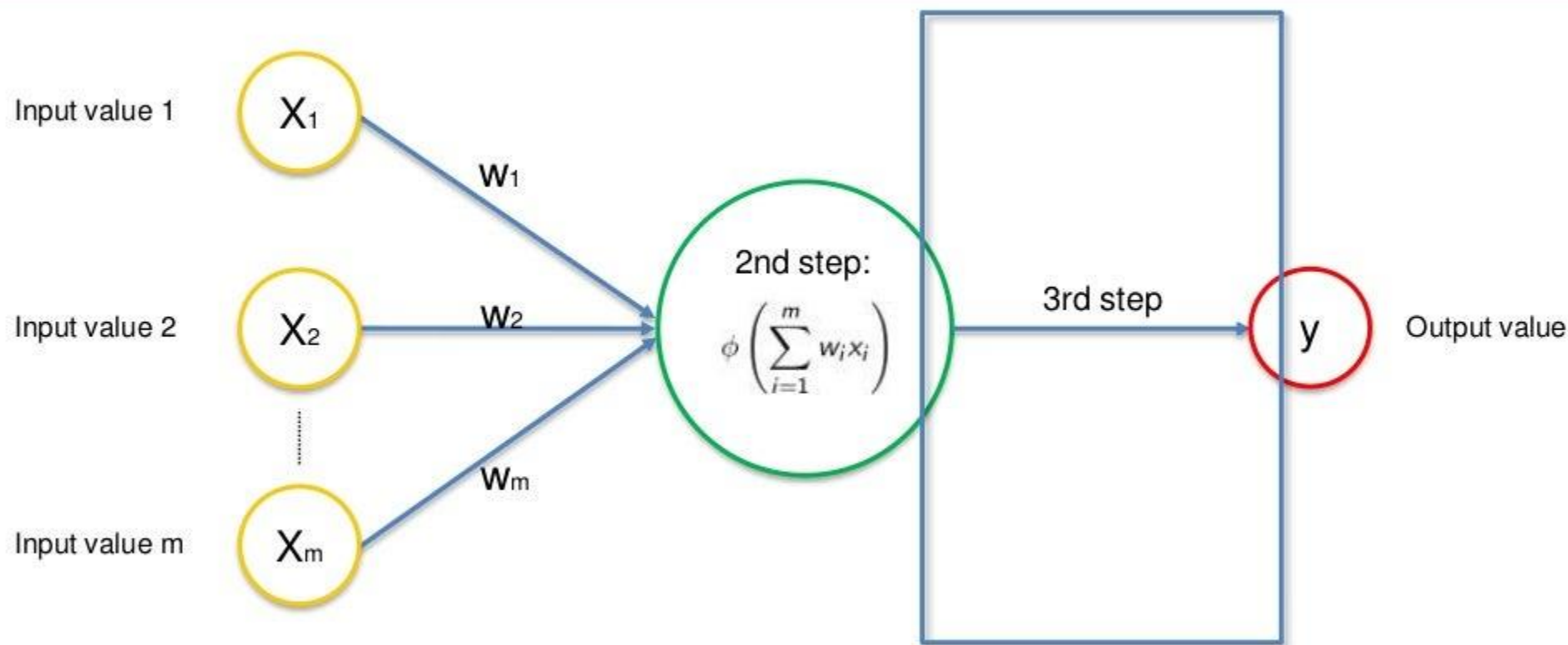
