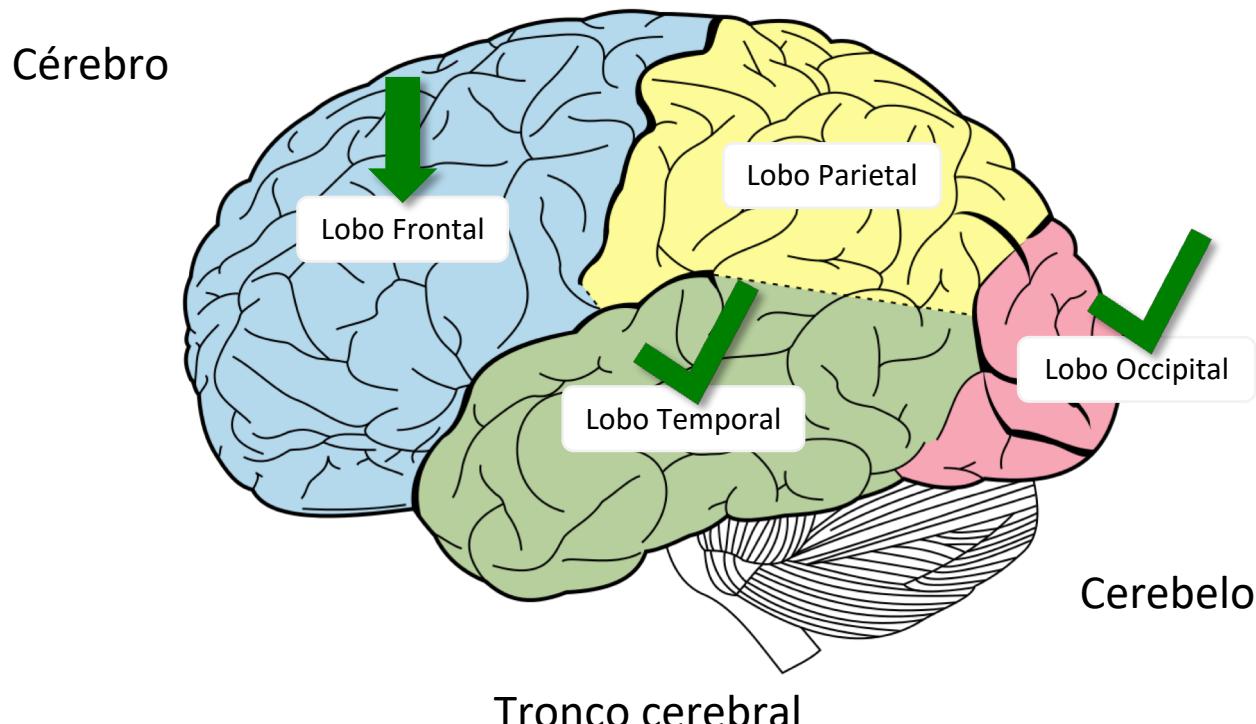


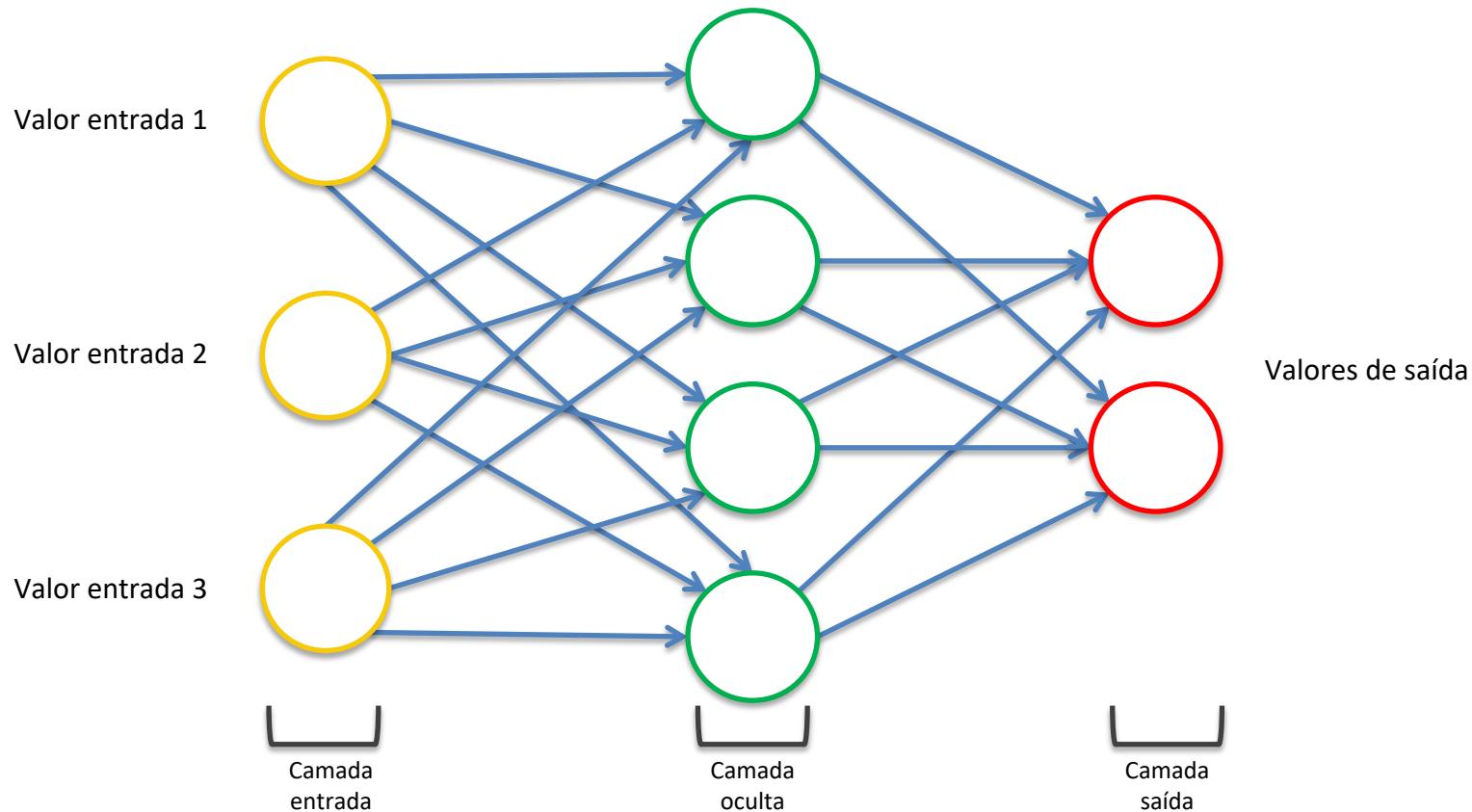
Redes Neurais Recorrentes (RNNs)

