Going deeper with LEGOs
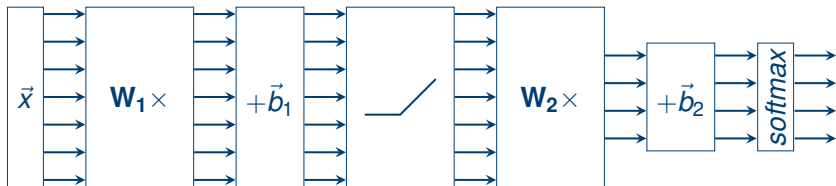
Making things *deep* is now easy...

$$C = \text{argmax}_j \left\{ \textit{softmax} \left(\mathbf{W_2} \textit{ReLU} \left(\mathbf{W_1} \vec{x} + \vec{b}_1 \right) + \vec{b}_2 \right)_j \right\}$$

N.B. the ReLU is *important*! A composition of linear functions is itself a linear function...