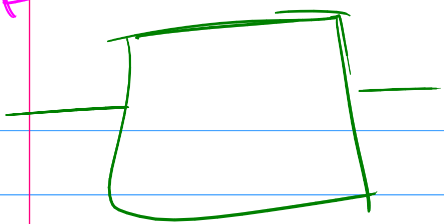


$\bar{x} \in \mathbb{R}^{3\cdot 10^6}$ $\quad\to\quad$ $y \in \{cat, dog\}$

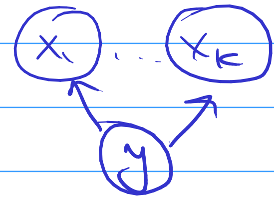
$p(y|\bar{x})$ – discr

$p(\bar{x}, y)$ – generative

$= p(y|\bar{x})\, p(\bar{x})$ $\quad \uparrow \mathbb{R}^{3\cdot 10^6}$

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

GEN **Naive Bayes** $\quad \bar{x} = (x_1, \ldots, x_K) \to y$



Новини → спорт, полит, фин.

y

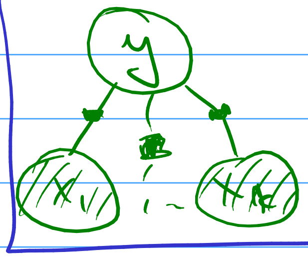
$$p(\bar{x}|y) = \prod_{k=1}^{K} p(x_k|y)$$

$$p(y|\bar{x}) \propto p(\bar{x}, y) = p(y) \cdot \prod_{k=1}^{K} p(x_k|y)$$

→ $y =$ „спорт"

$w_1 \quad w_2 \quad w_3$

DISCR **Logistic Regression**



$$p(y=c|\bar{x}) = softmax(e^{\bar{w}_c^T \bar{x}}) = \frac{e^{\bar{w}_c^T \bar{x}}}{\sum_{c'} e^{\bar{w}_{c'}^T \bar{x}}}$$

$$p(y=c|\bar{x}) = \frac{1}{Z(\bar{x})} e^{\bar{w}_c^T \bar{x}} = \frac{1}{Z(\bar{x})} (e^{w_1 x_1}) \cdot (e^{w_2 x_2}) \cdot \ldots \cdot (e^{w_K x_K})$$

$$\boxed{p_{LR}(y|\bar{x}) = \frac{1}{Z(\bar{x})} \cdot \prod_{i=1}^{K} f_k(y, x_k)}$$

NB и LR образуват generative–discr. pair

$D: \bar{x}$ $\qquad p(\bar{x}|\bar{\theta})$

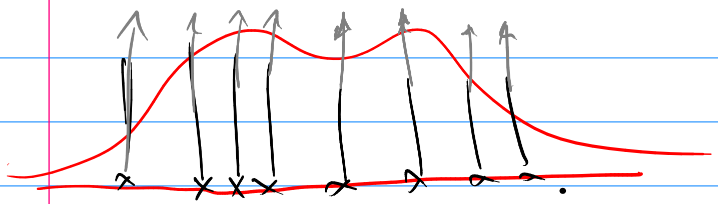
$$p_{model}(\bar{x}|\bar{\theta}) \approx p_{data}(\bar{x})$$

$$p(\bar{x}|\bar{\theta}) = \ell(\bar{x})$$

$$L(D) = \prod_{\bar{x} \in D} p(\bar{x}|\bar{\theta}) \to \max_{\bar{\theta}}$$

$p_{data}(\bar{x}) = \mathrm{Unif}(\{\bar{x}_1, \ldots, \bar{x}_N\})$

$1/N \quad 1/N$



$KL(p^{(\bar{x})} \| q^{(\bar{x})}) = -\int_x p^{(\bar{x})} \ln \frac{q^{(\bar{x})}}{p(\bar{x})} d\bar{x}$

$KL(p_{data} \| p_{model}) =$

$= -\int p_{data}(\bar{x}) \ln \frac{p_{model}(\bar{x})}{p_{data}(\bar{x})} d\bar{x}$

$= -\sum_n \frac{1}{N} \ln \frac{p_{model}(\bar{x}_n)}{1/N} =$

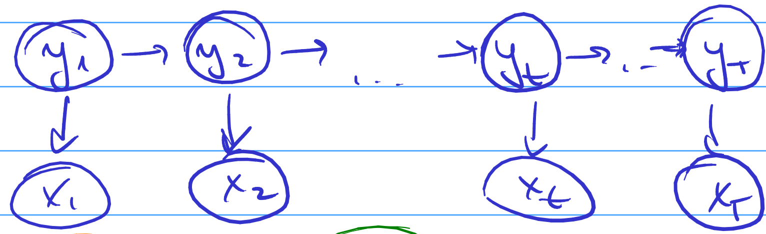
$= -\sum_n \frac{1}{N} \ln N - \underbrace{\frac{1}{N} \sum_n \ln p_{model}(\bar{x}_n)}_{p_{model}} \to \min$

