zoom.us Meeting View Edit Window Help
zoom 56% Thu 28 Dec 6:17 PM
You are screen sharing Stop Share

Collection of Base Models ← Ensemble Models

I/P → [M1 M2 M3] → O/P

Ensemble

↳ A Collection of things

1. Bagging } → Homogeneous
2. Boosting }
3. Stacking — Heterogeneous

f1 f2 f3 ... fd y

1
2
3
⋮
n = 1000

$D_n^{f_1 f_2 \cdots f_d}$

1. Bagging — Reduces Variance

↳ Base Models

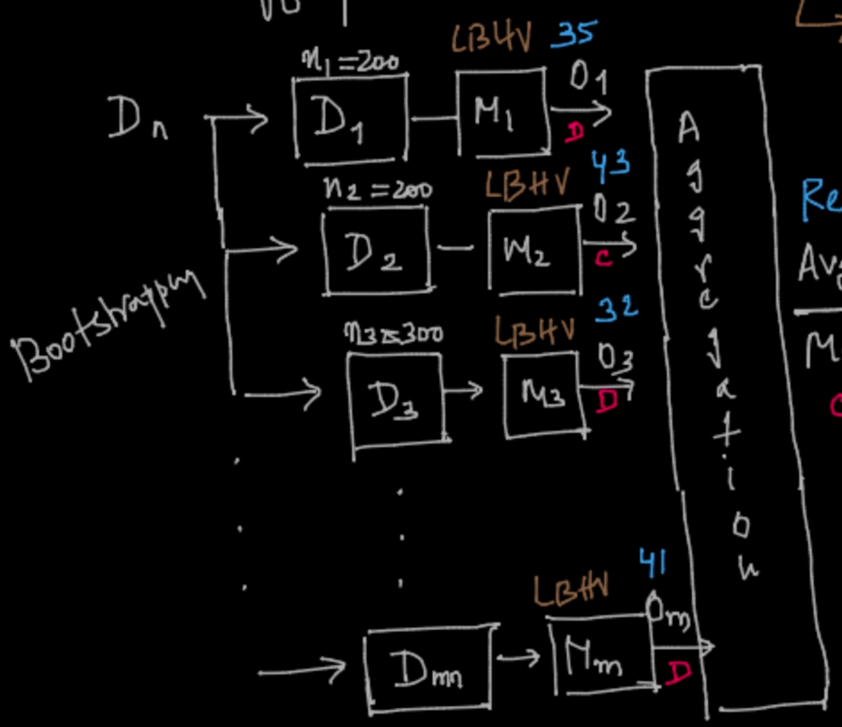
High Variance + Low Bias

Bootstrapping — Sampling with replacement

Bootstrapped Aggregation-

Bagging

Bias & Variance

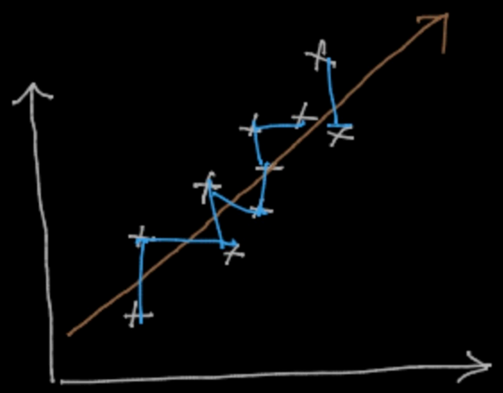


zoom.us Meeting View Edit Window Help
zoom 31% Thu 28 Dec 7:16 PM
You are screen sharing Stop Share

