

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

06 August 2013

13:55

$$P(X, Y) = \prod_{i=1}^N P(x_i, y_i) \quad (1)$$

$$P(X, Y | \theta) = \prod_{i=1}^N P(x_i, y_i | \theta) \quad (2)$$

$$P(Y | X) = \prod_{i=1}^N P(y_i | x_i) \quad (3)$$

$$P(Y | X, \theta) = \prod_{i=1}^N P(y_i | x_i, \theta) \quad (4)$$

approx $f: X \rightarrow Y$ $P(y|x)$

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

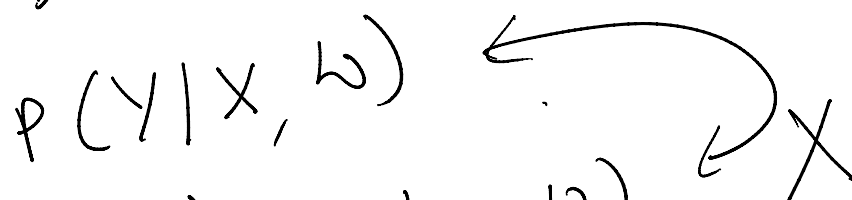
$$x = \langle x_1, \dots, x_n \rangle$$

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

(ii) Generative:

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

DISCRIMINATIVE



$$\begin{aligned}
 P(\mathcal{Y}|\mathcal{X}) &= \prod_{i=1}^N p(y_i | x_i, \omega) \\
 &= \prod_{i=1}^N \frac{1}{1 + e^{-y_i \omega^T x_i}} \\
 &= \sum_{i=1}^N \log(1 + e^{-y_i \omega^T x_i})
 \end{aligned}$$

GEN ERATIVE

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

$$\begin{aligned}
 &= \underset{\theta}{\text{ARGMAX}} P(\mathcal{X}, \mathcal{Y} | \theta) P(\theta) \\
 &= \underset{\theta}{\text{ARGMAX}} P(\theta) \prod_{i=1}^n P(x^i, y^i | \theta) \\
 &= \underset{\theta}{\text{ARGMAX}} P(\theta) \prod_{i=1}^n P(x^i | y^i, \theta) P(y^i | \theta)
 \end{aligned}$$

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

$$P(y = 0 | \theta) = 1 - \bar{\pi}$$

$$P(x | y = 1) = e^{-\frac{1}{2}(x - \mu_1)^T \Sigma^{-1}(x - \mu_1)}$$

$$P(x | y = 0) = e^{-\frac{1}{2}(x - \mu_0)^T \Sigma^{-1}(x - \mu_0)}$$

$$\bar{\pi} = \frac{\# 1}{\text{total}}$$

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

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

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

$$p(\theta = [\bar{\mu}, \mu_F, \mu_M, \Sigma]) = \text{Gauss}$$

$$p(y = \text{MALE} | x) \geq \frac{1}{2}$$

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

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

$$p(y | x) \text{ DJS C}$$

$$\theta_{\text{MAP}} = \underset{\theta}{\text{argmax}} p(\theta | x, y)$$

$$= \underset{\theta}{\text{ARGMAX}} p(x, y | \theta) p(\theta)$$

$$= \underset{\theta}{\text{ARG}} p(\theta) \prod_{i=1}^N p(x_i, y_i | \theta)$$

$$= \underset{\theta}{\text{ARGMAX}} p(\theta) \prod_{i=1}^N \underbrace{p(y_i | x_i, \theta)}_{\text{b}} \underbrace{p(x_i | \theta)}_{\text{b}}$$

$$\theta = [\omega, \omega_+]$$

$$p(y_i | x_i, \omega, \omega_+) \equiv p(y_i | x_i, \omega)$$

$y_i x_i^T \omega$

$$= \underset{\omega}{\text{ARGMAX}} p(\omega) \prod_{i=1}^N p(y_i | x_i, \omega)$$

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

$$\underset{\omega}{\text{ARG}} \sum \log(1 + e^{-y_i \omega^T x_i})$$

$$\bar{\lambda}, \mu_F, \mu_M, \sum$$

$$4, D, D, D^2$$