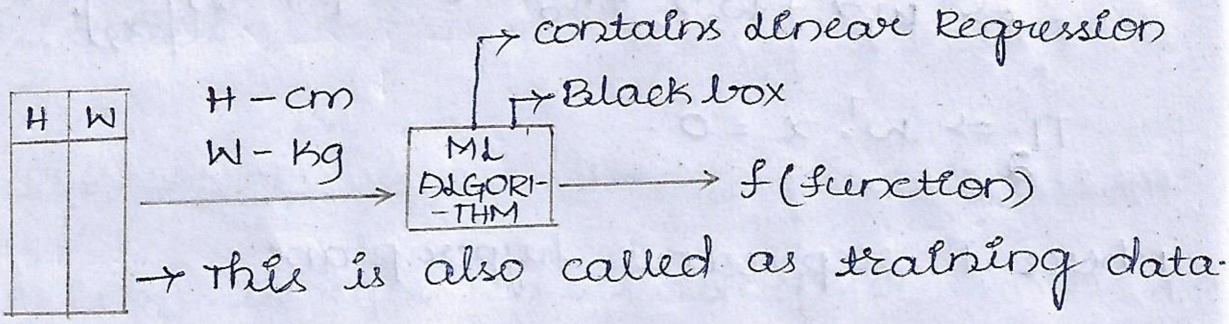
LINEAR REGRESSION:

It shows a linear relationship between a dependent and one or more independent variables.



↳ The entire data is kept in the black box which contains the ML Algorithm and learns the function.

The H is denoted as 'X' & W is denoted as 'y'.

$y \in \mathbb{R}^1$ → 1 denotes the number of dimensions.

Predictor → Weight (Kg) → Regression Algorithm.

$\square \rightarrow f \rightarrow \hat{y}$

future data point

EXAMPLE:

X	Y
10	100
0	0
5	25
1	1
2	4

$\longrightarrow$ ML ALGOR. $\downarrow$ REGRESSION $\longrightarrow f \rightarrow \hat{Y} = x^2$

$Y \in \mathbb{R}'$

X	Y
3	?

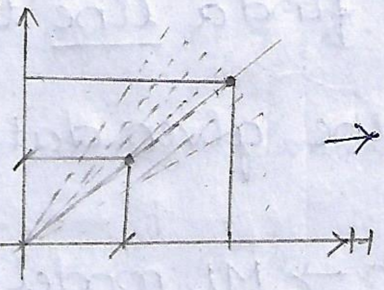
$\longrightarrow \hat{y} = x^2 \rightarrow \hat{y} = (3)^2$
$= 9.$

HOW LINEAR REGRESSION IS USEFUL TO FIND ML ALGORITHM?

A statistical method that is used for predictive analysis, makes predictions for continuous/real or numerical ~~values~~ variables.

H	W
X	y

$\longrightarrow$

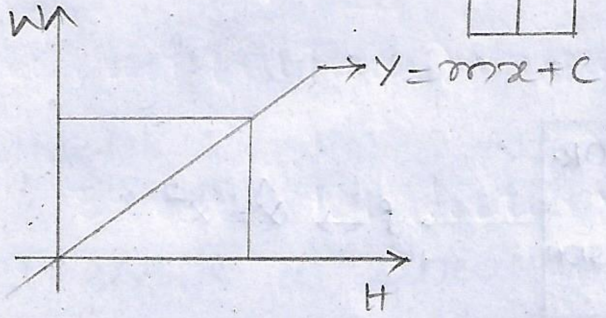


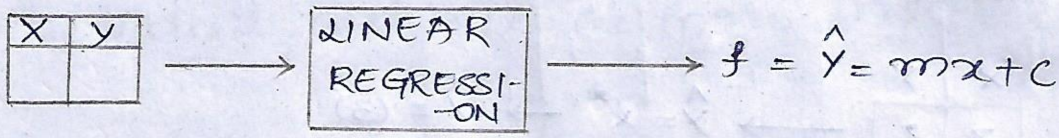
$\rightarrow$ As H$\uparrow$, weight tends to increase.

The data points which are plotted, needs to fit the line and from that weight can be

Predicted.

H	W
X	y





$\Rightarrow X \rightarrow f \rightarrow \hat{Y}$

$\downarrow$

$Y = mx + c$

STEPS:

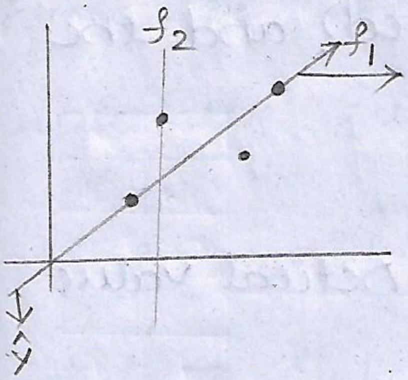
TASK - Trying to find a function that can exactly fit the data given to us.

Mathematically:-

TASK - Trying to find a line that can pass through the given data points

line → function → ML model

pass → fit



→ Trying to capture more variance.

→ Infinite lines can be passed, but we need to find the line which is the best fit for the data given to us.

Mathematically: $y = mx + c$

→ If we know m & c then $\hat{y}$ is found easily.

QUESTION: From above graph is f_2 better than f_1?

ANSWER: DEPENDS ON THE DATA POINTS

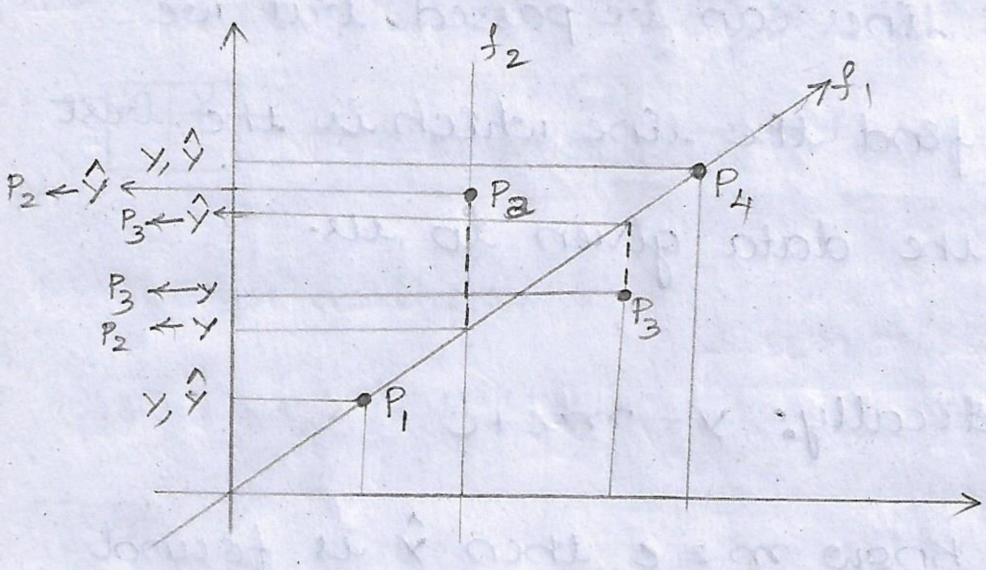
The best fit which gives least number of errors for the given data points.

ERROR:

The difference between the value which

has been computed (Predicted) value and the correct (actual) value.

Error = Predicted value - Actual value.



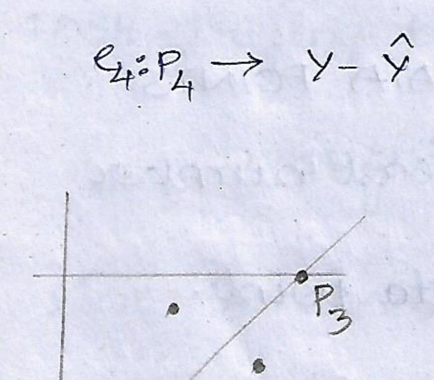
For each data point, error is

$e_1 : P_1 \rightarrow y - \hat{y} = 0$

$e_2 : P_2 \rightarrow y - \hat{y} = +ve$

$e_3 : P_3 \rightarrow y - \hat{y} = -ve$

$e_4 : P_4 \rightarrow y - \hat{y} = 0$



→

$e_1 \rightarrow P_1 \leftrightarrow y - \hat{y} = -ve$

$e_2 \rightarrow P_2 \leftrightarrow y - \hat{y} = +ve$

$e_3 \rightarrow P_3 \leftrightarrow y - \hat{y} = 0$

$e_4 \rightarrow P_4 \leftrightarrow y - \hat{y} = +ve$

From the above graphs, we need to look into the sum of the total errors.

$\Rightarrow f_1 \rightarrow \cancel{e_1}^{\to 0} + e_2 + e_3 + \cancel{e_4}^{\to 0} + e_5$

$\Rightarrow f_2 \rightarrow e_1 + e_2 + \cancel{e_3}^{\to 0} + e_4 + e_5$

so, there are errors having +ve & -ves.

* The best way is we need to look into the magnitude.

$$\rightarrow \sum (e_i)^2 = \sum (Y_i act - Y_i pred)^2$$

we need to minimize the total error.

Y_i predicted $\rightarrow \hat{Y} = mx + c$.

$$\Rightarrow m^*, c^* = \min_{c,m} \sum_{i=1}^{N} (Y_i act - \underbrace{\{mx_i + c\}}_{\text{need to find \& minimise the eq}^n})^2$$

* ↳ m^*, c^* are stated as optimal value (or) best values (or) ordinary least squares (OLS) (or) simple linear regression.

TASK : Try to find m, c such that it minimizes the sum of squared errors.

→ Squared Errors is also called as "RESIDUALS".

Mathematically,

$$m^*, c^* = \underset{m,c}{\arg\min} \underbrace{\boxed{\sum \left(Y_{act} - \{mx + c\} \right)^2}}_{\text{SUM OF SQUARED ERRORS}}$$

$$m^*, c^* = \underset{m,c}{\arg\min} \sum \left(Y_{act} - Y_{Pre} \right)^2$$

OPTIMIZATION :

It is the problem of finding a set of inputs to an objective function that results in a maximum or minimum function evaluation.

↳ It is a challenging problem that underlies many ML Algorithms, from fitting logistic regression models to training models.

OPTIMAL VALUE :

The minimum value of the objective function over the feasible region of an optimization problem.

WHY SQUARED ERROR IS BETTER THAN ABSOLUTE VALUE ?

Having a square as opposed to the absolute value function gives a nice continuous and differentiable function. (absolute value is not differentiable at zero)- which makes it the natural choice, especially in the context of estimation and regression analysis.

ORDINARY LEAST SQUARES:

It is a method in Linear Regression for estimating the unknown parameters by creating a model which will minimize the sum of the squared errors between the observed data and the predicted one.

→ The smaller the distance, the better model fits the data.