Bias → Training Error

↓ Bias → 0

Bias ↑

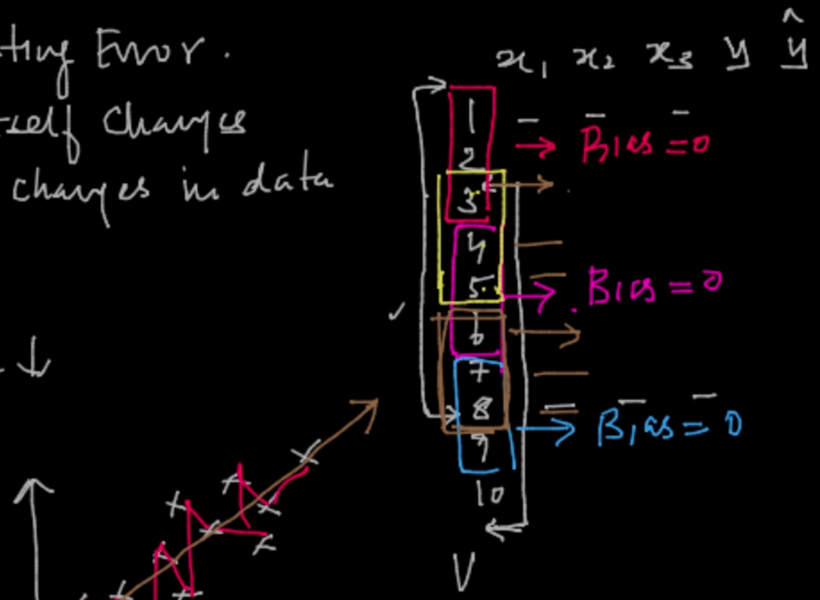


Variance — Testing Error.

↑ The model itself changes with small changes in data

Variance ↓

Bias ↓, Low Variance ↓



Bagging Reduces Variance → Base Models -

Low Bias & High Variance

Homogeneous Ensembles

Bagging Regressor (M_1 —]

10 ppl → Predicting House price-

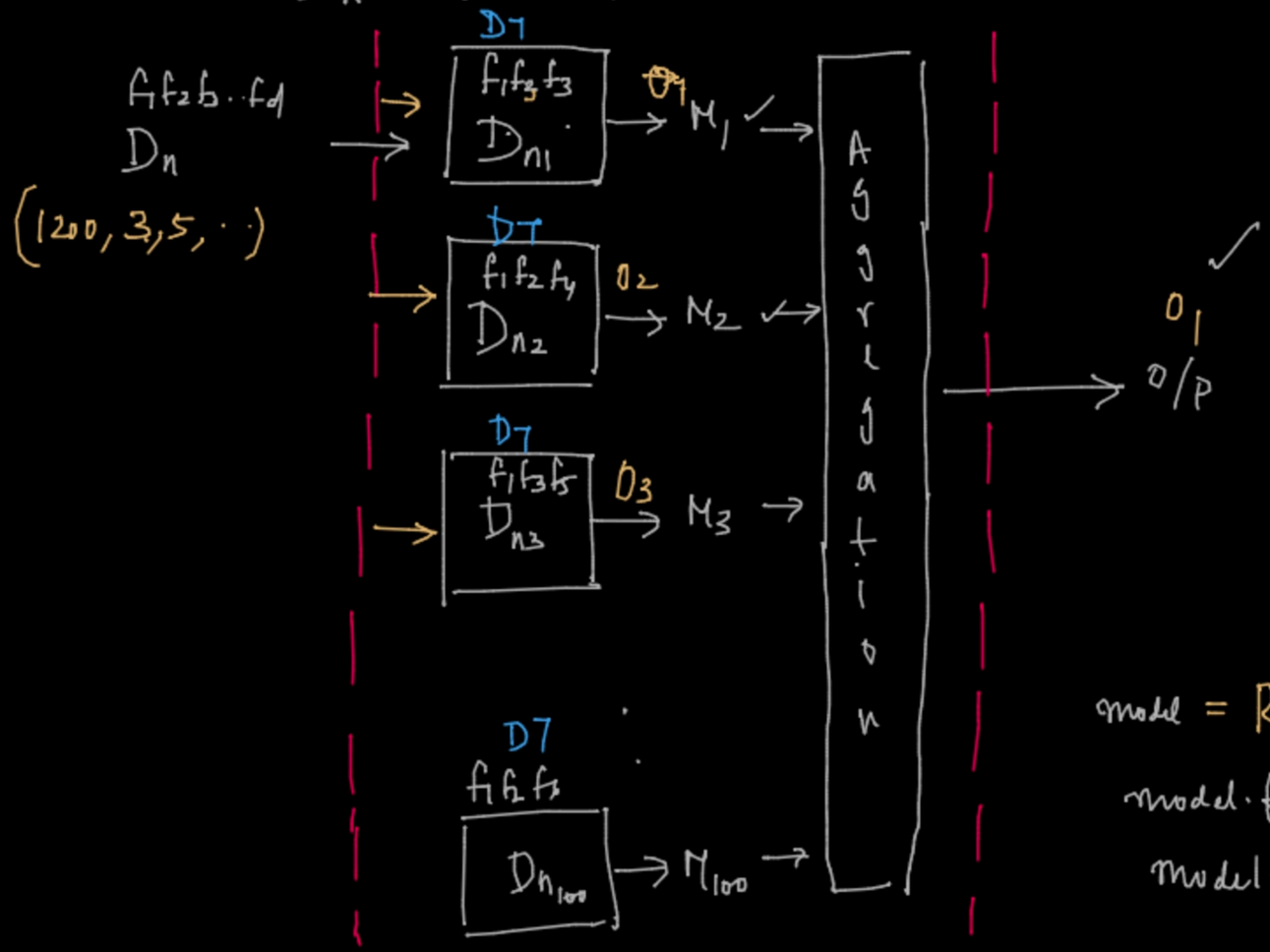
90% → 1 → BTM ✓
2 → Indranagar
3 → HSR
4 → BTM
5 → Indranagar
.
.

