Line Equation :- $y = \underbrace{m}_{\text{Slope}}x + \overset{\text{Intercept}}{c}$

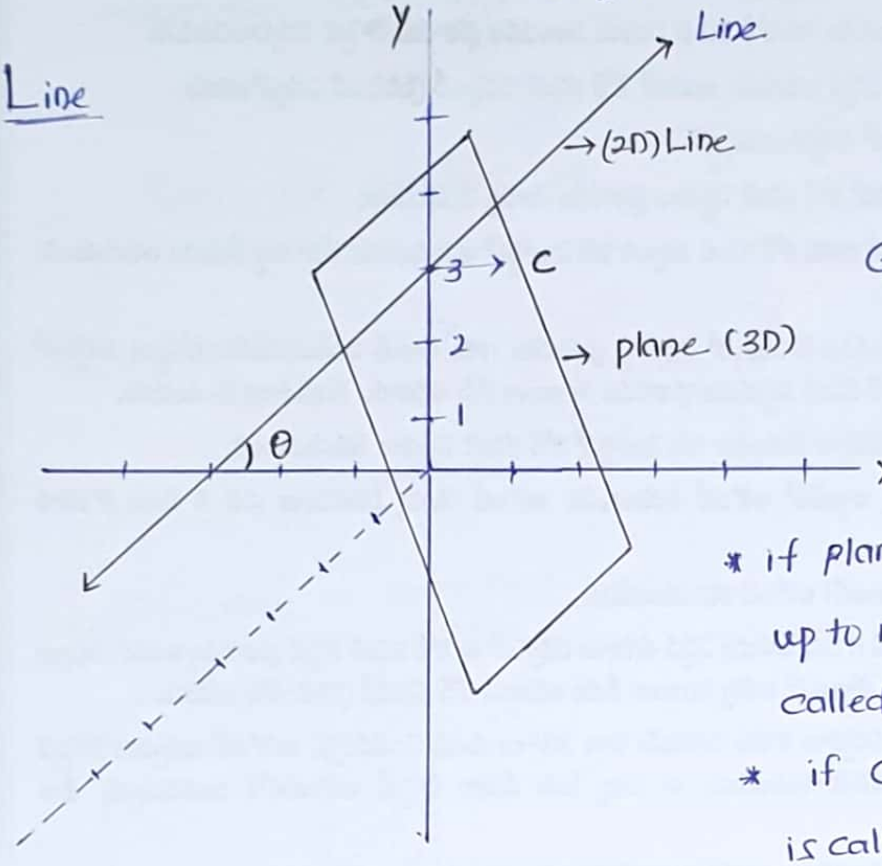
DOT PRODUCT :- $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

$$\vec{a}\cdot\vec{b} = \|\vec{a}\| \; \|\vec{b}\| \cdot \cos\theta_{a,b} \quad (or)$$

$$\vec{a}\cdot\vec{b} = a_1b_1 + a_2b_2 = \sum_{i=1}^{d} a_ib_i \quad (or)$$

$$\vec{a}\cdot\vec{b} = [a_1 a_2] \times \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \vec{a}^T \times \vec{b}$$

Line



c → Intercept at y axis

m → Slope with x axis

m = tanθ

General line eqn = $ax + by + c = 0$

$$by = -ax - c$$

$$y = \underbrace{\left[\frac{-a}{b}\right]}_{m} x - \underbrace{\left[\frac{c}{b}\right]}_{c}$$

* if plane more than 3D (Dimentionality up to n (4D, 5D, 6D ---- nD) is called as Hyperplane

* if Circle (2D) in 3-Dimentionality is called as 'Sphere'

if we have ~~line~~ [illegible]