Redes Neurais Recorrentes



Fonte: Wikipedia

Redes Neurais Recorrentes

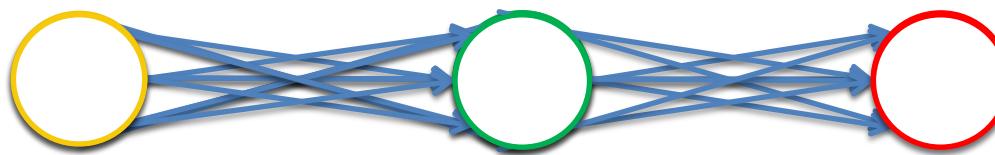


Redes Neurais Recorrentes

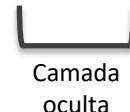
Valor entrada 1

Valor entrada 2

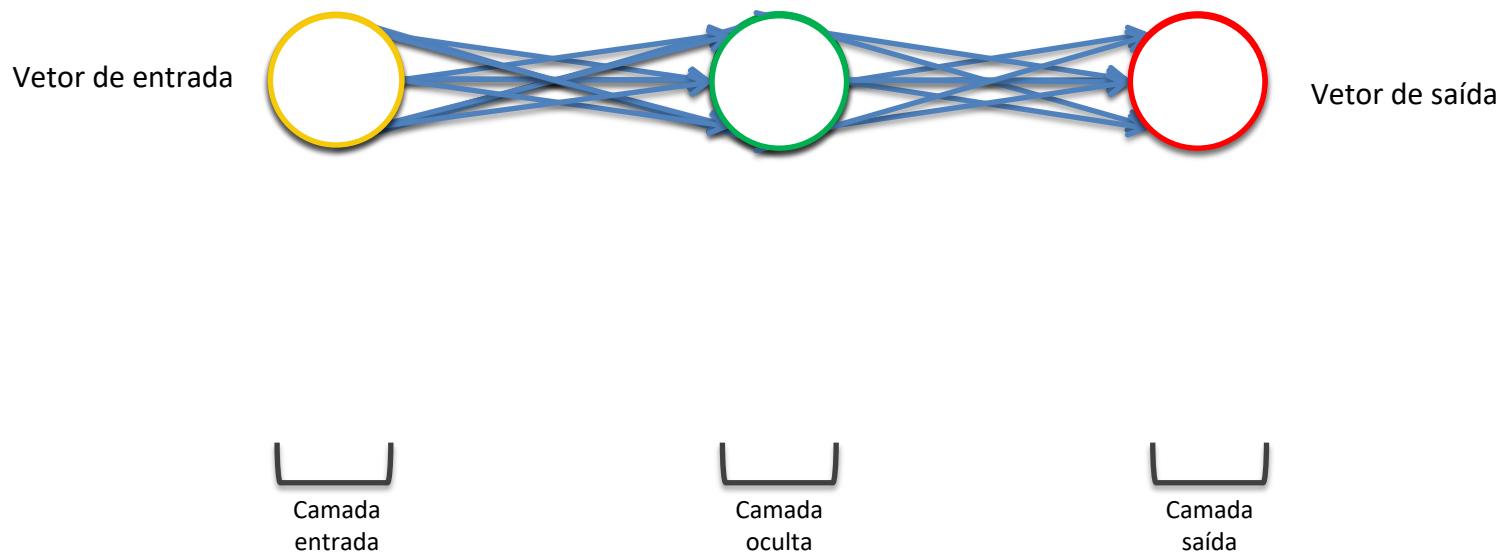
Valor entrada3



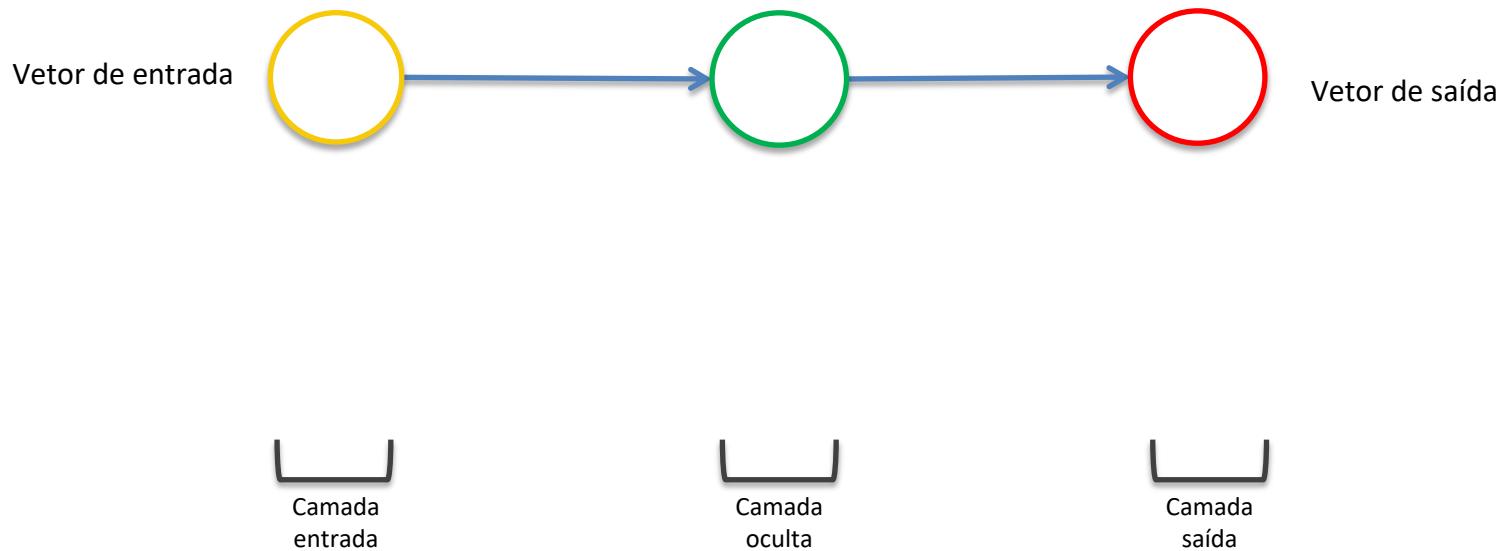
Valores de saída



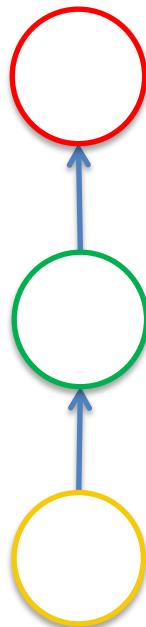
Redes Neurais Recorrentes



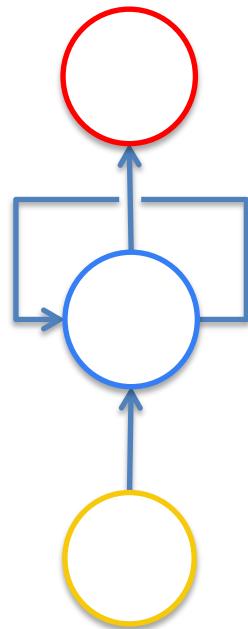
Redes Neurais Recorrentes



Redes Neurais Recorrentes



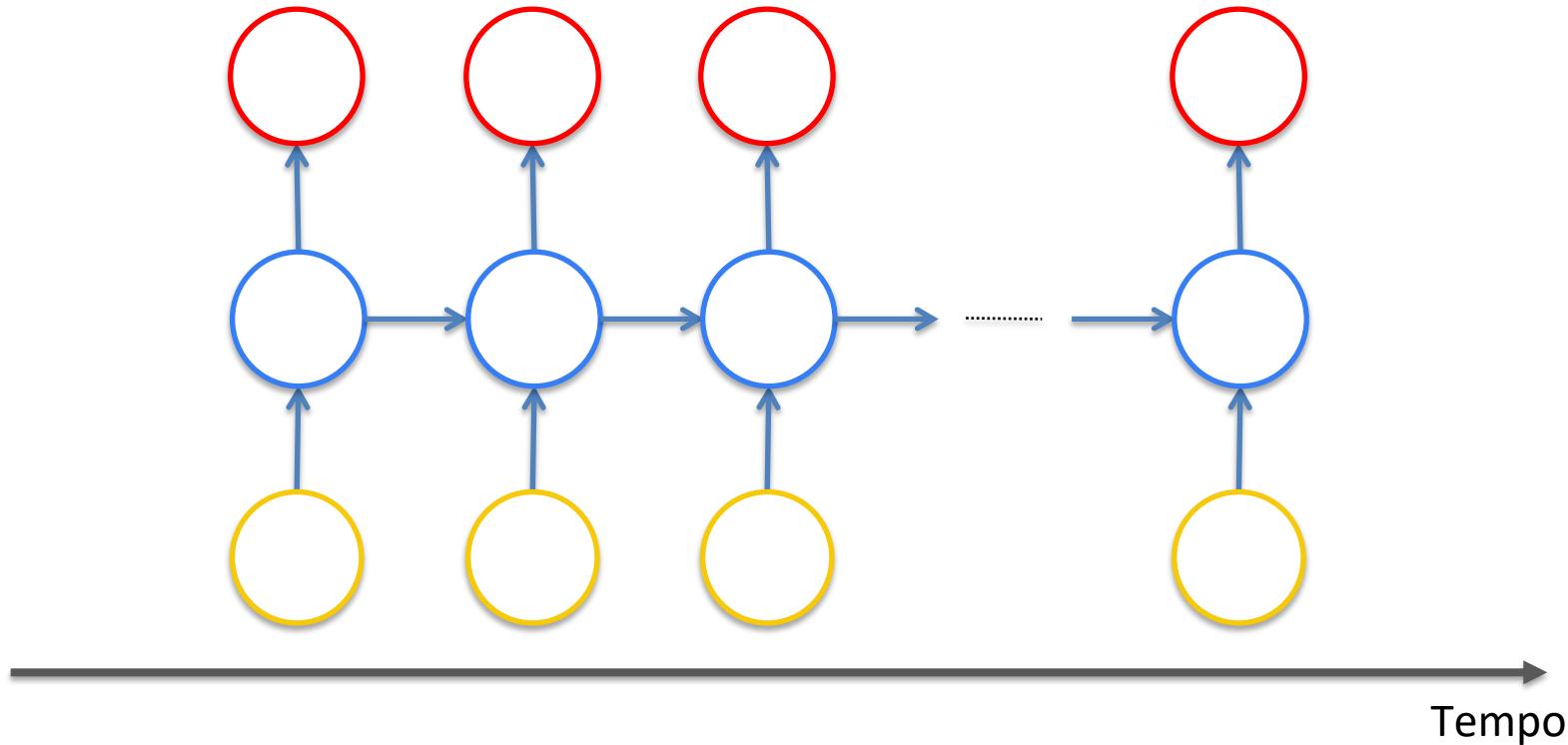
Redes Neurais Recorrentes



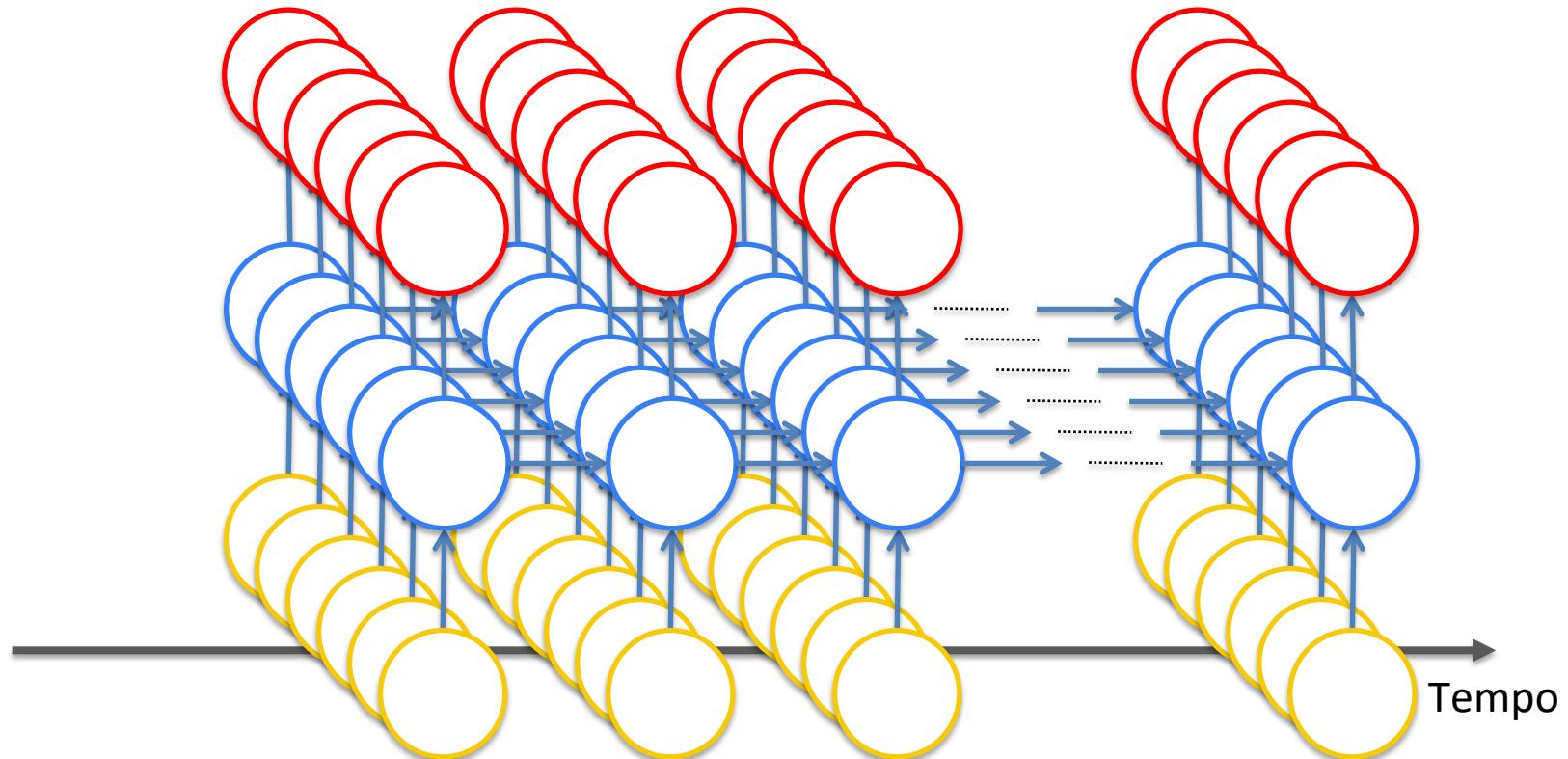
Redes Neurais Recorrentes



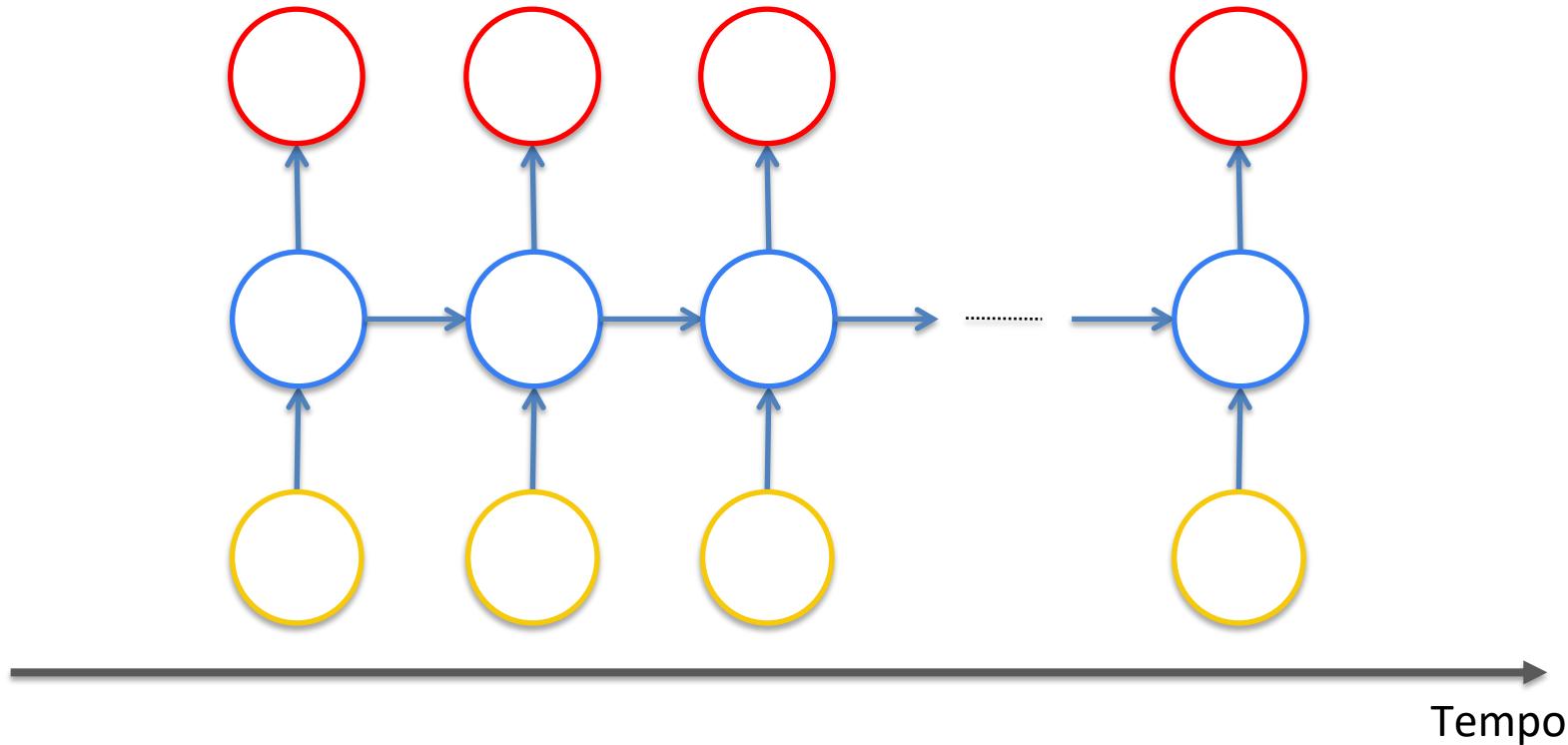
Redes Neurais Recorrentes



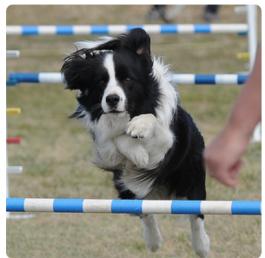
Redes Neurais Recorrentes



Redes Neurais Recorrentes