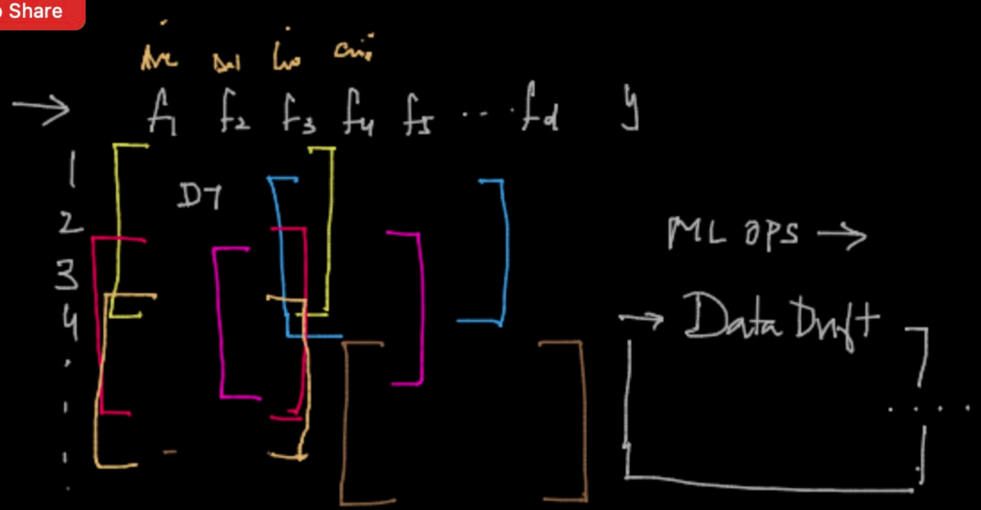
90%

zoom.us Meeting View Edit Window Help
zoom 46% Thu 28 Dec 6:39 PM
You are screen sharing Stop Share

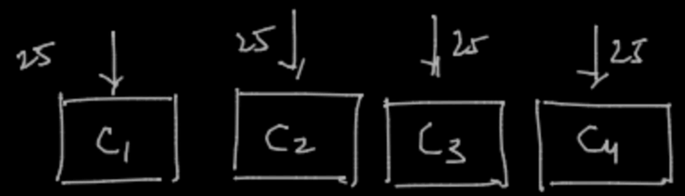
Random Forest → Bagging Technique

1. All Base models are Decision Trees ✓
2. Row sampling + Feature sampling
   ↳ only a subset of Features.





Parallelize the model.
→ Quad Core 25%.




```
model = RandomForest( n_estimators = 100 )
model.fit( X_train )
model.predict( )
```

zoom.us Meeting View Edit Window Help
zoom 30% Thu 28 Dec 7:17 PM
You are screen sharing Stop Share

