Subject: | Date:

Marginal Probability:

	Male	Female	Total
GOT	0.16	0.24	0.40
TBBT	0.2	0.05	0.25
Others	0.1	0.25	0.35
Total	0.46	0.54	1

ⓐ What is the marginal probability of people watching TBBT?

$P(TBBT) = 0.25$

ⓑ What is the joint probability of a person being female and liking TBBT?

$P(F \wedge T) = 0.05$

ⓒ What is the probability of a person liking GOT given that person is male?

$$P(GOT \mid Male) = \frac{P(GOT \wedge Male)}{P(Male)} = \frac{0.16}{0.46} = 0.347$$

ⓓ

Subject : Date :

Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Bigg| \quad P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$P(A \cap B) \rightarrow$ Join probability

$P(A) \rightarrow$ Marginal Probability

Hence. $P(A|B) \times P(B) = P(B|A) \times P(A)$

Then, $P(A|B) = \dfrac{\left(P(B|A) \times P(A)\right)}{P(B)}$ [Bayes' Theorem]

* Are Male viewers and GOT independent?

$P(M \wedge GOT) = 0.16$ $\qquad$ $P(A \cap B) = P(A) \times P(B)$

$P(M) = 0.46$

$P(GOT) = 0.40$

$P(M) \times P(GOT) = 0.46 \times 0.40 = 0.184$

Since $P(M \wedge GOT) \neq P(M) \times P(GOT)$, so not independent.

	Smart		Not Smart	
	Study	Not Study	Study	Not Study
Prepared	0.432	0.16	0.084	0.008
Not Premared	0.048	0.16	0.036	0.072

Is smart conditionally independent of prepared given study?

$$P(SM \wedge PRE | STU) = \frac{P(SM \wedge PRE \wedge STU)}{P(STU)} = \frac{0.432}{0.6} = 0.72$$

Subject: Date:

Again,

$$P(SM \wedge PrE \mid STU) = P(SM \mid STU) * P(PrE \mid STU)$$

$$P(SM \mid STU) = \frac{P(SM \wedge STU)}{P(STU)} = \frac{0.432 + 0.048}{0.6} = 0.8$$

$$P(PrE \mid STU) = \frac{P(PrE \wedge STU)}{P(STU)} = \frac{0.516}{0.6} = 0.85$$

$\therefore P(SM \mid STU) * P(PrE \mid STU) = 0.8 \times 0.85 = 0.68 \neq 0.72$

$\therefore$ Not conditionally independent

A person is brought in front of a jury. The jury finds the defendant guilty in 98% of the cases in which he committed a crime and it finds the defendant not guilty 97% of the cases when the defendant has not committed a crime. Only 0.8% of the population has committed a crime.

If a random person is found guilty by the jury what is more likely, criminal or not?

Subject : Date :

Bayes' Theorem

Naive Bayes

1

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Problem 1:

Bag A contains 3 red balls
4 black balls

Bag B contains 5 red balls
6 black balls

Given

One ball is drawn at random from one of the bags and $\boxed{\text{its red.}}$

Find the probability that it was from bag B.

Solution:

Bag A: 3Red and 4Black

Bag B: 5Red and 6Black

$P(\text{Bag B given that ball is red}) = P(B_B|Red)$

$P(B_A) = 1/2$, $P(B_B) = 1/2$

$P(Red|B_A) = 3/7$, $P(Red|B_B) = \frac{5}{11}$

$$= \frac{P(B_B) \cdot P(Red|B_B)}{P(B_A) \cdot P(Red|B_A) + P(B_B) \cdot P(Red|B_B)}$$

$$= \frac{1/2 \times 5/11}{(1/2 \times 3/7) + (1/2 \times 5/11)}$$

$$= 35/68$$

Subject: | Date:

Problem 2:

Given three identical boxes A, B, C.

A has $\longrightarrow$ 2 gold coins

B has $\longrightarrow$ 2 silver coins

C has $\longrightarrow$ 1 gold and 1 silver coin.

What is the probability that the other coin in the box is also of gold if a person choose a box at random and take out a coin which is $\boxed{\text{gold?}}$ given

$P(B) = 1/3$, $P(A) = 1/3$, $P(C) = 1/3$

$P(G|B) = \cancel{\frac{2}{2}}\ 0/2$, $P(G|A) = 2/2$, $P(G|C) = 1/2$

$$P(A|G) = \frac{P(A) \times P(G|A)}{P(A) \times P(G|A) + P(B) \times P(G|B) + P(C) \times P(G|C)}$$

$$= \frac{1/3 \times 2/2}{(1/3 \times 2/2) + (1/3 \times 0) + (1/3 \times 1/2)}$$

$$= \frac{2/6}{2/6 + 1/6}$$

$$= 1/3 \times 2$$

$$= 2/3$$

Subject: 　　　　 Date: -

Problem 3:

Bag X → contains 1 White, 2 Red, 3 Green balls

Bag Y → contains 2 White, 3 Red, 1 Green balls

Bag Z → contains 3 White, 1 Red, 2 Green balls

Two balls were choosen from the bag and it was 1W and 1R. Find the probability that the balls so drawn came from the Bag Y.

Let E = Getting 1W and 1R

$$P(BY|E) = \frac{P(Y) \times P(E|Y)}{P(X) \times P(E|X) + P(Y) \times P(E|Y) + P(Z) \times P(E|Z)}$$

$$= \frac{\frac{1}{3} \times ({}^2C_1 \times {}^3C_1) / {}^6C_2}{\frac{1}{3} \times \frac{({}^1C_1 \times {}^2C_1)}{{}^6C_2} + \frac{1}{3}\frac{({}^2C_1 \times {}^3C_1)}{{}^6C_2} + \frac{1}{3}\frac{({}^3C_1 \times {}^1C_1)}{{}^6C_2}}$$

$$= 6/11$$

Subject: | Date:

Problem 4:

Suppose that 5% of men and 0.25% of women have grey hair. A grey haired persons selected at random. What is the probability of this person being male?

There are equal number of males and females.

$P(G|M) = \frac{5\%}{100\%}$; $P(G|W) = \frac{0.25\%}{100\%}$

$$P(M|G) = \frac{P(G|M) \cdot P(M)}{P(G|M) \cdot P(M) + P(G|W) \cdot P(W)}$$

$$= \frac{5/100 \times 1/2}{\left(5/100 \times 1/2\right) + \left(\frac{0.25}{100} \times 1/2\right)}$$

$$= \frac{1/40}{1/40 + \left(\frac{1}{800}\right)}$$

$$= 20/21$$

Subject : Date :

Problem 5:

In a certain college, 4% of the boys and 1% of girls are taller than 1.8m. 60% of the students are girls. If a student is selected at random and is found to be taller than 1.8m. What is the probability that the student is a girl?

E = taller than 1.8m

$P(E/B) = \frac{4}{100}$, $P(E/G) = \frac{1}{100}$, $P(B) = \frac{40}{100}$, $P(G) = \frac{60}{100}$

$$P(G/E) = \frac{P(E|G) \cdot P(G)}{P(E|G) \cdot P(G) + P(E|B) \cdot P(B)}$$

$$= \frac{1/100 \times \frac{60}{100}}{\left(\frac{1}{100} \times \frac{60}{100}\right) + \left(\frac{4}{100} \times \frac{40}{100}\right)}$$

$$= 60/220$$

Subject: Date:

Problem 6:

If a machine is correctly setup, it produces 90% acceptable items. If it is incorrectly set up, it produces only 40% acceptable items. Past experience shows that 80% of the set ups are correctly done. After a setup the machine produces 2 acceptable items.

Find the probability that the machine is correctly setup.

A = getting 2 acceptable items

$$P(C|A) = \frac{P(A|C) \times P(C)}{P(A|C) \times P(C) + P(A|C') \times P(C')}$$

$P(C) = \frac{80}{100}$ $\quad P(C') = \left(1 - \frac{80}{100}\right)$

$= 0.8$ $\quad = 0.2$

$P(A|C) = \left(\frac{90}{100} \times \frac{90}{100}\right)$ $\quad P(A|C') = \left(\frac{40}{100} \times \frac{40}{100}\right)$

$= 0.81$ $\quad = 0.16$

$$P(C|A) = \frac{0.81 \times 0.8}{(0.81 \times 0.8) + (0.16 \times 0.2)}$$

$$= \frac{81}{85} = 0.95$$

Subject: Date:

Problem 7:

In a factory which manufactures bolts, machine A, B, C manufactures respectively 25%, 35%, 40% of the bolts. Output 5%, 4%, 2% are defective bolts. A bolt is drawn at random and found to be defective. What is the probability that it is manufactured by the machine B?

$$P(B|D) = \frac{P(D|B) \times P(B)}{P(D|B) \times P(B) + P(D|A) \times P(A) + P(D|C) \times P(C)}$$

$P(A) = 1/3$, $P(B) = 1/3$, $P(C) = 1/3$

$P \left(= \frac{25}{100} \right) = \frac{35}{100} = \frac{40}{100}$

$P(D|A) = 5/100$, $P(D|B) = 4/100$, $P(D|C) = 2/100$

$$P(B|D) = \frac{\left(4/100 \times \frac{35}{100}\right)}{\left(4/100 \times \frac{35}{100}\right) + \left(5/100 \times 25/100\right) + \left(2/100 \times \frac{40}{100}\right)}$$

$$= \frac{0.04 \times 0.35}{(0.04 \times 0.35) + (0.05 \times 0.25) + (0.02 \times 0.4)}$$

$$= 0.405797.$$

Scanned with CamScanner

Subject: Date:

Problem 8:

Companies B1, B2, B3 produces 30%, 45%, 25% of the cars respectively. It is known that 2%, 3% and 2% of these cars produced from are defected.

(a) What is the probability that a car purchased is defective?

