'Assignment -04'

Amirun Nahin

ID: 23201416

Sec: 18

Course: CSE422

Semester: Fall '2025

Date: 09-01-2026

Answer to the que. no- 1

(a)

Given x_1, x_3, y belong to a multinomial distribution. So, we can calculate the model parameter using the probability table;

	0	1
y	$4/10$	$6/10$

x_1	$y=0$	$y=1$
$x_1=0$	$3/4$	$1/6$
$x_1=1$	$0/4$	$2/6$
$x_1=2$	$1/4$	$3/6$

x_3	$y=0$	$y=1$
$x_3=0$	$1/4$	$4/6$
$x_3=1$	$3/4$	$2/6$

Now, x_2 belongs to a normal distribution. So, model parameter for x_2 are,

$$\mu_{01} = \frac{\sum_{i=1}^{N} 1\{y^{(i)}=0\} \cdot x_2^{(i)}}{\sum_{i=1}^{N} 1\{y=0\}}$$

$$= \frac{(28.5 + 29.2 + 19.2 + 27.6)}{4}$$

$$= 26.125$$

$$\sigma_0^2 = \frac{\sum_{i=1}^{N} 1\{y=0\} \cdot (x_2^{(i)} - \mu_0)^2}{\sum_{i=1}^{n} 1\{y=0\}}$$

$$= \frac{(28.5-26.125)^2 + (29.2-26.125)^2 + (19.2-26.125)^2 + (27.6-26.125)^2}{4}$$

$$= \frac{65.2275}{4}$$

$$= 16.307$$

$$\mu_1 = \frac{\sum_{i=1}^{n} 1\{y=1\} \cdot x_2^{(i)}}{\sum_{i=1}^{n} 1\{y=1\}}$$

$$= \frac{(25.4+21.1+18.5+22.8+28.9+21.8)}{6}$$

$$= 23.083$$

$$\sigma_1^2 = \frac{\sum_{i=1}^{n} 1\{y=1\} (x_2^{(i)} - \mu_1)^2}{\sum_{i=1}^{n} 1\{y=1\}}$$

$$= \frac{(25.4-23.083)^2 + (21.1-23.083)^2 + (22.8-23.083)^2 + (18.5-23.083)^2 + (28.9-23.083)^2 + (21.8-23.083)^2}{6}$$

$$= 10.978$$

So, model parameter for x_2 are $\sigma_0^2 = 16.307$, $\sigma_1^2 = 10.978$, $\mu_0 = 26.125$, $\mu_1 = 23.083$

(b)

We have to calculate, $P(y=1 \mid x_1=1, x_2=25.2, x_3=0)$

So, from Naive Bayes inference,

$$P(y=1 \mid x_1=1, x_2=25.2, x_3=0) = P(y=1) \cdot P(x_1=1 \mid y=1) \cdot P(x_2=25.2 \mid y=1) \cdot P(x_3=0 \mid y=1)$$

Now,

$P(y=1) = 6/10$

$P(x_1=1 \mid y=1) = 2/6$ [from 'a']

$P(x_3=0 \mid y=1) = 4/6$

for, $$P(x_2=25.2 \mid y=1) = \frac{1}{\sqrt{2\pi}\,\sigma_1} \cdot e^{-\frac{1}{2\sigma_1^2} \cdot (25.2-\mu_1)^2}$$

$$= \frac{1}{\sqrt{2\pi \times 10.978}} \cdot e^{-\frac{1}{2 \times 10.978} \times (25.2-23.083)^2}$$

$$= 0.0982$$

So,

$$P(y=1 \mid x_1=1, x_2=25.2, x_3=0) = 6/10 \times \frac{2}{6} \times (0.0982) \times 4/6$$

$$= 0.013$$

(Ans.)

(c)

Given, $x_1 = 2$, $x_2 = 22.2$, $x_3 = 1$.

