Boosting Algorithms

↳ Sequential Weak Learners

① Adaboost

Stump

Underfitting

Boosting { Low Bias ⇐ High Bias
Low Variance ⇐ Low Variance/High Variance

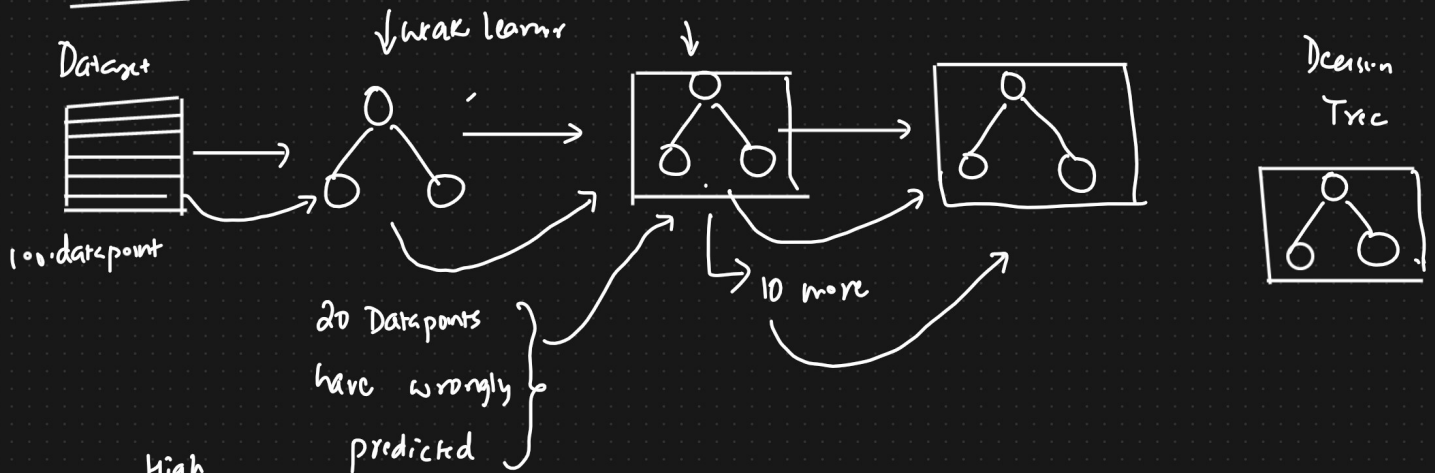
Training Data Acc ↓↓ 40%
Test Data Acc ↑↑ 45%

Weak Learners ⟶

Random Forest ⟶ Majority Voting Classifier [Classification]
Average of o/p [Regression]

Adaboost ⟶ Weak Learners ⟶ Add the o/p of the weak learners with some weights assigned to it

Adaboost



High ↓ ↓-ve

$$f = \alpha_1(M_1) + \alpha_2(M_2) + \alpha_3(M_3) + \cdots + \alpha_n(M_n)$$

$M_1, M_2, M_3 \cdots M_n$ ⟶ Weak Learners ⟶  Depth = 1

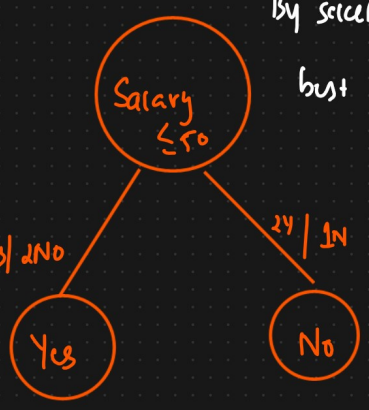
$\alpha_1, \alpha_2, \alpha_3 \cdots \alpha_n$ ⟶ Weights

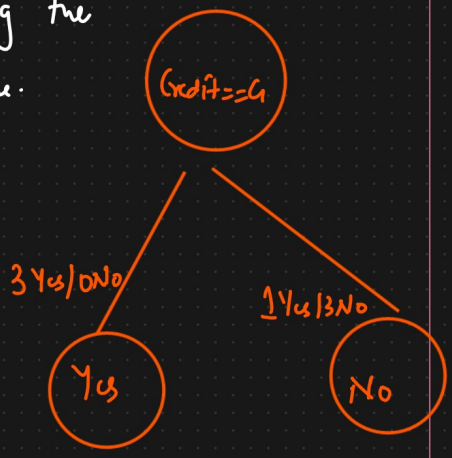
①

Salary	Credit	Approval	Weights
<=50K	B	No	1/7
<=50K	G	Yes	1/7
<=50K	G	Yes	1/7
>50K	B	No	1/7
>50K	G	Yes	1/7
>50K	N	Yes	1/7
<=50K	N	No	1/7

① We create Decision Tree Stump

By selecting the best one.





