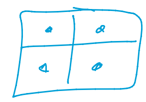
315 Lecture 24

Monday, April 20, 2020 5:44 PM

$$R(h) = \mathbb{E}_{XY}[L(Y, h(X))]$$
$$h^* = \underset{h}{\operatorname{argmin}}\, R(h)$$
$$= \underset{h}{\operatorname{argmin}}\, \mathbb{E}_{XY}[L(Y, h(X))]$$

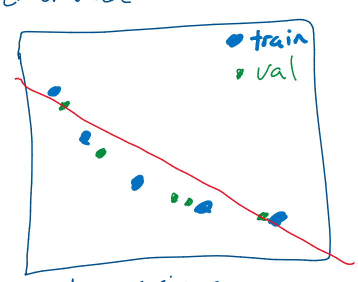
$$R(h) = E_{XY}[L(Y, h(X))]$$
$$= \sum_{x \in \mathcal{X}} \sum_{y \in \{0,1\}} p(x,y)\, L(y, h(x))$$
$$= \sum_{x} \sum_{y} p(y|x)\, p(x)\, L(y, h(x))$$
$$= \sum_{x} p(x) \sum_{y \in \{0,1\}} p(y|x)\, L(y, h(x))$$
$$= \sum_{x} p(x) \left[p(Y=0|x) L(0, h(x)) + p(Y=1|x) L(1, h(x)) \right]$$

$L(y, \hat{y})$ 

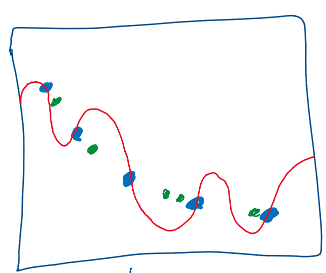
$$h^*(x) = \underset{\hat{y} \in \{0,1\}}{\operatorname{argmin}}\; p(Y=0|x) L(0, \hat{y}) + p(Y=1|x) L(1, \hat{y})$$
$$= \operatorname{argmin} \{ p(Y=0|x) L(0,0) + p(Y=1|x) L(1,0),\; p(Y=0|x) L(0,1) + p(Y=1|x) L(1,1) \}$$

$$h^*(x) = \begin{cases} 0 & p(Y=0|x) L(0,0) + p(Y=1|x) L(1,0) \le p(Y=0|x) L(0,1) + p(Y=1|x) L(1,1) \\ 1 & \text{o.w.} \end{cases}$$

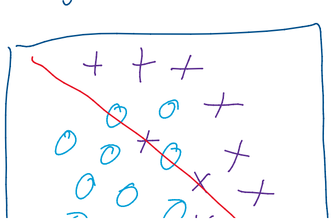
Generalize to train, val, test

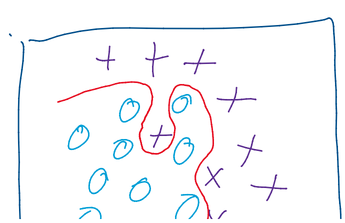


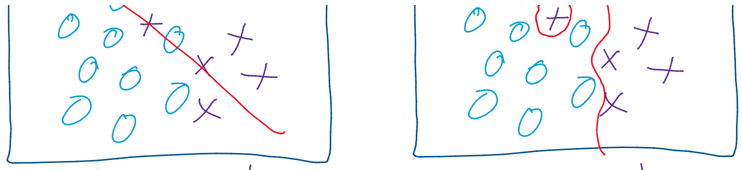
low variance
high bias



high variance
low bias







low complexity — high complexity

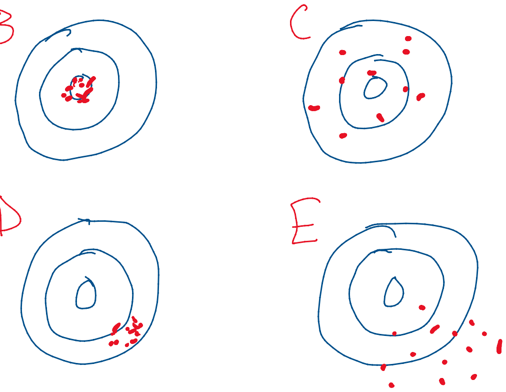
	h_A	h_B
true error	0.0	0.0
train error	0.25	0.1
val error	0.3	0.3
	high bias	low bias
	low variance	high variance

model complexity →

A Calamity

B — C

low bias



D — E

high bias

low variance — high variance

Poll 1: **which high var**

C E

Poll 2: **which high bias**

D E