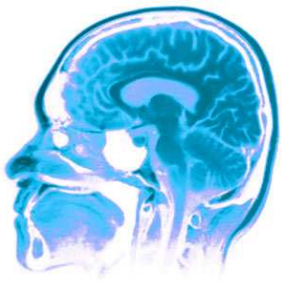
CPSC540

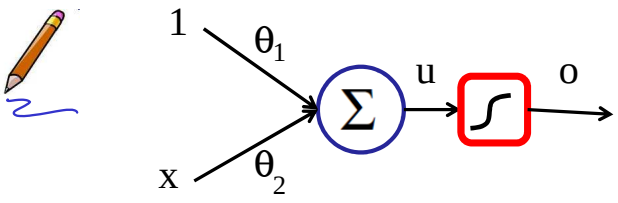


Feedforward Neural Networks

aka Multilayer Perceptrons

Nando de Freitas
October, 2011
University of British Columbia

MLP – 1 neuron

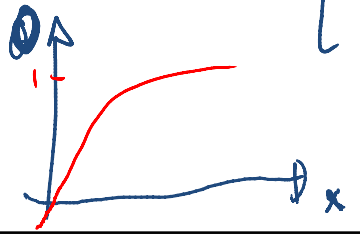


x	o
2.2	0
3.3	0
5	1

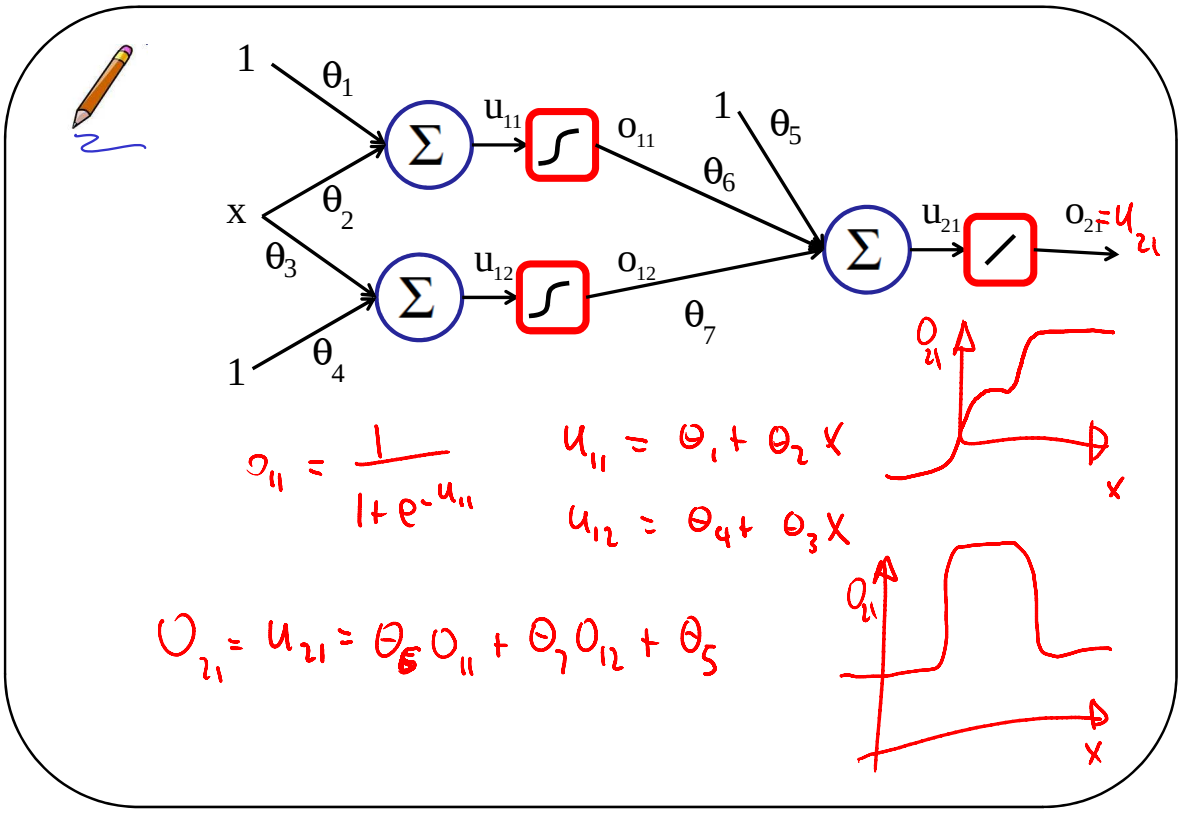
$$u = \theta_1 + \theta_2 x \qquad o = \frac{1}{1+e^{-u}}$$

