CS4103 Quiz-2

Solutions for questions on problem solving

Q2. Given the full joint distribution shown below,

	toothache		¬*toothache*	
	catch	¬*catch*	*catch*	¬*catch*
cavity	0.108	0.012	0.072	0.008
¬*cavity*	0.016	0.064	0.144	0.576

calculate *P(cavity | toothache ∨ catch)* up to two decimal places.

Ans. $$\boldsymbol{P(cavity \mid toothache \vee catch) = \frac{P(cavity \wedge (toothache \vee catch))}{P(toothache \vee catch)}}$$

$$= \frac{(0.108 + 0.012 + 0.072)}{(0.108 + 0.012 + 0.016 + 0.064 + 0.072 + 0.144)}$$

$$= \frac{0.192}{0.416} = 0.46 \text{ (up to two decimal places)}$$

Q6.

x1	x2	x3	y
L	F	F	F
L	T	T	T
M	F	F	F
M	T	T	T
H	F	F	T
H	F	T	T

i) Dimensionality of the given dataset: 3 (number of input features)

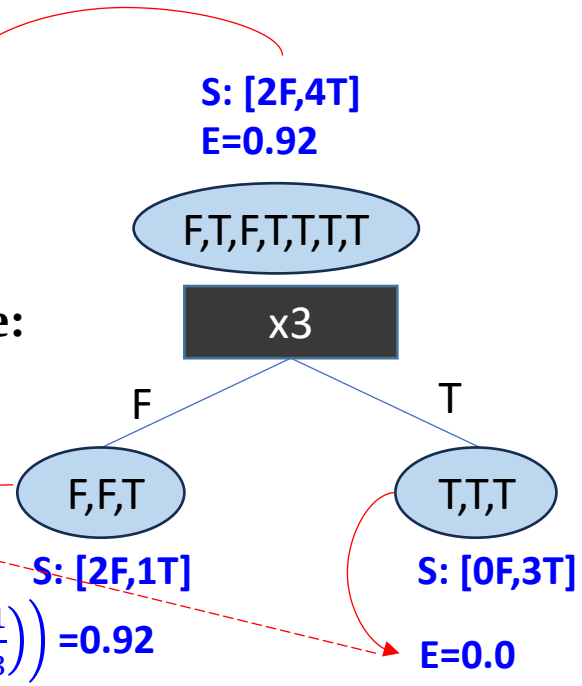
ii) Entropy before any split:

$$-\left(\left(\frac{2}{6}\right)\log_2\left(\frac{2}{6}\right)+\left(\frac{4}{6}\right)\log_2\left(\frac{4}{6}\right)\right)=\mathbf{0.92}$$

iii) Information gain after the split based on attribute x3 will be:

$$0.92-((0.5*0.92)+(0.5*0.0))$$

=0.46



iv)

HammingDist(<L,F,F>,<M,F,T>)= 2
HammingDist (<L,T,T>,<M,F,T>)= 2
HammingDist (<M,F,F>,<M,F,T>)= 1
HammingDist (<M,T,T>,<M,F,T>)= 1
HammingDist (<H,F,F>,<M,F,T>)= 2
HammingDist (<H,F,T>,<M,F,T>)= 1

The 3 nearest neighbors of <MFT> are: <M,F,F>, <M,T,T>, <H,F,T> which correspond to the class/label: F, T, and T, respectively.

Hence, as per 3-NN, the predicted class/label of <M,F,T> sample is T

v)

The total number of parameters (including bias) in the said neural network will be:

$$(3\times 6)+(6\times 1)+6+1=31$$

Associated with bias

vi)

Actual Labels $y=\{F,T,F,T,T,T\}$

Predicted Labels $\hat{y}=\{F,F,T,T,F,T\}$

Confusion matrix:

Actual \ Predicted	T (+ve)	F (-ve)
T (+ve)	2 [TP]	2 [FN]
F (-ve)	1 [FP]	1 [TN]

Precision= $\frac{TP}{TP+FP}=\frac{2}{3}=0.67$

Recall= $\frac{TP}{TP+FN}=\frac{2}{4}=0.50$