

Redes Neuronales Recurrentes

Francisco Cervantes

Octubre, 2019

HOY ...

- Motivación
- Notación
- Modelo de red neuronal recurrente
- Forward Propagation

Motivación

Reconocimiento del discurso



“El valor de una idea radica en el uso de la misma”

Clasificación de sentimientos

“La cinta consigue ser controversial en cada festival que se presenta y no deja a nadie indiferente”



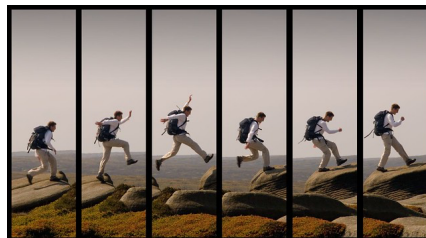
Traducción automática

¿Quieres cantar con migo?



Voulez-vous chanter avec moi?

Reconocimiento de acciones en video



Saltando

Notación

Utilicemos el siguiente ejemplo, para conocer la notación que utilizaremos al trabajar con redes neuronales recurrentes:

Supongamos que estamos abordando un problema de procesamiento de lenguaje natural en el que dado un texto debemos identificar el nombre de los personajes que intervienen en la historia. Por ejemplo:

Representación de datos de entrada y salida esperada:

	x:	Nikola	Tesla	y	Thomas	Edison	fueron	inventores.
Entrada:		$x^{<1>}$	$x^{<2>}$		$x^{<t>}$	$x^{<7>}$
Salida:	y:	1	1	0	1	1	0	0
		$y^{<1>}$	$y^{<2>}$		$y^{<t>}$	$y^{<7>}$

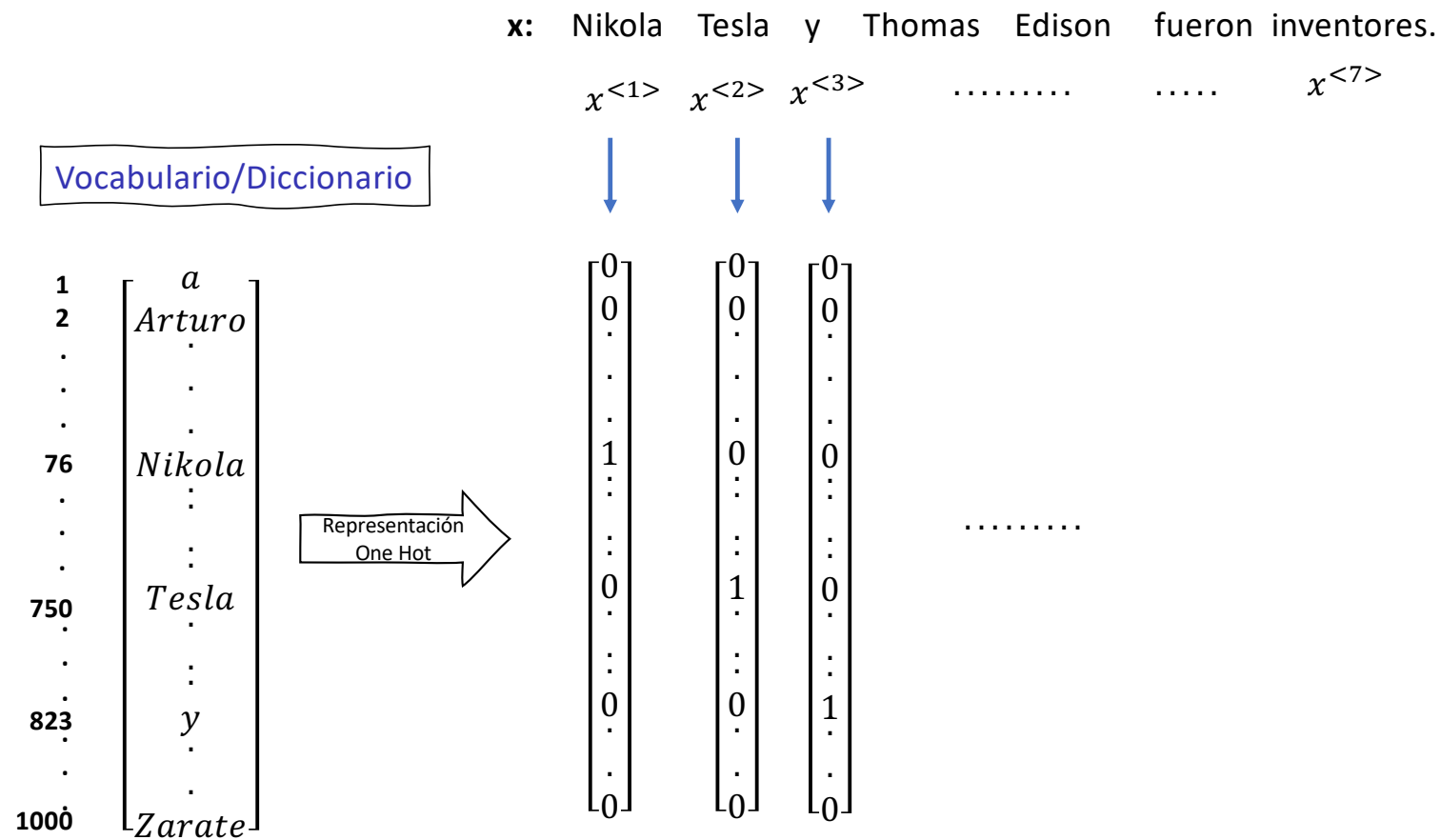
Notación

	x:	Nikola	Tesla	y	Thomas	Edison	fueron	inventores.
Entrada:		$x^{<1>}$	$x^{<2>}$	$x^{<t>}$	$x^{<7>}$	
Salida:	y:	1	1	0	1	1	0	0
		$y^{<1>}$	$y^{<2>}$	$y^{<t>}$	$y^{<7>}$	

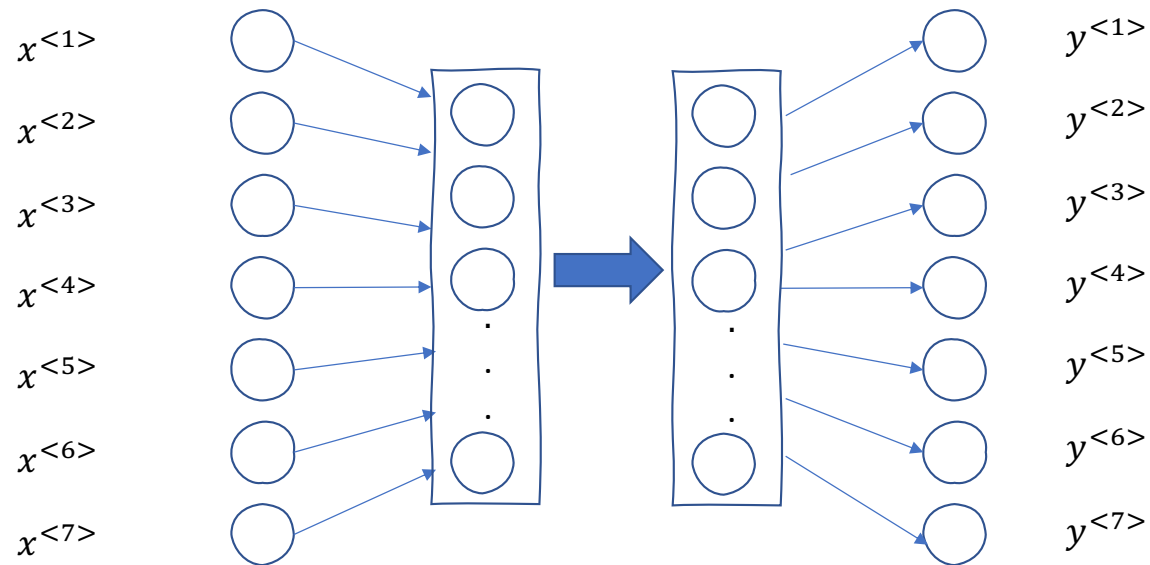
T_x	Longitud de la secuencia de entrada
T_y	Longitud de la secuencia de salida
$x^{(i)<t>}$	Elemento t de la secuencia de entrada del i -ésimo ejemplo
$T_x^{(i)}$	Longitud de la secuencia de entrada del i -ésimo ejemplo
$T_y^{(i)}$	Longitud de la secuencia de salida del i -ésimo ejemplo

Ahora tenemos que decidir
¿cómo representar cada
palabra de la secuencia?

Representación de palabras



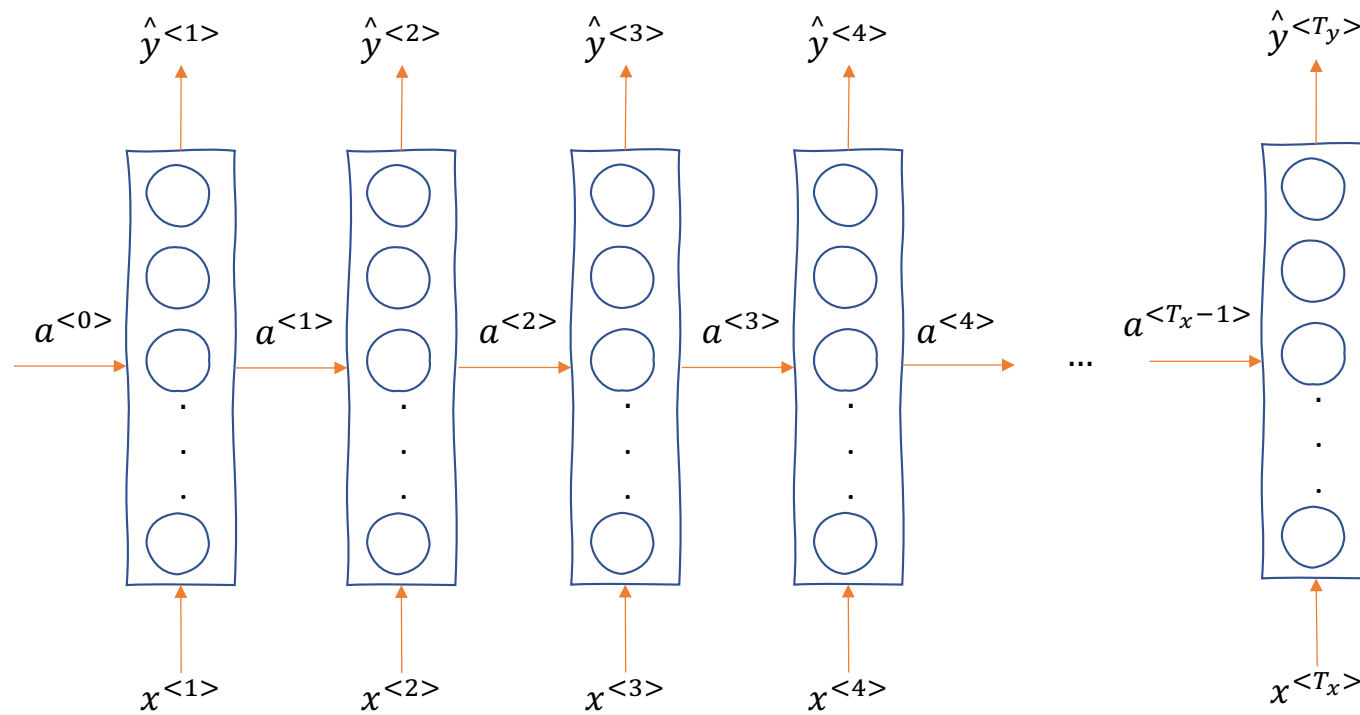
¿Por qué no utilizar una red neuronal clásica?



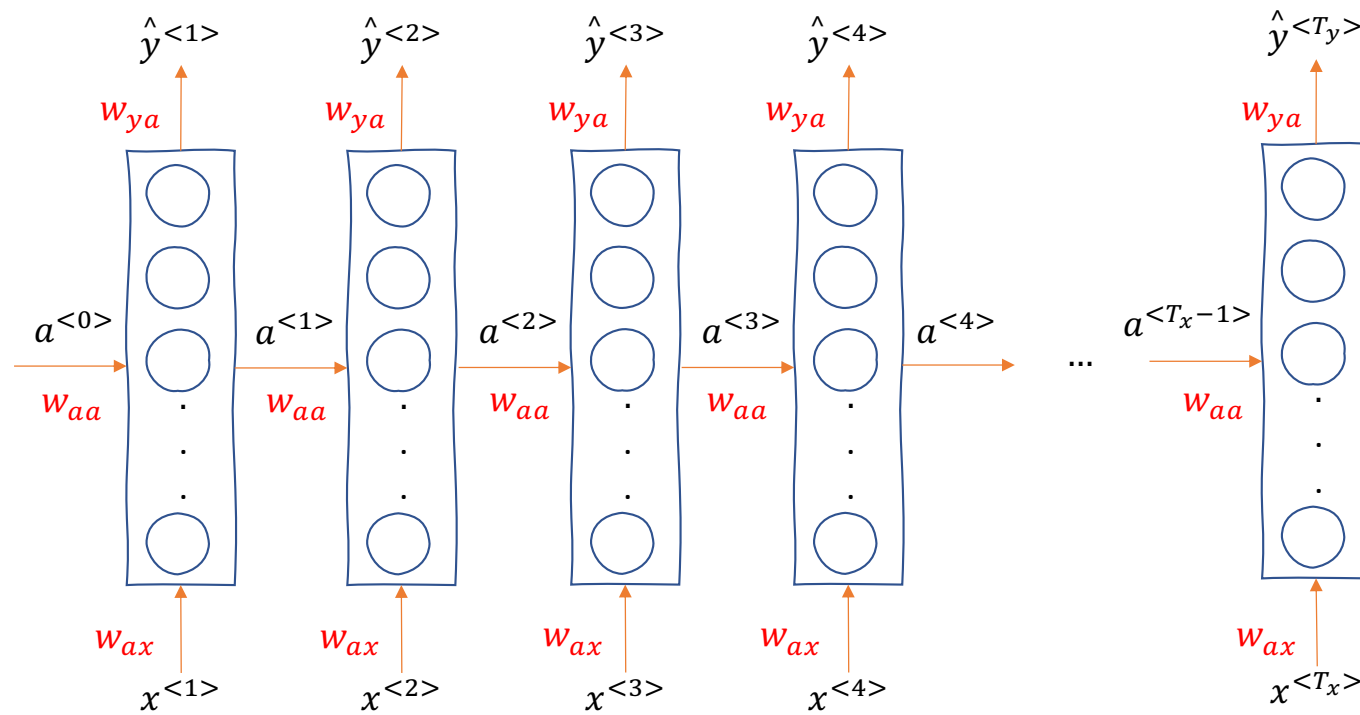
Problemas:

- La longitud de las secuencias de entrada/salida pueden variar para cada ejemplo.
- En una estructura clásica, no se comparten características aprendidas a través de diferentes posiciones del texto.

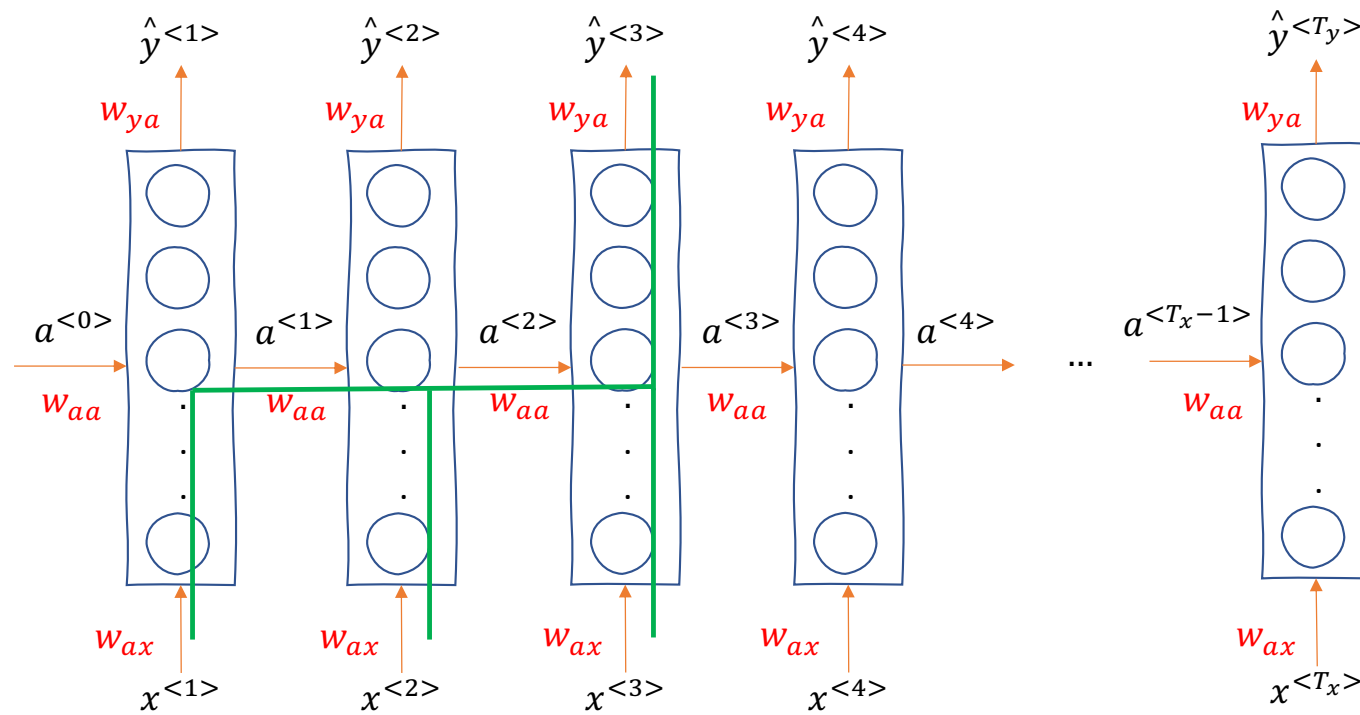
Representación de una red neuronal recurrente



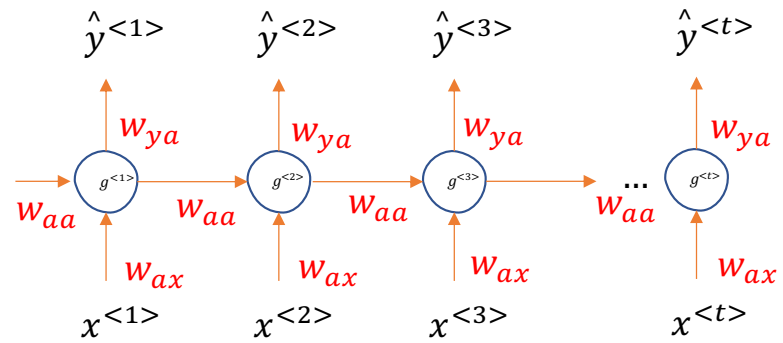
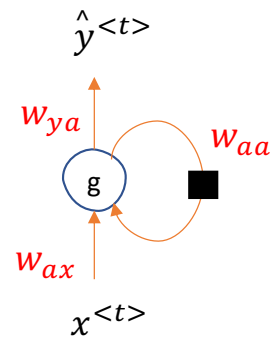
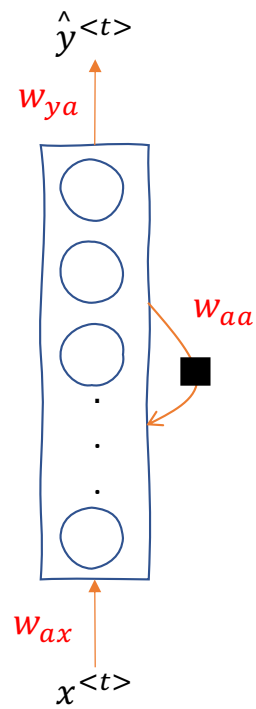
Representación de una red neuronal recurrente



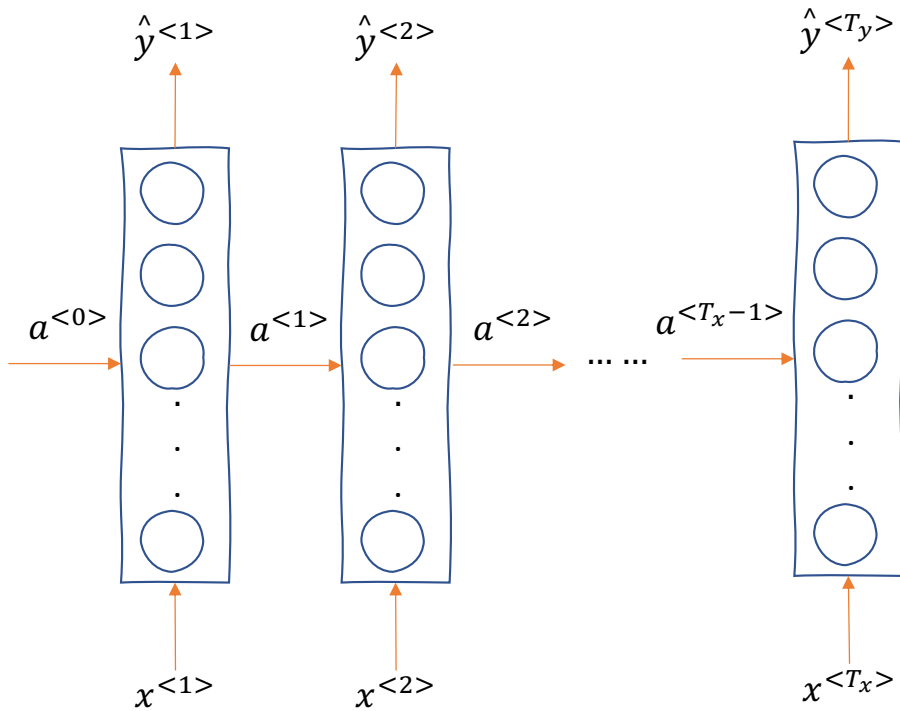
Representación de una red neuronal recurrente



Otras representaciones ...



Forward propagation



$$a^{<0>} = [0 \dots 0]$$

$$a^{<1>} = g_1(w_{aa}a^{<0>} + w_{ax}x^{<1>} + b_a)$$

$$\hat{y}^{<1>} = g_2(w_{ya}a^{<1>} + b_y)$$

$$a^{<t>} = \tanh(w_{aa}a^{<t-1>} + w_{ax}x^{<t>} + b_a)$$

$$\hat{y}^{<t>} = \text{softmax}(w_{ya}a^{<t>} + b_y)$$

