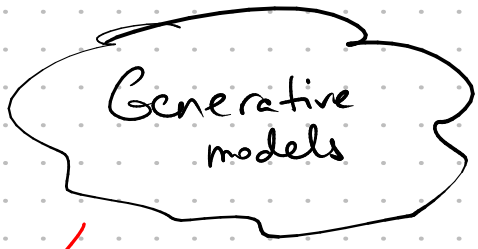
LR Discriminative $p(y|\bar{x})$ · ~~$p(\bar{x})$~~

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



→ Implicit density — GAN

$\bar{z} \sim N(\bar{0}, I)$ — Model — $\bar{x}_i \sim p(\bar{x})$

Explicit density

$p(\bar{x})$

NB: $p(\bar{x}, y) = p(y) \prod_i p(x_i|y)$

Tractable density

Flow-based models

$p(\bar{x}) \{ f_d \circ \dots \circ f_2 \circ f_1 \circ p_0 \}(\bar{x})$

Approximate density — $p(\bar{x}) \approx q(\bar{x}) = \prod \dots$ — VAG

Autoregressive models

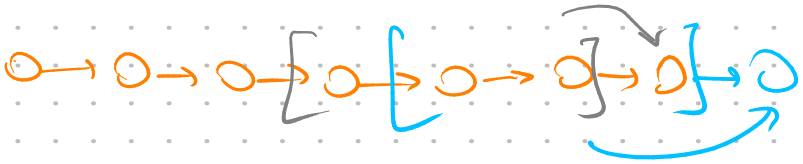
$$p(\bar{x}|\bar{\theta}) = p(x_1, x_2, \dots, x_d|\bar{\theta}) =$$

$$= p(x_1|\bar{\theta})\, p(x_2|x_1, \bar{\theta})\, p(x_3|x_1, x_2, \bar{\theta}) \dots p(x_d|x_1 \dots x_{d-1}, \theta)$$

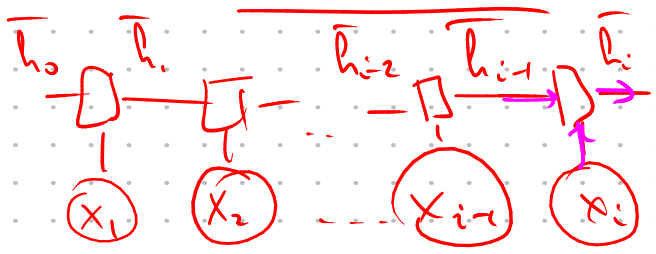
$$p(x_i|x_1, \dots, x_{i-1}, \bar{\theta}) \approx \dots$$

Sliding window

$$p(x_i|x_1 \dots x_{i-1}, \bar{\theta}) = p(x_i|x_{i-1}, \dots x_{i-k}, \bar{\theta})$$