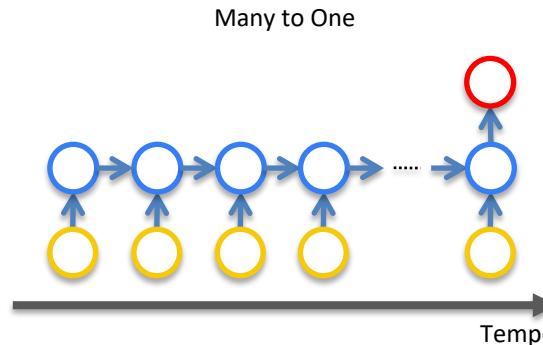
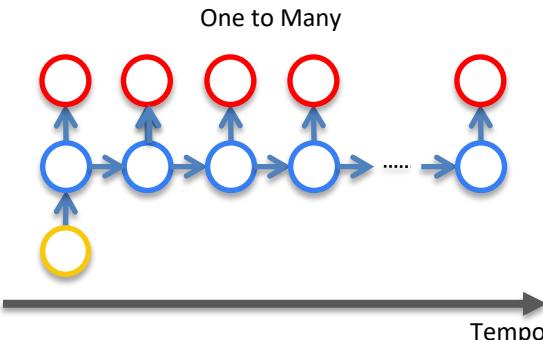


Redes Neurais Recorrentes



"black and white
dog jumps over
bar."

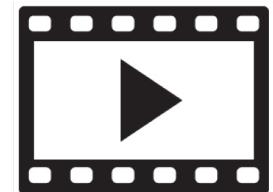
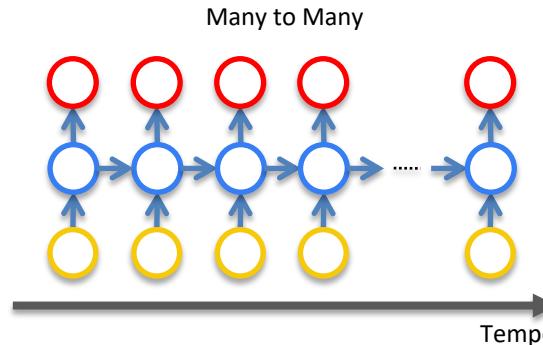
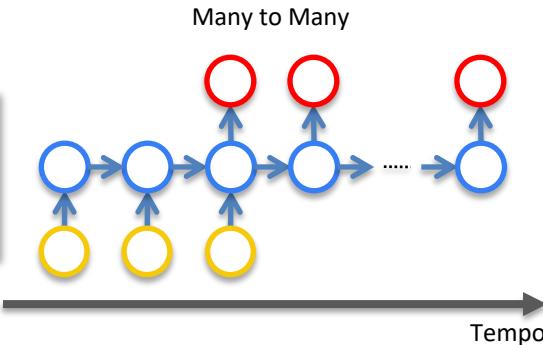
karpathy.github.io



"Thanks for a great
party at the
weekend, we really
enjoyed it!"

> sentiment: positive
score: 86%

dev.havenondemand.com



Referência: karpathy.github.io

Redes Neurais Recorrentes

Vídeo adicional:

Sunspring (filme, 2016)

- Diretor: Oscar Sharp
- Escrito por Benjamin

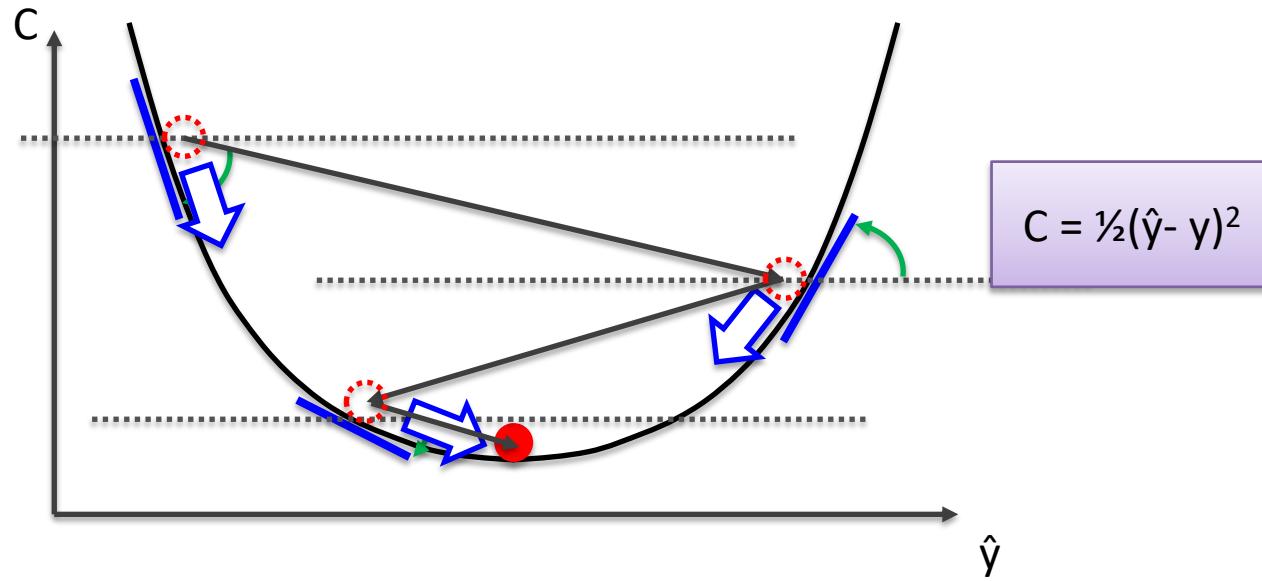


Link:

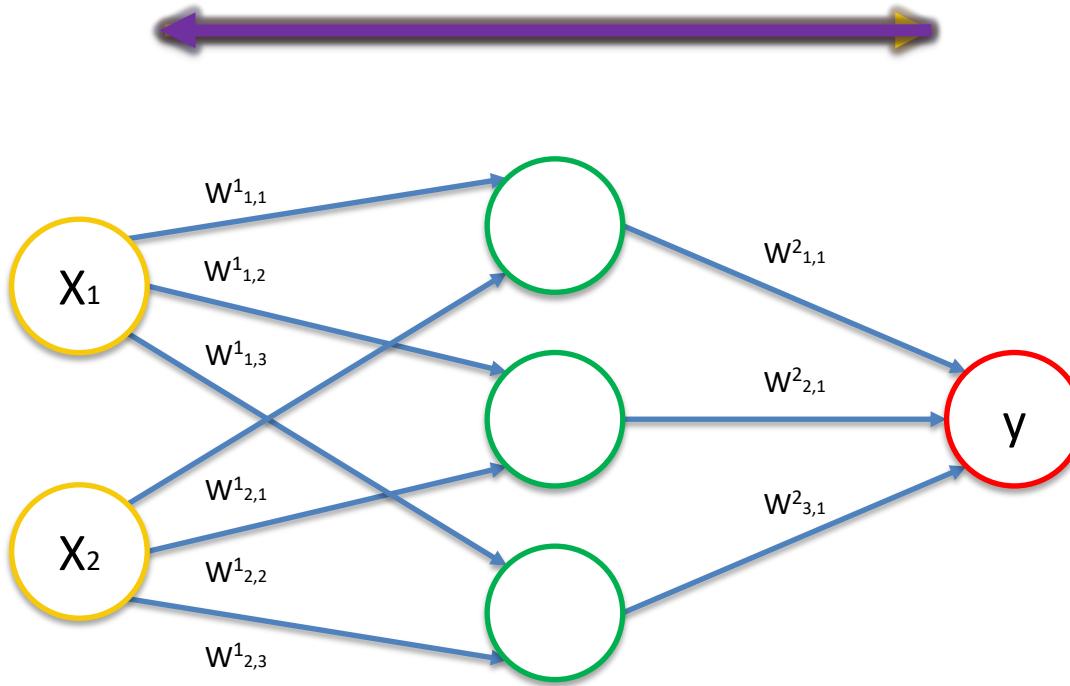
<https://arstechnica.com/the-multiverse/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/>

Vanishing Gradient (Gradiente Desaparecendo)

