wiki

(x, y)

learning

mapping will be learnt

Supervised learning

unsupervised

x (input) no outcome

[This learning The pattern from unlabelled data]

↳ clustering [similarity (or distance)

↳ ARM → Association rule mining
freq pattern mining

↳ Anomaly detection → Anomaly

↳ dimensionality reduction

↳ 10000

↳ 10, 20, 30

y → cat (2) classification

↳ patients has heart disease

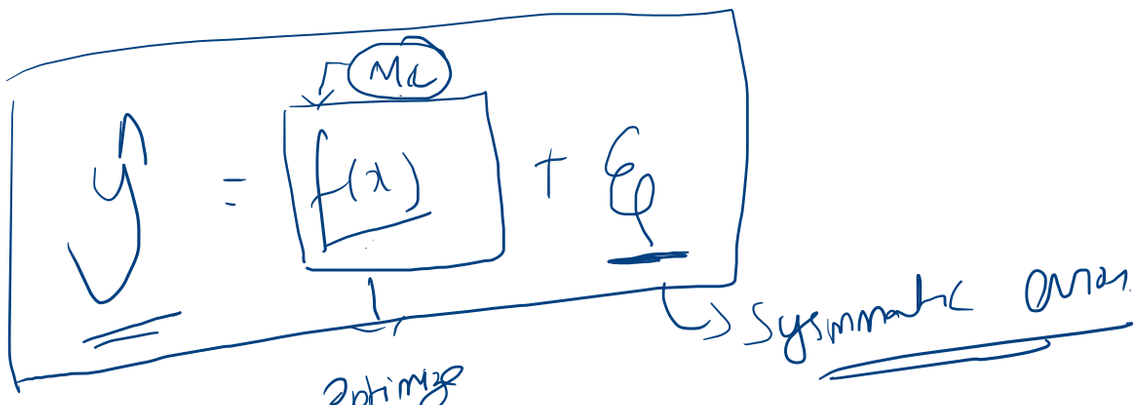
↳ has cancer (or) not.

y → cat (1) regression

↳ hours of study

(vs)

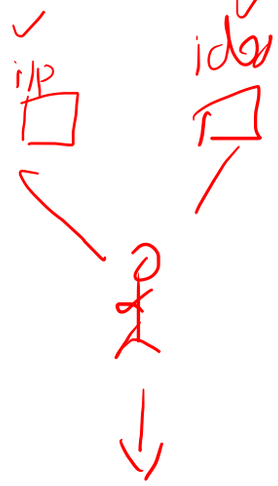
percentage scored



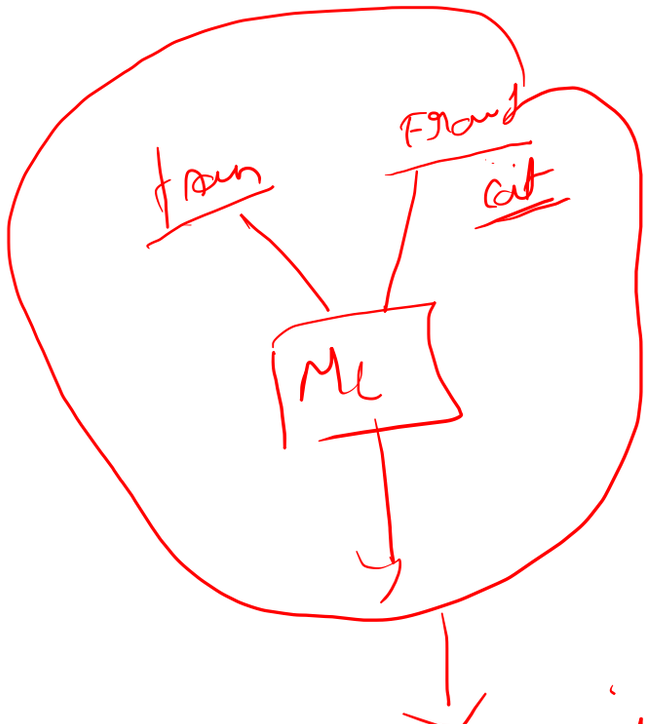
↑ bias
accuracy

ML
↳ error ↓

i/p ✓ idx ✓



Sup (or) unsup?



