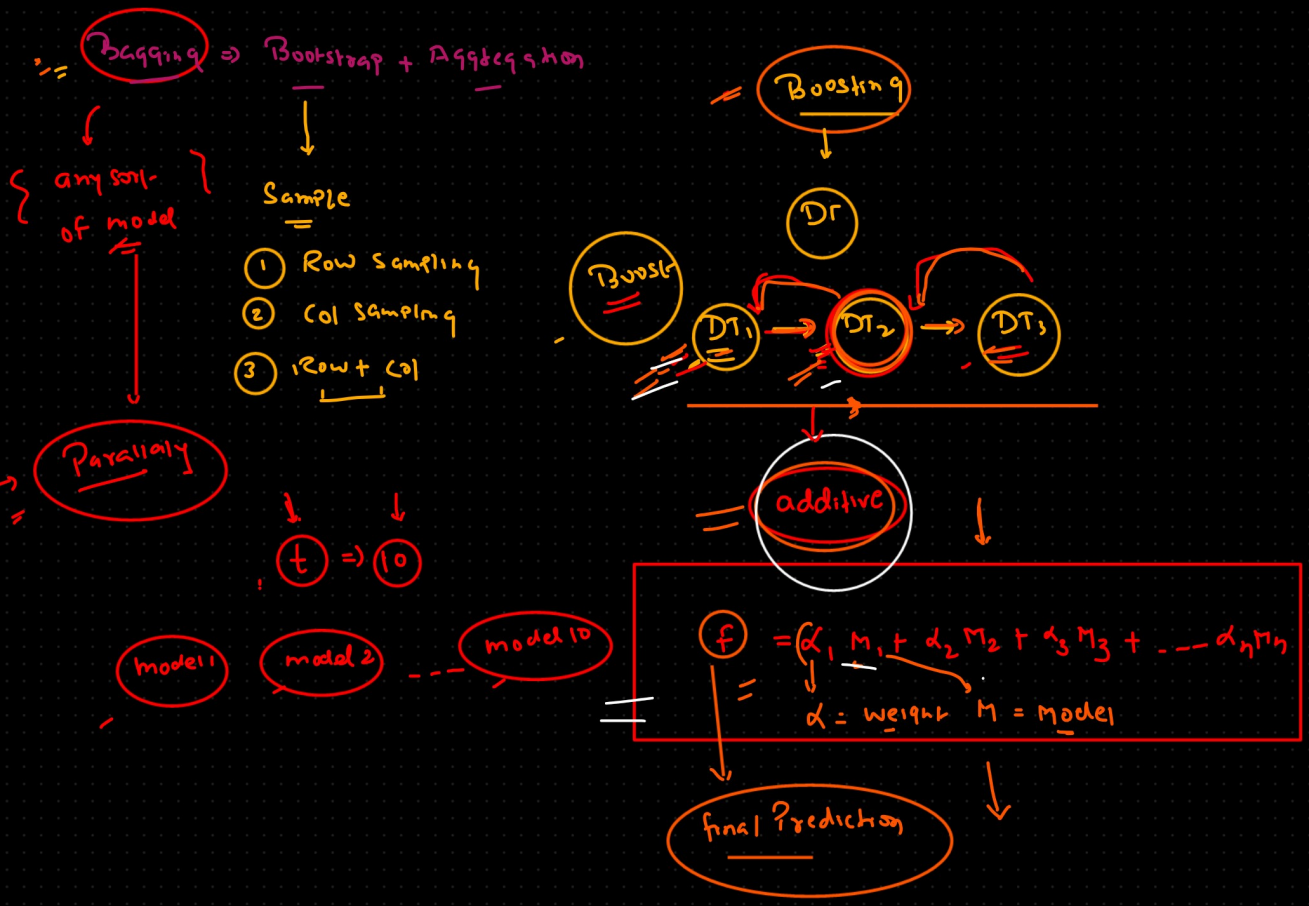
1. Voting
2. Stacking
3. Bagging → Random forest
4. Boosting
   1. AdaBoosting
   2. GBM (Gradient Boosting machine)
   3. XGB
   4. CatBoost

Bagging

1. Parallel technique
2. RF

Boosting

1. Sequential technique
2. AB, GB, XGB



Ensemble =>

Low ↓ Bais ↓ error

low ↓ variance ↓ erro

- Boosting => homogenous

- 10 => Unfitted

1000 => Overfitted

- Adaboosting

1. Regression
2. Classification

Data

Salary	Credit	Approval
≤ 50K	B	NO
≤ 50K	G	yes
≤ 50K	G	Yes
> 50K	B	NO
> 50K	G	Yes
> 50K	N	Yes
≤ 50K	N	NO

(ARF => Gini-index)

1. Base estimator

↓

DT

↓

max_Depth = 1

Stump

- Entropy & Gini-impurity -

⇓

information gain

{ Salary, Credit-score }

Credit score

<= G

(G) Yes

NO (BAD)

[3 YES | 0 NO]

[1 Yes | 3 NO]

