

FEBRUARY 2020

	1	2
3	4	5
6	7	8
9	10	11
12	13	14
15	16	17
18	19	20
21	22	23
24	25	26
27	28	29
M	T	W
T	F	S
S		

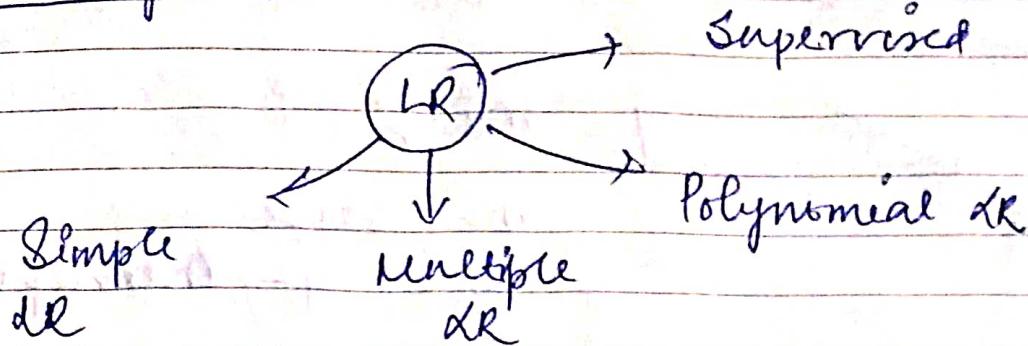
MONDAY  
JANUARY | 2020

27

WK 09 | 22-339

## → Machine Learning Algorithms :-

### i) Linear Regression :-



Simple → 1 i/p | 1 o/p

Suppose,  
cgpa | package  
6.66 | 3.01.

Multiple dl → One or  
more i/p | 1 o/p.

Cgpa | gender | 12<sup>th</sup> Marks | state | package

Polynomial ? - Suppose if your dataset is  
not ~~ps~~ linear, then we use  
polynomial Regression.

$x^1 x^2 \dots x^n \rightarrow$  Completely linear

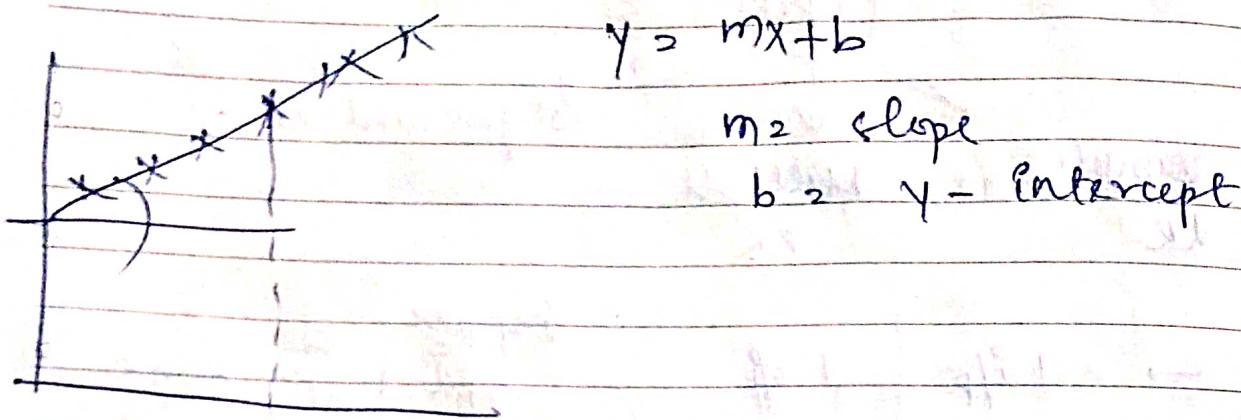
$\begin{matrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{matrix} \rightarrow$  Sort of linear

28

TUESDAY  
2020 | JANUARY

1	2	3	4	5
8	9	10	11	12
15	16	17	18	19
22	23	24	25	26
29	30	31		

- Why datasets are not completely linear or sort of linear.
- Real world datasets are not linear as they are affected by various factors. (these are the stochastic errors).



To handle the sort of linear data we'll draw a ~~sort of~~ best fit line.

→ Basically, finds the value of  $m$  &  $b$ , so that it makes less error.

