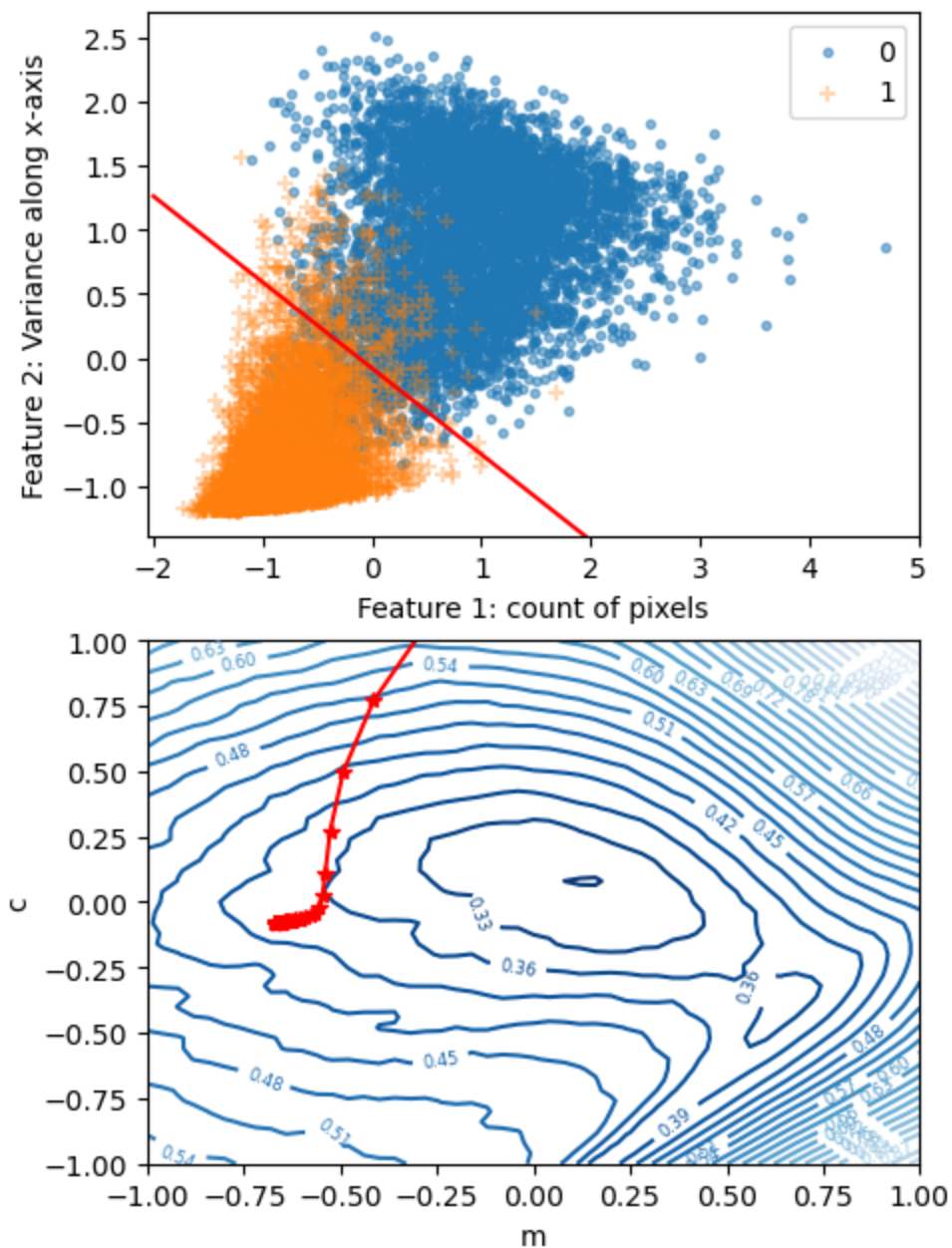


Multi layer Perceptron

All figures are from Chapter 3 of UDLBook. <https://github.com/udlbook/udlbook>

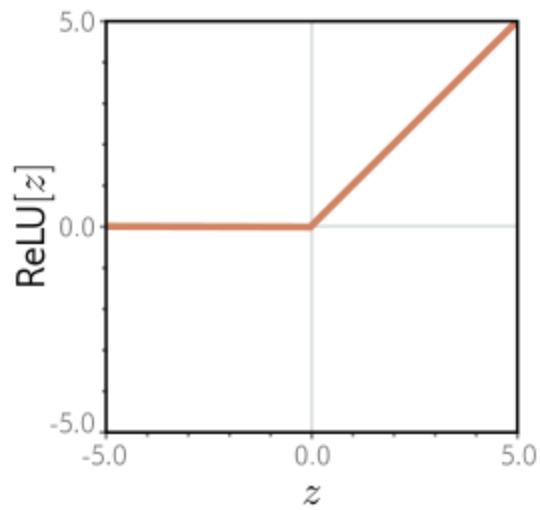
Recall the single layer perceptron