So,

$$P(y=0 \mid x_1=2, x_2=22.2, x_3=1) = P(x_1=2 \mid y=0) \cdot P(x_2=22.2 \mid y=0) \cdot P(x_3=1 \mid y=0) \cdot P(y=0)$$

$$P(y=1 \mid x_1=2, x_2=22.2, x_3=1) = P(x_1=2 \mid y=1) \cdot P(x_2=22.2 \mid y=1) \cdot P(x_3=1 \mid y=1) \cdot P(y=1)$$

Now,

$$P(x_2=22.2 \mid y=0) = \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{1}{2\sigma_0^2} \cdot (22.2-\mu_0)^2}$$

$$= \frac{1}{\sqrt{2\pi \times 16.307}} \times e^{-\frac{1}{2 \times 16.307} \times (22.2-26.125)^2}$$

$$= 0.0616$$

$$P(x_2=22.2 \mid y=1) = \frac{1}{\sqrt{2\pi}\,\sigma_1} \cdot e^{-\frac{1}{2\sigma_1^2} \cdot (22.2-\mu_1)^2}$$

$$= \frac{1}{\sqrt{2\pi \times 10.978}} \cdot e^{-\frac{1}{2 \times 10.978} \cdot (22.2-23.083)^2}$$

$$= 0.1162$$

from 'a'

$P(x_1=2 \mid y=0) = 1/4$, $P(x_3=1 \mid y=0) = 3/4$, $P(y=0) = 4/10$

$P(x_1=2 \mid y=1) = 3/6$, $P(x_3=1 \mid y=1) = 2/6$, $P(y=1) = 6/10$

Finally,

$$P(y=0 \mid x_1=2, x_2=22.2, x_3=1) = 1/4 \times (0.0616) \times 3/4 \times 4/10$$

$$= 0.00462$$

$$P(y=1 \mid x_1=2, x_2=22.2, x_3=1) = 3/6 \times (0.1162) \times 2/6 \times 6/10$$

$$= 0.01162$$

Since, $P(y=1 \mid x_1=2, x_2=22.2, x_3=1) > P(y=0 \mid x_1=2, x_2=22.2, x_3=1)$

the model inferes $y=1$. (Ans.)

(d)

For Laplace Smoothing, the model parameters are

$$\phi_{y_k} = \frac{1 + \sum_{i=1}^{n} 1\{y^{(i)}=1\}}{k + \sum_{i=1}^{n} 1\{y^{(i)}=1\}}$$

$$\phi_{x_j \mid y=0} = \frac{1 + \sum_{i=1}^{n} 1\{x_j^{(i)}=1 \wedge y=0\}}{k + \sum_{i=1}^{n} 1\{y=0\}}$$

$$\phi_{x_j \mid y=1} = \frac{1 + \sum_{i=1}^{n} 1\{x_j^{(i)}=1 \wedge y=1\}}{k + \sum_{i=1}^{n} 1\{y=1\}}$$

Here, in the formula K is the number of distinct values in x. So, probability tables after applying laplace smoothing,

for x_1

x_1	$y=0$	$y=1$
$x_1=0$	$\frac{3+1}{4+3} = \frac{4}{7}$	$\frac{1+1}{6+3} = \frac{2}{9}$
$x_1=1$	$\frac{0+1}{4+3} = \frac{1}{7}$	$\frac{2+1}{6+3} = \frac{3}{9}$
$x_1=2$	$\frac{1+1}{4+3} = \frac{2}{7}$	$\frac{3+1}{6+3} = \frac{4}{9}$

for x_3,

x_3	$y=0$	$y=1$
$x_3=0$	$\frac{1+1}{4+2} = \frac{2}{6}$	$\frac{4+1}{6+2} = \frac{5}{8}$
$x_3=1$	$\frac{3+1}{4+2} = \frac{4}{6}$	$\frac{2+1}{6+2} = \frac{3}{8}$

Since, x_3 follows a normal distribution, we don't have to apply laplace smoothing for x_2. (Ans.)