$b$  = offset, doesn't make the value to zero.

$(m, b)$

Closed  
form  
soln

Non-Closed  
form soln.

OLS

Scikit learn

uses OLS to find the value of  $m$  &  $b$ .

FEBRUARY 2020

	1	2
3	4	5
10	11	12
17	18	19
24	25	26
27	28	29
M	T	W
T	F	S
S		

WEDNESDAY  
JANUARY | 2020

29

WK 05 | 029-337

1) Closed form soln :- If we can create any formula & find the corresponding value of the variable, then that is closed form soln.

for ex.

$$ax^2 + bx + c = 0$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Ex: OLS

→ kind of direct formula

2) Non-closed form solns :- Approximation based solns.

Ex: Gradient descent

→ So, if there are less dimensions, we'll use, OLS (direct formula method / closed form soln), if we are working on higher dimensions, we'll use gradient descent.

→ Scikit Learn's Linear Regression uses OLS internally to find the values of m & b.

→ SGD Regressor uses, GD Gradient descent internally.

30

THURSDAY  
2020 | JANUARY

020-335 | WR CS

DECEMBER	
1	2
8	9
15	16
22	23
29	30
S	M

FEBRUARY	
3	4
10	11
17	18
24	25
M	

OLS :-

$$\therefore b = \bar{y} - m\bar{x}$$

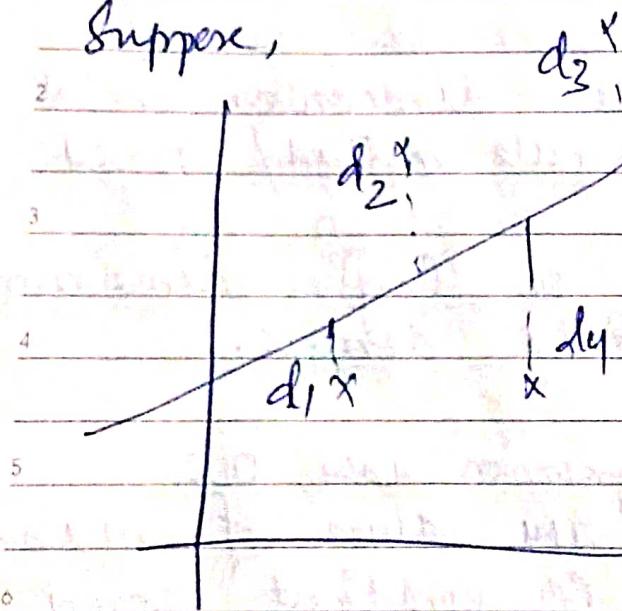
 $y = \text{package}$  $x = \text{cgpa}$ .

where,

$$m = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$\bar{x}$   $\nearrow$  mean values  
 $\bar{y}$  of  $x$  and  $y$ .

Suppose,



$$E = d_1 + d_2 + d_3 + \dots + d_n$$

$$E = d_1^2 + d_2^2 + d_3^2 + \dots$$

$$E = \sum_{i=1}^n d_i^2 + d_n^2$$

→ The idea of squaring is due to:

- To make the values positive/decision points

we don't use modulus as

R → modulus is not differentiable at origin.

→ we need to penalise the outliers to make a better model.