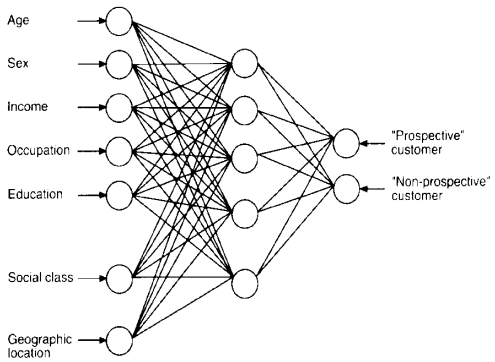
$$P(o_i|\theta_1,\theta_2,x_i) = \pi_i^{o_i}(1-\pi_i)^{1-o_i} = \begin{cases} P(o_i|\theta_1,\theta_2,x_i) = \pi_i \\ \end{cases}$$

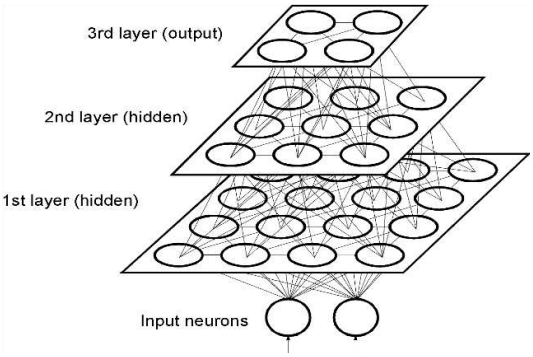
$$\pi_i = \frac{1}{1+e^{-x_i\theta_2-\theta_1}}$$

