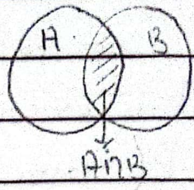
Date

Artificial Intelligence - Probability;

⇒ Probability = $\frac{\text{favourable}}{\text{Total}}$; Prob(I win in race with 3 others) = $\frac{1}{4}$

favourable (Me) → 1; 4 → Total People



⇒ Conditional probability :- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ → Joint; P(B) → Marginal

↳ Find Probability of A given that B already occured!!

So; let see:-

⇒ ~~Bayes~~ Dependent variable

$P(A \cap B) = P(A) \cdot P(B|A)$

Question:- Bag contains 5 red balls and 4 green balls.

Event (A) ⇒ {first ball is red}

Event (B) ⇒ {Second ball is red}.

$P(A) = 5/9$

~~P(B) = 5/9~~ Now; if dependency that ~~P(A∩B) =~~ if first ball already taken out and it is red ; Now; find second ball probability.

$P(B|A) = 4/8 = 1/2$

↳ Because 4 red ball left; and total 8 left.

Now;

$P(A \cap B) = 5/9 \times 1/2 = 0.2777.$

⇒ Means we don't replace our red ball hence Dependency occured.

Independent variable;

$P(A \cap B) = P(A) \cdot P(B)$

$P(A) = 5/9$

$P(B) = 5/9$

$P(A \cap B) = P(A) \cdot P(B)$.

$= 5/9 \cdot 5/9$

$P(A \cap B) = 0.3086.$

⇒ Means we replace red ball; means put back red ball again in the bag hence; No dependency occured!

⇒ Baye's Theorem:- $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$

Likelihood → P(B|A); Prior → P(A); Marginal → P(B); Posterior → P(A|B)

→ Single evidence case {No comparision}.

↳ Find probability of event "P(A|B)" when the given hypothesis is true!!

Multiple Evidence case ⇒ $P(A|B_1, B_2, \ldots B_n) = \frac{P(B_1|A)P(B_2|A) \ldots P(B_n|A)P(A)}{P(B_1) \cdot P(B_2) \ldots P(B_n)}$

{comparision, {Yes, No}.}

But evidence are Independ...

... Joint Probability Distribution:-

i) Move bottom to top.

ii) What is known keep on left side.

Date

Bayesian Belief Network- Example01;

Q01 ⇒ Probability that alarm has sounded but neither burgarly nor earthquake occurred but bot John and Marry call?

Using concept of Joint Probability;

P(Alarm ∩ Marry calls ∩ Jhon calls ∩ No Burglary ∩ No Earthquake)

↳ Now; we need to find relation in these 5 variables;

i) ~~A~~ Marry Calls given that Alarm rang

ii) Jhon Calls " " " "

iii) Alarm rangs given That No Burglary And No Earthquake.

Likelihood with its prior.

$\Rightarrow P(M/A)\,P(J/A)\,\cancel{P(\neg E, \neg B/A)}\,P(A/\neg B, \neg E).\,P(\neg B)\,P(\neg E)$

$= 0.70 \times 0.90 \times 0.001 \times 0.999 \times 0.998$

$P(\neg B, \neg E/A) = 0.00062$ Ans

Q-02 ⇒ Probability That Jhon calls?

Now; here we don't have info about marry, Earthquake, Burglary;

⇒ We skip Marry;

Now;

$P(\text{Jhon}) = P(J/A)P(A) + P(J/\neg A)P(\neg A)$

Now; Alarm depends upon burglary and earthquake; so we need Total probability; All 4;

TT
TF
FT
FF

and add all those.

So; Independent event Hence $P(B) \cdot P(E)$.

$$P(J) = P(J/A)\{P(A/B,E)*P(B,E) + P(A/\neg B,E)*P(\neg B,E) + P(A/B,\neg E)P(B,\neg E) + P(A/\neg B,\neg E)P(\neg B,\neg E)\}$$

+

$$P(J/\neg A)\{P(\neg A/B,E)*P(B,E) \ldots\ldots\}$$

$P(J) = 0.90 * 0.00252 + 0.05 * 0.9974$

$P(\text{Jhon}) = 0.0521.$

Date

Q-03 ⇒ Probability that there is burglary given that Jhon and Mary called?

Here: No information about the Earthquake and Alarm.

As Jhon and Mary both can Misinterpret alarm.

So, we need to cater all conditions!!

