

Revision - R^2



Unexplained Variance

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOTAL}}$$

Error = Residual
 $\sum_i^n (y_i - \hat{y}_i)^2$

$$\sum_i^n (y_i - \bar{y})^2$$

Total variance in the target

R^2 score

$R^2 \in (-\infty, 1]$

$-\infty$

$$\frac{SS_{RES}}{SS_{TOTAL}} = \frac{y}{x}$$

$$y > x, y/x > 1$$

$$R^2 = -ve$$

1

$$\frac{SS_{RES}}{SS_{TOTAL}} = \frac{0}{x}$$

$$R^2 = 1 - 0 = 1$$

0

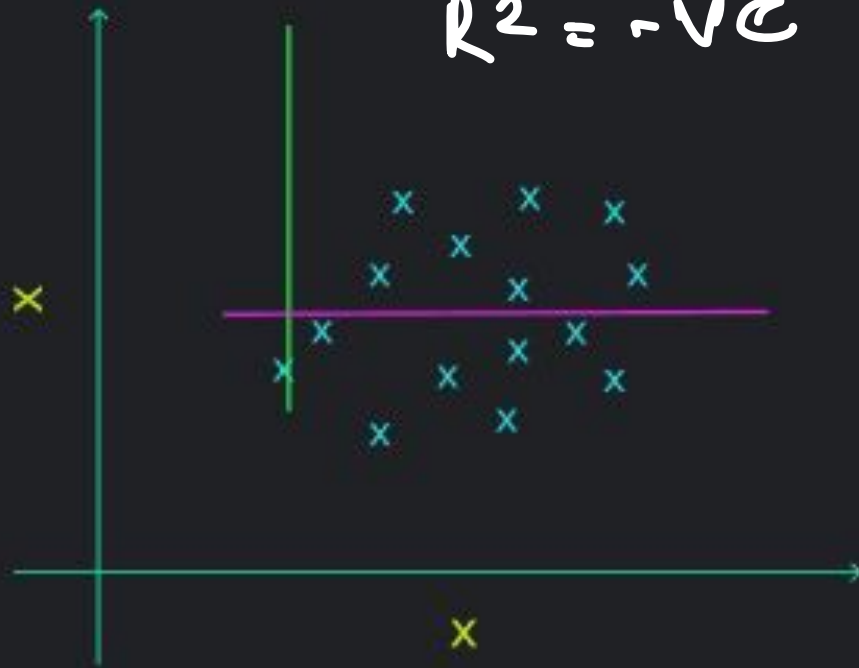
$$\frac{SS_{RES}}{SS_{TOTAL}} = \frac{x}{x}$$