FEBRUARY 2020

				1	2
3	4	5	6	7	8
10	11	12	13	14	15
17	18	19	20	21	22
24	25	26	27	28	29

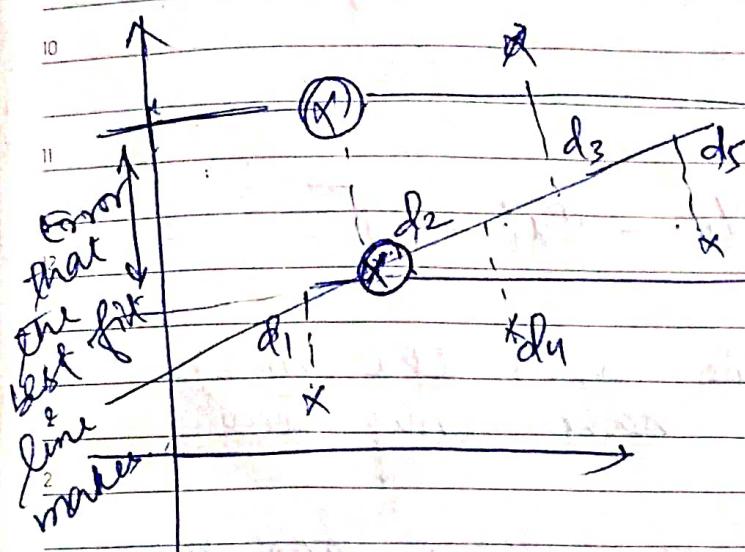
E OR S

$$E = \sum_{i=1}^n d_i^2$$

FRIDAY  
JANUARY | 2020

31

WK 05 | 03-335

→ Error / function  
Loss $(m, b) \rightarrow$ gives the value of  $m$  &  $b$ , that minimizes the value of the eqn.

but it is here  
& actual point  
should be here  
according to best  
fit line.

 $y_i^{\circ} \rightarrow$  prediction value

$$\text{So, } d_i = (y_i^{\circ} - \hat{y}_i)$$

(actual - predicted)

Thus, error formula becomes,

Total  
Error

$$E = \sum_{i=1}^n (y_i^{\circ} - \hat{y}_i)^2$$

we need  
to find  
such a valueor, we need to find such a line that gives  
that minimizes the value of the min<sup>m</sup> value  
the equation.

01

SATURDAY  
2020 | FEBRUARY

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
S	M	T	W	T	F	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S		

032-334 | W# 05

$$\text{Total Error} \rightarrow E = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\text{Avg. Total Error} \rightarrow E = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y}_i = mx_i + b$$

$$E(m, b) = \sum_{i=1}^n (y_i - mx_i - b)^2$$

The function depends upon the  $(m, b)$  values, as, we have only two options.

- 1) Either we fix the slope and move the intercept.
- 2) or fix the intercept & ~~move~~ change the value.

02

SUNDAY) To get the minimized value, for the loss function.

$y = f(x) \rightarrow$  &  $y$  is a function of  $x$ .

MARCH 2020

30	31	1	2	3	4	5	6	7	8
9	10	11	12	13	14	15			
16	17	18	19	20	21	22			
23	24	25	26	27	28	29			
M	T	W	T	F	S	S			

MONDAY  
FEBRUARY | 2020

WK 06 | 034-332

03

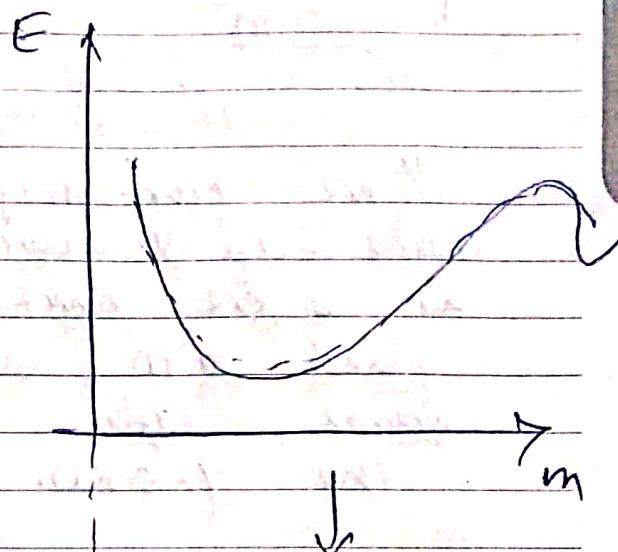
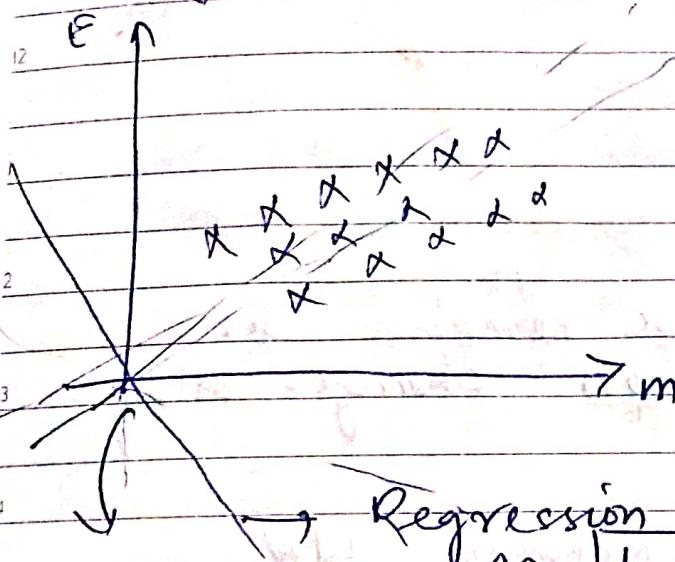
Suppose,  $b = 0$ , then the error function

