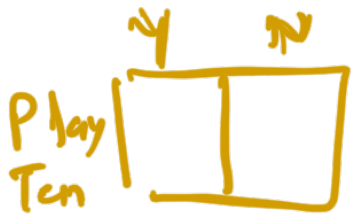
Distribution

- Normal
  - Bernoulli
  - Multinomial
  - Normal (Gaussian)



Bernoulli (Binomial)

$P(\text{success}) = p$

$P(\text{failure}) = q = 1-p$

$x = 1$ [success]

$x = 0$ [failure]

$$P(X = x) = p^x \cdot (1-p)^{1-x}$$

$$P(x) = \begin{cases} p & \text{if } x = 1 \\ q & \text{if } x = 0 \end{cases}$$

MultiNominal Distribution

— Discrete count (Categorical) (Range)

$P(X_1 = x_1, \dots X_k = x_k)$

$$P(X = x_1 \dots x_k) = \frac{n!}{x_1! \dots x_k!} \times p_1^{x_1} \dots p_k^{x_k}$$

eg.

BG	O	A	B	AB
P	0.4	0.42	0.10	0.02

Categorical Multiclass

$Y = \{O, AB, A, B\}$ → 4 = 4/20

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$

$4/20 \quad 6/20 \quad 5/20 \quad 5/20$

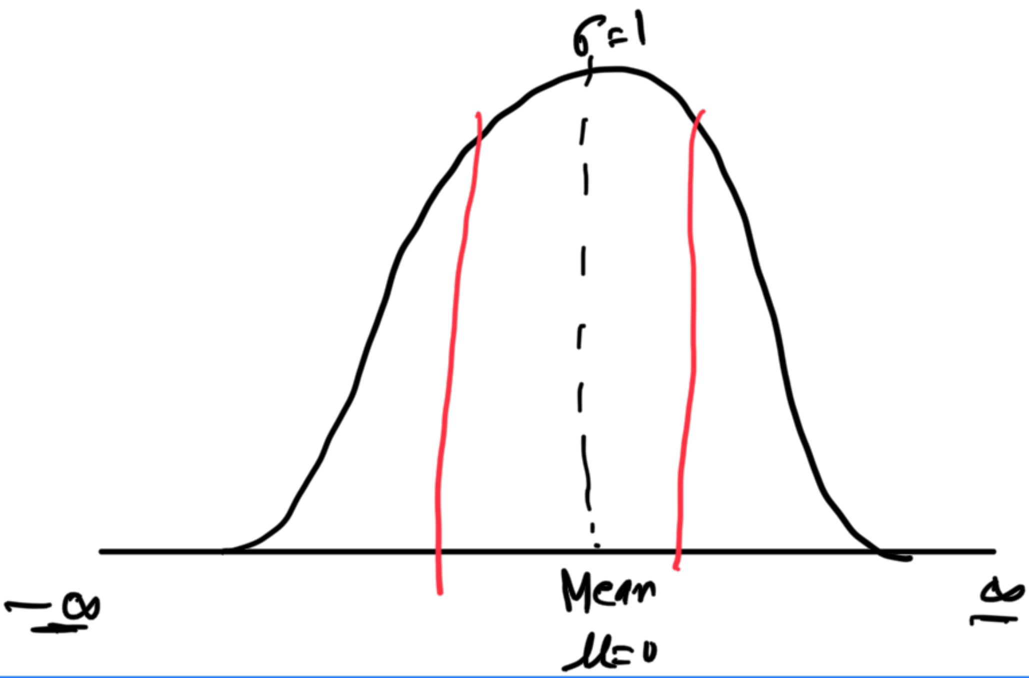
$6 = 6/20$
$5 = 5/20$
$5 = 5/20$
20

Gaussian – Normal

- Discrete
- Continuous

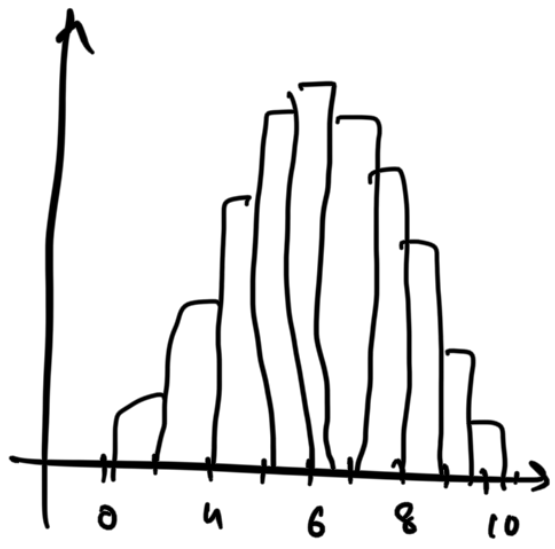
— prob density function (pdf)