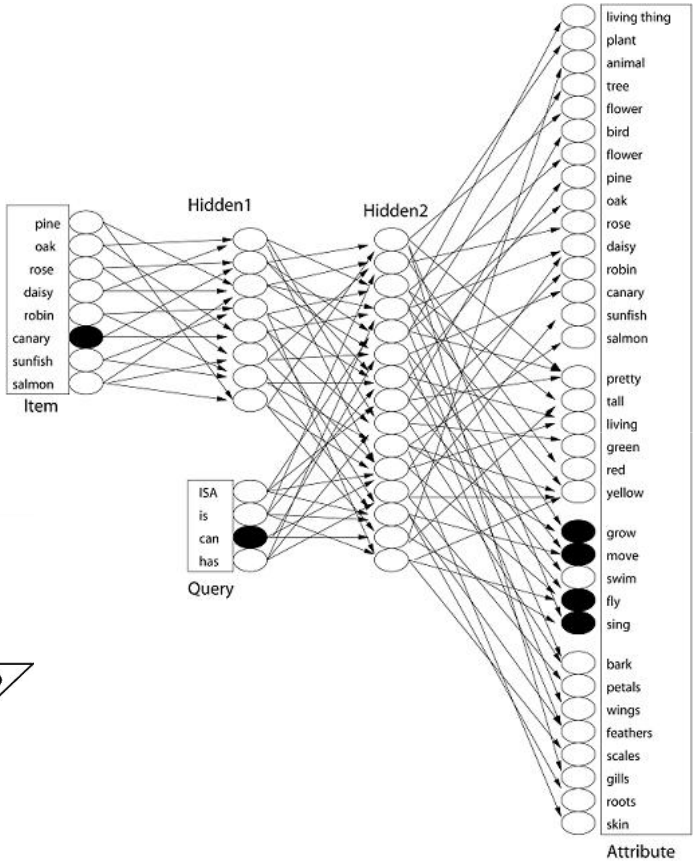


MLP – 3 neurons, 2 layers

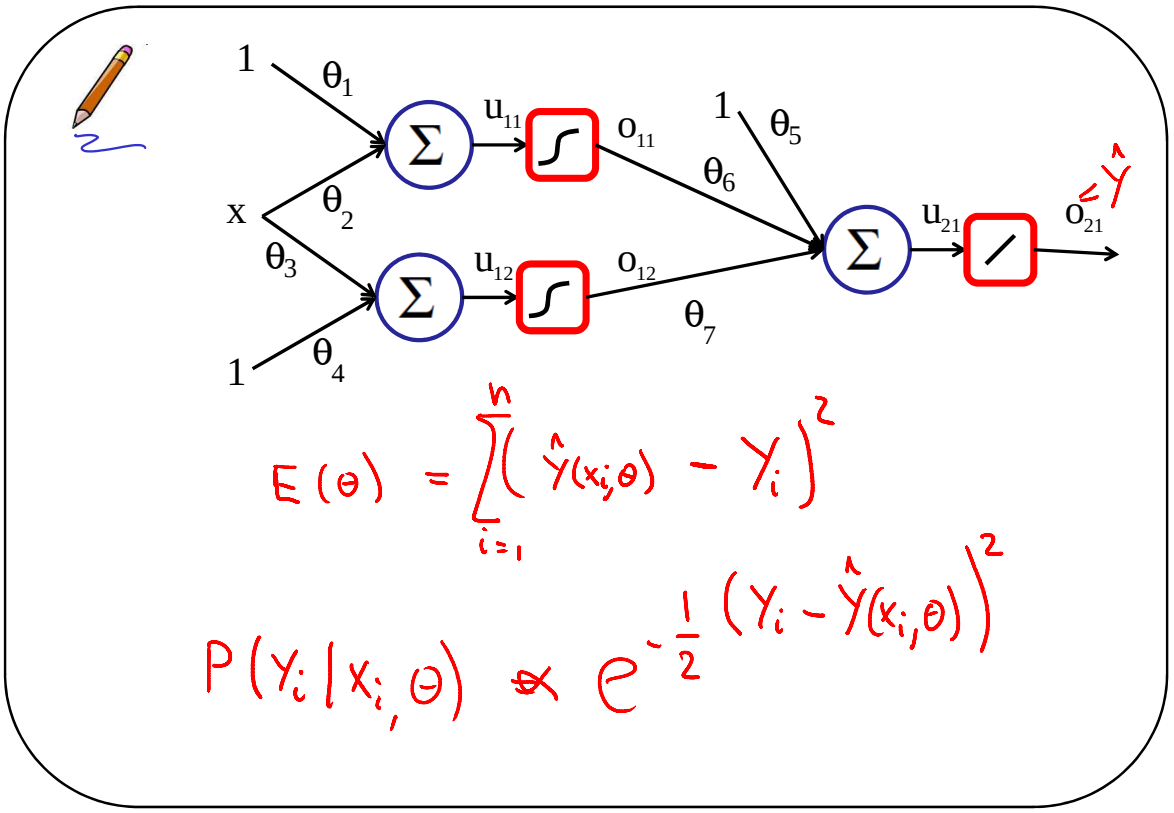








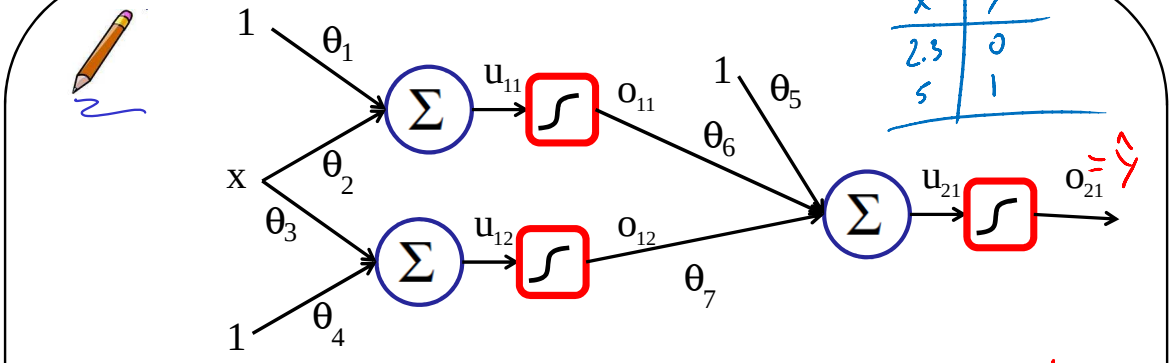
MLP – Regression



$$E(\theta) = \sum_{i=1}^{n} \left(\hat{y}(x_i;\theta) - y_i \right)^2$$

$$P(y_i \mid x_i, \theta) \propto e^{-\frac{1}{2}\left(y_i - \hat{y}(x_i;\theta)\right)^2}$$