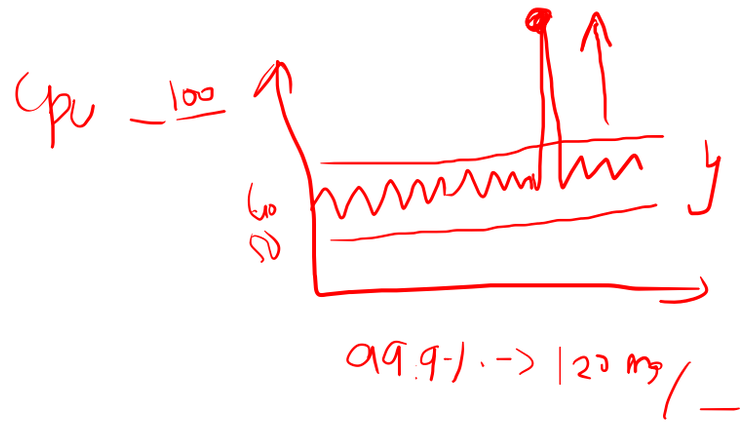
detected ✓ in

Fraud

bank?

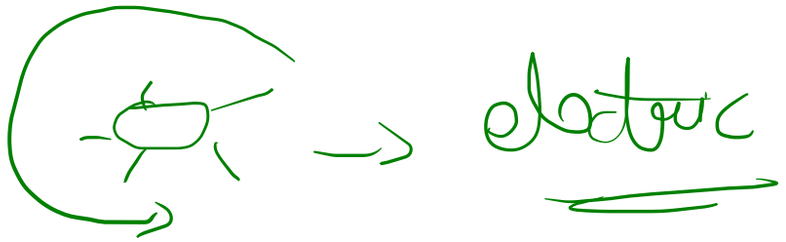
→ classification

AWS server



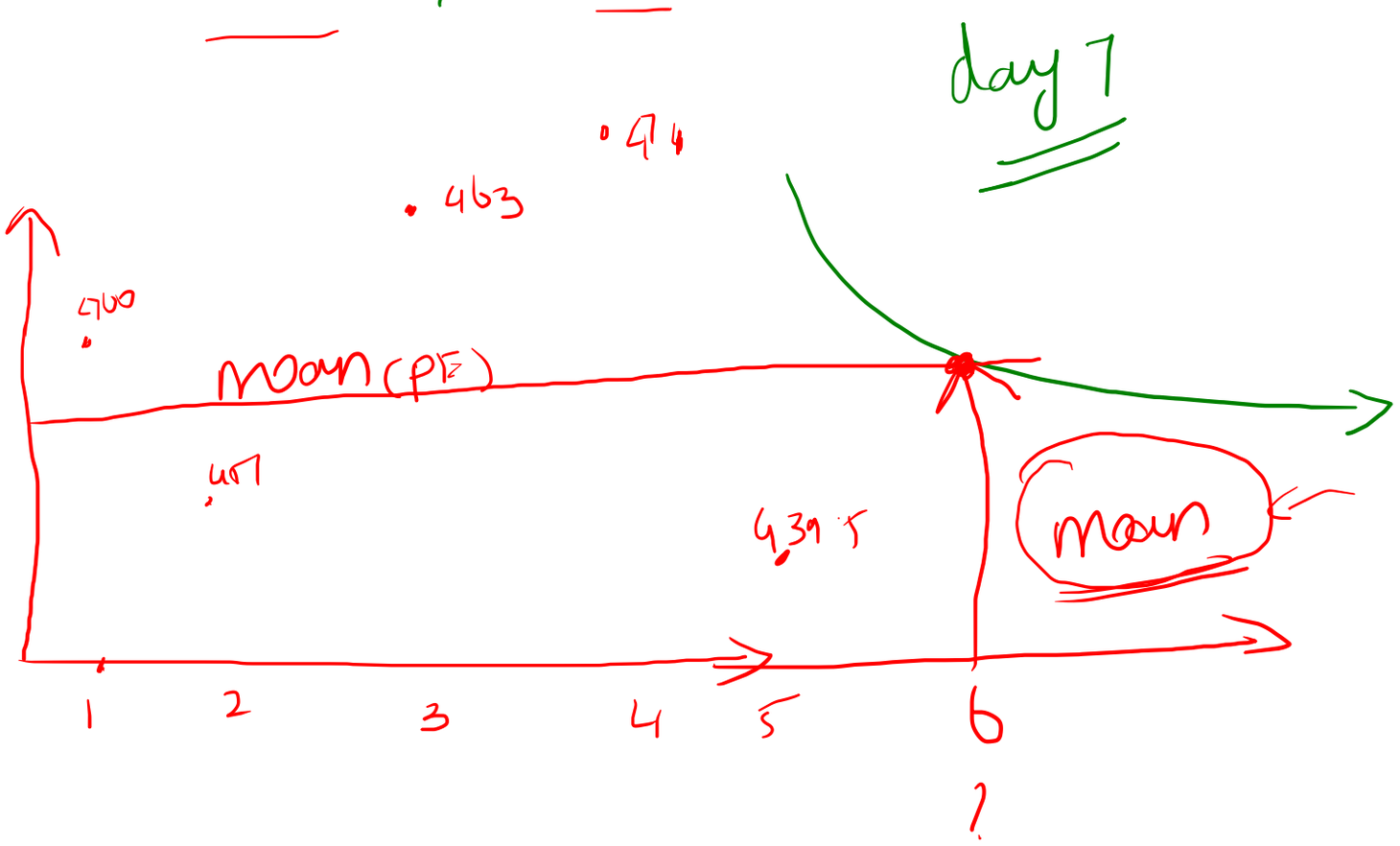
Anomaly → rare!