$\Leftarrow ?$ $\quad \downarrow p_{model}$ max

$\prod_n p_{model}(\bar{x}_n) \to \max$

GEN

Hidden Markov models



$p(x_1 \ldots x_T, y_1 \ldots y_T | \pi, A, B) =$

$= p(y_1) p(x_1|y_1) p(y_2|y_1) p(x_2|y_2) \ldots p(y_T|y_{T-1}) p(x_T|y_T)$

$\boxed{\pi_i = p(y_1 = i)}$

$= \prod_{i=1}^{n} \pi_i^{[y_1 = i]}$

$A, \quad a_{ij} = p(y_t = j | y_{t-1} = i)$

$B, \quad b_i(k) = p(x_t = k | y_t = i)$

DISCR

Conditional random fields CRF

HMM:

$p(x_1 \ldots x_T, y_1 \ldots y_T | \lambda) = e^{\sum_i [y_1 = i] \ln \pi_i + \sum_{i,j} \left(\sum_t [y_t = j, y_{t-1} = i] \right) \ln a_{ij} + \sum_{i,k} \left(\sum_t [y_t = i, x_t = k] \right) \ln b_i(k)}$

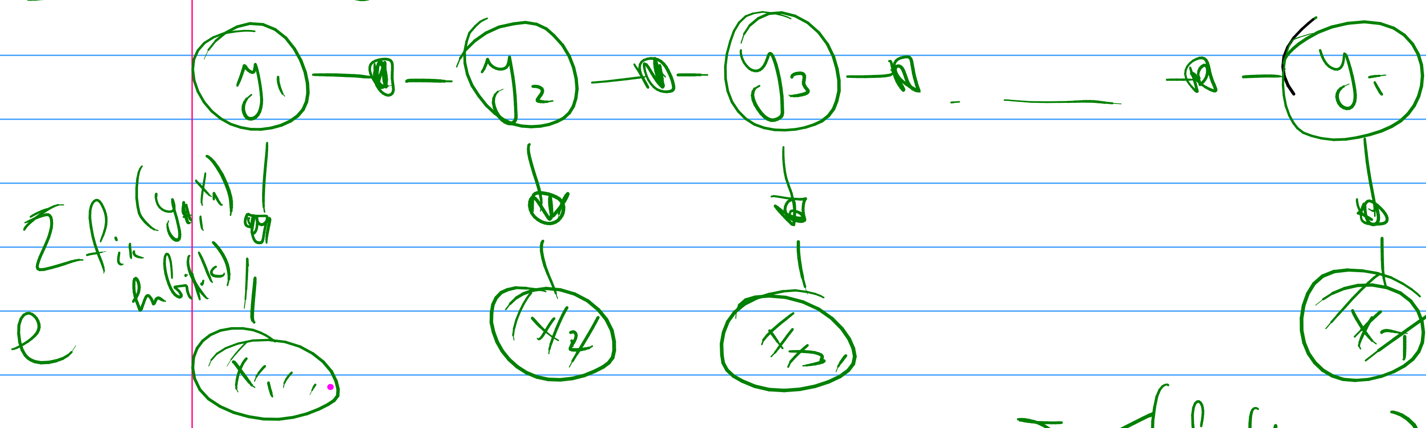
$$= e^{\sum_i \underbrace{f_i(y_1)}_{} \underbrace{\ln \pi_i}_{} + \sum \underbrace{f_{ij}(y_t, y_{t-1})}_{} \underbrace{\ln a_{ij}}_{} + \sum \underbrace{f_{ik}(y_t, x_t)}_{} \underbrace{\ln b_i(k)}_{}}$$

features

parameters

$e^{\sum_{ij} f_{ij}(y_1, y_2)\ln a_{ij}}$

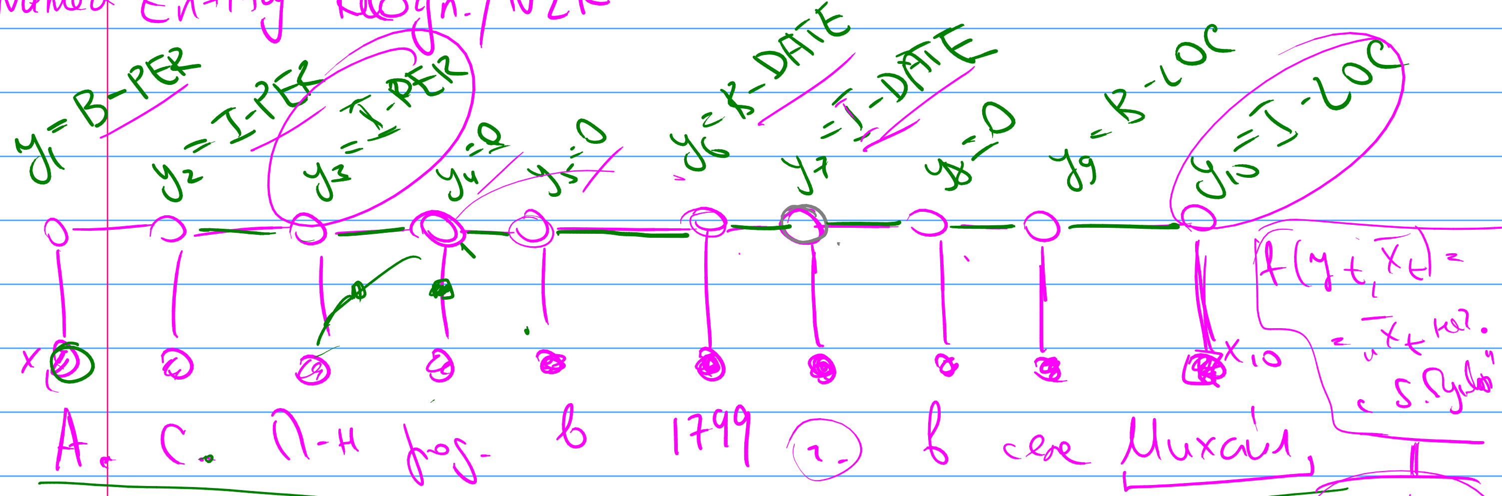
CRF



$$p(y_1 - y_T \mid x_1 - x_T) = \frac{1}{Z(\bar{x})} \cdot e^{\sum_t \sum_s \left(f_s(y_t, y_{t+1})\theta_s + f'_s(y_t, x_t)\theta'_s \right)}$$

$p(x_t \mid y_t)$, $p(y_t \mid y_{t-1})$

Named Entity Recogn./NER



$f(y_t, x_t) =$
$= x_t$ кап.
" S Ryl… "

$p(x_t \mid y_t)$

$p(x_t \mid y_{t-1}, y_t, y_{t+1})$

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

$$p(\bar{y} \mid \bar{x}) = \frac{1}{Z(\bar{x})} e^{\sum_s f_s(y_t, y_{t+1}, \bar{x}) \cdot \theta_s}$$