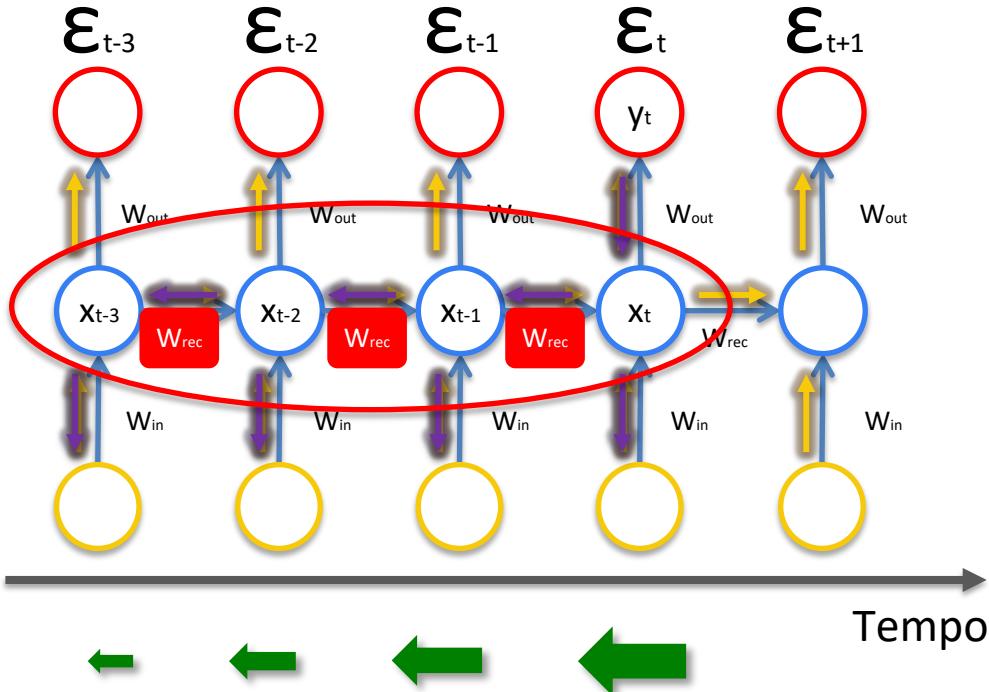
Vanishing Gradient Problem



Vanishing Gradient Problem



Vanishing Gradient Problem



$$\frac{\partial \mathcal{E}}{\partial \theta} = \sum_{1 \leq t \leq T} \frac{\partial \mathcal{E}_t}{\partial \theta} \quad (3)$$

$$\frac{\partial \mathcal{E}_t}{\partial \theta} = \sum_{1 \leq k \leq t} \left(\frac{\partial \mathcal{E}_t}{\partial \mathbf{x}_t} \frac{\partial \mathbf{x}_t}{\partial \mathbf{x}_k} \frac{\partial^+ \mathbf{x}_k}{\partial \theta} \right) \quad (4)$$

$$\frac{\partial \mathbf{x}_t}{\partial \mathbf{x}_k} = \prod_{t \geq i > k} \frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_{i-1}} = \prod_{t \geq i > k} \mathbf{W}_{rec}^T diag(\sigma'(\mathbf{x}_{i-1})) \quad (5)$$

$W_{rec} \sim \text{pequeno} \rightarrow$ Vanishing

$W_{rec} \sim \text{grande} \rightarrow$ Exploding

Fonte da fórmula: Razvan Pascanu et al. (2013)

Vanishing Gradient Problem

Soluções:

1. Exploding Gradient

- Truncated Backpropagation
- Penalidades
- Gradient Clipping

2. Vanishing Gradient

- Inicialização dos pesos
- Echo State Networks
- Long Short-Term Memory Networks (LSTMs)



Vanishing Gradient Problem

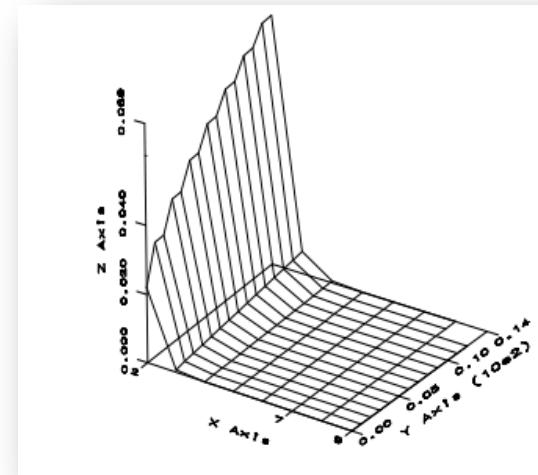
Leitura adicional:

*Untersuchungen zu dynamischen
neuronalen Netzen*

Sepp (Josef) Hochreiter (1991)

Link:

<http://people.idsia.ch/~juergen/SeppHochreiter1991ThesisAdvisorSchmidhuber.pdf>



Vanishing Gradient Problem

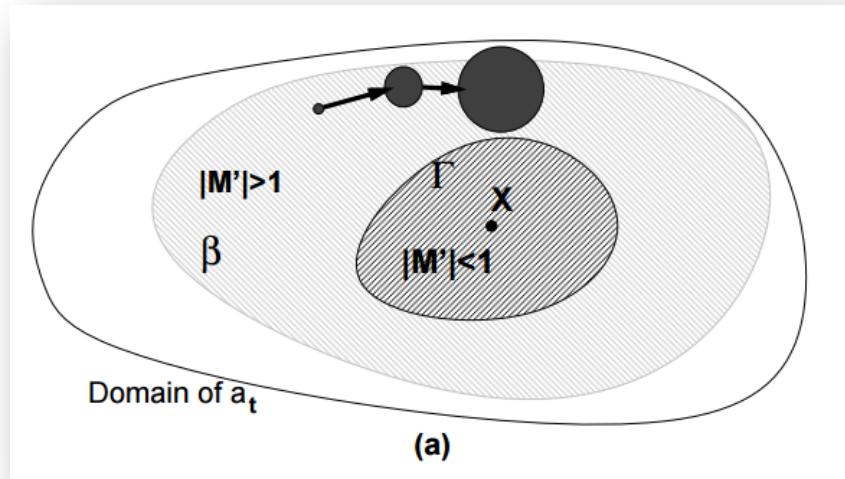
Leitura adicional:

*Learning Long-Term Dependencies
with Gradient Descent is Difficult*

Yoshua Bengio et al. (1994)

Link:

<http://www-dsi.ing.unifi.it/~paolo/ps/tnn-94-gradient.pdf>



Vanishing Gradient Problem

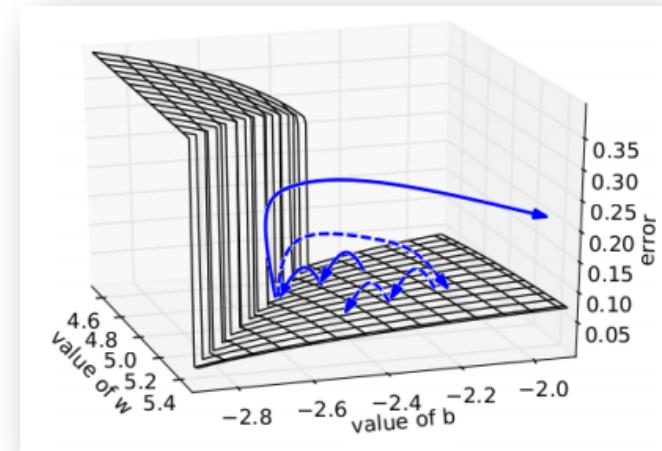
Leitura adicional:

On the difficulty of training recurrent neural networks

Razvan Pascanu et al. (2013)

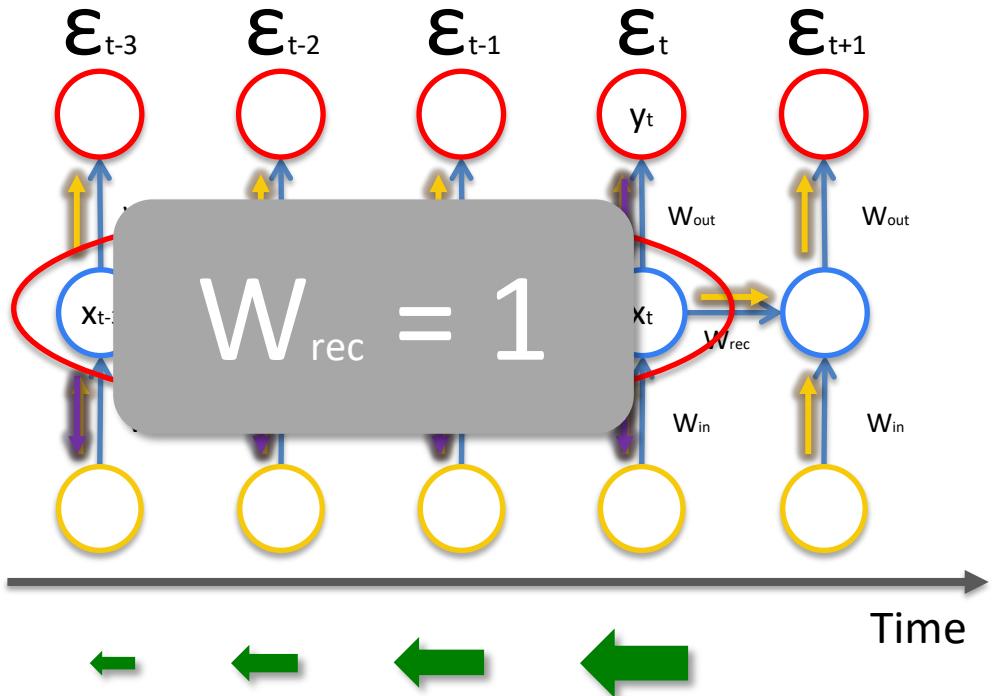
Link:

<http://www.jmlr.org/proceedings/papers/v28/pascanu13.pdf>



LSTMs

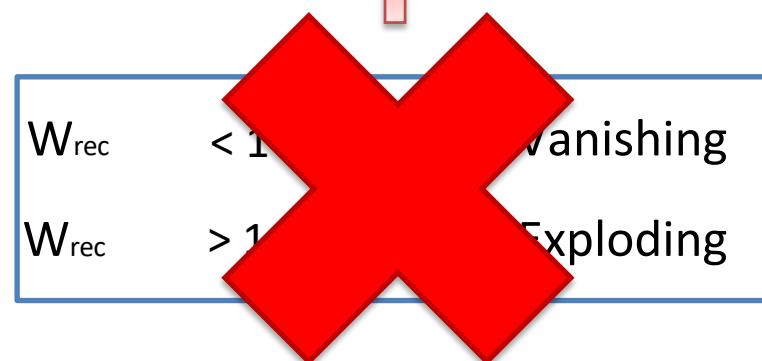
Long Short-Term Memory



$$\frac{\partial \mathcal{E}}{\partial \theta} = \sum_{1 \leq t \leq T} \frac{\partial \mathcal{E}_t}{\partial \theta} \quad (3)$$

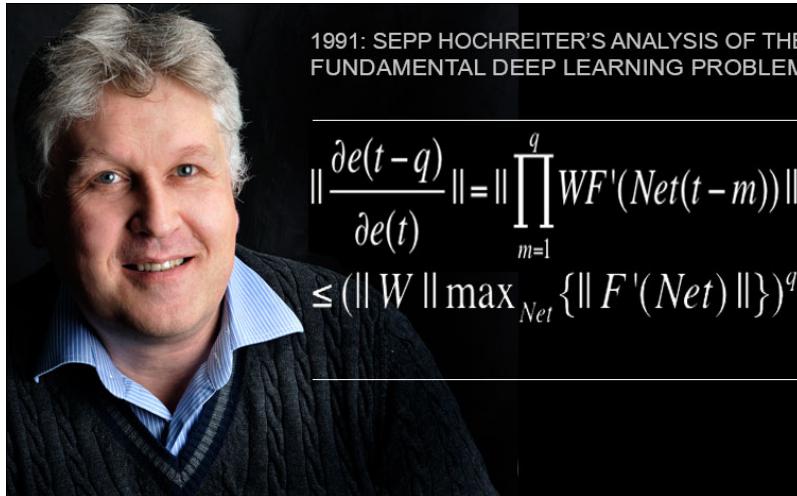
$$\frac{\partial \mathcal{E}_t}{\partial \theta} = \sum_{1 \leq k \leq t} \left(\frac{\partial \mathcal{E}_t}{\partial \mathbf{x}_t} \frac{\partial \mathbf{x}_t}{\partial \mathbf{x}_k} \frac{\partial^+ \mathbf{x}_k}{\partial \theta} \right) \quad (4)$$

$$\frac{\partial \mathbf{x}_t}{\partial \mathbf{x}_k} = \prod_{t \geq i > k} \frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_{i-1}} = \prod_{t \geq i > k} \mathbf{W}_{rec}^T diag(\sigma'(\mathbf{x}_{i-1})) \quad (5)$$



