Lec 3: Generative vs Discriminative, Tuesday, 6th Aug

06 August 2013
13:55

$$\boxed{P(X,Y) = \prod_{i=1}^{n} p(x_i, y_i) \quad +1}$$

$$P(X,Y|\theta) = \prod_{i=1}^{n} p(x_i, y_i|\theta) \quad +3$$

$$P(Y|X) = \prod_{i=1}^{n} p(y_i|x_i) \quad 0$$

$$P(Y|X,\theta) = \prod_{i=1}^{n} p(y_i|x_i,\theta) +$$

approx $f: X \to Y$ $\quad p(y|x)$

$$D = \{\langle x^i, y^i\rangle\}$$

$$x = \langle x_1, \dots x_n\rangle$$

(i) Discriminative: approx
$p(y|x)$ directly from data

(ii) Generative:

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)}$$

DISCRIMINATIVE

$$p(Y|X, w) \leftarrow X$$

$P(Y|X)$

$$= \prod_{i=1}^{n} P(y_i | x_i, w)$$

$$= \prod_{i=1}^{n} \frac{1}{1+e^{-y_i w^t x_i}}$$

$$= \sum_{i=1}^{n} \log(1+e^{-y_i w^t x_i})$$

GENERATIVE

$$\theta_{MAP} = \underset{\theta}{\text{ARGMAX}}\, P(\theta | X, Y)$$

$$= \underset{\theta}{\text{ARGMAX}}\, P(X, Y | \theta) P(\theta)$$

$$= \underset{\theta}{\text{ARGMAX}}\, P(\theta) \prod_{i=1}^{n} P(x^i, y^i | \theta)$$

$$= \underset{\theta}{\text{ARGMAX}}\, P(\theta) \prod_{i=1}^{n} P(x^i | y^i, \theta) P(y^i | \theta)$$

$$P(y = M | \theta) = \bar{\pi}$$

$$P(y = F | \theta) = 1 - \bar{\pi}$$

$$P(X | y = M) = e^{-\frac{1}{2}(x - \mu_M)^t \Sigma^{-1} (x - \mu_M)}$$

$$P(X | y = F) = e^{-\frac{1}{2}(x - \mu_F)^t \Sigma^{-1} (x - \mu_F)}$$

$$\bar{\pi} = \frac{\# B}{\ldots}$$

$$\bar{\pi} = \frac{\# B}{\# B + \# G}$$

$$\mu_M = \frac{1}{\# B} \sum_{i \in BOYS} x_i$$

$$\mu_F = \frac{1}{\# G} \sum_{i \in GIRLS} x_i$$

$$p\left(\theta = [\bar{\pi}, \mu_F, \mu_M, \Sigma] \right) = \text{CONST}$$

$$p(y = \text{MALE} \mid X) \geq 1/2$$

$$\frac{p(X \mid y)\, p(y)}{p(x)}$$

$$\frac{e^{-1/2 (x - \mu_M) \Sigma^{-1} (x - \mu_M)} \bar{\pi}}{(\quad) + \left(e^{-1/2 (x - \mu_F) \Sigma^{-1} (x - \mu_F)} (1 - \bar{\pi}) \right)}$$

$p(Y \mid X)$ DISC

$$\theta_{MAP} = \underset{\theta}{\text{ARGMAX}}\ p(\theta \mid X, Y)$$

$$= \underset{\theta}{\mathrm{ARGMAX}}\ p(X, y \mid \theta)\, p(\theta)$$

$$= \mathrm{ARG}\ p(\theta) \prod_{i=1} p(x, y$$

$$= \underset{\mathrm{MAX}}{\mathrm{ARG}}\ p(\theta) \prod_{i=1}^{n} \underbrace{p(y_i \mid X_i, \theta)}_{,w}\, \underline{\underline{p(x_i \mid \theta)}}$$

$$\theta = [w, w_\perp]$$

$$p(y_i \mid x, w, w_\perp) \equiv p(y_i \mid x, w)$$

$$y_i x_i^t w$$

$$= \underset{w}{\mathrm{ARGMAX}}\ p(w) \prod_{i=1}^{n} p(y_i \mid x_i, w)$$

$$p(x_i, y_i \mid \theta)$$

$$\boxed{\underset{w}{\mathrm{Arg}} \sum \log\left(1 + e^{-y_i w^t x_i}\right)}$$

$$\pi, \mu_F, \mu_M, \Sigma$$

$$1, D, D, D^2$$