$$p(\theta|D) \propto \underline{p(D|\theta)\,p(\theta)} \longrightarrow \max, \ \theta_{MAP}$$

Supervised learning: $\bar{x} \in X \longrightarrow y \in Y$

$p(y|\bar{x})$?

$$p(y|\boxed{\bar{x}}, \bar{w}) = \mathcal{N}(y|\bar{w}^T\bar{x}, \sigma^2)$$

Discriminative

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

$\cancel{p(\bar{x}, y) = p(y|\bar{x})\,p(\bar{x})}$

Log. regr.:

$$p(y|\bar{x}, \bar{w}) = \sigma(\bar{w}^T\bar{x})$$

Generative

$p(\bar{x}, y)$

$$p(y|\bar{x}) = \frac{p(\bar{x}, y)}{p(\bar{x})} \propto p(\bar{x}, y)$$

LDA/QDA

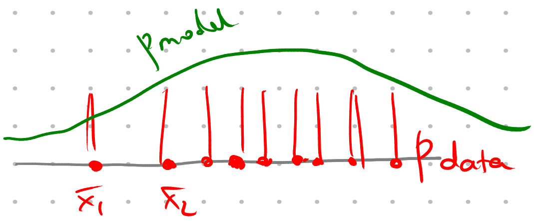
$$p(\bar{x}|C_1) = \mathcal{N}(\bar{x}|\bar{\mu}_1, \Sigma_1)$$
$$p(\bar{x}|C_2) = \mathcal{N}(\bar{x}|\bar{\mu}_2, \Sigma_2)$$

$$p(\bar{x}, y) = p(y)\,p(\bar{x}|y)$$

$p(\bar{x})$ $\quad D = \{\bar{x}_1, \dots, \bar{x}_N\}$ $\quad p_{model}(\bar{x}|\bar{\theta})$

$$\bar{\theta}_{ML} = \arg\max \prod_{\bar{x} \in D} p_{model}(\bar{x}|\bar{\theta})$$

$$\bar{\theta}_{MAP} = \text{——} \,''\, \text{——} \cdot p(\bar{\theta})$$



$$\theta^* = \arg\min KL(p_{data} \| p_{model}) =$$

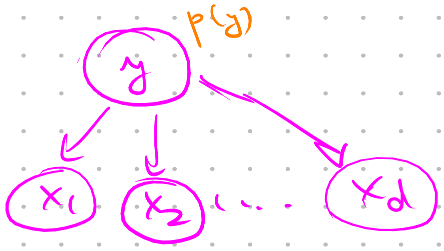
$$= \arg\min \int p_{data}(\bar{x}) \cdot \log \frac{p_{data}(\bar{x})}{p_{model}(\bar{x})}\, d\bar{x} =$$

$$= \arg\min \sum_{\bar{x} \in D} \frac{1}{N} \cdot \left(\cancel{\log \frac{1}{N}} - \log p_{model}(\bar{x})\right)$$

$$= \arg\max \sum_{\bar{x} \in D} \log p_{model}(\bar{x}) = \theta_{ML}$$

Naive Bayes

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



$$p(\bar{x}, y) = p(y) \cdot \prod_{i=1}^{d} p(x_i|y)$$

$y \in \{1, \dots, K\}, \quad x_i \in \{1, \dots, M\}$

$$\theta_k = p(y=k) \; ; \qquad \theta_{imk} = p(x_i = m | y = k)$$

$$\bar{\theta} = \{\theta_1 \dots \theta_K, \; \theta_{imk}, \; i = 1 \dots d, \; m = 1 \dots M, \; k = 1 \dots K\}$$

$$\hat{\theta}_k = \frac{\#\{y=k\} + 1}{N + K} \quad \Bigg| \quad \hat{\theta}_{imk} = \frac{\#\{m \text{ бгдес. на } i\text{-и пр. в кл. } k\} + 1}{\#\{\text{всеко сл на } i\text{-и пр. в кл. } k\} + M}$$

$$p(D|\bar{\theta}) = \prod_{n=1}^{N} \left(p(y_n) \prod_{i=1}^{d} p(x_{ni}|y_n) \right) =$$

$$= \prod_{n=1}^{N} \prod_{k=1}^{K} \left(\underline{p(y_n=k)} \prod_{i} p(x_{ni}|y_n=k) \right)^{[y=k]} =$$

$$= \prod_{n=1}^{N} \prod_{k=1}^{K} \left(\theta_k \cdot \prod_{i=1}^{d} \prod_{m=1}^{M} \left(p(x_{ni}=m|y_n=k) \right)^{[x_{ni}=m]} \right)^{[y_n=k]} =$$

$$= \prod_{n=1}^{N} \prod_{k=1}^{K} \left(\theta_k \prod_{i=1}^{d} \prod_{m=1}^{M} \theta_{imk}^{[x_{ni}=m]} \right)^{[y_n=k]}$$

$\theta_k \longrightarrow \max$

$$\sum_{k=1}^{K} \#[y_n=k] \cdot \theta_k$$

$\sum \theta_k = 1$

$$\log p(D|\theta) = \sum_{n=1}^{N} \sum_{k=1}^{K} [y_n=k] \cdot \theta_k \; +$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{i=1}^{d} \sum_{m=1}^{M} [y_n=k] \cdot [x_{ni}=m] \; \theta_{imk} \xrightarrow[\theta]{} \max$$

$$\text{u} \quad \sum_{k} \sum_{i} \left[\sum_{m} \#[y_n=k, x_{ni}=m] \; \theta_{imk} \right]$$

$$\log p(D|\theta) = \underbrace{\sum_{k=1}^{K} f_k(D)\,\theta_k}_{} + \underbrace{\sum_k \sum_i \sum_m f_{imk}(D)\cdot \theta_{imk}}_{}$$

$$f_k(D) = \#[y_n = k] \qquad f_{imk}(D) = \#[y_n = k,\ x_{ni} = m]$$

<u>Logistic regression</u>

$$p(y=k\,|\,\bar{x}, \bar{\theta}) = \frac{e^{\bar{\theta}_k^T \bar{x}}}{\sum_s e^{\bar{\theta}_s^T \bar{x}}} = \frac{1}{Z(\bar{x})}\cdot e^{\sum_{i=1}^{d} \theta_{ki} x_i}$$

$$p(D|\bar{\theta}) = \prod_{n=1}^{N} p(y_n\,|\,\bar{x}_n, \bar{\theta}) = \prod_{n=1}^{N} \left(\frac{1}{Z(\bar{x}_n)} \cdot e^{\sum_{i=1}^{d} \theta_{ki} x_{ni}} \right)$$

$[y_n = k]$

const

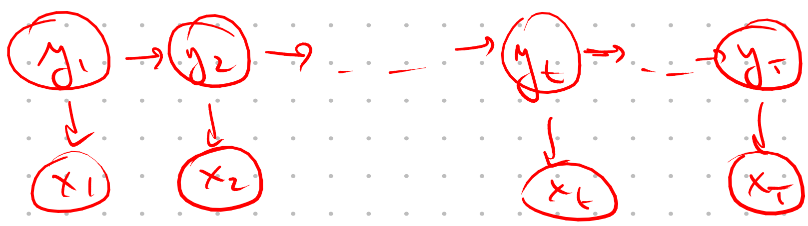
$\bar{\varphi}(\bar{x}_n)$

$$\log p(D|\bar{\theta}) = -\sum_{n=1}^{N} \log Z(\bar{x}_n) + \sum_{n=1}^{N} \sum_{i=1}^{d} [y_n = k]\, x_{ni}\, \theta_{ki}$$

новые признаки

Generative-discriminative pair

<u>HMM</u>



$$p(\bar{x}, \bar{y}) = p(y_1)\,p(x_1|y_1) \cdot p(y_2|y_1)\,p(x_2|y_2)\,p(y_3|y_2) \;\text{— —}\; p(y_T|y_{T-1})\,p(x_T|y_T)$$

