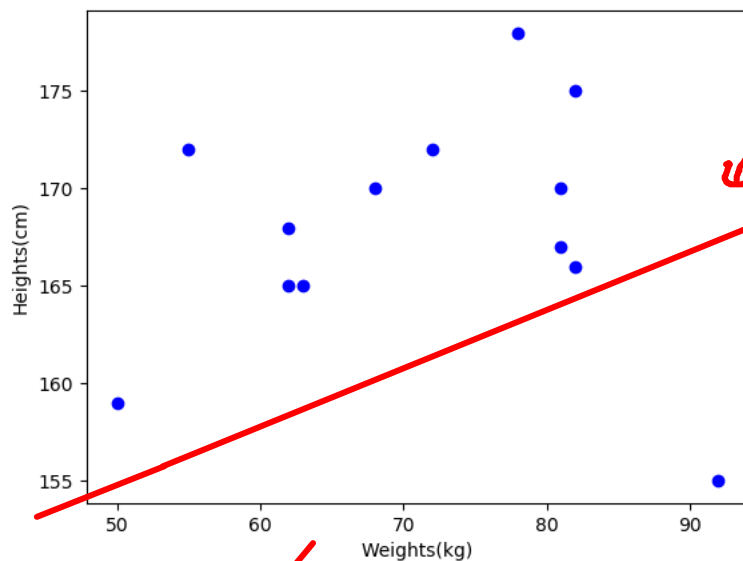
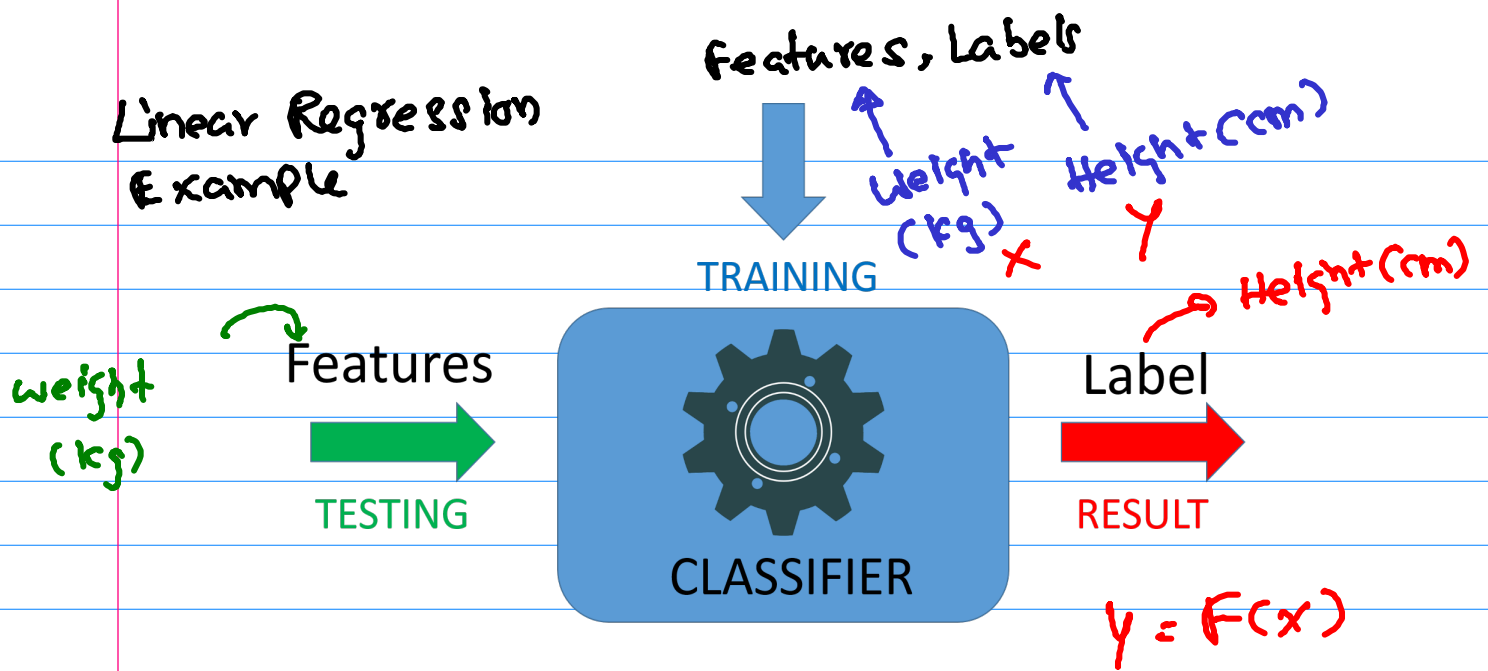
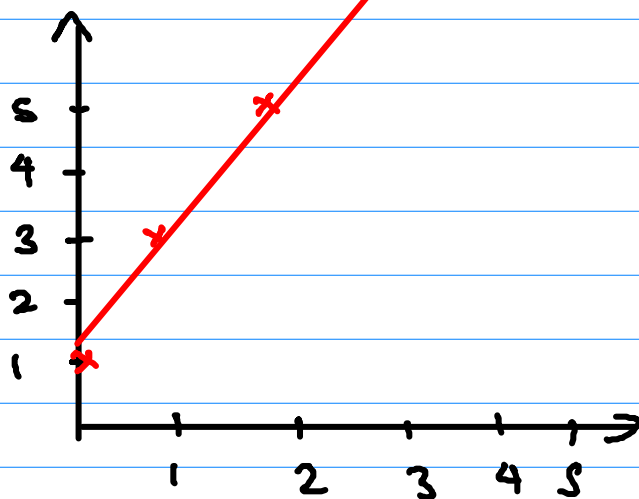