$$R^2 = 1 - 1 = 0$$

a

Extremely
Bad

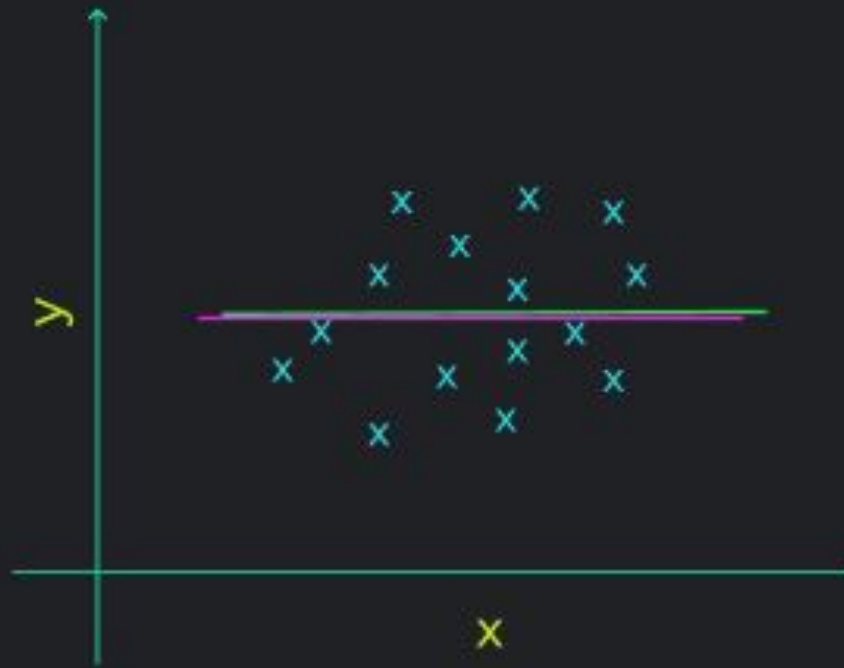
$$R^2 = -\infty$$



b

As bad as mean

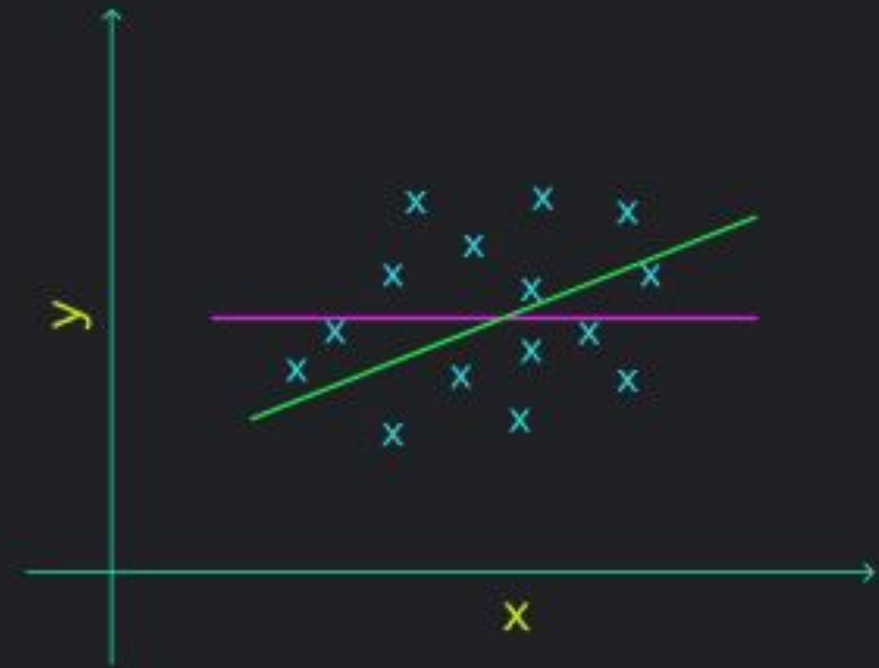
$$R^2 \rightarrow 0$$



c

Good


$$R^2 > 0$$



Extremely Bad - Worse than mean

Bad

Good

 Predicted Model

 Baseline Model

Model Interpretability

Linear Regression - linear comb. of $\langle x_1 x_2 x_3 \dots x_d \rangle$

$$\text{Model } \hat{y}^i = w_1 x_1 + w_2 x_2 + w_3 x_3 + \dots + w_d x_d + w_0$$

$$\text{Model } \hat{y}^i = w_0 + w_1 x_1 + (-10000) \text{ age} + (10) \text{ edometer} + w_n x_n$$

Insurance Coverage

- 1) sign
- 2) Magnitude

Direction of movement

Cases as per sign

Constants

$$y = 2x_1 + 3x_2 + 5$$

1. -ve weight of feature

$x \uparrow \rightarrow y \downarrow$
 $x \downarrow \rightarrow y \uparrow$

2. +ve weight of feature

$x \uparrow \rightarrow y \uparrow$
 $x \downarrow \rightarrow y \downarrow$

3. $Wt = 0$

$x \uparrow \rightarrow$ No change in y
 $x \downarrow \rightarrow$ No change in y



Cases as per magnitude

Importance of a feature.

Let's take e.g of age & odometer

x_1

Age wt. = -10000

x_2

~~Odometer~~ wt. = 10

Insurance

Age of vehicle was increased by 1 year

$$x_1 \rightarrow x_1 + 1 ; y \rightarrow y - 10000$$

Insurance coverage was updated by 1 Re

$$x_2 \rightarrow x_2 + 1 ; y \rightarrow y + 10$$

$$x_2 \rightarrow x_2 + 1000 ; y \rightarrow y + 10,000$$

LARGER THE ABSOLUTE VALUE, MORE IMPORTANT THE FEATURE

So, is age >>>>> ~~odometer imp~~??

Insurance

>>>>>>



Scale unit

Age [0, 15] - w_1

lovenage [12, 102] = w_2

$w_1 \gg \gg \gg \gg w_2$

To compare the weights \rightarrow perform standardization.



AGE - [1, 15] Years ✓

Instance ~~ODO~~ - [5000, 250000] Kms

Solution

Feature Scaling