$$l = f(x) = \mathbf{w}^\top \mathbf{x} + w_0$$

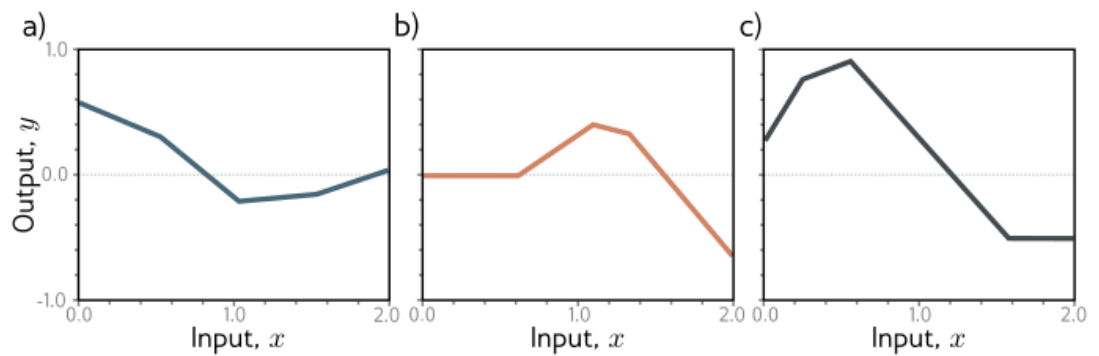
Multi Layer Perceptrons

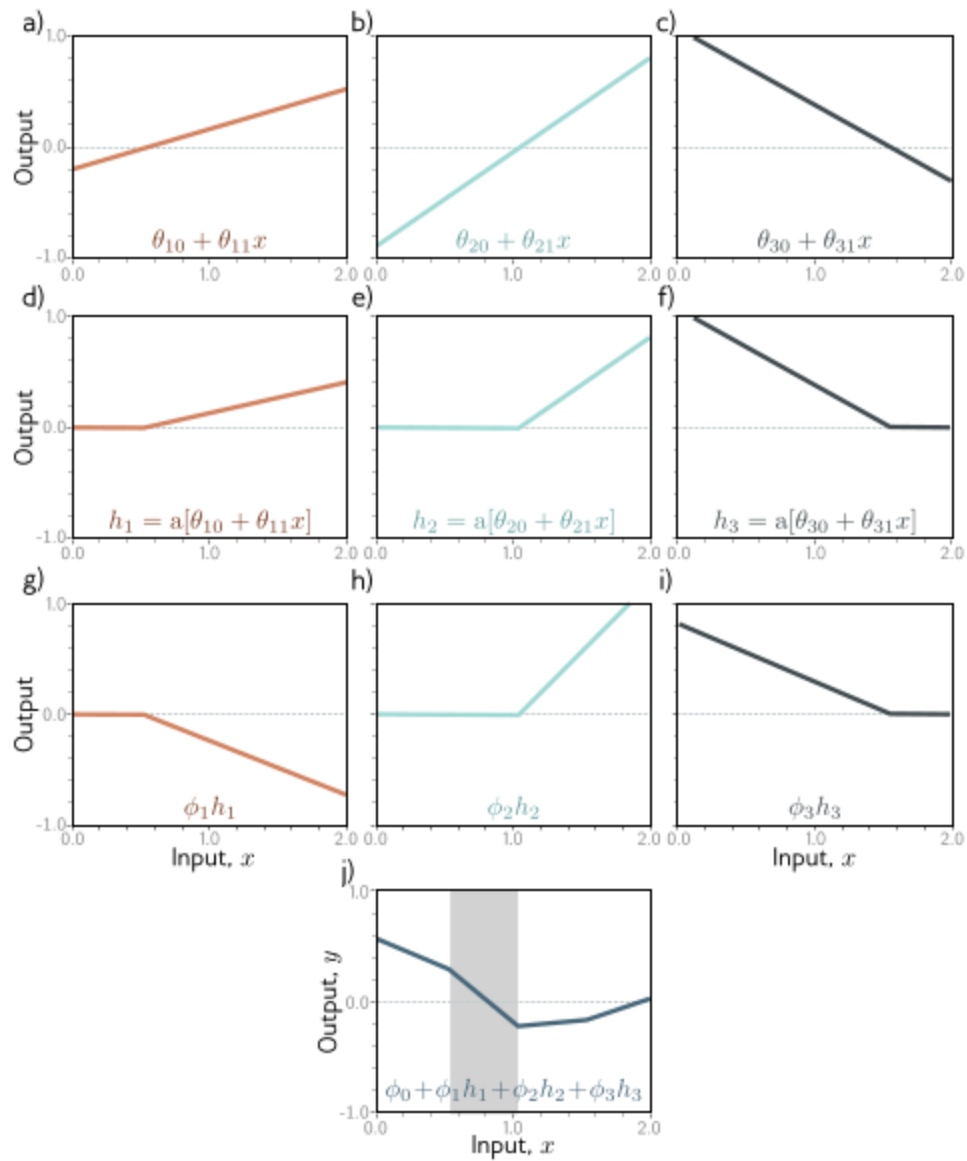
ReLU activation function $\text{ReLU}(z) = \max\{0, z\}$