DT => Stump

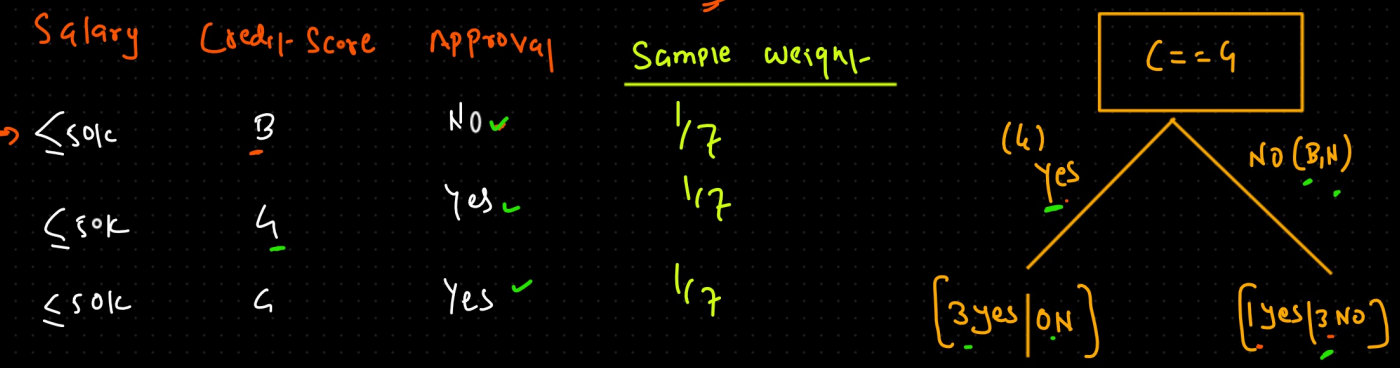
max_depth = 1

(2) Performance of the stump

$\left(\frac{1}{\text{no. of Rows}}\right)$

$(1/n)\ (1/7)$

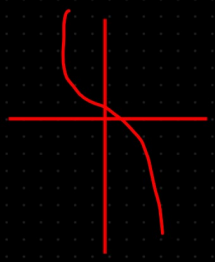
Salary	Credit Score	Approval	Sample weight
→ ≤ 50lc	B	No ✓	1/7
≤ 50K	G	Yes ✓	1/7
≤ 50lc	G	Yes ✓	1/7
> 50lc	B	No ✓	1/7
> 50lc	G ✓	Yes ✓	1/7
> 50lc	N	Yes	1/7
≤ 50lc	N	No	1/7



1/7

several

TE = total error



Performance of the stump = $\left\{\frac{1}{2}\ln\left[\frac{1-TE}{TE}\right]\right\}$

Story

$= \frac{1}{2}\log_e\left[\frac{1-1/7}{1/7}\right]$

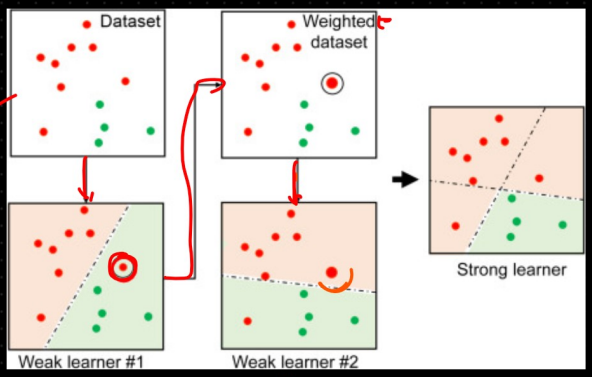
$= \frac{1}{2}\log_e\left[\frac{6}{7}\times\frac{7}{1}\right]$

$= 1/2\ (\log_e(6)) \Rightarrow 0.896$

POS = 0.896

$\alpha_1 \Rightarrow DT_1 \rightarrow$ Stump1

$$\text{final} = \alpha_1(M_1) + \alpha_2(M_2) + \alpha_3(M_3) \ldots \alpha_n(M_n)$$



3 Update the weight of incorrect & correct classified point-

- For correct classified point-

- Weight x e^{-POF}

1/7

incorrect classified point-

= weight x e^{POF}

1/7

POF = 0.896

C.C. $\Rightarrow \frac{1}{7} \times e^{-0.896} = 0.058$

I.C $\Rightarrow \frac{1}{7} \times e^{0.896} = 0.349$

Salary	Credit Score	Approval	Sample weight	Update weight
≤ 50k	B	No	1/7	0.058 / 0.697
≤ 50k	4	Yes	1/7	0.058 / 0.697
≤ 50k	4	Yes	1/7	0.058 / 0.697
> 50k	B	No	1/7	0.058 / 0.697
> 50k	4	Yes	1/7	0.058 / 0.697
> 50k	N	Yes	1/7	0.349 / 0.697
≤ 50k	N	No	1/7	0.058 / 0.697