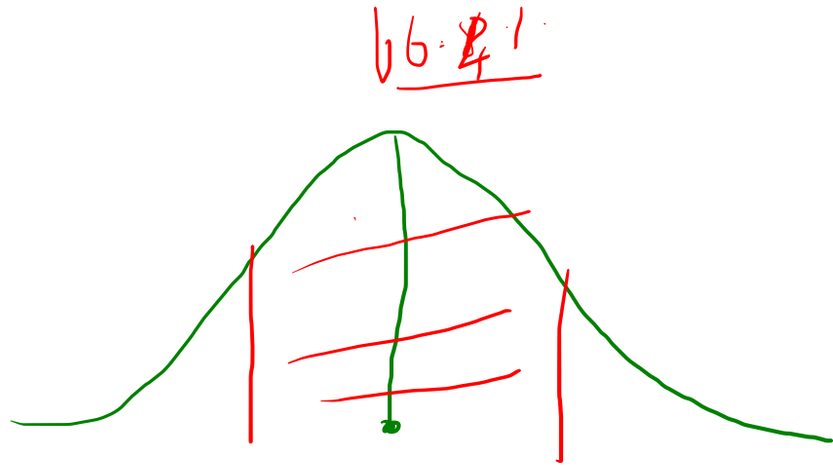
PE → power generation in power plant

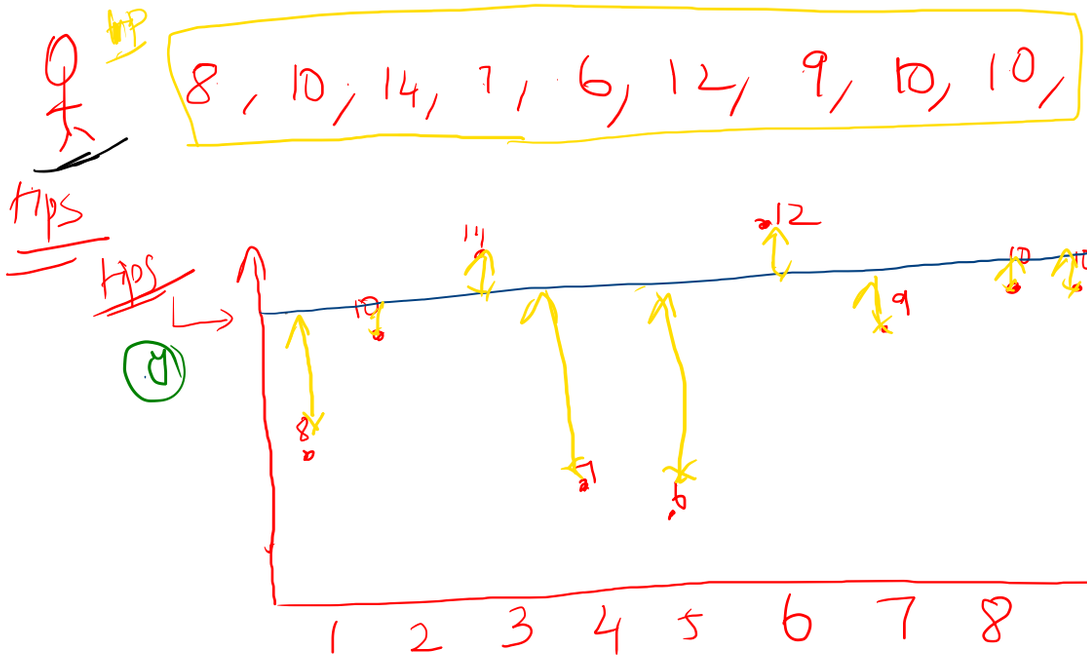


→ 460.2, 451.4, 463.7, 471.7,

439.5, 481.3, ? mean(past day vals)







tips

mean(tips)

11.2

ŷ predict	Actual (y)	Error (residue)
10 → 4.2	9	+ 2.2
11 → 11.2	12	-0.8

predict → reduce error

$SST = SSE$

Sum of Square total

Sum of squared error $= \sum (y - \hat{y})^2$

Minimize Error

$SST = SSR + SSE$

$SST = 0$

$SST \equiv SSE$

	total bill (x)	87	110	135	85	52	80	150	75	100
	tip (y)	8	10	14	7	6	12	9	10	18

(1)

ML Algorithm

tip

?

total bill

supervised

↳ regression

tip y

correlation

↳ linearly increasing

↓

