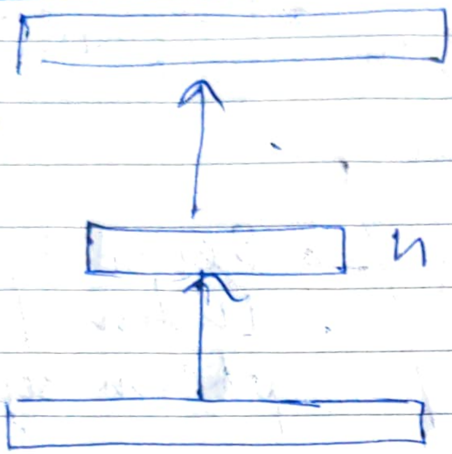
Auto Encoders:



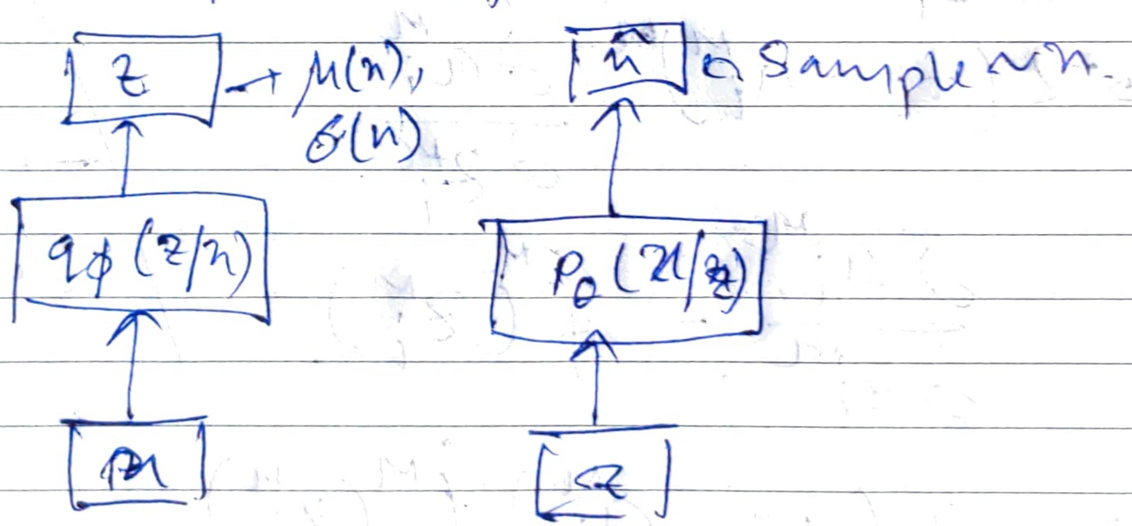
$\hat{x}_i = f(w^* h + c)$

$h = g(w x_i + b)$

x_i

VAE:

- Encoder
- Decoder
- Regularization loss function



Training Data

Sample using $\mu(x)$ & $\sigma(x)$

loss function:

$$-D_{KL}\left[q_\phi(z|x) \,\|\, p_\theta(z)\right] + E_{q_\phi(z|x)}\left[\log(p_\theta(x|z))\right]$$

regularizer. $\hookrightarrow N(0, I)$

data reconstr. loss.

$q_\phi(z|x) \to \mu(x), \sigma(x)$ Gaussian

constrain to be aligned.

$\log(p_\theta(x|z))$

$\hookrightarrow$ output of Decoder.