Here, we will find whether there is burglary or not burglary

written here as it is true that burglary occured

$$P(B/J,M) = \alpha P(B) \sum_a \left\{ P(J/A)P(M/A) \sum_e \left[P(A|B,E)P(E) \right] \right\}$$

$\sum_a$ cather both A and ¬A

$\sum_e$ cater both E and ¬E.

$$= \alpha P(B) \sum_a \left\{ P(J|A)P(M/A) \left[P(A|B,E)P(E) + P(A/B,\neg E)P(\neg E) \right] \right\}$$

Now open for A and ¬A.

$$= \alpha P(B) \left\{ P(J/A)P(M/A)\{P(A|B,E)P(E) + P(A|B,\neg E)P(\neg E)\} + P(J/\neg A)P(M/\neg A)\{P(\neg A|B,E)P(E) + P(\neg A|B,\neg E)P(\neg E)\} \right\}$$

Then put values and we do same with $P(\neg B|J,M)$.

jahan B hai wahan ¬B kardo.

So; $P(B|J,M) = \alpha * 0.00059$

$P(\neg B|J,M) = \alpha * 0.0015$

Now, finding α;

formula:- $\alpha = \dfrac{1}{(P(B,J,M) + P(\neg B,J,M))}$ ⇒ Both Numerator part.

$$\Rightarrow \alpha = \frac{1}{0.00059 + 0.0015} \Rightarrow 478.5$$

Now, put α value;

$P(B|J,M) = 478.5 * 0.00059 = 0.28$

$P(\neg B|J,M) = 478.5 * 0.0015 = 0.72$

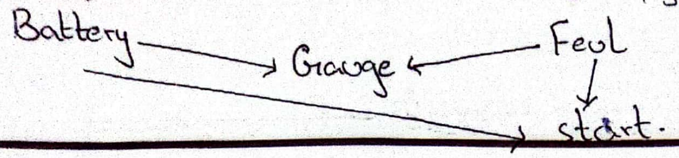
+ 1.00

1.00

Thus means our answer is write

Thus means there is **no** burglar in the home!!

$P(B=bad)=0.1$ $P(F=empty)=0.2$



Date

$P(Gauge) =$

Gauge	Battery	Feul	Value
Empty	Good	Not Empty	0.1
"	Good	Empty	0.8
"	Bad	Not Empty	0.2
"	Bad	Empty	0.9

$P(Start) =$

Start	Battery	Feul	Value
No	Good	Not Empty	0.1
"	Good	Empty	0.8
"	Bad	Not Empty	0.9
"	Bad	Empty	1.0.

1) $P(B=Good, F=Empty, G=Empty, S=Yes)$.

$P(0.8)P(0.9) \times P(0.2) \times P(0.2) = 0.0288$

2) Given that Battery is bad, Compute the probability that car will start.

$P(\text{Battery}=\text{Bad} \vee \text{Car Start}=\text{Yes})$ → No info about feul.

$= P(S=Yes \mid B=Bad, Feul=Empty) + P(S=Yes \mid B=Bad, F=Not\ Empty)$

$P(Yes \mid Bad, Feul)P(Feul) + P(Yes \mid Bad, \neg Feul)P(\neg Feul)$.

$= (0 \times 0.2) + (0.1 \times 0.8)$

$= 0.08$

{why we don't have P(Bad), As because it is already given}.

X — X — X — X.

~~Mileage~~

3) $P(\cancel{B=Good}, \cancel{F=Empty}, \cancel{G}$

~~Engine~~ ~~Air Conditioner~~

~~Car Value~~

* Used in:- Binary Classifications
i) Spam Detecting/Filtering
ii) Sentimental Analysis
iii) Classifying Articles.

Date

{W → white ; S → short ; Y → Yes ; G → green ; T → tall ; N → No}

NAIVE BAYES CLASSIFIER:-

No	Color	Legs	Height	Smelly	Species
1	W	3	S	Y	M
2	G	2	T	N	M
3	G	3	S	Y	M
4	W	3	S	Y	M
5	G	2	S	N	H
7	W	2	T	N	H
8	W	2	T	N	H
9	W	2	S	Y	H

$P(M) = 4/8 = 0.5$
$P(H) = 4/8 = 0.5$

Find probability:- {Color = G ; Legs = 2 ; Height = T ; Smelly = No}.

Color	M	H
White	2/4	3/4
Green	2/4	1/4

Legs	M	H
3	3/4	0/4
2	1/4	4/4

Height	M	H
Tall	3/4	2/4
Short	1/4	2/4

Smelly	M	H
Yes	3/4	1/4
No	1/4	3/4

• $P(M \mid \text{New Instance}) = P(M) * P(\text{Color: Green} \mid M) * P(\text{legs}=2 \mid M) * P(\text{Height: tall} \mid M) * P(\text{Smelly}=\text{No} \mid M)$.

$= 0.5 * 2/4 * 1/4 * 3/4 * 1/4$

$P(M \mid \text{New Instance}) = 0.0117$.