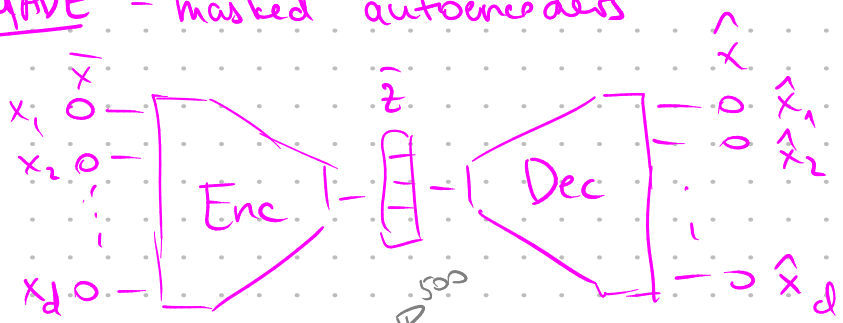
→ Hidden state



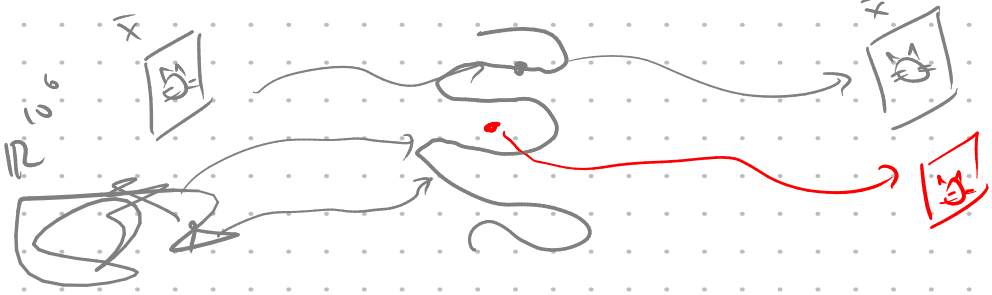
$$p(x_i|x_1 \dots x_{i-1}, \bar{\theta}) = p(x_i|\bar{h}_{i-1}, \bar{\theta})$$

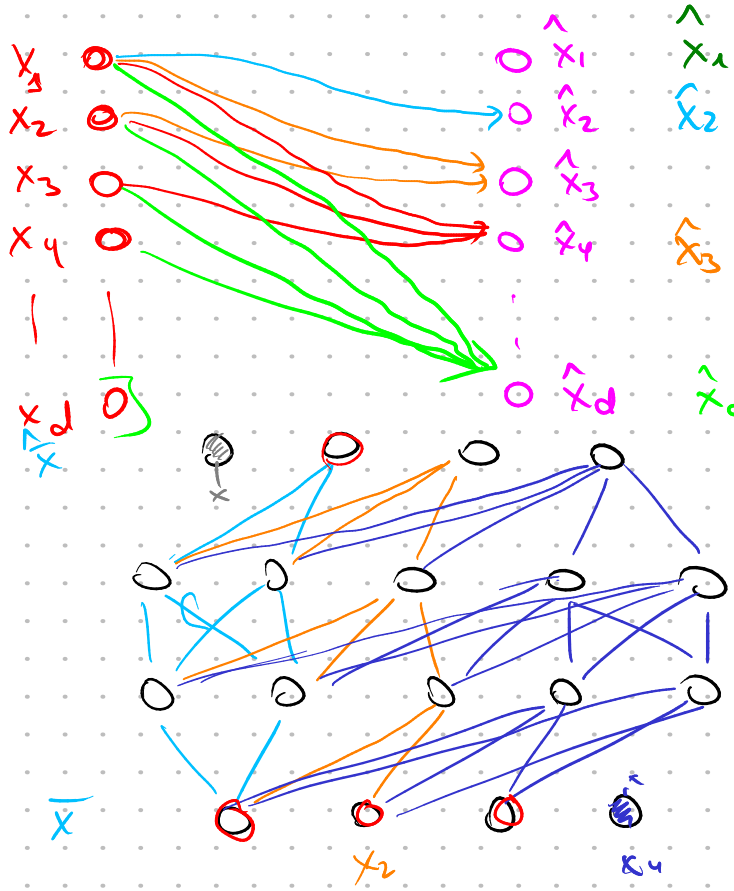
где $\bar{h}_i = f(x_i, \bar{h}_{i-1}, \bar{\theta})$

