MOTHER

$(p_i, t_i) \longrightarrow$

BABY

$\underbrace{(\varphi(p_i), t_i)}_{\text{memorize}}$

t_i'

Evaluation

$(p_j, t_j) \xrightarrow{\text{hide } t_j} t_j'$

F_{XY}

m times

$\{(x_i, y_i)\}_{i=1 \text{ to } m}$ $\mathcal{D}$

$D = \{(X_i, Y_i)\}_{i=1 \text{ to } m}$

$\mathcal{T}$ (test net) $\longrightarrow y_j'$

NN model

Storing $\mathcal{D}$ (TRAINING)

y_j'

$$P\left[Y \neq g_m^{NN}(X) \,/\, X = x, X' = x' \right]$$

$$= P\left[Y = 0, Y' = 1 \,/\, X = x, X' = x' \right] + P\left[Y = 1, Y' = 0 \,/\, X = x, X' = x' \right]$$

$$= P\left[Y = 0 \,/\, X = x \right] P\left[Y' = 1 \,/\, X = x, X' = x' \right] + \;''$$

$$= \left(1 - \eta(x)\right)\left(\eta(x')\right) + \eta(x)\left(1 - \eta(x')\right) \xrightarrow[m \to \infty]{} \underline{2\eta(x)\left(1 - \eta(x)\right)}$$

$X' \longrightarrow X$ as $m \to \infty$

$\eta(X') \longrightarrow \eta(X)$ as $m \to \infty$

$$R^{NN}(x) = 2\eta(x)(1-\eta(x))$$

$$\underbrace{E\left[R^{NN}(x)\right]}_{R^{NN}} = 2E\left[\eta(x)(1-\eta(x))\right] = 2E\left[\tilde{\eta}(x)(1-\tilde{\eta}(x))\right]$$

$$= 2E[\tilde{\eta}(x)]\left(1-E[\tilde{\eta}(x)]\right) - 2\,\mathrm{var}(\tilde{\eta}(x))$$

By adding and subtracting $2\left(E[\tilde{\eta}(x)]\right)^2$.

$$\leq 2R^*(1-R^*) \leq 2R^*$$

Also, $R^* = E[\tilde{\eta}(x)]$,

where $\tilde{\eta}(x) \equiv \min(\eta(x), 1-\eta(x))$

In the lecture, I just assumed $\tilde{\eta}(x) = \eta(x)$.

The above is the generic derivation.

LECTURE-5

k is odd

$$R^{kNN}(x) = \eta(x) \sum_{j=0}^{\lfloor k/2 \rfloor} {}^kC_j \,(\eta(x))^j \,(1-\eta(x))^{k-j} + (1-\eta(x)) \sum_{j=0}^{\lfloor k/2 \rfloor} {}^kC_j \,(\eta(x))^{k-j} \,(1-\eta(x))^j$$

$$R^{kNN}(x) = \eta(x)\, P\left[\sum_{i=1}^{k} Z_i < 0\right] + (1-\eta(x))\, P\left[\sum_{i=1}^{k} Z_i > 0\right]$$

$$= \eta(x) + \underbrace{(1-2\eta(x))}_{>0}\, P\left[\sum_{i=1}^{k} Z_i > 0\right]$$

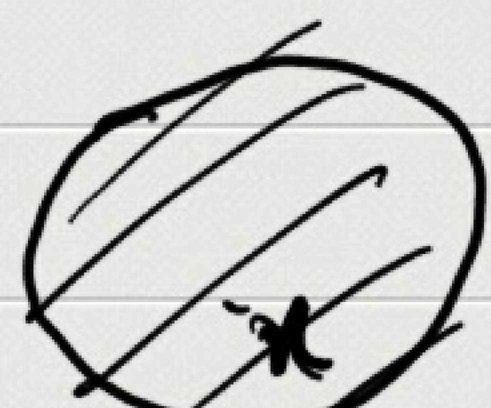
$\eta(x) < 1/2$

$k \uparrow$, then $\downarrow$

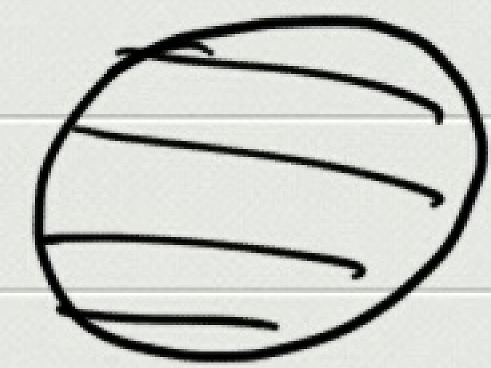
$$\therefore\ R^{(2k+1)NN} \le R^{(2k-1)NN} \quad \forall k \in \mathbb{N}$$

counter example with finite 'm'!

$f_{X|Y}(x|1)$



$f_Y(1) = 0.5$



$f_{X|Y}(x|2)$

$f_Y(2) = 0.5$

Non-deterministic Models

KNN → Estimate distributions from samples of the D

$F_{Y|X}$ ~ distribution form

Non-parametric

↓

training → store D

inference → find nearest neighbours.

Parametric

↓

training = Parameter estimation

Inference = obvious

Deterministic Models

eg: linear model

quadratic model.

EXAMPLE OF ANOTHER MODEL

$$f(x) = \sum_{i=1}^{m} \alpha_i y_i \underbrace{k(x_i, x)}_{\text{closeness of } x_i \text{ to } x}$$

Model = $\{ f \}$

training $\longrightarrow$ obtaining α.

Inference $\longrightarrow$ evaluate $f(x)$.

Deterministic modeling.

Parametric.