• $P(H \mid \text{NewInstance}) = 0.047$

$0.047 > 0.0117$

Hence New Instance = H

We can normalize the value by dividing by Total probability;

$P(M \mid \text{Instance}) = \dfrac{0.0117}{0.0117 + 0.047} = 0.1993$; $P(H \mid \text{Instance}) = \dfrac{0.047}{0.047 + 0.0117} = 0.8006$.

Hence;

Teacher's Signature

Date

⇒ Naive Bayes - Text Classification:-

• Includes = {Sentiment analysis, Spam detection, language identification}

Formula:- +1 smoothing:- $P(w_i|c) = \dfrac{count(w_i, c) + 1}{\left(\sum_{w \in V} count(w, c)\right) + |V|}$

word occurence in either class (+/-)

+1 → smoothing

$\sum$ count → Sum of Length of sentence. In Given sentence

$|V|$ → Unique word present in docs

Note:- If word not present in the document then we simply drop it.

Question:-

Category	Document
Spam	Send us your password
Spam	Review us
Spam	Send us your account
Spam	Send your password
Non-Spam	password review
Non-Spam	send us your review
?	review us now
?	review account

P(Spam) = 4/6 = 0.6666

P(Non-Spam) = 2/6 = 0.3333

count (w_i, c) ⇒

count(Review, spam) = 1

count (us, spam) = 3

count (now, spam) = 0

Vocabulary = 6

$\sum_{w \in V}$ count(words, spam) = 13

P(Review | spam) = (1+1)/13+6 = 0.1052

P(us | spam) = (3+1)/13+6 = 0.2105

P(spam) P(w|spam) = 0.6666 × 0.1052 × 0.2105 = 0.014

Same Do with non-spam.

Then select that whose probability is lowest.

{We will drop now as it is not present neither in spam nor in the non-spam.

If we are not using +1 smoothing techique then:-

Formula ⇒ P(spam) × P(x | spam)

{ here also we will ignore word in target which not exist in any document

Remember:- We stop for further calculation when we ...
Example:- Cloudy — Yes {Day - (3,7,12,13)} } hence we stop there and donot calculate further for Cloudy attribute!!

Date

DECISION TREE:

Used for both classification and Regression tasks.

Steps to make decision tree:- i) Calculate Entropy
ii) Calculate Information Gain;

Gatesmasher:-

Entropy of complete dataset:-

$S[+9, 5-] = -\frac{9}{14} \log_2 9/14 - 5/14 \log_2 5/14 = 0.940.$

Now find Entropy for all Attributes in Dataset;
{Temperature, Weather, Humidity, Wind};

i) Weather:-

Yes No

Entropy of Sunny = $[+2, 3-] = -2/5 \log_2 2/5 - 3/5 \log_2 3/5 = 0.97$

Entropy of Cloudy = $[+4, 0-] = -4/4 \log_2 4/4 - 0/4 \log_2 0/4 = 0$

Entropy of ~~Rain~~ Wind = $[+3, 2-] = -3/5 \log_2 3/5 - 2/5 \log_2 2/5 = 0.97.$

Step02:- Information Gain for Weather:-

= Entropy of whole dataset − {All Entropy of that column (weather)} * Data/Total

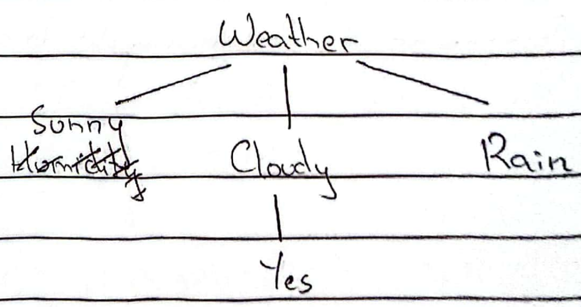
= $0.940 - 5/14\, Ent(S) - 4/14\, Ent(C) - 5/14\, Ent(R)$

= $0.940 - 0.346 - 0 - 0.344$

IG (Weather) = 0.248.

Now; we find IG for other attributes: {Temp, Humidity, Wind}
After calculation select highest IG as root and repeat process;

IG(Weather) = 0.248
IG(Temp) = 0.029
IG(Humidity) = 0.15
IG(Wind) = 0.0478



Now; which attribute to choose next?
⇒ Only pick {Sunny} data and check Entropy for all remaining attributes which are {Temp, Humidity, Wind}.
and same with Rain

{ Cloud → Yes ; As because in ... there are all Yes Hence we reach our first destination! }

Teacher's Signature

1	Sunny	Hot	High	Weak	No
2	"	Hot	Normal	Strong	Yes
8	"	Mild	⋮	⋮	No
9	"	Cool			⋮

→ Total value = 5

Date

Entropy of whole dataset which is only Weather = Sunny; [+2, 3-] = 0.97

Entropy of other attributes:

For Temperature :- Hot [+0, -2] = $-0/2 \log_2 0/2 - 2/2 \log_2 2/2 = 0$

Mild [+1, 1-] = 1.0

Cool [+1, 0-] = 0.

