

Roll: 18204016

CSE-6119

Ans! to the Ques Q2

~~Q2(a)~~

We can tell the right numbers of clusters using

Elbow method.

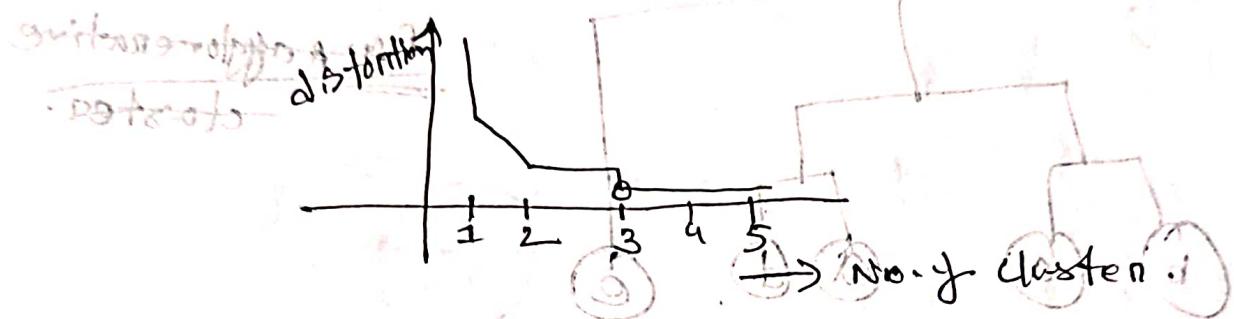


Fig: Elbow finder method.

o'clock environmental DPA 18/3

In Elbow finder method we take no. of cluster as when the ~~first~~ Elbow is found & from where the remaining no. of cluster gives no ~~other~~ elbow Elbow & it gradually reaches in total 0's.

In the fig we can see no. 3 is the no. of cluster we can use. If in first and 2<sup>nd</sup> of 3<sup>rd</sup> elbow

Agglomerative cluster.



The phidiasian planesh 18/3 and in the next  
we have to turn to a meeting

6/11/2020

	Age	Income
a	27	90
b	29	90
c	29	61
d	28	60
e	92	115

6/10/2020

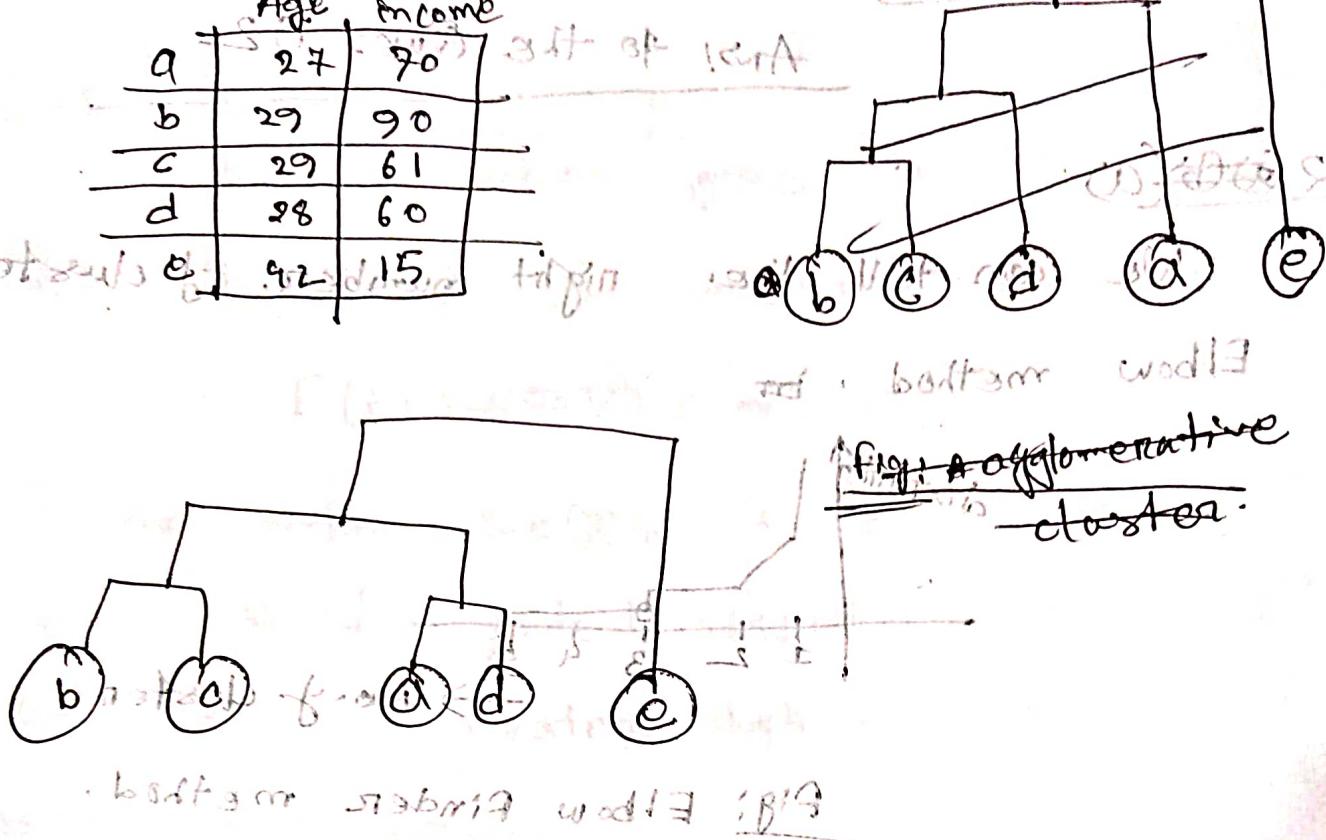
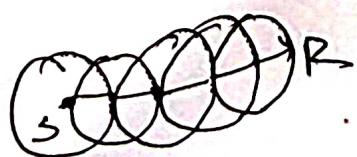