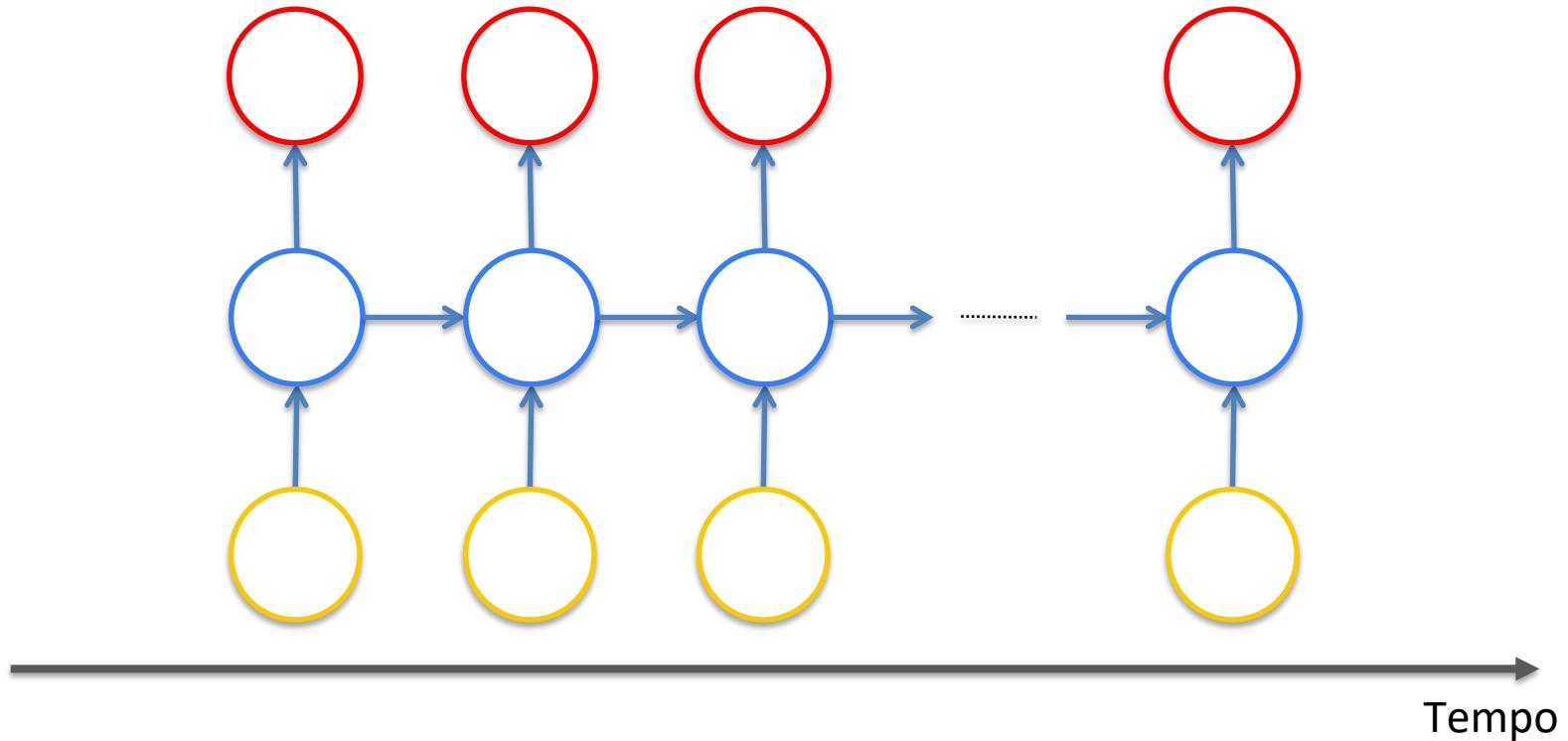
Fonte da fórmula: Razvan Pascanu et al. (2013)

