

for σ_{an} interest

Low - 6

High - 4

OI

$$p = \text{high} = 4/10 \quad \left| \begin{array}{cc} p(\text{high \& up}) & p(\text{high \& down}) \\ 2/4 & 2/4 \end{array} \right| \quad \left| \begin{array}{c} \text{GI} \\ 1 - \left[\left(\frac{2}{4} \right)^2 + \left(\frac{2}{4} \right)^2 \end{array} \right] \end{array} \right.$$

$$p = \text{low} = 6/10 \quad \left| \begin{array}{cc} p(\text{low \& up}) & p(\text{low \& down}) \\ 2/6 & 4/6 \end{array} \right| \quad \left| \begin{array}{c} \text{GI} \\ 1 - \left[\left(\frac{2}{6} \right)^2 + \left(\frac{4}{6} \right)^2 \end{array} \right] \end{array} \right.$$

$$\therefore \text{GI} : \begin{array}{cc} 1/2 = 0.5 & p(\text{high}) \\ 0.45 & p(\text{low}) \end{array}$$

$$\text{Weighted avg. of GI} = \frac{4}{10} \times 0.5 + \frac{6}{10} \times 0.45$$

$$= 0.47$$

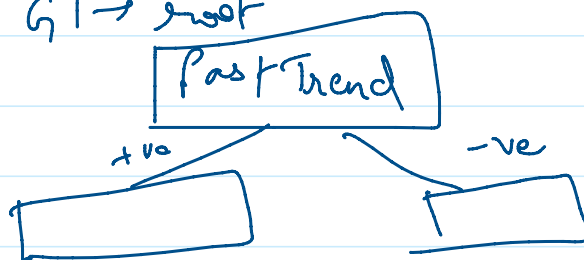
for Trading Vol.

$$\begin{array}{ccc|c} p(\text{high}) & p(\text{high \& up}) & p(\text{high \& down}) & \text{GI} \\ = 7/10 & = 4/7 & = 3/7 & 1 - \left[\left(\frac{4}{7} \right)^2 + \left(\frac{3}{7} \right)^2 \right] = 0.48 \\ p(\text{low}) & p(\text{low \& up}) & p(\text{low \& down}) & \\ = 3/10 & = 0/3 & = 3/3 & 1 - \left[0^2 + 1^2 \right] = 0 \end{array}$$

$$\text{Weighted avg of GI} = \frac{7}{10} \times 0.48 + \frac{3}{10} \times 0$$

$$= 0.336 \approx 0.34$$

Lowest GI \rightarrow root





Now, again for +ve:

Open Interest			
P(High)	r(High & up)	r(High & down)	GI $1 - [1^2 \times 0^2] = 0$ $1 - \left[\left(\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)^2\right] = 0.5$
$\frac{2}{6}$	$\frac{2}{6}$	0	
P(Low)	r(Low & up)	r(Low & down)	
$\frac{4}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	

Weighted GI

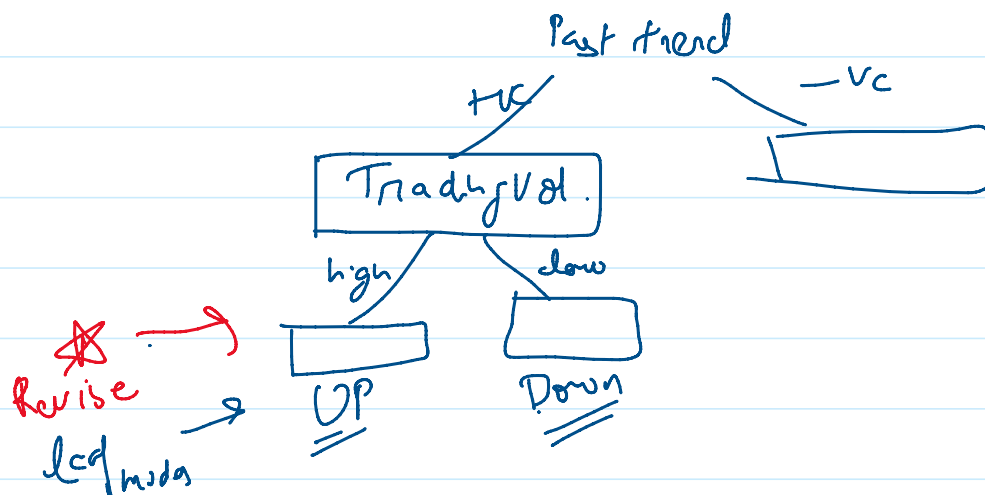
$$= 0 \times \frac{2}{6} + 0.5 \times \frac{4}{6}$$

$$= 0.33$$

for IV:

P(High)	up	down	GI
$\frac{4}{6}$	$\frac{4}{4} = 1$	0	0
P(Low)	up	down	
$\frac{2}{6}$	0	1	0

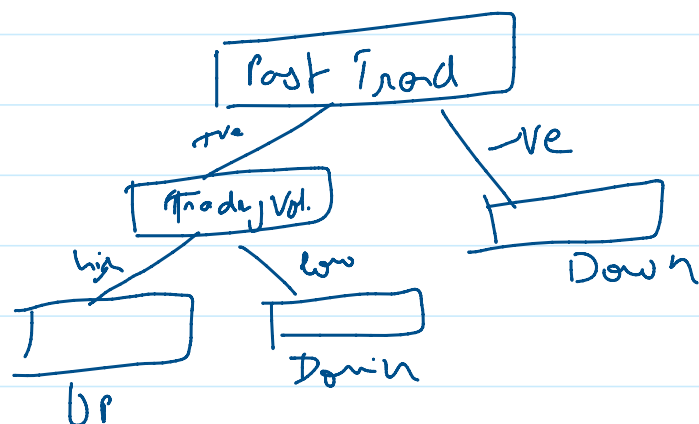
$$WGI = 0$$



Now, for -ve,

only 4 -ve all are down Target.

★ Revise :-
 "This point
 only important"



Penning: 9

Regression Tree:

$$\text{Avg of Golf Players} = \frac{25 + 30 + \dots + 30}{14} = 39.78$$

$$\text{SSE} = (39.78 - 25)^2 + \dots + \dots$$

$$= 1216.07$$

Outlook

Golf for Sunny: Avg = 35.2

$$\text{SSE for Sunny Outlook} = (35.2 - 25)^2 + \dots + \dots$$

$$= 302.64$$

Golf for Overcast Avg = 46.25

$$\text{SSE for overcast outlook} = (46 - 46.25)^2 + \dots + \dots$$

$$= 48.72$$

Golf for Rainy Avg = 39.2

$$\text{SSE for rainy outlook} = (45 - 39.2)^2 + \dots + \dots$$

$$\begin{aligned} \text{SSE for new outlook} &= (45 - 39.2)^2 + \dots + \dots \\ &= 590.78 \end{aligned}$$

$$\underline{\text{SSE}_{\text{Drop}}} : \text{Original SSE} - \text{calc SSE}$$

$$= 1216.67 - 302.64 - 48.72 - 590.78$$

$$\text{SSE}_{\text{drop}} =$$

for outlook

Similarly calc. SSE_{drop} for all features

→ Select highest SSE_{drop}

"For Regression tree we can stop when less than
 ≤ 14 instances are present"