Information Gain for Temperature = Entropy (Sunny) − 2/5 Ent(Hot) − 2/5 (Mild) − 1/5 (Cool)

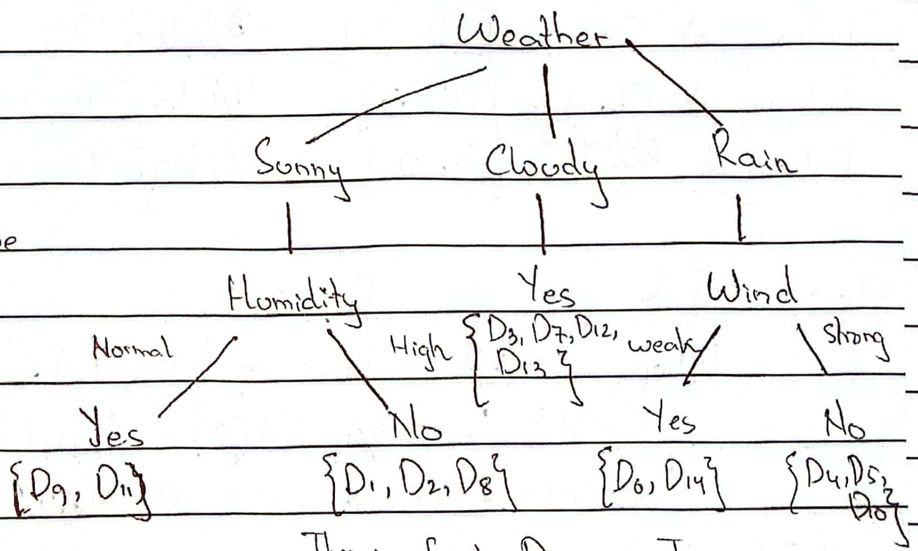
= 0.97 − 2/5 (0) − 2/5 (1) − 1/5 (0).

IG (Temperature) = 0.57.

Find other attributes IG w.r.t Sunny = {Humidity, ~~Wind~~ Wind}

IG (Humidity) = 0.97 ✓

IG (~~Temp~~ Wind) = 0.01



Same With Rain; As we done with Sunny!!

This is final Decision Tree;

Question:- Where does the temperature attribute?

Ans:→ Multiple reason can be there:-

i) Due to low IG ; it is not significantly impacting the decision.

ii) We have seen in example that our all days are set {D1 − D14} based on attributes for each Day

So we stop by further calculations! as we achieved on target lable {No, Yes}

iii) Pre-pruning or Post-pruning due to attribute not contributing in Information Gain!

To find best splitting attribute:-
i) Highest the gain best attribute
ii) lowest the gini Index best attribute.

Date

GINI INDEX Splitting Attribute Decision Tree.

First of all Formal :- $Gini = 1 - \sum_{i=1}^{n} (p_i)^2$

$Gini\ Index = \sum_{v \in (T,F)} \frac{|S_v|}{|S|} Gini(S_v).$

Weekend	Weather	Parents	Money	Decision.
W_1	Sunny	Yes	Rich	Cinema
W_2	Sunny	No	R	Tennis
⋮	Windy	Y	R	C
⋮	Rainy	Y	Poor	C
⋮	R	N	R	Stay In
⋮	R	Y	P	C
⋮	W	N	P	C
⋮	W	N	R	Shopping
⋮	W	Y	R	C
W_{10}	S	N	R	T

i) Gini for overall collection / Dataset;

Decision ⇒ {C = 6 ; T = 2 ; Stay = 1; Shopping = 1}.

$Gini(S) = 1 - \left[\left(\frac{6}{10}\right)^2 + \left(\frac{2}{10}\right)^2 + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^2\right] = 0.58$

ii) ⇒ Gini Index for Money :-

Poor = 3 value → all correspond to Cinema;

$Gini(S) = 1 - \left[\left(3/3\right)^2\right] = 0.$

⇒ Rich = 7 value → In which {Tennis = 2; Cinema = 3; Stay-In = 1; Shopping = 1

$Gini(S) = 1 - \left[\left(2/7\right)^2 + \left(3/7\right)^2 + \left(1/7\right)^2 + \left(1/7\right)^2\right] = 0.694.$

•) Gini Index for Money :- $0 * (3/10) + 0.694 * (7/10) = 0.486.$

Now, Same with Gini Index for Weather!!