Only the first model predicts 'y'
All subsequent models predicts the errors from the previous

Boosting
- Reduces Bias ✓
- Sequential Model
- Additive Model

Stage '0': M_0

$D_n \Rightarrow \{x_i, y_i\}_{i=1}^{n}$ $\quad \hat{y} = f_0(x)$ $\quad \varepsilon_0 = y - \hat{y} = y - f_0(x)$

Stage 1: M_1

$\{x_i, \varepsilon_{0i}\}_{i=1}^{n}$ $\quad \hat{\varepsilon}_0 = f_1(x)$ $\quad \varepsilon_1 = \varepsilon_0 - \hat{\varepsilon}_0 = y - f_0(x) - f_1(x)$

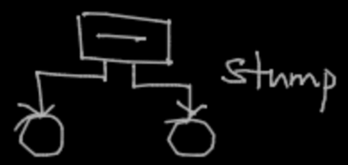
Stage 2 M_2

$\{x_i, \varepsilon_{1i}\}_{i=1}^{n}$ $\quad \hat{\varepsilon}_1 = f_2(x)$ $\quad \varepsilon_2 = \varepsilon_1 - \hat{\varepsilon}_1 = y - f_0(x) - f_1(x) - f_2(x)$

$\Downarrow$

$y = f_0(x) + f_1(x) + f_2(x) \longrightarrow$ Additive.

Decision stump



f_1	f_2	f_3	y	$\hat{y}$	ε_0
x_i	$\leftarrow x_1 \rightarrow$		y_1		
x_2	$\leftarrow x_2 \rightarrow$		y_2		
x_3					
x_n	$\leftarrow x_n \rightarrow$		y_n		

Base Models
High Bias & Low Variance
LBLV.

ADABOOST ✓
XGBOOST ✓
- M_estimators = ?
- Each Tree →

Opera File Edit View History Bookmarks Developer Window Help
You are screen sharing Stop Share
zoom 25% Thu 28 Dec 7:28 PM

Stage 0 — $f_0(x)$

Area	y	$\hat{y}$	ε_0
1200	35	22	13
2300	53	34	19
1700	47	38	9
3000	96	81	15
3500	94	82	12

Stage 1 — $f_1(x)$

Area	ε_0	$\hat{\varepsilon}_0$	ε_1
1200	13	8	5
2300	19	10	9
1700	9	2	7
3000	15	6	9
3500	12	5	7

Stage 2 — $f_2(x)$

Area	ε_1	$\hat{\varepsilon}_1$	ε_2
1200	5	1	4
2300	9	3	6
1700	7	2	5
3000	9	3	6
3500	7	4	3

$$y = f_0(x) + f_1(x) + f_2(x) + \cdots$$
$$= 22 + 8 + 1 + -0.5$$
$$= 31$$

−1 −0.8

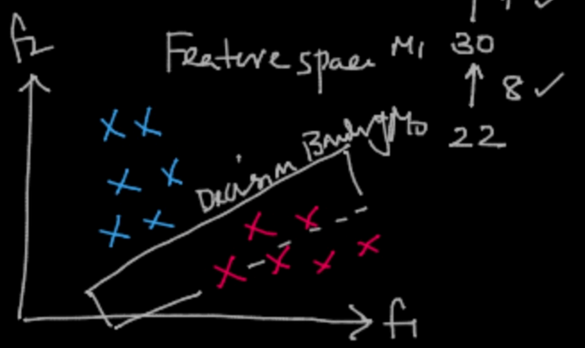
36 → 35.2 → 34.95

M_2 31.8 ↑ 0.5 M_2 31 ↑ 1 ✓

30 ↑ 8 ✓ 22

Classification

"Partition the feature space into pure regions belonging to each class"