Answer to the que. no-2

Given, the general equation for a simple perceptron.

$$w_0 + w_1 x_1 + w_2 x_2 = 0$$

Now, assumming that, there exists a solution to the XOR problem solved by a simple perceptron. Now, XOR truth table,

X_1	X_2	Output
0	0	0
0	1	1
1	0	1
1	1	0

Now, from the step function,

$$S(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ 1 & \text{if } z > 1 \end{cases}$$

So, for $(0,0)$, $\quad w_0 \leq 0$ — (i)

for $(0,1)$, $\quad w_0 + w_2 > 0$ — (ii)

for $(1,0)$, $\quad w_0 + w_1 > 0$ — (iii)

for $(1,1)$, $\quad w_0 + w_1 + w_2 \leq 0$ — (iv)

Now, adding (i) and (iv)

$$w_0 + w_0 + w_1 + w_2 \leq 0 + 0$$

$$\Rightarrow 2w_0 + w_1 + w_2 \leq 0$$

and, adding (ii) and (iii)

$$w_0 + w_1 + w_0 + w_2 > 0 + 0$$

$$\Rightarrow 2w_0 + w_1 + w_2 > 0$$

So, we have both $2w_0 + w_1 + w_2 \leq 0$ and $2w_0 + w_1 + w_2 > 0$ from the same XOR table which is a contradiction.

Therefore, our assumption that a simple perceptron can solve the XOR problem is wrong, and so, it cannot solve the XOR problem.

(Ans.)

Answer to the question no-3

(a)

Given,

$$a_i = f_i\left(\sum w_j^{(i)} a_j\right)$$

input, $a_0 = x$, $f_2(x) = ReLU(x)$, $f_1(x) = f_3(x) = f_4(x) = \sigma(x)$

$w_0^1 = 1$, $w_0^3 = 2$, $w_1^2 = 1.5$, $w_3^2 = 1$, $w_1^4 = 2.5$, $w_2^4 = 1.2$, $w_3^4 = 0.8$

Now,

$$a_0 = x.$$

$$z_1 = w_0^1 a_0 = 1 \times x = x$$

$$z_3 = w_0^3 a_0 = 2 \times x = 2x$$

$$z_2 = w_1^2 a_1 + w_3^2 a_3 = 1.5 a_1 + 1 a_3 = 1.5 a_1 + a_3$$

$$z_4 = w_1^4 a_1 + w_2^4 a_2 + w_3^4 a_3 = 2.5 a_1 + 1.2 a_2 + 0.8 a_3$$

So,

$\therefore$

$$a_0 = x$$

$$a_1 = \sigma(x)$$

$$a_2 = ReLU(1.5 a_1 + a_3)$$

$$a_3 = \sigma(2x)$$

$$a_4 = \sigma(2.5 a_1 + 1.2 a_2 + 0.8 a_3)$$

Therefore, output of the neural network,

$$\hat{y} = a_4 = \sigma(2.5 a_1 + 1.2 a_2 + 0.8 a_3)$$

(Ans)

(b)

The squared loss function,

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{n} (\hat{y} - y)^2$$

Now,

$$\frac{\partial J(\theta)}{\partial w_3^4} = \frac{\partial J(\theta)}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_4} \cdot \frac{\partial z_4}{\partial w_3^4}$$

So,

$$\frac{\partial J(\theta)}{\partial \hat{y}} = (\hat{y} - y)$$

$$\frac{\partial \hat{y}}{\partial z_4} = \sigma(z_4)\,(1 - \sigma(z_4))$$

$$\frac{\partial z_4}{\partial w_3^4} = a_3$$

So,

$$\frac{\partial J(\theta)}{\partial w_3^4} = (\hat{y} - y) \cdot \sigma(z_4)\,(1 - \sigma(z_4)) \cdot a_3$$