Date

Weather:-

→ Sunny = 3 values → {Cinema = 2; Tennis = 1};

$$Gini(S) = 1 - \left\{ \left(\tfrac{2}{3}\right)^2 + \left(\tfrac{1}{3}\right)^2 \right\} = 0.444.$$

→ Now same for $\underset{\uparrow}{\overset{4\,value}{}}$Windy and Rainy; ↱ 3 values

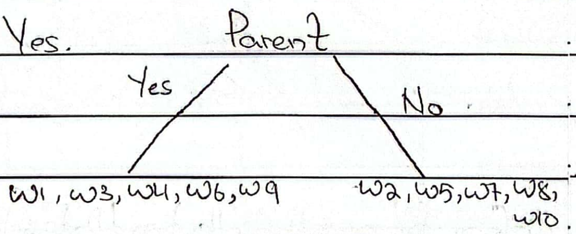
Gini (Windy) = 0.374 ; Gini (Rainy) = 0.444.

→ Gini Index for Weather:- $0.444 * (3/10) + 0.374 (4/10) + 0.444 * (3/10)$
$= 0.416.$

→ Directly calculated Gini Index for Parents = 0.36.

So, Parents Gini Index = 0.36 is lowest hence best attribute!!

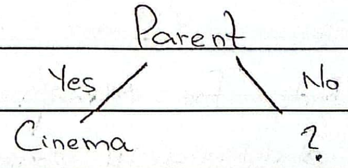
Now; Split Dataset into Parent = No and Yes.



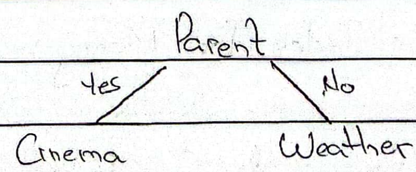
Repeat all process Aggain!!.

→ We can see from Dataset that when Parents are Yes than Decision = Cinema.
So; we achieved our first Destination hence Not to calculate on Parent = Yes.

So; let do: The No Data will seem as:-



Weekend	Weather	Parents	Money	Decision.
W2	S	No	R	Tennis
W5	R	:	R	Stay-In
W7	W	:	P	Cinema
W8	W	:	R	Shopping
W10	S	No	R	Tennis



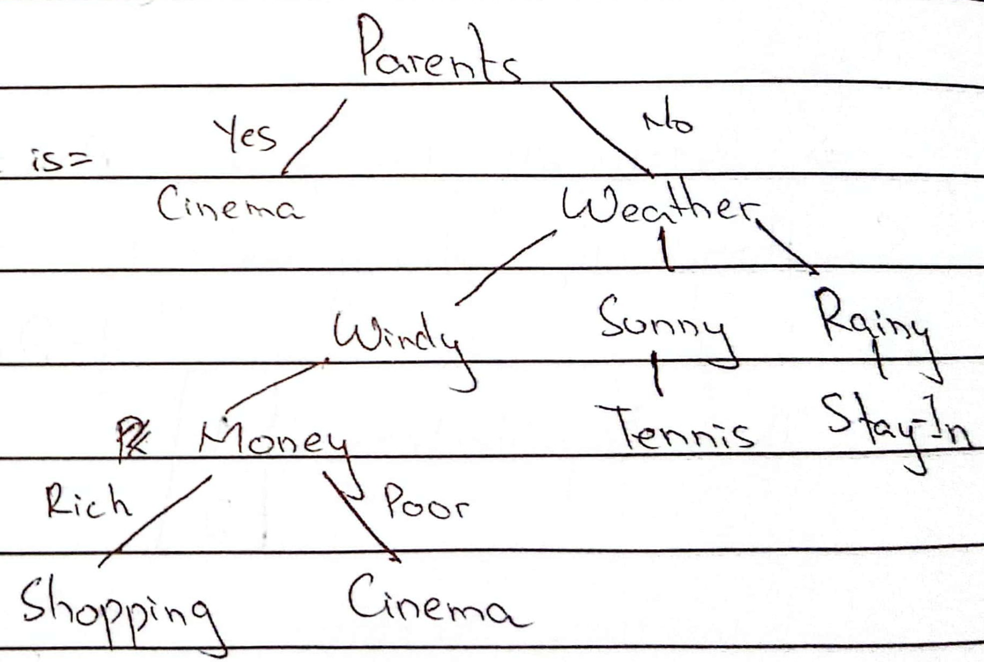
→ Gini Index for weather: = 0.2 ✓
Gini Index for Money = 0.5

Now, there is catch ; we didn't get to our destination:

Date

We see that when {Sunny = Tennis ; Rainy = Stay-In}
but when {Windy = {Cinema or Shopping}} ⇒ Hence need to reduced it.

So; here the Final Decision Tree is =



X — X — X — X

DIMENSIONALITY REDUCTION.