$P(B_1) = \frac{30}{100}$, $P(B_2) = 45/100$, $P(B_3) = 25/100$

$P(D|B_1) = 2/100$, $P(D|B_2) = 3/100$, $P(D|B_3) = 2/100$

$$P(D) = P(D|B_1)\,P(B_1) + P(D|B_2)\,P(B_2) + P(D|B_3)\,P(B_3)$$

$$= \left(\frac{2}{100} \times \frac{30}{100}\right) + \left(\frac{3}{100} \times \frac{45}{100}\right) + \left(\frac{2}{100} \times \frac{25}{100}\right)$$

$$= (0.02 \times 0.3) + (0.03 \times 0.45) + (0.02 \times 0.25)$$

$$= 0.0245$$

(b) If a car purchased is found out to be defective, What is the probability that this car is produced by Company B1?

$$P(B1|D) = \frac{P(D|B1) \times P(B1)}{P(D|B1) \times P(B1) + P(D|B_2) \times P(B_2) + P(D|B_3) \times P(B_3)}$$

Subject: Date:

$P(B_1) = \frac{30}{100}$, $P(B_2) = \frac{45}{100}$, $P(B_3) = \frac{25}{100}$

$P(D|B_1) = \frac{2}{100}$, $P(D|B_2) = 3/100$ $P(D|B_3) = 2/100$

$$P(B_1|D) = \frac{2/100 \times \frac{30}{100}}{\left(2/100 \times \frac{30}{100}\right) + \left(\frac{3}{100} \times \frac{45}{100}\right) + \left(2/100 \times \frac{25}{100}\right)}$$

$$= \frac{0.02 \times 0.3}{(0.02 \times 0.3) + (0.03 \times 0.45) + (0.02 \times 0.25)}$$

$$= \frac{6 \times 10^{-3}}{6 \times 10^{-3} + 0.0135 + 5 \times 10^{-3}}$$

$$= \frac{6 \times 10^{-3}}{0.0245}$$

$$= 0.244897$$

Problem: 9

A lot of IC chips contains 2% defective chips. Each is tested before delivery. The tester itself is not totally reliable. Probability of tester says the chip is good when it is really good is 0.95 and the probability of tester says the chip is defective when it is actually defective is 0.94. If a tested device is indicated to be defected, what is the probability that it is actually defective.

Subject: Date:

$\cancel{P(T|D) = \frac{P(D|T) \times P(T)}{P(D|T) \times P(T)}}$

T = tester say good

T' = tester say defective

$P(D) = 2/100$; $P(D') = 1 - 2/100$

$P(T|D') = 0.95$ $P(\neg T|D) = 0.94$

$P(\neg T|\neg D) = 1 - P(T|D')$
$= 1 - 0.95$
$= 0.05$

$$P(D|T') = \frac{P(\neg T|D) \times P(D)}{P(\neg T|D) \times P(D) + P(\neg T|\neg D) \times P(\neg D)}$$

$$= \frac{0.94 \times 0.02}{(0.94 \times 0.02) + (0.05 \times 0.98)}$$

$$= 0.27.$$

Subject: | Date:

Naive Bayes

Dataset

Day	Outlook	Temperature	Humidity	Windy	Play Tennis
1	Sunny	Hot	H	W	No
2	Sunny	Hot	H	S	No
3	Overcast	Hot	H	W	Yes
4	Rain	Mild	H	W	Yes
5	Rain	Cool	N	W	Yes
6	Rain	Cool	N	S	No
7	Overcast	Cool	N	S	Yes
8	Sunny	Mild	H	W	No
9	Sunny	Cool	N	W	Yes
10	Rain	Mild	N	W	Yes
11	Sunny	Mild	N	S	Yes
12	Overcast	Mild	H	S	Yes
13	Overcast	Hot	N	W	Yes
14	Rain	Mild	H	S	No

$P(\text{Play Tennis} = \text{Yes}) = \frac{9}{14} = 0.64$

$P(\text{Play Tennis} = \text{No}) = 5/14 = 0.36$

Outlook	Y	N
Sunny	2/9	3/5
Overcast	4/9	0
Rain	3/9	2/5

Hum.	Y	N
H	3/9	4/5
N	6/9	1/5

Windy	Y	N
W	6/9	2/5
S	3/9	3/5

Temp	Y	N
Hot	2/9	2/5
Mild	4/9	2/5
Cool	3/9	1/5

Subject: | Date:

(Outlook = sunny, Temp = cool, Hum = high, Wind = Strong)

$$V_{NB} = \underset{v_j \in \{yes, No\}}{\arg\max} \; P(v_j) \prod_i P(a_i | v_j)$$

$$= \underset{v_j \in \{yes, No\}}{\arg\max} \; P(v_j)$$

V_{NB} (yes):

$P(yes)\, P(sunny|Y)\, P(cool|Y)\, P(high|Y)\, P(strong|Y)$

$= 0.64 \times 0.222 \times 0.3333 \times 0.3333 \times 0.3333$

$= 0.00526$

V_{NB} (No):

$P(No)\, P(sunny|N)\, P(cool|N)\, P(high|N)\, P(strong|N)$

$= 0.36 \times 0.6 \times 0.2 \times 0.8 \times 0.6$

$= 0.020736$

$$V_{NB}(Yes) = \frac{V_{NB}(Yes)}{V_{NB}(Yes) + V_{NB}(No)} = \frac{0.00526}{0.00526 + 0.020736}$$

$$= 0.2023$$

$$V_{NB}(No) = \frac{V_{NB}(No)}{V_{NB}(No) + V_{NB}(Yes)} = \frac{0.020736}{0.00526 + 0.020736}$$

$$= 0.79766$$

Subject: Date:

Dataset

No	Color	Legs	Heights	Smelly	Species
1	White	3	S	Y	M
2	Green	2	T	N	M
3	Green	3	S	Y	M
4	White	3	S	Y	M
5	Green	2	S	N	H
6	White	2	T	N	H
7	White	2	T	N	H
8	White	2	S	Y	H

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

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

Color	M	H
W	2/4	3/4
G	2/4	1/4

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

Height	M	H
T	3/4	2/4
S	1/4	2/4

Smell	M	H
Y	3/4	1/4
S	1/4	3/4

(Color = Green ; Legs = 2, Height = Tall, Smelly = No)

$P(M) = P(M) * P(G|M) * P(2|M) * P(Tall|M) * P(No|M)$

$= 0.5 \times 2/4 \times 1/4 \times 3/4 \times 1/4 = 0.0117$

$(H) = P(H) * P(G|H) * P(2|H) * P(Tall|H) * P(No|H)$

$= 0.5 \times 1/4 \times 1 \times 2/4 \times 3/4 = 0.047$

Subject : | Date :

Decision Tree

Day	Outlook	Temp	Humidity	Wind	Play Tennis
01	Sunny	Hot	HA	Weak	No
02	Sunny	Hot	HA	Strong	No
03	Overcast	Hot	HA	Weak	Yes
04	Rain	Mild	HA	Weak	Yes
05	Rain	Cold	NR	Weak	Yes
06	Rain	Cold	NR	Strong	No
07	Overcast	Cold	NR	Strong	Yes
08	Sunny	Mild	HA	Weak	No
09	Sunny	Cold	NR	Weak	Yes
10	Rain	Mild	NR	Weak	Yes
11	Sunny	Mild	NR	Strong	Yes
12	Overcast	Mild	HA	Strong	Yes
13	Overcast	Hot	NR	Weak	Yes
14	Rain	Mild	HA	Strong	No

Entropy : $\sum_{i=1}^{n} -P \log_2 P$

$E(\text{Decision}) = -P(\text{Yes}) \log_2 P(\text{Yes}) + \left(-P(\text{No}) \log_2 P(\text{No})\right)$

$= -\frac{9}{14} \log_2\left(\frac{9}{14}\right) + \left(-\frac{5}{14} \log_2\left(\frac{5}{14}\right)\right)$

$= 0.9402859$

Subject :

Date :

[Outlook]

$E(\cancel{Yes}\,|\,\text{Sunny}) \Rightarrow$

$-P(\text{yes}|\text{Sunny})\log_2(\text{yes}|\text{Sunny}) - P(\text{No}|\text{Sunny})\log_2(\text{No}|\text{Sunny})$

$= -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - P\ \frac{3}{5}\log_2\left(\frac{3}{5}\right)$

$= 0.971$

$E(\text{Overcast}) \Rightarrow$

$-P(\text{yes}|\text{overcast})\log_2(\text{yes}|\text{overcast}) - P(\text{No}|\text{overcast})\log_2(\text{No}|\text{overcast})$

$= -4/4\log_2(4/4) - 0/4\log_2(0/4)$

$= 0$

$E(\text{Rain}) \Rightarrow$

$-P(\text{yes}|\text{Rain})\log_2(\text{yes}|\text{Rain}) - P(\text{No}|\text{Rain})\log_2(\text{No}|\text{overcast})$

$= -3/5\log_2(3/5) - 2/5\log_2(2/5)$

$= 0.971.$

Information Gain: [Outlook]

$E(\text{Decision}) - P(\text{Sunny})\times E(\text{Sunny}) - P(\text{Rain})\times E(\text{Rain}) - P(OC)\times E(OC)$

$= 0.9402 - (5/14 \times 0.971) - (5/14 \times 0.971) - (4/14 \times 0)$

$= 0.246.$

Subject : Date :

[Temp]

E(Hot) ⇒

$-P(\text{yes}|\text{Hot})\log_2(\text{yes}|\text{Hot}) - P(\text{No}|\text{Hot})\cdot\log_2(\text{No}|\text{Hot})$

$= -\frac{2}{4}\log_2\left(\frac{2}{4}\right) - \frac{2}{4}\log_2\left(\frac{2}{4}\right)$

$= 1.0$

E(Mild) ⇒

$-P(\text{yes}|\text{Mild})\log_2(\text{yes}|\text{Mild}) - P(\text{No}|\text{Mild})\log_2(\text{No}|\text{Mild})$