* Above 3D upto n.D is called as Hyper Sphere

~~Hp~~ Hyper plane (4D) Equation :-

General Eq — $ax + by + c = 0 \Rightarrow m = \frac{-a}{b}$; $c = \frac{-c}{b}$

(4D)-Hyperplane — $w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_0 = 0$

G. Eq. can be written as — $w_1x_1 + w_2x_2 + w_0 = 0$

$$m = \frac{-w_1}{w_2} \; ; \; c = \frac{-w_0}{w_2}$$

if hyperplane (3D) - $w_1x_1 + w_2x_2 + w_3x_3 + w_0 = 0$

In this plane we have two Slopes (m) and One Intercept

$$\left.\begin{aligned} m_1 &= \frac{-w_1}{w_2} \text{ wrt to } x_1 \text{ (plane)} \\ m_2 &= \frac{-w_3}{w_2} \text{ wrt to } x_3 \text{ plane} \end{aligned}\right\} y = \underline{m_1}x_1 + \underline{m_2}x_2 + C$$

$$c = \frac{-w_0}{w_2} \text{ (Intercept with 'y' axis)}$$

from this we can write as - $\boxed{\sum_{i=1}^{d} w_i x_i + w_0 = 0}$

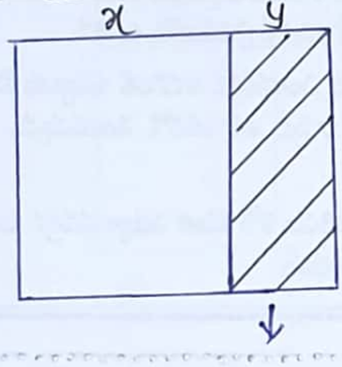
Simply we can Say → $\pi_d : w^T \cdot x + w_0 = 0$

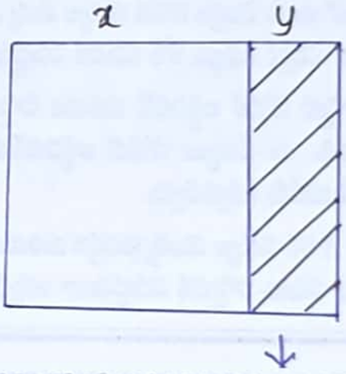
↳ No of Dimentions

if we have 5D → $\pi_5 : \underset{\text{Coefficent}}{w^T \begin{bmatrix} \quad \end{bmatrix}} \; \overset{x}{\begin{bmatrix} \\ \\ \end{bmatrix}} + w_0 = 0$

→ Dimensions

Classification vs regression :-





y → discreate (or) Categorical variable	y → Continious Variable
$y_i \in \{0,1\}$	$y_i \in$ Any Real No. (R)
it is Countable finet Set	it is infinet Set

Note → After E.D.A We have to do Data preparation.

Ex:

x	y
	Red
	Green
	Blue
	Black
	white

if we observe this → Data is Categorical

So we can Say classification problem

$y \in \{Red, Green, Blue, Black, white\}$

Red → 0, Green → 1, Blue → 2, Black → 3, white → 4

Linear Regression :- it is a machine learning algorithm on Supervised learning. it performs the task to predict a dependent variable value (y) based on a given indipendent variable value (x) this technique finds out a linear relationship between X (input) and y (output)

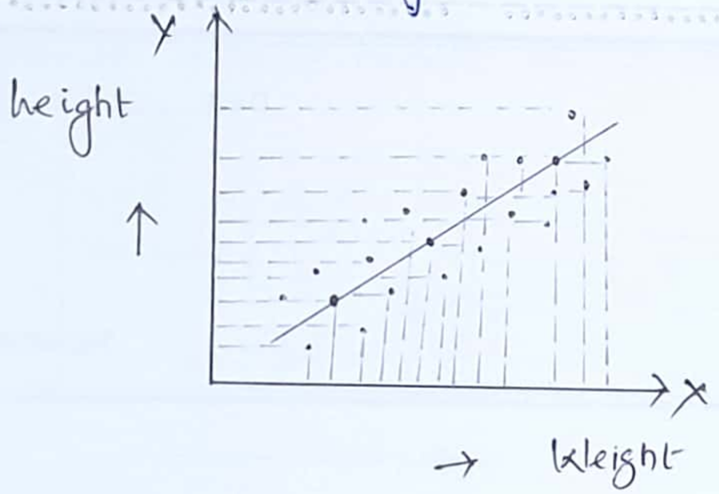
Ex: if we take a heights & weight of a Students

