lec - 50

ML - Algorithms

Linear regression

Why we study it

① It is easy to understand
(we can understand only by watching)

② It is foundation of algorithm
(whichever we study in linear regression we are also study in further algo)

LR → Supervised ML Algo
↓
regression. (i.e output is numerical)

LR
- Simple LR
- Multiple LR
- Polynomial LR

→ we also do regulization & different different new LR

Simple LR

1 input col | 1 output col

for ex:- A student is placed & his pacakage based on cgpa

cgpa	pacakage
6.66	3.01

↳ input ↳ output

MLR

more than 1 input col | 1 output col

for eg:- cgpa | gender | pacakage
2 input col ↳ one output col.

Polynomial LR

→ when our model is not linear we use polynomial LR.

Simple linear regression

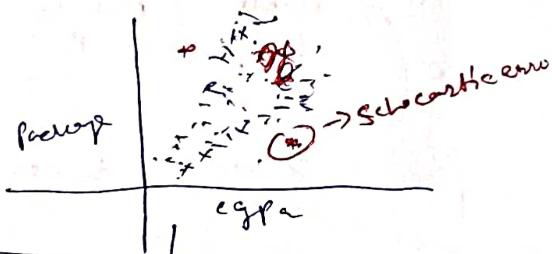
we have a $\boxed{\text{model}}$
↳ we give it cgpa
↳ It gives Package

cgpa	package
7.1	3.5
4.7	1.2
⋮	⋮

NOTE:- Suppose we have previous data so, we can see the data and based on the data we can tell what is the Package of given student's cgpa.

Now let suppose/ there is/ no data.
→ best way to do we can evaluate the Avg package of Previous but this is not good bcs we cannot tell what student capability to get that package
It is vary between good & bad student.

now, on the basis of given data we can draw graphs to plot this data

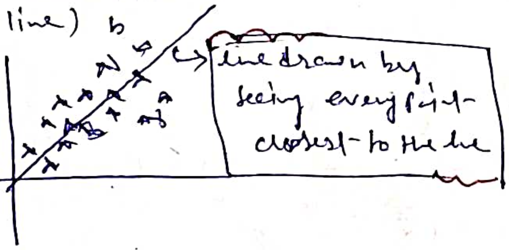


↳ It is not completely linear it is sort of linear data.

(bcs It is real world database so it is varies on real word database)

Note - If there is sort of linear data there also we can draw a line (is called best fit line) b

Why it is called so bcs this line makes minimum mistakes



~~Stotast~~ Stocheastic error -

The error which we cannot determine

for ex:- a student had good interview but how can we mathematically say how the good or bad the interview was

Linear regression

In linear regression it find a line (best fit line)

i.e $y = mx + b$

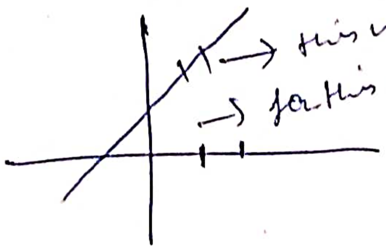
↑ for this line it find the value of m & b

What is human understandable weitage.

m → weightage (i.e how much salary depend on cgpa

if m is very large that means salary highly depend on cgpa

m " " less " " " less " " "



Package = $m \times cgpa + b$

↳ m is the weitage of cgpa on the package.

$y = mx + \boxed{b}$ → b is offset

let $b = 0$

$y = mx$

↳ package

↳ experince

let $m = 0$ $y = 0$

i.e if the person is fresher then he/she get 0 salary which is wrong.

So, (b) is offset

class - 51

now how to find the value of m & b

(m, b)

there are two ways

Closed form solution

Non closed form solution

↳ In mathematics C-F-S expressed using finite number of standard operation

It may contain constant variables certain well known operation (+, −, ÷, ×)

& func (nth root, log, trigo, hyperbolic)

but not limit, continu, diff

ex: $y = ax^2 + bx + c$

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

but in the N-C-F-S we are approximate to find the value

So, (m, b) → finding there are two method

① ↳ direct formula (this technique known as $\boxed{OLS}$ (ordinary least squares))

② ↳ approximation known as (Gradient descent)

↳ generally used in higher dimension

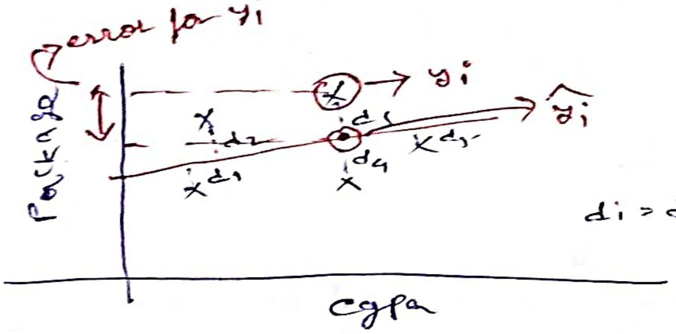
formula for m & b

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

y - package
x - cgpa

$\bar{x}$, $\bar{y}$ — mean values

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



d_i = distances between point & the line (Prediction)

total Error = $d_1 + d_2 + \dots + d_5$

we want error as positive

$E = d_1^2 + d_2^2 + \dots + d_5^2$.

we use not mod ∵
reason 1 : outlier
Jo der hai usko positive karta hai
reason 2 = we differentiate it later

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

↳ Error function / Loss function / represented also by J

So, we need to find the value of E that minimize m & b

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

avg error

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

↳ Prediction
↳ actual

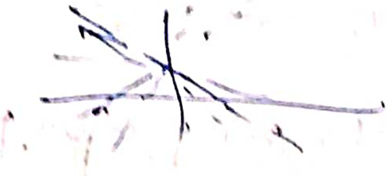
$$\hat{y}_i = m x_i + b$$

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

these are constant

So, error depend on m & b

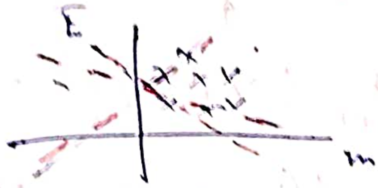
how? fix m b↑↓ fix b m↑↓



or both m↑↓, b↑↓

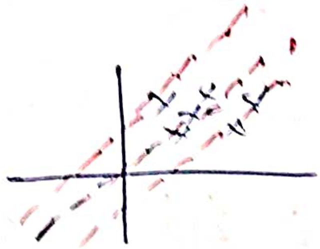
let assume $b = 0$

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

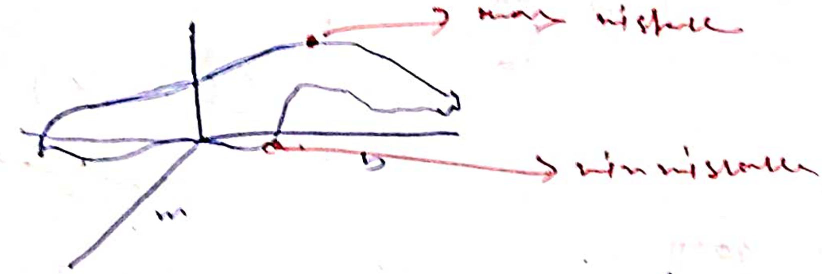


let assume $m = 1$

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



$E(m,b)$ both (m, b) is independent



now note :- we simply find maxima & minima for the graph

$$\frac{\partial E}{\partial m} = 0 \qquad \frac{\partial E}{\partial b} = 0$$

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

$$= \sum -2(y_i - m x_i - b) = 0$$

$$= \sum (y_i - m x_i - b) = 0$$

dividing both side by n

$$\sum \frac{b}{n} = \frac{nb}{n} = b \qquad = \sum \Big(\underbrace{\frac{y_i}{n}}_{\text{mean } \bar{y}} - m \underbrace{\frac{x_i}{n}}_{\text{mean } \bar{x}} - \frac{b}{n} \Big) = 0$$

$\therefore \quad b = \bar{Y} - m\bar{X}$

$$\frac{\partial E}{\partial m} = \sum \frac{\partial}{\partial m} (y_i - mx_i - b)^2 = 0$$

$$= \cancel{\sum -2 (y_i - mx_i - b) \cdot m} = 0$$

$$= \cancel{\sum m (y_i - mx_i - b)} = 0$$

Putting value of b

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

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

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

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

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

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