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

$= -4/6\log_2(4/6) - 2/6\log_2(2/6)$

$= 0.918295$

E(Cold) ⇒

$-P(\text{yes}|\text{Cold})\log_2(\text{yes}|\text{Cold}) - P(\text{No}|\text{Cold})\log_2(\text{No}|\text{Cold})$

$= -\frac{3}{4}\log_2\left(\frac{3}{4}\right) - \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)$

$= 0.811278$

Information gain:

$E(D) - P(\text{Hot})\times E(\text{Hot}) - P(\text{Mild})\times E(\text{Mild}) - P(\text{Cold})\times E(\text{Cold})$

$= 0.9402 - \left(\frac{4}{14}\times 1\right) - \left(6/14 \times 0.918\right) - \left(\frac{4}{14}\times 0.811278\right) = 0.0289$

Subject:

Date:

[Humidity]

$E(High) \Rightarrow$

$-P(yes|High)\, Log_2(yes|High) - P(No|High)\, Log_2(No|High)$

$= -\frac{3}{7} \log_2\left(\frac{3}{7}\right) - \left(\frac{4}{7}\right)\log_2\left(\frac{4}{7}\right)$

$= 0.985228136$

$E(Normal) \Rightarrow$

$-P(yes|Normal)\, Log_2(yes|Normal) - P(No|Normal)\, log_2(No|Normal)$

$= -\frac{6}{7} \log_2\left(\frac{6}{7}\right) - \frac{1}{7} \log_2\left(\frac{1}{7}\right)$

$= 0.59167277$

Information Gain:

$E(D) - P(High) \times E(High) - P(Normal) \times E(Normal)$

$= 0.94028 - \left(\frac{7}{14} \times 0.9852\right) - \left(\frac{7}{14} \times 0.59167\right)$

$= 0.1516$

Subject: Date:

[Wind]

E(Weak) ⇒

$-P(\text{Yes}|\text{Weak}) \log_2(\text{Yes}|\text{Weak}) - P(\text{No}|\text{Weak}) \log_2(\text{No}|\text{Weak})$

$= -\frac{6}{8}\log_2\left(\frac{6}{8}\right) - \frac{2}{8}\log_2\left(\frac{2}{8}\right)$

$= 0.8112781$

E(Strong) ⇒

$-P(\text{Yes}|\text{Strong}) \log_2(\text{Yes}|\text{Strong}) - P(\text{No}|\text{Strong}) \log_2(\text{No}|\text{Strong})$

$= -\frac{3}{6}\log_2\left(\frac{3}{6}\right) - \frac{3}{6}\log_2\left(\frac{3}{6}\right)$

$= 1.0$

Information Gain:

$E(D) - P(\text{Weak}) \times E(\text{Weak}) - P(\text{Strong}) \times E(\text{Strong})$

$= 0.940289 - \left(\frac{8}{14} \times 0.81127\right) - \left(\frac{6}{14} \times 1\right)$

$= 0.0478.$

Subject :

Date :

Summary:

$\boxed{\text{Gain(Outlook)} = 0.246}$ Gain (Humidity) = 0.1516

Gain (Temp) = 0.0289 Gain (Wind) = 0.0478

Outlook:

Sunny: 01, 02, 08, 09, 11 ⟶ X

Overcast: 03, 07, 12, 13 ⟶ Y

Rain: 04, 05, 06, 10, 14 ⟶ Z

$\boxed{X}$ Outlook = Sunny

Day	Temp	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes

$\boxed{\text{Temp}}$

$$E(\text{Sunny}) = -P(\text{Yes})\log_2 P(\text{Yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)$$

$$= 0.97$$

$$E(\text{Hot}) = -P(\text{Yes}|\text{Hot})\log_2 P(\text{Yes}|\text{Hot}) - P(\text{No}|\text{Hot})\log_2 P(\text{No}|\text{Hot})$$

$$= -\frac{0}{2}\log_2(0/2) - P(2/2)\log_2(2/2)$$

$$= 0$$

Subject : Date :

$E(Mild) \Rightarrow -P(yes|Mild)\log_2(yes|Mild) - P(No|Mild)\log_2(No|Mild)$

$= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

$E(Cold) \Rightarrow -P(yes|Cold)\log_2(yes|Cold) - P(No|Cold)\log_2(No|Cold)$

$= -\frac{1}{1}\log_2(\frac{1}{1}) - (\frac{0}{1})\log_2(\frac{0}{1})$

$= 0$

Information Gain:

$E(S) - P(Hot) \times E(Hot) - P(Mild) \times E(Mild) - P(Cold) \times E(Cold)$

$= 0.97 - 0 - 1 \times \frac{2}{5} - 0$

$= 0.570.$

Humidity

$E(High) \Rightarrow -P(yes|High)\log_2(yes|High) - P(No|High)\log_2(No|High)$

$= -\frac{0}{3}\log_2(\frac{0}{3}) - (\frac{3}{3})\log_2(\frac{3}{3})$

$= 0$

$E(Normal) = -P(yes|Normal)\log_2(yes|Normal) - P(No|Normal)\log_2(No|Normal)$

$= -(\frac{2}{2})\log_2(\frac{2}{2}) - (\frac{0}{2})\log_2(\frac{0}{2})$

$= 0$

IG:

$E(S) - P(H) \times E(H) - P(N) \times E(N) = 0.97 - 0 - 0$

$= 0.97$

Subject: Date:

$E(Weak) \Rightarrow -P(Yes|Weak)\log_2(Yes|Weak) - P(No|Weak)\log_2(No|Weak)$

$= -\frac{1}{3}\log_2(\frac{1}{3}) - \frac{2}{3}\log_2(\frac{2}{3})$

$= 0.9182$

$E(Strong) \Rightarrow -P(Yes|Strong)\log_2(Yes|Strong) - P(No|Strong)\log_2(No|Strong)$

$= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

Information Gain: $\boxed{\text{Wind}}$

$E(S) - P(W) \times E(W) - P(S) \times E(S)$

$0.971 - \frac{3}{5} \times 0.9182 - \frac{2}{5} \times 1$

$\Rightarrow 0.02008$

Summary:

Gain(Temp): 0.570

$\boxed{\text{Gain(Humidity): 0.97}}$

Gain(Wind) = 0.02008

$E(Hot) = 0$

$E(Mild) = -\frac{2}{3}\log_2(\frac{2}{3}) - \frac{1}{3}\log_2(\frac{1}{3})$

$= 0.91829$

$E(Cold) = -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2})$

$= 1.0$

$E(Hot) = 0.0$

$\boxed{2}$ outlook = Rain

Day	Temp	Humidity	Wind	PlayTennis
D4	Mild	High	Weak	Yes
D5	Cold	Normal	Weak	Yes
D6	Cold	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No

$E(Rain) = -\frac{3}{5}\log_2(\frac{3}{5}) - \frac{2}{5}\log_2(\frac{2}{5})$

$= 0.970950$

$E(High) = -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{2}) = 1.0$

$E(Normal) = -\frac{2}{3}\log_2(\frac{2}{3}) - \frac{1}{3}\log_2(\frac{1}{3}) = 0.9182$

Subject :

Date :

Information Gain: [Temp]

$E(Rain) - P(Hot) \times E(Hot) - P(Mild) \times E(Mild) - P(cold) \times E(cold)$

$= 0.970 - 0 - (3/5 \times 0.918) - (2/5 \times 1.0)$

$= 0.0192$

Information Gain: [Humidity]

$E(Rain) - P(High) \times E(High) - P(Normal) \times E(Normal)$

$= 0.970 - (2/5 \times 1.0) - (3/5 \times 0.9182)$

$= 0.0192$

$E(Weak) = -3/3 \log_2(3/3) - (0/3 \log_2(0/3)) = 0.0$

$E(strong) = -0/2 \log_2(0/2) - 2/2 \log_2(2/2) = 0.0$

Information Gain: [Wind]