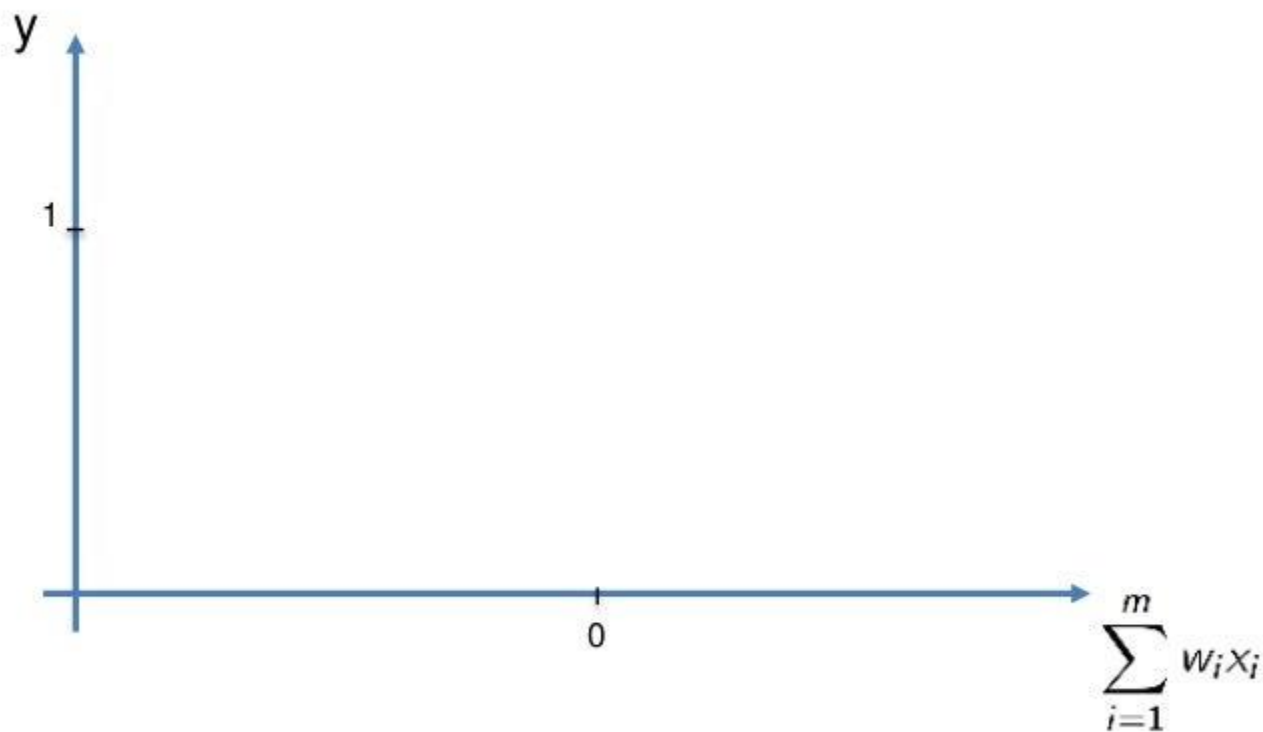