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

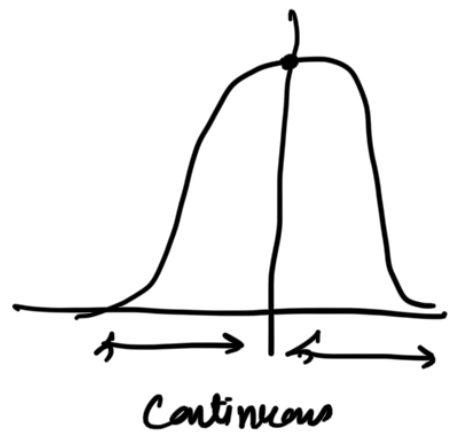


Data →	Discrete ✓		Continuous
	x_1	p_1	
	x_2	p_2	pdf = $f(x)$
	x_3	p_3	
	⋮	⋮	
	x_n	p_n	
	$\sum_{i=1}^{n} p_i = 1$		$\int_{-\infty}^{\infty} f(x)\,dx = 1$
μ	$\sum_{i=1}^{n} x_i p_i$		$\int_{-\infty}^{\infty} x \cdot f(x)\,dx$
σ			

Variance σ^2 $\sum_{i=1}^{} (x_i - \mu) p_i$ $\int_{-\infty}^{} (x-\mu) f(x)\, dx$



