



$$\begin{aligned}
 &= \sigma \left( w_{1,0}a_0^{(0)} + w_{1,1}a_1^{(0)} + \dots + w_{1,n}a_n^{(0)} + b_1^{(0)} \right) \\
 &= \sigma \left( \sum_{i=1}^n w_{1,i}a_i^{(0)} + b_1^{(0)} \right)
 \end{aligned}$$

$$\begin{pmatrix} a_1^{(1)} \\ a_2^{(1)} \\ \vdots \\ a_m^{(1)} \end{pmatrix} = \sigma \left[ \begin{pmatrix} w_{1,0} & \dots & w_{1,n} \\ w_{2,0} & \dots & w_{2,n} \\ \vdots & \ddots & \vdots \\ w_{m,0} & \dots & w_{m,n} \end{pmatrix} \begin{pmatrix} a_1^{(0)} \\ a_2^{(0)} \\ \vdots \\ a_n^{(0)} \end{pmatrix} + \begin{pmatrix} b_1^{(0)} \\ b_2^{(0)} \\ \vdots \\ b_m^{(0)} \end{pmatrix} \right]$$

$$\mathbf{a}^{(1)} = \sigma \left( \mathbf{W}^{(0)} \mathbf{a}^{(0)} + \mathbf{b}^{(0)} \right)$$

$$\mathbf{y} = \sigma (\mathbf{W}\mathbf{x} + \mathbf{b}), \quad \text{for a single hidden layer.}$$