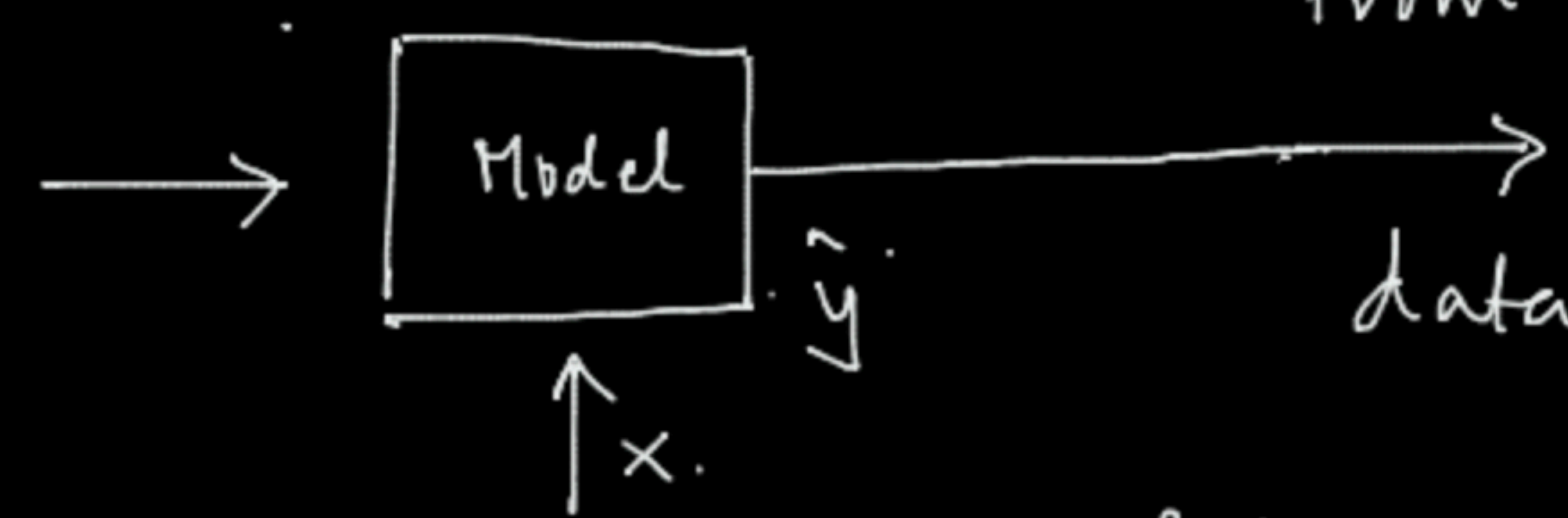
← Machine Learning →

Supervised Learning

Ind. dep

	X	y
Training phase	1	2
	2	5
	3	6
	4	9
	5	10
	6	13
Testing	7	14
	8	17

↳ Predictions.



Machine Learning.

↓

Extracted from data

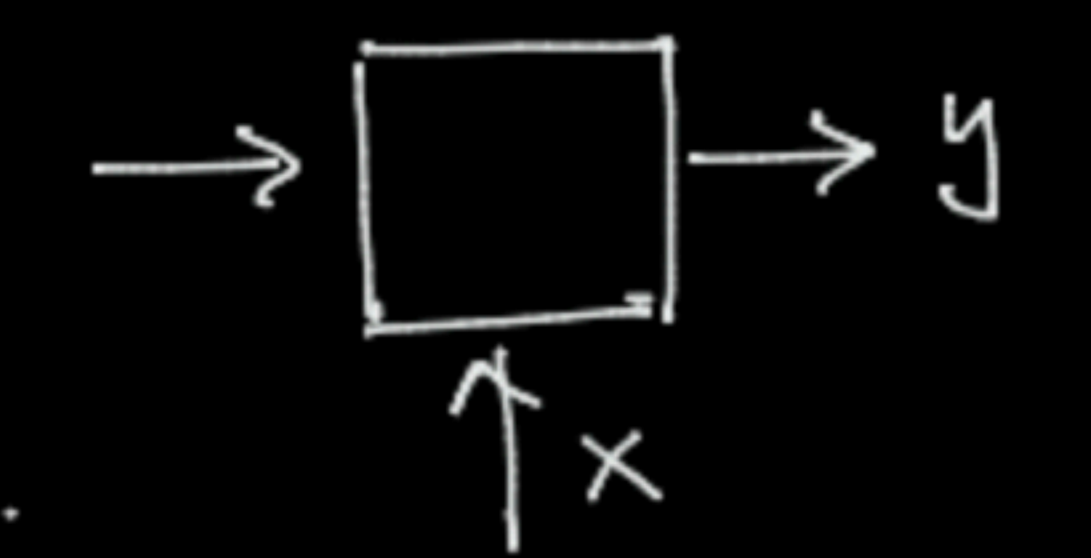
Traditional Programming.

if x is odd.