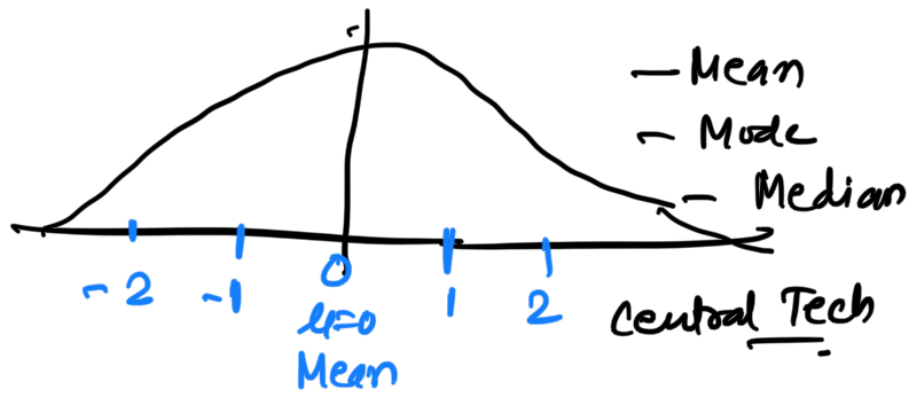
Discreate

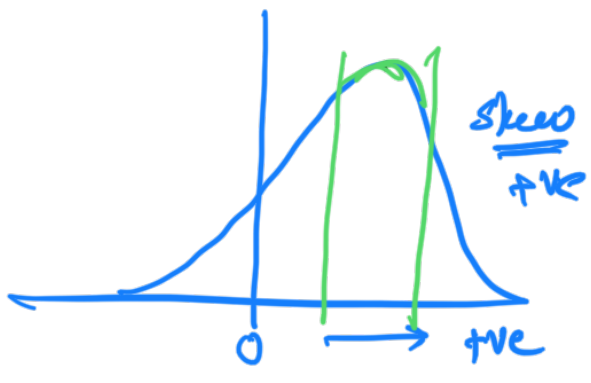


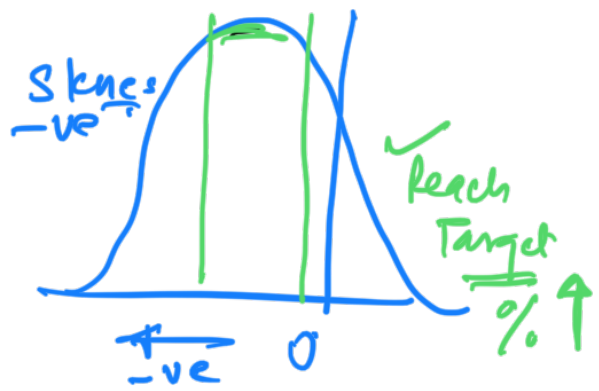
Continuous

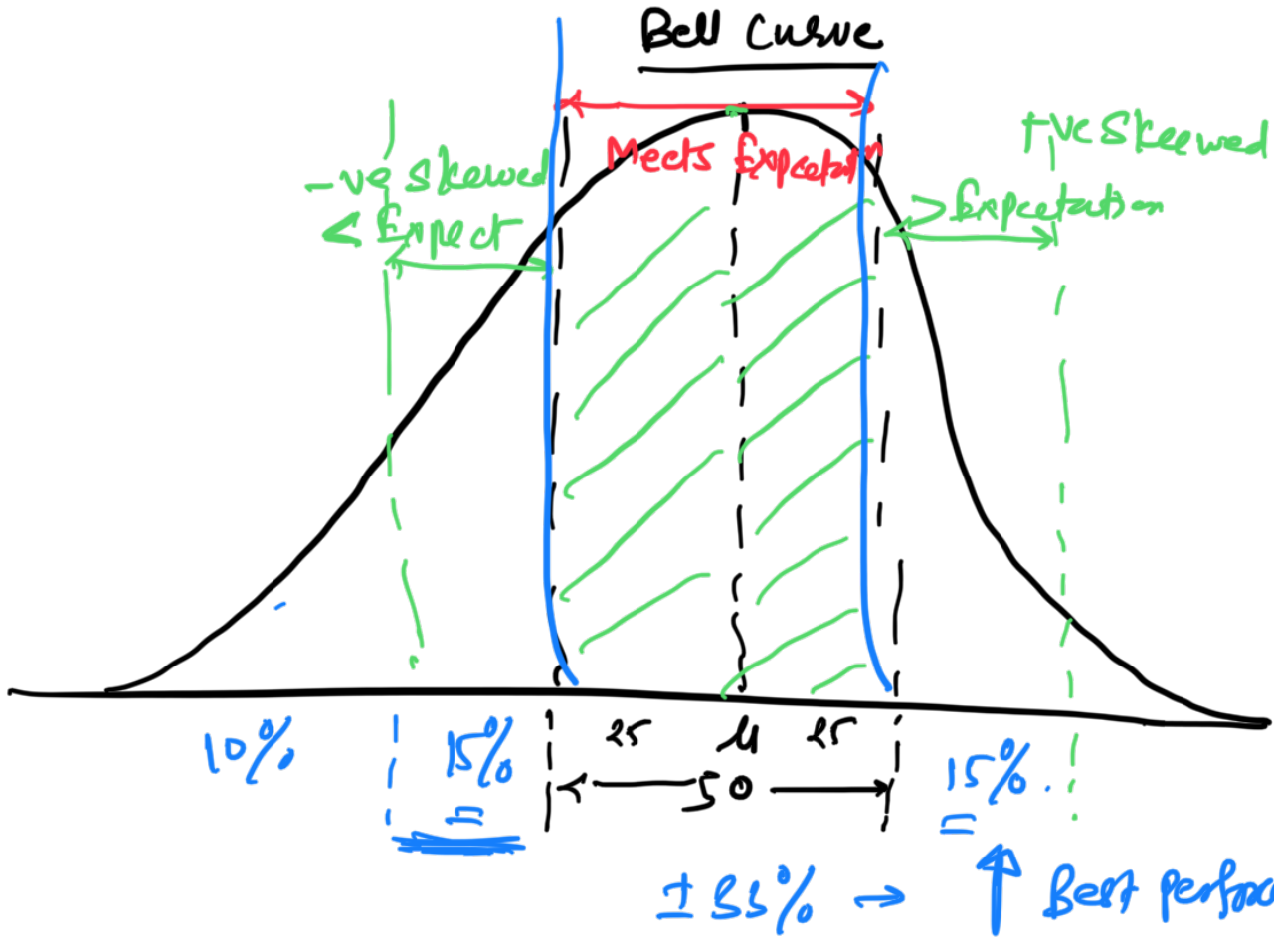
Statistics →

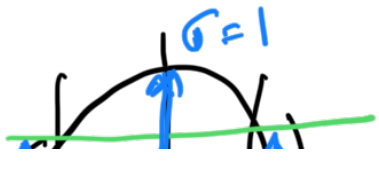
Cage → data

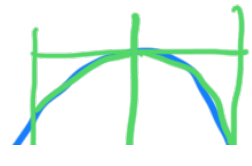


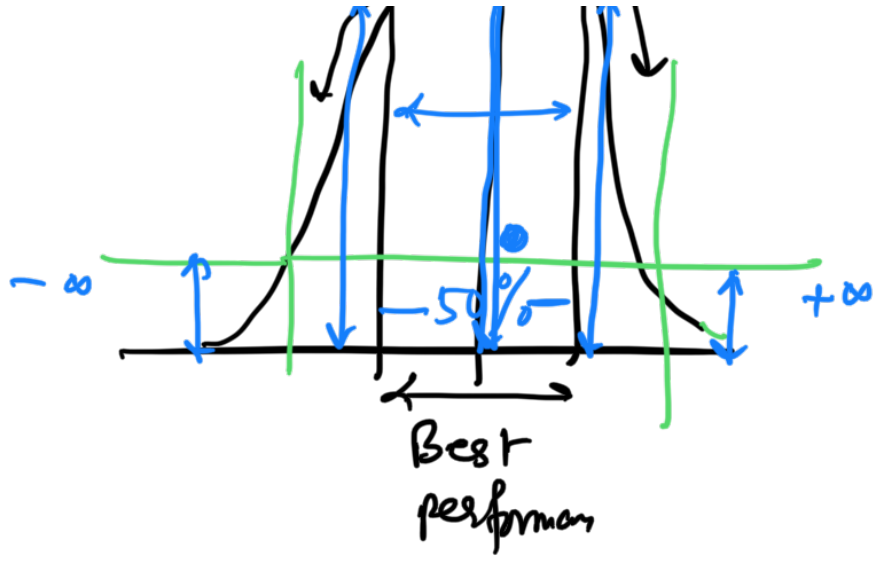