Fig: Agglomerative clustering.

2(b) density reachability: (density reachability is)

called if we can reach through a core point 's' to 'r' by some point in it.



bottom up

Fig: density reachability.

## ii) density connectivity:

If two points 'P' & 'Q' connects to a point 'm' through the border of point 'm' from both 'P' & 'Q' then this is called density connectivity.

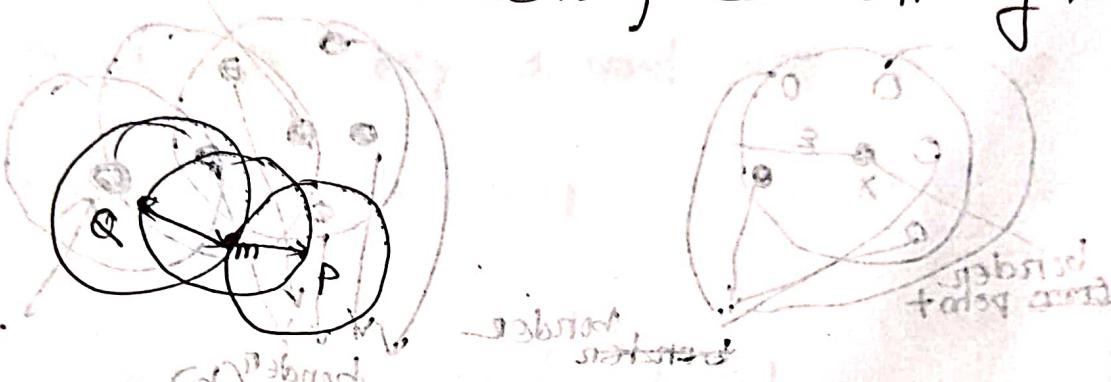


Fig: density connectivity (ii)

identify core point, border point, noise point —

1. core point: Core point is a point from where the <sup>no. of</sup> 'eps'-neighbor found such that the no. of 'eps'-neighbor is equal or greater than minpts.