Again,

$$\frac{\partial J(\theta)}{\partial w_3^2} = \frac{\partial J(\theta)}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^4} \cdot \frac{\partial z^4}{\partial a^2} \cdot \frac{\partial a^2}{\partial z^2} \cdot \frac{\partial z^2}{\partial w_3^2}$$

$$= (\hat{y} - y) \cdot \sigma(z^4)\,(1 - \sigma(z_4)) \cdot w_2^4 \cdot ReLU'(z_2) \cdot a_3$$

and for,

$$\frac{\partial J(\theta)}{\partial w_0^1} = \frac{\partial J(\theta)}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^4} \cdot \frac{\partial z^4}{\partial a_1} \cdot \frac{\partial a_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_0^1}$$

$$= (\hat{y}-y) \cdot \sigma(z_4)(1-\sigma(z_4)) \cdot w_1^4 \cdot \sigma(z_1)(1-\sigma(z_1)) \cdot a_0$$

$$= (\hat{y}-y) \cdot \sigma(z_4)(1-\sigma(z_4)) \cdot w_1^4 \cdot \sigma(z_1)(1-\sigma(z_1)) \cdot x$$

(Ans.)

(c)

We know.

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$\text{ReLU}(z) = \max(0, z)$$

Now, for $x = 0.7$.

$$a_0 = x = 0.7$$

$$a_1 = \sigma(0.7) = \frac{1}{1+e^{-0.7}} = 0.668$$

$$a_3 = \sigma(2 \times 0.7) = \frac{1}{1+e^{-2\times 0.7}} = 0.802$$

$$a_2 = \text{ReLU}((1.5 \times 0.668) + 0.802)$$

$$= \text{ReLU}(1.804)$$

$$= \max(0, 1.804)$$

$$= 1.804$$

Now, $a_4 = \sigma((2.5 \times 0.668) + (1.2 \times 1.804) + (0.8 \times 0.802))$

$$= \sigma(4.4764)$$

$$= \frac{1}{1+e^{-4.4764}}$$

$= 0.989$

So, for $x = 0.7$, output of neural network is 0.989.

Now, For $x = 1.5$.

$a_0 = x = 1.5$

$a_1 = \sigma(1.5) = \frac{1}{1+e^{-1.5}} = 0.818$

$a_3 = \sigma(2 \times 1.5) = \frac{1}{1+e^{-(2\times1.5)}} = 0.953$

$a_2 = ReLU((1.5 \times 0.818) + (0.953))$

$= ReLU(2.18)$

$= max(0, 2.18)$

$= 2.18$

So, $a_4 = \sigma((2.5 \times 0.818) + (1.2 \times 2.18) + (0.8 \times 0.953))$

$= \sigma(5.4234)$

$= \frac{1}{1+e^{-5.4234}}$

$= 0.996$

$\therefore$ for $x = 1.5$, output of neural network is 0.996.

(Ans.)

(d)

for Epoch 1:

training set $\{(0, 0.98), (-0.5, 0.89), (1.2, 0.52)\}$

for $(0, 0.98)$,

$$a_0 = 0$$

$$a_1 = \sigma(0) = 0.5$$

$$a_3 = \sigma(2\times 0) = 0.5$$

$$a_2 = \text{ReLU}((1.5\times 0.5) + (0.5)) = \text{ReLU}(1.25)$$

$$= \max(1.25, 0)$$

$$= 1.25$$

$$\hat{y} = a_4 = \sigma((2.5\times 0.5) + (1.2\times 1.25) + (0.8\times 0.5))$$

$$= 0.959$$

so,

$$w_3^4 = w_3^4 - \eta \frac{\partial J(\theta)}{\partial w_3^4}$$

$$= 0.8 - 0.1((0.959 - 0.98)\times 0.959\times(1-0.959)\times 0.5)$$

$$= 0.80004$$

$$w_3^2 = w_3^2 - \eta \frac{J(\theta)}{\partial w_3^2}$$

$$= 1 - 0.1\{(0.959-0.98)\times 0.959\times(1-0.959)\times 1.2\times 1\times 0.5\}$$

$$= 1.00005$$

$$w_0^1 = w_0^1 - \eta \frac{J(\theta)}{\partial w_0^1}$$

$$= 1 - 0.1\{(0.959-0.98)\times 0.959\times(1-0.959)\times 2.5\times 0.5\times(1-0.5)\times 0)$$

$$= 1$$

for $(-0.5, 0.89)$

$a_0 = -0.5$

$a_1 = \sigma(-0.5) = 0.3775$

$a_3 = \sigma(2 \times -0.5) = 0.269$

$a_2 = ReLU((1.5 \times 0.3775) + (0.269 \times 1.00005)) = 0.83525$

$a_4 = \sigma((2.5 \times 0.3775) + (1.2 \times 0.83525) + (0.8 \times 0.269))$

$\Rightarrow \hat{y} = 0.897$

$\therefore\ w_3^4 = 0.80004 - 0.1((0.897 - 0.89) \times 0.897 \times (1 - 0.897) \times 0.269)$

$= 0.80002$

$w_3^2 = 1.00005 - 0.1\{(0.897 - 0.89) \times 0.897 \times (1 - 0.897) \times 1.2 \times 1 \times 0.269\}$

$= 1.00003$

$w_0^1 = 1 - 0.1\{(0.897 - 0.89) \times 0.897 \times (1 - 0.897) \times 2.5 \times 0.3775 \times (1 - 0.3775) \times (-1.5)\}$

$= 1.000019$

for $(1.2, 0.52)$,

$a_0 = 1.2$

$a_1 = \sigma(1.2 \times 1.000019) = 0.769$

$a_3 = \sigma(1.2 \times 2) = 0.917$

$a_2 = ReLU((1.5 \times 0.769) + (0.917 \times 1.00003)) = 2.0705$

$a_4 = \sigma((2.5 \times 0.769) + (1.2 \times 2.0705) + (0.8 \times 0.917))$

$\Rightarrow \hat{y} = 0.9942$

$\therefore\ w_3^4 = 0.80002 - 0.1\{(0.9942-0.52)\times 0.9942\times(1-0.9942)\times 0.917\}$

$= 0.7998$

$w_3^2 = 1.00003 - 0.1\{(0.9942-0.52)\times 0.9942\times(1-0.9942)\times 1.2\times 1\times 0.917\}$

$= 0.99702$

$w_0^1 = 1.000019 - 0.1\{(0.9942-0.52)\times 0.9942\times(1-0.9942)\times 2.5\times 0.769\times(1-0.769)$
$\times 1.2\}$

$= 0.9999$

Now, <u>for Epoch 2</u> :

for $(0, 0.98)$,

$a_0 = 0$

$a_1 = \sigma(0.9999\times 0) = 0.5$

$a_3 = \sigma(2\times 0) = 0.5$

$a_2 = \text{ReLU}((1.5\times 0.5) + (0.99702\times 0.5))$

$= 1.24851$

$a_4 = \sigma((2.5\times 0.5) + (1.2\times 1.24851) + (0.7998\times 0.5))$

$\Rightarrow \hat{y} = 0.95883$

so, $w_3^4 = 0.7998 - 0.1\{(0.95883-098)\times 0.9588\times(1-0.9588)\times 0.5\}$

$= 0.79984$

$$w_3^2 = 0.99702 - 0.1\{(0.95883-0.98)\times 0.95883\times(1-0.95883)\times 1.2\times 1\times 0.5\}$$

$$= 0.99707$$

$$w_0^1 = 0.9999 - 0.1\{(0.95883-0.98)\times 0.95883\times(1-0.95883)\times 2.5\times 0.5\times(1-0.5)\cdot 0\}$$

$$= 0.9999$$

for $(-0.5, 0.89)$.