$y = 2x$.

else

$y = 2x + 1$.



dep. indep.

← Features of tumour → ← y → ↳ target.

X_{100} . . . . . .

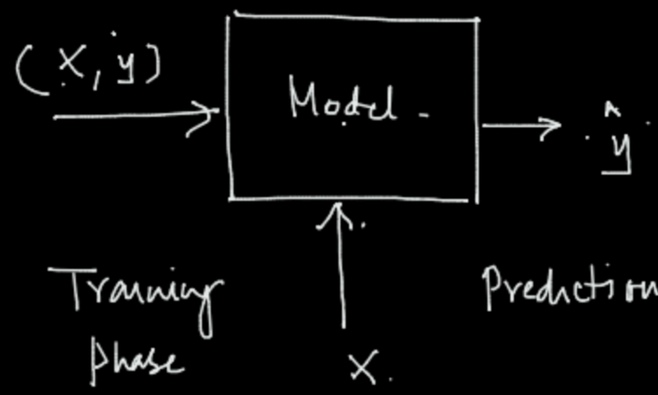
X_1	X_2	y
1	2	M
2	3	B

X	y
Features	Target
Indep. Var.	dep Var

Machine Learning

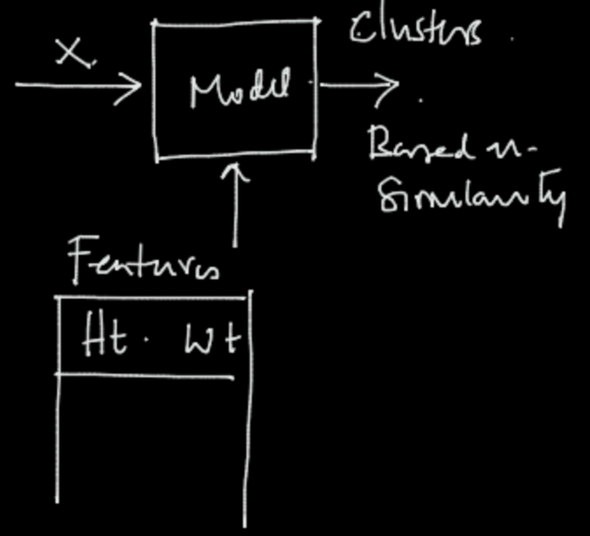
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Labelled Data



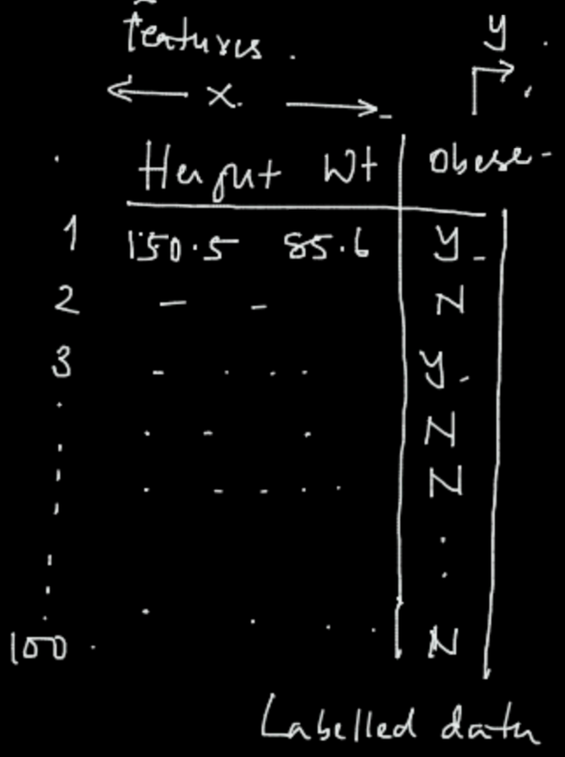
Labelled Data → Target Value is present in the training data.

— Features + Target.



Unlabelled Data → No target / Nothing to predict.

↳ Target is not available in the training data



Labelled data

Machine Learning

- Supervised Learning
  - Regression
    - Predict Continuous data.
  - Classification
    - predict categorical data.
- Unsupervised Learning
  - Clustering → Recommendations.

Nr	Nr	Cat	Cat	y
Area	No. of Bed	location	Builder	Price → Continu-

Target.