$$KL\left[N(\mu(x), \Sigma(x)) \,\|\, N(0, I)\right] + \|x - f(z)\|^2.$$

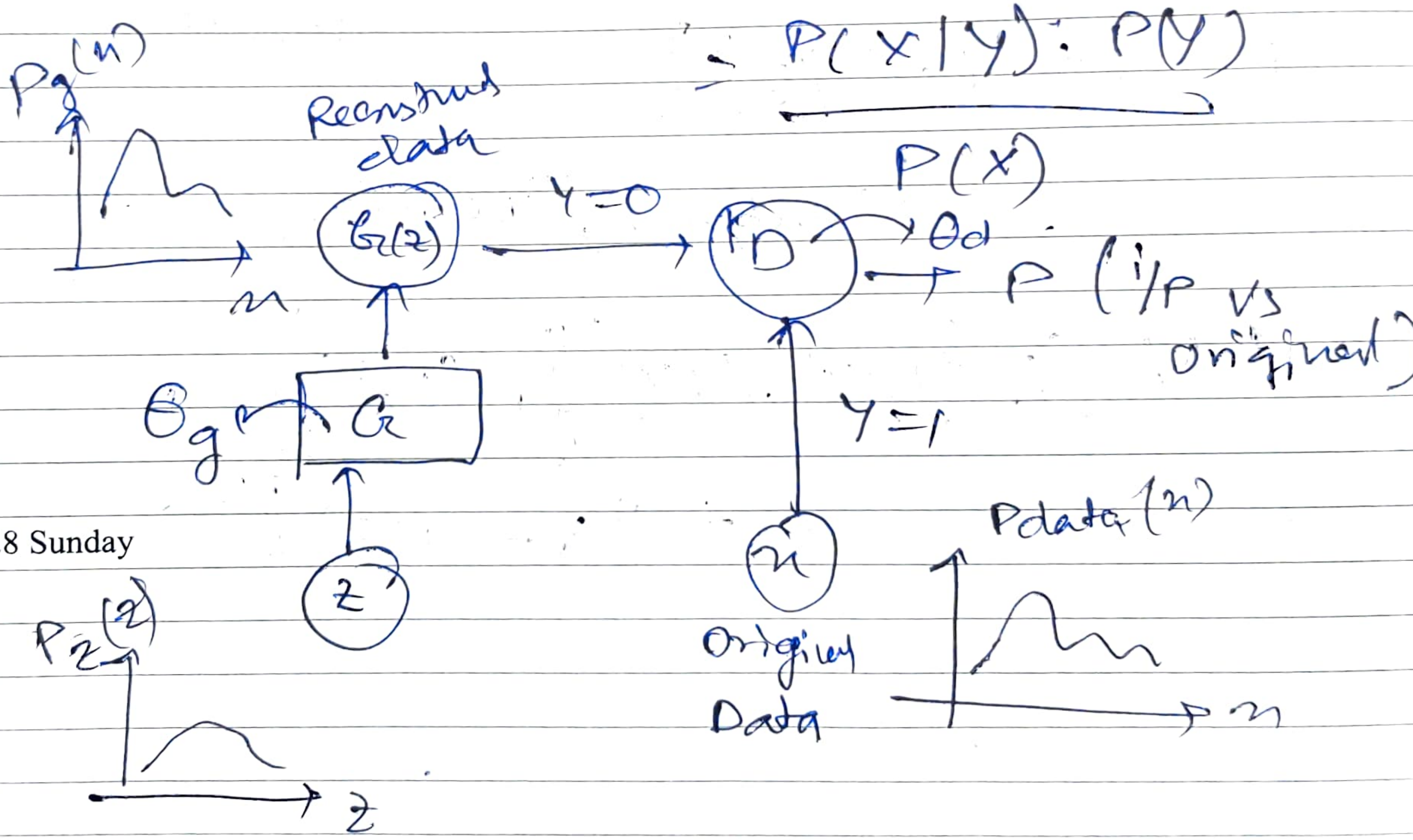
GANs

Generator : $P(X, Y) = P(X|Y) \cdot P(Y)$

Discriminator. $P(Y | X = x)$

Bays term $\rightarrow$ $P(Y|X) = \dfrac{P(X, Y)}{P(X)}$

$= \dfrac{P(X|Y) \cdot P(Y)}{P(X)}$



28 Sunday

Minimax Game:

Value function:

$$\min_{G} \max_{D} V(G,D) = E_{x\sim P_{data}}\left[\ln(D(x))\right] + E_{z\sim P_z}\left[\ln(1-D(G(z)))\right]$$