The Activation Function

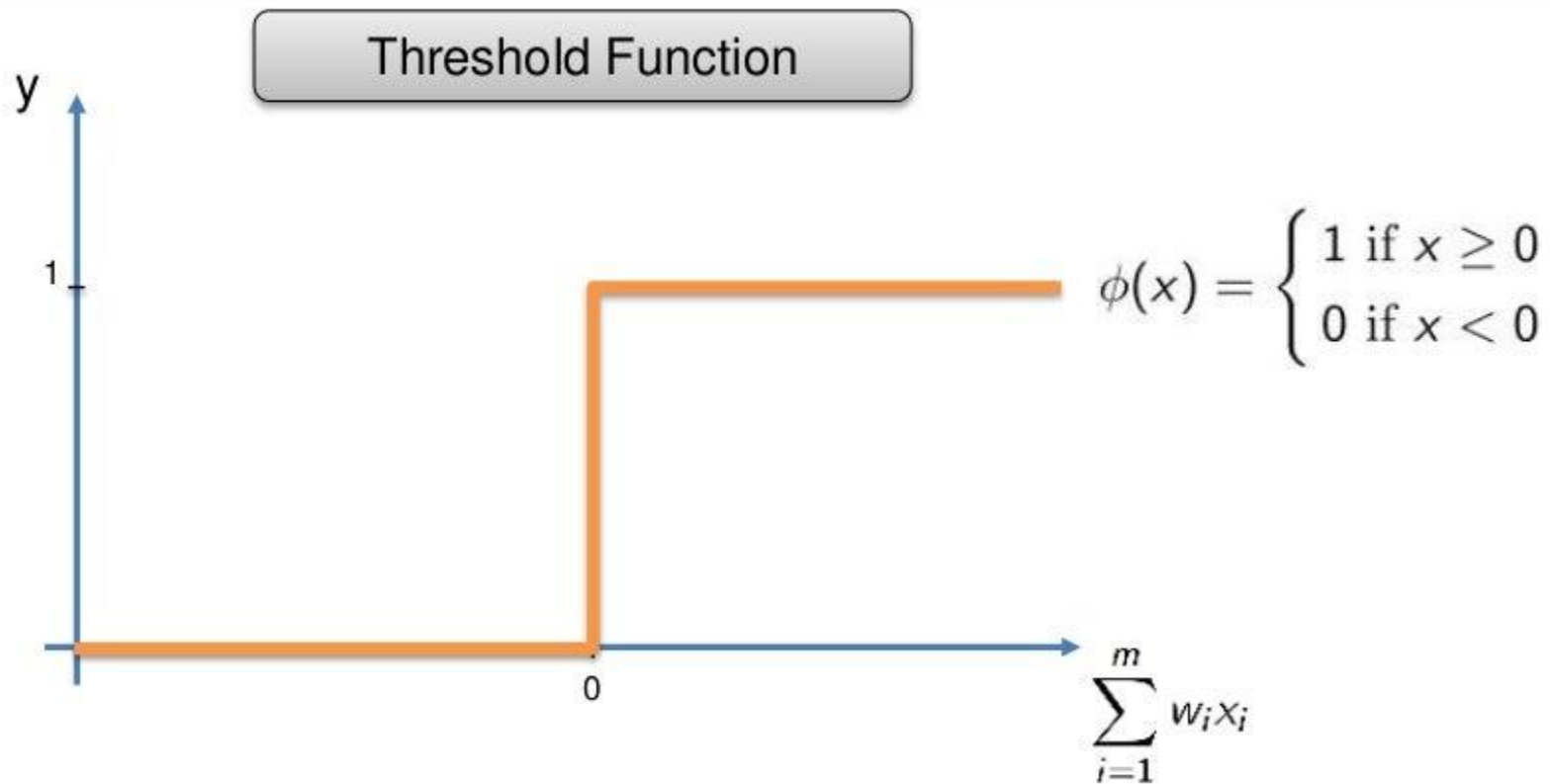
The Activation Function