$$a_0 = -0.5$$

$$a_1 = \sigma(0.9999\times -0.5) = 0.3775$$

$$a_3 = \sigma(2\times -0.5) = 0.269$$

$$a_2 = \mathrm{ReLU}((1.5\times 0.3775) + (0.99707\times 0.269)) = 0.83446$$

$$a_4 = \sigma((2.5\times 0.3775) + (1.2\times 0.83446) + (0.79984\times 0.269)).$$

$$\Rightarrow \hat{y} = 0.8966$$

So, $w_3^4 = 0.79984 - 0.1\{(0.8966-0.89)\times 0.8966\times(1-0.8966)\times 0.269\}$

$$= 0.79982$$

$$w_3^2 = 0.99707 - 0.1\{(0.8966-0.89)\times 0.8966\times(1-0.8966)\times 1.2\times 0.269\}$$

$$= 0.99705$$

$$w_0^1 = 0.9999 - 0.1\{(0.8966-0.89)\times 0.8966\times(1-0.8966)\times 2.5\times 0.3775 \times(1-0.3775)\times(-0.3)\}$$

$$= 0.9999$$

for $(1.2, 0.52)$,

$$a_0 = 1.2$$

$$a_1 = \sigma(0.9999 \times 1.2) = 0.7685$$

$$a_3 = \sigma(2 \times 1.2) = 0.917$$

$$a_2 = \text{ReLU}((1.5 \times 0.7685) + (0.99705 \times 0.917))$$

$$= 2.067$$

$$a_4 = \sigma((2.5 \times 0.7685) + (1.2 \times 2.067) + (0.79982 \times 0.917))$$

$$\Rightarrow \hat{y} = 0.99414$$

so, $w_3^4 = 0.79982 - 0.1\{(0.99414 - 0.52) \times 0.99414 \times (1 - 0.99414) \times 0.917\}$

$$= 0.79956$$

$$w_3^2 = 0.99705 - 0.1\{(0.99414 - 0.52) \times (0.99414) \times (1 - 0.99414) \times 1.2 \times 0.917\}$$

$$= 0.99674$$

$$w_0^1 = 0.9999 - 0.1\{(0.99414 - 0.52) \times 0.99414 \times (1 - 0.99414) \times 2.5 \times 0.7685 \times (1 - 0.7685) \times 1.2\}$$

$$= 0.9984$$

Finally, after 2 epoch,

$$w_3^4 = 0.79956$$

$$w_3^2 = 0.99674$$

$$w_0^1 = 0.9984$$

(Ans:)

Ans. to the que. no-04

(a)

Total people = 50+40+30 = 120

total votes for A = 25+30+5 = 60

total votes for B = 15+5+10 = 30

total votes for C = 10+5+15 = 30

So,

$$P(A) = \frac{60}{120} = 0.5$$

$$P(B) = \frac{30}{120} = 0.25$$

$$P(C) = \frac{30}{120} = 0.25$$

$$\therefore H(V) = -\left[P(V=A)\log_2 P(V=A) + P(V=B)\log_2 P(V=B) + P(V=C).\log_2 P(V=C)\right]$$

$$= -(0.5\times\log_2 0.5 + 0.25\times\log_2 0.25 + 0.25\log_2 0.25)$$

$$= 1.5$$

(Ans.)

(b)

$$H(V|Age) = P(Age=18\text{-}30) + H(V|Age=18\text{-}30) + P(Age=31\text{-}50) \cdot H(V|Age\ 31\text{-}50) +$$
$$P(Age=50+) \cdot H(V|Age=50+)$$

Now,

$$H(V|Age=18\text{-}30) = -\{P(V=A|Age=18\text{-}30)\log_2 P(V=A|Age=18\text{-}30) +$$
$$P(V=B|Age=18\text{-}30)\log_2 P(V=B|Age=18\text{-}30) +$$
$$P(V=C|Age=18\text{-}30)\log_2 P(V=B|Age=18\text{-}30)$$