$$E(m, b) = \sum_{i=1}^n (y_i^o - mx_i^o - b)^2$$

becomes



$$E(m) = \sum_{i=1}^n (y_i^e - mx_i^o)^2$$



Rotates along the origin.

after we rotate the best fit line at  $b = 0$  along origin.

Now, let's if we make,  $m = \text{constant}$ .

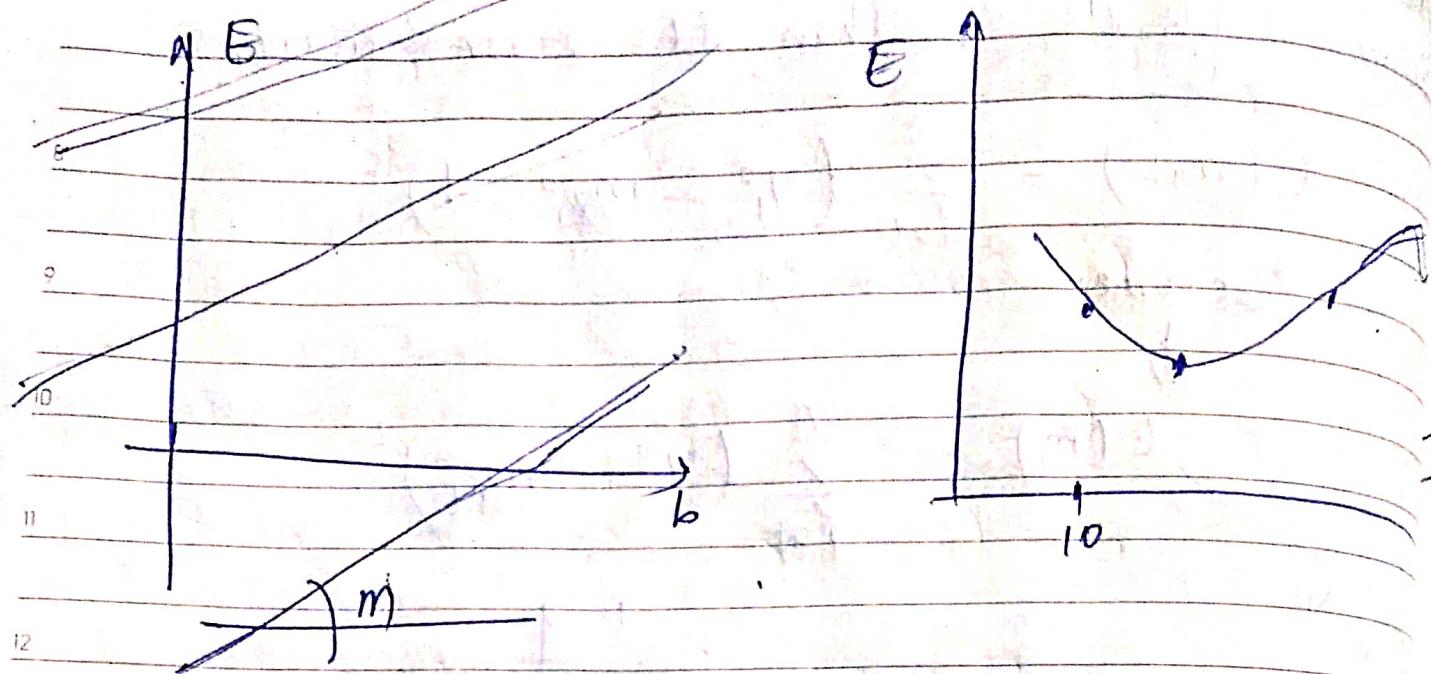
$$E(b) = \sum_{i=1}^n (y_i^o - x_i^o - b)^2$$