Linear Regression Example



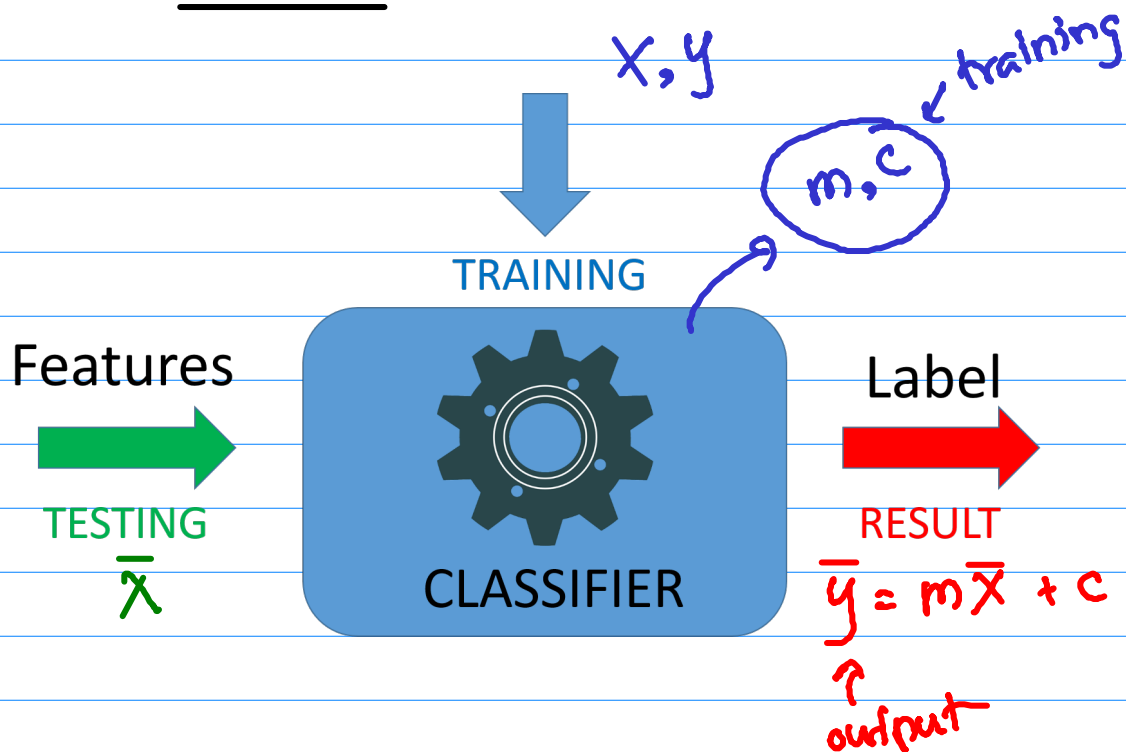
$$y = mx + c$$

$$m = 2, c = 1$$



x	y
0	1
1	3
2	5
3	7

How Linear Regression Works?