$$= -\{25/50 \log_2 25/50 + 15/50 \log_2 15/50 + 10/50 \log_2 10/50\}$$

$$= 1.485$$

$$H(V|Age=31=50) = -\{P(V=A|Age=31\text{-}50) \cdot \log_2 P(V=A|Age=31\text{-}50) +$$
$$P(V=B|Age=31\text{-}50)\log_2 P(V=B|Age=31\text{-}50) +$$
$$P(V=C|Age=31\text{-}50)\log_2 P(V=C|Age=31\text{-}50)$$

$$= -\{30/40 \log_2 30/40 + 5/40 \log_2 5/40 + 5/40 \log_2 5/40\}$$

$$= 1.0613$$

$$H(V|Age=50+) = -\{P(V=A|Age=50+) \cdot \log_2 P(V=A|Age=50+) +$$
$$P(V=B|Age=50+) \cdot \log_2 P(V=B|Age=50+) +$$
$$P(V=C|Age=50+) \cdot \log_2 P(V=C|Age=50+)\}$$

$$= -\{5/30 \log_2 5/30 + 10/30 \log_2 10/30 + 15/30 \log_2 15/30\}$$

$$= 1.459$$

So, $H(V|Age) = \left(\frac{50}{120} \times 1.485\right) + \left(\frac{40}{120} \times 1.0613\right) + \left(\frac{30}{120} \times 1.459\right)$

$= 1.3373$ (Ans.)

(c)

$I(V, Age) = H(V) - H(V|Age)$

$= 1.5 - 1.3373$

$= 0.1627$ (Ans.)

(d)

$I(V, i) = H(V) - H(V|i)$

Now,

$H(V) = -\{P(V=A).\log_2 P(V=A) + P(V=B)\log_2 P(V=B) + P(V=C).\log_2 P(V=C)\}$

$= -\{ 60/120 \log_2 60/120 + 45/120 . \log_2 45/120 + 15/120 \log_2 15/120 \}$

$= 1.4056$

Now,

$H(V|i) = P(i=l)\, H(V|i=l) + P(i=m)\, H(V|i=m) + P(i=h) . H(V|i=h)$

So, $H(V|i=l) = -\{P(V=A|i=l).\log_2 P(V=A|i=l) + P(V=B|i=l)\log_2 P(V=B|i=l) + P(V=C|i=l) . \log_2 P(V=C|i=l)\}$

$$= -\left(\frac{10}{60}\log_2\frac{10}{60} + \frac{40}{60}\log_2\frac{40}{60} + \frac{10}{60}\log_2\frac{10}{60}\right)$$

$$= 1.2516$$

$$H(v|i=m) = -\{P(v=A|i=m)\cdot\log_2 P(v=A|i=m) + P(v=B|i=m)\cdot\log_2 P(v=B|i=m) +$$
$$P(v=C|i=m)\cdot\log_2 P(v=C|i=m)\}$$
$$= -\left\{\frac{20}{30}\log_2\frac{20}{30} + \frac{5}{30}\log_2\frac{5}{30} + \frac{5}{30}\log_2\frac{5}{30}\right\}$$

$$= 1.2516$$

$$H(v|i=h) = -\{P(v=A|i=h)\cdot\log_2 P(v=A|i=h) + P(v=B|i=h)\cdot\log_2 (v=B|i=h) +$$
$$P(v=C|i=h)\cdot\log_2 P(v=C|i=h)\}$$
$$= -\left\{\frac{30}{30}\log_2\frac{30}{30} + 0 + 0\right\}$$

$$= 0$$

So, $H(v|i) = \left(\frac{60}{120} \times 1.2516\right) + \left(\frac{30}{120} \times 1.2516\right) + \left(\frac{30}{120} \times 0\right)$

$$= 0.9387$$

finally, $I(v|i) = H(v) - H(v|i)$

$$= 1.4056 - 0.9387$$

$$= 0.4669$$

(Ans.)

(e)

We got,

$$I(v, age) = 0.1627$$

$$I(v, i) = 0.4669$$

Since, $I(v, i) > I(v, Age)$

"Interest in politics" should be the root node of the decision tree.

(Ans.)