



Human-Centered Data & AI



Vinicius Caridá

Data Science Manager, Itaú Unibanco
MBA Professor, FIAP
GDE – Machine Learning



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



Amir Jalilifard

Data Scientist Senior, PicPay
Ph.D. Candidate - UFMG



@amir jalilifard



@Machine Learning TV



@Jalilifard



@amir jalilifard

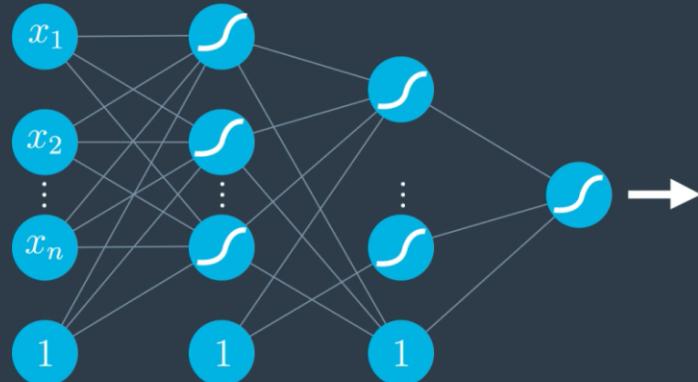
“

Redes Neurais

Recapitulando...

Redes Neurais

Multi-layer Perceptron



PREDICTION

$$\hat{y} = \sigma W^{(3)} \circ \sigma W^{(2)} \circ \sigma \circ W^{(1)}(x)$$

ERROR FUNCTION

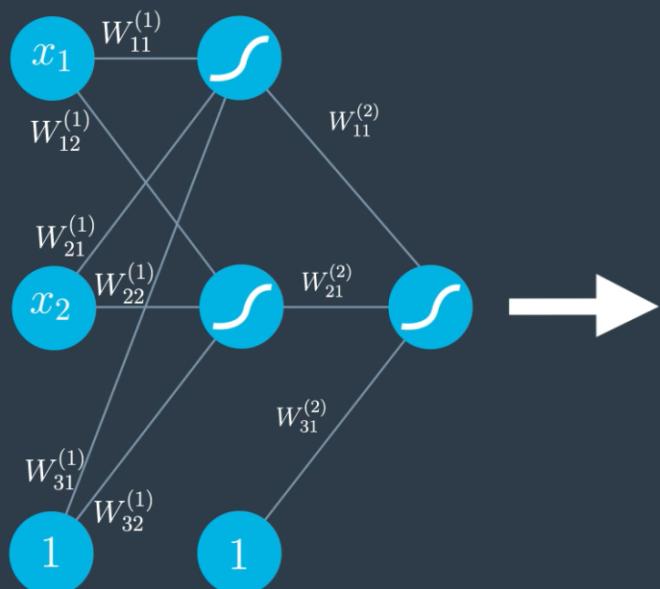
$$E(W) = -\frac{1}{m} \sum_{i=1}^m y_i \ln(\hat{y}_i) + (1 - y_i) \ln(1 - \hat{y}_i)$$

GRADIENT OF THE ERROR FUNCTION

$$\nabla E = \left(\dots, \frac{\partial E}{\partial w_j^{(i)}}, \dots \right)$$

Redes Neurais

Backpropagation



$$\hat{y} = \sigma W^{(2)} \circ \sigma \circ W^{(1)}(x)$$

$$W^{(1)} = \begin{pmatrix} W_{11}^{(1)} & W_{12}^{(1)} \\ W_{21}^{(1)} & W_{22}^{(1)} \\ W_{31}^{(1)} & W_{32}^{(1)} \end{pmatrix} \quad W^{(2)} = \begin{pmatrix} W_{11}^{(2)} \\ W_{21}^{(2)} \\ W_{31}^{(2)} \end{pmatrix}$$

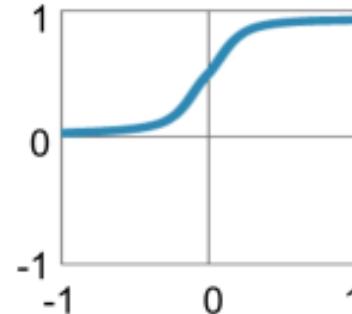
$$\nabla E = \begin{pmatrix} \frac{\partial E}{\partial W_{11}^{(1)}} & \frac{\partial E}{\partial W_{12}^{(1)}} & \frac{\partial E}{\partial W_{11}^{(2)}} \\ \frac{\partial E}{\partial W_{21}^{(1)}} & \frac{\partial E}{\partial W_{22}^{(1)}} & \frac{\partial E}{\partial W_{21}^{(2)}} \\ \frac{\partial E}{\partial W_{31}^{(1)}} & \frac{\partial E}{\partial W_{32}^{(1)}} & \frac{\partial E}{\partial W_{31}^{(2)}} \end{pmatrix}$$

$$W'_{ij}^{(k)} \leftarrow W_{ij}^{(k)} - \alpha \frac{\partial E}{\partial W_{ij}^{(k)}}$$

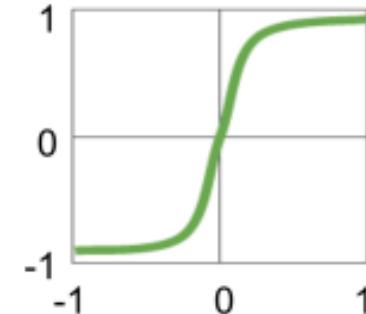
Redes Neurais

Traditional Non-Linear Activation Functions

Sigmoid

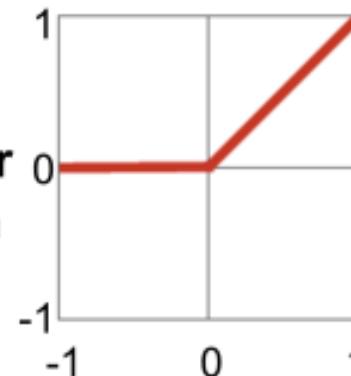


Hyperbolic Tangent

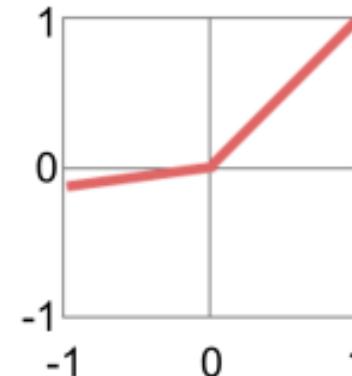


Modern Non-Linear Activation Functions

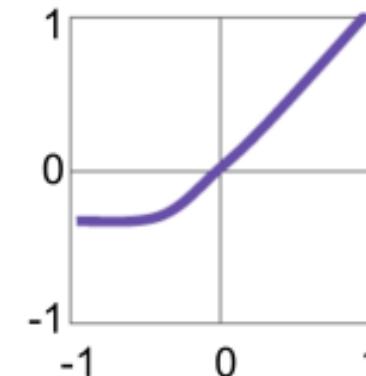
Rectified Linear Unit (ReLU)



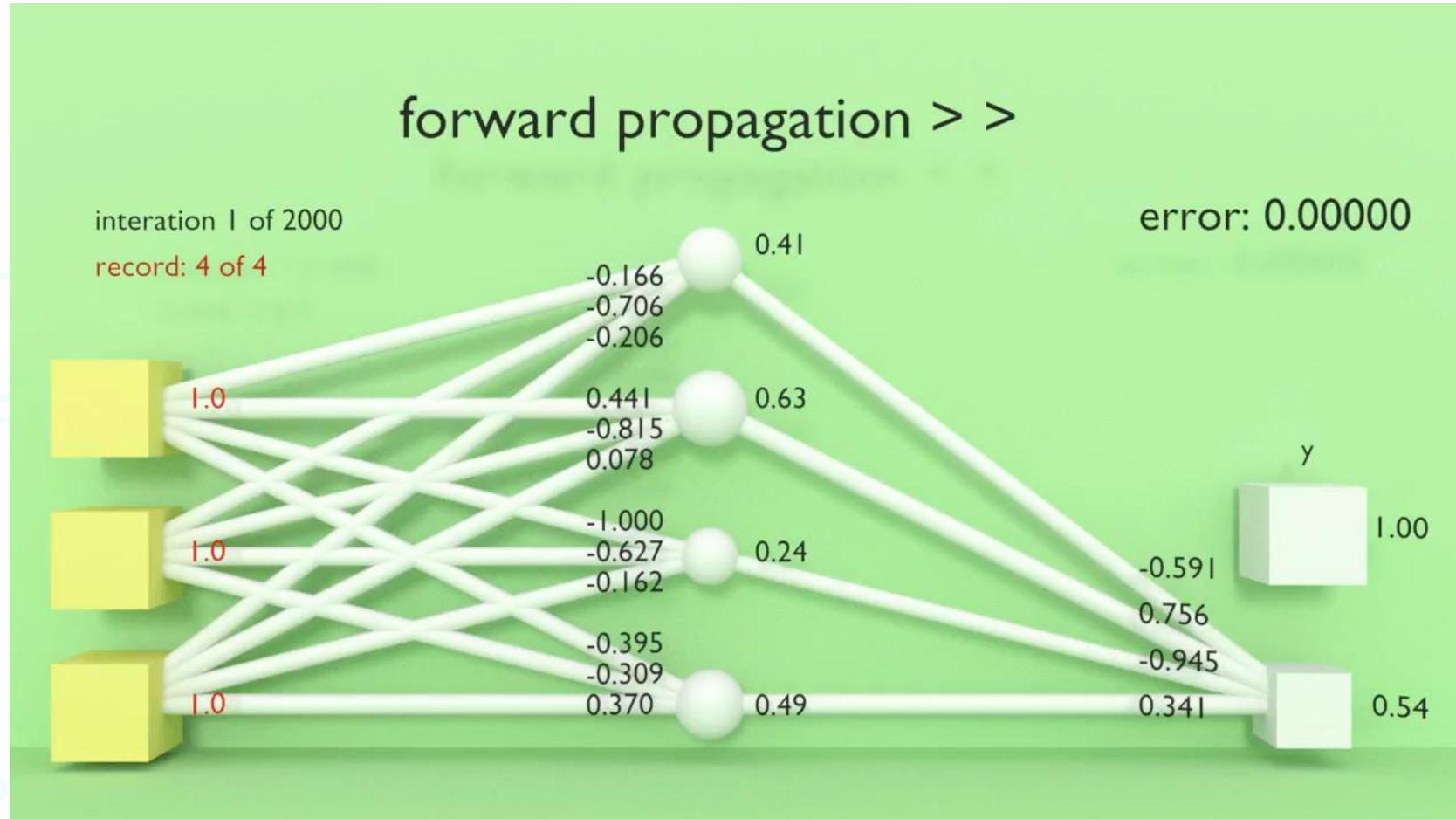
Leaky ReLU



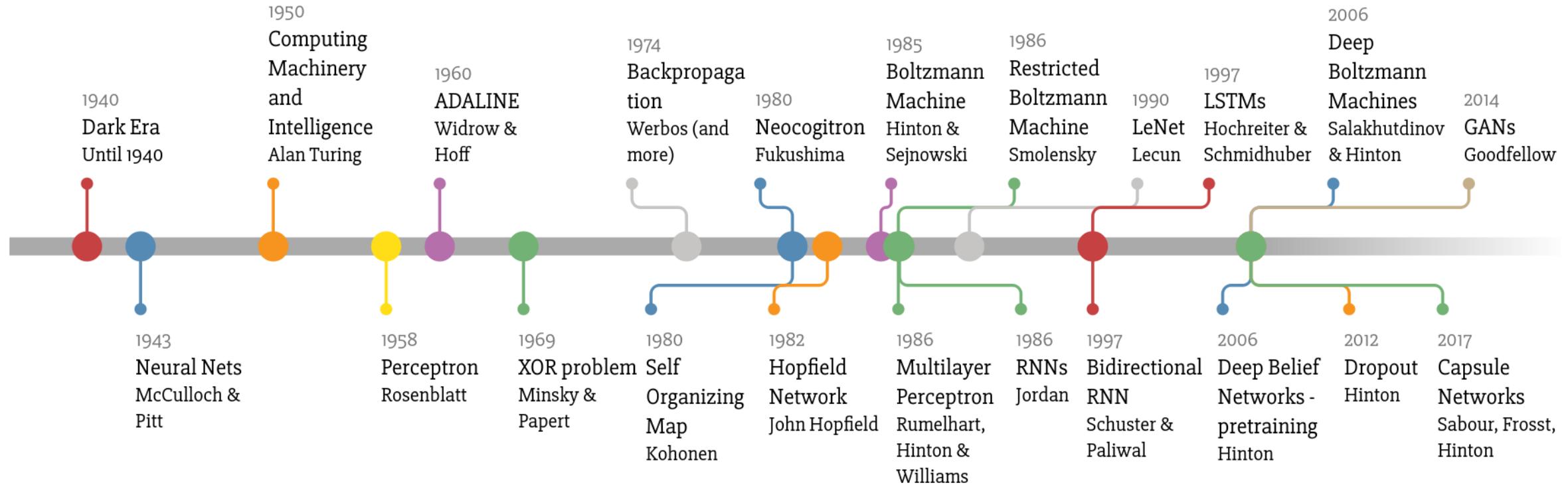
Exponential LU



Redes Neurais



Artificial Neural Networks



“

Redes Neurais Recorrentes (RNN)

Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Apple pie



Burger



Chicken

Redes Neurais Recorrentes (RNN)



Weather



Redes Neurais Recorrentes (RNN)