04

TUESDAY  
2020 | FEBRUARY

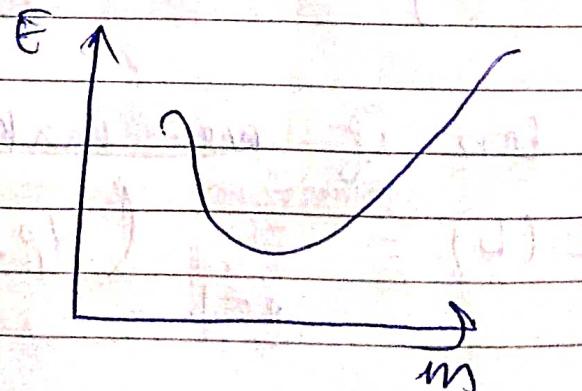
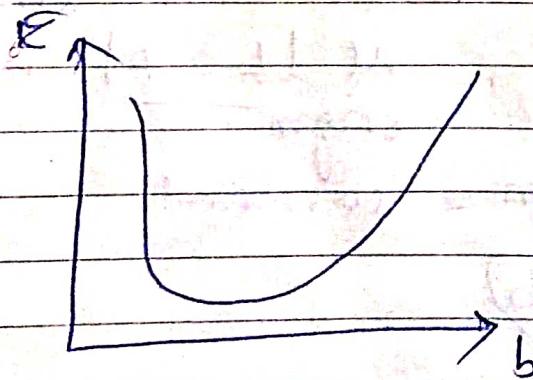
035-331 | WK 06

5	6
12	13
19	20
26	27
S	M



↓ we can only move up & down as we have fixed ' $b$ ' & ' $m$ ', so, slope remains the same, only intercepts changes along the  $y$ -axis.

So, roughly, the relation b/w  $E \propto b$  &  $E \propto m$  is:



MARCH 2020

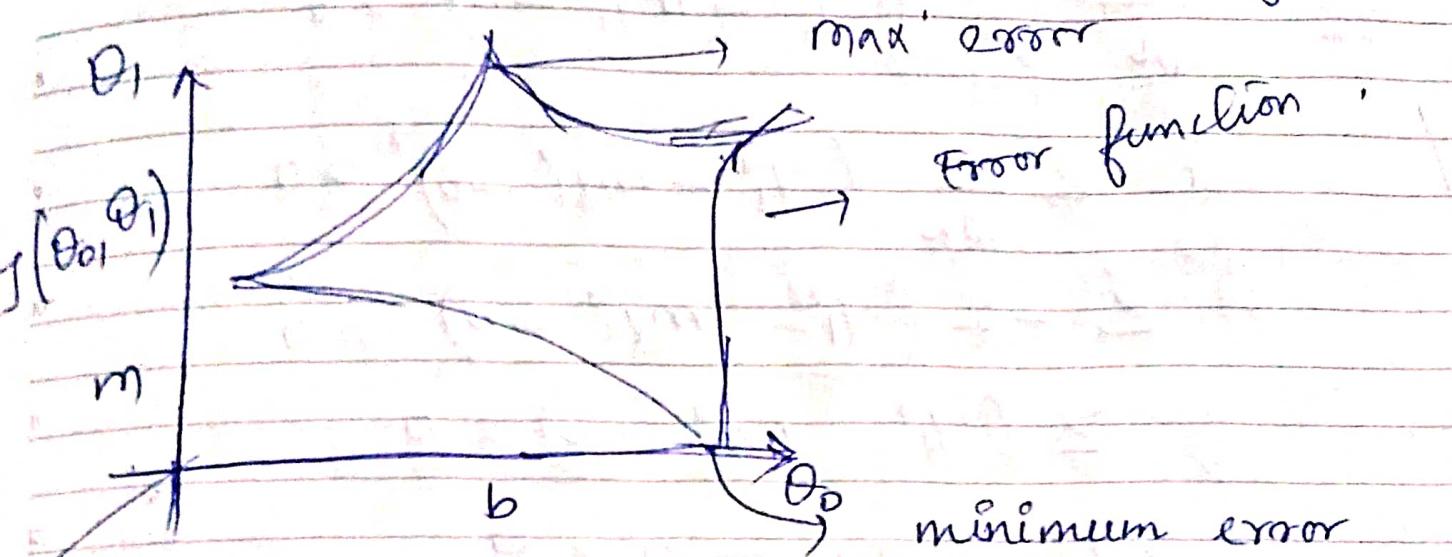
1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	
16	17	18	19	20	21	22	
23	24	25	26	27	28	29	
30							

WEDNESDAY  
FEBRUARY | 2020

05

WK.06 | 036-320

Let's plot a 3-D plot, including the two.



