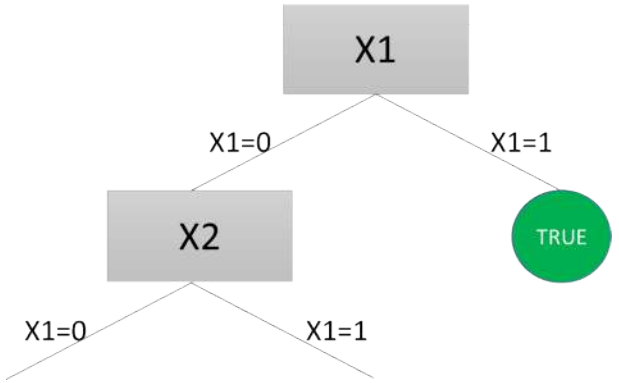
	x1	x2	x3	OR (x1,x2,x3)
5	1	0	0	TRUE
6	1	0	1	TRUE
7	1	1	0	TRUE
8	1	1	1	TRUE

P(y=True)=4/4=1

P(y=False)=0/4=0

E(y)= - 1*log(1) - 0*log(0) = 0

Entropy is zero which means all the labels are equal, the leaf is "True"



When the x2=0:

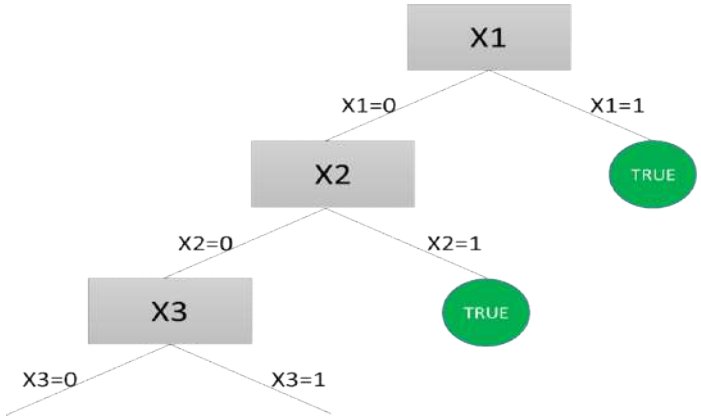
	x1	x2	x3	OR (x1,x2,x3)
1	0	0	0	FALSE
2	0	0	1	TRUE

P(y=True)= 1/2

P(y=False)= 1/2

E(y)= - 1/2*log(1/2) - 1/2*log(1/2) = 1

We have only one node to complete. So we choose X3.



When the x2=1:

	x1	x2	x3	OR (x1,x2,x3)
3	0	1	0	TRUE
4	0	1	1	TRUE

P(y=True)= 2/2=1

P(y=False)= 0/2=0

E(y)= - 1*log(1) - 0*log(0) = 0

Entropy is zero which means all the labels are equal, the leaf is "True"

	x1	x2	x3	OR (x1,x2,x3)
1	0	0	0	FALSE

	x1	x2	x3	OR (x1,x2,x3)
2	0	0	1	TRUE

Entropy is 0 for both of them.