Long Short-Term Memory

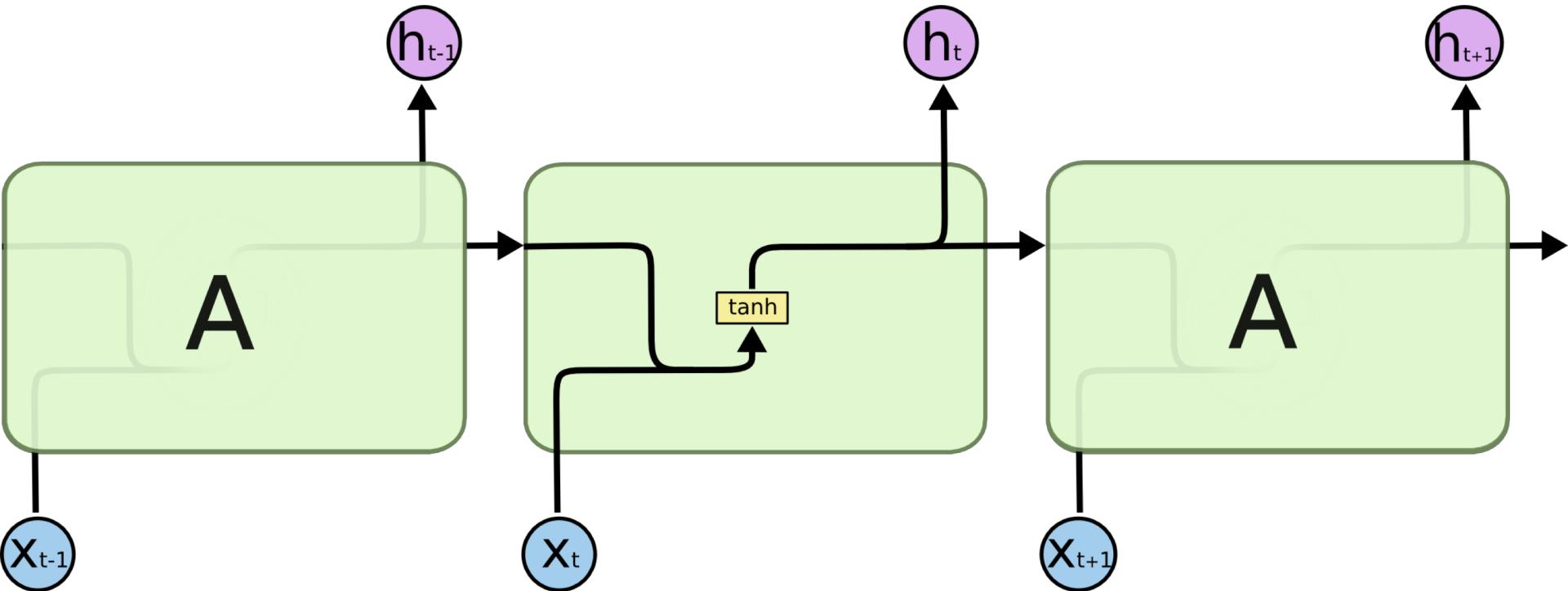


Fonte das imagens: people.idsia.ch, ics.usi.ch

Long Short-Term Memory

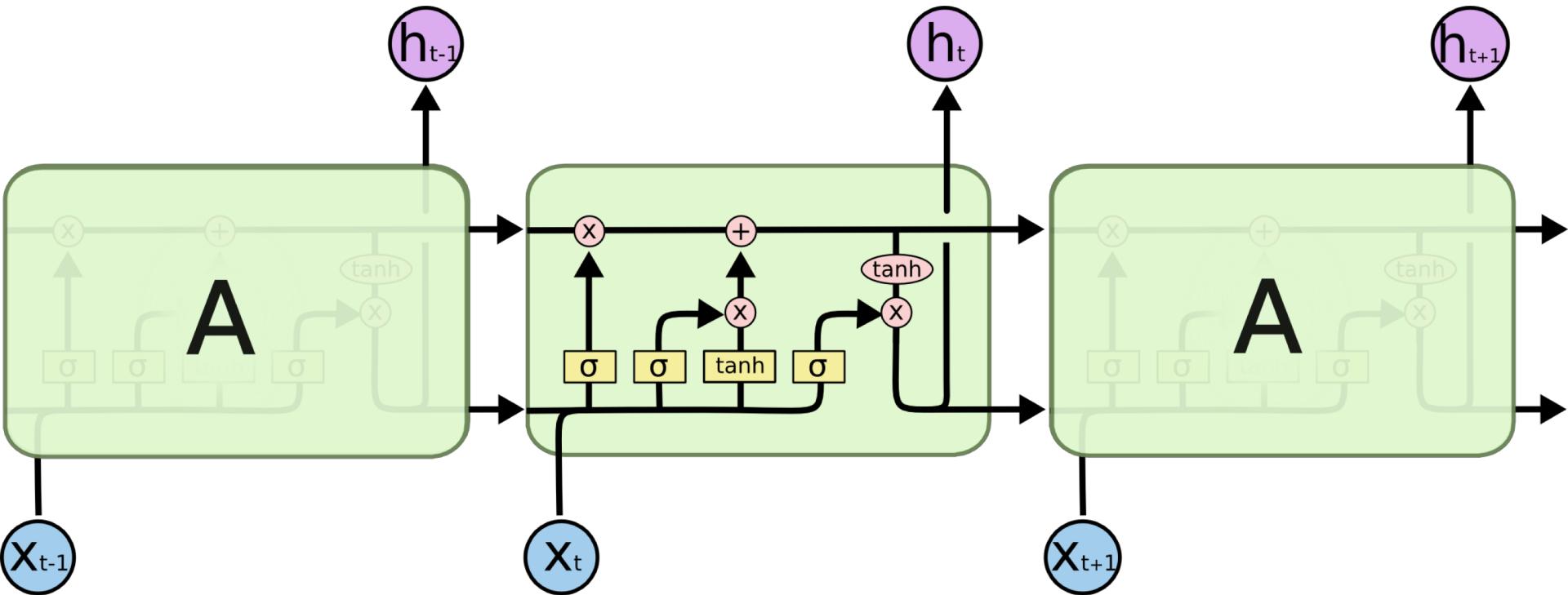


Long Short-Term Memory



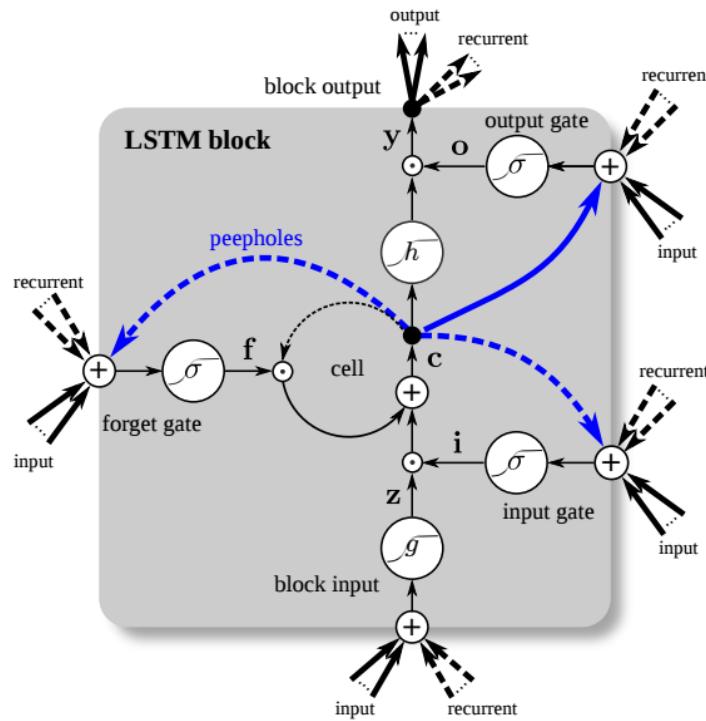
Fonte da imagem: colah.github.io

Long Short-Term Memory



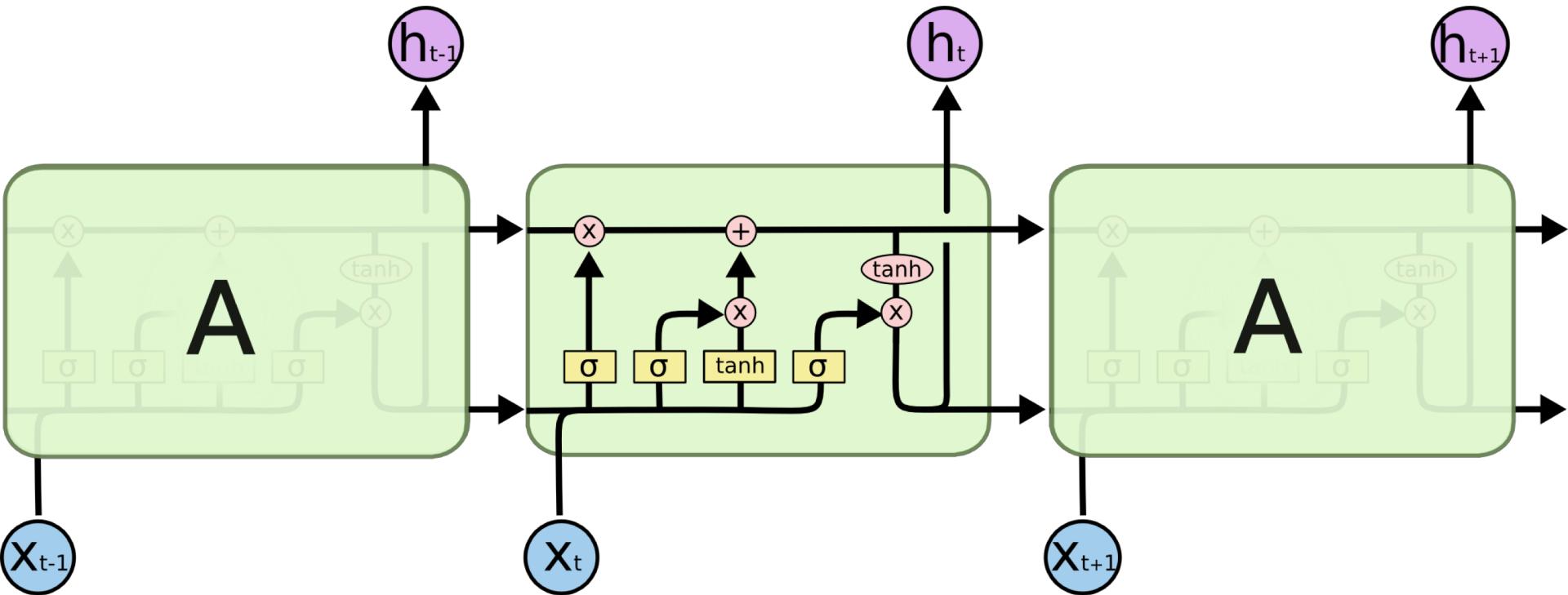
Fonte da imagem: colah.github.io

Long Short-Term Memory



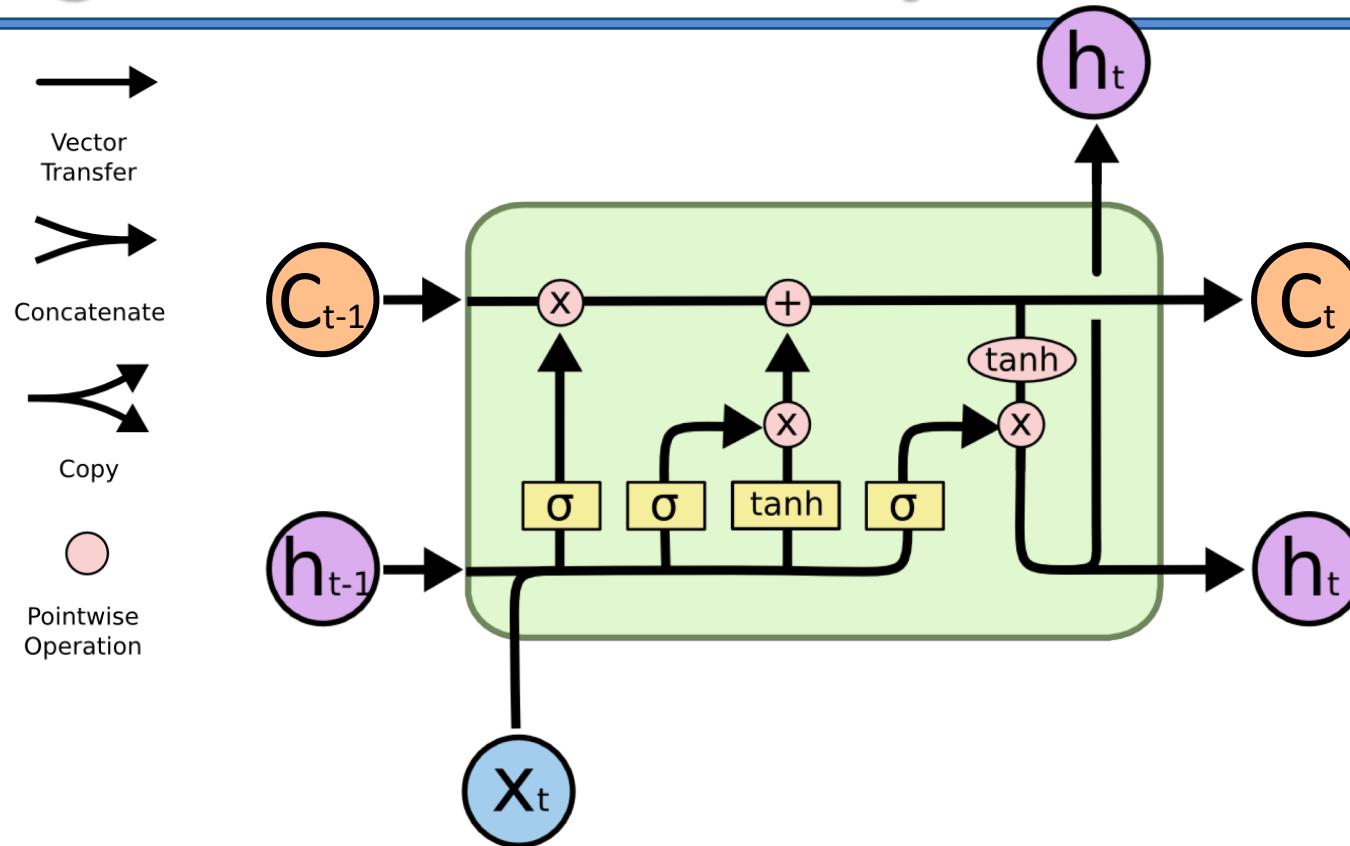
Fonte da imagem: arxiv.org/pdf/1503.04069.pdf