~~These value of B_0 and B_1 is found by:-~~

~~$B_0, B_1 = \dfrac{\Sigma((x-\bar{x})(y-\bar{y}))}{\Sigma(x-\bar{x})^2}$~~ ~~; $y/x \rightarrow$ value from table~~

~~; $\bar{x}/\bar{y} \rightarrow$ mean of the whole table column value.~~

~~$B_0 \rightarrow$ put value $\{\hat{x}, \hat{y}, B_1\}$ and find value!!~~

Formula: $y = \hat{a}x + \hat{b}$;

x	y	x^2	y^2	xy
-1	-1	1	1	1
1	2	1	4	2
2	3	4	9	6
4	3	16	9	12
Σ 6	7	22	23	21

Formula:-

$$\hat{a} = \frac{(\Sigma x)(\Sigma y) - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2}$$

$$\hat{b} = \frac{(\Sigma x)(\Sigma xy) - (\Sigma y)(\Sigma x^2)}{(\Sigma x)^2 - n\Sigma x^2}$$

standard deviation $= S_x = \sqrt{\dfrac{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2}{n-1}}$

$$S_y = \sqrt{\frac{\Sigma y^2 - \frac{1}{n}(\Sigma y)^2}{n-1}}$$

Regression Correlation:- $r = \left(\dfrac{S_x}{S_y}\right)\hat{a}$ OR $r = \dfrac{\Sigma xy - \frac{1}{n}(\Sigma x)(\Sigma y)}{(n-1)S_x S_y}$

X — X — X.

Parameter Estimation:-

$$\hat{B}_1 = \frac{\sum_{i=1}^{n} X_i Y_i - \dfrac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{n}}{\sum_{i=1}^{n} X_i^2 - \dfrac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}$$ In the Slides

$$\hat{B}_0 = \bar{Y} - \hat{B}_1\bar{X}$$

Date ______

QUESTION-01 (a):

$$a = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$

Given:-

Age = x_1

Blood Pressure = x_2

Cholestrol level = y.

$$b_1 = \frac{(\Sigma x_2^2)(\Sigma x_1 y) - (\Sigma x_1 x_2)(\Sigma x_2 y)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

$$b_2 = \frac{(\Sigma x_1^2)(\Sigma x_2 y) - (\Sigma x_1 x_2)(\Sigma x_1 y)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

Patient	Y	x_1	x_2	x_1x_1	x_2x_2	x_1x_2	x_1y	x_2y	yy
1	190	45	120	2025	14400	5400	8550	22800	36100
2	200	50	130	2500	16900	6500	10000	26000	40000
3	210	55	140	3025	19600	7700	18550	29400	44100
4	220	60	150	3600	22500	9000	13200	33000	48400
5	230	65	160	4225	25600	10400	14950	36800	52900
6	240	70	170	4900	28900	11900	16800	40800	57600
7	250	75	180	5625	32400	13500	18750	48000	62500
8	260	80	190	6400	36100	15200	20800	49400	67600
9	270	85	200	7225	40000	17000	22950	54000	72900
10	280	90	210	8100	44100	18900	25200	58800	78400
Σ =	2350	675	1650	47625	280500	115500	162750	39600	560500

Necessary Terms to solve :-

$\Sigma x_1^2 = \Sigma x_1 x_1 - (\Sigma x_1)(\Sigma x_1)/N \Rightarrow 47625 - (675)(675)/10 = 2062.5$

$\Sigma x_2^2 = \Sigma x_2 x_2 - (\Sigma x_2)(\Sigma x_2)/N \Rightarrow 280500 - (1650)(1650)/10 = 8250.$

$\Sigma x_1 y = \Sigma x_1 y - (\Sigma x_1)(\Sigma y)/N \Rightarrow 162750 - (675)(2350)/10 = 4125$

$\Sigma x_2 y = \Sigma x_2 y - (\Sigma x_2)(\Sigma y)/N \Rightarrow 396000 - (1650)(2350)/10 = 8250$

$\Sigma x_1 x_2 = \Sigma x_1 x_2 - (\Sigma x_1)(\Sigma x_2)/N \Rightarrow 115500 - (675)(1650)/10 = 4125.$

$$b_1 = \frac{(8250)(4125) - (4125)(8250)}{(2062.5)(8250) - (4125)}, \quad b_2 = \frac{(2062.5)(8250) - (4125)(4125)}{(2062.5)(8250) - (4125)}$$

$b_1 = 0$ $\qquad$ $b_2 = 0;$

Teacher's Signature

Date

• Regression. $a = \bar{Y} + b_1 x_1 + b_2 x_2$.