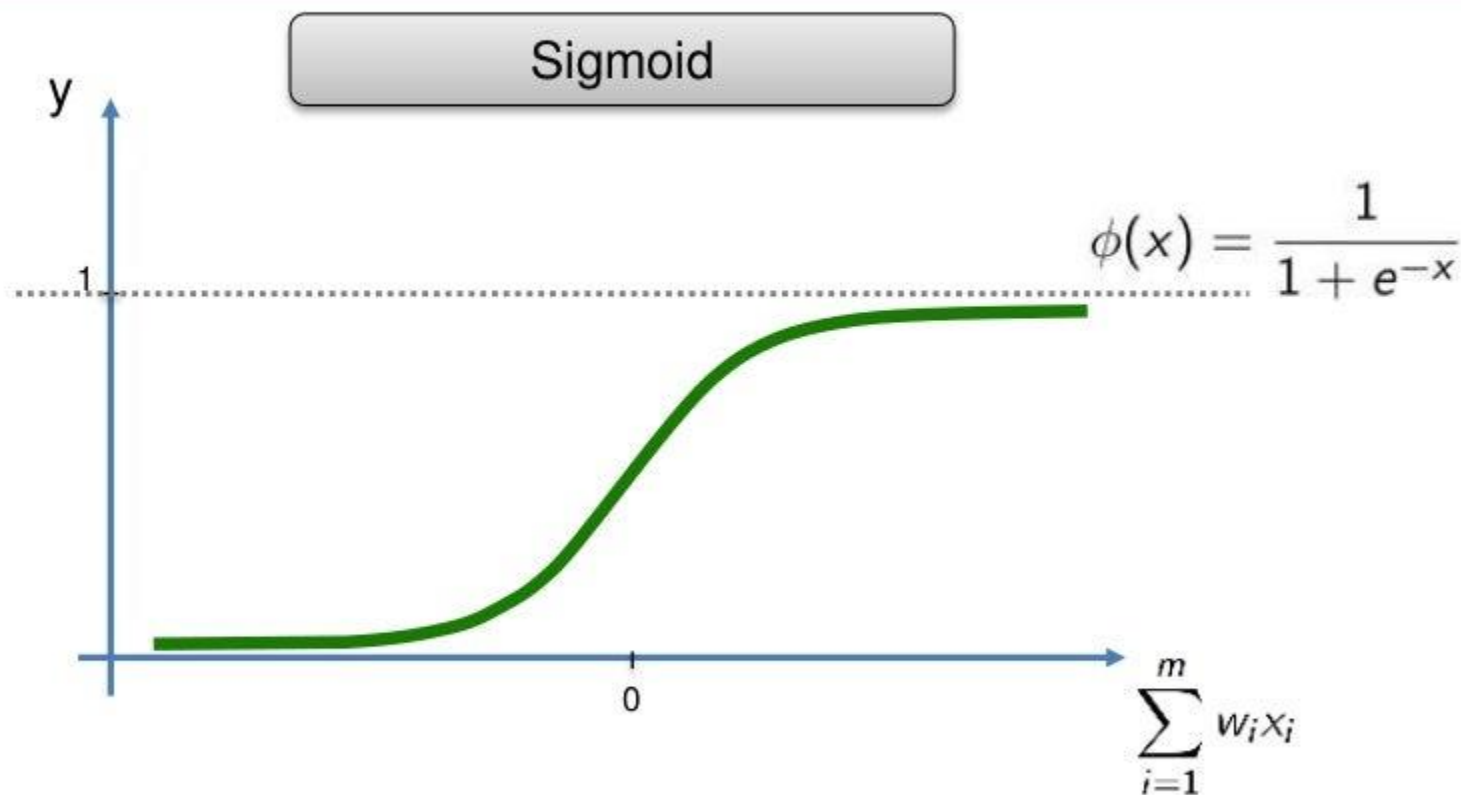
The Activation Function



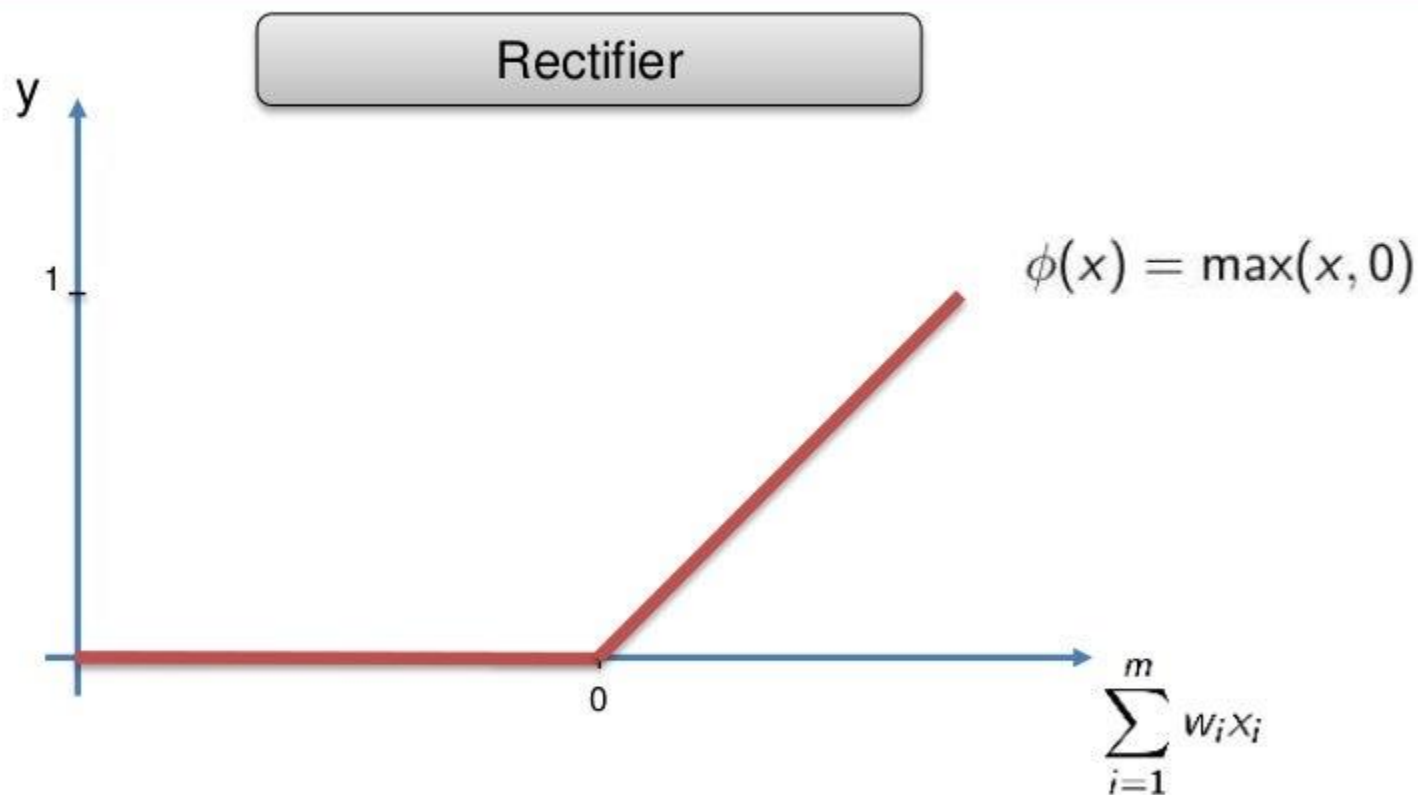
The Activation Function