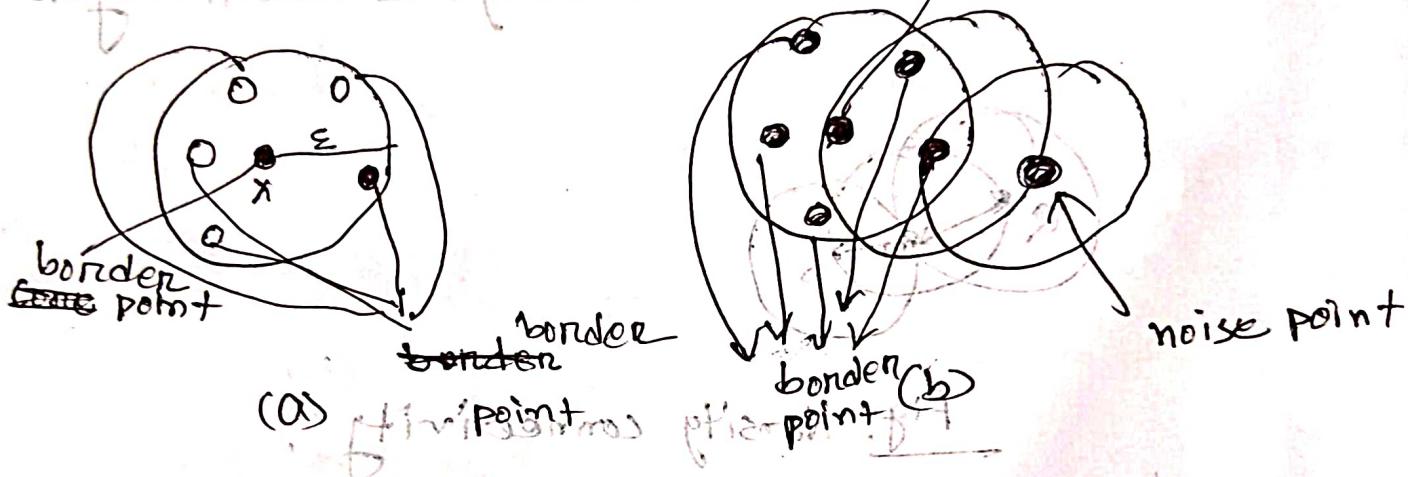
2. border point: Border point is a point from where the no. of 'eps'-neighbor found i.e., the no. of 'eps'-neighbor is less than the minpts.

3. noise point: A point that is not either

boundary point or border point is known as

~~border point~~ noise point (in English)

Given, Min Pts = 6, ~~elements~~ ball & ~~not~~ melt



## Ans! to Ques! No: 01

1(b)

(a) b

start with activities for model

models cause over fitting in dt!

when decision tree generates decision  
model with only 1 leaf node.

to stop it we can do pruning of decision tree

Avoid overfitting!

1. define no. of tree depth.

2. define [no. of features sample] that

will be divided further.

3. define no. of min sample to split.

effecient

pruning of

Tree based model works on both categorical

and

regression data when relation is between  
data has higher information gain, lower entropy.