$\cancel{p(\bar{x})}$

Gen models

Explicit density

Implicit density

$\boxed{p(\bar{x})}$

Naive Bayes

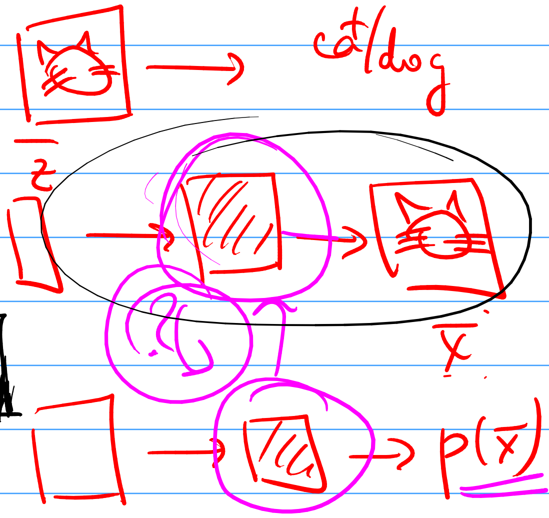
$$p(\bar{x},y)=p(y)\prod p(x_t|y)$$

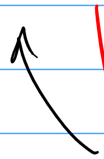
HMM

$$p(\bar{x},\bar{y})=p(y_1)p(y_2|y_1) \ldots$$

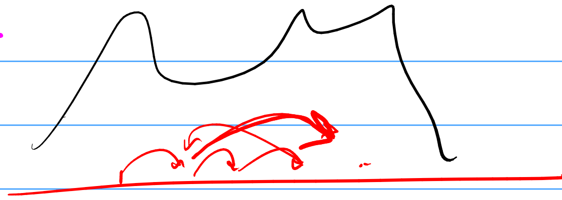
GAN

$\boxed{\text{GAN}}$

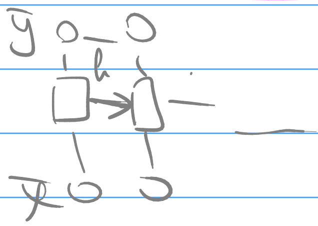




MCMC



RNN

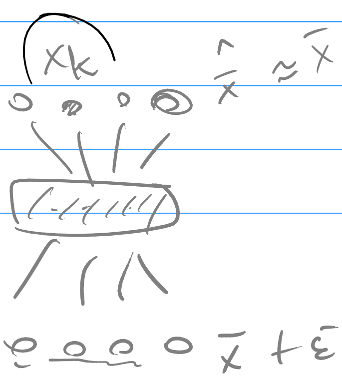


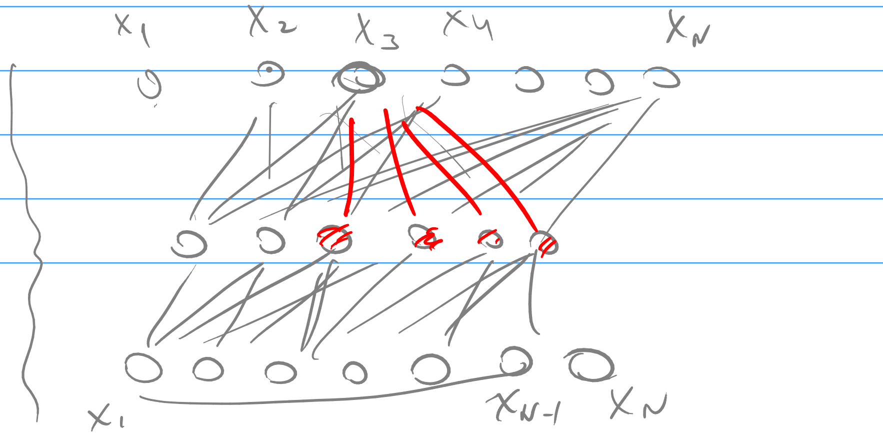
$$p(\bar{y}|\bar{x}) = p(y_1|x_1)\, p(y_2|y_1,x_1,x_2)\, p(y_3|y_2,y_1,x_1,x_2,x_3)$$

$$p(\bar{y}_k|\bar{x}_1 \ldots \bar{x}_{k-1}, \bar{y}_1, \ldots \bar{y}_{k-1}, \bar{x}_k) \approx p(\bar{y}_k|\bar{x}_k, \bar{h}_k)$$

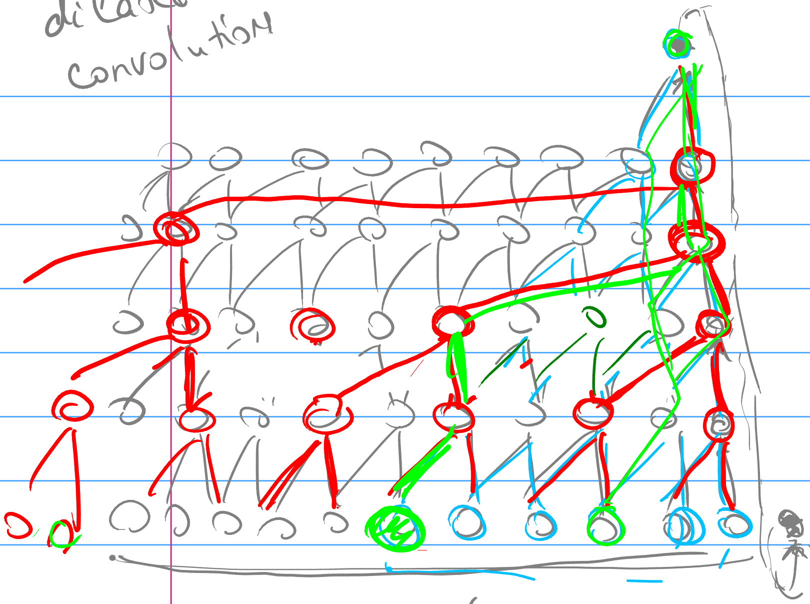
$$p(\bar{h}_k|\bar{x}_{k-1}, \bar{h}_{k-1})$$

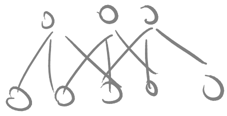
$$p(\bar{x}) = p(x_1)\,p(x_2|x_1) \ldots p(x_k|x_1 \ldots x_{k-1}) \ldots$$

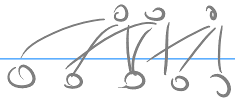




dilated convolution







16000/s