$E(Rain) - P(Weak) \times E(Weak) - P(Strong) \times E(Strong)$

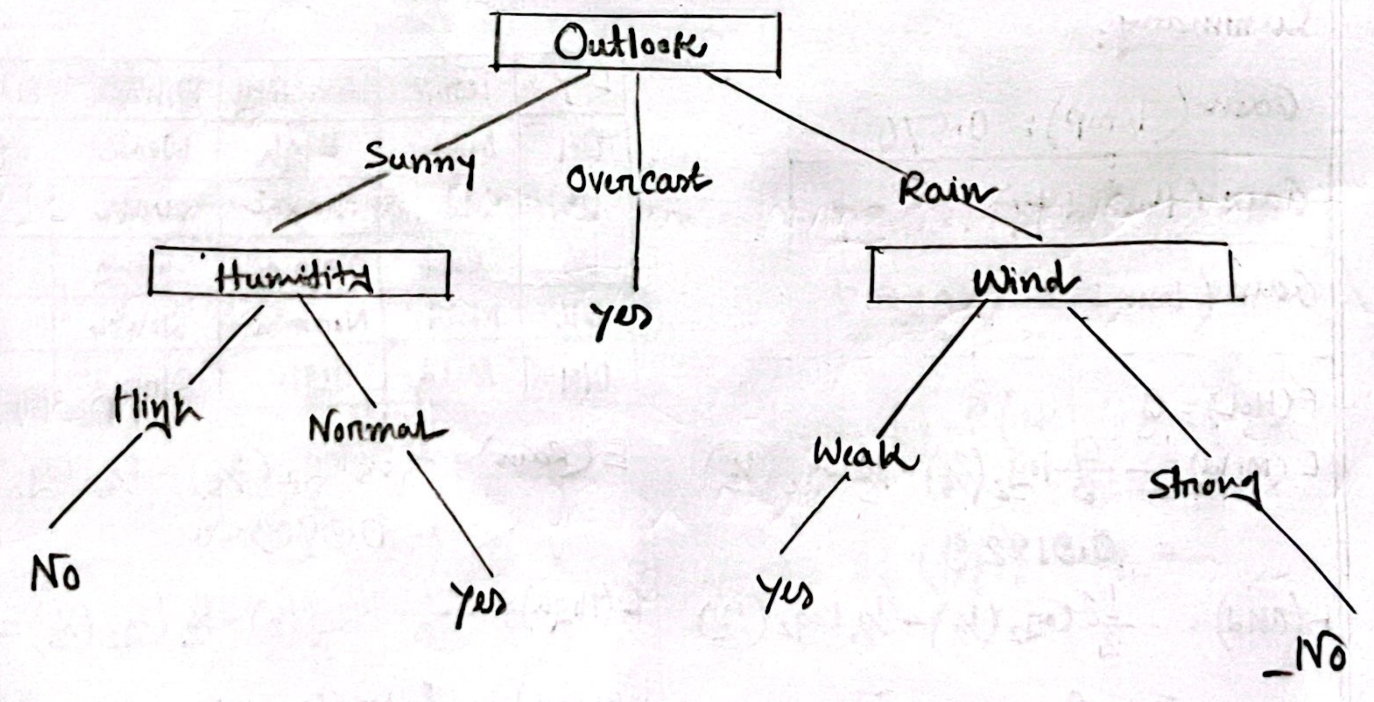
$= 0.97 - 0 - 0$

$= 0.97$

Temp = 0.0192

Humidity = 0.0192

$\boxed{\text{Wind} = 0.97}$



Subject: | Date:

Problems

(a) A patient went to the hospital for a malaria test. The doctor informed him that their test can successfully diagnose malaria positive given that patient is actually malaria positive 94% of the time.

Also, the probability of having no malaria and getting a malaria negative test result is 4%.

Meanwhile, 27% of people in general who come for tests are malaria positive. Now if the patient is already diagnosed malaria negative, then calculate the probability of the patient actually being malaria negative.

Given that,

P(Positive test | Actually positive) = 0.94

P(Negative test | Actually negative) = 0.04

P(Actually positive) = 0.27

P(Actually negative) = 1 - 0.27 = 0.73

P(Actually negative | Negative test)

$$= \frac{P(\text{Negative test} \mid \text{Actually negative}) \times P(\text{Actually negative})}{P(\text{Negative test})}$$

Subject: Date:

P(Negative test) = P(Negative test | Actually negative) × P(Actually neg) + P(Negative test | Actually positive) × P(Actually positive)

P(Negative test | Actually positive) = ~~1 − 0.04~~ 1 − 0.94
= ~~0.96~~ 0.06

∴ P(Actually negative | Negative test)

$$= \frac{(0.04 \times 0.73)}{(0.04 \times 0.73) + (0.06 \times 0.27)}$$

$$= \frac{0.0292}{0.0292 + 0.0162}$$

$$= 0.64317$$

(b) Is being malaria positive and having a positive test result independent of each other?

We know, $P(A \cap B) = P(A) \times P(B)$
$= 0.27 \times 0.94$
$= 0.2538$

Assume,
A = Malaria positive
B = Malaria positive test result

Given, $P(A) = 0.27$.

Subject:

Date:

$P(B) = P(\text{positive test})$

$= P(B|A) \times P(A) + P(B|A') \times P(A')$

$= P(\text{positive test} | \text{Actually positive}) \times P(\text{Actually positive})$

$+ P(\text{positive test} | \text{Actually negative}) \times P(\text{Actually negative})$

$= (0.94 \times 0.27) + ([1 - 0.04] \times 0.73)$

$= 0.2538 + \cancel{0.0438}\ 0.7008$

$= \cancel{0.2976}\ 0.9546$

$P(A \cap B) = 0.2538$

$P(A) = 0.27$

$P(B) = 0.9546$

$P(A) \times P(B) = 0.27 \times 0.9546$

$= 0.257742$

$P(A \cap B) \neq P(A) \times P(B)$

$\therefore$ They are not independent.

Subject: Date:

Covid-19 test all over the World aren't 100% accurate. A patient is actually positive in 85% of the cases When the test comes out to be positive. A person is actually positive in 10% of the cases When the test comes out to be negative.

(a) Of all the people who tested for Covid-19, 70% of them actually had the disease. If 1000 people participated in the tests, calculate the probability of a person's test results being positive.

P(Actually positive | positive test) = 0.85

P(Actually positive | negative test) = 0.10

P(Actually positive) = 0.70

Total people = 1000

P(Actually Positive | positive test)

$$= \frac{P(\text{positive test} \mid \text{Actually Positive}) \times P(\text{Actually Positive})}{P(\text{positive test})}$$

$$= \frac{\overset{x}{\cancel{0.85}} \times 0.70}{P(\text{positive test})} = 0.85 = \frac{0.70x}{P(\text{positive test})}$$

Subject: | Date:

$$P(\text{positive test}) = \frac{0.70x}{0.85} \quad \text{— (B)}$$

$$P(\text{Actually positive} \mid \text{negative test}) = 0.10$$

$$0.10 = \frac{P(\text{negative test} \mid \text{Actually positive}) \times P(\text{Actually positive})}{P(\text{negative test})}$$

$$0.10 = \frac{(1-x) \times 0.70}{1 - P(\text{positive test})}$$

$$0.10 = \frac{(1-x) \times 0.70}{1 - \frac{0.70x}{0.85}} \quad \text{or,} \quad 0.10\left(1 - \frac{0.70x}{0.85}\right) = (1-x) \times 0.70$$

or, $0.085 - 0.70x = 0.595 - 0.595x$

or, $x = \frac{0.51}{0.525} \approx 0.9714$

$\therefore P(\text{positive test} \mid \text{Actually positive}) = 0.9714$

$$P(\text{positive test}) = \frac{0.70 \times (0.9714)}{0.85}$$

$$= 0.8$$

$\therefore P(\text{positive test}) = 0.80$

total population = 1000

$\therefore$ people tested positive = (1000×0.80)

$= 800$ people.

Subject: | Date:

Consider two medical tests, A and B, for a virus. Test A is 95% effective at recognizing the virus when it is present, but has a 10% false positive rate (including that the virus is present, when it is not). Test B is 90% effective at recognizing the virus, but has a 5% false positive rate. The two test use independent methods of identifying the virus. The virus is carried by 1% of all the people. Say that a person is tested for the virus using only one of the tests. And the test comes back positive for carrying the virus.

Which test returning positive, is more indicative of someone really carrying the virus?

$$P(\text{Virus} \mid \text{Positive}) = \frac{P(\text{positive} \mid \text{virus})\, P(\text{virus})}{P(\text{positive})}$$

$$P(\text{positive}) = P(\text{positive} \mid \text{virus})\, P(\text{virus}) + P(\text{positive} \mid \text{No virus})\, P(\text{No virus})$$

For test A:

$$\frac{0.95 \times 0.01}{(0.95 \times 0.01) + (0.1 \times 0.99)} = 0.0876$$

For test B:

$$\frac{0.90 \times 0.01}{(0.90 \times 0.01) + (0.05 \times 0.99)} = 0.1538$$

Hamid Book Bainding

Scanned with CamScanner

Subject: | Date:

Linear Regression

$y = a_0 + a_1 x$] same
$y = b + mx$]

Linear Regression equation is given by $y = a_0 + a_1 x + \text{error}$

where

$$a_1 = \frac{(\overline{xy}) - (\bar{x})(\bar{y})}{\overline{x^2} - (\bar{x})^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$

x_i (week)	y_i (Sales in Thousands)
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

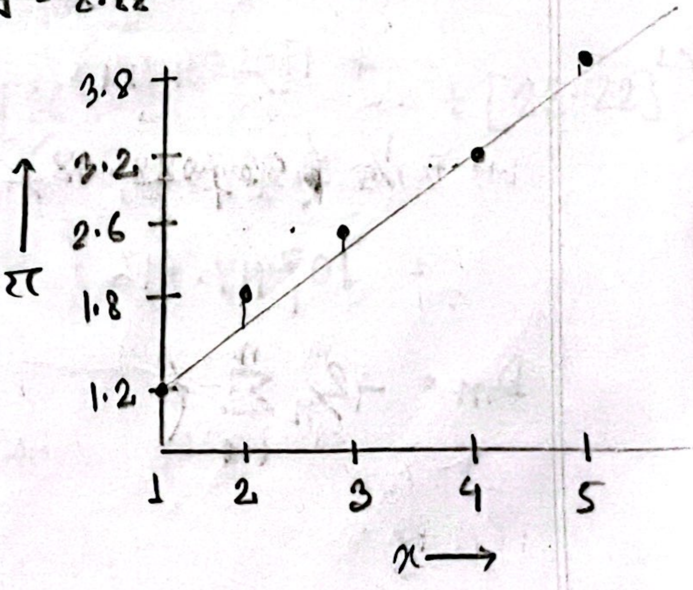
	x_i (week)	y_i (Sales in Thousands)	x_i^2	$x_i y_i$	Prediction
	1	1.2	1	1.2	1.2
	2	1.8	4	3.6	1.86
	3	2.6	9	7.8	2.52
	4	3.2	16	12.8	3.18
	5	3.8	25	19	3.84
sum	15	12.6	55	44.4	
Avg	3	2.52	11	8.88	

Now, $\bar{x} = 3$, $\bar{y} = 2.52$, $\overline{x_i^2} = 11$, $\overline{xy} = 8.88$

$$a_1 = \frac{(8.88) - (3 \times 2.52)}{(11) - (3)^2} = 0.66$$

$$a_0 = 2.52 - (0.66 \times 3) = 0.54$$

$$\boxed{\begin{aligned} \therefore y &= a_0 + a_1 x \\ &= 0.54 + 0.66x \end{aligned}}$$



