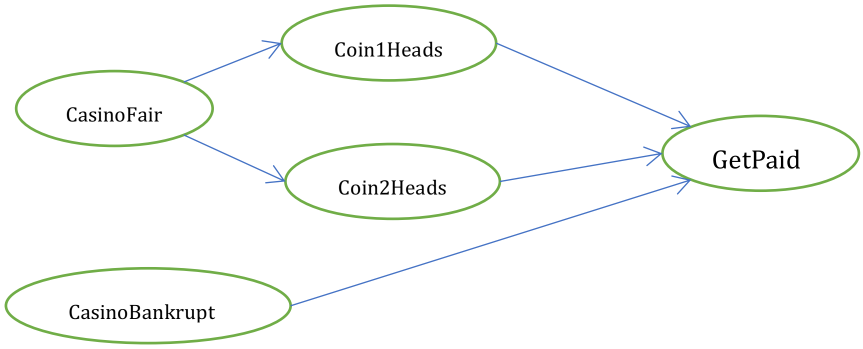
1.

Solutions:

I)



CasinoFair:

CasinoFair	p
T	0.4
F	0.6

CasinoBankrupt:

CasinoBankrupt	p
T	0.1
F	0.9

Coin1Heads:

CasinoFair	Coin1Heads	p
T	T	0.5
T	F	0.5
F	T	0.3
F	F	0.7

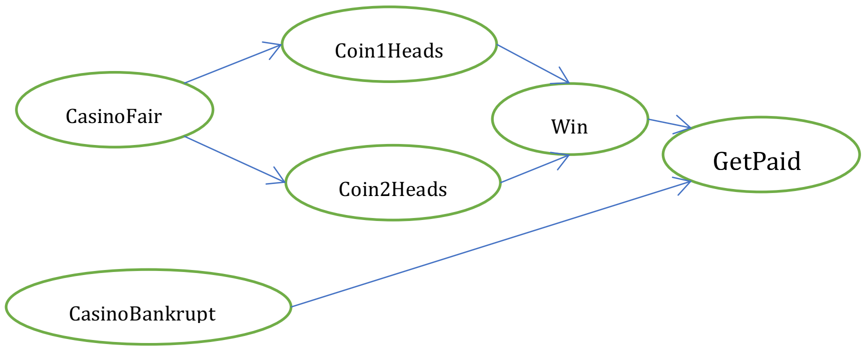
Coin2Heads:

CasinoFair	Coin2Heads	p
T	T	0.5
T	F	0.5
F	T	0.3
F	F	0.7

GetPaid:

CasinoBankrupt	Coin1Heads	Coin2Heads	GetPaid	p
T	T	T	T	0.2
T	T	T	F	0.8
...	...	...	...	0
F	T	T	T	0.8
F	T	T	F	0.2
...	...	...	...	0

or



Win:

Coin1Heads	Coin2Heads	Win	p
T	T	T	1
...	...	...	0

GetPaid:

CasinoBankrupt	win	GetPaid	p
T	T	T	0.2
T	T	F	0.8
...	...	...	0
F	T	T	0.8
F	T	F	0.2
...	...	...	0

II) P(CasinoFair, CasinoBankrupt, Coin1Heads, Coin2Heads, GetPaid)=
P(CasinoFair)P(CasinoBankrupt)P(Coin1Heads|CasinoFair)P(Coin2Heads|CasinoFair)
P(GetPaid|Coin1Heads, Coin2Heads, CasinoBankrupt)

Or

P(CasinoFair, CasinoBankrupt, Coin1Heads, Coin2Heads, Win, GetPaid)=
P(CasinoFair)P(CasinoBankrupt)P(Coin1Heads|CasinoFair)P(Coin2Heads|CasinoFair)
P(Win|Coin1Heads, Coin2Heads)P(GetPaid|Win, CasinoBankrupt)

III) Markov Blanket of variable Coin2Heads={CasinoFair, CasinoBankrupt, Coin1Heads, GetPaid}

Or

Markov Blanket of variable Coin2Heads={CasinoFair, Coin1Heads, Win}

IV) 0.74

$P(GetPaid = true|Coin1Heads = true \wedge Coin2Heads = true)$

$= P(GetPaid| Coin1Heads = true \wedge Coin2Heads = true \wedge CasinoBankrupt = true) * P(CasinoBankrupt = true) + P(GetPaid| Coin1Heads = true \wedge Coin2Heads = true \wedge CasinoBankrupt = false) * P(CasinoBankrupt = false)$

$= 0.2 * 0.1 + 0.8 * 0.9$

$= 0.74$

2.

