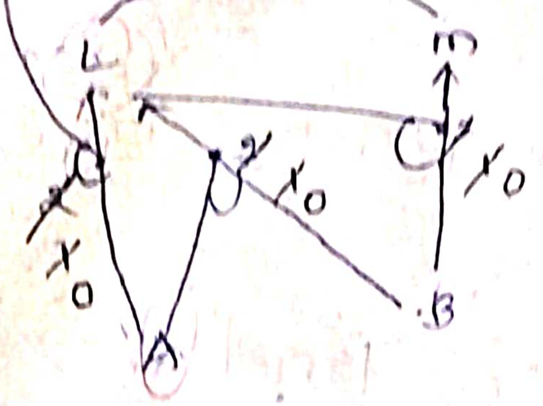
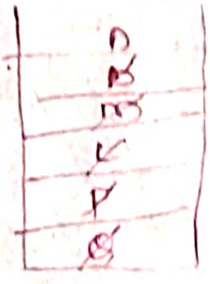


AB

Backward chaining:

- Start from goal and work backward
- It will take the goal node first and satisfy the result



Stack

BALMPQ

Already know as true

- first Order logic

Regression:-

Ex:-

chocolate % of coco	price
10	35
20	55
30	40
35	100
40	60
50	90
60	110
70	130

$\Sigma x = 315$

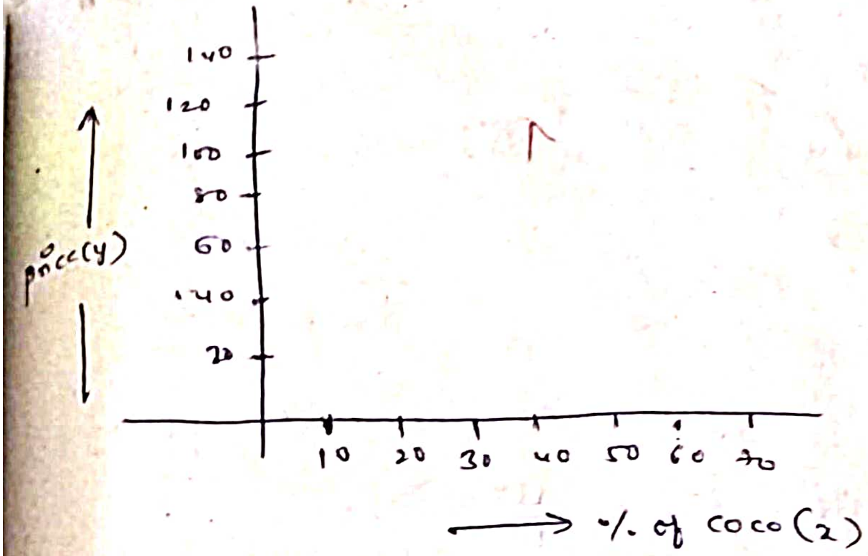
$\Sigma x^2 = 15225$

$\Sigma y = 620$

$\Sigma y^2 = 56550$

$\Sigma xy = 28750$

Which Brand is Overpriced?
what should be the fair price?



$\hat{\alpha} = \bar{y} - \beta \bar{x} = 16.98$

$\hat{\beta} = \dfrac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = 1.54$

$y = 16.98 + 1.53x$

$y = 16.98 + 1.53 \times 35$

$= 70.52$

Classification:-

(x, y)

^ categorical

(class)

- Decision Tree
- Naive Bayes
- Bayesian Network

Evaluation of classification models:

- Confusion matrix
- Evaluation matrix
  - Accuracy
  - Error Rate
  - Precision
  - Recall
  - F1- Score

Confusion matrix - Binary output Actual

		0	1
Predicted	0	TN	FN
	1	FP	TP

Example:-

Eg:-

(165)

		Actual	
		-ve	+ve
Predected	-ve	100	5
	+ve	10	50

Accuracy: $\frac{TP+TN}{TP+FP+TN+FN} = \frac{150}{165} = 0.9$

Precision: $\frac{TP}{TP+FP} = \frac{50}{60} = 0.83$ // no. of correct prediction out of total +ve predi

Recall: $\frac{TP}{TP+FN} = \frac{50}{55} = 0.90$ // out of all positive cases how many you are able to detect

$$F_1 = 2 * \frac{P*R}{P+R}$$

$$= \frac{2 * 0.83 * 0.9}{0.83+0.9}$$

$$= 0.86.$$

Eg2:

Actual	Doctor1 prediction	Doctor2 prediction.
0	0	0
1	1	0
0	0	0
0	0	0
0	1	0
0	1	0
0	0	0
0	0	0
0	0	0
	0	0

Accuracy

Doctor1 $= \frac{7}{10} = 0.7$

Doctor 2 $= \frac{8}{10} = 0.8$

Actual

	-ve	+ve
Doc1 -ve	6	1
+ve	2	1

Doctor1	Doctor 2
$P = 0.3$	$P = 0$
$R = 0.5$	$R = 0$
$F_1 = 0.375$	$F_1 =$

Bayesian Classification:- (It compares two probability)

↳ Naive Bayes classifier

(X, Y)

features → X; class label, categorical variable → Y

$X = \{x_1, x_2, \cdots x_n)$ $Y = \{c_1, c_2, c_3\}$

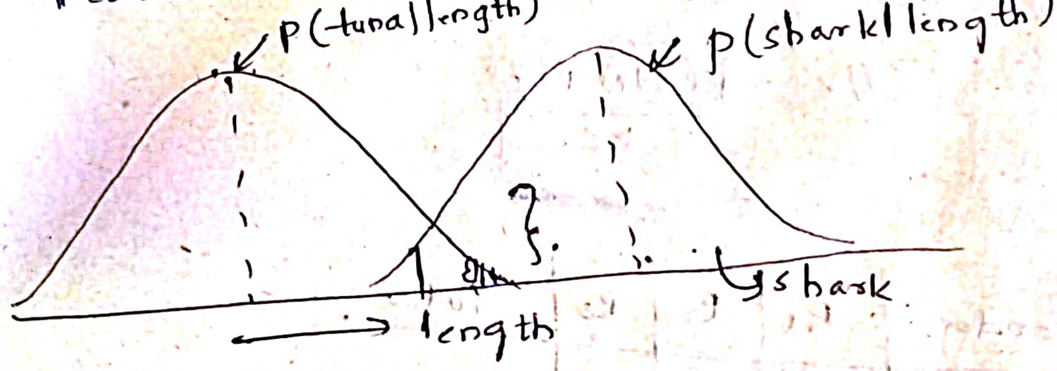
$$x \longrightarrow \boxed{f(x)} \longrightarrow y$$

length shark / tuna?

eg: Classify fish

Two types: shark, Tuna

Feature: length

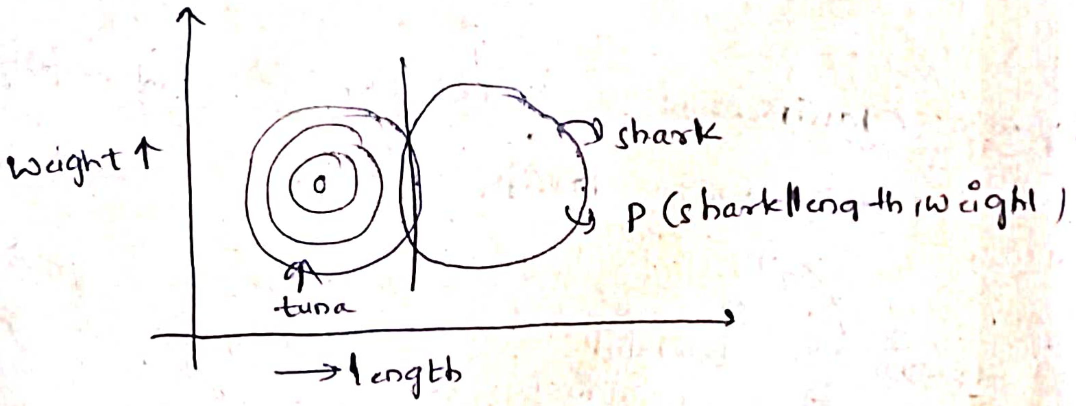


P (shark|length) > P (tuna| length)
→ shark

P (tuna|length) > P (shark| length)
→ tuna

Features:- length, weight

Tuna: small length, light weight
Shark: long length, heavy weight



$$P(Y/x) = \frac{P(X \land Y) \rightarrow \text{joint}}{P(X) \rightarrow \text{marginal}}$$

$$\rightarrow P(X \land y) = P(Y|x)\,P(X)$$

$$P(x|Y) = \frac{P(X \land Y)}{P(Y)}$$

$$\rightarrow P(X \land Y) = P(x|Y)\,P(Y)$$

$$\boxed{P(Y/x) = \frac{P(x|Y)\,P(Y)}{P(x)}}$$

↑ x ↓

Gender	R₁	R₂	R₃	Total
m	30	80	90	200
F	20	40	40	100
Total	50	120	130	300

Joint probability:-

$$P(X = m \land y = R_1) = \frac{30}{300}$$

$$P(X=F \wedge Y=R_2) = \frac{40}{300}$$

Marginals:-

$$P(X=m) = \frac{200}{300}$$

$$= P(X=m \wedge Y=R_1) + P(X=m \wedge Y=R_2) + P(X=m \wedge Y=R_3)$$

$$= \frac{30}{300} + \frac{80}{300} + \frac{90}{300} = \frac{200}{300}$$

Conditionals:- $P(Y=R_1 | X=m) = \frac{30}{200}$

$$\frac{P(Y=R_3 | X=F)}{Y(Y=R_3)} = \frac{40}{130}.$$

Naive Bayes Classifier:-

$P(X \wedge Y)$ - joint

$P(X|Y)$ - conditional

$P(X)$ - marginal

$Y = \{0, 1\}$ $\quad X = \{x_1, x_2\}$

$$P(Y=0 | X=\{x_1, x_2\}) = \frac{P(Y=0 \wedge X)}{P(X)}$$

$$P(X \wedge Y) = P(X|Y) P(Y) = P(Y|X) P(X)$$

$$P(X/Y) = \frac{P(X \wedge Y)}{P(Y)}$$

$$= \frac{P(Y/X) P(X)}{P(Y)}$$

$$P(Y/X) = \frac{P(X \wedge Y)}{P(X)}$$

$$= \frac{P(X|Y) P(Y)}{P(X)}$$

Naive Bayes Assumption (Conditional Independence):

$$P(X = \{x_1, x_2\} | Y=0\})$$

$$= P(X=x_1 | Y=0) * P(X=x_2 | Y=0)$$

↳ Assume x_1 & x_2 are Independent given y

Independence:

$$P(x_1 . x_2)$$

$$= P(x_1) P(x_2)$$

Play Golf :-

Day	Outlook	Temp	Humidity	Wind	Play
1	Sunny	Hot	High	Weak	No
2	S	H	H	Strong	N
3	Overcast	H	H	W	Yes
4	Rainy	mild	H	W	Y
5	R	Cold	Normal	W	Y
6	R	C	N	S	N
7	O	C	N	S	Y
8	S	m	H	W	N
9	S	C	N	W	Y
10	R	m	N	W	Y
11	S	m	N	S	Y
12	O	m	H	S	Y
13	O	H	N	W	Y
14	R	m	H	S	N

Day = { Rainy, cold | High, strong }

play = yes/No ?

P(Yes | R, C, H, S) = ?

P(No | R, C, H, S) = ?

$$P(Y | R, C, H, S) = \frac{P(R, C, H, S | Y) P(Y)}{P(R, C, H, S)}$$

$$= \frac{P(R, C, H, S | Y) P(Y)}{P(R, C, H, S | Y) P(Y) + P(R, C, H, S | N) P(N)}$$

Conditional tables:

Outlook. P(outlook|play)

	Yes	No	P(·|yes)	P(·|No)
S	2	3	2/9	3/5
O	4	0	4/9	0
R	3	2	3/9	2/6
Total	9	5	1	1

Temp: P(Temp|play)

	Yes	No	P(·|yes)	P(·|No)
Hot	2	2	2/9	2/5
mild	4	2	4/9	2/5
Cold	3	1	3/9	1/5
Total	9	5	1	1

Humidity P(Humidity|play)

	Yes	No	P(·|yes)	P(·|No)
High	3	4	3/9	4/9
Normal	6	1	6/9	1/5
Total	9	5	1	1

Wind. P(wind|play)

	Yes	No	P(·|yes)	P(·|No)
Strong	6	2	6/9	2/5
weak	3	3	3/9	3/5
Total	9	5	1	1

$P(Yes) = 9/14$

$P(No) = 5/14$

$$\Rightarrow \frac{P(R/y) \cdot P(C/y) \cdot P(H/y) P(S/y) \cdot P(y)}{P(R/y) P(C/y) \cdot P(H/y) P(S/y) \cdot P(y) + P(R/N)\, P(C/N) \cdot P(H/N) \cdot P(S/N) \cdot P(N)}$$

$$= \frac{(3/9)(3/9)(3/9)(6/9)(9/14)}{(3/9)(3/9)(3/9)(6/9)(9/14) + (2/6)(1/5)(4/9)(2/5)(5/14)}$$

$$\simeq \frac{0.0158}{0.0201}$$

$$= 0.7860$$

Decision Tree:-

Entropy
↳ for a distribution

freq: $[9+, 5-]$

prob: $[9/14, 5/14]$

Entropy: $-\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2 5/14$

Purity

$[14+ ; 0-]$ pure → Entropy: 0 Low ↓ Certainity ↑

$[10+, 4-]$

$[7+, 7-]$ unpure → Entropy: 1 high ↑ Certainity ↓

$$-\frac{7}{14}\log_2\frac{7}{14} - \frac{7}{14}\log_2\frac{7}{14}$$

$$= -1/2\log_2\frac{1}{2} - 1/2\log_2\frac{1}{2}$$

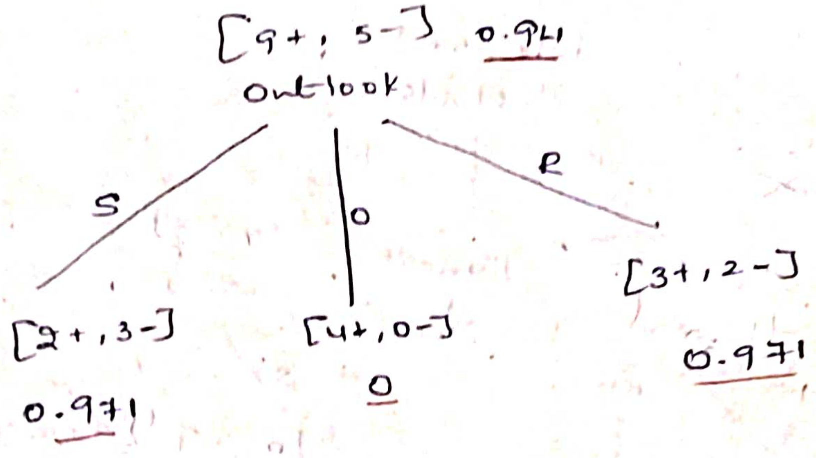
$$= -\log_2 1/2$$

$$= -(\log_2 1 - \log_2 2)$$

$$= -\log_2 1 + \log_2 2$$

$$= 0 + 1$$

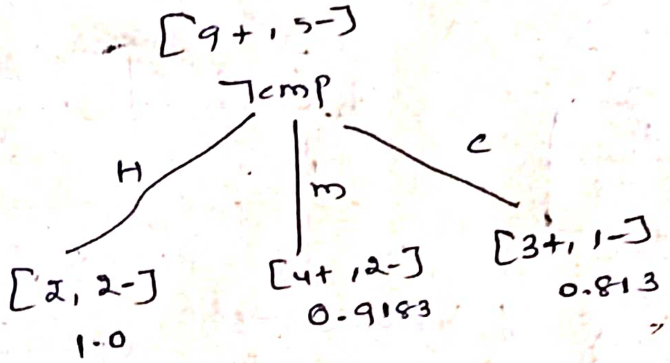
$$= 1$$



Average Entropy: $\frac{5}{14} \times 0.971 + \frac{4}{14} \times 0 + \frac{5}{14} \times 0.971 = 0.693$

Gain: 0.94 - 0.693

= 0.246

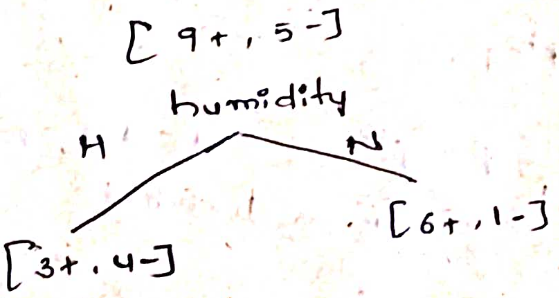


Average Entropy: $\frac{9}{14} \times 1.0 + \frac{6}{14} \times 0.9183 + \frac{9}{14} \times 0.8113$

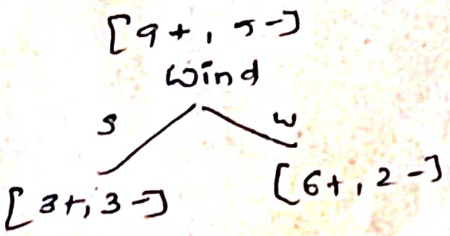
= 0.910

Gain: 0.94 - 0.910

= 0.03

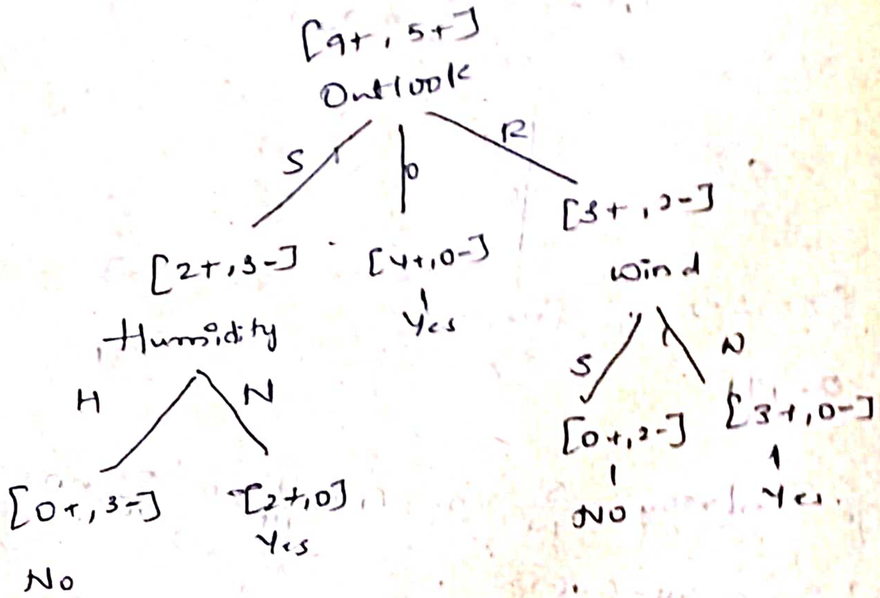


Gain: 0.1516



Gain: 0.0478

Multineuron will handle non-linear function



Neural Networks -

Data

x_1	x_2	y
0.1	0.3	0.03

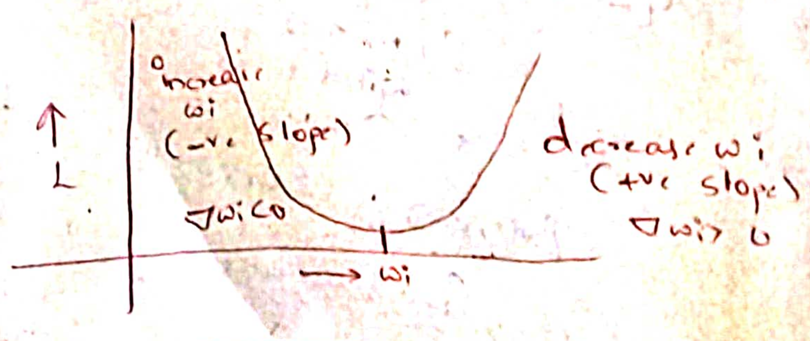
Forward Propagation: $\begin{cases} Z = b + w_1x_1 + w_2x_2 \\ \hat{y} = \sigma(z) = \frac{1}{1+e^{-z}} \end{cases}$ Parameters: (b, w1, w

Loss/Errors

1. Mean Square error
2. Cross Entropy

$$L = \frac{1}{2}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \rightarrow$$ calculate for all Examples in a batch.

Backpropagation (Using Gradient Descent):



Gradient rule:

$w_{i+1} = w_i - \eta . \nabla w_i$

↳ learning rate.

Gradient calculation:

$$\nabla w_1 = \frac{\partial z}{\partial w_1} = \underbrace{\frac{dz}{\partial \hat{y}}}\cdot\frac{\partial \hat{y}}{\partial z}\cdot\frac{\partial z}{\partial w_1}$$

$L = f(\hat{y})$

$\hat{y} = f(z)$

$z = f(w_1, w_2, b)$

$$\nabla w_2 = \frac{\partial z}{\partial w_2} = \underbrace{\frac{\partial z}{\partial \hat{y}}}\cdot\frac{\partial \hat{y}}{\partial z}\cdot\frac{\partial z}{\partial w_2}$$

$$\nabla b = \frac{\partial z}{\partial b} = \underbrace{\frac{\partial z}{\partial \hat{y}}}\cdot\frac{\partial \hat{y}}{\partial z}\cdot\frac{\partial z}{\partial b}$$

$$\frac{\partial z}{\partial w_1} = \frac{\partial}{\partial w_1}(w_1 x_1 + w_2 x_2 + b) = x_1$$

$$\frac{\partial z}{\partial w_2} = x_2$$

$$\frac{dz}{b} = 1$$

$$\frac{\partial z}{\partial \hat{y}} = \frac{1}{2} * 2(y-\hat{y}) * (-1)$$

$$= -(y-\hat{y})$$

$$\frac{\partial \hat{y}}{\partial z} = \frac{\partial}{\partial z}\left(\frac{1}{1+e^{-z}}\right) = \frac{-1}{(1+e^{-z})^2}\cdot e^{-z}(-1) = \frac{1}{(1+e^{-z})}\cdot\frac{e^{-z}}{(1+e^{-z})}$$

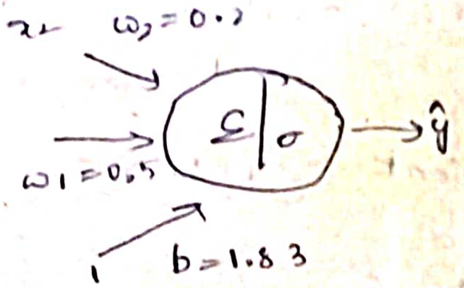
$$= \frac{1}{1+e^{-z}}\cdot\frac{1-1+e^{-z}}{(1+e^{-z})} = \frac{1}{1+e^{-z}}\left(1-\frac{1}{1+e^{-z}}\right)$$

$$= \sigma(z)(1-\sigma(z)) = \hat{y}(1-\hat{y}).$$

$$\boxed{\begin{aligned}\nabla w_1 &= -(y-\hat{y})\hat{y}(1-\hat{y}).x_1\\ \nabla w_2 &= -(y-\hat{y}).\hat{y}(1-\hat{y}).x_2\\ \nabla b &= -(y-\hat{y}).\hat{y}(1-\hat{y}).1\end{aligned}}$$

Data:-

x_1	x_2	y
0.1	0.3	0.03



$z = w_1 x_1 + w_2 x_2 + b$

$= 0.5 \times 0.1 + 0.2 \times 0.3 + 1.83 = 1.94$

$\hat{y} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-1.94}} = 0.874352143$

Loss calculation:- $Z = \frac{1}{2} (0.03 - 0.874352143)^2$

$= 0.356465271$

$\frac{dZ}{d\hat{y}} = -(y-\hat{y}) = -(0.03 - 0.874352143)$

$= 0.844352143$

$\frac{\partial \hat{y}}{\partial z} = \hat{y}(1-\hat{y}) = 0.874352143 (1 - 0.874352143)$

$= 0.013803732$

$\frac{dz}{\partial w_1} = x_1 = 0.1$

$\frac{\partial z}{\partial w_2} = x_2 = 0.3$

$\frac{\partial z}{\partial b} = 1$

$\nabla w_1 = \frac{\partial Z}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial w} = 0.84 \times \text{--} * 0.01 * 0.1$

$= 0.001165$

$\nabla w_2 = 0.00349656$

$b = 0.01165521$

New parameter Values:-

$w_1 = 0.5 - 1 * 0.001165$

$= 0.498835$

$w_2 = 0.2 - 1 * 0.00349652$

$= 0.19965$

$b = 1.83 - 1 * 0.01165521$

$= 1.818344789$

New $\hat{y}$:-

$z = w_1x_1 + w_2x_2 + b$

$= 0.498835(0.1) + 0.19965(0.3) + 1.818344789$

$= 1.928123289$

$$\hat{y} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-1.9281}} = \frac{1}{1.1454} = 0.873042.$$