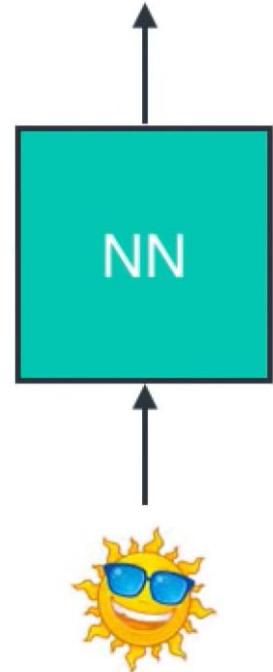
Weather



Redes Neurais Recorrentes (RNN)



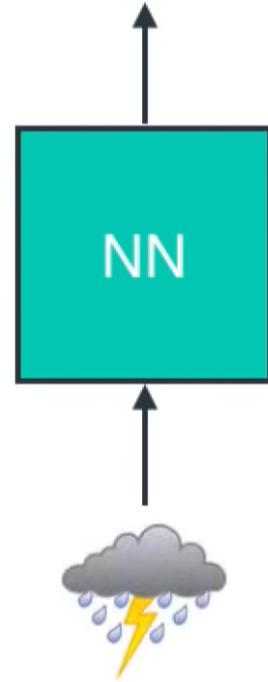
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)

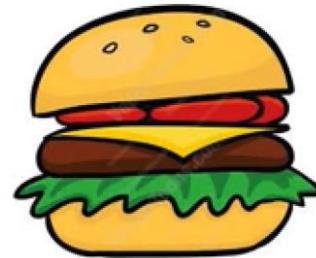


Redes Neurais Recorrentes (RNN)

Vectors



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Redes Neurais Recorrentes (RNN)

Vectors



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

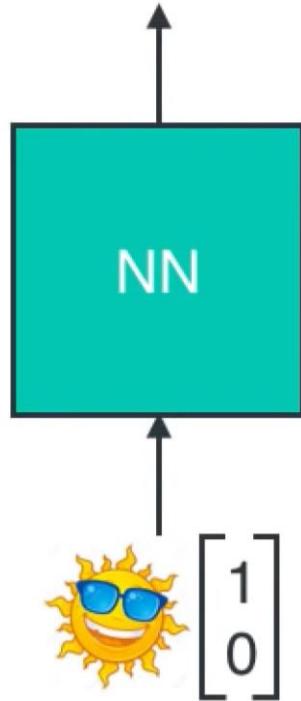


$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

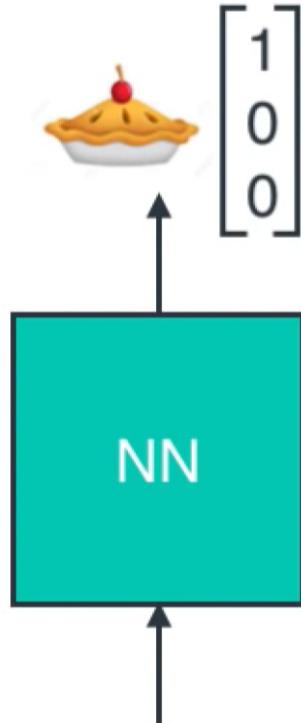


$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

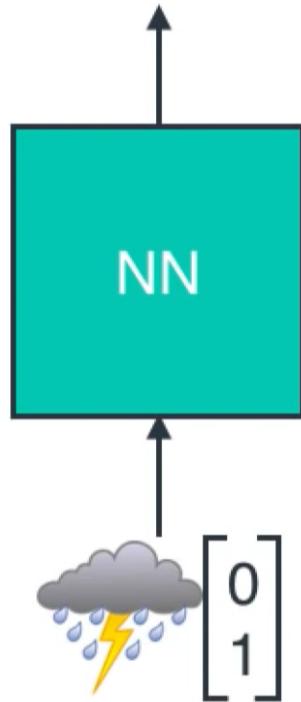
Redes Neurais Recorrentes (RNN)



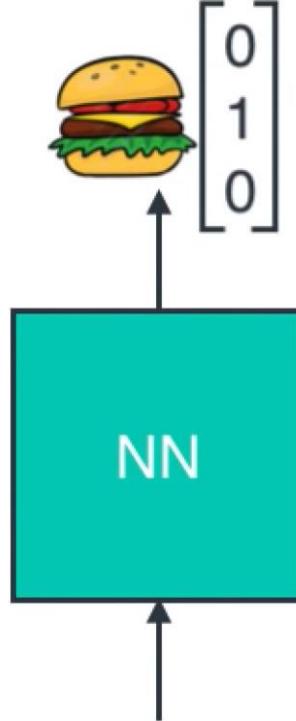
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

=

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ ☀️}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ 🌧️}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \text{ ☀️}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{☀️}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{⚡}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{🥧}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{sun with sunglasses}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{cloud with lightning}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \quad \text{cloud with rain}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{apple pie}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{☀️}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{⚡}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{⚡} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{🍔}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{🥧}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{☀️}$$

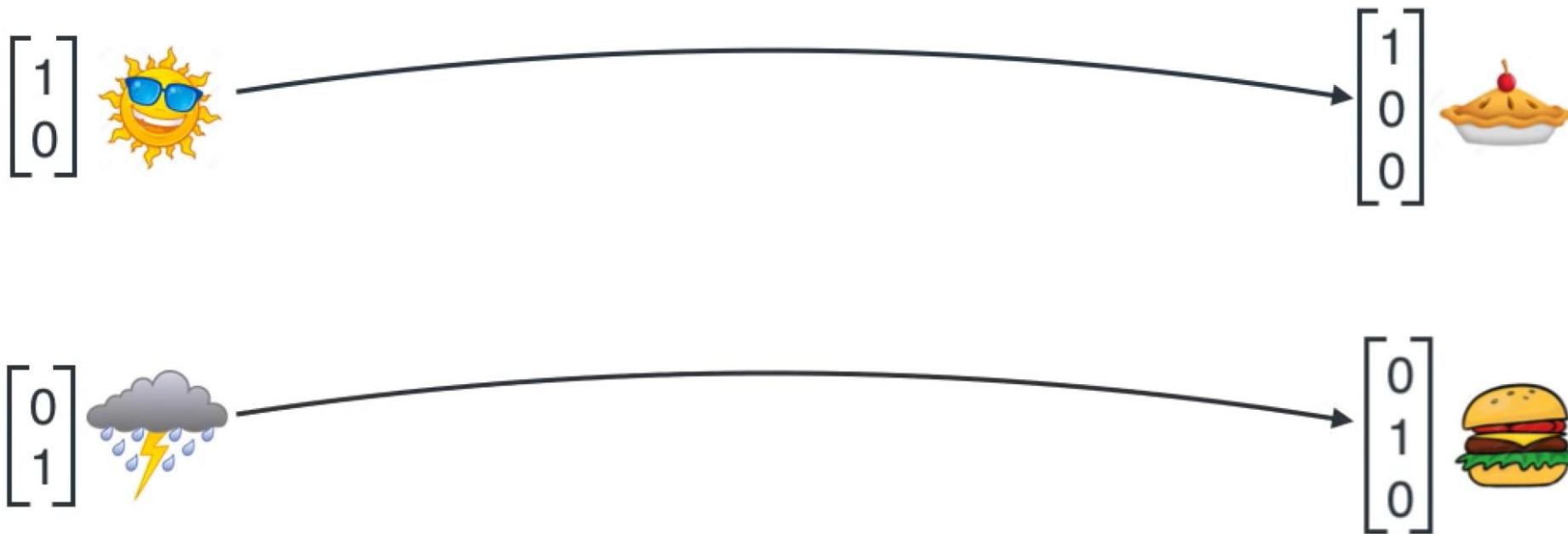
$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{⚡️}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{⚡️} =$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{🥧}$$

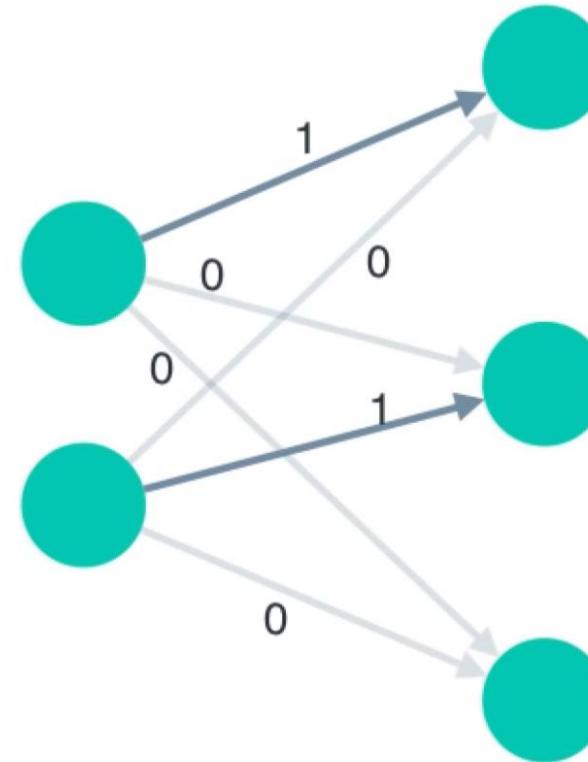
$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{🍔}$$

Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

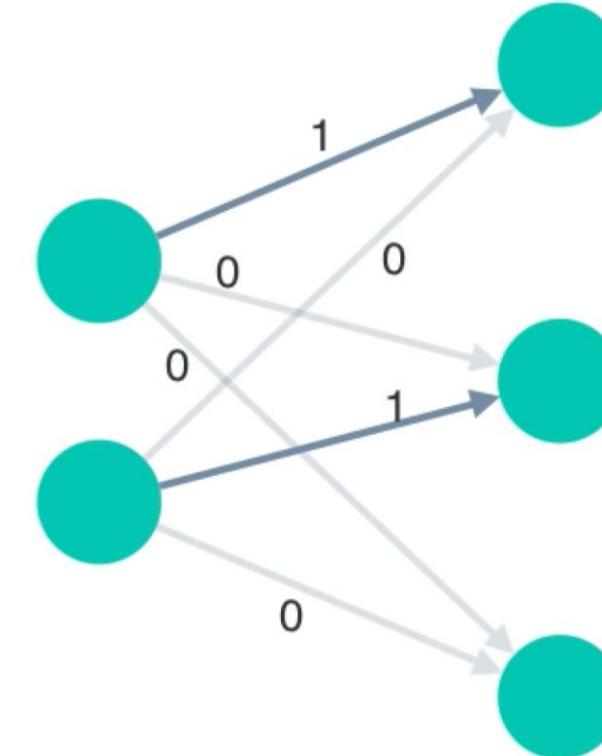


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

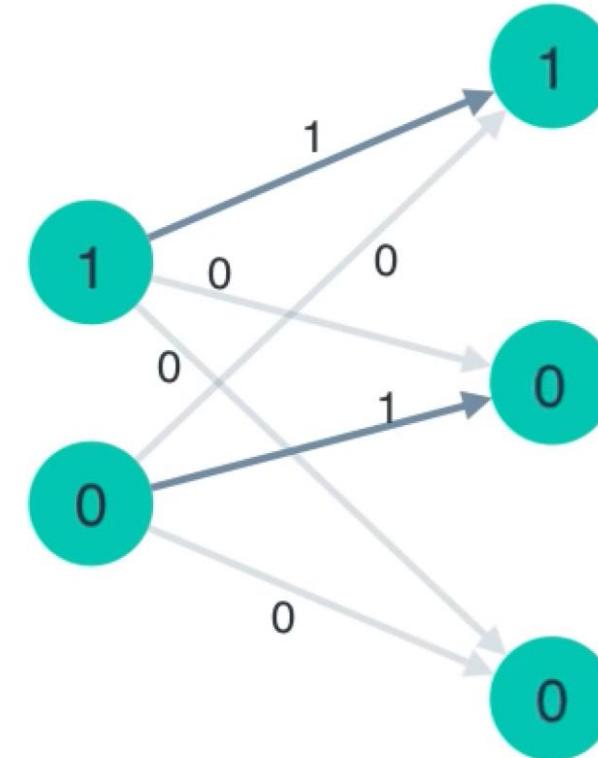


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

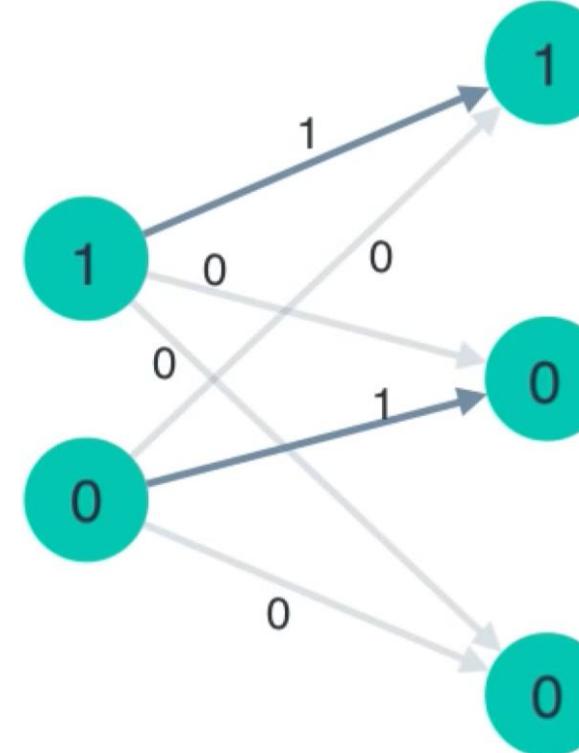


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



Redes Neurais Recorrentes (RNN)