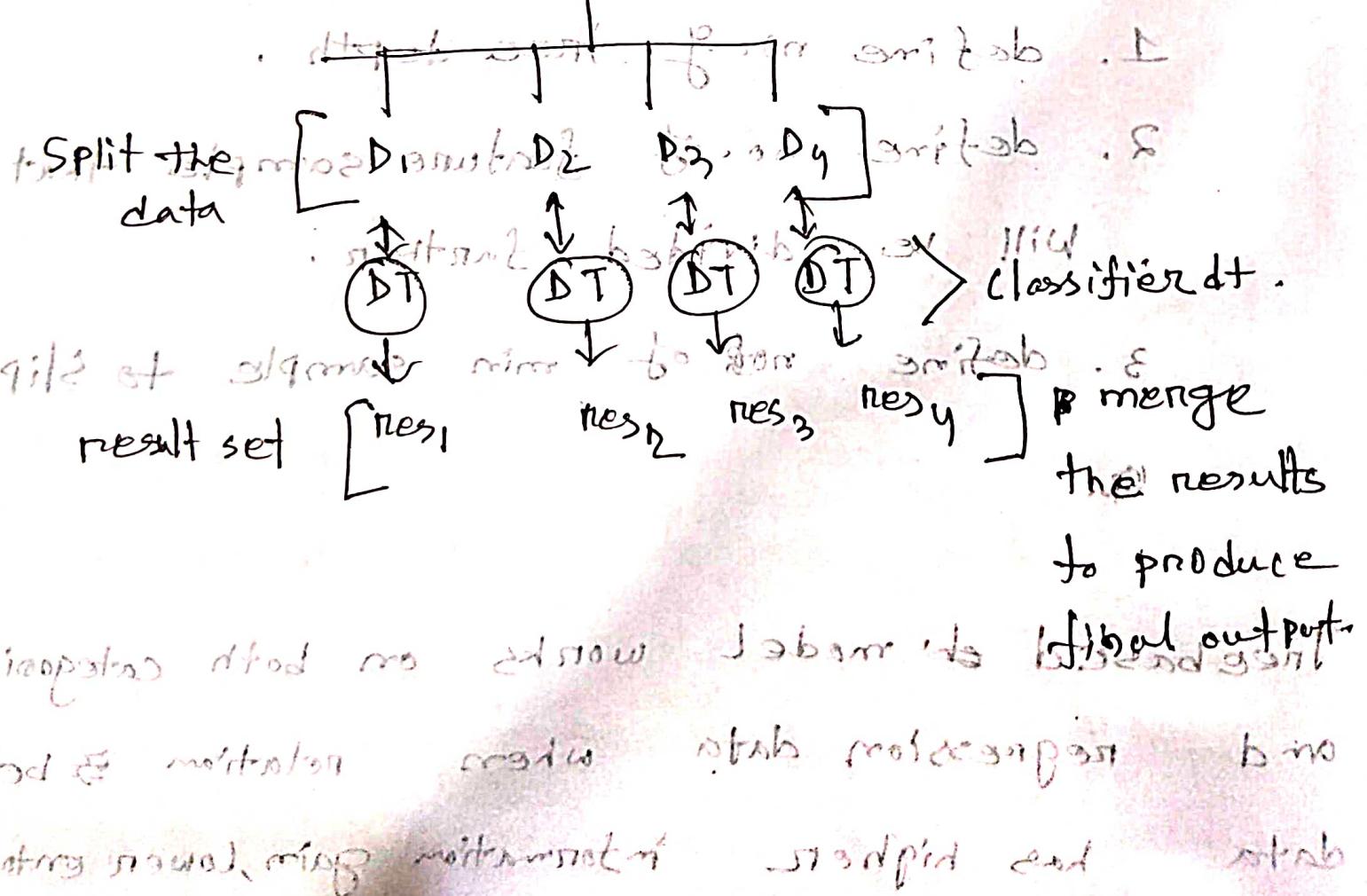
Q1 : Q1 result of QnA

1(a)

Random forest divides data into (d) 1  
 subset of data then apply the decision  
 tree based algorithm on each partition  
 of the data and finally merge  
 the output to produce model output.

Random Forest

with bias



A	B	C	Y
0	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

Hence, 1 = T  
0 = F.

Gini score calculate as  $(P^V + Q^V)$ .

Hence,  
 $P$  = probability of successfull  
 $Q$  = probability of failure.

Roll No. 18204016

Code - CSE-6110

Date: 9.8.2021

Ans: to the Ques. No. 0.4

4(a)

pooling on convolution Neural Network:

pooling in CNN contains few steps

1. define window size. (usually  $(2 \times 2) / (3 \times 3)$ )

2. define window stride (usually 2)

3. the window will pass through out the image by the defined stride and take the maximum value every time from the window.

4. Taking the maximum value from the window it build a new smaller image compared to input image.

pooling uses CNN to pool out the maximum value from the defined window size in the input image.

• transformation will repeat again for global max.