weight x	height y

→ Continious variable

So i need to do regression task

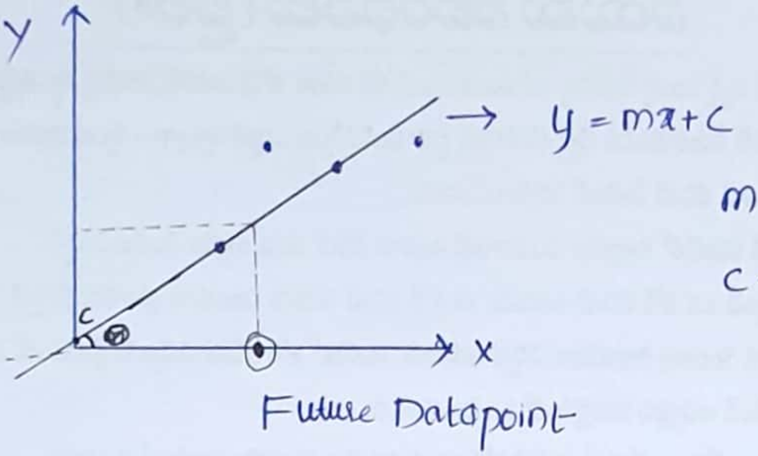
plot a Scatterplot diagram :-



H(↑) ⇒ W (↑)

positive Correlation

Suppose if i fit a line b/w the datapoints, i can easily predict height of a person by given weight. am able to predict the future data point becouse of that best fit Line.



m → Slope On xaxis.

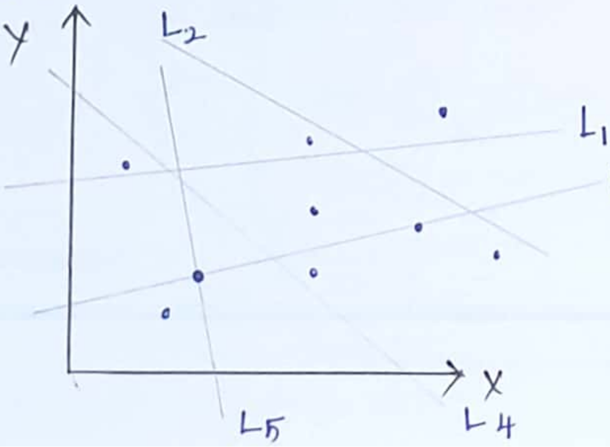
c – Contenent ~~O~~n Y

We know that line is (y=mx+c), if i want the line i want to know the m^*, c^* values Otherwise we can't that line.

Suppose if i have m & c values i can draw a line (y=mx+c) by Using this line. if any future data point Comes Simply we can predict it by Using this line as above Shown.

How to find a line that best fit the given data ?

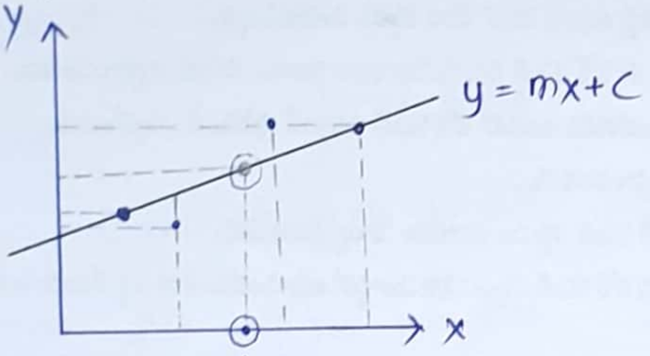
A line passes through most of datapoints in a given dataset Generally we Consider it as a best fit line of that dataset.



L_3 ⇒ This line<L_3> passes through most of Data points as compared to other Lines.

The above Shown figure we draw a line based on data points there are Somany lines that passes through data.

