Rua Cristiano Vianna 104, ap 21
São Paulo

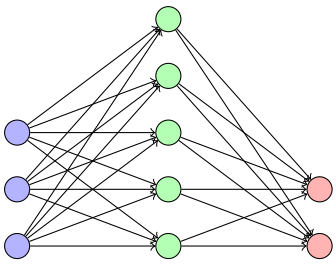
+55 35 99754-9882
caiolagana@gmail.com

April 17, 2024

Neural Networks in a Nutshell

Definition. Let w^l_{jk} be a collection of l matrices whose indexes jk are such that the number of columns of the l^{th} matrix equals the number of lines of the $(l+1)^{\text{th}}$ matrix. So, for example, the following set is a valid collection:

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}, \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$



Definition. Let $M_1, M_2, \cdots, M_N$ be a collection of matrices such that the number of columns of the ith matrix equals the number of rows of the $(i+1)$th matrix. Let W_i and B_i be ... Each matrix M acts on an input vector a as $M(a) = \sigma(w \cdot a + b)$. Then a Neural Network (NN) is the composition of matrices $T = (M_N \circ M_{N-1} \circ \cdots \circ M_1)$.

Definition. For a given input vector x, the prediction vector is defined as $y(x) = T(x)$.

Cordialmente,

Caio Laganá Fernandes