For, $x = 7$

$y = 0.54 + 0.66 \times 7$

$= 5.16$

Subject :

Date :

Linear Regression Gradient Descent

Loss function $= \sum_{i}^{n} (y_{actual} - y_{pred})^2$

Mean Square Error (MSE) $= \frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2$

x_i	y_i
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

Lets assume that, $\boxed{y = 10x + 300}$

in $y = mx+b$,

$m = 10$, $b = 300$, learning Rate $= 0.001$

$MSE = \frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2$ [n = total row]

$= \frac{1}{n} \sum_{i=0}^{n} (y_i - [mx+b])^2$

$MSE = \frac{1}{5} \Sigma (1.2 - 310)^2 + (1.8 - 320)^2 + (2.6 - 330)^2 + (3.2 - 340)^2$
$+ (3.8 - 350)^2$

$= \frac{1}{5} \Sigma (95357.4 + 101251.24 + 107190.76 + 113434.24$
$+ 119854.44)$

$= \frac{1}{5} (537088.08)$

$= 107417.616$

$D_m = -\frac{2}{n} \sum_{i=0}^{n} ($

Subject: Date:

Problem Solving

$h_\theta(x) = \theta_0 + \theta_1 x$

Assume that, $\theta_0 = 1$ and $\theta_1 = 1.5$

X	Y	Pred_Y
7	11	11.5
2	4	4
17	25	26.5
9	15	14.5
4	7	7
11	16.5	17.5
12	19	19
6	10.2	10
1	2.3	2.5
3	5.1	5.5
2.5	4	4.75
3.8	7	6.7
5.6	11	9.4
14	22	22

Cost Function:

$$\frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

Gradient Decent

$$\frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

For, θ_0

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)$$

$$\theta_1 = \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right) \times x_i$$

Cost Function

$$\frac{1}{2 \times (14)} \sum \left([11.5 - 11]^2 + [4-4]^2 + [26.5 - 25]^2 + \cdots + [22-22]^2 \right)$$

$$\frac{1}{28} \times 7.2025$$

$$= 0.25366071$$

Subject: Date:

1st Iteration

Assume that $\theta_0 = 1$ and $\theta_1 = 1$

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^{2m}$$

$$\theta_0 = \theta_0 - \alpha \times \left[\frac{1}{14}(8-11) + (3-4) + (18-25) \ldots\ldots + (15-22)\right]$$

$$\theta_0 = \theta_0 - \alpha \times \left[\frac{1}{14} \times -47.2\right]$$

$$\theta_0 = 1 - 0.01 \times (-3.3714)$$

$$= 1.0337$$

$$\theta_1 = \theta_1 - \frac{\alpha}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)\, x_i$$

$$= 1 - \frac{0.01}{14} \left[7(8-11) + (3-4) \times 2 + (18-28) \times 17 \ldots\ldots\right]$$

$$= 1 - 0.01 \times (-33.6821)$$

$$= 1.336$$

x	y	Pred-y
7	11	8
2	4	3
17	25	18
9	15	10
4	7	5
11	16.5	12
12	19	13
6	10.2	7
1	2.3	2
3	5.1	4
2.5	4	3.5
3.8	7	4.8
5.6	11	6.6
14	22	15

Subject: Date:

~~Neural Network~~
Probability Problem

1. In the table below, you are given a dataset containing 9 rows and 3 feature. Using naive bayes determine the most likely value of if $X_1 = 1, X_2 = a, X_3 = q$.

X_1	X_2	X_3	Y
1	a	p	0
2	b	r	1
3	b	p	1
3	c	q	1
2	c	r	0
1	b	q	1
2	a	p	0
3	a	r	1
3	b	q	0

$P(Y=0 \mid X_1 = 1, X_2 = a, X_3 = q)$

For $Y=0$, rows = 0, 4, 6, 8

For $Y=1$, rows = 1, 2, 3, 5, 7

$P(X_1 = 1 \mid Y=0) = \frac{1}{4}$

$P(X_2 = a \mid Y=0) = \frac{2}{4}$

$P(X_3 = q \mid Y=0) = \frac{1}{4}$

$P(Y=0) = 4/9$

$P(Y=1) = \frac{5}{9}$

Subject: Date:

$$P(Y=0 \mid X_1=1, X_2=a, X_3=q) = \frac{4}{9} \times \cancel{\tfrac{5}{8}} \times \frac{1}{4} \times \frac{2}{4} \times \frac{1}{4}$$

$$= \frac{1}{72} \approx 0.0138$$

Again;

$P(X_1=1 \mid Y=1) = \frac{1}{5}$ $\quad P(X_2=a \mid Y=1) = \frac{1}{5}$

$P(X_2=q \mid Y=1) = \frac{2}{5}$

$$\therefore P(Y=1 \mid X_1=1, X_2=a, X_3=q) = \frac{5}{9} \times \frac{1}{5} \times \frac{2}{5} \times \frac{1}{5}$$

$$= \frac{2}{225} \approx 8.889 \times 10^{-3}$$

$0.0138 > 8.889 \times 10^{-3}$

$\therefore Y = 0$ [Ans]

2. $X = \{A, B, C\}$. Assume $P(A) = 0.5$, $P(B) = 0.3$.

Determine $P(C)$ and $P(A \cup B)$

We know that,

$$P(A) + P(B) + P(C) = 1$$

or, $0.5 + 0.3 + P(C) = 1$

or, $P(C) = 0.2$.

$P(A \cup B) = P(A) + P(B) = 0.5 + 0.3 = 0.8$

Subject: Date:

3. Assume 2 coins are tossed simultaneously.

Event A = The 1st coin comming up head.

Event B = The 2nd coin comming up tail.

What is $P(A \cap B)$?

Sample Space, S = {HH, HT, TH, TT}

When the first coin is head, {HH, HT}

When the 2nd coin is tail, {HT, TT}

$\therefore P(A \cap B) = \{HT\}$ | Total space = 4, output = 1

$\therefore P(A \cap B) = \frac{1}{4} = 0.025.$

4. (i) Assume Event A & B are absolute independence and $P(B|A) = 0.5$

What is the value of x?

	A		A'	
	B	B'	B	B'
C	0.1	0.2	0.2	Y
C'	X	0.1	0.1	0.

$P(B|A) = \frac{P(A \cap B)}{P(A)} \cdot \frac{P(B \cap A)}{P(A)} = 0.5$

Now,

$\frac{P(A \cap B)}{P(A)} = 0.5$

From the table, $P(B \cap A) = 0.1$

$\therefore P(A) = \frac{0.1}{0.5} = 0.2$

$P(A) = P(B) + P(B')$

$0.2 = 0.1 + X + 0.2 + 0.1$

$\therefore X = -0.2$

Subject: Date:

given that,

$P(B|A) = 0.5$

$P(A\cap B) = P(C\cap A\cap B) + P(C'\cap A\cap B)$

$= 0.1 + X$

$P(A) = P(C\cap A\cap B) + P(C\cap A\cap B') + P(C'\cap A\cap B) + P(C'\cap A\cap B')$

$= 0.1 + 0.2 + X + 0.1$

$= 0.4 + X$

$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{0.18+X}{0.4+X} = 0.5$

$0.5 = \frac{0.1+X}{0.4+X}$

or, $0.2 + 0.5X = 0.1 + X$

or, $0.2 - 0.1 = X - 0.5X$

or, $0.1 = 0.5X$

or, $X = 0.2$

(b) using $X = 0.2$ determine Y.

$0.1 + 0.2 + 0.2 + Y + 0.2 + 0.1 + 0.1 + 0.1 = 1$

or, $1 + Y = 1$

or, $Y = 0$

$\therefore y = 0$

Subject: | Date:

ⓒ Determine $P(A|B\cap C)$

$$P(A|B\cap C) = \frac{P(A\cap B\cap C)}{P(B\cap C)}$$

Here,
$P(A\cap B\cap C) = 0.1$
$P(B\cap C) = 0.1 + 0.2$
$= 0.3$

$$= \frac{0.1}{0.3} = 0.333$$

	Pandamic		No Pandamic		Total
	Online	Offline	Online	Offline	
Public University	0.142	0.037	0.165	0.072	
Private University	0.103	0.146	0.217	0.118	
Total					1

ⓐ Is pandamic Conditionality Independent of Public University Given that Online Class?

For Conditional Independence:

$$P(A|B,C) = P(A|C)$$

given that part

$$P(A|B,C) = \frac{P(A\cap B\cap C)}{P(B\cap C)}$$

$$P(A|C) = \frac{P(A\cap C)}{P(C)}$$

$$P(\text{Pand}|\text{Pub, Online}) = \frac{P(\text{Pand}\cap\text{Pub}\cap\text{online})}{P(\text{Pub}\cap\text{online})}$$

$$= \frac{0.142}{0.142 + 0.165}$$

$$= 0.46254.$$

$\boxed{\text{Not conditionally Independent.}}$

$$\cancel{P(\text{Pub}} \; P(\text{Pand}|\text{online}) = \frac{P(\text{Pand}\cap\text{Online})}{P(\text{online})} = \frac{0.142 + 0.103}{0.627}$$

$$= 0.391$$

Subject: Date:

(b) Is Private Uni Independent of Online Class?

$P(\text{Priv} \cap \text{Online}) = P(\text{Priv}) \cdot P(\text{online})$

$P(\text{Private} \cap \text{online}) = 0.103 + 0.217 = 0.32$ —— (a)

$P(\text{Priv}) = 0.103 + 0.146 + 0.217 + 0.118 = 0.584$

$P(\text{Online}) = 0.142 + 0.103 + 0.165 + 0.217 = 0.627$

$P(\text{Private}) \cdot P(\text{online}) = 0.584 \times 0.627$

$= 0.366168$ —— (b)

Not Independent

(c) Find the marginal probability of offline class

$P(\text{Public} \Rightarrow (0.037 + 0.146 + 0.072 + 0.118)$

$= 0.373$

Subject :		Date :		

ⓐ $h_w(x) = w_1 x + w_0$

here compare the equation with $y = mx + c$

$$m = w_1 \text{ and } c = w_0$$

w_1 = slope of the line.

w_0 = the intercept point.