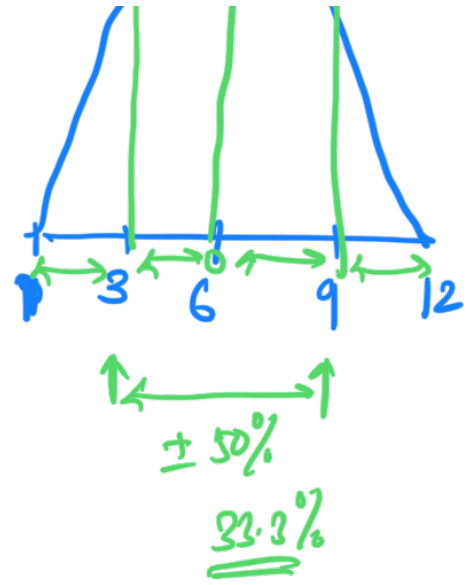


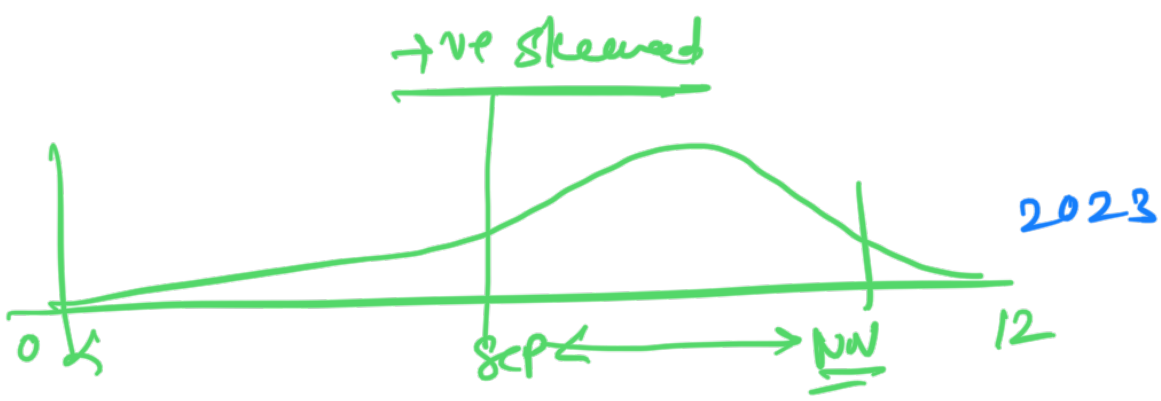


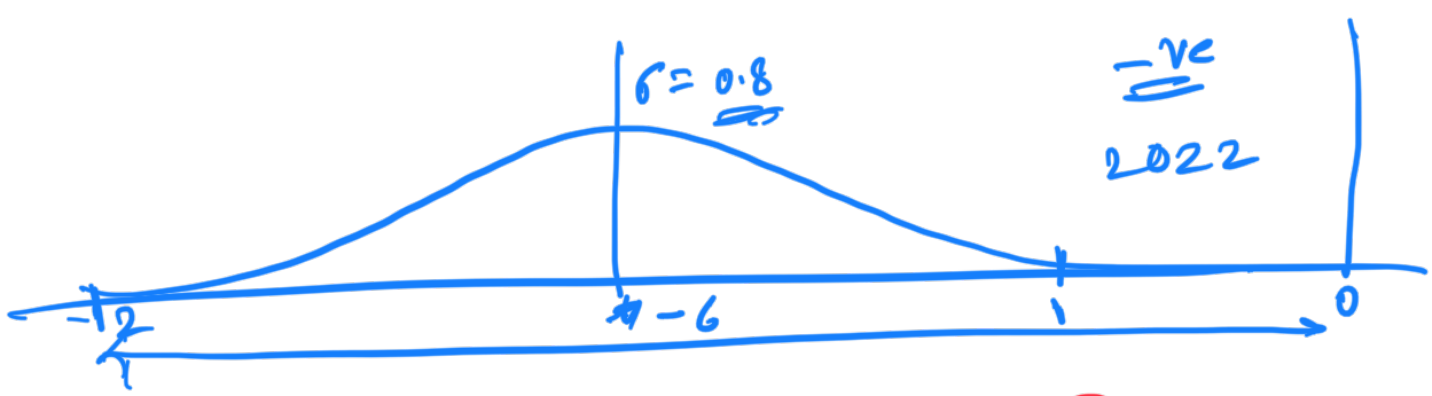




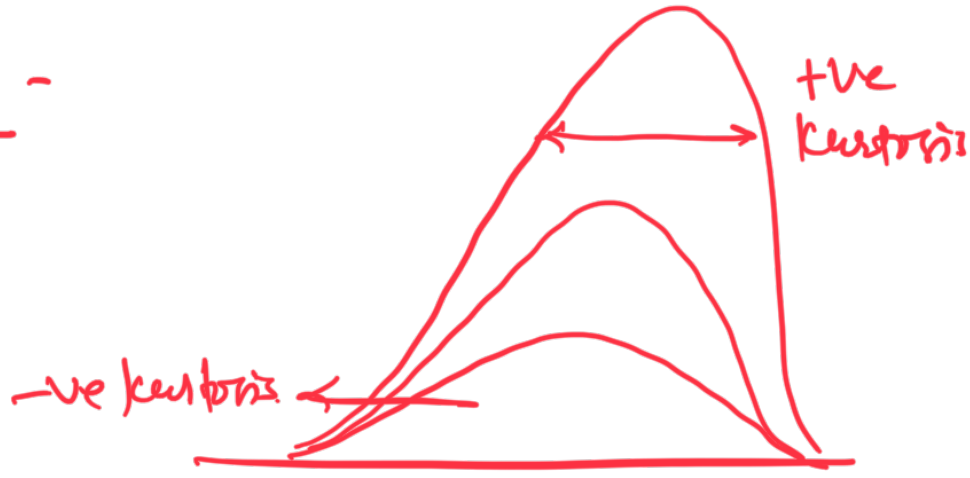


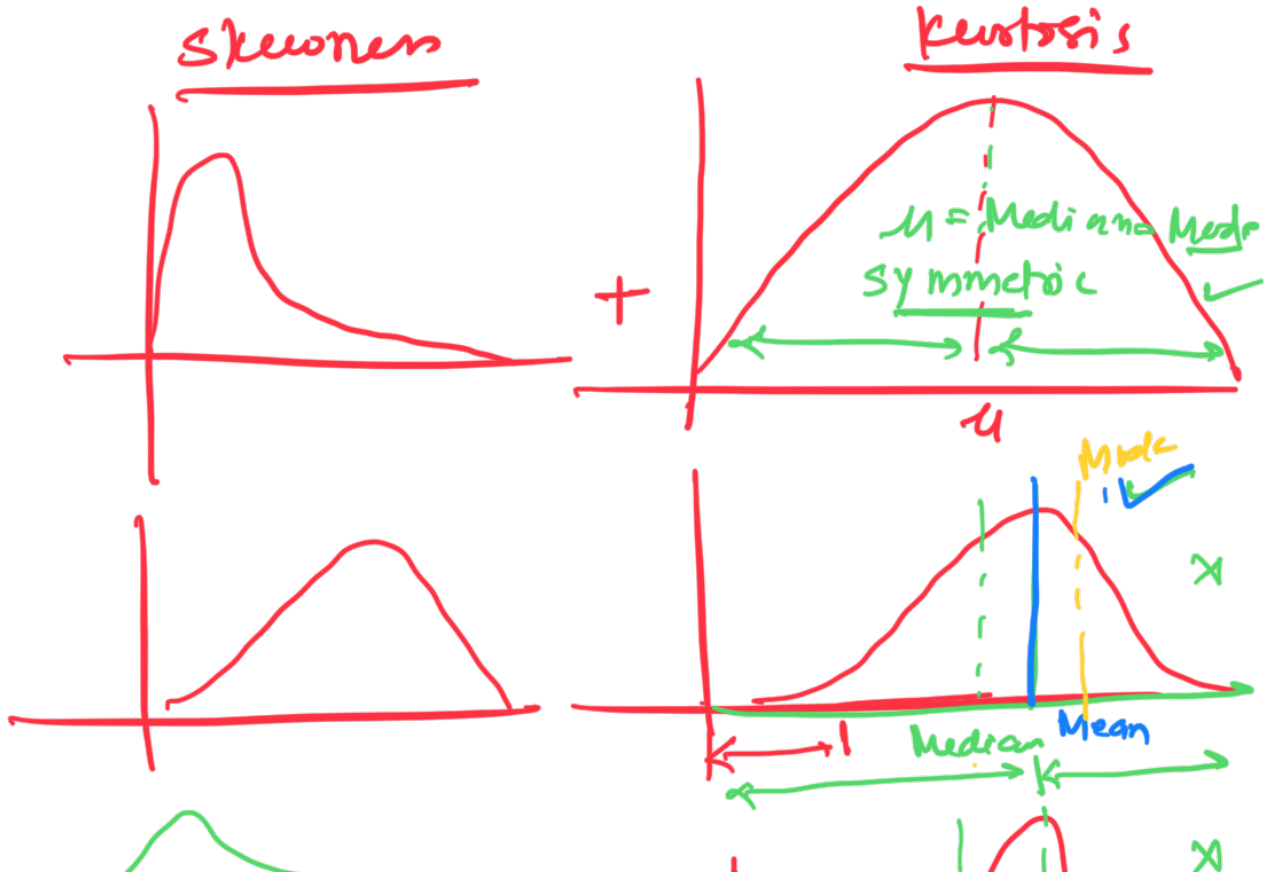






Decision Tree -

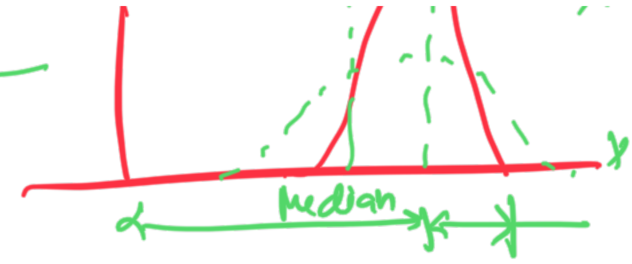




x Mean → ✓

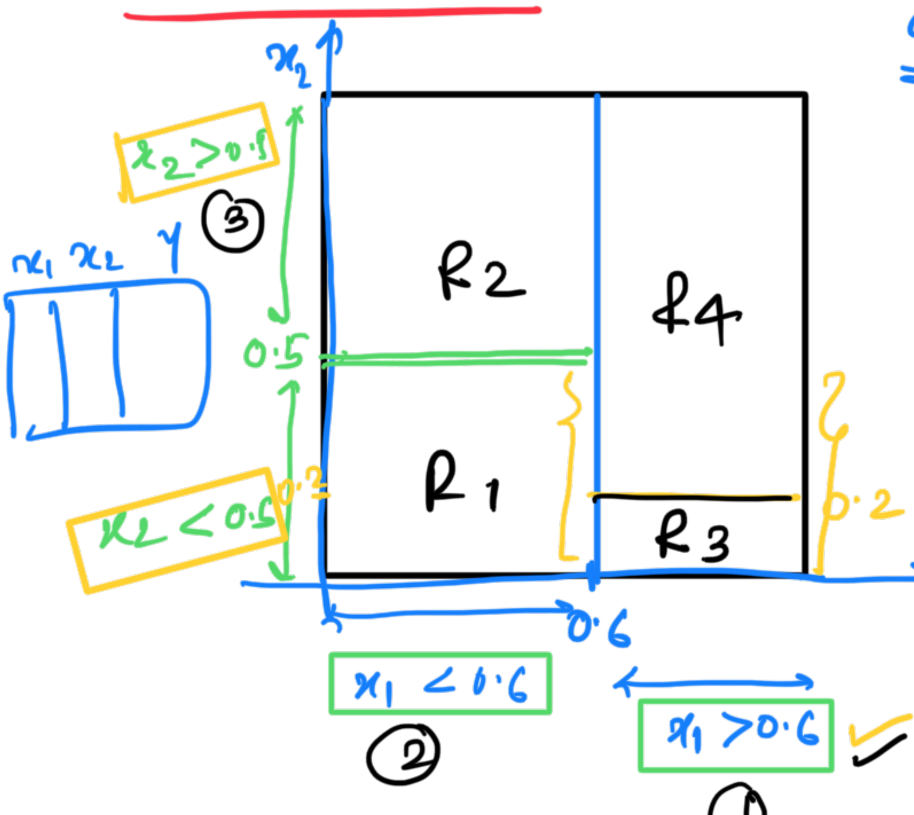
✓ Median → ✓

Mode



Mean < Mode < Median → I skewed

Decision Tree



Objective - Partition of I/P space

Divided any interval regions into subregion

↓

Decision Tree

① $R = \{R_1, R_2, R_3, R_4\}$

$x_2 = 0.5$

④ $x_2 = 0.2$ ✓

Eg Payoff Tables

	Growing	Declining
Mutual funds	53	-5
Bonds	20	
Stocks	70	-13

Probability: 0.4 (Growing), 0.6 (Declining)

Growing — (0.4)

Declining — (0.6)

II Calculate feature Value

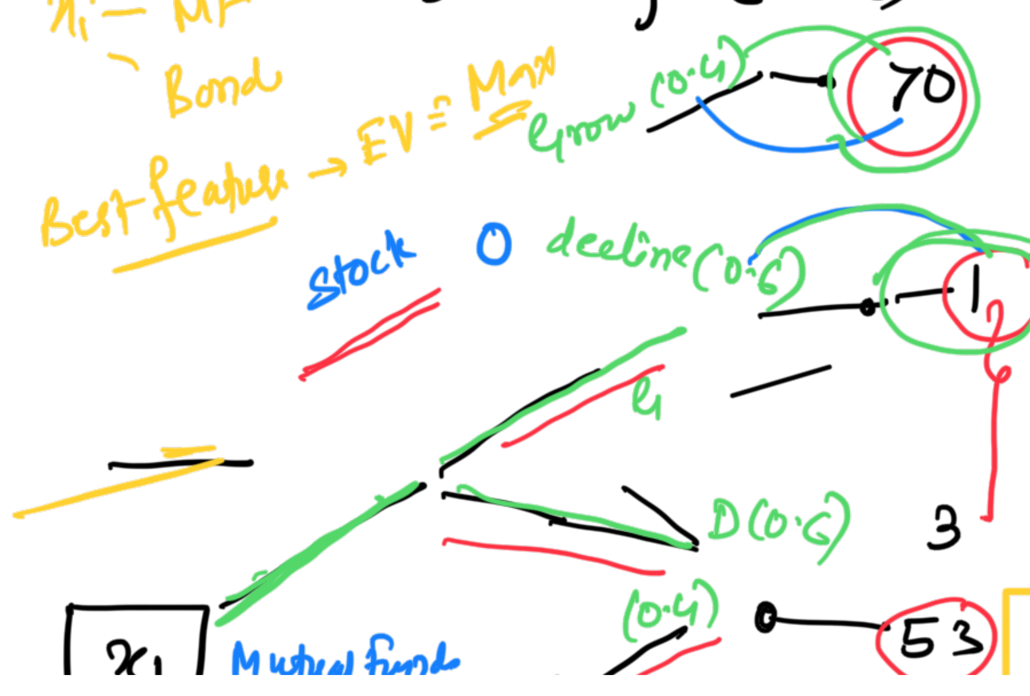
MF = 53×0.4 + -5×0.6

70×0.4 + -13×0.6 = 22.2 → 18

Bonds ⇒ 20

x_1 — Stocks, MF, Bond

Best feature → EV = Max



+ 0.6 × -13

EV(Stock) = 0.4×70

20.2 ✓

EV(MF) = 0.4×53

+0.6 × (-5)

= 18

-5

Bonds

Bonds

20

(same) EV(Bonds) =

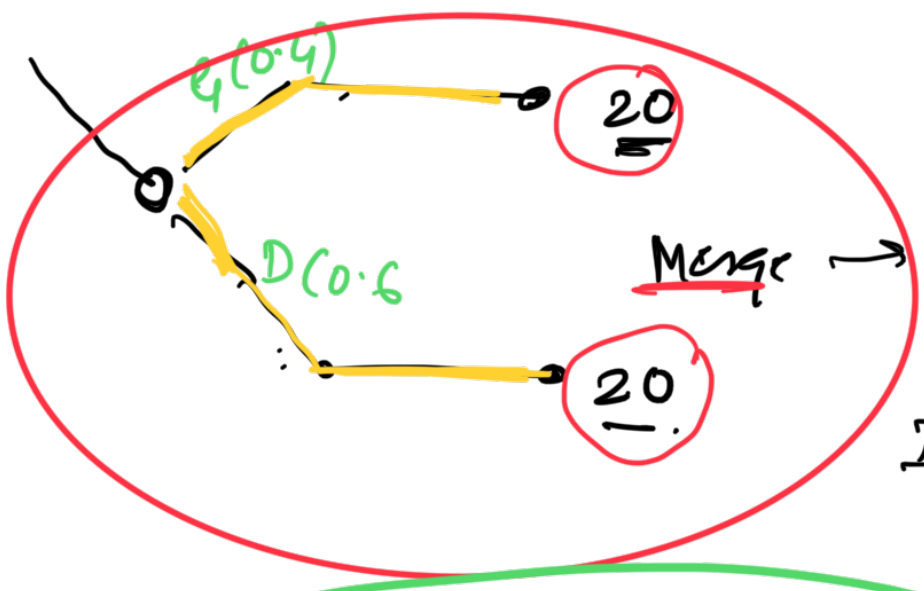
1 × 20 = 20

9/D

(0.6+0.4) = 1

Optimization of our Dec

Tree Pruning



Merge → Common Rule

If-then Rule

Result ⇒

20.02

Stocks

Growing (0.4) = 70

EV (Stock) = 0.4×70 + 0.6×(-13)

= 20.02

Decl (0.6) = -13

x1

Mutual funds

18.2

Growing (0.4) → 53

Decline (0.6) → -5

EV(MF) = 0.4×53 + 0.6×(-5)