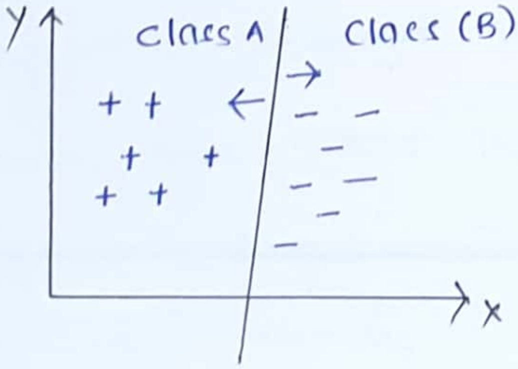
if we observe line 3 that Contain maximum no of points lying with that line as Compared to rest of line So Simply we can Say that the line 3 is Best fit line.



The Above Shown figure Says line passes through points, all which in data Set, by this we can predict the values of 'y' by using this line. ~~If we~~ Incase if we have future datapoint on x axis we have to find $y_{predict}$ value by using this line we Simply predict. if incase future data point is Realnumber (Continious)

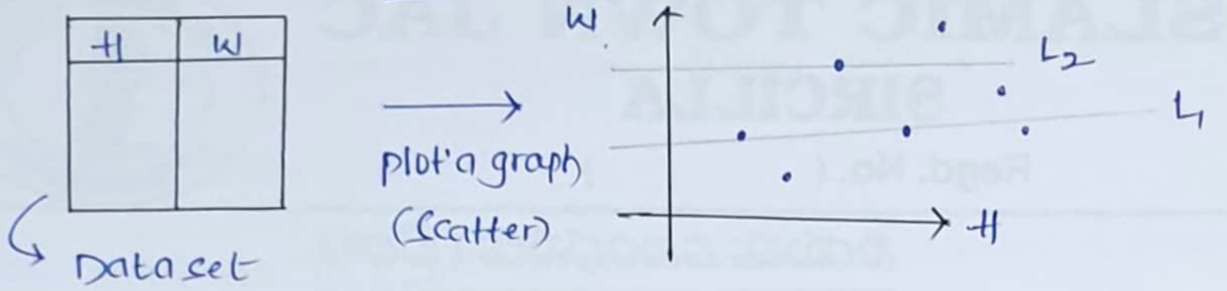
Note — i) In Regression we have to fit the Line with data points

ii) In classification we have to Separate the data by line



* class (A),(B) Separated by a line

* Process to find a line that BEST FIT the Given data :—

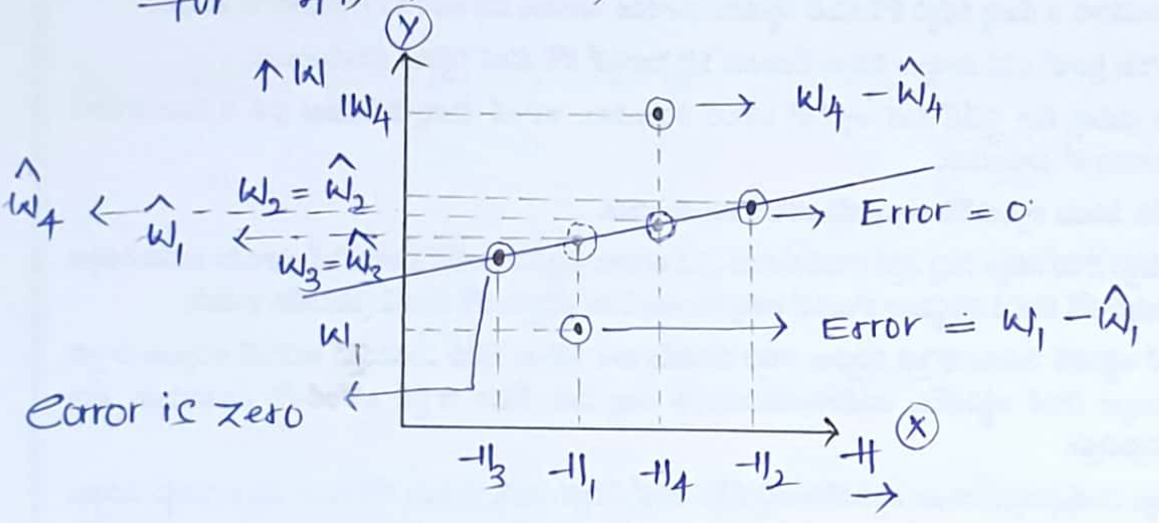


from the above graph we can Say line(L_1) is Bestfit line as Compared to L_2 but how to practically do you Come up with relationship that help us which line is better