MLP – Binary classification $\hat{y} \in [0,1]$



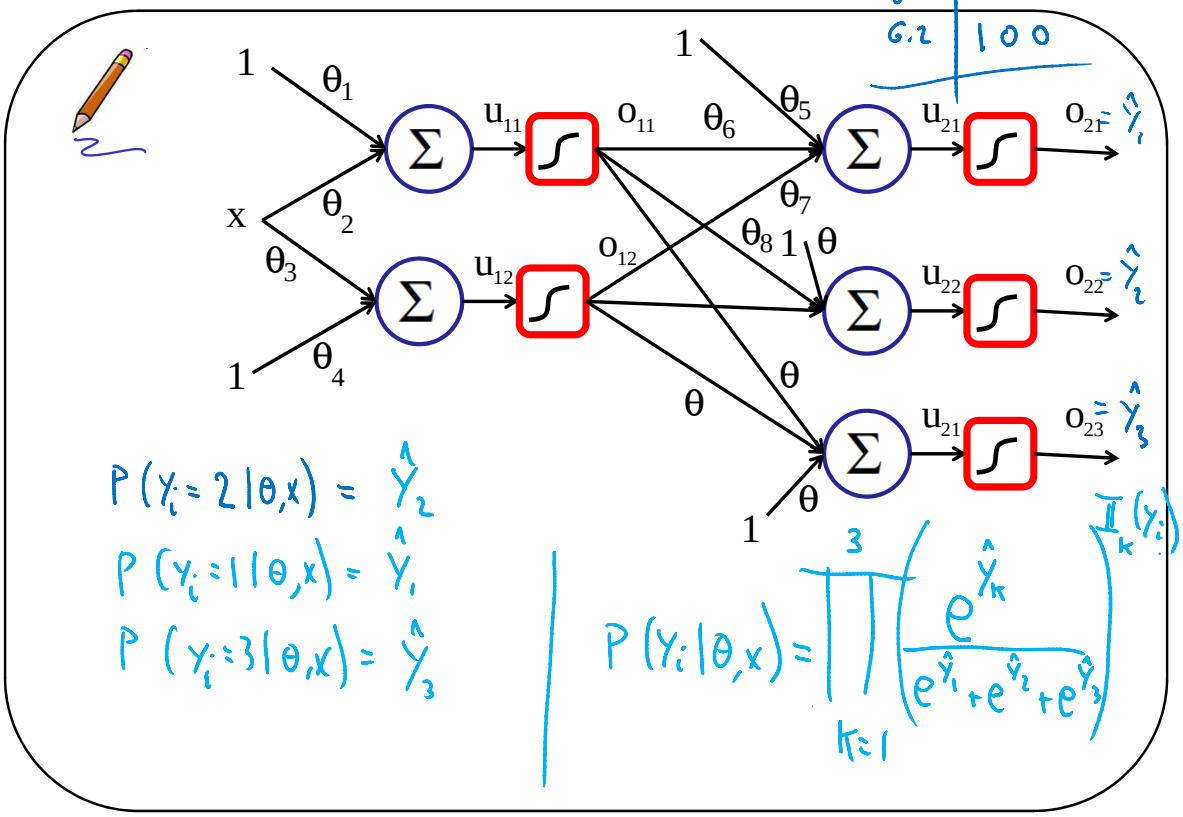
x	y
2.3	0
5	1

$$P(y_i \mid x_i, \theta) = \left[\hat{y}(x_i;\theta)\right]^{y_i} \left[1 - \hat{y}(x_i;\theta)\right]^{1-y_i}$$

$$P(y_{1:n} \mid x_{1:n}, \theta) = \prod_{i=1}^{n} P(y_i \mid x_i, \theta)$$

$$E(\theta) = -\sum_{i=1}^{n} y_i \log \hat{y}(x_i;\theta) + (1-y_i) \log\left(1 - \hat{y}(x_i;\theta)\right)$$