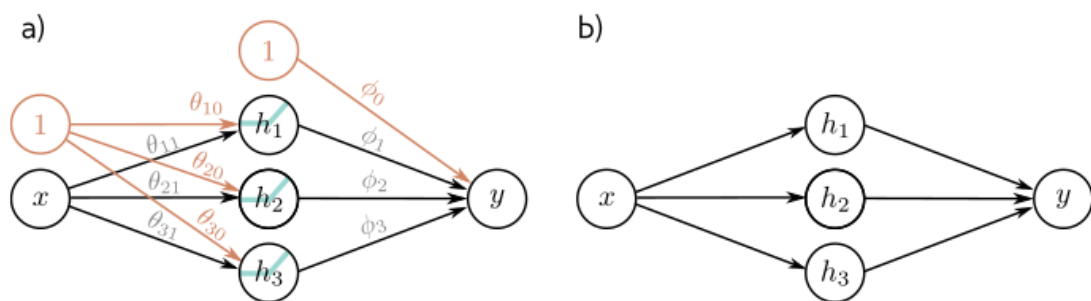
Two layer Perceptron

$$y = f(x) = w_2^\top \text{ReLU}(W_1 x) + w_{20}$$

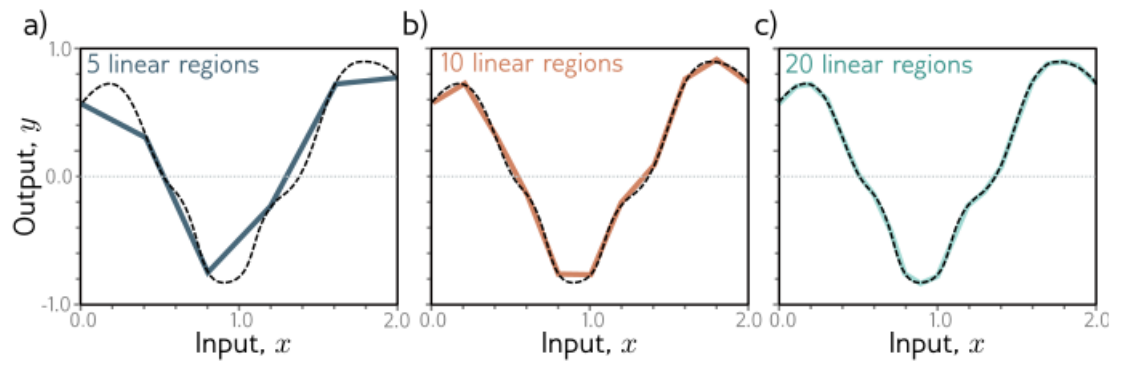




Depicting Neural Networks