Long Short-Term Memory



Fonte da imagem: colah.github.io

Long Short-Term Memory

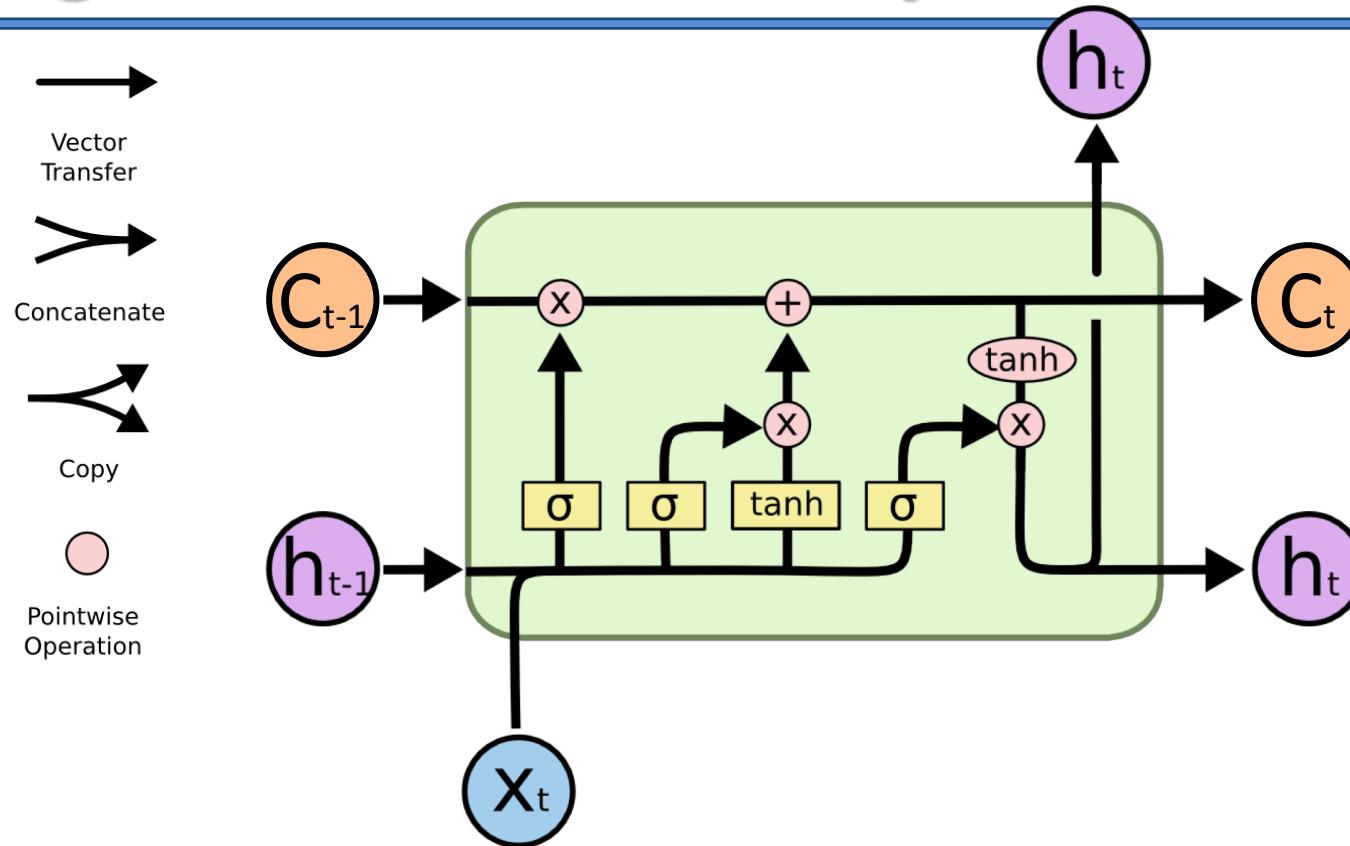


Fonte da imagem: colah.github.io

Long Short-Term Memory



Long Short-Term Memory

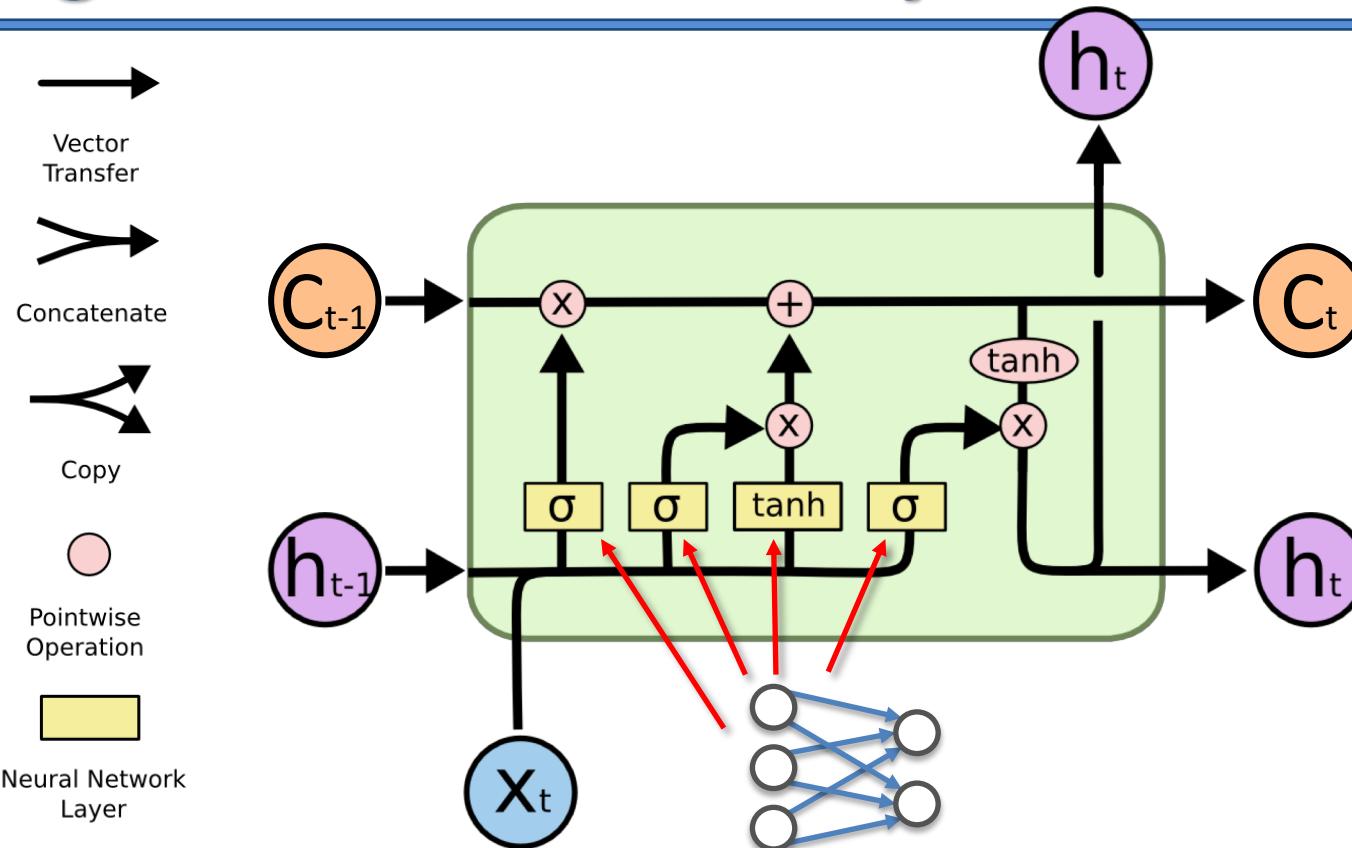


Fonte da imagem: colah.github.io

Long Short-Term Memory

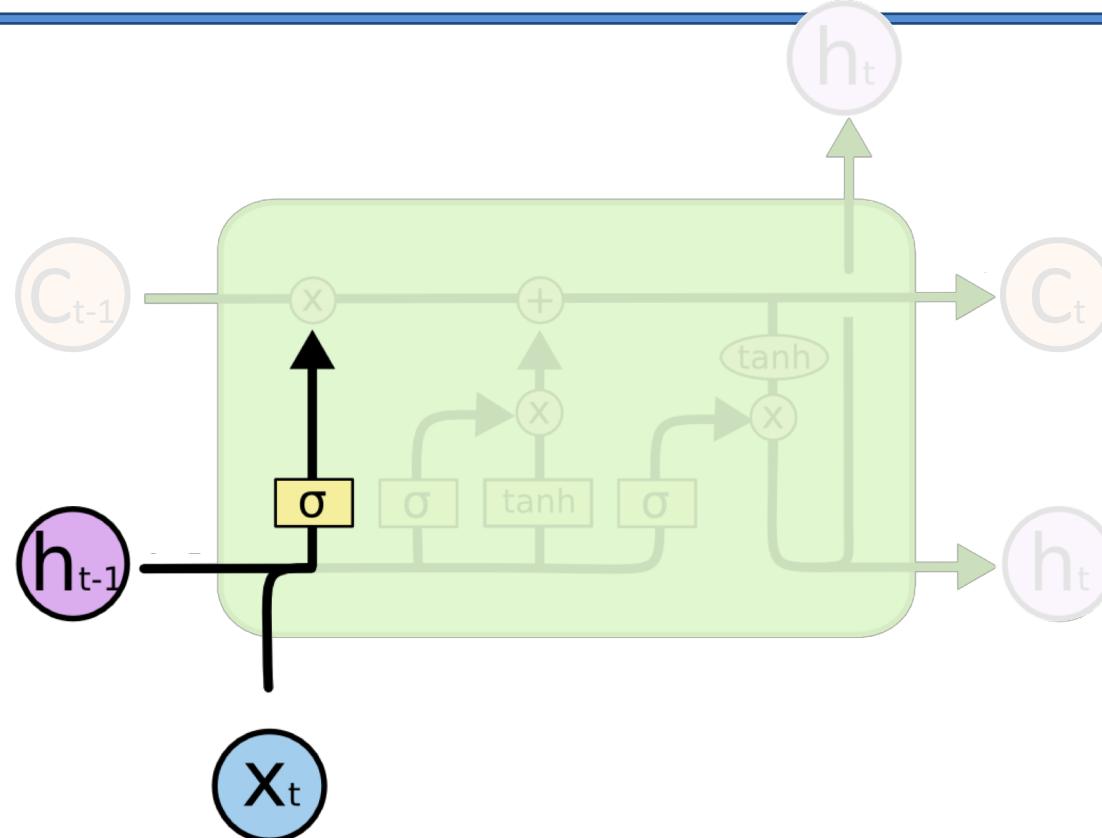


Long Short-Term Memory



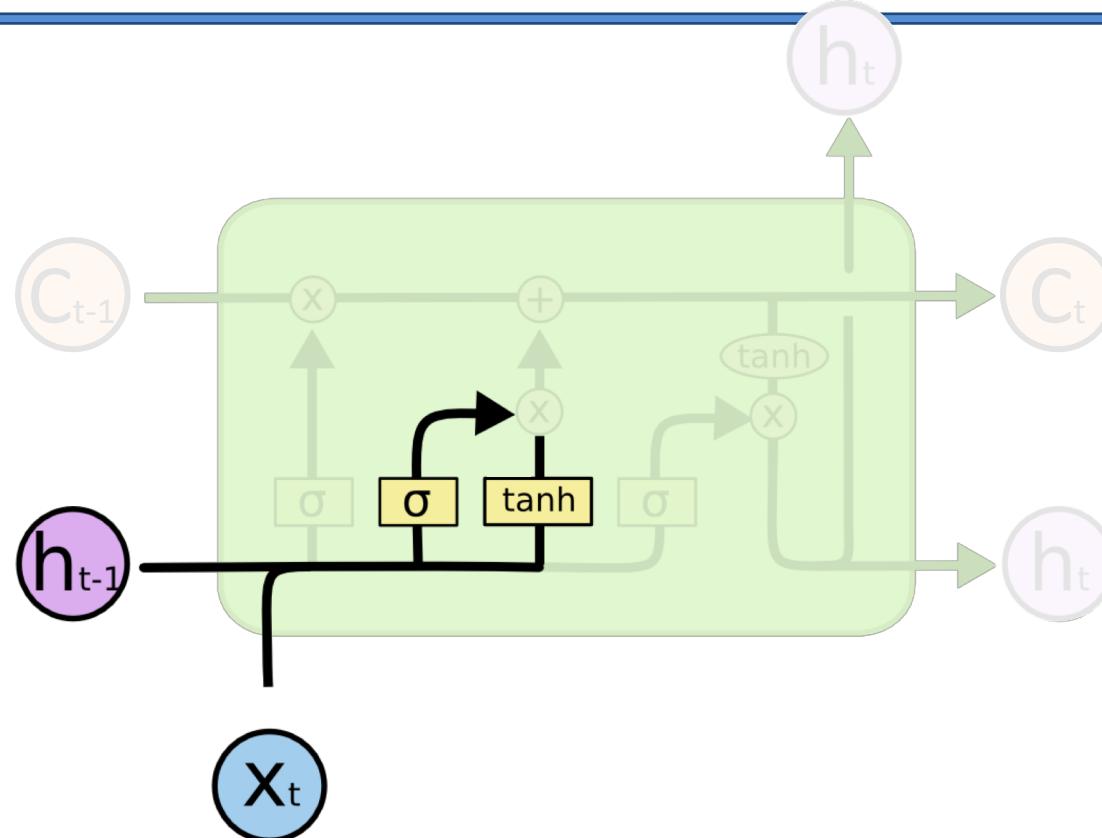
Fonte da imagem: colah.github.io

Long Short-Term Memory



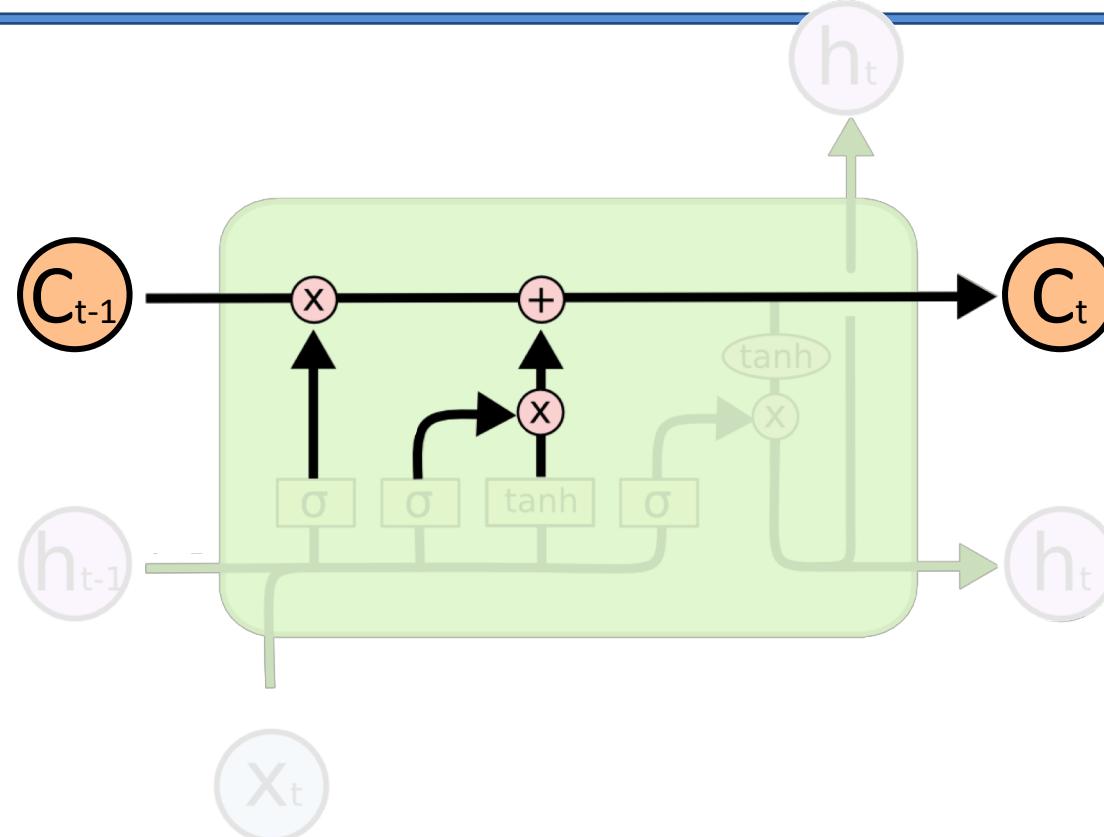
Fonte da imagem: colah.github.io

Long Short-Term Memory



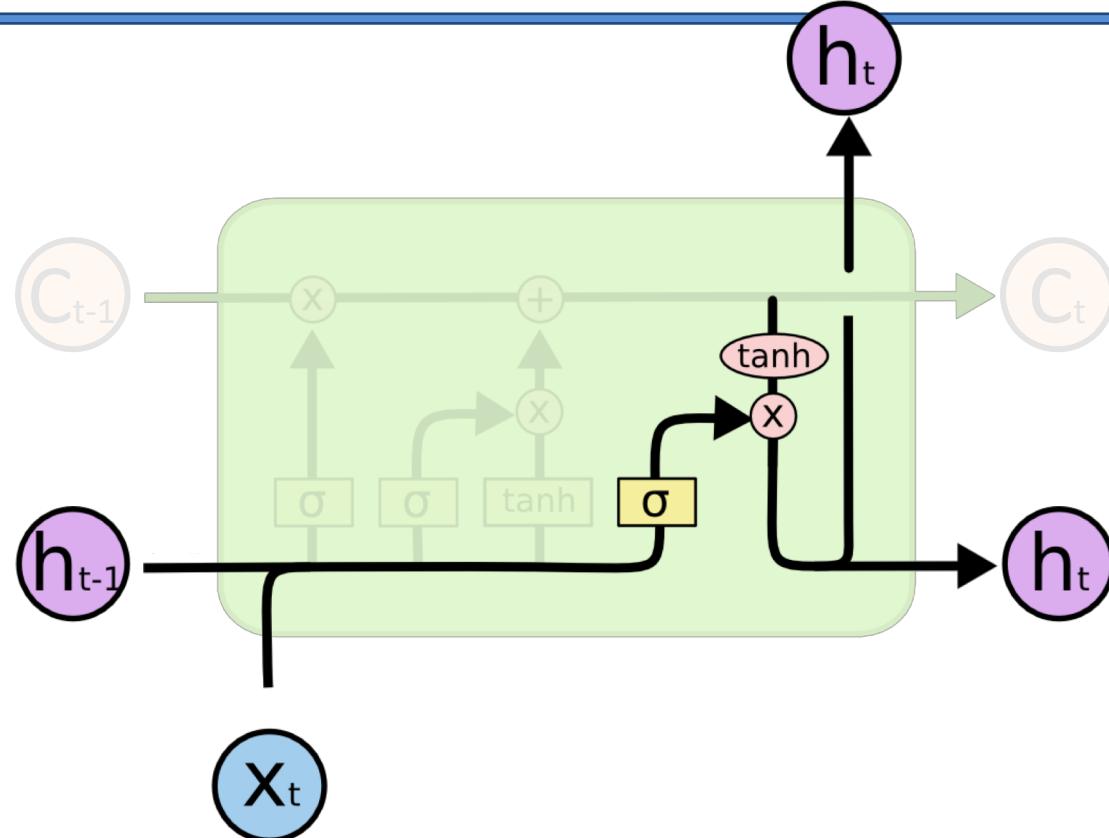
Fonte da imagem: colah.github.io

Long Short-Term Memory



Fonte da imagem: colah.github.io

Long Short-Term Memory