MLP – Multiclass



x	y
3	0 0 1
6.2	1 0 0

$P(y_i = 2 \mid \theta, x) = \hat{y}_2$

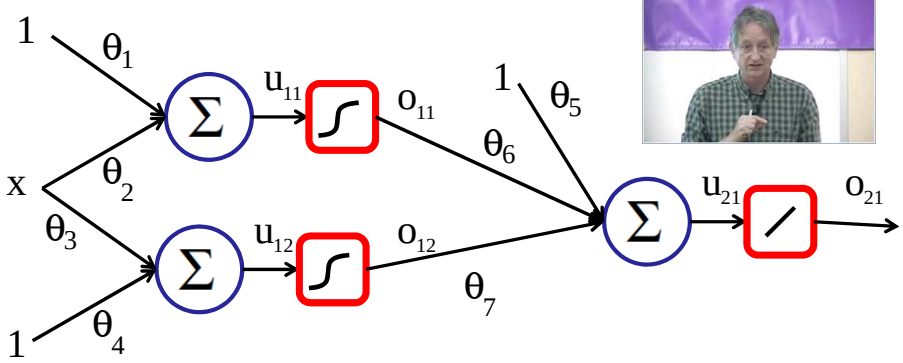
$P(y_i = 1 \mid \theta, x) = \hat{y}_1$

$P(y_i = 3 \mid \theta, x) = \hat{y}_3$

$$P(y_i \mid \theta, x) = \prod_{k=1}^{3} \left(\frac{e^{\hat{y}_k}}{e^{\hat{y}_1} + e^{\hat{y}_2} + e^{\hat{y}_3}} \right)^{\mathbb{I}_k(y_i)}$$