Binary Cross Entropy:

$$\mathcal{L} = -\sum y\ln\hat{y} + (1-y)\ln(1-\hat{y})$$

where, $y=1$, $\hat{y} = D(x) \Rightarrow \mathcal{L} = \ln[D(x)]$

$y=0$, $\hat{y} = D(G(z)) \Rightarrow \mathcal{L} = \ln[1-D(G(z))]$

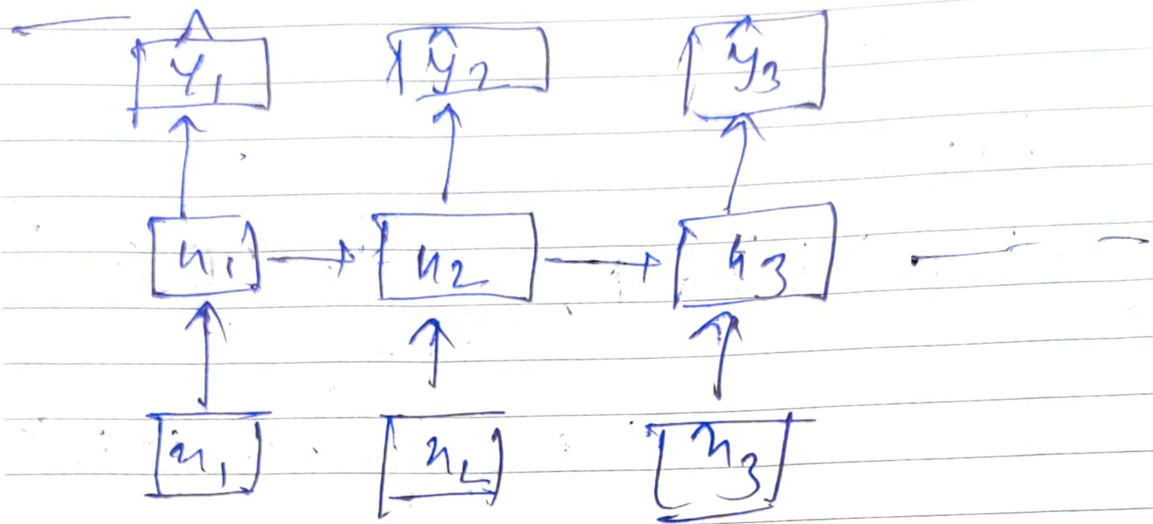
Adding, $\mathcal{L} = \ln[D(x)] + \ln[1-D(G(z))]$

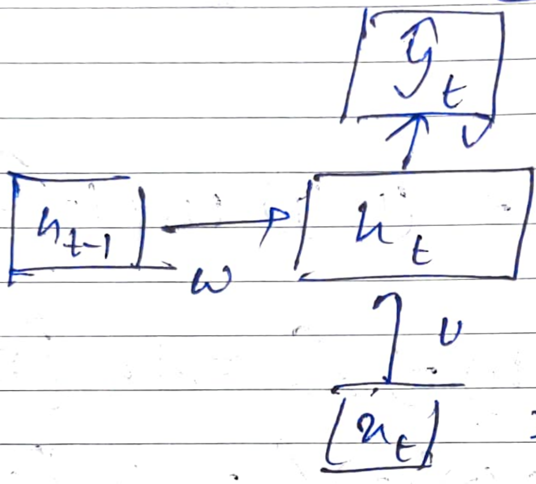
$$E(\mathcal{L}) = E[\ln D(x)] + E(\ln[1-D(G(z))])$$

$$= \sum P_{data}(x)\ln[D(x)] + \sum p_z(z)\ln(1-D(G(z)))$$

$$= \int P_{data}(x)\ln(D(x))dx + \int P_z(z)\ln(1-D(G(z)))dz$$

RNN.



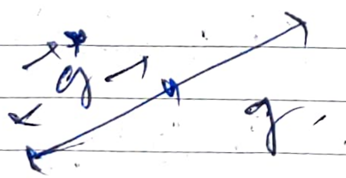


$h_t = f_w(h_{t-1}, x_t)$

$\Rightarrow h_t = \tanh(\overbrace{W_{hh}}^{W} h_{t-1} + \underbrace{W_{xh}}_{U} x_t + b)$

$\rightarrow \hat{y}_t = g(\underbrace{W_{hy}}_{V} h_t)$

Exploding Grad : Grad Clipping.



$$\vec{g}^* = \frac{\vec{g}}{\|\vec{g}\|} G_{max}$$

Vanishing Grad: GRU, LSTM.

LSTM:

Vanilla RNN

from GRU: $h_t = f \odot h_{t-1} + (1-f) \odot g$

Memory cell:

↳ c_t forget input

$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

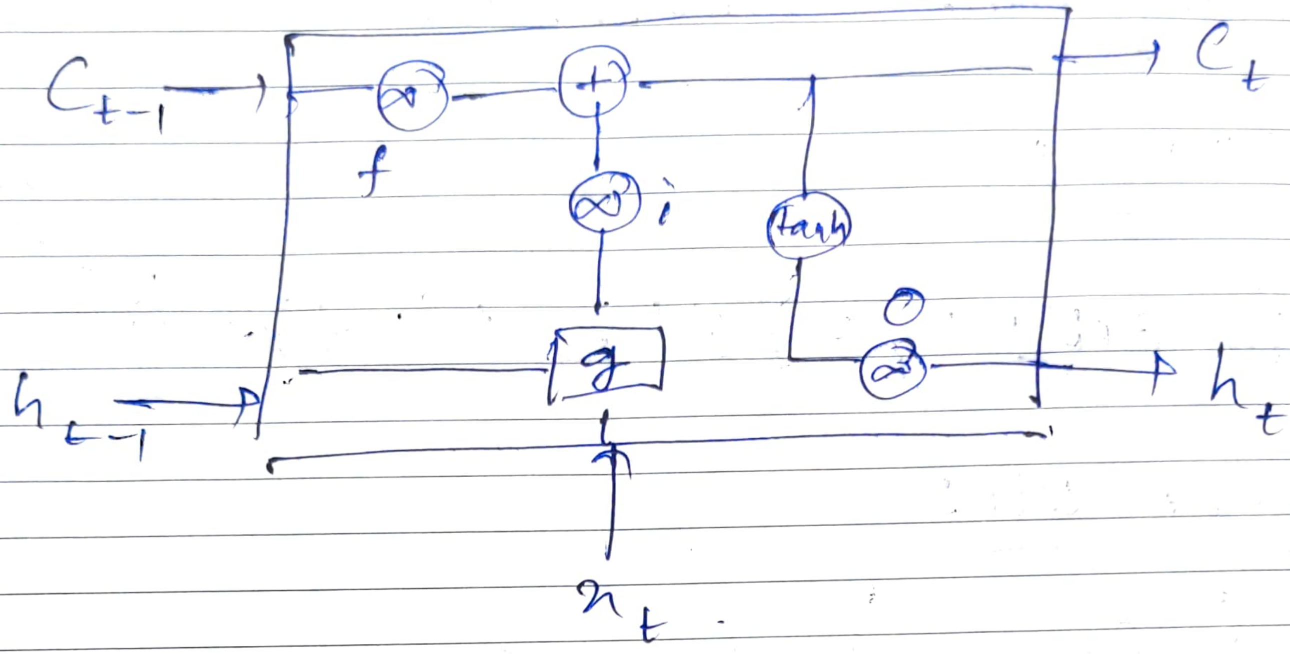
output $o \in [0,1]$

$$o = \sigma(z_o)$$ $\cancel{z_o = W_o h_{t-1} + U_h + b}$

$$f = \sigma(z_f)$$ $z_h = W_h h_{t-1} + U_h x_t$

$$i = \sigma(z_i)$$

$$g = \tanh(z_g)$$



C_{t-1}
C_t
f
i
tanh
o
g
h_{t-1}
h_t
x_t