Apple pie

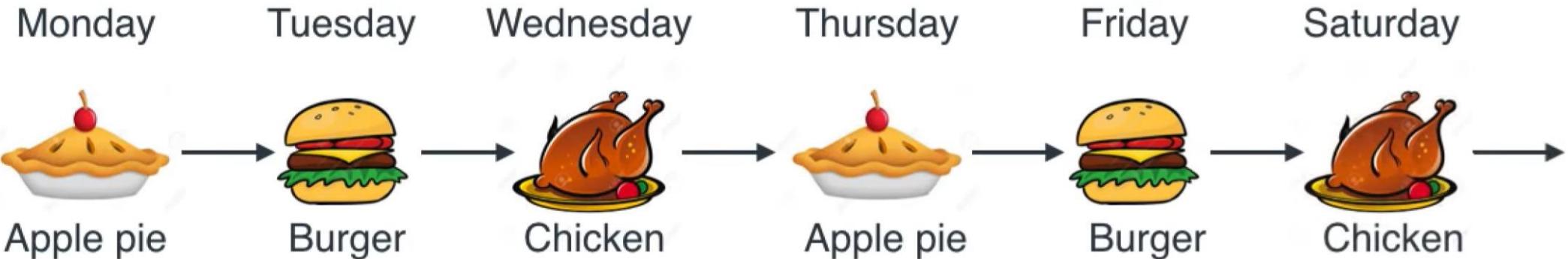


Burger

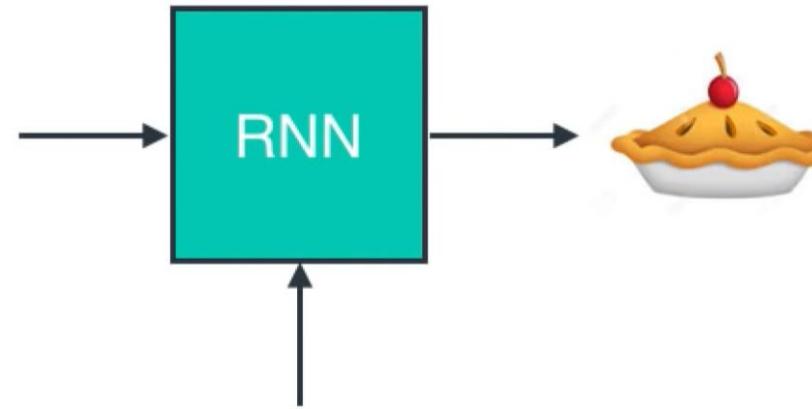


Chicken

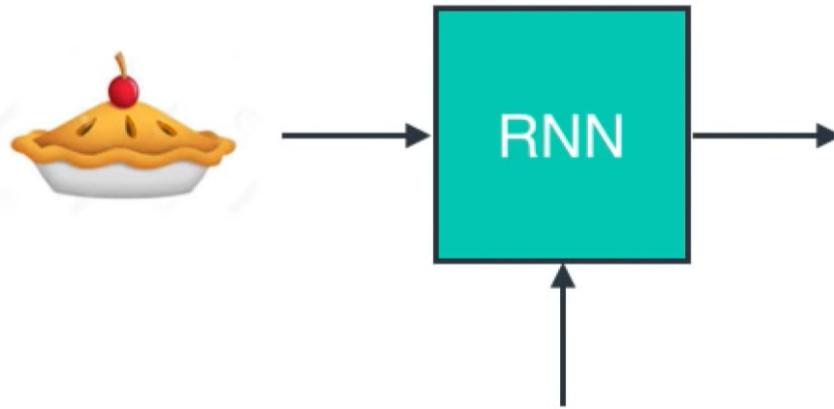
Redes Neurais Recorrentes (RNN)



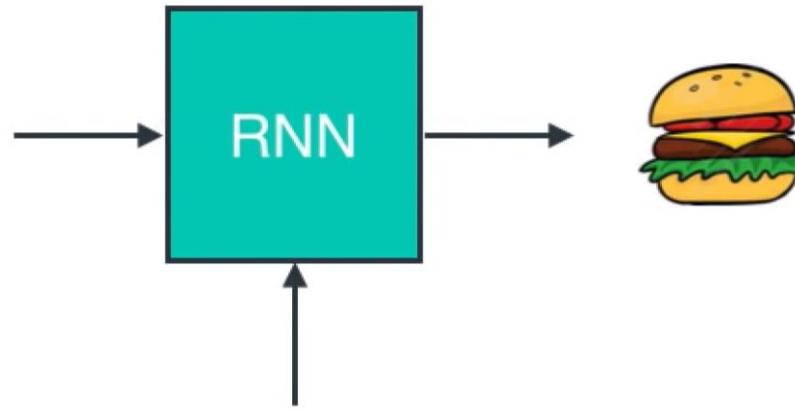
Redes Neurais Recorrentes (RNN)



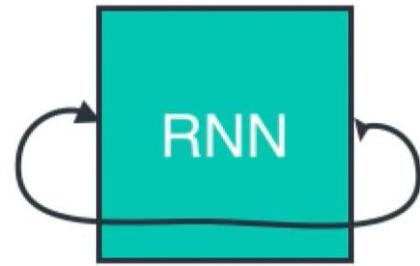
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

=

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{Pie emoji}$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{Hamburger emoji}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{Roasted Turkey emoji}$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{Hamburger emoji}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$


$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

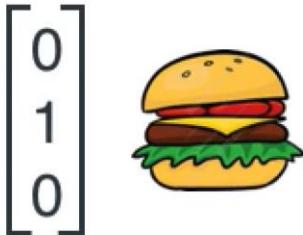
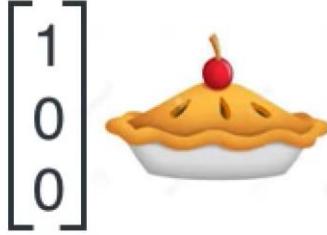

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$


$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

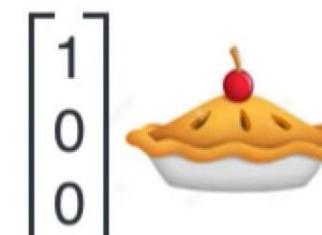
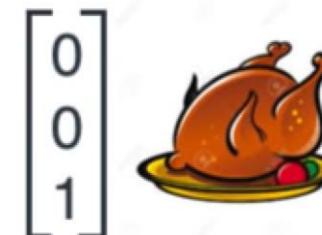
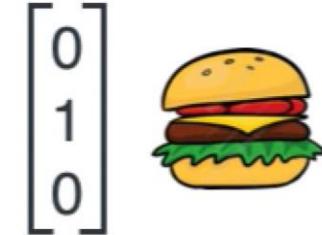

Redes Neurais Recorrentes (RNN)



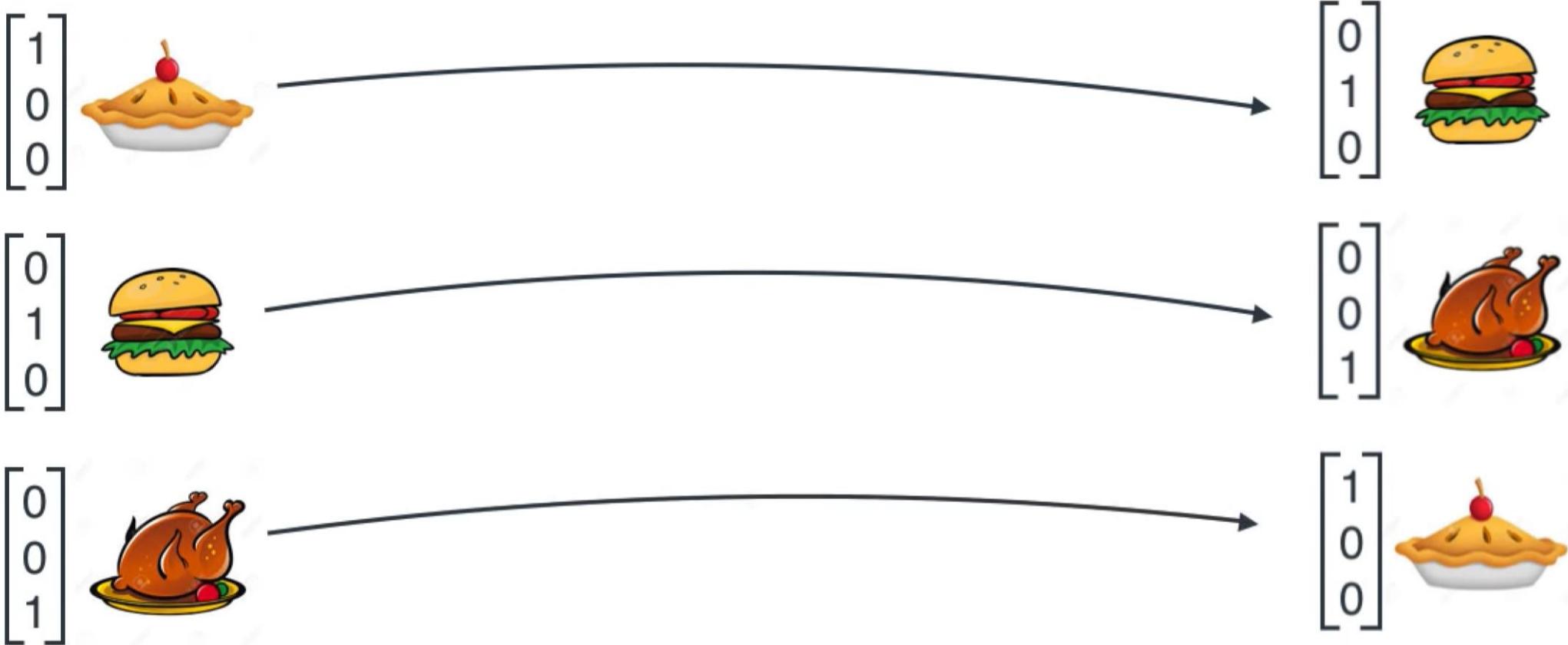
$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



=



Redes Neurais Recorrentes (RNN)

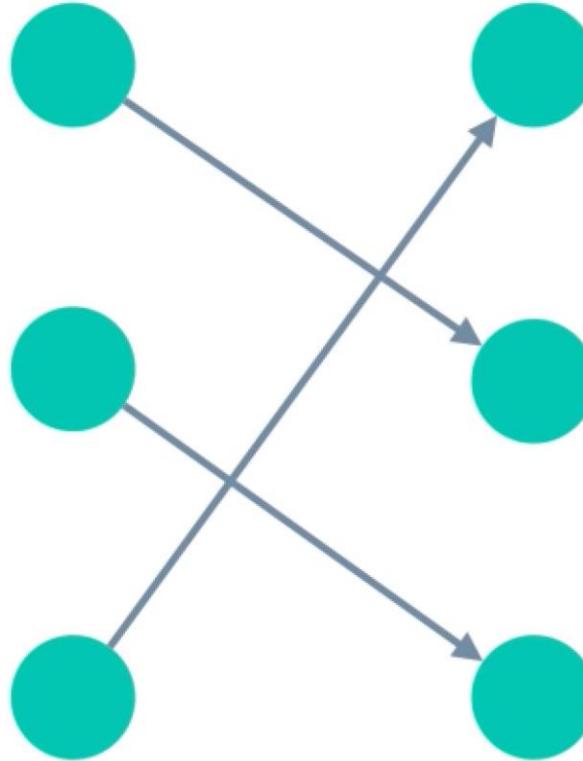


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

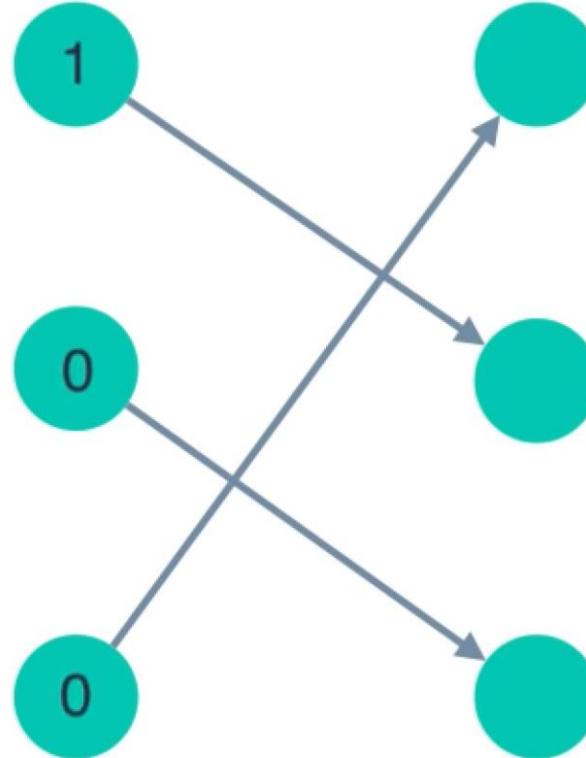


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

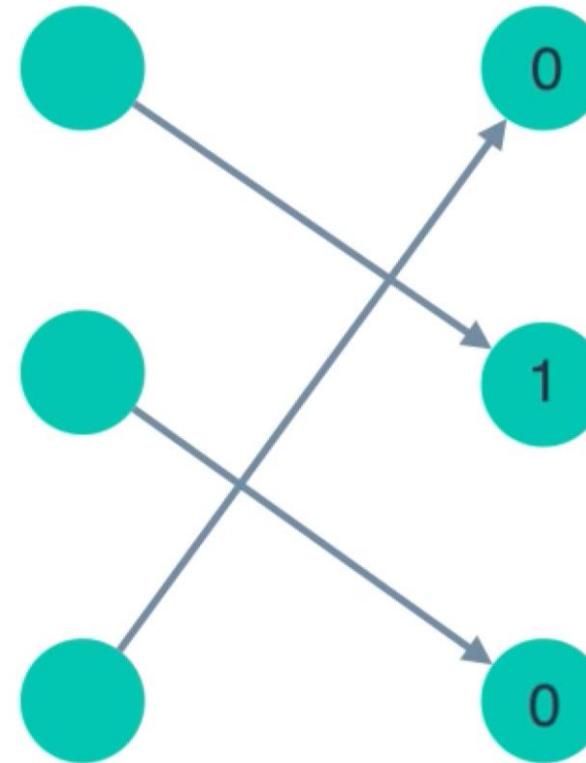


Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



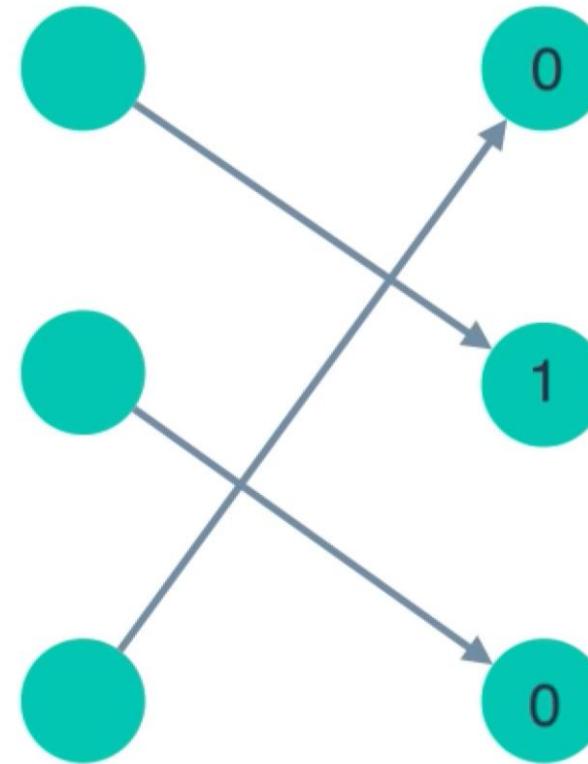
$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