Universal Approximation Theorem



Multivariate outputs

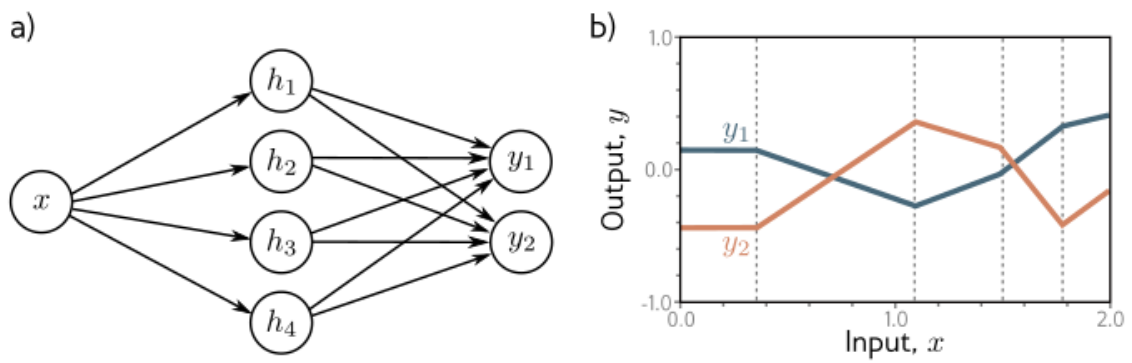
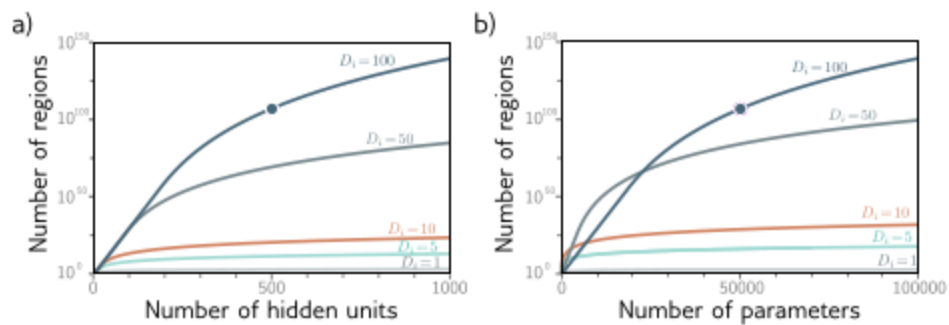
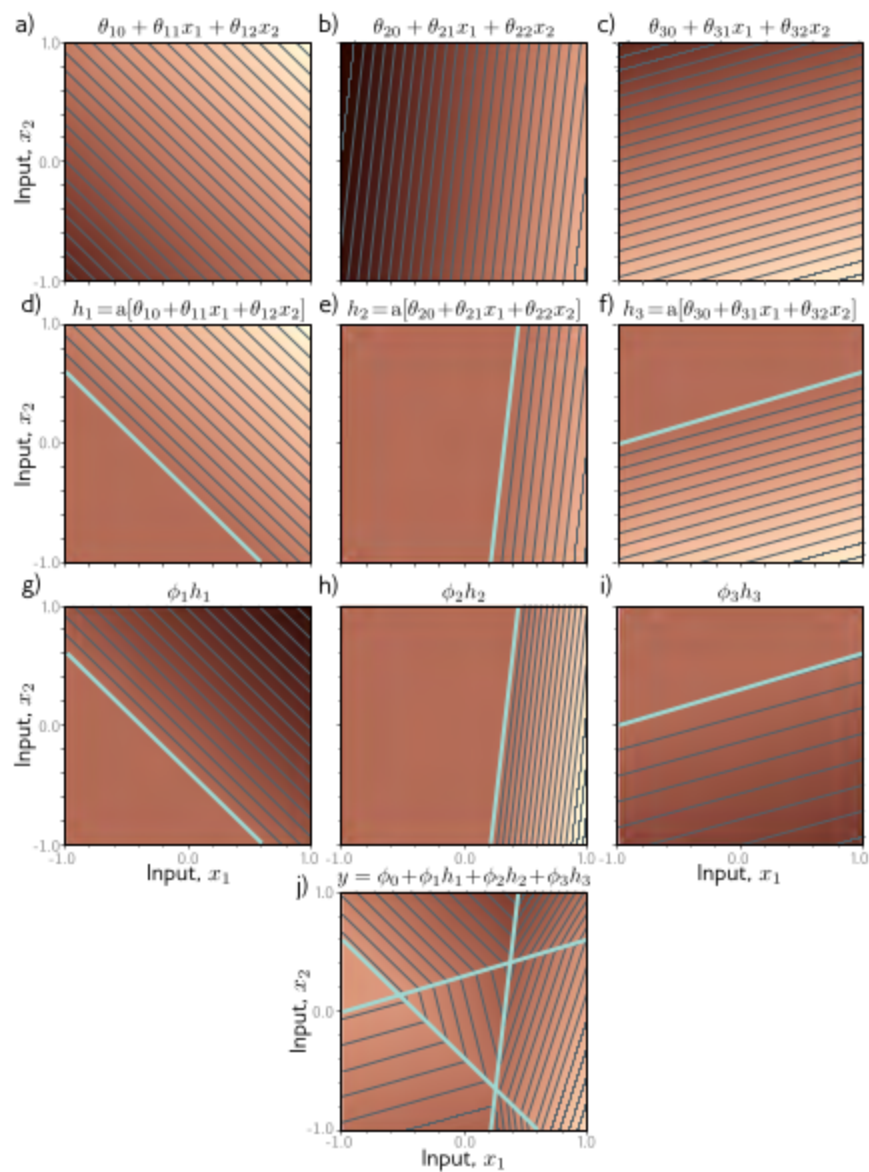