↳ Classification

Finder File Edit View Go Window Help

5 / 5

← Simple linear Regression →.

only one Feature. ← | ↓ | → Target is cont.

$y = f(x)$.

↳ linear fn.

x. y.

Area → Price.

Exp → Salary.

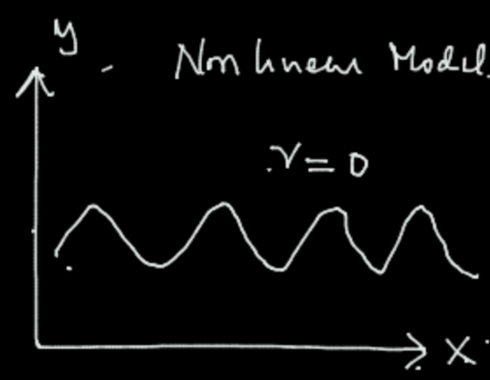
1. Relationship between x & y.

(a) Scatter plot.

(b) Strength of the Relationship between x - y (r).

Correlation Co.eff. (r) → how strong is the LINEAR Relationship between x, y

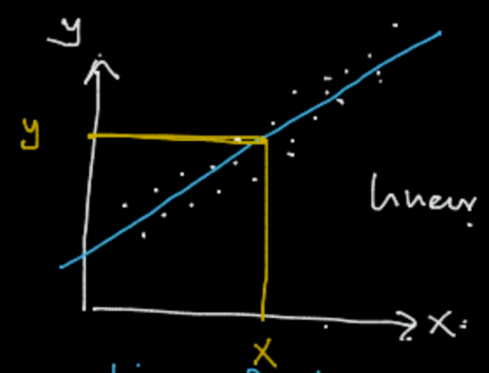
→ −1 to +1



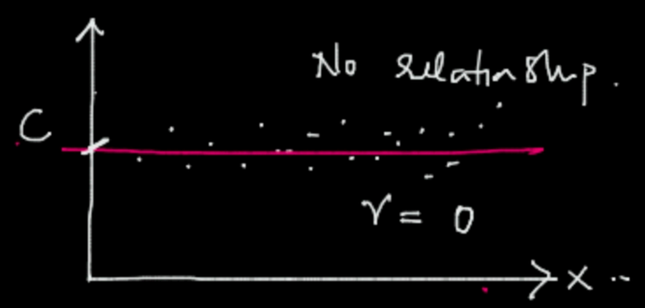
Non linear Relationship

✓

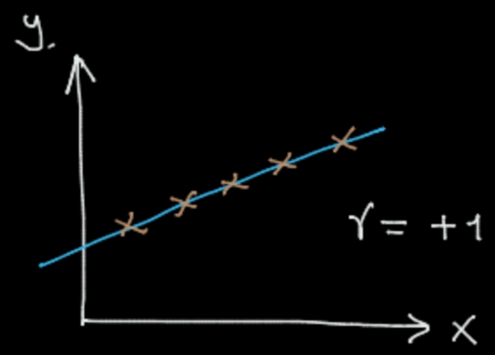
$y = mx + c$ → St. line



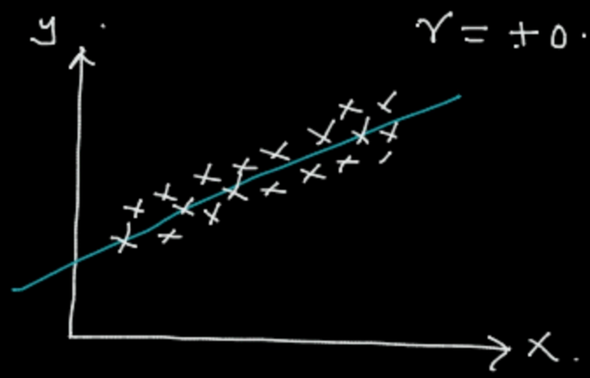
Linear Relationship.

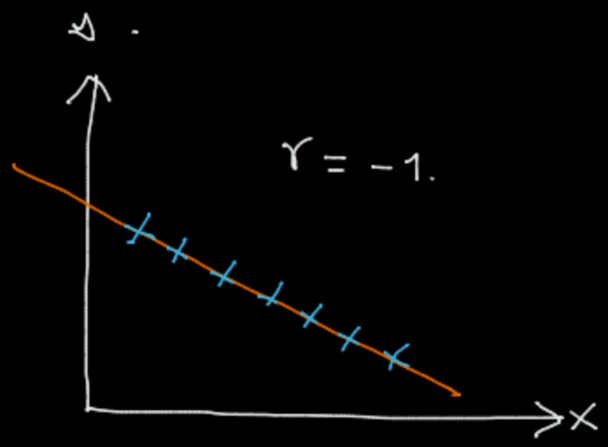


$y = c$.



