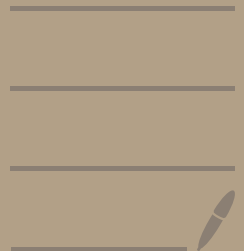
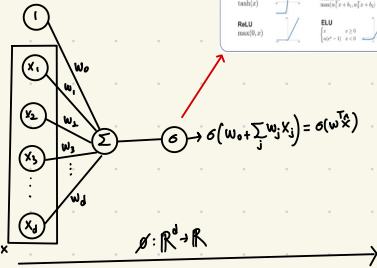


Deep Learning

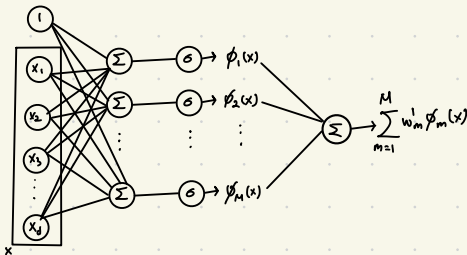


Neural Networks

Today	
Sigmoid $\sigma(x) = \frac{1}{1 + e^{-x}}$	Leaky ReLU $\max(0, x)$
tanh $\tanh(x)$	Maxout $\max(w_1x + b_1, w_2x + b_2)$
ReLU $\max(0, x)$	ELU $\begin{cases} x & \text{if } x \geq 0 \\ \alpha e^{-x} & \text{if } x < 0 \end{cases}$



Multi-Layer-Perceptron (MLP) - One Hidden layer



$$f(x) = \sum_{m=1}^M w_m^1 \phi_m(x)$$

• The problem with fully connected networks (MLP)

- A 256x256 (RGB) image \Rightarrow ~200k dimensional input x

- A fully connected network would need a very large # of parameters, very likely to overfit.

- Generic deep network also doesn't capture the 'natural' invariances we expect in images.

\Rightarrow