= 1 = 0.697

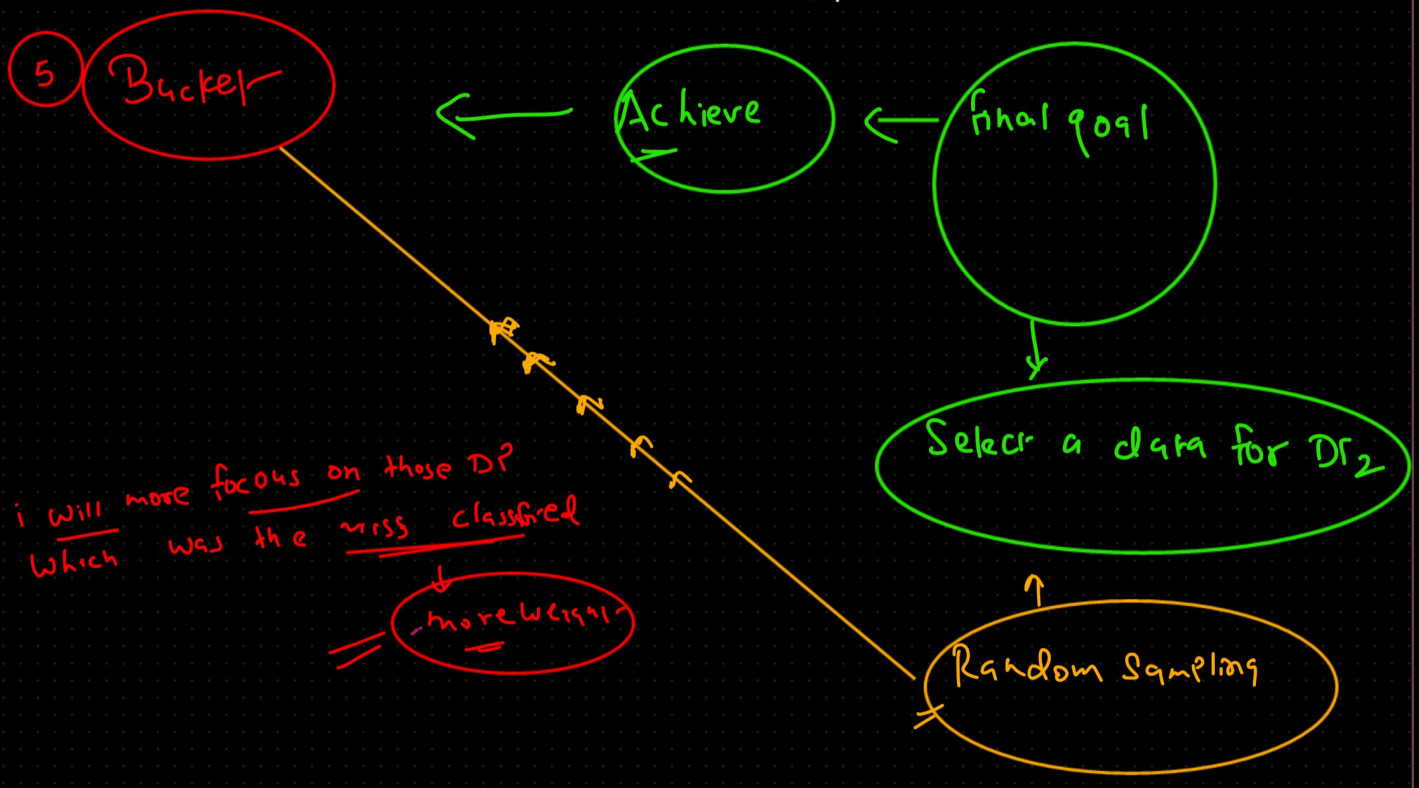
④ Normalize the weight.

[Addition = 1]

↑

{ addition for all the weight will be = 1 }

Salary	Credit Score	Approval	Sample weight	Update weight	Normalize weight
≤ 50k	B	No	1/7	0.58 / 0.697	0.083
≤ 50k	4	Yes	1/7	0.058 / 0.697	0.083
≤ 50k	4	Yes	1/7	0.058 / 0.697	0.083
> 50k	B	No	1/7	0.058 / 0.697	0.083
> 50k	4	Yes	1/7	0.058 / 0.697	0.083
> 50k	N	Yes	1/7	0.349 / 0.697	= 0.50
≤ 50k	N	No	1/7	0.058 / 0.697	0.083



0-1

Salary	Credit Score	Approval	Sample weight	update weight	Normalize weight	Bucket (Bin assignment)
			1/7	0.058	0.083	0 - 0.083
			1/7	0.058	0.083	0.083 - 0.16
			1/7	0.058	0.083	(0.16 - 0.24)
			1/7	0.058	0.083	(0.24 - 0.32)
			1/7	0.058	0.083	0.32 - 0.40
			1/7	0.349	0.500	0.40 - 0.90
			1/7	0.058	0.083	0.90 - 0.96

6th

0.6

Random Sampling ⇒ DT_2

(0 - 1) ⇒ D.P ⇒ 7

[0.1, 0.4, 0.7, 0.05, 0.8, 0.3, 0.2]

Salary Credit APPROVAL ⇒ DT_2

3
6
6
1
6
4
3

⇒ Importance