Bagging

df

Sdf   Sdf Sdf Sdf   Sdf

1   1   0   0   0

$$\underline{\underline{0}}$$

Boosting

df

Sdf
|
err

Sdf-2
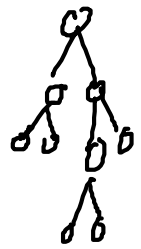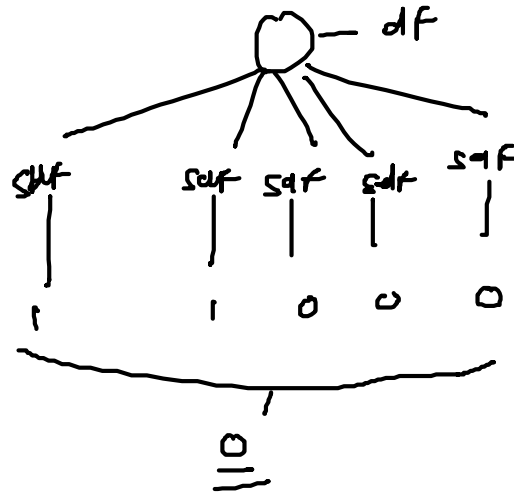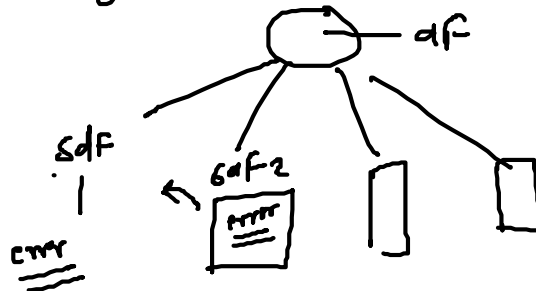
SAMME
(Stagewise Additive Multimodeling using Multiclass
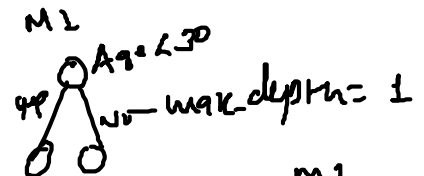Exponential Loss Function)

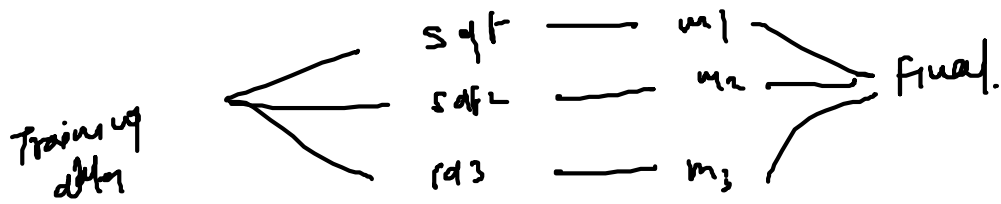→ Adaptive boosting

↘ Adaboost

→ Decision stumps
level of tree = 1

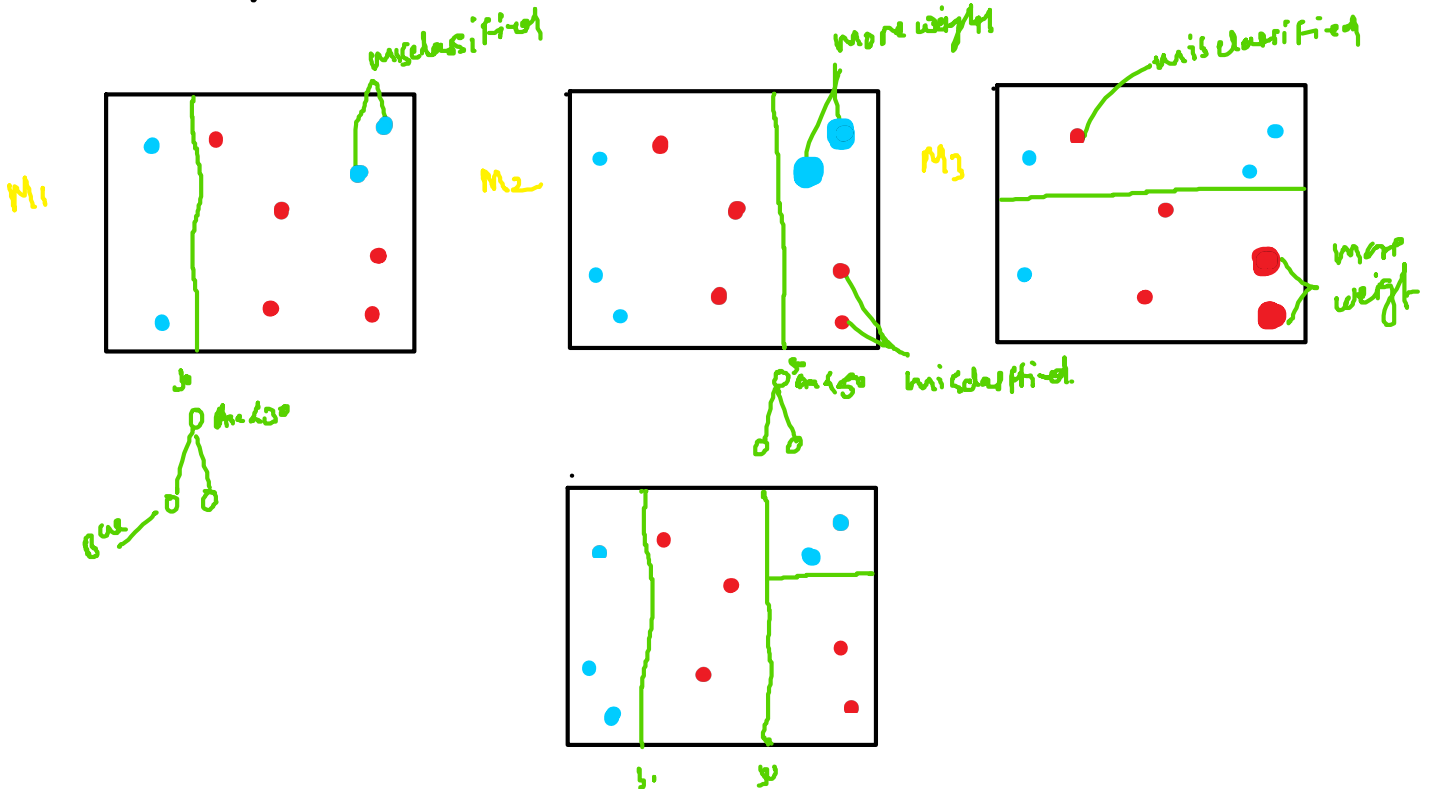Forest of stumps
A tree with a single rootnode &
2 leaf nodes