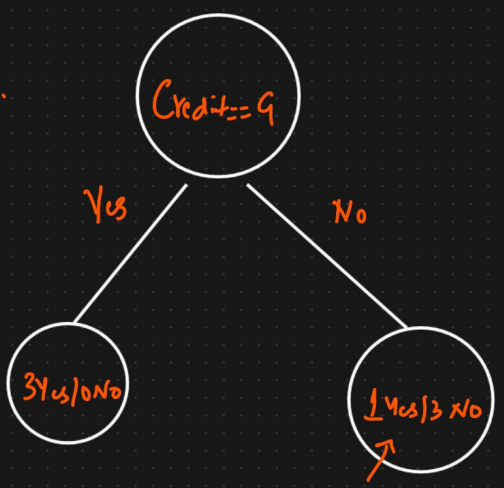
Entropy or Gini

$$H(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$

②

Salary	Credit	Approval	Weights
<=50K	B	No	1/7
<=50K	G	Yes	1/7
<=50K	G	Yes	1/7
>50K	B	No	1/7
>50K	G	Yes	1/7
>50K	N	Yes	1/7
<=50K	N	No	1/7

B => Bad
G => Good
N => Normal



② Calculate the Total Error [Add the weights of wrong data point]

T.E = 1/7

③ Performance of Stump $= \frac{1}{2} \ln\left[\frac{1-TE}{TE}\right] = \frac{1}{2} \ln[6] \approx 0.896$

$$f = \alpha_1(m_1) + \alpha_2(m_2) + \alpha_3(m_3) + \ldots \alpha_n(m_n)$$

$f_1 = 0.896$

④ Update the weight for correctly and Incorrectly data points

Salary	Credit	Approval	weights	Update weights
<=50K	B	No	1/7 (0.14) →	0.058
<=50K	G	Yes	1/7	0.058
<=50K	G	Yes	1/7	0.058
>50K	B	No	1/7	0.058
>50K	G	Yes	1/7	0.058
→ >50K	N	Yes	1/7 <	0.349
<=50K	N	No	1/7	0.058

10K, 20K

For correctly classified point

= weight * $e^{-\text{Performance}}$

= $\frac{1}{7} * e^{-(0.896)}$

= 0.058

For Incorrect classified point

= weight * $e^{\text{Performance}}$

= $\frac{1}{7} * e^{(0.896)}$

= 0.349

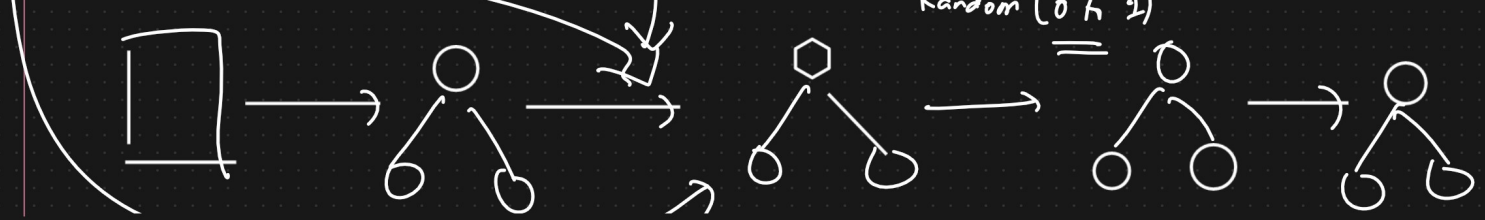
⑤ Normalize weights and Asign Bins

0-1 → 0.12 0.56 0-1 0.09

Salary	Credit	Approval	weights	Update weights	Normalized weight	Bins Assignment
<=50K	B	No	1/7 (0.14) →	0.058 ÷ 0.697	0.08	0 – 0.08
<=50K	G	Yes	1/7	0.058 ÷ 0.697	0.08	0.08 – 0.16
<=50K	G	Yes	1/7	0.058 ÷ 0.697	0.08	0.16 – 0.24
>50K	B	No	1/7	0.058	0.08	0.24 – 0.32
>50K	G	Yes	1/7	0.058	0.08	0.32 – 0.40
→ >50K	N	Yes	1/7 <	0.349	0.50	→ 0.40 – 0.90
<=50K	N	No	1/7	0.058	0.09	0.90 – 1
			1	0.697	≈ 1	

for Ran(0,1)

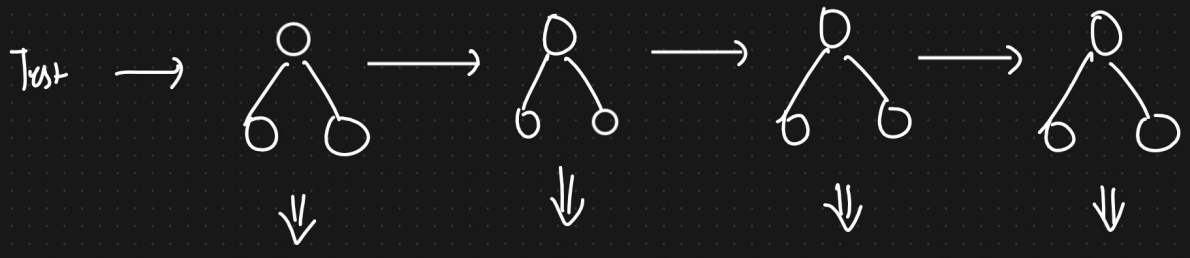
Random (0 to 1)



$$\alpha_1(M_1) + \alpha_2(M_2) + \alpha_3(M_3) + \cdots \alpha_n(M_n)$$

④ Final Prediction

Test $(\leq 50x, G)$



↳ $\alpha_1 = 0.896$ $\alpha_2 = 0.650$ $\alpha_3 = 0.38$ $\alpha_4 = 0.20$

$$f = \alpha_1(M_1) + \alpha_2(M_2) + \alpha_3(M_3) + \alpha_4(M_4)$$

$$= 0.896(\text{Yes}) + 0.650(\text{No}) + 0.38(\text{Yes}) + 0.20(\text{No})$$

$$= 1.2(\text{Yes}) + 0.85(\text{No})$$

⇓

Yes