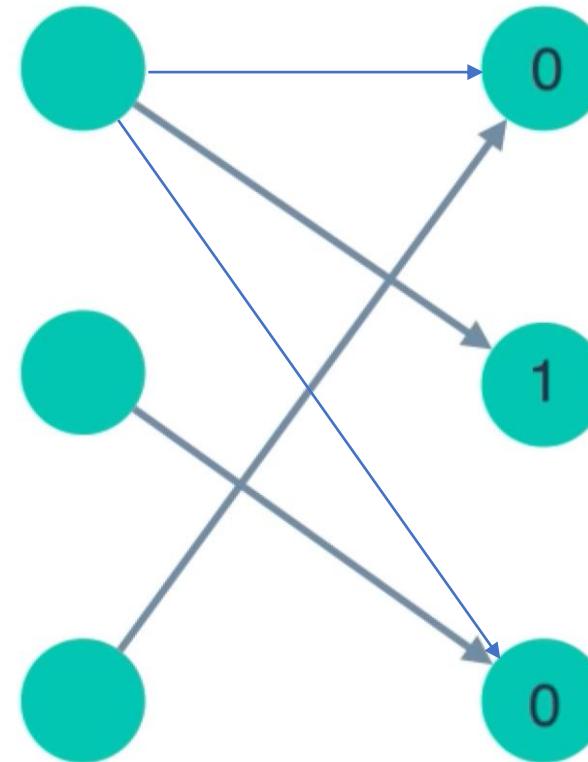
$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



Redes Neurais Recorrentes (RNN)



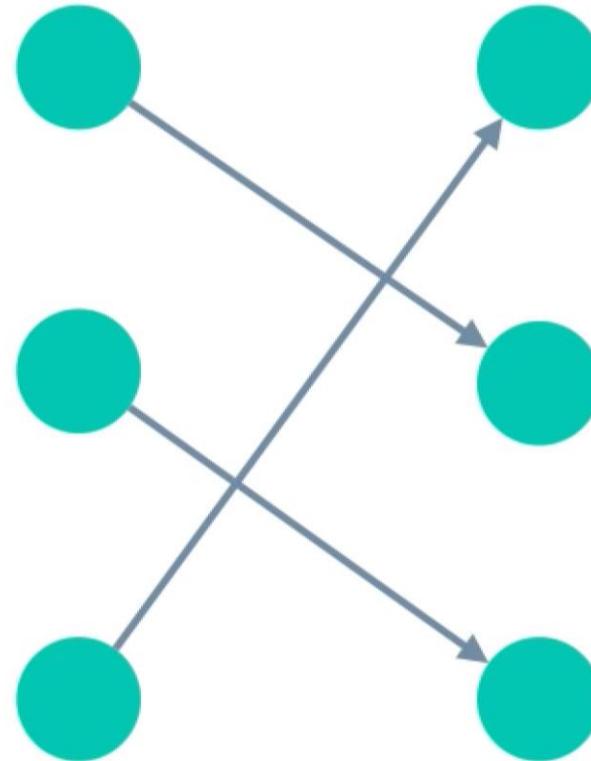
$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



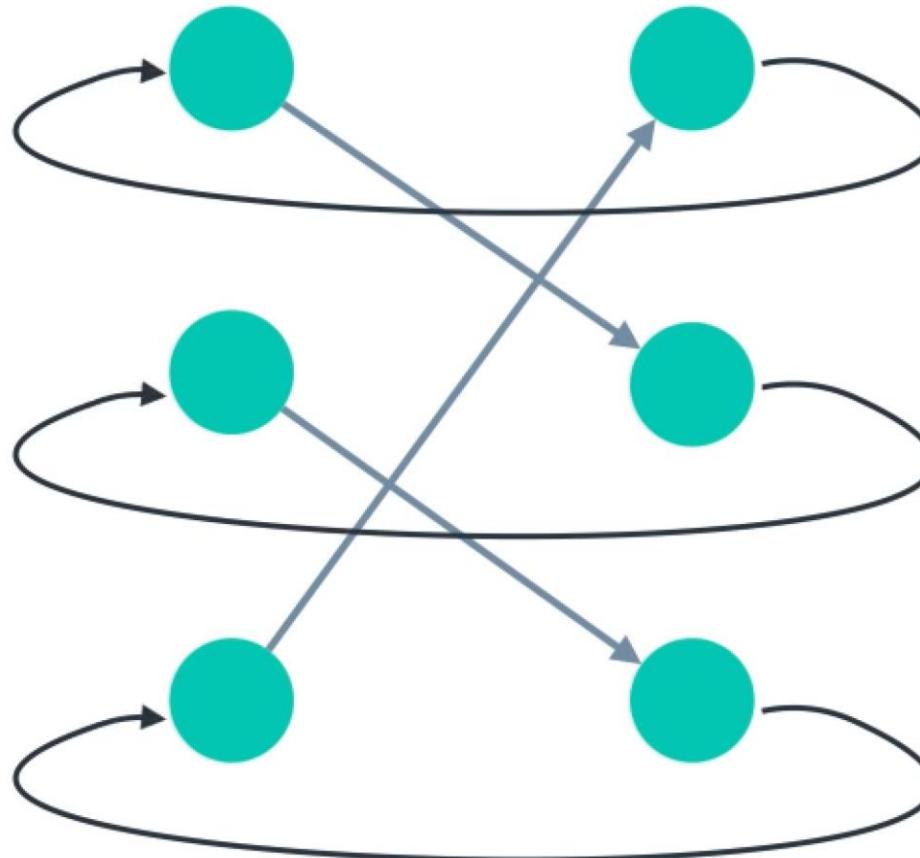
$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)

Weather



Sunny

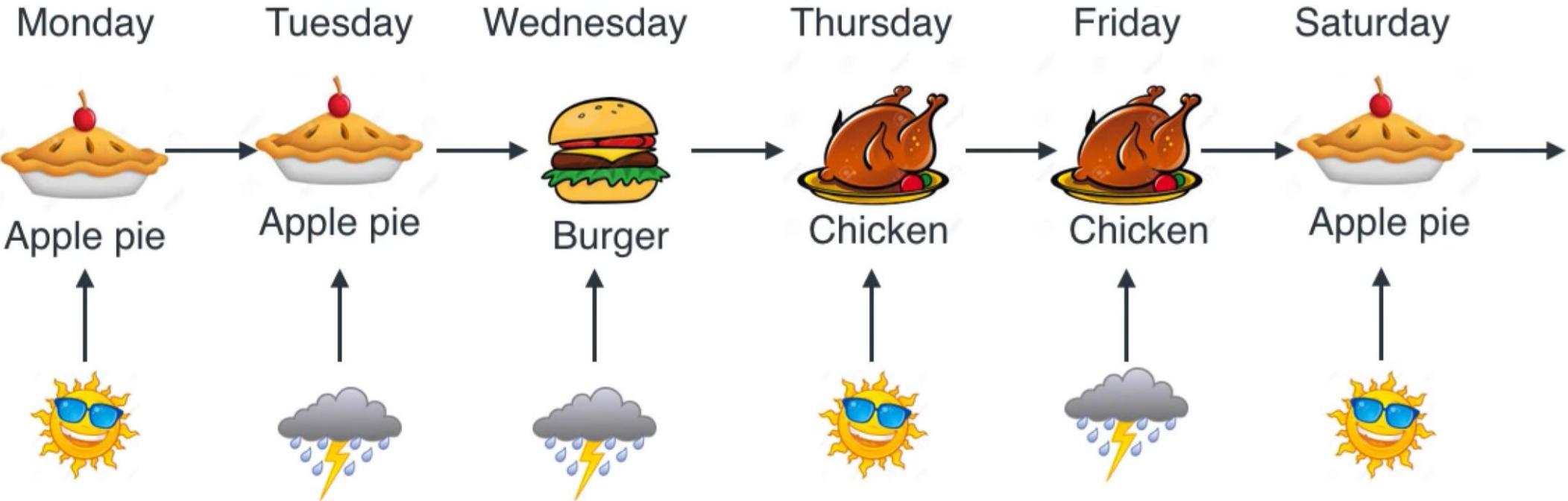
Same as yesterday



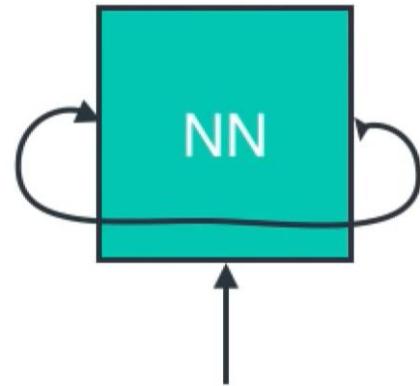
Rain

Next dish

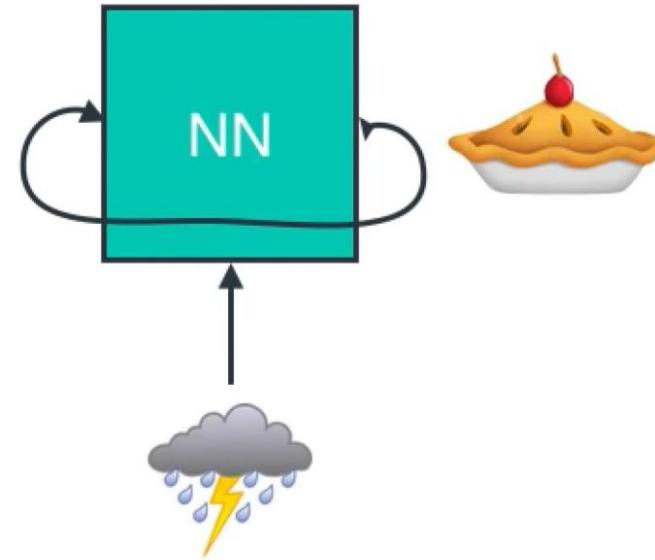
Redes Neurais Recorrentes (RNN)



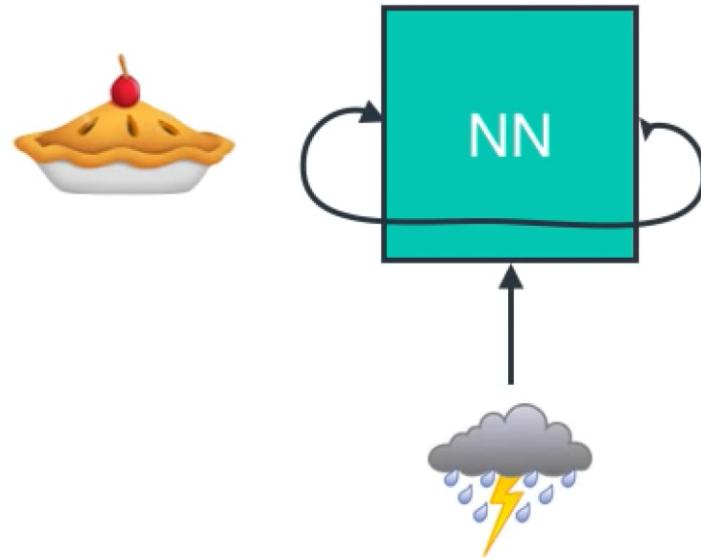
Redes Neurais Recorrentes (RNN)



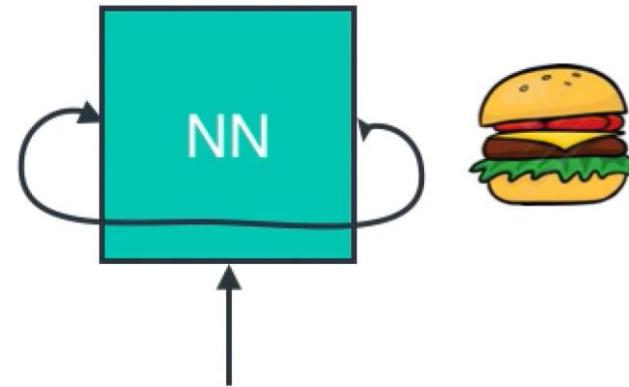
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



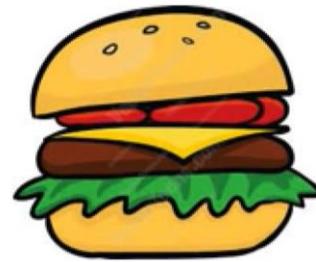
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

+

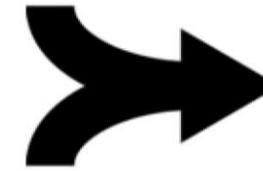
$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Food

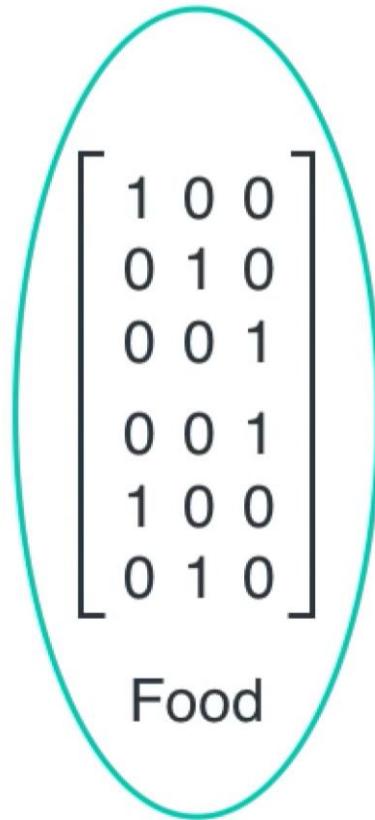
Add

Weather

Merge



Redes Neurais Recorrentes (RNN)

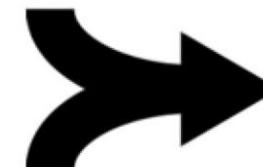


+

Add

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Weather



Merge

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Food

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

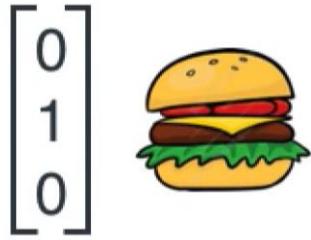
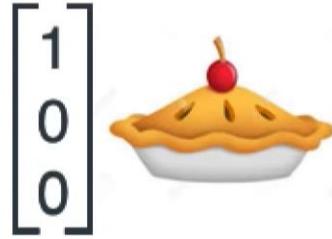

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$


$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

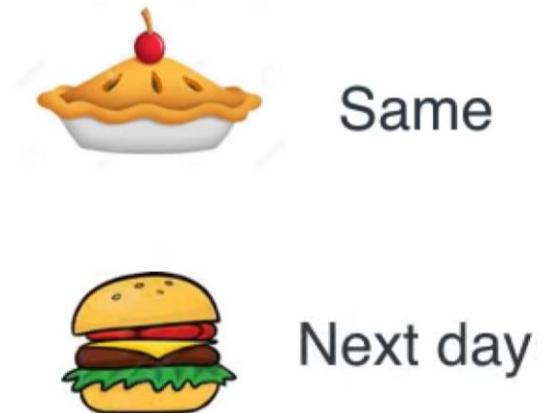

Food

Redes Neurais Recorrentes (RNN)

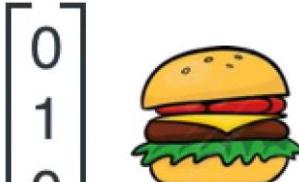


$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$


Food



Redes Neurais Recorrentes (RNN)


$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} =$$

Food



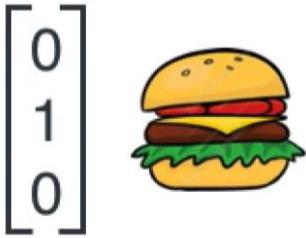
$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$



Same

Next day

Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Food



$$= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$



Same



Next day

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

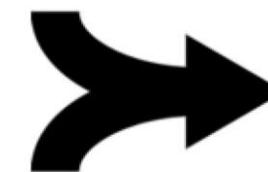
Food

+

Add

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Weather



Merge

Redes Neurais Recorrentes (RNN)

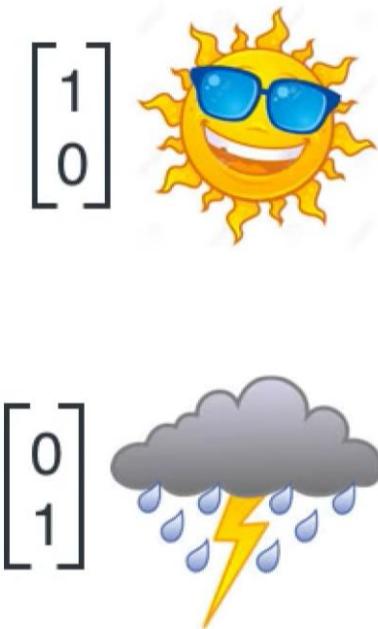
$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \hline 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Weather

Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \hline 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \hline 0 \\ 0 \\ 0 \end{bmatrix}$$

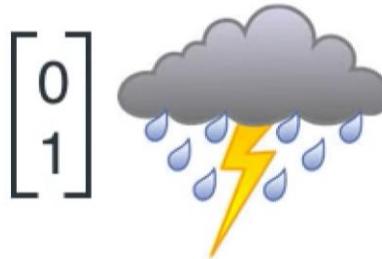
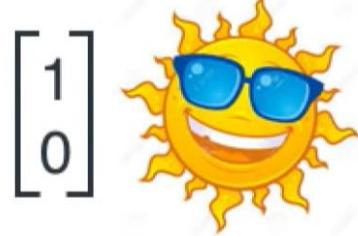
=




Same

Next day

Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \hline 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \text{Weather}$$


$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 1 \\ 1 \end{bmatrix}$$



Same

Next day

Redes Neurais Recorrentes (RNN)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

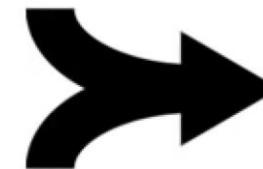
Food

+

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \hline 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Weather

Add



Merge

Redes Neurais Recorrentes (RNN)



Add

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\\\ \hline 0 \\ 1 \\ 0 \end{bmatrix}$$



Same



Next day

Redes Neurais Recorrentes (RNN)



Add

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \end{bmatrix}$$



Same



Next day

+

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 1 \\ 1 \end{bmatrix}$$



Same

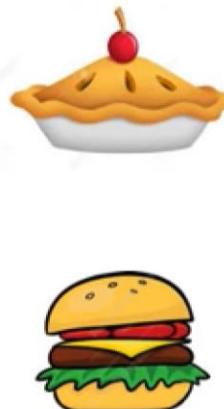
Next day

Redes Neurais Recorrentes (RNN)

Add



$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \end{bmatrix}$$



Same



Next day

+

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 1 \\ 1 \end{bmatrix}$$



Same

Next day

=

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \hline 1 \\ 2 \\ 1 \end{bmatrix}$$

Redes Neurais Recorrentes (RNN)

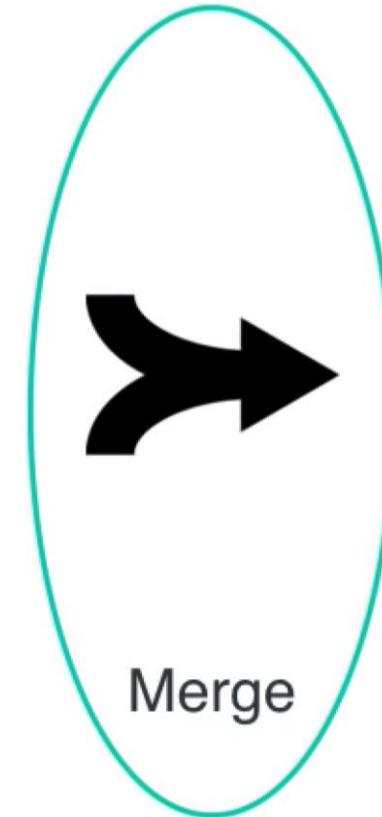
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Food

+

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \hline 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Weather



Redes Neurais Recorrentes (RNN)



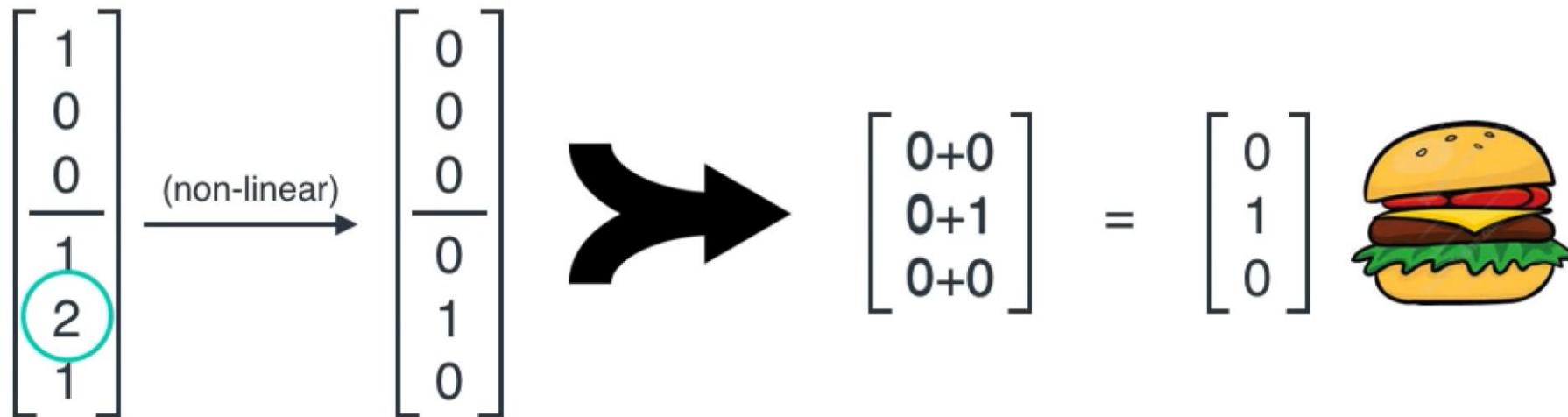
Merge

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \hline 1 \\ 2 \\ 1 \end{bmatrix} \xrightarrow{\text{(non-linear)}} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \end{bmatrix}$$

Redes Neurais Recorrentes (RNN)



Merge



Redes Neurais Recorrentes (RNN)

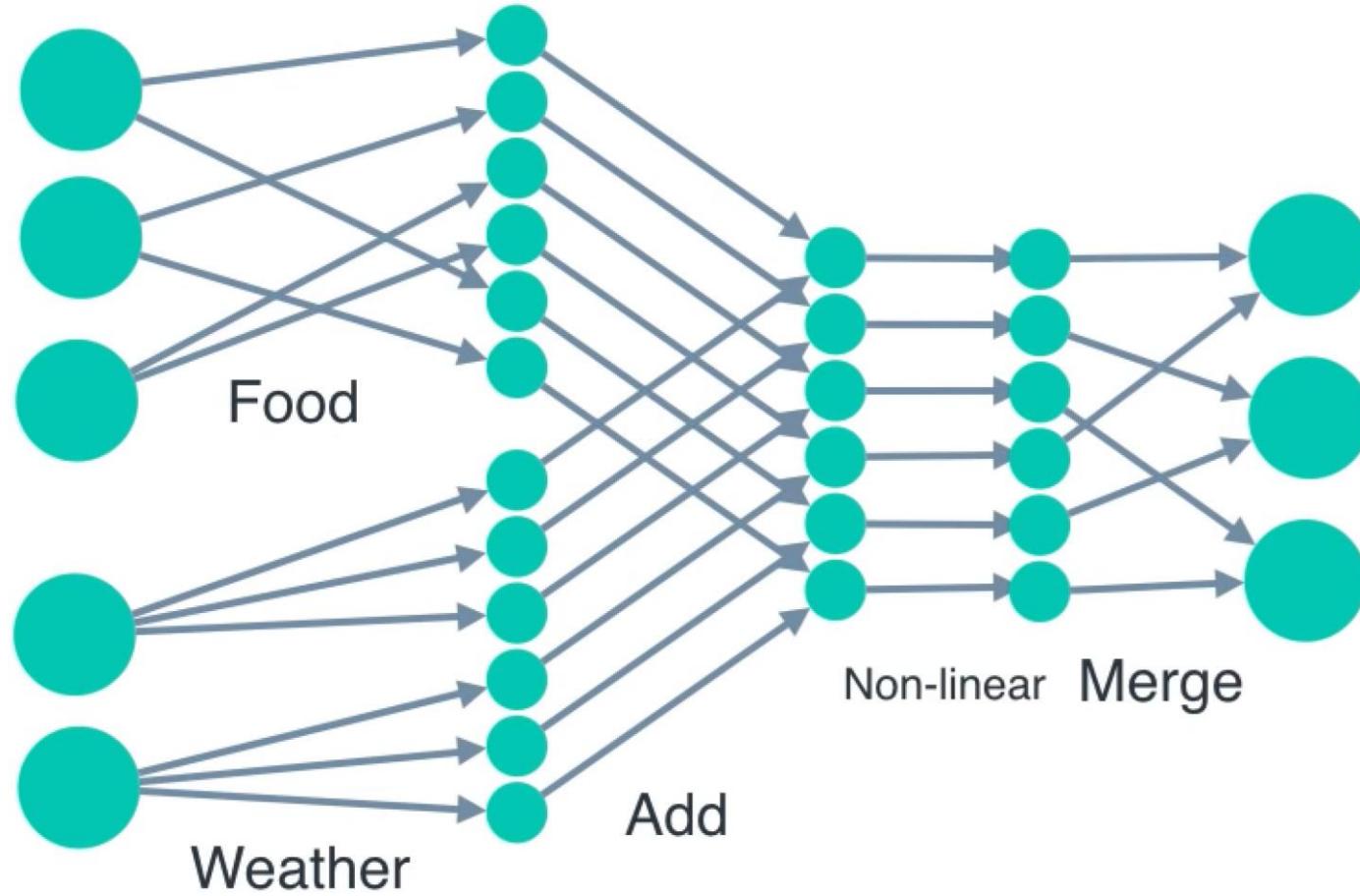


Merge

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \end{array} \right] = \left[\begin{array}{c} 0+0 \\ 0+1 \\ 0+0 \end{array} \right] = \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right] \text{Hamburger}$$

Merge

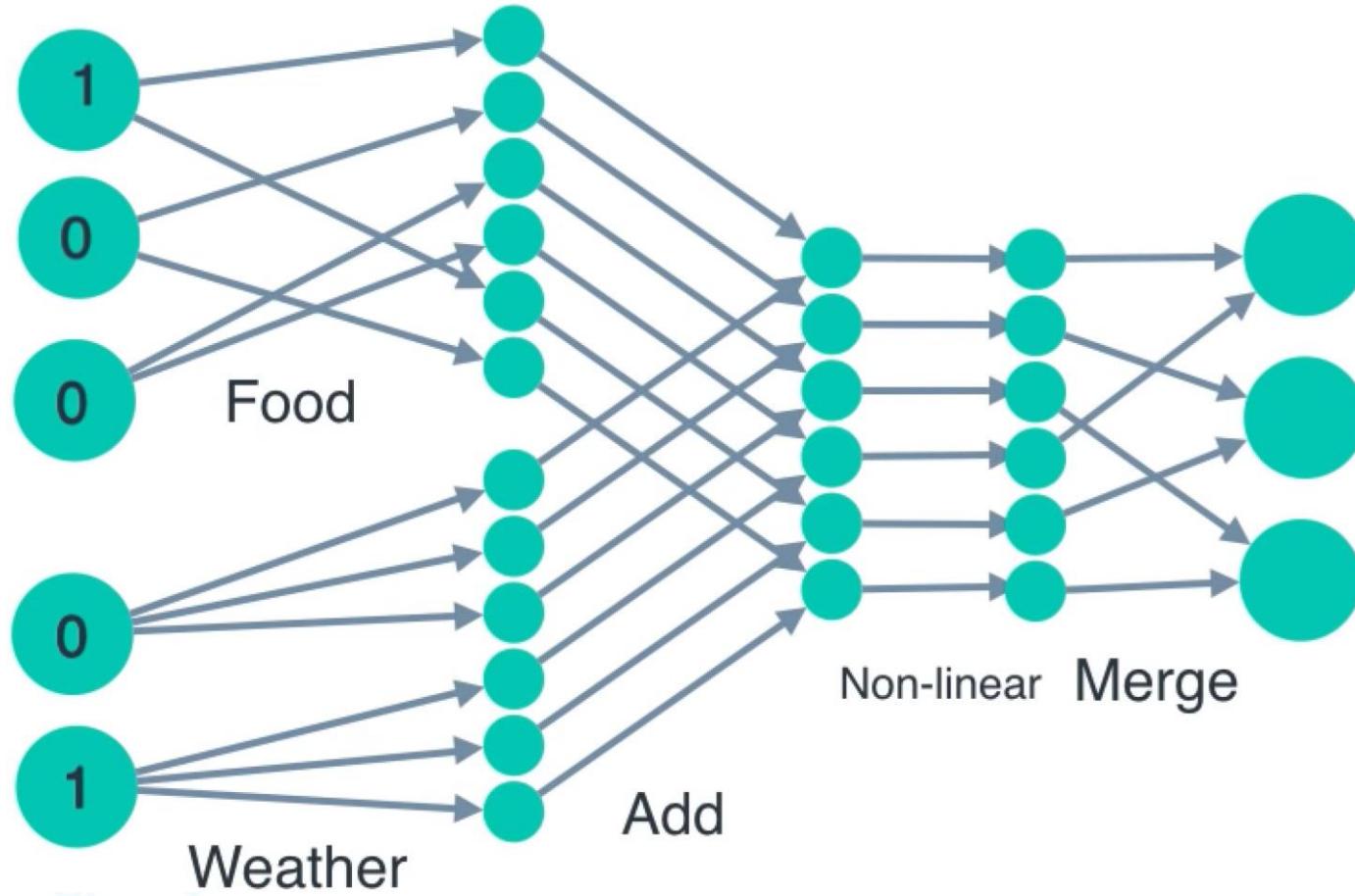
Redes Neurais Recorrentes (RNN)



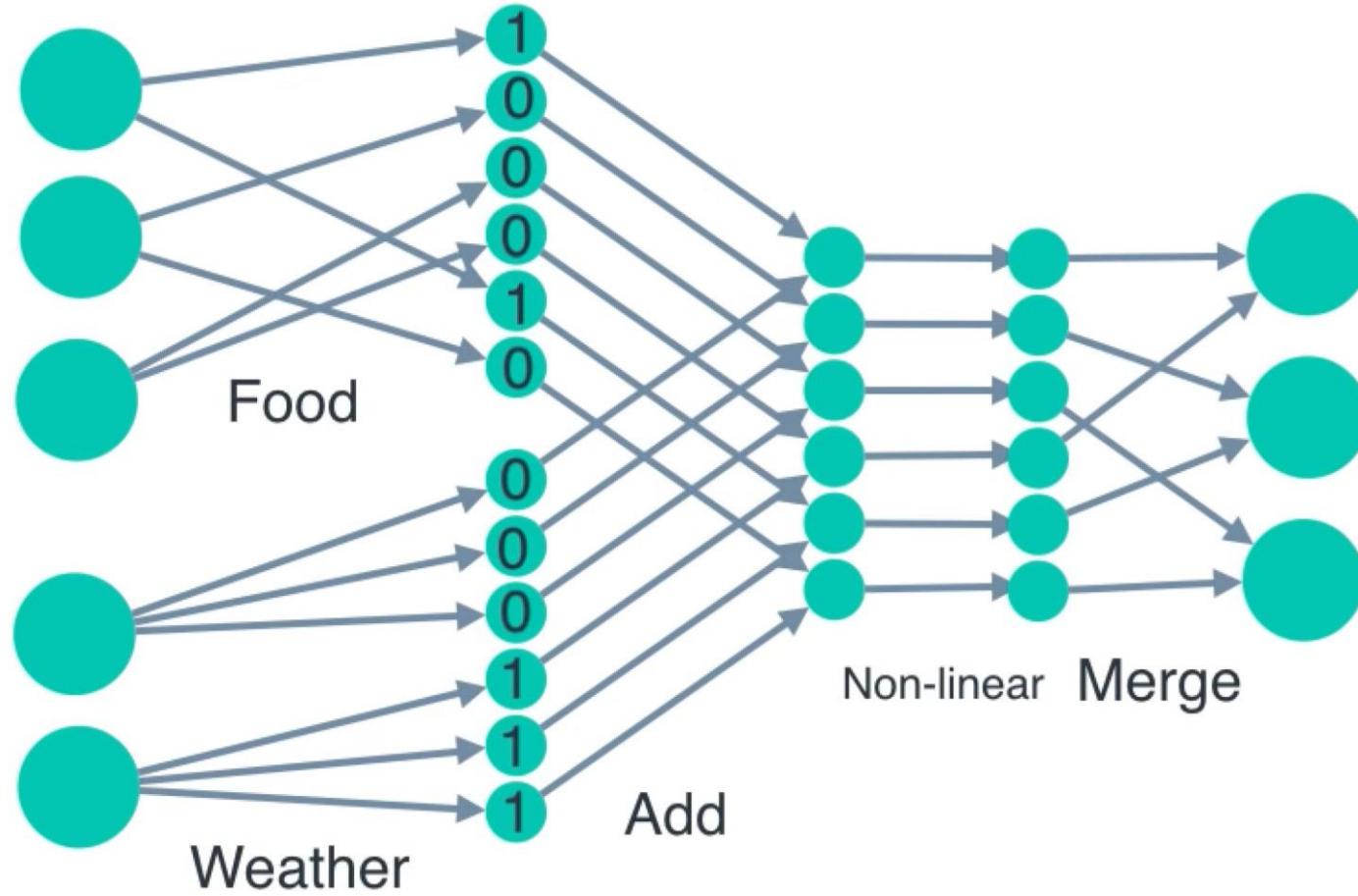
Redes Neurais Recorrentes (RNN)



$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

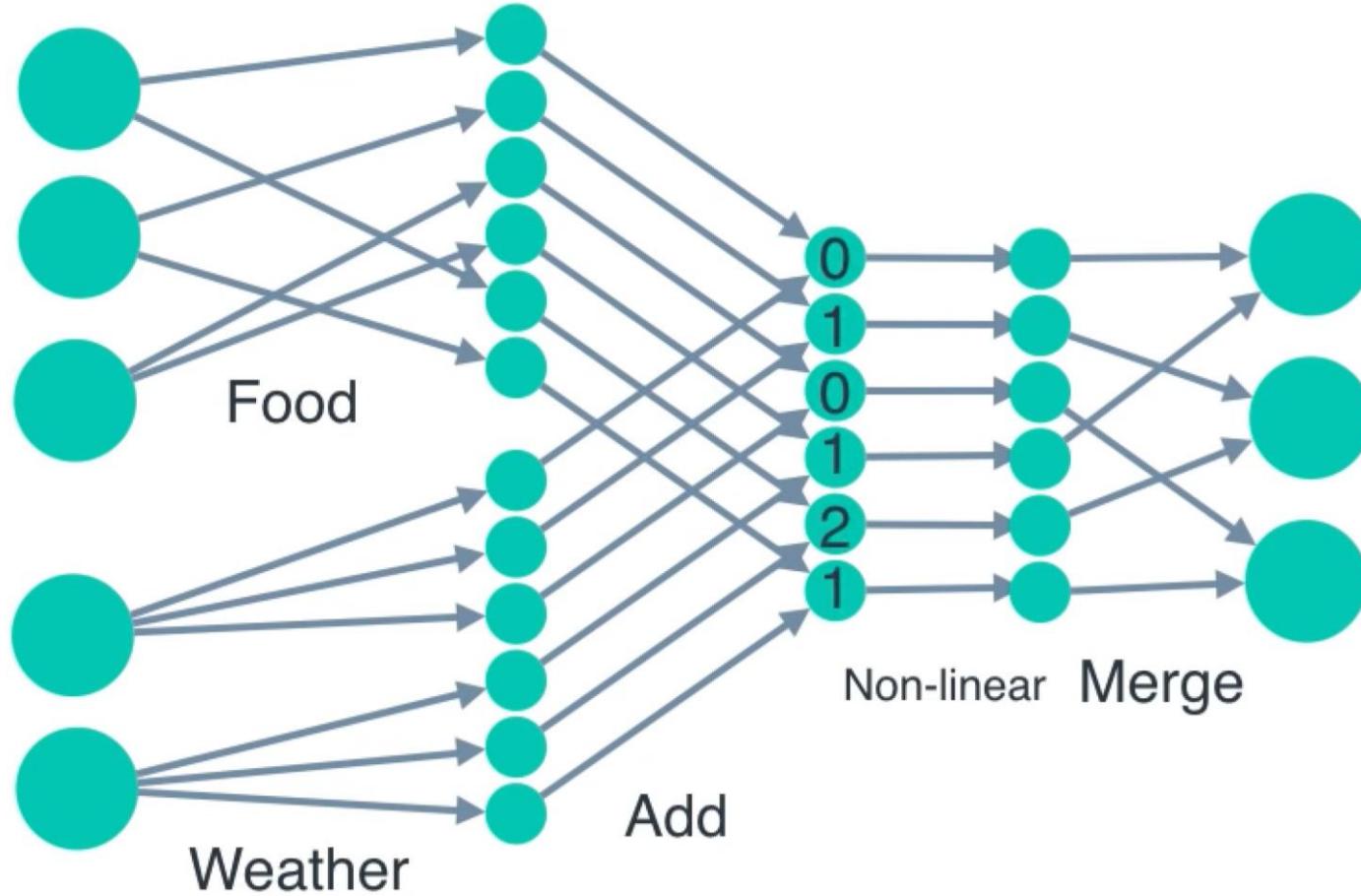


Redes Neurais Recorrentes (RNN)

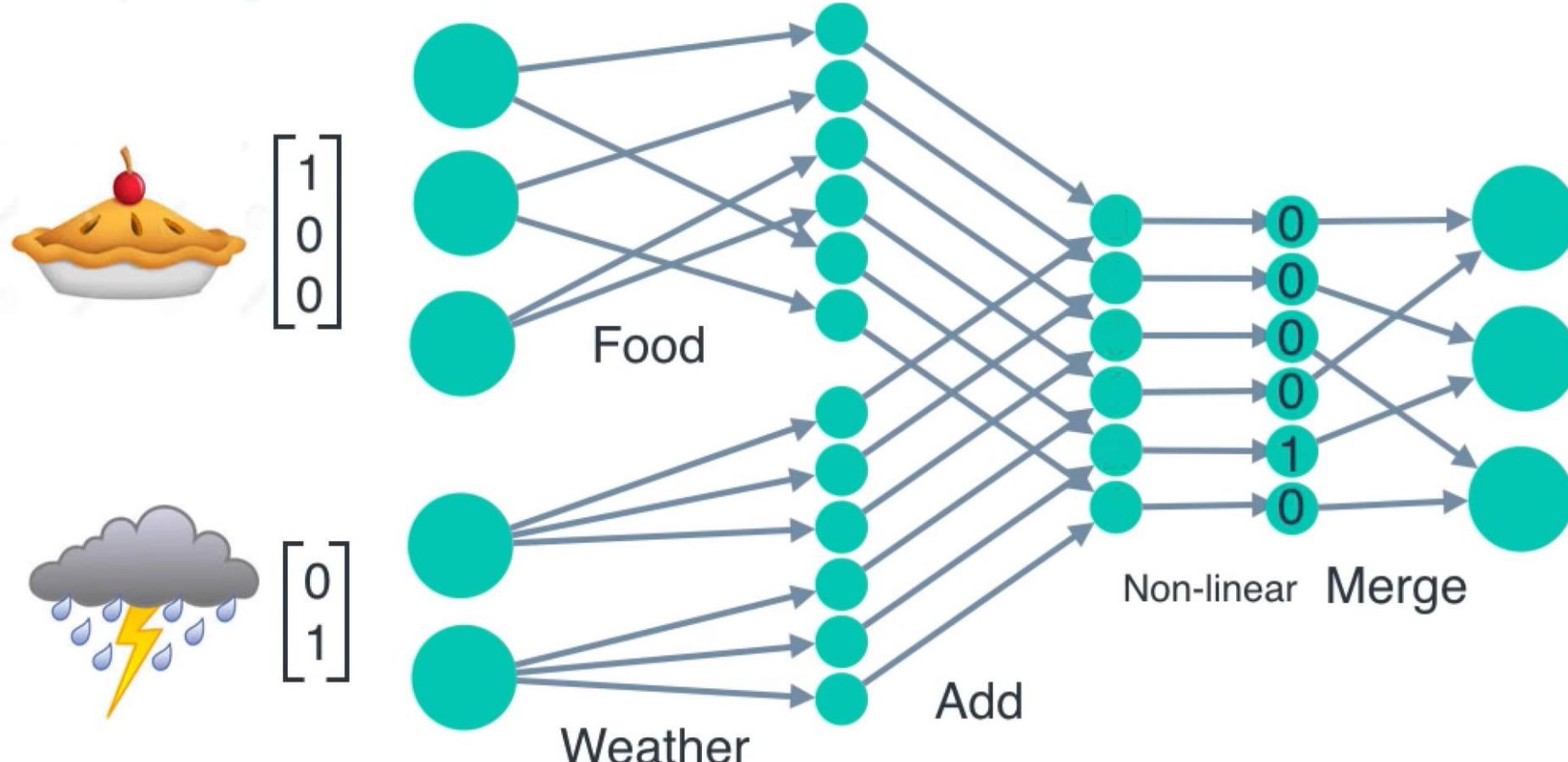

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$


Redes Neurais Recorrentes (RNN)

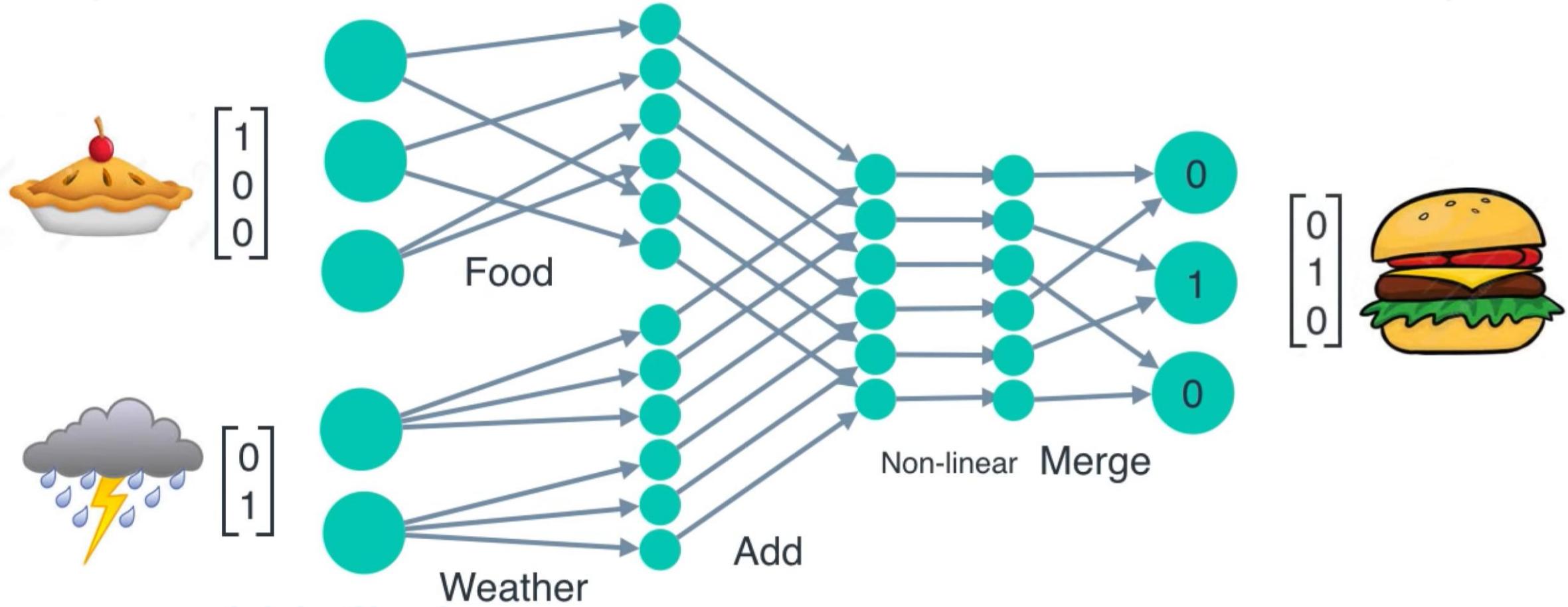

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$


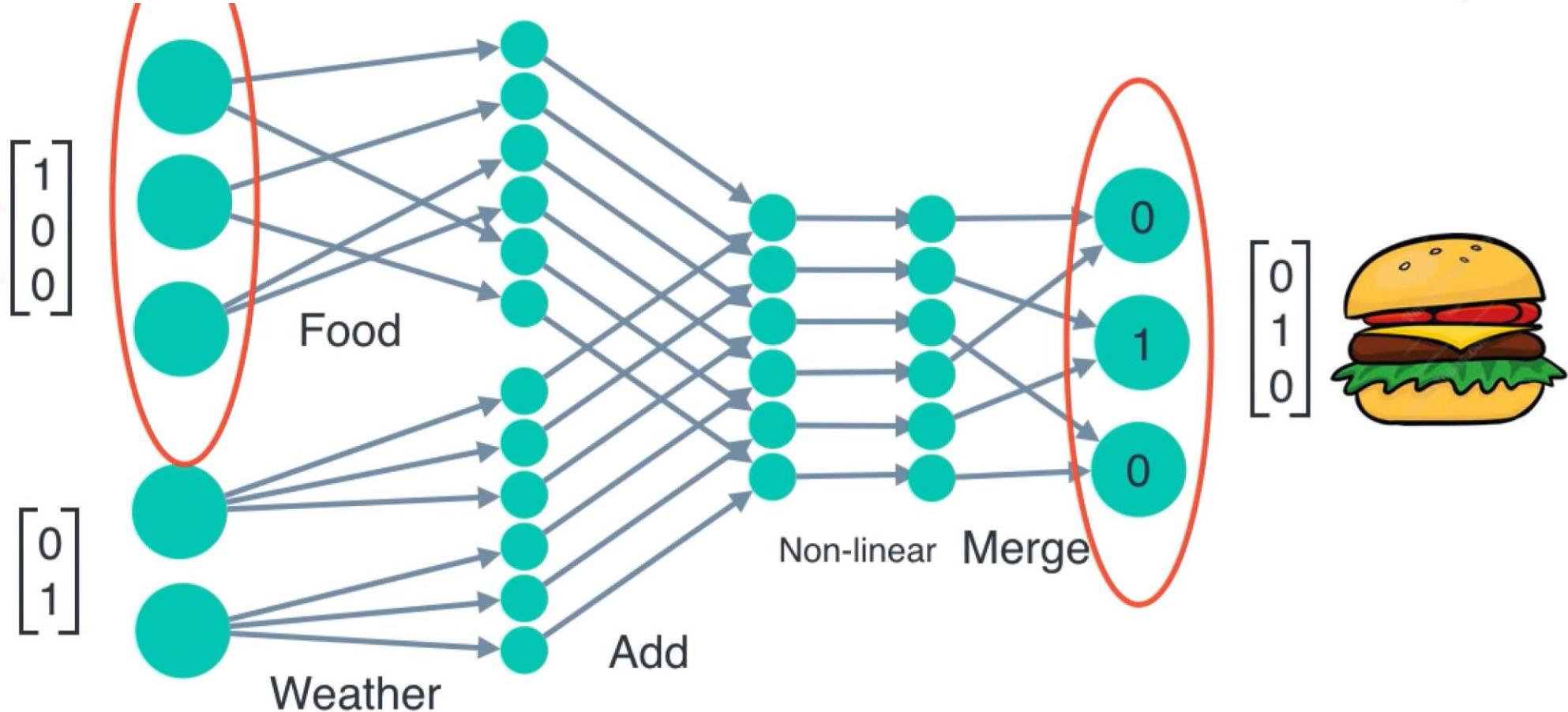
Redes Neurais Recorrentes (RNN)



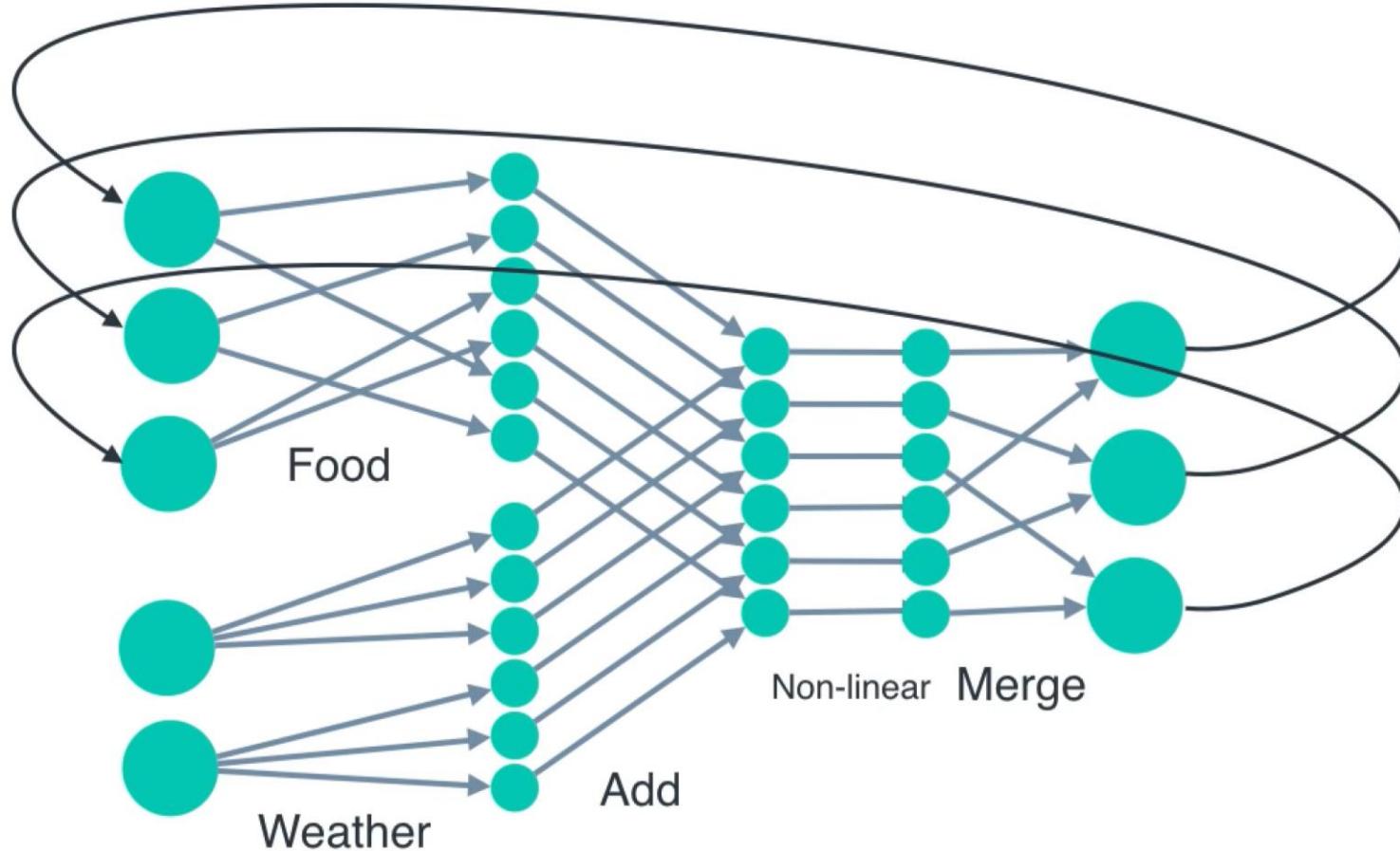
Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)



Redes Neurais Recorrentes (RNN)

$$h_t = f_W(h_{t-1}, x_t)$$

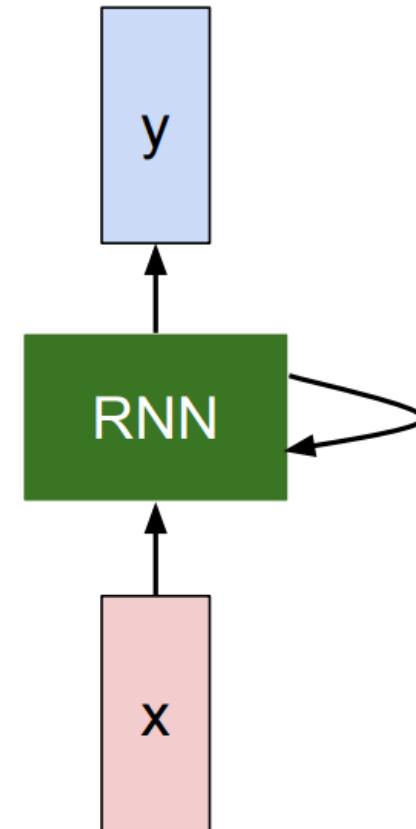
Podemos processar uma sequência de vetores x aplicando uma fórmula de recorrência a cada etapa de tempo (*time step*).

h_t - novo estado;

f_w - alguma função com parâmetros;

h_{t-1} - estado antigo;

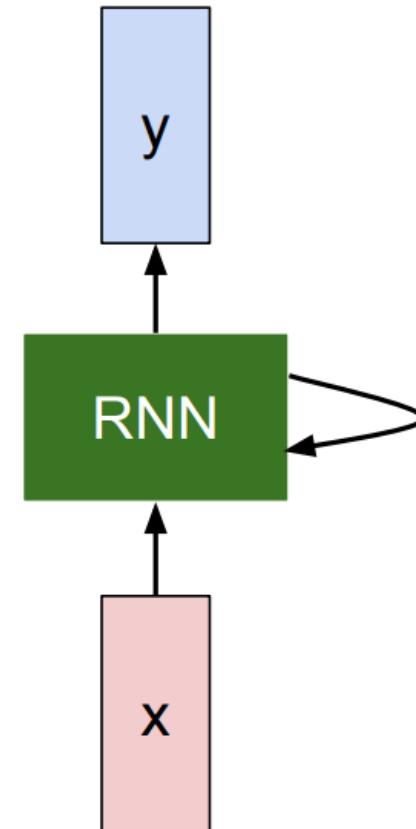
x_t – vetor de entrada em algum *time step*.



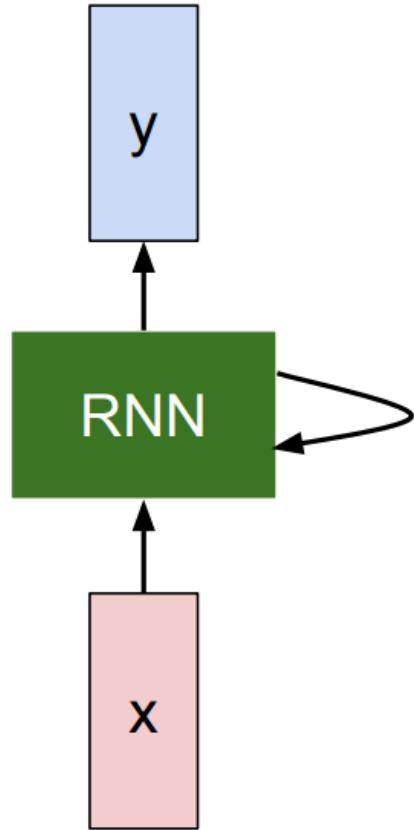
Redes Neurais Recorrentes (RNN)

Podemos processar uma sequência de vetores x aplicando uma fórmula de recorrência a cada etapa de tempo (*time step*).

$$h_t = f_W(h_{t-1}, x_t)$$



Redes Neurais Recorrentes (RNN) “Vanilla”



$$h_t = f_W(h_{t-1}, x_t)$$

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

$$y_t = W_{hy}h_t$$

O estado consiste em um único vetor "oculto" h:

RNN e LSTMs

Multilayer RNNs

$$h_t^l = \tanh W^l \begin{pmatrix} h_t^{l-1} \\ h_{t-1}^l \end{pmatrix}$$

$h \in \mathbb{R}^n$ $W^l [n \times 2n]$

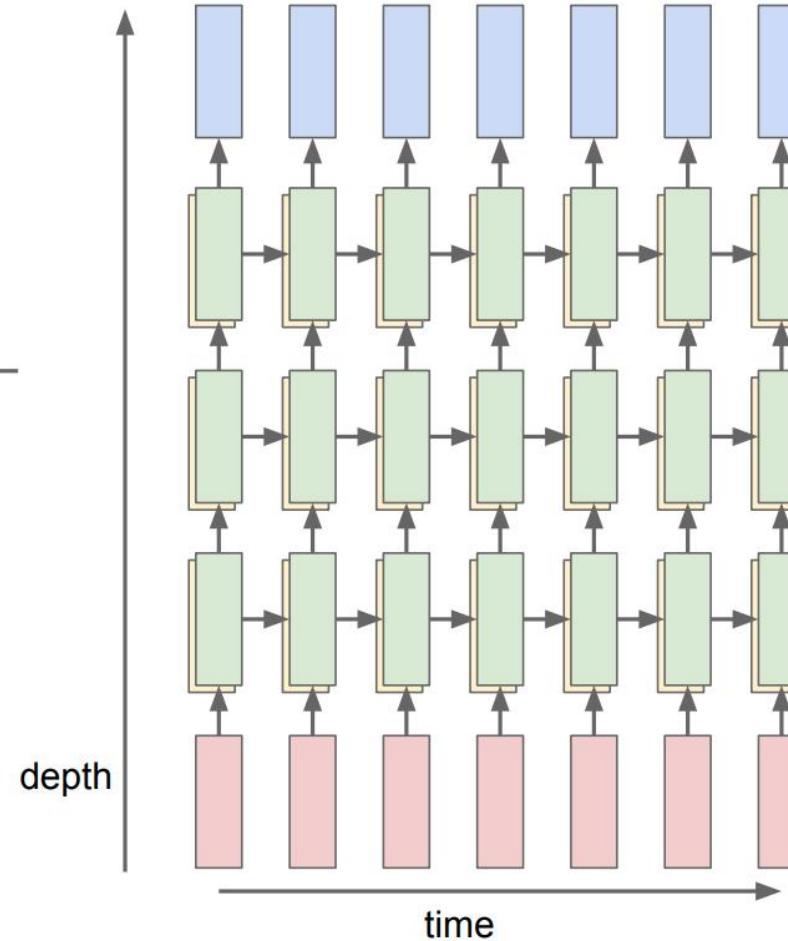
LSTM:

$$W^l [4n \times 2n]$$

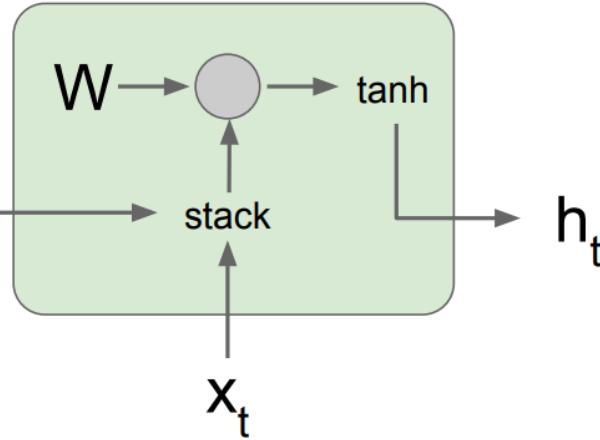
$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \text{sigm} \\ \text{sigm} \\ \text{sigm} \\ \tanh \end{pmatrix} W^l \begin{pmatrix} h_t^{l-1} \\ h_{t-1}^l \end{pmatrix}$$

$$c_t^l = f \odot c_{t-1}^l + i \odot g$$

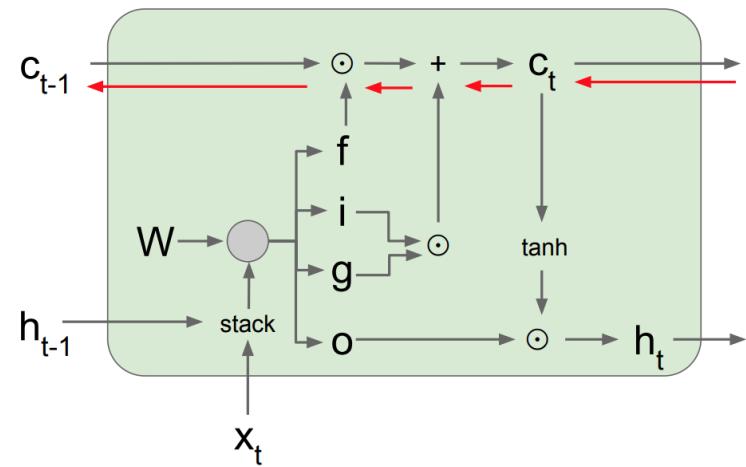
$$h_t^l = o \odot \tanh(c_t^l)$$



RNN e LSTMs



$$\begin{aligned}
 h_t &= \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \\
 &= \tanh \left(\begin{pmatrix} W_{hh} & W_{hx} \end{pmatrix} \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right) \\
 &= \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right)
 \end{aligned}$$

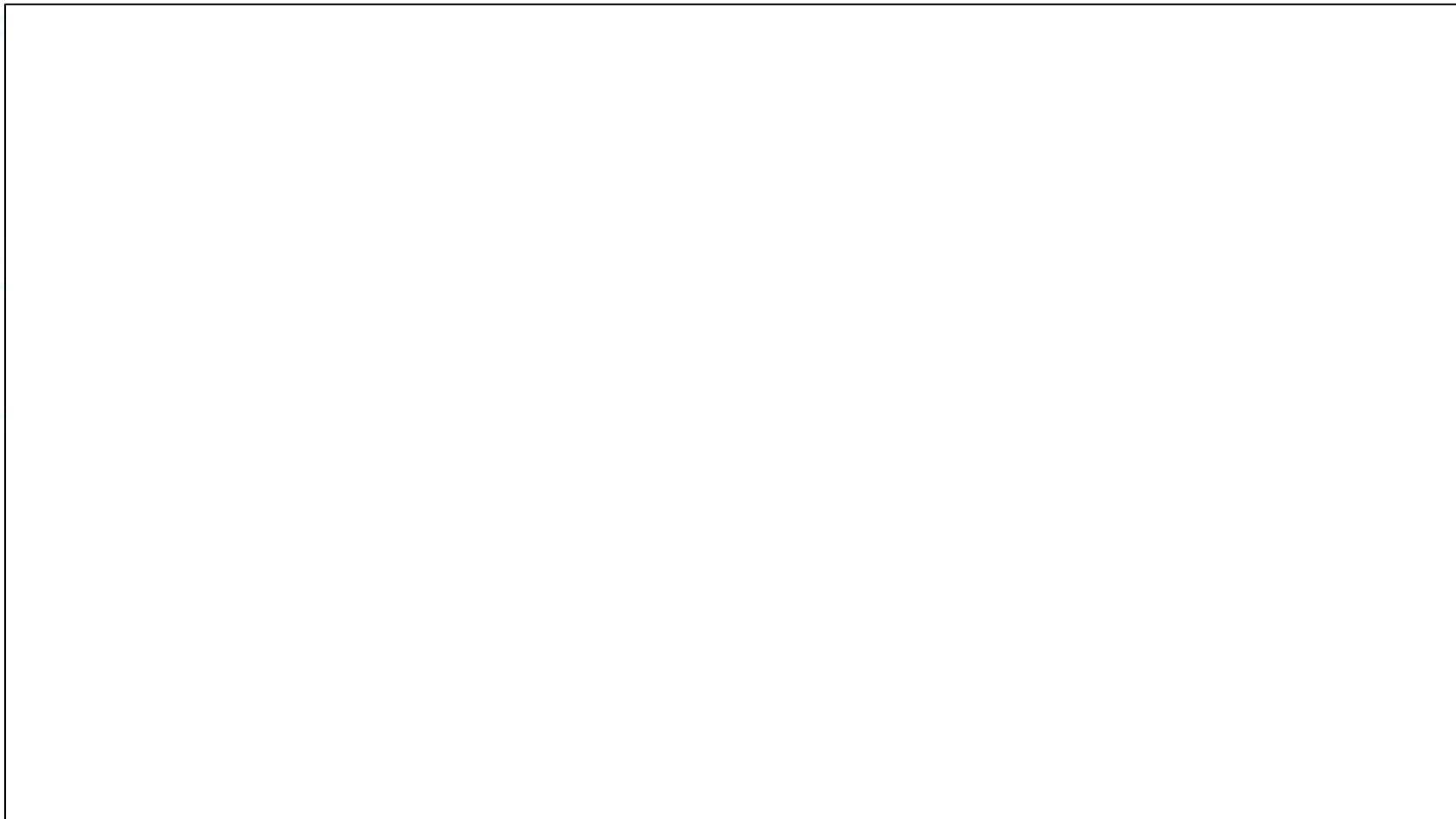


$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$

$$\begin{aligned}
 c_t &= f \odot c_{t-1} + i \odot g \\
 h_t &= o \odot \tanh(c_t)
 \end{aligned}$$

- f: Forget gate, se deseja apagar célula;
- i: Input gate, se deve ou não gravar na célula;
- g: Gate gate (?), quanto deve se escrever na célula;
- o: Portão de saída, célula do quanto se deve reveler.

- Long Short-Term Memory (LSTM)
- Applications



Thanks !



Vinicius Fernandes Caridá

vfcarida@gmail.com



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida