ML-2

Homework Assignment - 2

Shantanu Chandra
(12048461)

Q1.

(1) Mutual Inform. of $I(X:Y)$ is defined as :

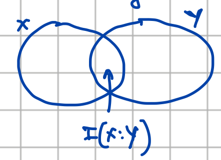
$$I(X:Y) = KL\left(p(X,Y) \,\|\, p(X)\cdot p(Y)\right)$$

Using definition of KL :

(discrete) $$I(X:Y) = \sum_x \sum_y p(x,y) \ln \frac{p(x,y)}{p(x)\,p(y)}$$

(continuous) $$I(X:Y) = \int_x \int_y p(x,y) \ln \frac{p(x,y)}{p(x)\,p(y)} \, dx\, dy$$

This intuitively says how much one random variable tells us about the other, OR in our case how much we reduce uncertainty of X if we knew Y. This implies that if X & Y are dependent, how much information is shared between them and can be inferred in the form of a Venn Diag. :



Conditional mutual information $I(X:Y|Z)$ is defined as :

$$I(X:Y|Z) = KL\left(p(x,y,z) \,\|\, p(x|z)\,p(y|z)\,p(z)\right)$$

Using definition of KL :

(discrete) $$I(X:Y|Z) = \sum_z p(z) \sum_x \sum_y p(x,y|z) \ln \frac{p(x,y|z)}{p(x|z)\,p(y|z)}$$

(continuous) $$I(X:Y|Z) = \int_z p(z) \int_x \int_y p(x,y|z) \ln \frac{p(x,y|z)}{p(x|z)\,p(y|z)} \, dx\, dy$$

This intuitively measures the mutual inform. between 2 distributions/random variables given a third. Hence, if both the distributions are independent of the third, $CMI(X:Y|Z) = MI(X:Y)$

// Ans.

2

x	y	z	p(x,y,z)
0	0	0	0.192
0	0	1	0.144
0	1	0	0.048
0	1	1	0.216
1	0	0	0.192
1	0	1	0.064
1	1	0	0.048
1	1	1	0.096

x	p(x)
0	0.6
1	0.4

y	p(y)
0	0.592
1	0.408

z	p(z)
0	
1	

x	y	p(x,y)
0	0	0.336
0	1	0.264
1	0	0.256
1	1	0.144

where we have used,

$$p(x) = p(x, y=0) + p(x, y=1)$$
$$p(y) = p(y, x=0) + p(y, x=1)$$
$$p(x,y) = p(x,y,z=0) + p(x,y,z=1)$$

Now,

$$I(X:Y) = KL(p(x,y) \,||\, p(x)\cdot p(y)) = \sum_x \sum_y p(x,y) \ln \frac{p(x,y)}{p(x)\,p(y)}$$

an example of this calculation for $x=0$, $y=0$:

$$I(x=0 : Y=0) = 0.336 \times \ln \frac{(0.336)}{(0.6)(0.592)} = 0.336 \times (-0.0555)$$
$$= -0.01867$$

Ans.

Doing this for all x, y we get:

x	y	I(x,y)
0	0	-0.0187
0	1	0.0199
1	0	0.02
1	1	-0.0180

Ans.

Thus, $I(X:Y) = \sum I(x,y)$
$= 0.0032 > 0$

Proved

$I(X:Y) > 0$ implies that there is some shared information between the distributions $p(x)$ & $p(y)$. This in turn means that the random variables X & Y are dependent

Ans.

3