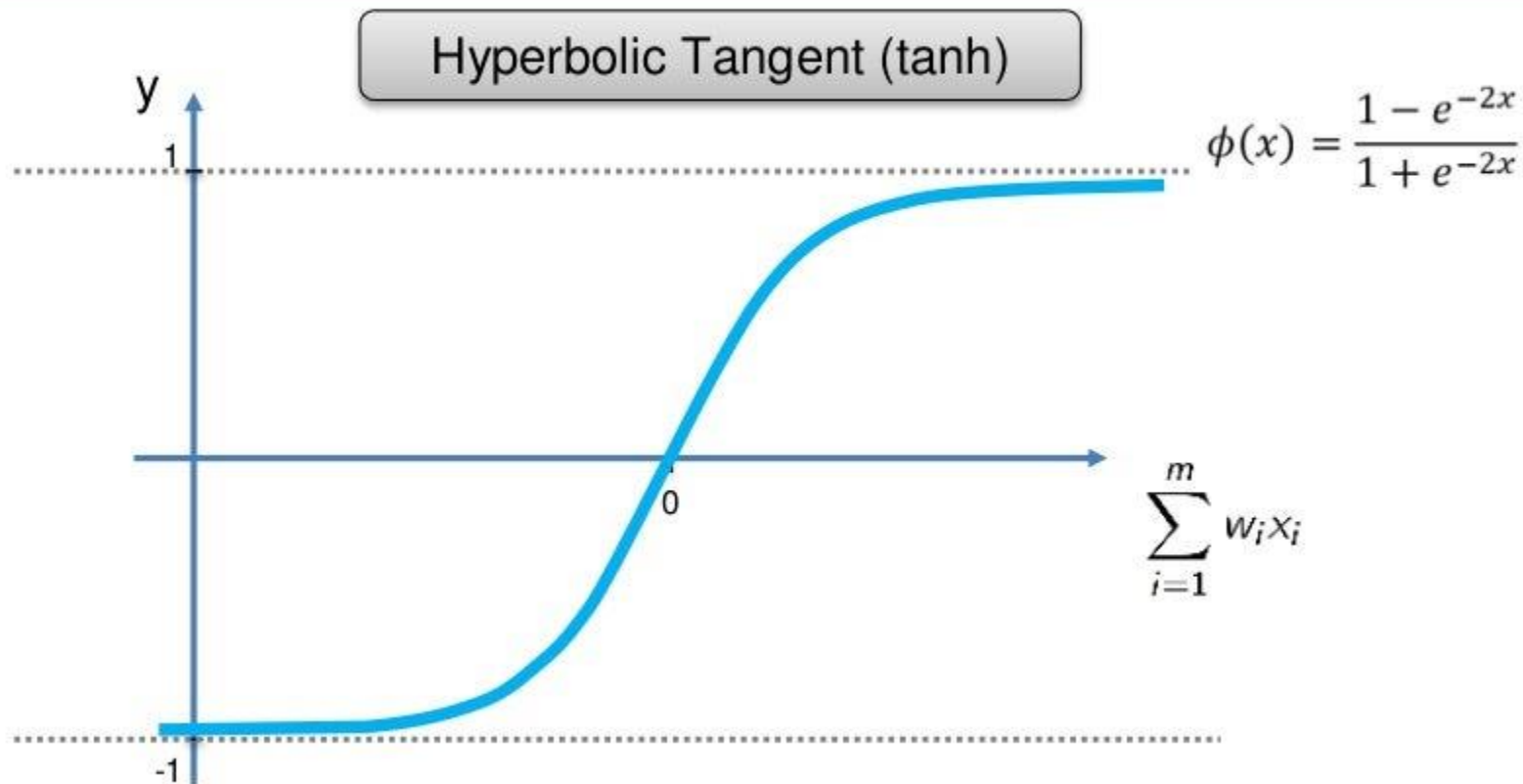
The Activation Function



The Activation Function