• This program can not handle the S

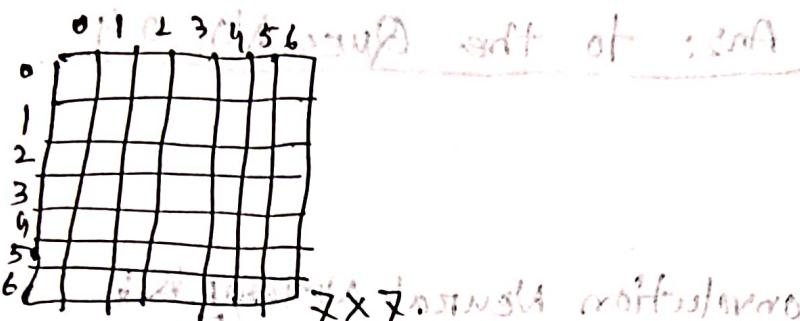
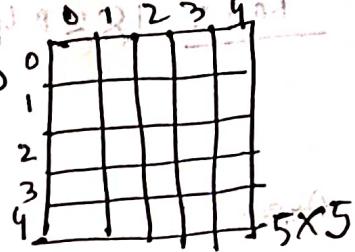
Input

0 1 2 - 3 2 5 - 3 6 8 0

2 1 0 9 8 0 1 2 3 4 9

In CNN, if input size  $3 \times 5$  that is

& if the filter size  $3 \times 3$  that is,



input

(6) P

$3 \times 3$  kernel convolution on profit

filter note with erosion using in profit

$(6 \times 6) \times (6 \times 6)$  pixels) will become 1 unit & then the output size will be the size of input

image that is  $3 \times 3$  because the input image size is

less than the filter.

• maximum diff more sum more (yellow)

addition diff more sum minimum diff greatest up

4(b)  $\rightarrow$  sum of all squares war is Hadamard product

The difference between traditional Neural Network

& Recurrent Neural Network (RNN) are written

below the two loops are also below

→ Traditional Neural Network has more inputs & more

1. Inputs of every layer are independent here.

2. It does not have any memory cell.

3. As in (NN) it doesn't have any memory cell

It doesn't ~~contains~~ stores output, ~~every~~ every

day.

storing lot of words in text environment ) off course

4. Example: CNN, perceptron, back propagation

at 1/3rd program is end . Johnson 2011-12 UG

at first does next years work together 2012  
Recurrent Neural Network (RNN)

1. Inputs of every layer are not independent as traditional  
Neural Network. Here what stored from one

2. It has memory cell to store output for next  
time use.

3. As it can stores the output for next time

input it works very well in text generation.

Ex: RNN, LSTM, etc. mostly in NLP

• PGM stores no storing in RNN . And

• It stores in RNN . PGM good for image

• text etc

1133. program you said forget it. (UN)il eA.E  
RNN works better in text data.

In text, data analysing it is important to  
know the previous text or data to generate  
next data or text. ~~next happening and continuing~~ RNN ; ~~glanced~~.

RNN ~~works~~ model has memory cell to  
store output from every layer such that the  
next time it can be used as the input.

So In text data ~~RNN~~ unlike in CNN or traditional  
neural Network of RNN works better if .  
~~not~~

Ex: 1. "I have bought a \_\_\_\_". See what  
will come first after ~~it~~ context and if eA.E  
will answer first in the blank space ~~now~~ if ~~twice~~  
Here in the blank space RNN can generate  
pen or book or other ~~the~~ things related to  
buy. If it predicts or generate play.  
we can not buy play. so it is irrelevant to  
the context.

Ans. to the Ques NO: 03

3(a) If the learning rate is too low then the model could take longer time to train on the given trained dataset.  
It will be worst if it comes in a large dataset.

If the learning rate is too ~~low~~<sup>high</sup> then the model could overlook the optimum Error point hence it will give bad performance on the test data.

