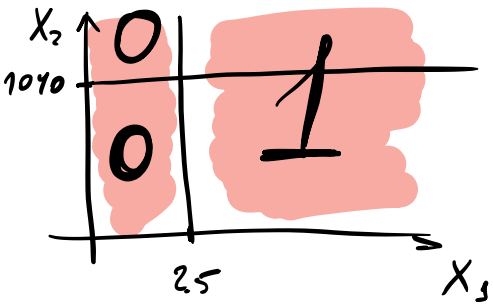
	X_1	X_2	y^{class}	y^{reg}
→	1	1000	0	701
#2	2	1050	0	805
	3	7000	1	904



→ [$X_1 < 2.5$]

да

нет

[$X_2 > 1040$]

#3

да

нет

#2

#1

$y = 700$

$y = 800$

X_1	y
1	0
2	1
3	0
4	1
5	1

[$X_1 > 3.5$]

да

X_1	y
4	1
5	1

класс 1

X_1	y
1	0
2	1
3	0

[$X_1 < 1.5$]

да

X_1	y
1	0

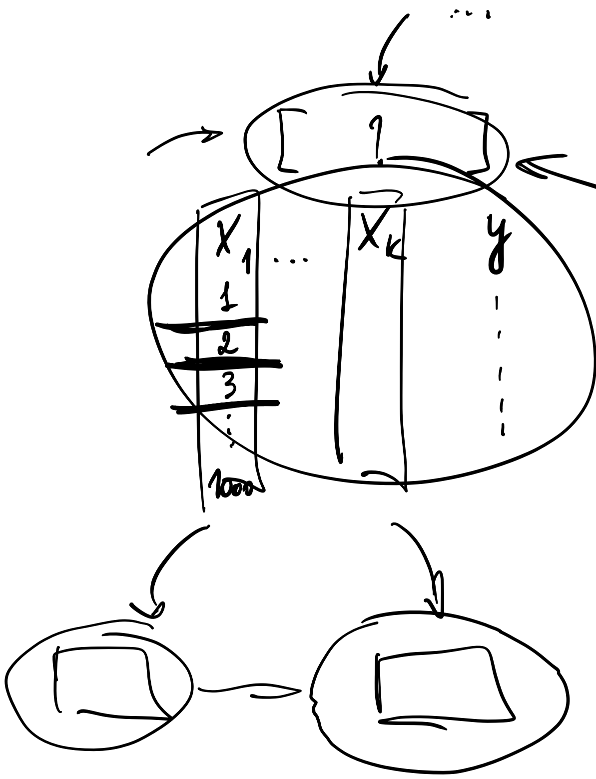
X_1	y
2	1
3	0

0

0

2 1

3 0



класиф.

$$H = \sum_k p_k \log p_k$$

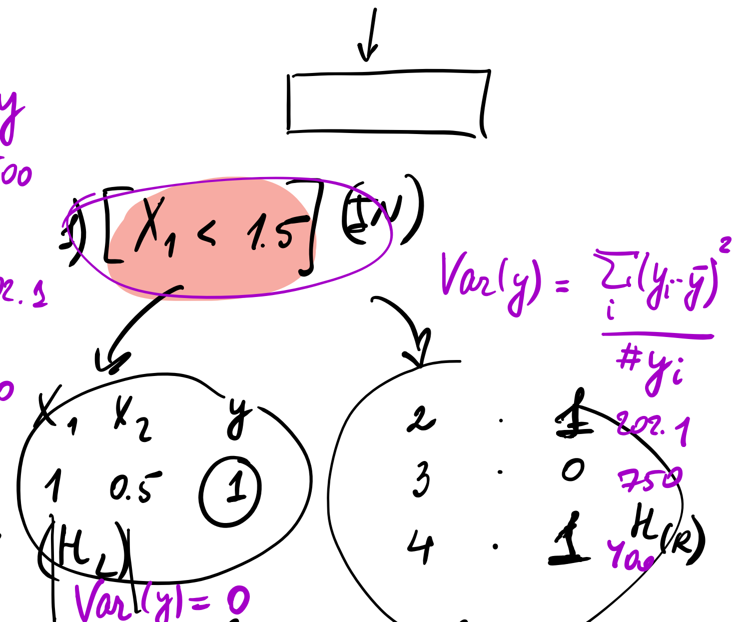
регрес.