To find the value of maxima & minima, we use,  $\Rightarrow$  differentiation, of maxima & minima.

minima  $\Rightarrow 0$ , (we know)

So, we will calculate the derivative for maxima & minima for  $E(m) \& E(b)$ .

$$\frac{dE}{dx} = 0$$

$$f(x,y) \rightarrow \frac{\partial E}{\partial m} = 0, \frac{\partial E}{\partial b} = 0$$

we get two eqns,  
& hence calculate (m, b).

MARCH

APRIL

06

THURSDAY  
2020 | FEBRUARY

037-329 | WK 06

5
12
19
26
3

$$\frac{\partial E}{\partial b} = \frac{\partial}{\partial b} \sum_{i=1}^n (y_i - mx_i - b)^2 \geq 0$$

$$\Rightarrow \sum \frac{\partial}{\partial b} (y_i - mx_i - b)^2 \geq 0$$

$$\Rightarrow \sum -2(y_i - mx_i - b) = 0$$

$$\Rightarrow \sum (y_i - mx_i - b) = 0$$

$$\Rightarrow \sum y_i - \sum mx_i - \sum b = 0$$

$$\Rightarrow \frac{\sum y_i}{n} - \frac{\sum mx_i}{n} - \frac{\sum b}{n} = 0$$

( $\because n$  is the no. of students  
on samples)

$$\Rightarrow \bar{Y} - \bar{mx} - \frac{nb}{n} = 0$$

$$\Rightarrow \bar{Y} - \bar{mx} = b$$

$$\Rightarrow b = \bar{Y} - \bar{mx}$$

MARCH 2020

						1
30	31	4	5	6	7	8
2	3	11	12	13	14	15
9	10	18	19	20	21	22
16	17	24	25	26	27	28
23	24	31				29
M	T	W	T	F	S	S

FRIDAY  
FEBRUARY | 2020

07

WK 06 | 03B-32B

fitting the terms in  $b$  in the loss/error eqn —

$$E = \sum (y_i^o - mx_i^o - \bar{y} + m\bar{x})^2$$

$$\frac{\partial E}{\partial m} = \sum \frac{\partial}{\partial m} (y_i^o - mx_i^o - \bar{y} + m\bar{x})^2 = 0$$

$$\Rightarrow \sum 2(y_i^o - mx_i^o - \bar{y} + m\bar{x})(-x_i^o + \bar{x}) = 0$$

$$\Rightarrow \sum -2(y_i^o - mx_i^o - \bar{y} + m\bar{x})(x_i^o - \bar{x}) = 0$$

$$\Rightarrow \sum (y_i^o - mx_i^o - \bar{y} + m\bar{x})(x_i^o - \bar{x}) = 0$$

$$\Rightarrow \sum [(y_i^o - \bar{y}) - m(x_i^o - \bar{x})](x_i^o - \bar{x}) = 0$$

$$\Rightarrow \sum [(y_i^o - \bar{y})(x_i^o - \bar{x}) - m(x_i^o - \bar{x})^2] = 0$$

$$\Rightarrow \sum (y_i^o - \bar{y})(x_i^o - \bar{x}) = m(x_i^o - \bar{x})^2 = 0$$

$$\Rightarrow \sum (y_i^o - \bar{y})(x_i^o - \bar{x}) = m \sum (x_i^o - \bar{x})^2$$

$$m = \frac{\sum (x_i^o - \bar{x})(y_i^o - \bar{y})}{\sum (x_i^o - \bar{x})^2} = \frac{\sum (x_i^o - \bar{x})(y_i^o - \bar{y})}{\sum (x_i^o - \bar{x})^2}$$

$$m = \frac{\sum_{i=1}^n (x_i^o - \bar{x})(y_i^o - \bar{y})}{\sum_{i=1}^n (x_i^o - \bar{x})^2}$$