So here we are trying to fit a line as good as possible. —

for 2D → we use line &

for 4D → we have to use Hyper plane.



For the point H_2 the actual value of weight is ω_1 Suppose if i want to predict H_1 by Using line i have to Scale the point on line $y=mx+c$ and predict the weight on y-axis as $\hat{\omega}_1$

if we abserve there is error associated with the point H_1

when Compared to Actual with predicted value as —

$$W_1 - \hat{W}_1$$

* For the point $\#_2$ there is no Error because the point is lying on line $(y=mx+c)$, So here Actual value & predicted value is Same, Error = 0.

* For point $\#_4$ there is an Error because the actual & predicted value is not Same.

$$\text{point } \#_1 \rightarrow w_1 - \hat{w}_1 = -ve$$
$$\text{point } \#_2 \rightarrow w_2 - \hat{w}_2 = 0$$
$$\#_3 \rightarrow w_3 - \hat{w}_3 = 0$$
$$\#_4 \rightarrow w_4 - \hat{w}_4 = +ve$$

In Order to get total Error made by this line we have to add all points.

$$\begin{aligned} \rightarrow \#_1 &\rightarrow -\cancel{ve} \\ \#_2 &\rightarrow 0 \\ \#_3 &\rightarrow 0 \\ \#_4 &\rightarrow +\cancel{ve} \\ \hline & \quad 0 \rightarrow \text{Total Error} \end{aligned}$$

* +ve magnitude is equal to −ve magnitude get cancelled.

Hence the total Error made by this line is zero, but if we See the line there is an Error associated with the line In order to avoid it, we have to do the Square for all Errors and we have to add it.

$$(w_1-\hat{w}_1)^2 + (w_2-\hat{w}_2)^2 + (w_3-\hat{w}_3)^2 + (w_4+\hat{w}_4)^2$$

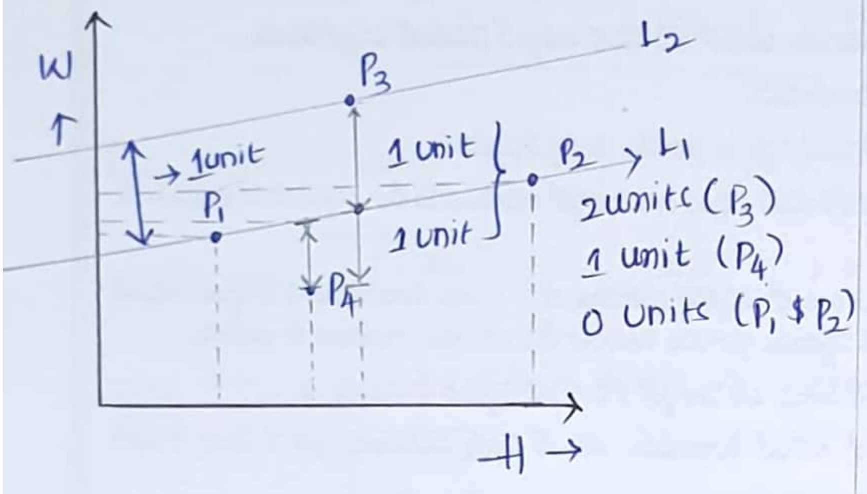
In Simply we can write as — $\sum_{i=1}^{4} (w_i - \hat{w}_i)^2$

Where 'n' total NO of points, we have to generalize this as —

y then Eqn = $\boxed{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$

↳ Total Error made by line

Mathematically :-



$L_1 \Rightarrow$

$\Rightarrow P_1 \Rightarrow (y_1 - \hat{y}_1) = 0$

$P_2 \Rightarrow (y_2 - \hat{y}_2) = 0$

$P_3 \Rightarrow (y_3 - \hat{y}_3) = 2$

$P_4 \Rightarrow (y_4 - \hat{y}_4) = 1$

$Eqn \Rightarrow \sum_{i=1}^{4} (y_i - \hat{y}_i)^2$

$\Rightarrow (0)^2 + (0)^2 + (2)^2 + (1)^2$

$= 4 + 1 = \underline{5}.$

According to $L_2 \Rightarrow$

$P_1 = (y_1 - \hat{y}_1)^2 = (1)^2$

$P_2 = (y_2 - \hat{y}_2)^2 = (1)^2$

$P_3 = (y_3 - \hat{y}_3)^2 = 0$

$P_4 = (y_4 - \hat{y}_4)^2 = (2)^2$

$Eqn = (1)^2 + (1)^2 + (0)^2 + (2)^2$

$= 4 + 1 + 1$

$= 6$,,

∴ Total Error made by line (L_1) is 5 and (L_2) is 6 so L_1 has min Error as compared to L_2, Hence we will go through L_1 as Best-fit Line for this dataset.

$$\boxed{\text{Min} \sum_{i=1}^{4} (y_i - \hat{y}_i)^2}$$

So we want to minimize the total Error (or) Sum Squared Error

∴ Minimum Error is the Best Fit line.

The line which gives minimum Error is the BEST FIT LINE

$$\boxed{\min \sum \text{Error}^2}$$

$$\boxed{\min \Sigma (\text{actual}_i - \text{predicted}_i)^2}$$

$$\boxed{\min \Sigma |\text{actual}_i - \text{predicted}_i|}$$

The Square value will get magnitized, the absolute value will not get magnified, By doing Square we can magnified value and by doing absolute we can't magnify value.

$$\boxed{m^*, c^* = \min \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \rightarrow * \text{ Represent Optimal value}$$

→ A. value → p. value.

So here predicted value of $y \rightarrow (\hat{y}_i)$ is given by the line $(y = mx + c)$

$$\boxed{m^*, c^* = \underset{m,c}{\arg} \min \left\{ \sum_{i=1}^{n} (y_i - \hat{y}_i(mx_i + c))^2 \right\}}$$

if we know (m,c) values, then any future data point comes.

We can easily predict the value by pointing all values in above eqn.

Note :- We Select the line in Such way which gives the minimum Error.

By using Gradient Descent → Used in ML, DL

- By Using linear Regression we will get the BEST FIT Line
- While using linear Regression our data doesn't contain Outliers
- In above eqn if we find argument m, c that gives best fit line with minimum Error.

Eqn — $$\boxed{m^*, c^* = \text{args}_{m,c} \min \left\{ \sum_{i=1}^{n} (y_i - (mx_i + c))^2 \right\}}$$

is also called as Ordinary least Square (OLS), [illegible]

$$\min \left\{ \sum_{i=1}^{n} (y_i - (mx_i + c))^2 \right\}$$

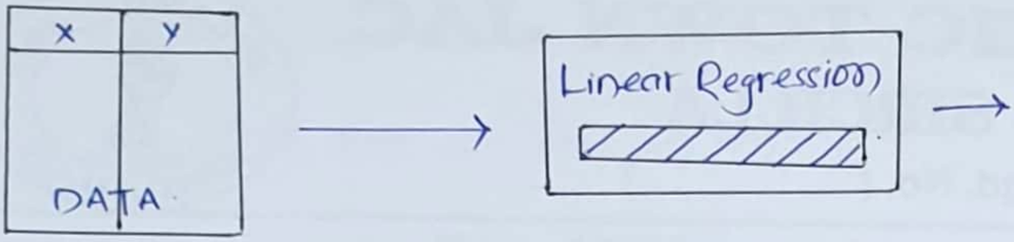
or

$$\min \left\{ \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \right\}$$

→ Error / Residual Error

- We have to do regression task only when input & Output both are mapping and output is Continious in Supervised Learning.
- All the points which are lying on line ($y = mx + c$) that gives error as zero, and Best fit line gives minimum Error
- The points which are not lying on line will give Error and we call it as (Training Error) $(y_i - \hat{y}_i)$
- if we dont Square the error, the total Error made by the line is zero, but by vizualizing the points there is an error.
- In Eqn x_i, y_i is data given in training data.