$\cancel{a_{ij}}$ $\quad$ $\cancel{b_j(k)}$

$$\log p(D|\theta) = \sum_n \log p(\bar{x}_n, \bar{y}_n|\bar{\theta}) =$$

$$= \sum_{n=1}^{N} \log p(y_{n1}) + \sum_{n=1}^{N} \sum_{t=1}^{T-1} \log p(y_{n,t+1}|y_{n,t}) + \sum_n \sum_{t=1}^{T} \log p(x_{nt}|y_{nt})$$

$y_t = i$ $\qquad x_t = k$

$y_t = 1, \dots, M \quad , \quad x_t = 1, \dots, K$

$[\ldots]$

$$= \sum_{n=1}^{N} \sum_{i=1}^{M} [y_{n1}=i] \underbrace{\left(\log p(y_{n1}=i)\right)}_{\theta_i} + \sum_{n} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \overbrace{\left(\log p(y_{n,t+1}=j \mid y_{n,t}=i)\right)}^{\theta_{ij}}$$

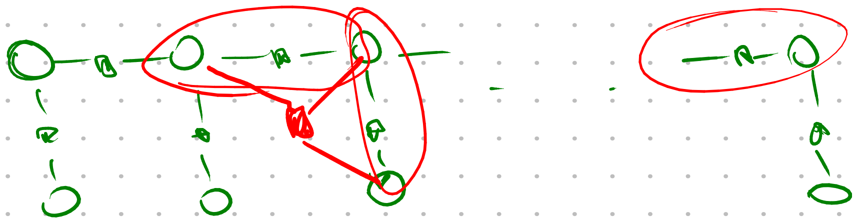
$$+ \sum_{n} \sum_{t} \sum_{i=1}^{N} \sum_{k=1}^{K} [x_{nt}=k][y_{nt}=i] \underbrace{\left(\log p(x_{nt}=k \mid y_{nt}=i)\right)}_{\theta_{ik}}$$

$$\log p(D|\bar{\theta}) = \sum_i \left[\overbrace{\sum_n [y_{n1}=i]}\right] \theta_i$$

$$+ \sum_{i=1}^{M} \sum_{j} \left(\underbrace{\sum_{n=1}^{N} \sum_{t=1}^{T-1} [y_{n,t}=i][y_{n,t+1}=j]} \right) \theta_{ij}$$

$$+ \sum_{i}^{M} \sum_{k}^{K} \left(\underbrace{\sum_n \sum_t [y_{nt}=i][x_{nt}=k]} \right) \theta_{ik} \quad \xrightarrow[\theta]{} \max$$

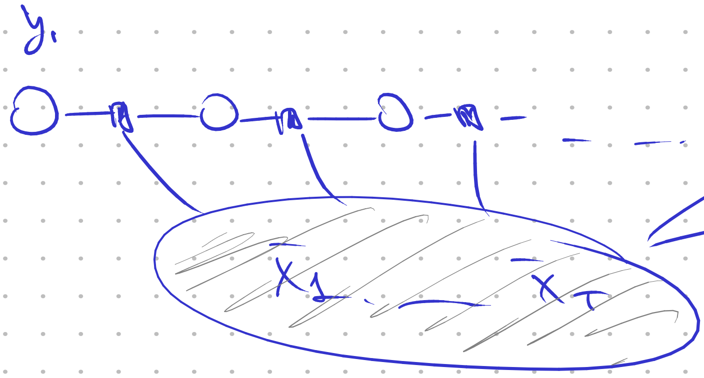
Conditional random field, CRF $\qquad p(\bar{y}|\bar{x})$



$$\log p(D|\theta) = \sum_n \left[\sum_i f_i(y_1)\, \theta_i + \sum_{i,j} \left[\sum_t f(y_t, y_{t+1}) \right] \theta_{ij} + \right.$$

$$\left. + \sum_i \sum_k \left[\sum_t f(y_t, x_t) \right] \theta_{ik} \right] + \text{const}$$

$$\sum_{i,j,k} \sum_t f(y_{t-1}, y_t, x_t)\, \theta_{ijk}$$



$p(\bar{y} \mid \bar{x})$

$f(y_t, y_{t+1}, \bar{x})$

NER