* $m, c \rightarrow$ Gradient, intercept of the Best Fit line drawn for the train data and train target (x, y)

$$y = m * x + c$$

$$m = \frac{\bar{x} \cdot \bar{y} - \overline{xy}}{(\bar{x})^2 - \overline{x^2}}$$

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

	x	y	xy	x^2
1	x_1	y_1	$x_1 y_1$	x_1^2
2	x_2	y_2	$x_2 y_2$	x_2^2
...				
.				
n				
	$\sum x$		$\sum xy$	$\sum x^2$

$$\bar{x^2} = \frac{\sum x^2}{n}$$

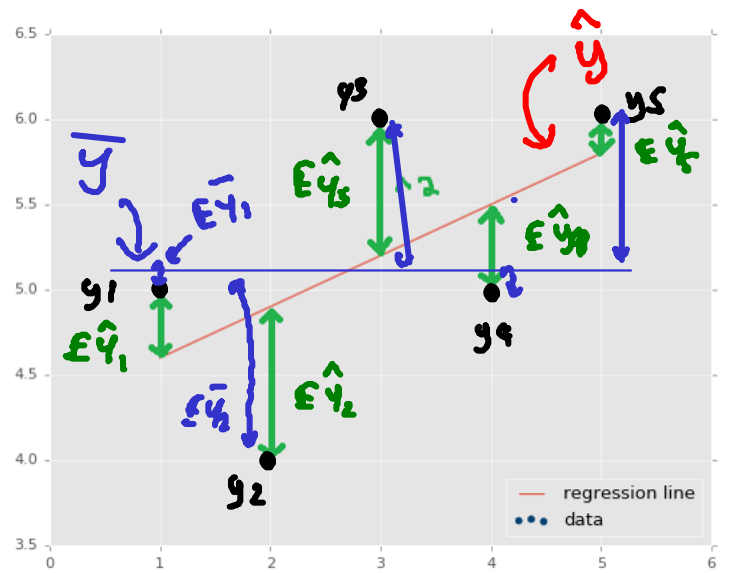
$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{y} = \frac{\sum y}{n}$$