When Data given to linear Regression :-



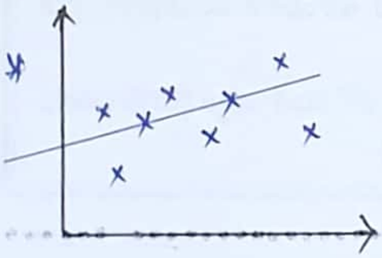
This Gives the line with minimum Error by optimizing m, c. and there is a technique (Algorithem) to optimizing m & c values.

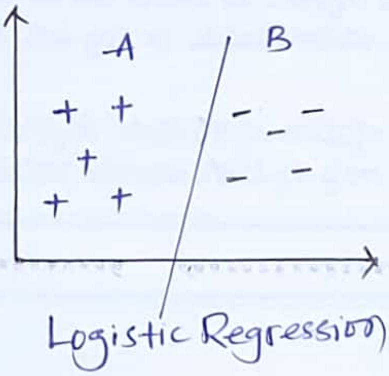
When Data Given to Logistic Regression :-

* Logistic Regression is Used for classification (Descrete / ~~Continous~~ categorical) Type of tasks.

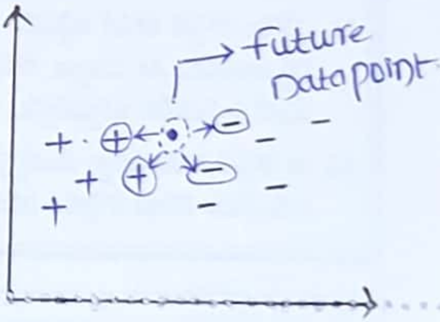
* there is an algorithem technique for optimizing m, c values in Logistic.

* In this we draw a line that separates the classes.


Linear Regression



Logistic Regression



KNN Neighbour

→ In Case of KNN Neighbour, if any future Datapoint Come, then we finding the k-nearest neighbour to that point by Using the Euclidian distance.

→ In KNN there is no Separation of the classes.

→ In logistic regression we draw a line that separate class, if in case the point lying on line we can't say anything.

Lineal Regression vs Logistic Regression

<table>
<tr><td>1)</td><td>Superviced Learning Technique</td><td>Superviced Learning.</td></tr>
<tr><td>2)</td><td>Regression Taske</td><td>Classification Tasks</td></tr>
<tr><td>3)</td><td>O/p variable is Continious</td><td>O/p variable is Descreate</td></tr>
<tr><td>4)</td><td><table><tr><th>X</th><th>Y</th></tr><tr><td>x_1</td><td>y_1</td></tr><tr><td>x_2</td><td>y_2</td></tr><tr><td>⋮</td><td>⋮</td></tr><tr><td>x_n</td><td>y_n</td></tr></table>
$D_n = \{(x_i, y_i)_{i=1}^{n} \mid y_i \in R\}$</td><td>$D_n = \{(x_i, y_i)_{i=1}^{n} \mid y_i \in \{+1, -1\}\}$
Where – D = Dataset
n = no of Datapoints.</td></tr>
<tr><td>5)</td><td>Task – Find a line (m^*, c^*) that BEST FIT the training Data.
↳ Minimizing the Sum of Squared Errors.
$\min \sum_{i=1}^{n} (y_i - (m x_i + c))^2$</td><td>5 find a line that BEST SEPARATE +ve's from –ve's.
$\{(x_i, y_i)_{i=1}^{n} \mid y_i \in \{+1, -1\}\}$</td></tr>
</table>