MADE - masked autoencoders



$$L = \|\hat{x} - \bar{x}\| \to \min$$





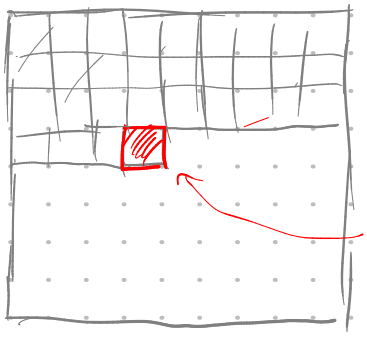
$$\hat{x}_1 \sim p(\hat{x}_1)$$

$$\hat{x}_2 \sim p(\hat{x}_2 \mid x_1)$$

$$\|\hat{x} - \bar{x}\| \to \min$$

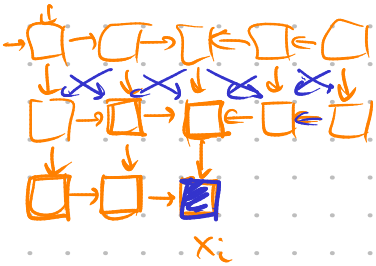
$$\hat{x}_3 \sim p(\hat{x}_3 \mid x_1, x_2)$$

$$\hat{x}_d \sim p(\hat{x}_d \mid x_1 \ldots x_{d-1})$$

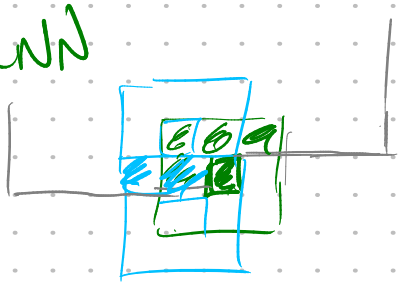


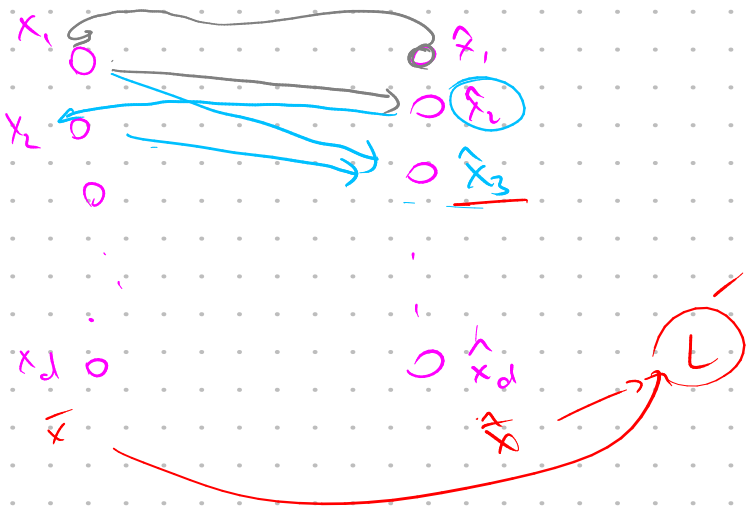
$$x_i \sim p(x_i \mid x_1 \ldots x_{i-1}, \theta)$$

PixelRNN

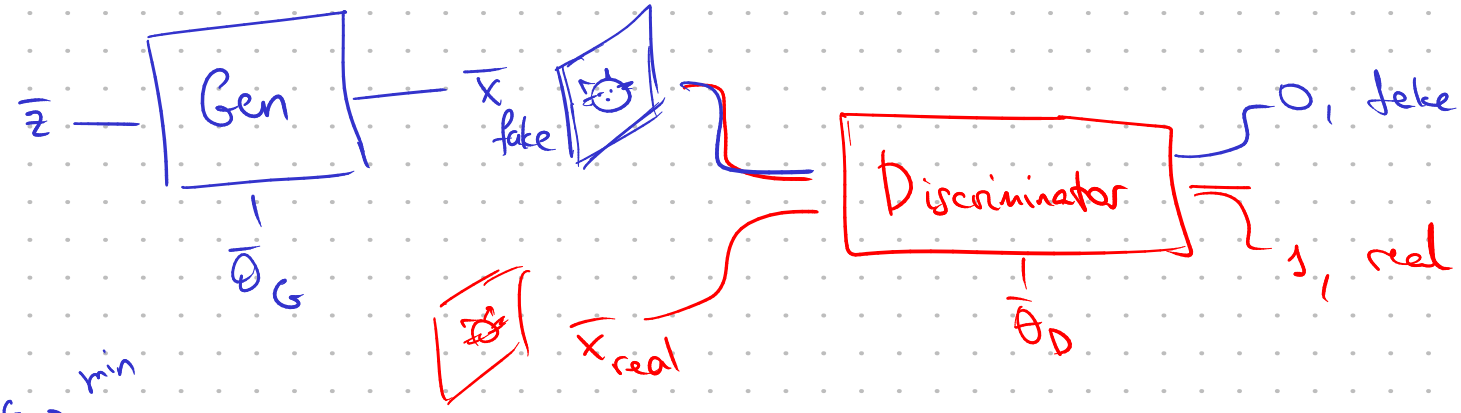


Pixel CNN





Generative Adversarial Networks



$$\bar{\theta}_G \xrightarrow{\min} L_D = \underbrace{\mathbb{E}_{\bar{x}_{real} \sim D}\left[\log D(\bar{x})\right]} + \underbrace{\mathbb{E}_{\bar{x}_{fake} \sim p_g(\bar{x})}\left[\log\left(1 - D(\bar{x})\right)\right]} \xrightarrow[\bar{\theta}_D]{\max}$$

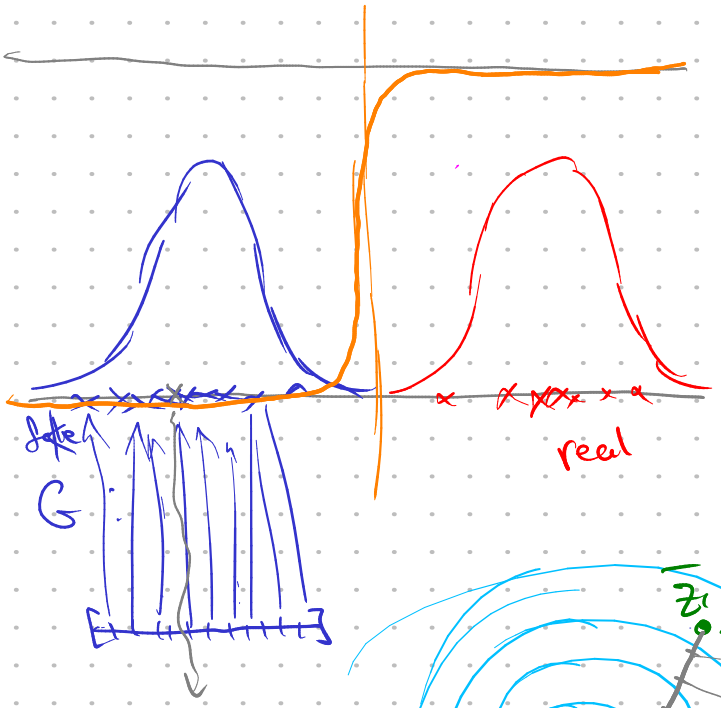
$$L_G = \mathbb{E}_{\bar{z} \sim p(\bar{z})}\left[\log\left(1 - D(G(\bar{z}))\right)\right] \xrightarrow[\bar{\theta}_G]{\min}$$

$$\min_{\bar{\theta}_G} \max_{\bar{\theta}_D} L(\bar{\theta}_D, \bar{\theta}_G)$$

- fix $\bar{\theta}_G$, train $\bar{\theta}_D$
- fix $\bar{\theta}_D$, train $\bar{\theta}_G$

fix G.

$$D_G^*(\bar{x}) = \frac{p_{data}(\bar{x})}{p_{data}(\bar{x}) + p_g(\bar{x})} \Rightarrow G^* : p_G(\bar{x}) = p_{data}(\bar{x})$$



$$L_G = -\mathbb{E}_{\bar{z}}\left[\log D(G(\bar{z}))\right]$$