= 18.2

Bonds

20

G, D = (0.6+0.4) = 1

20

E(Bonds) = 20×1 = 20

Subset feature

x1 →	Feature Gro	Dcl
Stocks	70	-13
MF	53	-5
Bonds	20	20
Prob	0.4	0.6 ⇒1

x1 {Stock, MF, Bonds}

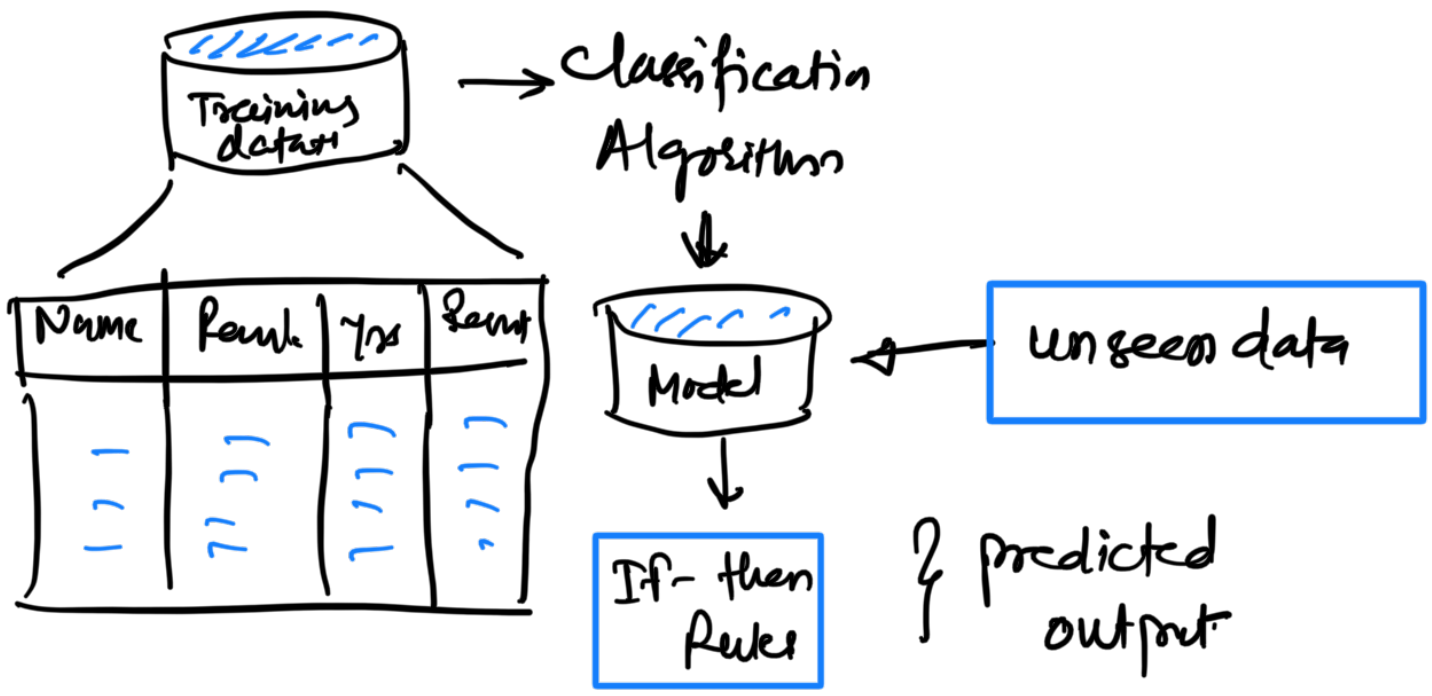
Stocks = 20.02 → Stock > MF & Bonds

MF = 18

Bonds = 20

Important feature

Purpose → Identify the most important features in subset list



Root Node → Start of DT

Node → Condition spe

Age → <30, 30-40, >40

Ex

Age	Income	Student	credit_rating	buys_computer

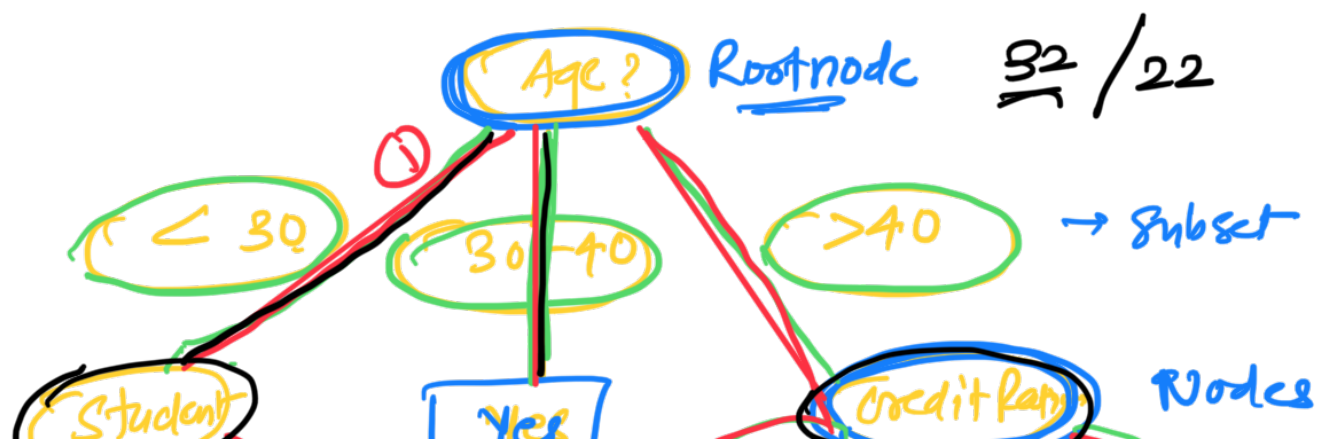
x_1 Age = {<30, 30-40, >40}

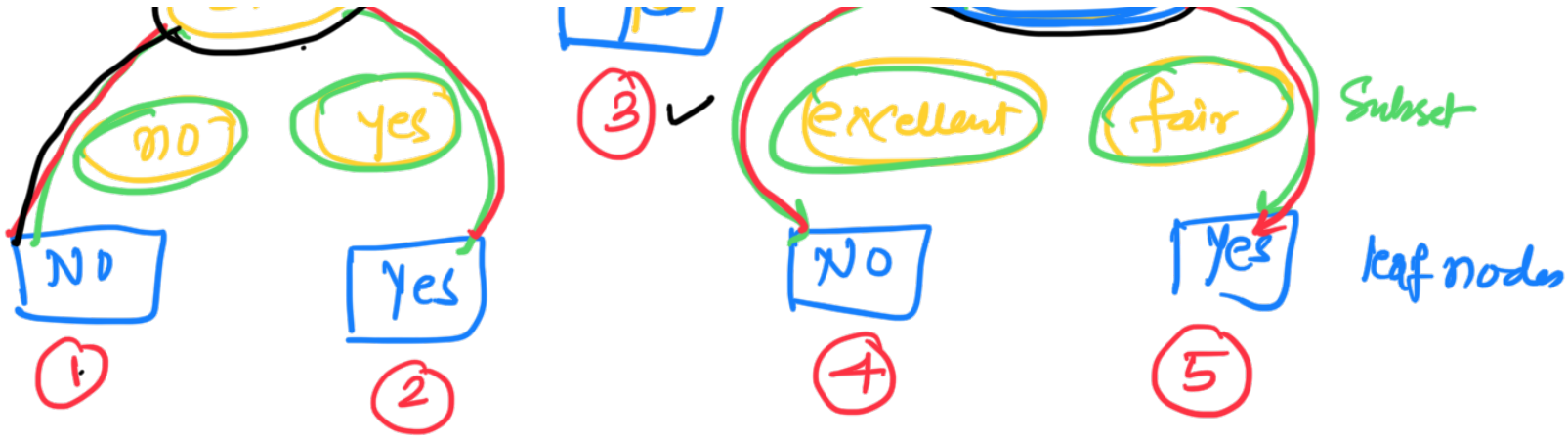
x x_2 Income = {high, medium, low}

x_3 Student = {yes, no}

x_4 credit = {fair, excellent}

$y \leftarrow$ buy = {yes, no} < yes – P(1) / no – P(0)





⟶ Age = 32, Student = no, Income = high.
unknown

Credit rating = fair ⇒ buys = ? yes
NO