Variance

усл. перем.

$X_5 < 5.5$

X_1	X_2	y	y
1	0.5	1	500
[illegible]			
2	0.17	1	202.1
3	0.78	0	750
4	0.91	1	400



$$Var(y) = \frac{\sum_i (y_i - \bar{y})^2}{\# y_i}$$

$$H_{IN} = \sum_k p_k \log p_k = \frac{3}{4} \log \frac{3}{4} + \frac{1}{4} \log \frac{1}{4}$$

$\frac{1}{2}$ | ln

→ $p_1 = 3/4 \quad p_0 = 1/4$

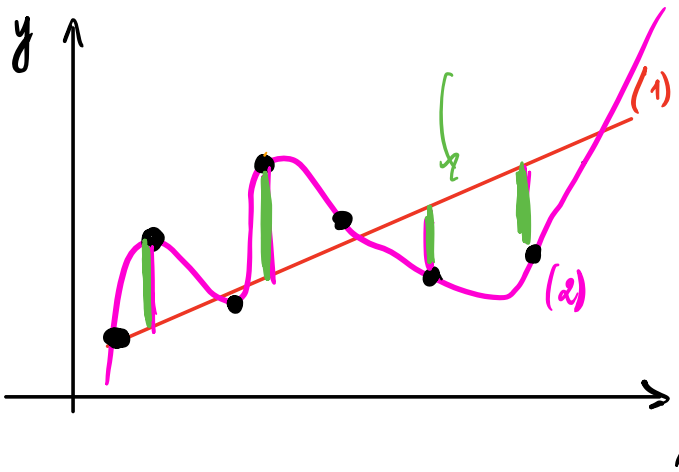
$$H_L = 1 \cdot \log 1 = 0$$

→ $p_{1L} = 1 \quad p_{0L} = 0$

$$H_R = 2/3 \log 2/3 + \frac{1}{3} \log \frac{1}{3}$$

$p_{1R} = \frac{2}{3} \quad p_{0R} = \frac{1}{3}$