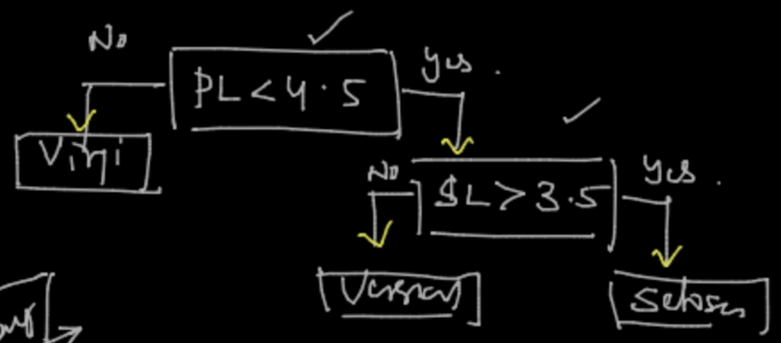
linear Model

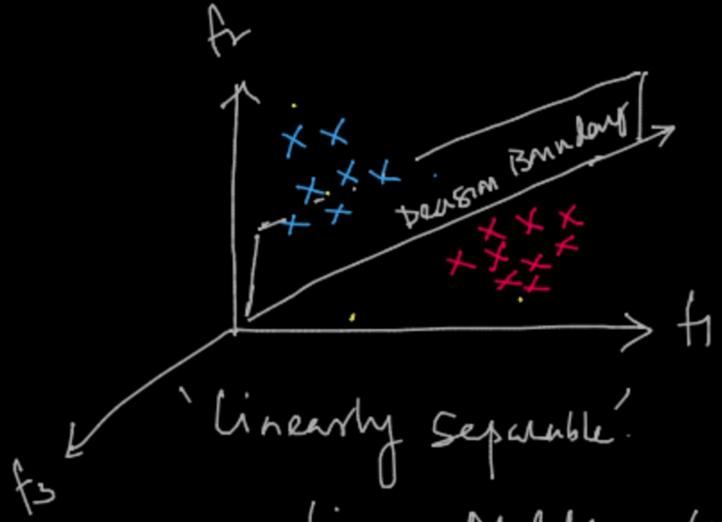
'Linearly Separable' — linear surface

5 / 6

Opera File Edit View History Bookmarks Developer Window Help
zoom 23% Thu 28 Dec 7:32 PM
You are screen sharing Stop Share

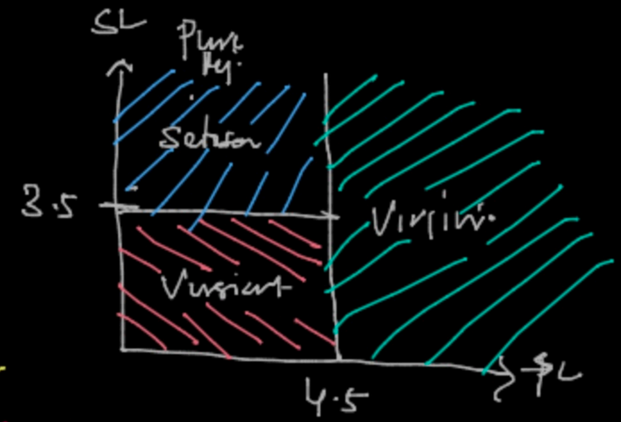
PL SL species

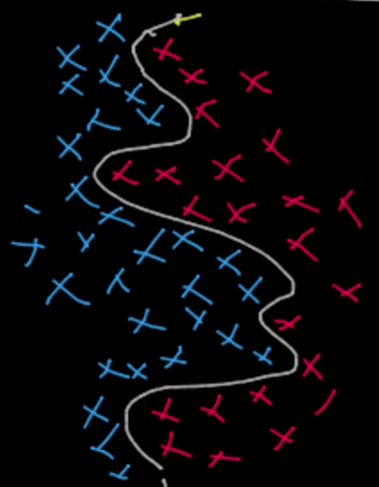




'Linearly Separable'

Linear Models → Logistic Regression





↓ Non Linear Decision Boundary

6 / 6