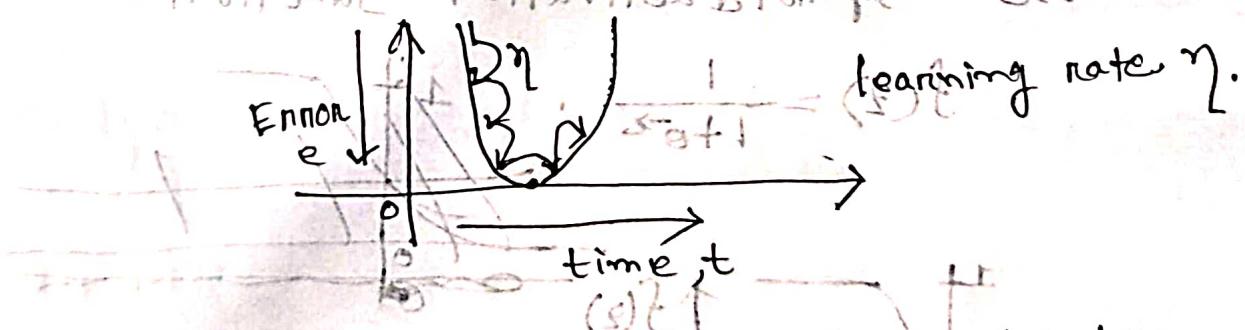
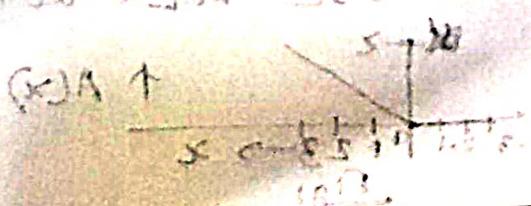
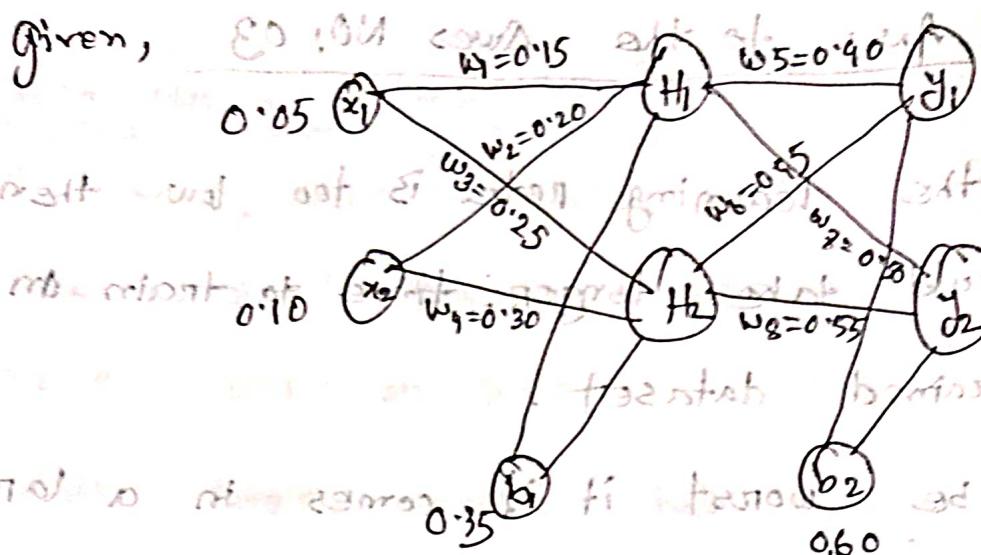


Fig: when learning rate is too high.

When learning rate is too high, the error does not decrease and the steps might not



(x)A + s = 0, constant



Target value:

$$\begin{cases} T_1 = 0.01 \\ T_2 = 0.99 \end{cases}$$

learning rate  $\eta = 0.25$ .

### i) Activation function:

As the output is from 0 to 1, so we should use sigmoid activation function.

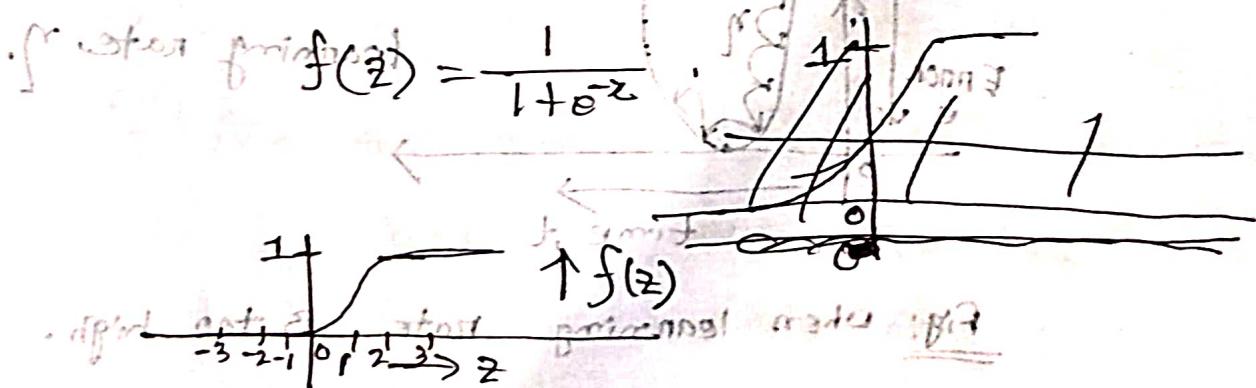
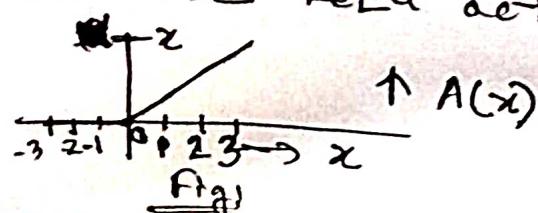