MARCH

APRIL

08

SATURDAY  
2020 | FEBRUARY

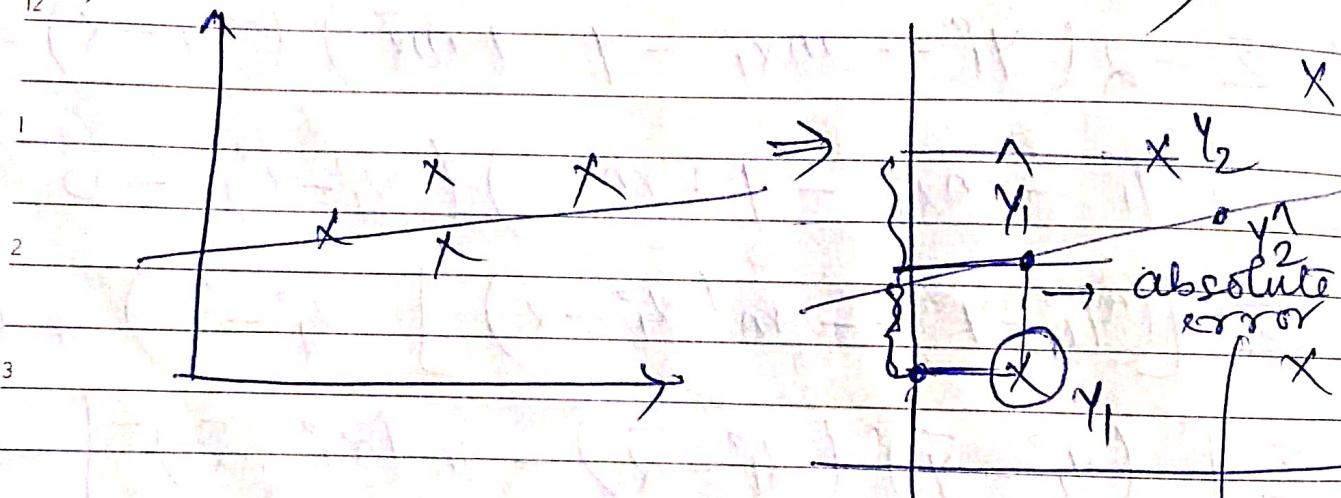
039-327 | WK 06

5	6	7
12	13	14
19	20	21
26	27	28
S	M	T

## Regression Metrics :-

- 8) 1) MAE
- 9) 2) MSE
- 10) 3) RMSE
- 11) 4) R<sup>2</sup> Score
- 12) 5) Adjusted Score.

11) MAE :— (Mean absolute Error)



09

SUNDAY

$$|y_1 - \hat{y}_1| + |y_2 - \hat{y}_2| + \dots + |y_n - \hat{y}_n|$$

W/F Consider  
the sign.

$$+$$

$$\text{mae} = \sum_{i=1}^n |y_i - \hat{y}_i|$$

MARCH 2020						
SUN	MON	TUE	WED	THU	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				
M	T	W	T	F	S	S

MONDAY  
FEBRUARY | 2020

10

WK 07 | 041-325

### Advantage :-

- MAE should have min error.

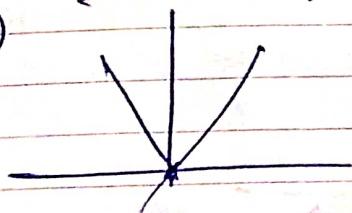
mae up = model up  
unit unit

same unit.

- Robust to outliers (Can handle outliers easily)

### Disadvantage :-

- Not differentiable at zero.