+ve correlation : x ↑ y ↑.





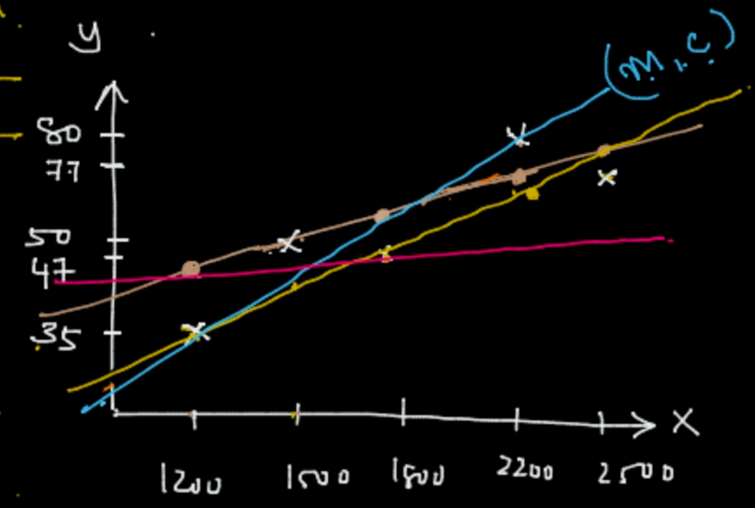
Good linear Model.

$r \geq \pm 0.85$.

$r = +0.6$.

X. Area	y. Price	$\hat{y}$	$(y-\hat{y})$	$(y-\hat{y})^2$	$\lvert y-\hat{y}\rvert$	$\hat{y}$
1200	35	45	−10	100	10	35
1500	50	50	0	0	0	39
1800	47	52	−5	25	5	0
2200	80	75	+5	25	5	75
2500	77	79	−2	4	2	70

MAE — Mean Absolute Error



$\hat{y} \leftarrow mx+c$

MSE → Mean Squared Error

RMSE → $\sqrt{\text{MSE}}$

→

← Metrics for evaluating Regression Model →

1. MSE ✓
2. RMSE ✓
3. MAE ✓
4. R^2 Value

" Ordinary Least Squares (OLS)"

y	$\hat{y}$	$y-\hat{y}$	$(y-\hat{y})^2$
35	35	0	0
50	39	11	121
47	47	0	0
80	75	5	25
77	70	7	49

MSE

Area y $\hat{y}$ $y-\hat{y}$ $(y-\hat{y})^2$

↳ Mean squared Error ✓ ✓
RMSE ✓
Mean Absolute Error ✓
R _ Squared

} unstandardised metric.

↳ No specific range.

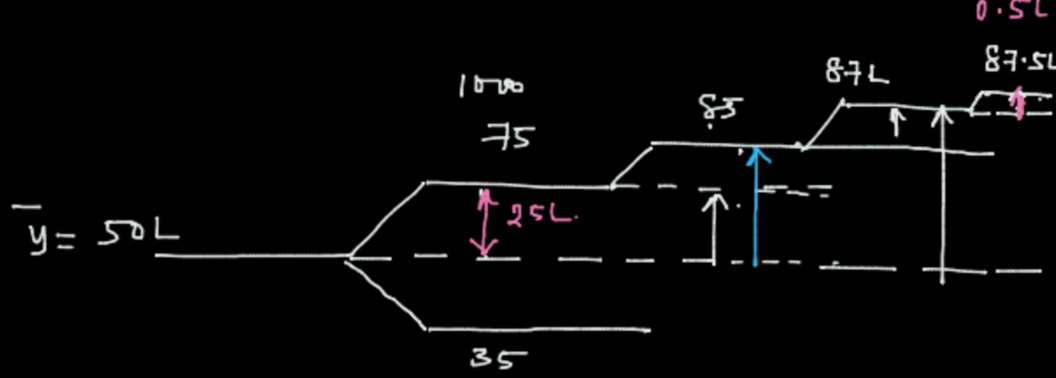
RMSE → 1000.

150 → 1150 ×

± 1000 Rs

4, 7w ✓
→ 5, 46, 750
5, 75.

Area	y
—	50
—	63
—	35
—	47
⋮	⋮
—	95
	$\bar{y}$ = 50L



R^2 → Percentage of Variance explained by the model

Total Var ⇒ ↑ Explained Var + unexplained Vara ↓.

R^2 = Explained Var / Total Variance.

600. $R^2 \geq 85\%$.

R^2 = Explained Var / Explaind Var + unexplained Var = 0

⇒ 1