Solutions:

(i) In this case each $p_i = 1/256$ and the ensemble entropy summation extends over 256

such equiprobable grey values, so $H = -(256)(1/256)(-8) = 8$ bits.

(ii) Since all humans are in this category (humans $\subset$ mammals), there is no uncertainty

about this classification and hence the entropy is 0 bits.

(iii) The entropy of this distribution is $-(1/4)(-2) - (1/4)(-2) - (1/2)(-1) = 1.5$ bits.

(iv) By the definition of median, both classes have probability 0.5, so the entropy is 1 bit.

3. (2pt) Let p(x, y) be as shown in the table below.

X\Y	0	1	2
0	1/12	1/6	1/12
1	1/6	1/6	1/6
2	0	1/12	1/12

Find

(a) H(X), H(Y),

(b) H(X,Y)

(c) H(Y|X)

(d) I(X;Y)

(e) Draw a Venn diagram for the quantities in (a) through (d)

Solution:

(a) P(x=0)=1/12+1/6+1/12=1/3, p(x=1)=1/6+1/6+1/6=1/2, P(x=3)=1/12+1/12=1/6

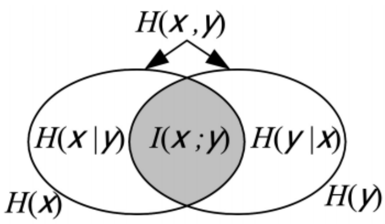
So, H(X) = -1/3log1/3 – 1/2log1/2 -1/6log1/6 ≈1.46
P(Y=0) = 1/4, P(Y=1)=5/12, P(Y=2) = 1/3
So, H(Y) = -1/4log1/4 – 5/12log5/12 – 1/3log1/3 ≈1.55

(b) H(X,Y)= -1/12log1/12 -1/6log1/6-1/12log1/12
-1/6log1/6-1/6log1/6-1/6log1/6
-1/12log1/12-1/12log1/12
≈2.918

(c) H(Y|X) = H(X, Y)-H(X) ≈1.459
或者
H(Y|X) = -1/12log1/4-1/6log1/2-1/12log1/4
-1/6log1/3-1/6log1/3-1/6log1/3
-0-1/12log1/2-1/12log1/2
≈1.459

(d) I(X;Y) = H(X) + H(Y) – H(X, Y) ≈0.096

(e)

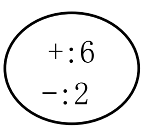


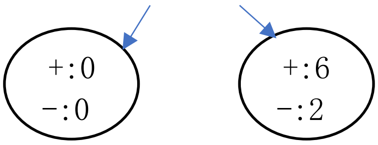
4. (1pt) We have a dataset in the following table where A, B denote attributes and Y denotes labels. We want to build a decision tree to classify them according to Y.

Y	A	B
−	1	0
−	1	0
+	1	0
+	1	0
+	1	1
+	1	1
+	1	1
+	1	1

Which attribute should be selected for the next split? Give your explanation.
1) A
2) B
3) A or B (tie)
4) Neither

Solution:
按 A 划分：



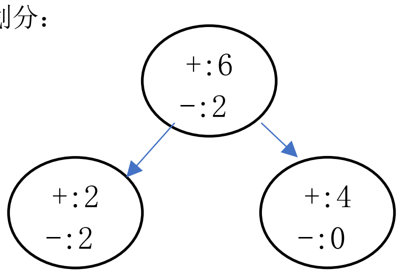


H(Y)=-1/4log1/4-3/4log3/4=0.81

H(Y|left)=0, H(Y|right)=H(Y)=-1/4log1/4-3/4log3/4=0.81

IG(Y; A) = 0.81-0.81x1=0

按 B 划分：



H(Y) = -1/4log1/4-3/4log3/4 = 0.81,

H(Y|left) = -1/2log1/2-1/2log1/2=1, H(Y|right) = 0

IG(Y; B) = 0.81 – (1x1/2+0x1/2)=0.31

IG(Y;B)>IG(Y;A)，所以按 B 划分