Fig: Sigmoid Activation function.

for hidden layer we can use ReLU activation function.  $A(x) = \max(0, x)$

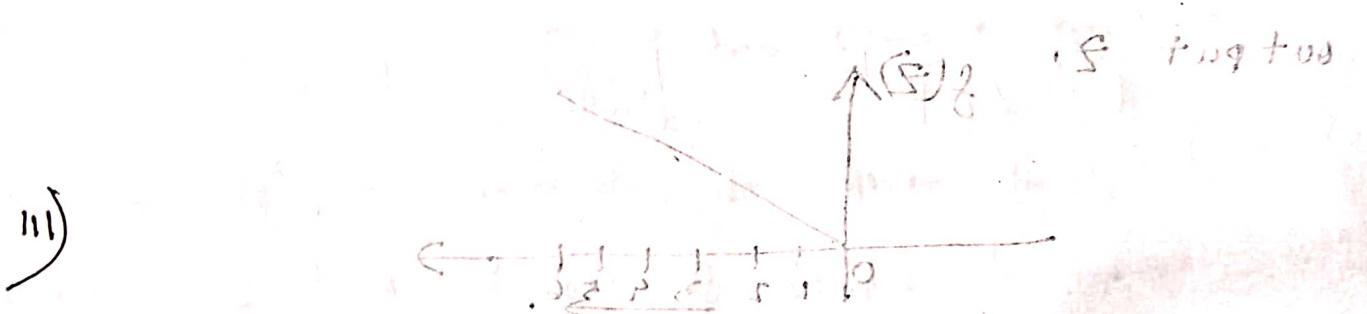


As the ReLU function gives the max value from  $\max(0, x)$ . It could help to get better result.

iii) Calculation of Error :

$$E = \frac{[\text{Actual output} - \text{Expected output}]}{\text{Expected output}}$$

$$E = \sqrt{\sum (\text{Actual output} - \text{Expected output})^2}$$



iii)

3(b)

ReLU Activation! Re ( $\text{ReLU}(z)$ )

from  $(0 \text{ to } \infty)$ , it defines as —

$$f(z) = \begin{cases} 0 & z \leq 0 \\ \text{Max}(0, z) & z > 0 \end{cases}$$

~~if the given value ( $z$ ) to it is less or equal to ( $0$ ) then it will give output =  $0$ ; if the given value ( $z$ ) is greater than  $0$ , it will give output  $z$ .~~

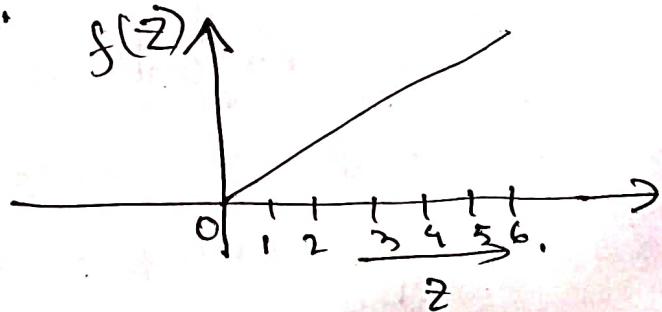


Fig 1 ReLU Activation function.