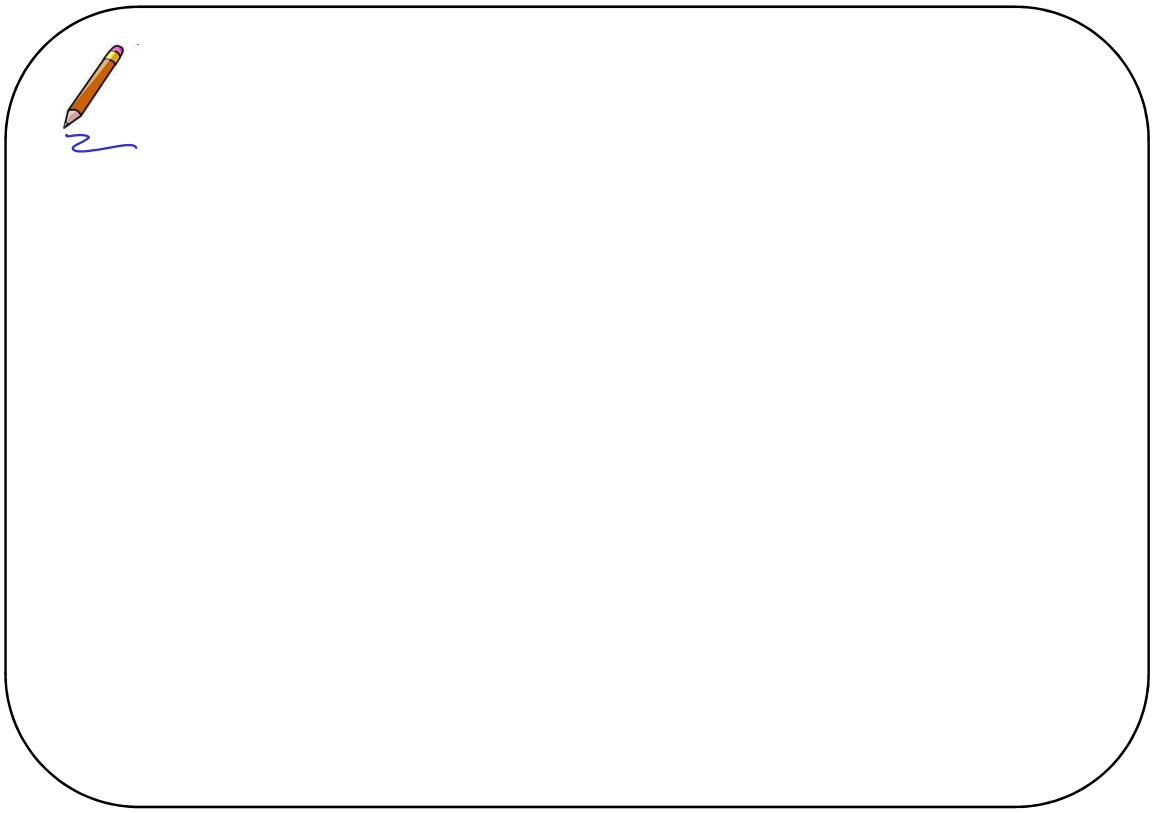
MLP - Multiclass

Backpropagation

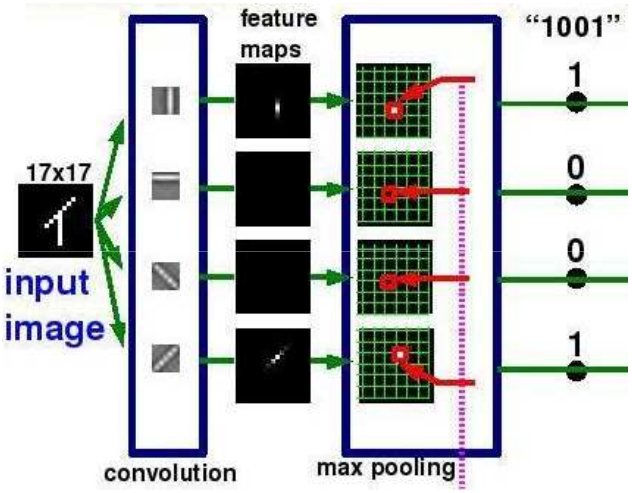


Backpropagation

Backpropagation

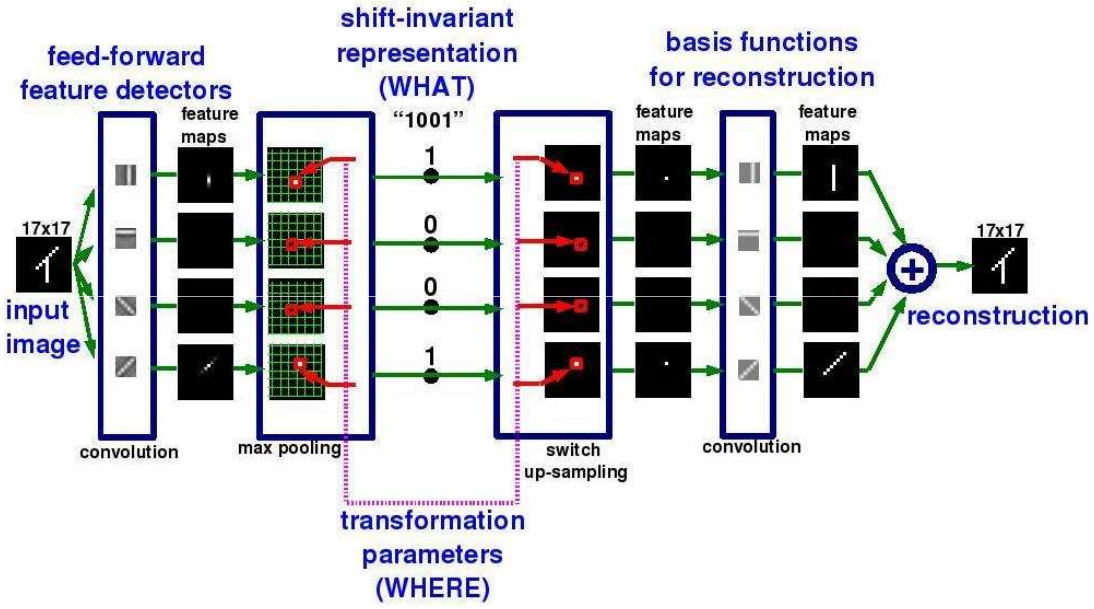


Convolutional/pooling architectures



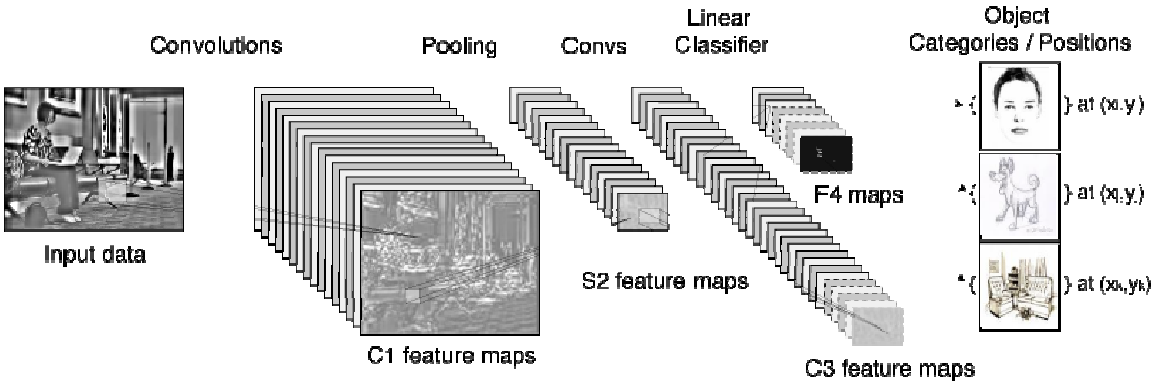
[Marc'Aurelio Ranzato & Yann LeCun]

Autoencoders



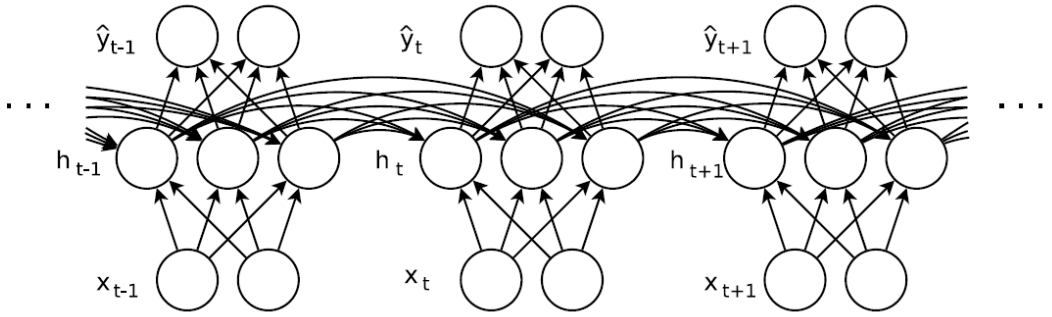
[Marc'Aurelio Ranzato & Yann LeCun]

Convolutional/pooling architectures



[Yann LeCun]

Recurrent neural networks

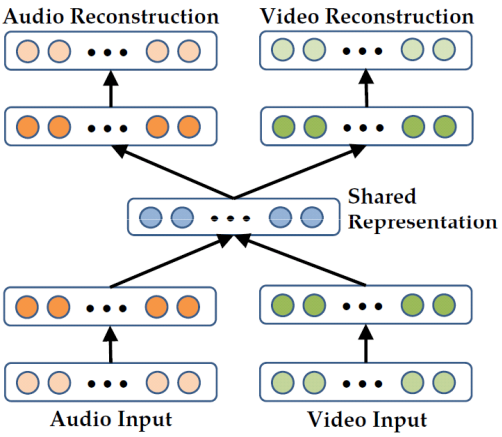


phrase “The meaning of life is”:

The meaning of life is the tradition of the ancient human reproduction: it is less favorable to the good boy for when to remove her bigger. In the show’s agreement unanimously resurfaced. The wild pasteured with consistent street forests were incorporated by the 15th century BE. In 1996 the primary rapford undergoes an effort that the reserve conditioning, written into Jewish cities, sleepers to incorporate the .St Eurasia that activates the population. Mar??a Nationale, Kelli, Zedlat-Dukastoe, Florendon, Ptu’s thought is. To adapt in most parts of North America, the dynamic fairy Dan please believes, the free speech are much related to the

[James Martens & Ilya Sutskever]

Deep learning examples



(a) ICDAR test image (b) Text detector scores

(c) ICDAR test image (d) Text detector scores

[Andrew Ng’s group]