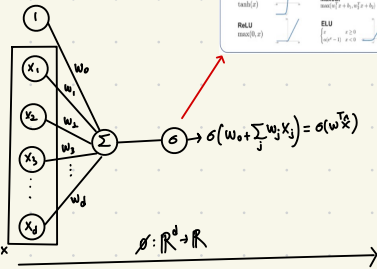
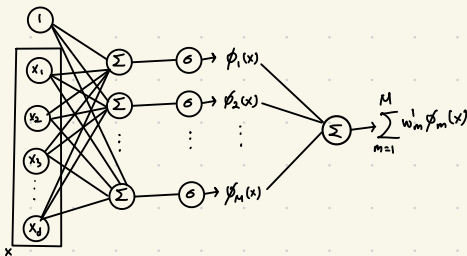


Neural Networks

Today	
Sigmoid $\sigma(x) = \frac{1}{1 + e^{-x}}$	Leaky ReLU $\max(0, x) + \alpha \min(0, x)$
tanh $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	Maxout $\max_j (w_j^T x + b_j)$
ReLU $\max(0, x)$	ELU $\begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases}$



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



$$f(x) = \sum_{m=1}^M w_m^1 \phi_m(x) = \sum_{j=1}^d W_{jm}^0 x_j$$

• The problem with fully connected networks (MLP)

- A 256×256 (RGB) image $\Rightarrow \approx 200k$ 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