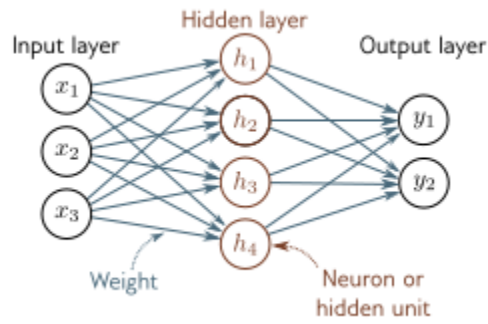
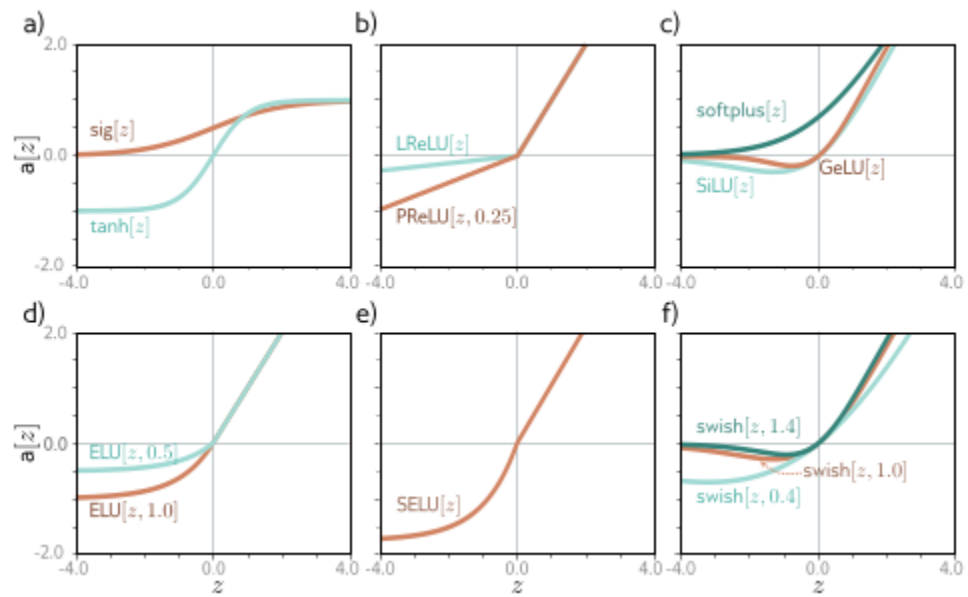


Figure 3.6.3: A neural network with 4 hidden nodes and 2 output nodes.





Activation functions



In []: