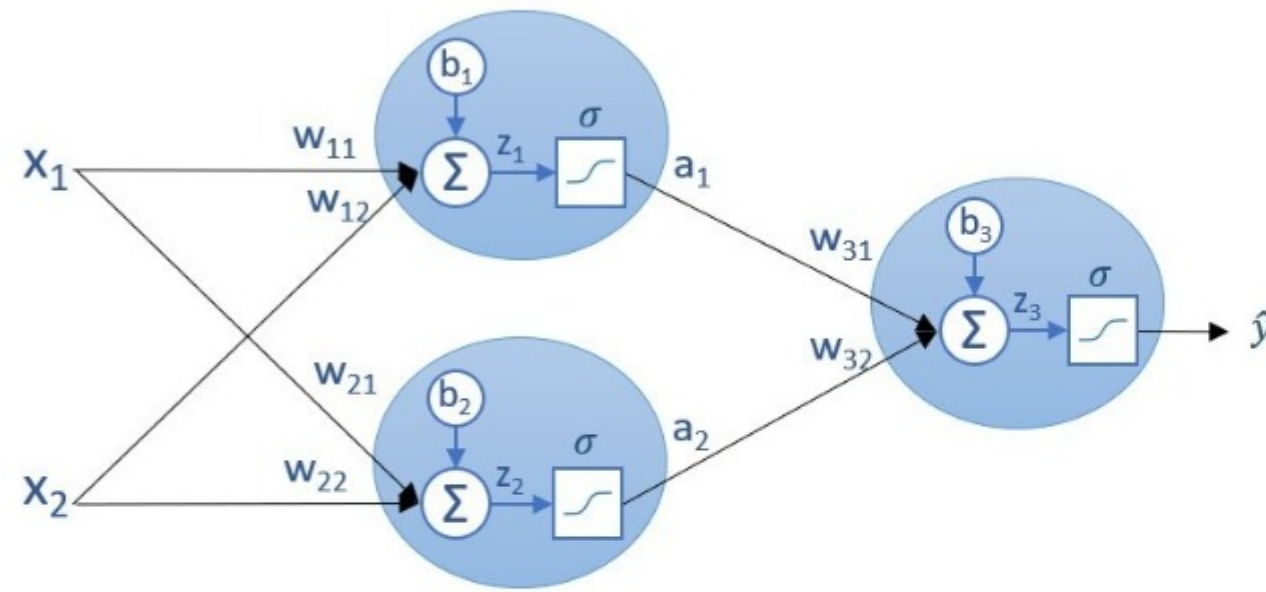


Utilizar varias neuronas con funciones de activación lineal no mejora la expresividad de la red, ya que sigue estando limitada a resolver únicamente problemas linealmente separables:

$$f(x) = \frac{e^x - 1}{e^x + 1}$$

$$Y' = F1(X1W1 + b1)$$



$$Y' = F2(\mathbf{X2W2} + b2)$$

$$Y' = F2(F1(\mathbf{X1W1} + \mathbf{b1}) W2 + B2)$$

$$\begin{aligned} F1(x) = c1x + d1 \quad \rightarrow \quad F(x) = F2(F1(x)) &= \\ &= c2(c1x + d1) + d2 \\ &= c2c1x + c2d1 + d2 \end{aligned}$$

$$F2(x) = c2x + d2$$