ⓑ given that,

$$Loss(h_w) = \sum_{j=1}^{N} \left(y_i - (w_1 x_j + w_0)\right)^2$$

$$y \rightarrow \frac{\partial Loss}{\partial w_1} = -2 \sum x_j (y_j - h_w(x_j))$$

$$x \rightarrow \frac{\partial Loss}{\partial w_0} = -2 \sum (y_i - h_w(x_j))$$

Initially $w_1 = 15$, $w_0 = 0$

Iteration	$h(w)$	Loss	x	y	w_0	w_1
1	$-410.8x - 132$	-66	132	425.8	-132	-410.8
2	$31495.3x - 11726.2$	5959.08	-11918.16	-31906.096	11786.16	31495.3

$h_w(x) = 15x + 0$

$h_{w_1}(x) = 15(1.6) = 24$

$h_{w_2}(x) = 15(2.0) = 30$

$h_{w_3}(x) = 15(2.5) = 37.5$

$h_{w_4}(x) = 15(3.0) = 45$

$h_{w_5}(x) = 15(3.5) = 52.5$

Here, error = Actual − Prediction

$e_1 = (30 - 24) = 6$

$e_2 = (27 - 30) = -3$

$e_3 = (24 - 37.5) = -13.5$

$e_4 = (22 - 45) = -23$

$e_5 = (20 - 52.5) = -32.5$

Subject : | Date :

$$\frac{\partial Loss}{\partial w_0} = -2\sum (y_i - h_w(x))$$

$$= -2(6-3-13.5-23-32.5)$$

$$= -132$$

$$\frac{\partial Loss}{\partial w_1} = -2\sum x_i (y_i - h_w(x))$$

$$= -2\left[(1.6\times 6) + (2.0\times -3) + (2.5\times -13.5) + (3\times -23) + (3.5\times -32.5)\right]$$

$$= -2[9.6 - 6.0 - 33.75 - 69 - 113.75]$$

$$= -2(-212.9)$$

$$= 425.8$$

Now,

$$w_0 = w_0 - \alpha \frac{\partial Loss}{\partial w_0}$$

$$= 0 - 1\times(132)$$

$$= -132$$

$$w_1 = w_1 - \alpha \frac{\partial Loss}{\partial w_1}$$

$$= 15 - 1\times(425.8)$$

$$= -410.8$$

Final equation : $-410.8x - 132$

Subject : | Date :

Again,

$hw(x) = -410.8x - 132$

$hw_1(x) = -410.8(1.6) - 132 = -789.28$

$hw_2(x) = -410.8(2.0) - 132 = -953.6$

$hw_3(x) = -410.8(2.5) - 132 = -1159$

$hw_4(x) = -410.8(3.0) - 132 = -1364.4$

$hw_5(x) = -410.8(3.5) - 132 = -1569.8$

Error = Actual - Prediction Value

$e_1 = (30 + 789.28) = 819.28$

$e_2 = (27 + 953.6) = 980.6$

$e_3 = (24 + 1159) = 1183$

$e_4 = (22 + 1364.4) = 1386.4$

$e_5 = (20 + 1569.8) = 1589.8$

$$\frac{\partial Cost}{\partial w_0} = -2\sum(y_i - hw(x))$$

$$= -2(819.22 + 980.6 + 1183 + 1386.4 + 1589.8)$$

$$= -11918.16$$

$$\frac{\partial Cost}{\partial w_1} = -2\sum x(y_i - hw(x))$$

$$= -2[(1.6 \times 819.28) + (2.0 \times 980.6) + (2.5 \times 1183) + (3.0 \times 1386.4) + (3.5 \times 1589.8)]$$

$$= -2[1310.848 + 1961.2 + 2957.5 + 4159.2 + 5564.3]$$

$$= -31906.096$$

Hamid Book Bainding

Subject : | Date :

Now,

$$w_0 = w_0 - \alpha \frac{\partial Loss}{\partial w_0}$$

$$= -132 - 1(-11918.16)$$

$$= 11786.16$$

$$w_1 = w_1 - \alpha \frac{\partial Loss}{\partial w_1}$$

$$= -410.8 - 1(-31906.096)$$

$$= -410.8 + 31906.096$$

$$= 31495.296$$

Final equation: $31495.296x + 11786.16$

ⓒ from question 2;

$$31495.296x + 11786.16$$

for $x = 3$, $y = 31495.296(3) + 11786.16$

$$= 106272.048 \text{ mpg}$$

Actual = 22 mpg

Predict = 106272.048 mpg

difference = $(106272.048 - 22)$ mpg

$$= 106250.048 \text{ mpg}$$

Here learning rate $\alpha = 1$ causes a misslead of prediction. The rate must in range of 0.01 to 0.001 for better output.

Subject : Date :

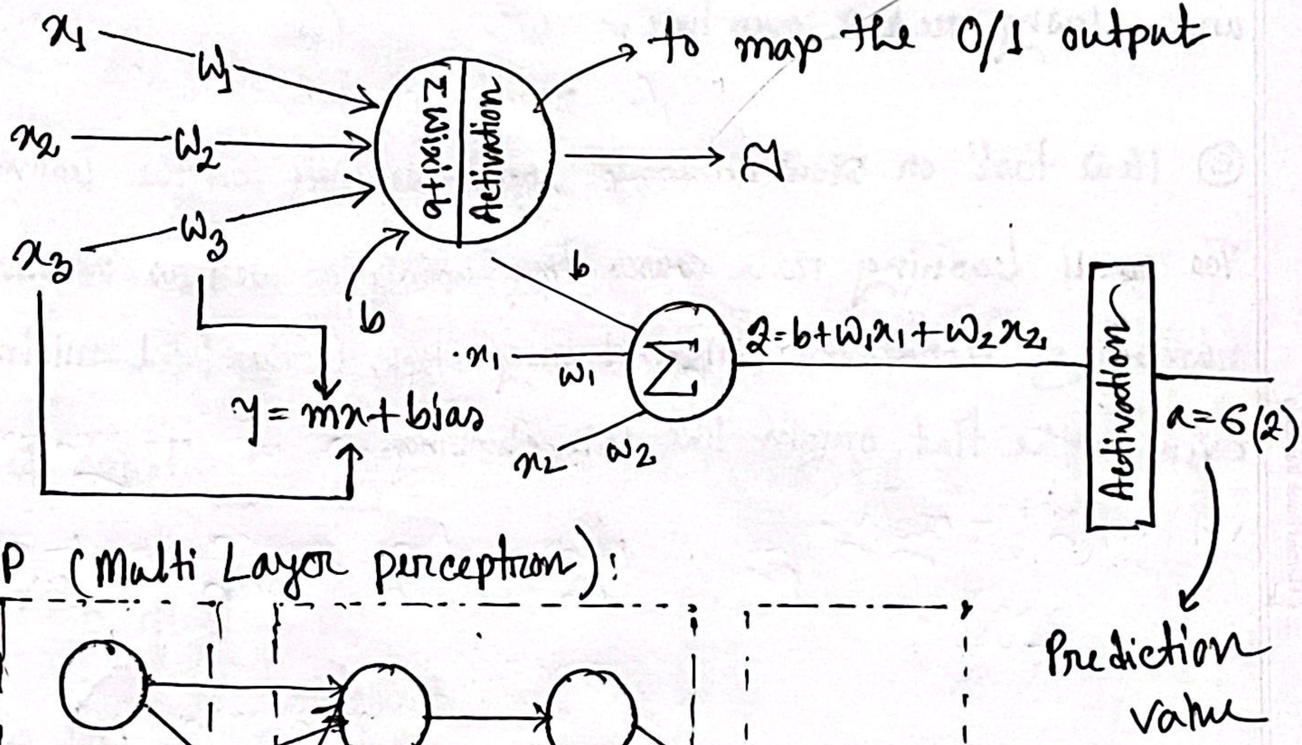
(d) Large learning rate causes of divergence of the output. Also a large learning rate may cause a bad prediction. On the other hand, a decaying learning rate start with a high value to jump quickly and slowly reduce over-time.

(e) How fast or slow a model learn, depends on the learning rate. Too small learning rate causes tiny update as well as increase the number of iteration. Also, it may stock in the local minima or even in the flat origin like hill climbing.

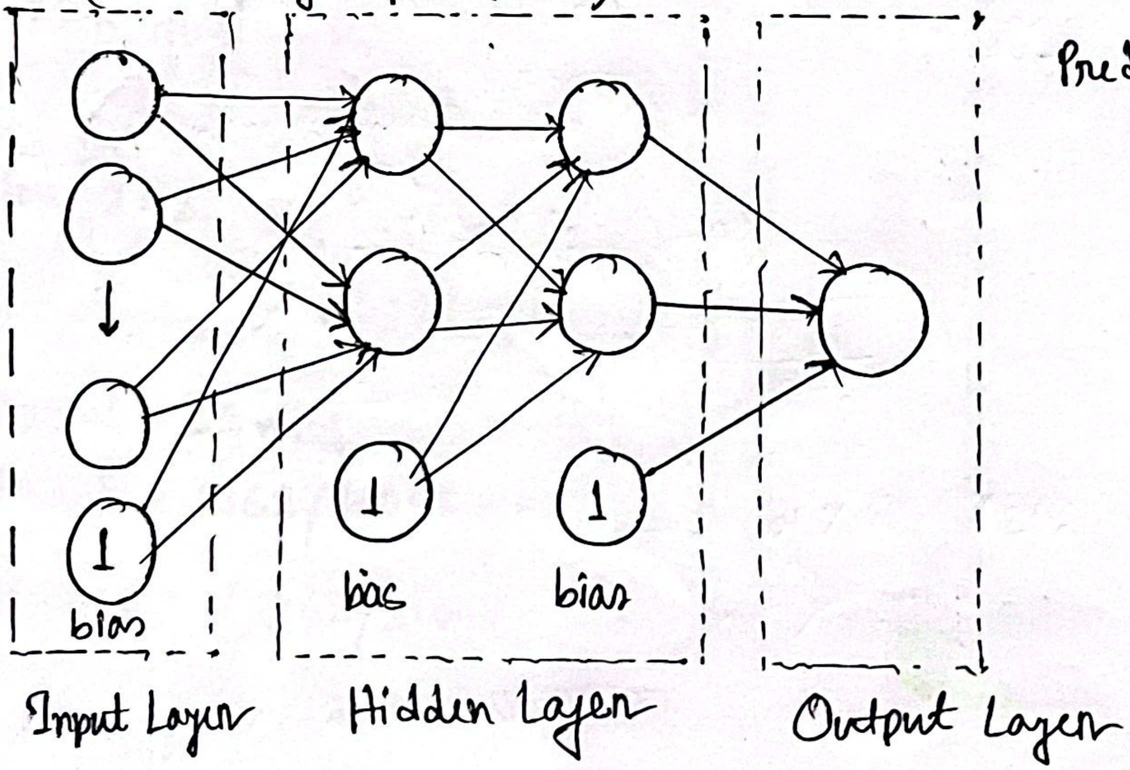
Neural Network

What is a perceptoron?

One layer neural network that can classify things into two parts.

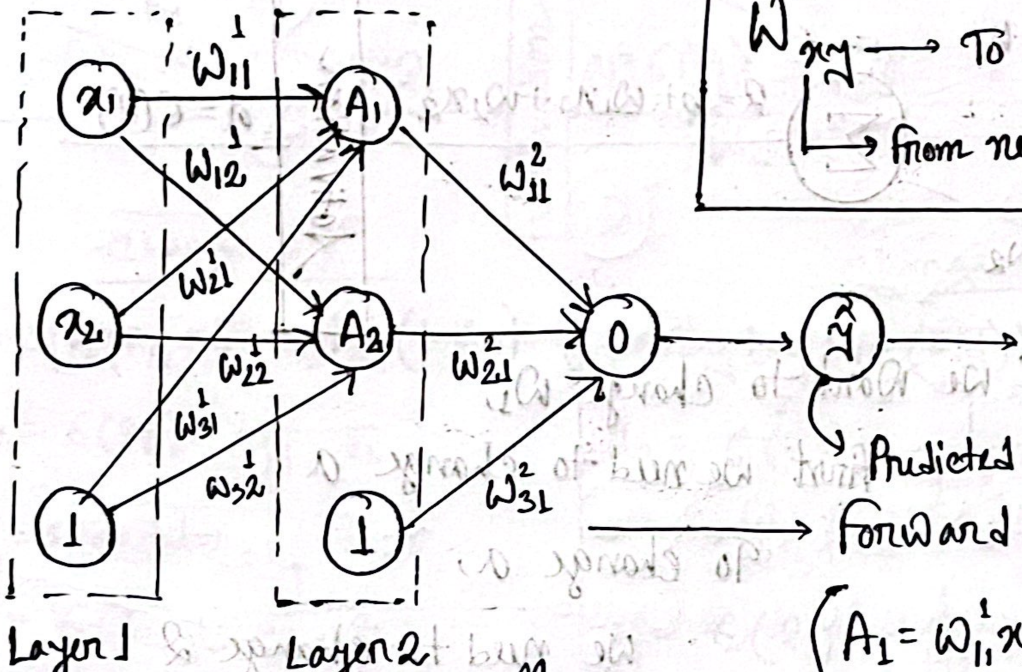


MLP (Multi Layer perceptron):



Forward Propagation

W^{2}_{xy} → Layer number (2), To neuron (y), from neuron (x)



Predicted value

→ Forward Propagation

$$\sum_{i=1}^{n} W_i x_i + b = \begin{cases} A_1 = W^1_{11}x_1 + W^1_{21}x_2 + W^1_{31} \\ A_2 = W^1_{12}x_1 + W^1_{22}x_2 + W^1_{32} \\ O = W^2_{11}\sigma(A_1) + W^2_{21}\sigma(A_2) + W^2_{31} \end{cases}$$

$$\underbrace{\sigma(x) = \frac{1}{1+e^{-x}}}_{\text{Sigmoid Function}}$$

Also called as the activation function

→ $\boxed{A_1/A_2}$ A_1 or A_2 or ---- A_n

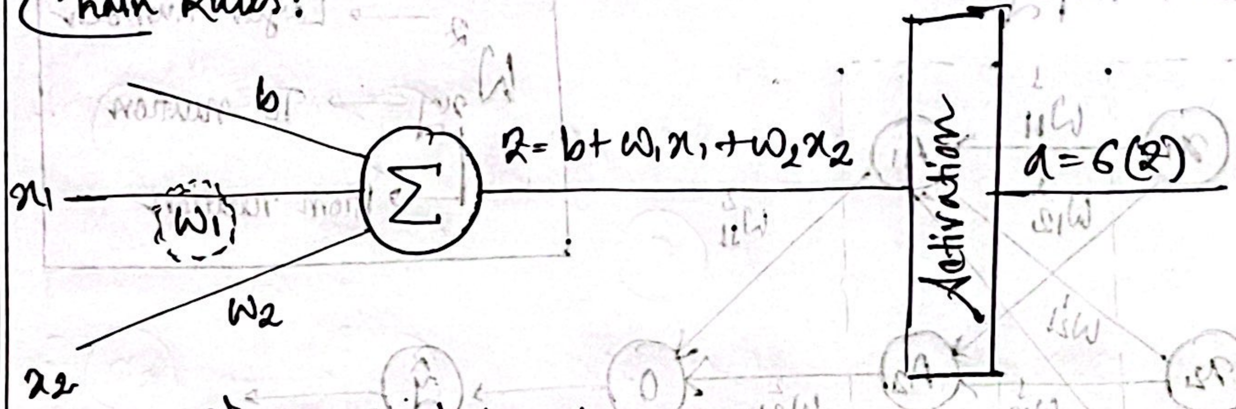
$$\hat{y} = \sigma \cdot \begin{bmatrix} W^2_{11} & W^2_{21} & W^2_{31} \end{bmatrix} \cdot \sigma \begin{bmatrix} W^1_{11} & W^1_{21} & W^1_{31} \\ W^1_{12} & W^1_{22} & W^1_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

$$\hat{y} = \sigma \cdot W^2 \cdot \sigma W^1 x$$

- W^1 → Weight of Layer 1
- W^2 → Weight of Layer 2
- σ → Sigmoid Function

Subject: Date:

Chain Rules:



If we want to change w_1,

first we need to change a

To change a,

we need to change z

According to the chain rules:

$$\frac{dL}{dw_1} = \frac{dz}{dw_1} \times \frac{da}{dz} \times \frac{dL}{da}$$

$$\underbrace{L = -y\log a - (1-y)\log(1-a)}_{\text{LOSS}}$$

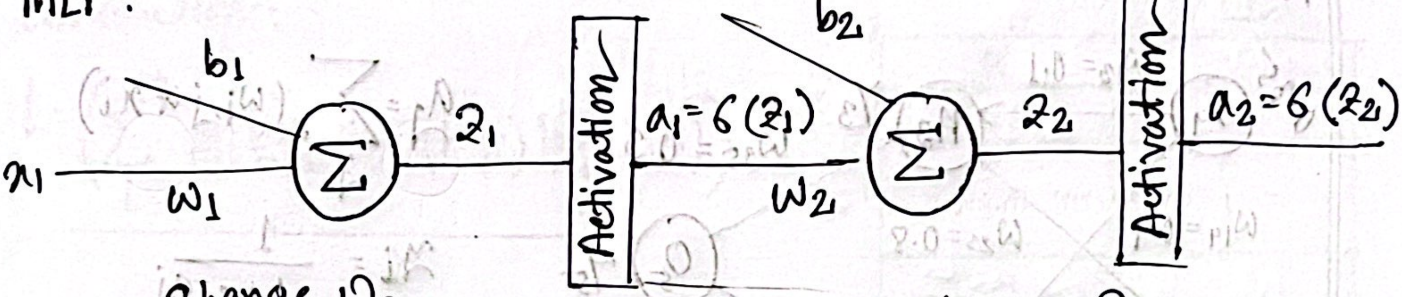
Why this dependency?

$L = -y\log a - (1-y)\log(1-a)$

$\quad \rightarrow a = \sigma(z) = \dfrac{1}{1+e^{-z}}$

$\quad\quad \rightarrow z = b + x_1w_1 + x_2w_2$

MLP :



<u>change w_2</u>

$L = -y \log a_2 - (1-y)\log(1-a_2)$

$a_2 = \sigma(z_2) = \frac{1}{1+e^{-z_2}}$

$z_2 = a_1 w_2 + b_2$

$a_1 = \sigma(z_1) = \frac{1}{1+e^{-z_1}}$

$z_1 = x_1 w_1 + b_1$

$$\frac{dL}{dw_2} = \frac{dz_2}{dw_2} \times \frac{da_2}{dz_2} \times \frac{dL}{da_2}$$

<u>change w_1</u>

$L = -y \log a_2 - (1-y)\log(1-a_2)$

$a_2 = \sigma(z_2) = \frac{1}{1+e^{-z_2}}$

$z_2 = a_1 w_2 + b_2$

$a_1 = \sigma(z_1) = \frac{1}{1+e^{-z_1}}$

$z_1 = x_1 w_1 + b_1$

$$\frac{dL}{dw_1} = \frac{dz_1}{dw_1} \times \frac{da_1}{dz_1} \times \frac{dz_2}{da_1} \times \frac{da_2}{dz_2} \times \frac{dL}{da_2}$$

$\boxed{w_1}$ $\boxed{w_2}$

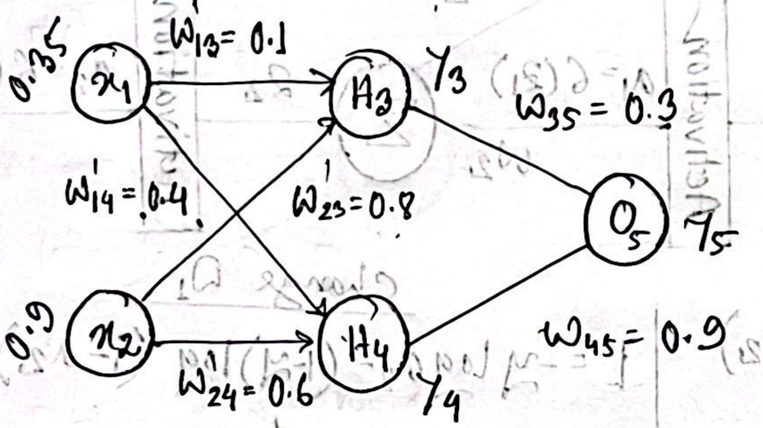
Here w_1 and w_2 are independent. Change in w_1 doesnot effect in w_2. But does effect a_1 and z_1.

For w_1, w_2 & w_3:

$$\frac{dL}{dw_1} = \frac{dz_1}{dw_1} \times \frac{da_1}{dz_1} \times \frac{dz_2}{da_1} \times \frac{da_2}{dz_2} \times \frac{dz_3}{da_2} \times \frac{da_3}{dz_3} \times \frac{dL}{da_3}$$

Subject :	Date :

Problems:



$$a_j = \sum_i (w_{i,j} * x_i)$$

$$y_i = \frac{1}{1+e^{-a_i}}$$

$$Y_5 = 0.5$$

$$a_1 = (W_{13} \times x_1) + (W_{23} \times x_2)$$
$$= (0.1 \times 0.35) + (0.8 \times 0.9)$$
$$= 0.755$$

$$y_3 = \frac{1}{1+e^{-0.755}}$$
$$= 0.68$$

$$a_2 = (W_{14} \times x_1) + (W_{24} \times x_2)$$
$$= (0.4 \times 0.35) + (0.6 \times 0.9)$$
$$= 0.68$$

$$y_4 = \frac{1}{1+e^{-0.68}}$$
$$= 0.6637$$

$$a_3 = (W_{35} \times Y_3) + (W_{45} \times Y_4)$$
$$= (0.3 \times 0.68) + (0.9 \times 0.6637)$$
$$= 0.801$$

$$Y_5 = \frac{1}{1+e^{-x}} = \frac{1}{1+e^{-0.801}} = 0.69 \text{ [output]}$$

Error $= Y_{target} - Y_{pred} = (0.5 - 0.69) = -0.19$

Subject: Date:

Loss function:

$$L = \frac{\left(\sum_{i=1}^{3} -y_i \log_{ai} a_i - (1-y_i)\log(1-a_i)\right)}{3}$$

X	Y
0.1	0
0.2	0
0.3	1
0.4	

Assume that,

$W = 0.7$ and $b = 0.1$

$z = 0.7x + 0.1$ $[y = mx + b]$

$$L = -0\log\left(\frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right) - (1-0)\log\left(1 - \frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right)$$

$$- \boxed{} - \boxed{}$$

$$= -\log\left(1 - \frac{1}{1+e^{-(0.7\times0.1+0.1)}}\right) - \log\left(1 - \frac{1}{1+e^{-(0.7\times0.2+0.1)}}\right)$$

$$- \log\left(\frac{1}{1+e^{-(0.7\times0.3+0.1)}}\right)$$

$$= -\log(1-0.54) - \log(1-0.55) - \log 0.57$$

$$= 0.928$$

Avg loss value $= \left(\frac{0.928}{3}\right) = 0.309$

Subject : Date :

$$\frac{dL}{dW} = \frac{\left(\sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i\right) x_i\right)}{3}$$

$$W = W - \alpha \frac{dL}{dW}$$

$$\frac{dL}{db} = \frac{\left(\sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i\right)\right)}{3}$$

$$b = b - \alpha \frac{dL}{db}$$

$$\frac{dL}{dW} = \left(\left(\frac{1}{1+e^{-(0.7\times 0.1+0.1)}} - 0\right)0.1 + \left(\frac{1}{1+e^{-(0.7\times 0.2+0.1)}} - 0\right)0.2\right.$$

$$\left. + \left(\frac{1}{1+e^{-(0.7\times 0.3+0.1)}} - 1\right)0.3\right) / 3$$

$$= \frac{(0.054 + 0.112 - 0.127)}{3}$$

$$= 0.013$$

$$\frac{\partial L}{\partial b} = \left(\left(\frac{1}{1+e^{-(0.7\times 0.1+0.1)}} - 0\right) + \left(\frac{1}{1+e^{-(0.7\times 0.2+0.1)}} - 0\right)\right.$$

$$\left. + \left(\frac{1}{1+e^{-(0.7\times 0.3+0.1)}} - 0\right)\right) / 3$$

Subject: Date:

$$= \frac{(0.54 + 0.56 - 0.42)}{3}$$

$$= 0.23$$

if learning rate, $\alpha = 1$.

updated weight, $w = w - \alpha \frac{dL}{dw}$

$$= 0.7 - 1(0.013)$$

$$= 0.687$$

updated bias, $b = b - \alpha \frac{dL}{db}$

$$= 0.1 - 1(0.23)$$

$$= -0.13$$

Old $z = 0.7x + 0.1$

new $z = 0.687x - 0.13$

L for new z will be, $L = 0.86$

Avg Loss value $= \frac{0.86}{3}$

$= 0.2866$ [new]

Avg Loss value $= 0.309$ [previous]

Conclusion error/Loss slightly reduce from 0.309 to 0.2866.

Loss function:

Logistic Regression:

$$\frac{\partial Loss}{\partial W_j} = (y_{pred} - y_{act}) \cdot x_j$$

$$\frac{\partial Loss}{\partial b} = (y_{pred} - y_{act})$$

$$\frac{\partial L}{\partial W} = \sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i \right) x_i$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n} \left(\frac{1}{1+e^{-x}} - y_i \right)$$

Update weight:

$$W = W - \alpha \frac{d Loss}{dW}$$

Update bias:

$$W = W - \alpha \frac{d Loss}{db}$$

Sigmoid function: $\frac{1}{1+e^{-x}}$

$$\frac{d}{dx}\left(\frac{1}{1+e^{-x}} \right) = \frac{e^{-x}}{1+e^{-x}} \rightarrow \frac{1}{1+e^{-x}} \cdot e^{-x}$$

$$\sigma(x) \quad (1 - \sigma(x))$$

or, $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$

Subject: Date:

Gradient Descent Loss function:

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - (mx+c))^2$$

$$\frac{\partial L}{\partial m} = -\frac{2}{n}\sum x_i(y_{actual} - y_{pred})$$

$$\frac{\partial L}{\partial c} = -\frac{2}{n}\sum(y_{actual} - y_{pred})$$

Actual Data Table

→ Create an equation and predict the output

→ Calculate the loss function

→ Update w an bias

→ $w = w - \alpha \frac{dL}{dw}$

$b = b - \alpha \frac{dL}{db}$

MES/SSR:

Actual Table

→ Prediction

→ Find Loss $\frac{\sum(\text{actual} - \text{predict})^2}{n}$

→ .

Subject: Date:

$$E(\text{Decision}) = -P(\text{yes})\log_2 P(\text{yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right)$$

$$= 0.940.$$

Student:

$$E(\text{student} \mid \text{yes}) =$$

$$E(\text{Decision}) = -P(\text{yes})\log_2(\text{yes}) - P(\text{No})\log_2 P(\text{No})$$

$$= -\frac{4}{10}\log_2\left(\frac{4}{10}\right) - \frac{6}{10}\log_2\left(\frac{6}{10}\right)$$

$$= 0.971$$

x_1

$$E(\text{yes} \mid a_1) = -\frac{4}{5}\log_2\left(\frac{4}{5}\right) - \left(\frac{1}{5}\right)\log_2\left(\frac{1}{5}\right) \cancel{- \frac{5}{5}\log_2(\ )}$$

$$\cancel{E(\text{No} \mid a_1) = -\frac{5}{5}}$$

Subject : Date :

$$E(D) = -P(Y)\log_2 P(Y) - P(N)\log_2 P(N)$$

$$= -\frac{4}{10}\log_2\left(\frac{4}{10}\right) - \frac{6}{10}\log_2\left(\frac{6}{10}\right)$$

$$= 0.971$$

X1:

$$\underset{a_1}{E(X1)} = -P(Y|X_1)\log_2 P(Y|X_1) - P(N|X_1)\log_2 P(N|X_1)$$

$$= -\frac{4}{5}\log_2\left(\frac{4}{5}\right) - \frac{1}{5}\log_2\left(\frac{1}{5}\right) = 0.721$$

$$E(a_2) = -P(Y|a_2)\log_2 P(Y|a_2) - P(N|a_2)\log_2(N|a_2)$$

$$= -\frac{0}{5}\log_2\left(\frac{0}{5}\right) - \frac{5}{5}\log_2\left(\frac{5}{5}\right) = \cancel{0.361}\ 0$$

$$IG = 0.971 - 0.721\, P(a_1) - 0.361\, P(a_2)$$

$$= 0.971 - 0.721\ \frac{5}{10} - \cancel{0.361}\ \frac{5}{10}\ 0$$

$$= 0.610$$