→ Weak learner | base learner

M1
Age <30
yes   no — max depth = 1

M2
M | F
yes   no
m   F

Bagging

Training data
- Sdf — m1
- Sdf2 — m2 — Final.
- sd3 — m3

Boosting

Training data — Sdf — m1
→ move out to error — m2 — final
weighted sample — m2
weighted — m3

M1 — misclassified

M2 — more weight — misclassified

M3 — misclassified — more weight
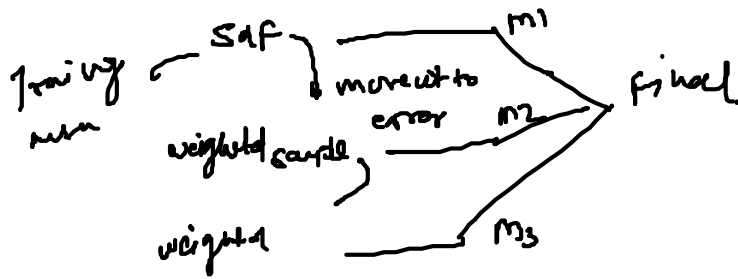
Dataset

	x next af						
Sr.no.	Age	gender	BP	Target	wt	Age pd pred	update	normalize wt
1	63	1	122	1	0.14	1	0.08	0.094
2	37	1	100	0	0.14	0	0.08	0.09
3	41	0	85	0	0.14	1	0.22	0.26
4	57	1	125	1	0.14	1	0.08	0.09

more weight

normalize wt

3	41	0	85	0	0.14	1				
4	52	1	125	1	0.14	1	0.08	0.09		market.
5	57	0	130	0	0.14	0	0.08	0.09		
6	85	1	80	1	0.14	0	0.22	0.16		
7	52	0	95	1	0.14	1	0.08	0.09		

Misclassified
0.84 (circled)
1

1) weight: $\frac{1}{n} = \frac{1}{7} = 0.14$

2) select 1 feature, lowest Gini / highest IG.

Age ≥ 45 (tree diagram)

$$\text{Total error} = \frac{2}{7} = 0.28$$

3) performance of stump:

$$\alpha = \frac{1}{2} \times \log_e\left(\frac{1-TE}{TE}\right)$$

$$= \frac{1}{2} \times \log_e\left(\frac{1-0.28}{0.28}\right)$$

$$= \frac{1}{2} \times \log_e(2.57)$$

$$= \frac{1}{2} \times 0.94$$

$$\alpha = 0.47$$

4) new weight

Misclassified = +
correctly classified = −

new weight = old weight $\times e^{\pm(\alpha)}$

new weight (+) = $0.14 \times e^{+0.47}$

$$= 0.14 \times 1.6$$

$$= 0.22$$

new weight (−) = $0.14 \times e^{-0.47}$

$$= 0.14 \times 0.62$$

$$= 0.08$$

normalise the wts

$(-)$ $\dfrac{0.68}{0.84}$: 0.094

$(+)$ $\dfrac{0.n}{0.84}$ = 0.16

hi