The Activation Function

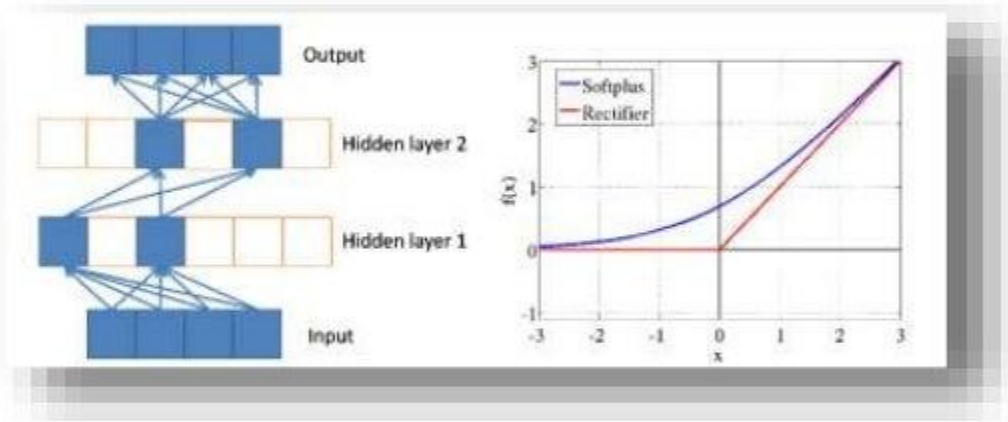


The Activation Function

Additional Reading:

*Deep sparse rectifier
neural networks*

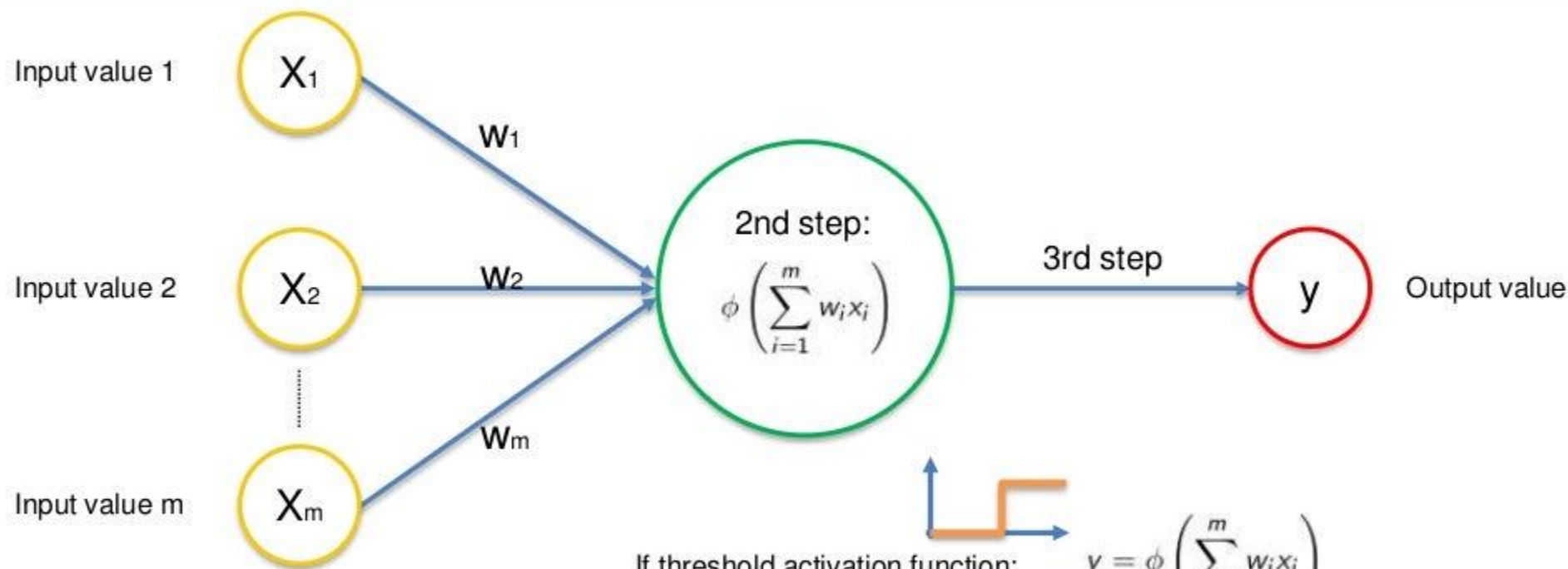
By Xavier Glorot et al. (2011)



Link:

<http://jmlr.org/proceedings/papers/v15/glorot11a/glorot11a.pdf>

The Activation Function



Assuming the DV is binary ($y = 0$ or 1)

If threshold activation function:



$$y = \phi \left(\sum_{i=1}^m w_i x_i \right)$$

If sigmoid activation function:



$$\mathbb{P}(y = 1) = \phi \left(\sum_{i=1}^m w_i x_i \right)$$

The Activation Function