$a = \frac{2350}{10} + 0 + 0 \Rightarrow a = 235$ Ans

• $S_{x_1} = \sqrt{\frac{\Sigma(x_1 - \bar{x}_1)^2}{n-1}}$; $S_{x_2} = \sqrt{\frac{\Sigma(x_2 - \bar{x}_2)^2}{n-1}}$

$\Rightarrow \sqrt{\frac{\Sigma x_1^2 - 1/n(\Sigma x_1)^2}{n-1}} = \sqrt{\frac{47625 - 1/10(675)}{10-1}} = S_{x_1} = 15.13.$

$S_{x_2} = \sqrt{\frac{\Sigma x_2^2 - 1/n(\Sigma x_2)^2}{n-1}} = \sqrt{\frac{880500 - 1/10(7650)^2}{10-1}} = S_{x_2} = 30.27$

$S_y = \sqrt{\frac{5603500 - 1/10(2350)^2}{10-1}} = 30.27 \rightarrow S_y$ Ans

• Formula:- $r = \left(\frac{S_x}{S_y}\right)\hat{a}$

$\Rightarrow$ Between X_1 and X_2 ; $r = \left(\frac{S_{x_1}}{S_{x_2}}\right)\hat{a} \Rightarrow r = 117.46.$

$\Rightarrow$ Between X_1 and Y ; $r = 117.46.$

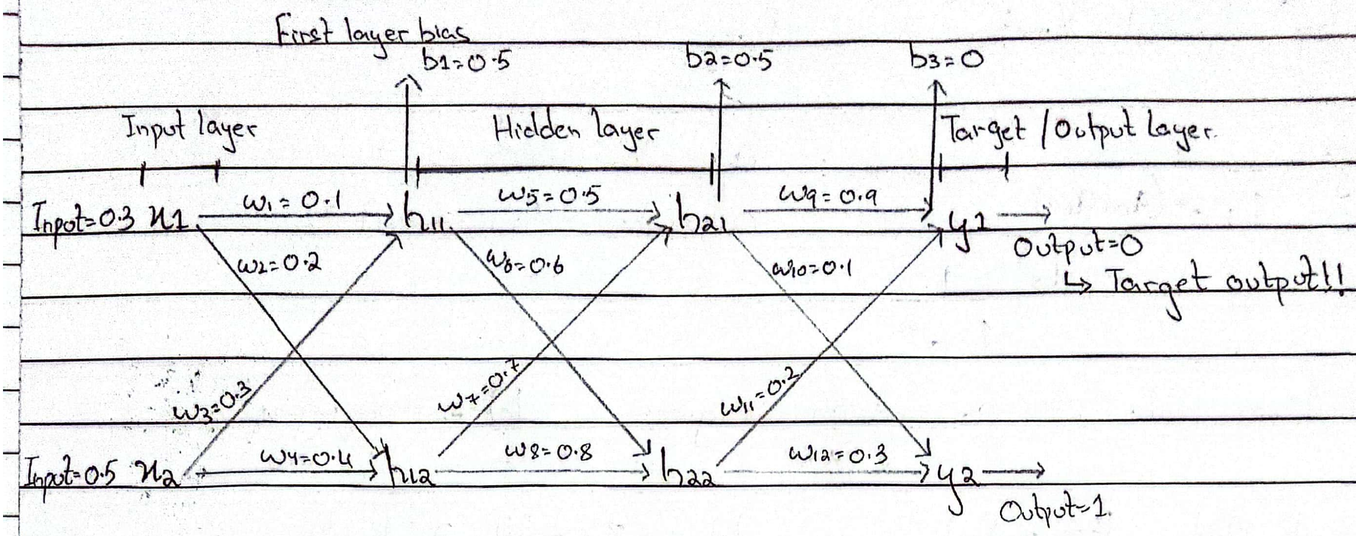
$\Rightarrow$ Between X_2 and Y ; $r = 235.0.$ Ans

• For $r^2 \Rightarrow$ Between X_1 and X_2. ; $r^2 = 13796.85$

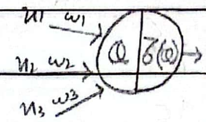
Between X_1 and Y ; $r^2 = 13796.85$

Between X_2 and Y ; $r^2 = 55225$ Ans

FEED FORWARD NEURAL NETWORK:-



⇒ A sigmoid function is calculated = $\sigma(Q) = \dfrac{1}{1+e^{-Q}}$ → Output ⇒



⇒ $Q = w_1x_1 + w_2x_2 + \cdots\cdots w_nx_n + b_1$

∵ w_n = weights.
x_n = Input
b_1 = Bias.

Lets; start calculation:-

•) $h_{11} = w_1x_1 + w_3x_2 + b_1$.
$h_{11} = 0.1 * 0.3 + 0.3 \times 0.5 + 0.5$
$h_{11} = 0.68$.

•) $h_{12} = w_2 \times x_1 + w_4 \times x_2 + b_1$
$h_{12} = 0.2 \times 0.3 + 0.4 \times 0.5 + 0.5$
$h_{12} = 0.76$

$\sigma(h_{11}) = \dfrac{1}{1+e^{-0.68}} = 0.66$

$\sigma(h_{12}) = \dfrac{1}{1+e^{-0.76}} = 0.68$

•) $h_{21} = w_5 \times \sigma(h_{11}) + w_7 \times \sigma(h_{12}) + b_2$
$h_{21} = 0.5 \times 0.66 + 0.7 \times 0.68 + 0.5$
$h_{21} = 1.306$.

•) $h_{22} = w_6 \times \sigma(h_{11}) + w_8 \times \sigma(h_{12}) + b_2$.
$h_{22} = 0.6 \times 0.66 + 0.8 \times 0.76 + 0.5$
$h_{22} = 1.504$

$\sigma(h_{21}) = 0.786$

$\sigma(h_{22}) = 0.818$

•) $y_1 = w_9 \times \sigma(h_{21}) + w_{11}\, \sigma(h_{22}) + b_3$
$y_1 = 0.9 \times 0.786 + 0.2 \times 0.818 + 0$
$y_1 = 0.1636$.

•) $y_2 = w_{10} \times \sigma(h_{21}) + w_{12} \times \sigma(h_{22}) + b_3$
$y_2 = 0.1 \times 0.786 + 0.3 \times 0.818 + 0$.
$y_2 = 0.324$

$\sigma(y_1) = 0.54$

$\sigma(y_2) = 0.58$

Date

⇒ Now calculate loss/Error:-

E/L = Mean Squared Error.

$$= \frac{1}{2}\left[(y_{A_1}' - y_{T_1}')^2 + (y_{A_2} - y_{T_2})^2\right]$$

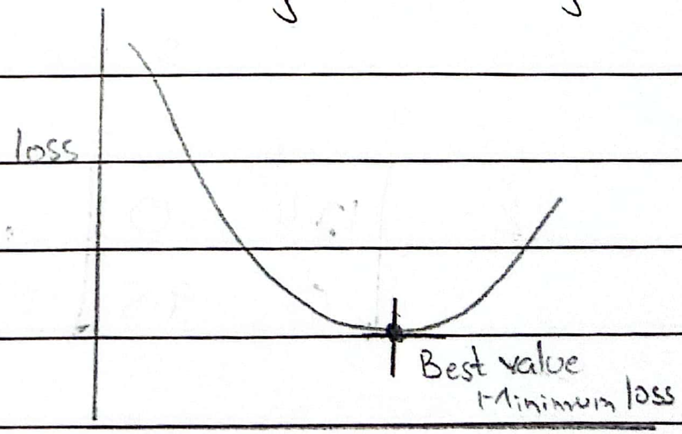
$$= 1/2\left[(0.54 - 0)^2 + (0.58 - 1)^2\right].$$

$$= 1/2\,(0.2916 + 0.1764) \quad = 0.468/2 = 0.234$$

So; this is our whole one iteration;

We will do same iteration until loss is minimize;

⇒ But we can't change our nodes or input so we change our weight



Date

K-Mean Clustering algorithm Disadvantage:-
i) Requires to specify the number of cluster (k) in advanced.
ii) We need to select the initial centroids for each of the cluster.

So; to cater these two problems we introduce k-mean ++ Algorithm.
step-01:- select randomly one point as centroid.
step-02:- for each data point compute its distance to all the centroids.
step-03:- Select next centroid as the maximum distance from all centroids.

Now; Question arises how to find best K value:-
Elbow Method:- Every centroid means k:{1,2,3....} we calculate wcss;
Step 01:- Formula:- WCSS =

$$\sum_{C_k}^{C_n} \left(\sum_{d_i \text{ in } C_i}^{d_m} \text{distance}(d_i, C_k)^2 \right).$$

↳ distance² of every point in a cluster.

step-02:- Map on the graph.
X-axis = k-value ; y-axis = WCSS
step-03:- The point we (k-value) we find that no difference is occuring when increasing k-value thats the best k-value (No of cluster).

X — X — X

Example:- K-mean:-

Initial Centroids:-
A1: (2,10)
B1: (5,8).

Data points			Distance to 2 10	Distance to 5 8	Cluster	New Cluster
A1	2	10	0.00	3.61	1	
A2	8	4	8.49	5.00	2	
A3	7	5	7.07	3.61.	2	
A4	5	8	3.61	0.00	2	

After 1st iteration; we will find mean of each cluster;
B1 = (5,8) :- points = {A2, A3, A4}; mean:- $x \Rightarrow \frac{8+7+5}{3}$; $y \Rightarrow \frac{4+5+8}{3}$.
↳ (6.6, 5.6).

A1 = (2,10):- points = {A1} ;- mean = (2,10).

So; new points = B1(6.6, 5.6); A1 = (2,10).
we will do second iterate;
if previous and new clustering value that is point those previously lie in cluster after new points lies within those only