positive correlation

40 50 60 70 80 90 100 110 120 130

total bill x

Linear regression

line → $y = mx + c$

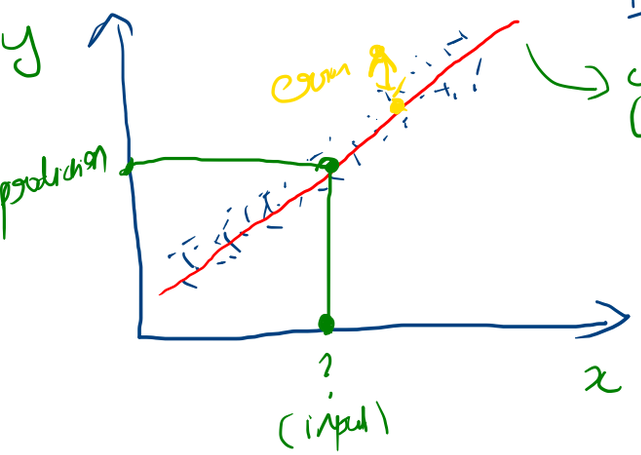
factor

total bills

no relation

Linear regression

$f(x) = \theta_0 + \theta_1 x$



$y = mx + c$

$y = \theta_0 + \theta_1 x$

θ_0 → intercept

θ_1 → coefficient

slope / coefficient

$mx + c$

θ → m → slope / coefficient

y → intercept

$x = 0$

if y = best value

intercept

which best line?

Error in line 3 is less

$\min \sum (y - \hat{y})^2$

Linear regression

↳ simple linear regression

↳ multiple linear regression

least square method

↳ Normal equation

↳ Machine learning math

↳ Gradient descent

↳ SGD