$$H_{sum} = \frac{1}{4} H_L + \frac{3}{4} H_R$$



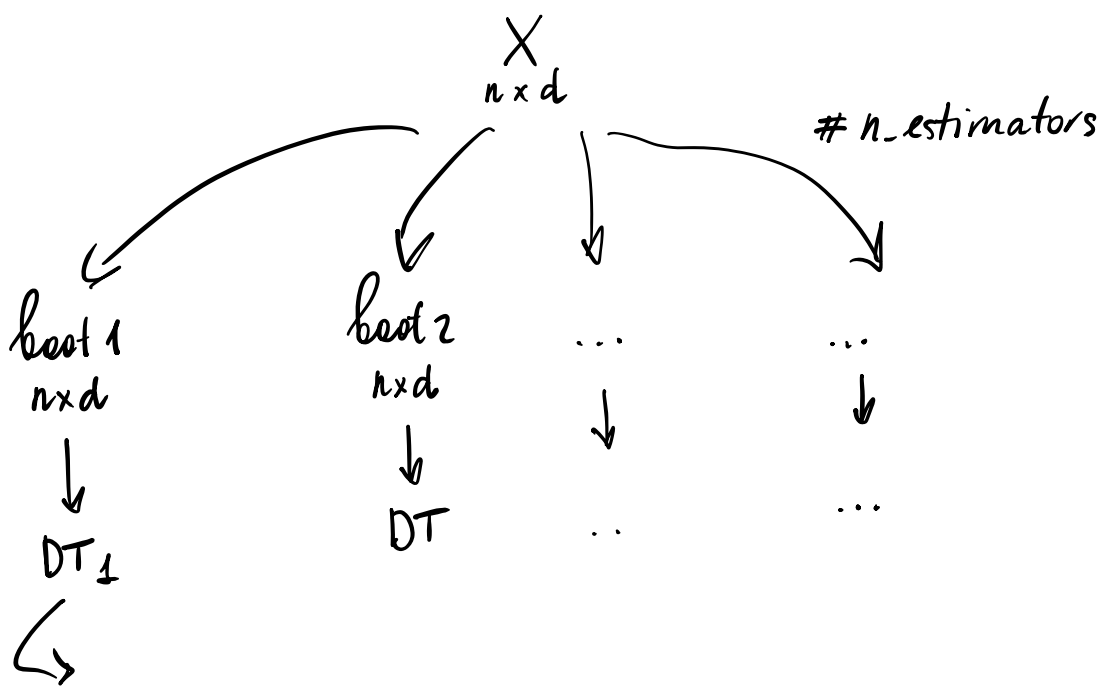
Bias ↑

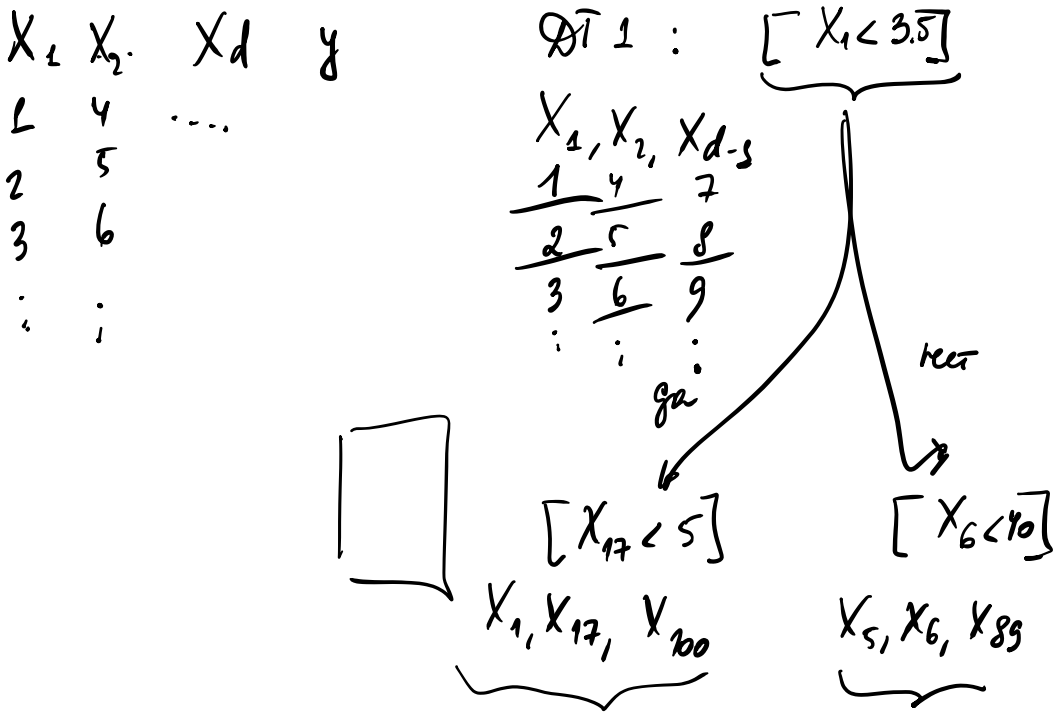
(1) $\hat{y} = \hat{w}_0 + \hat{w}_1 X_1$ Var ↓ (не переобуч.)

(2) $\hat{y} = \hat{w}_0 + \hat{w}_1 X_1 + \hat{w}_2 X_1^2 + \dots + \hat{w}_8 X_1^8$

Var ↑ Bias ↓ (переобуч.)

модель	Bias	Var	Переоб.
простая	высок.	низк.	×
сложная	низк	высок	✓




X_1 X_2 X_d y
1 4 ....
2 5
3 6
DT 1 : $[X_1 < 3.5]$
X_1, X_2, X_{d-3}
1 4 7
2 5 8
3 6 9
ja
nein
$[X_{17} < 5]$
$[X_6 < 10]$
X_1, X_{17}, X_{100}
X_5, X_6, X_{89}