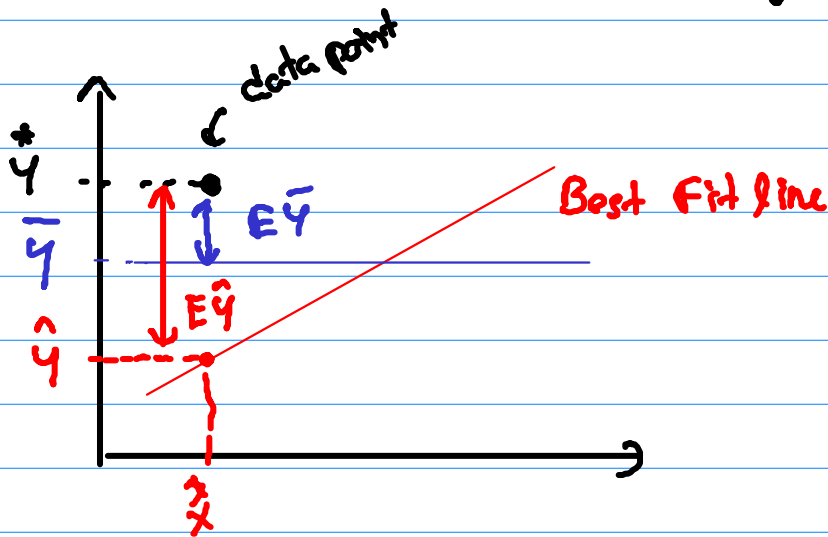
$$\overline{xy} = \frac{\sum xy}{n}$$

R^2 value

$$r^2 = 1 - \frac{SE_{\hat{y}}}{SE_{\bar{y}}}$$

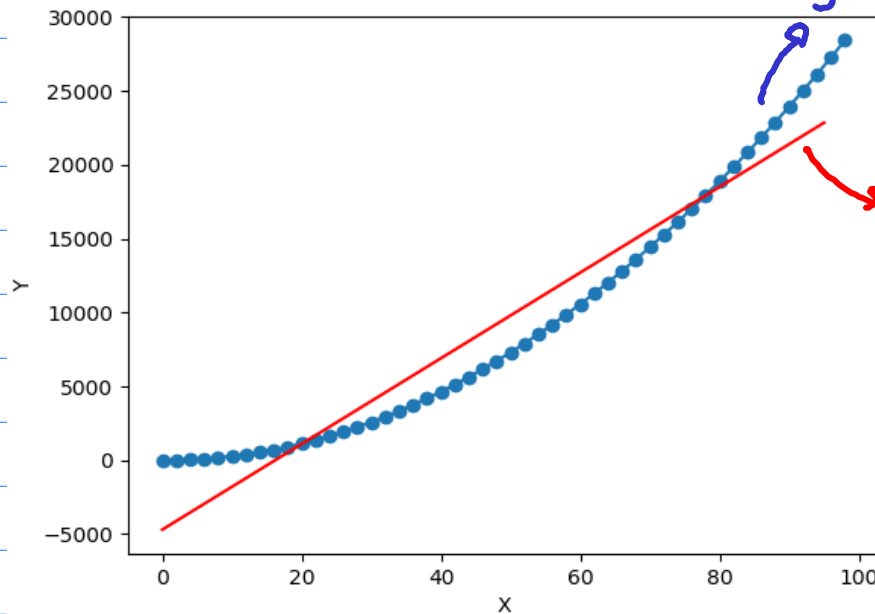


• $\bar{y} = \frac{\sum y}{n}$ (mean)



Linear Regression for Non Linear Data

$$y = a_1x + a_2x^2 + a_0$$
$$y = 3x^2 - 4x + 5$$



$$y = mx + c$$

Polynomial function

$$y = a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n + a_0$$

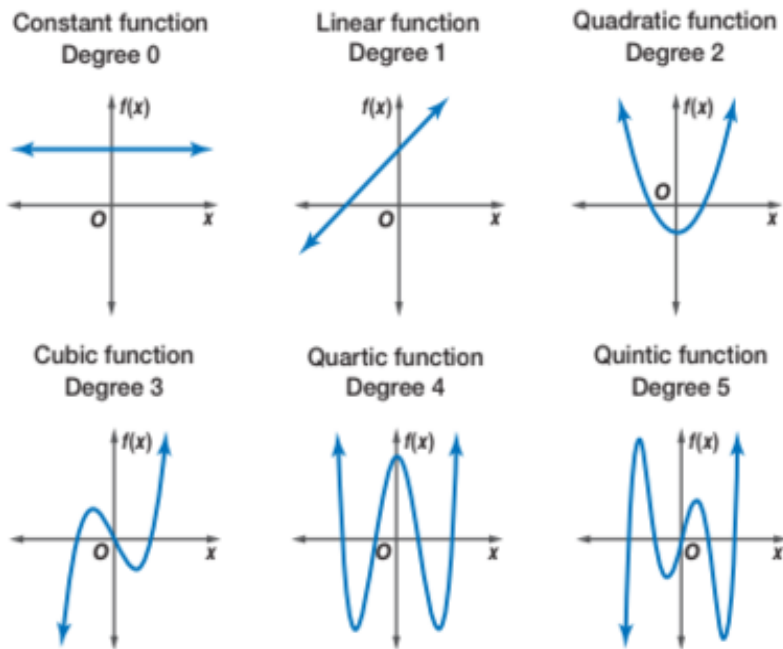
↗ constant (c)

$$\text{degree} = 2 \rightarrow y = a_1x + a_2x^2 + a_0$$

$$d = 3 \rightarrow y = a_1x + a_2x^2 + a_3x^3 + a_0$$

$$d = 1 \rightarrow y = a_1x + a_0 \quad (a_1 = m, a_0 = c)$$

$$d = 0 \rightarrow y = a_0 \quad y = mx + c$$



```
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
```

```
poly=PolynomialFeatures(degree=2,include_bias=False)
x_new=poly.fit_transform(x)
```

$x \rightarrow d=1$

x_1
x_2
x_3
x_4
\vdots

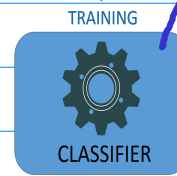
$x_{new} \rightarrow d=2$

x_1	x_1^2
x_2	x_2^2
x_3	x_3^2
\vdots	\vdots

$d=1$

Features
TESTING

\bar{x}



Label
RESULT

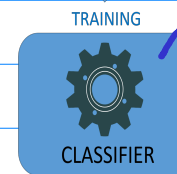
$$\bar{y} = a_1 \bar{x} + a_0$$

$(x, x^2), y$

$d=2$

Features
TESTING

\bar{x}



Label
RESULT

a_1, a_2, a_0

$$y = a_1 \bar{x} + a_2 \bar{x}^2 + a_0$$

Polynomial
Regression

Linear Regression for cardio dataset