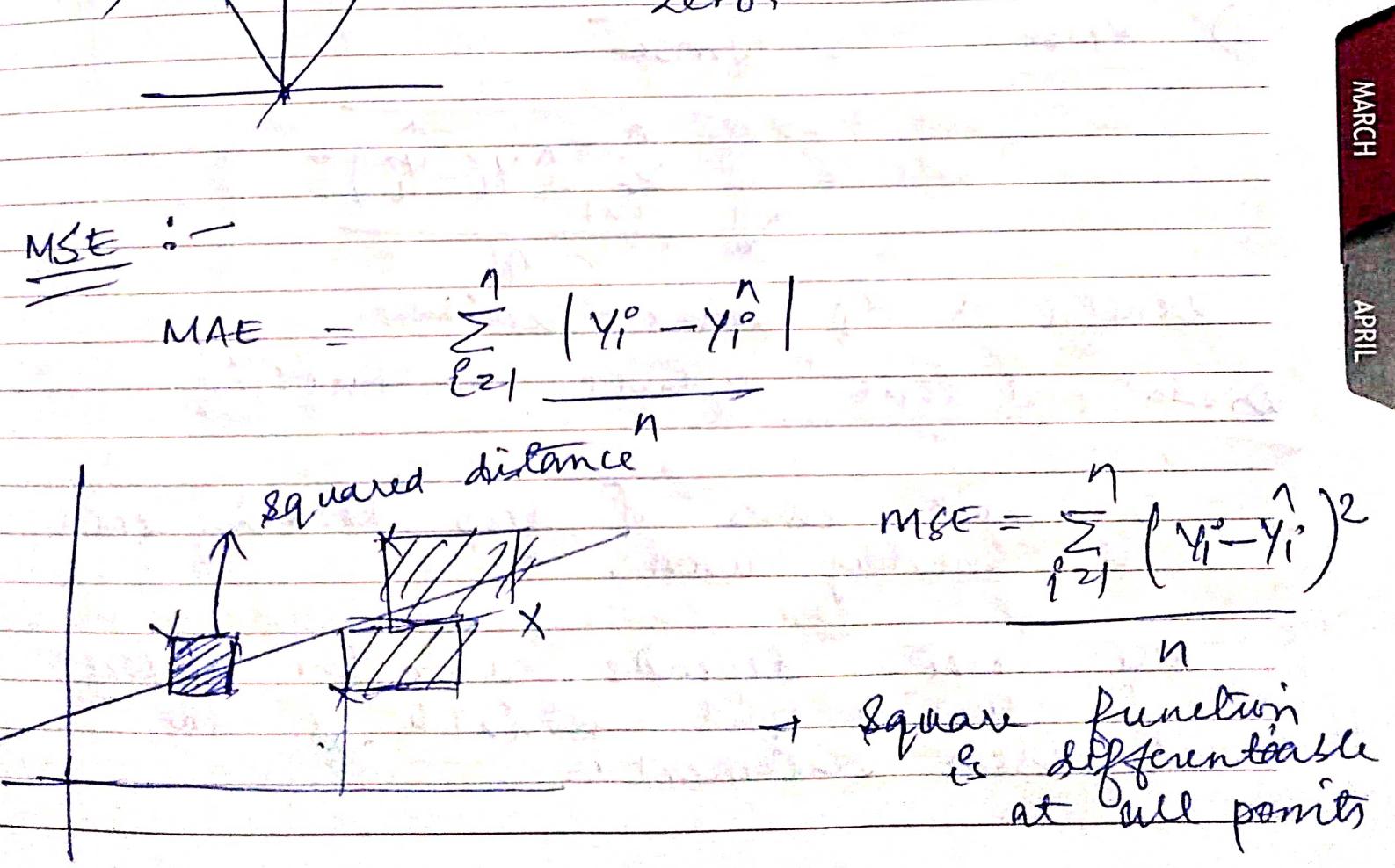
### MSE :-

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

squared distance

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

→ Square function is differentiable at all points



11

TUESDAY  
2020 | FEBRUARY

042-324 | WK 07

5
12
13
20
21
22
23
24
25
26
27
28
29
30
31

The square represents  $\rightarrow (y_i - \hat{y})^2$

→ Advantage :-

1) The loss function is differentiable.

$$y - (\text{lpa}) \rightarrow y_2 \rightarrow (\text{lpa})^2$$

2) Penalize the outliers too much.

N of Robust to outliers, as the square will be too much.

(2) RMSE:  $\underline{\underline{= \sqrt{MSE}}}$

$$\underline{\underline{= \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}}$$

Benefit : 1) Same unit

Disadvantage that Robust to outliers

In most cases of deep learning RMSE is mostly used.

MSE, MAE, depends upon the context of the output variable & the problem statement.