Fonte da imagem: colah.github.io

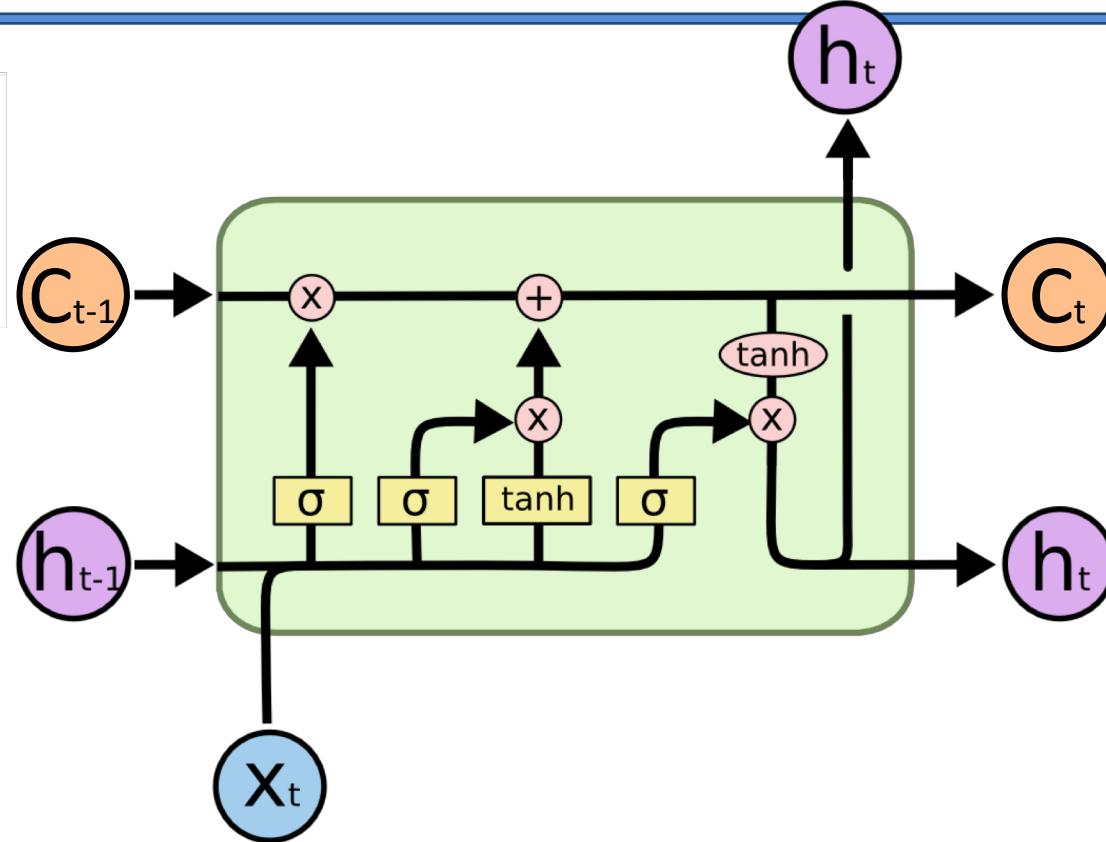
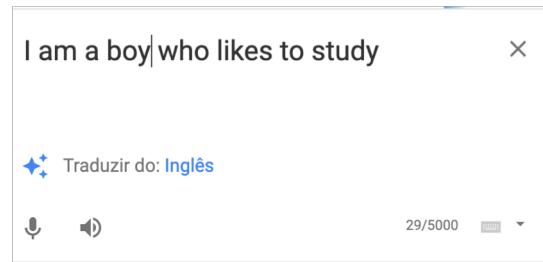
Long Short-Term Memory

DETECTAR IDIOMA INGLÊS PORTUGUÊS ▾ ↔ PORTUGUÊS INGLÊS ESPANHOL ▾

I am a boy who likes to study × Eu sou um garoto que gosta de estudar

Traduzir do: Inglês

Long Short-Term Memory



Long Short-Term Memory

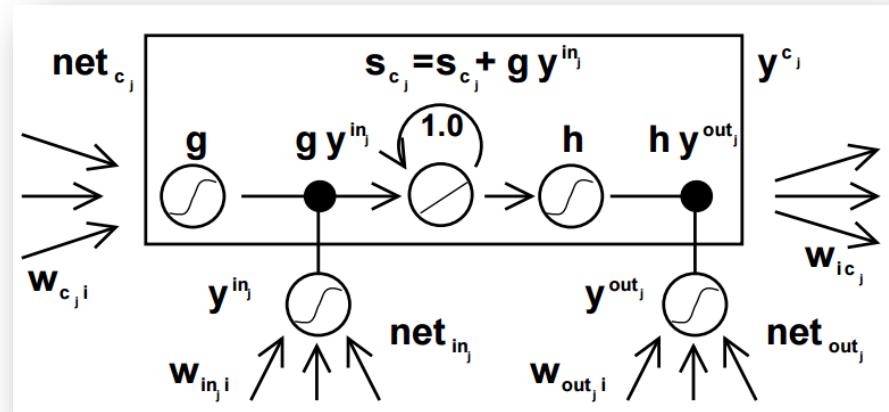
Leitura adicional:

Long Short-Term Memory

Sepp Hochreiter &
Jurgen Schmidhuber (1997)

Link:

<http://www.bioinf.jku.at/publications/older/2604.pdf>

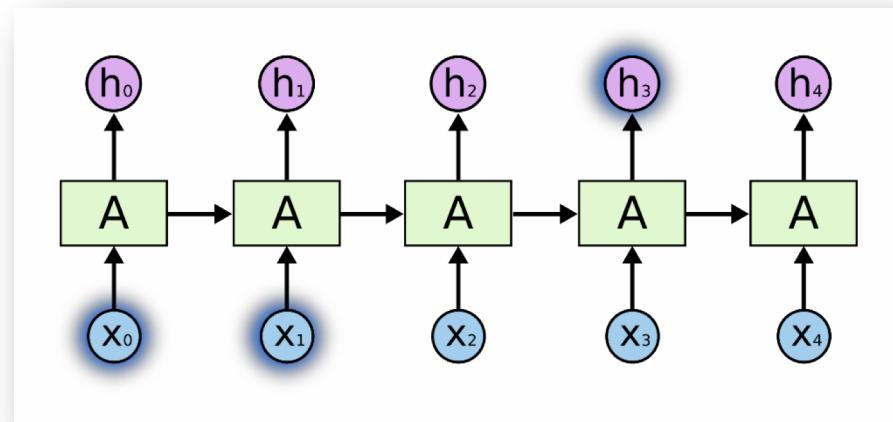


Long Short-Term Memory

Leitura adicional:

Understanding LSTM Networks

Christopher Olah (2015)



Link:

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Long Short-Term Memory

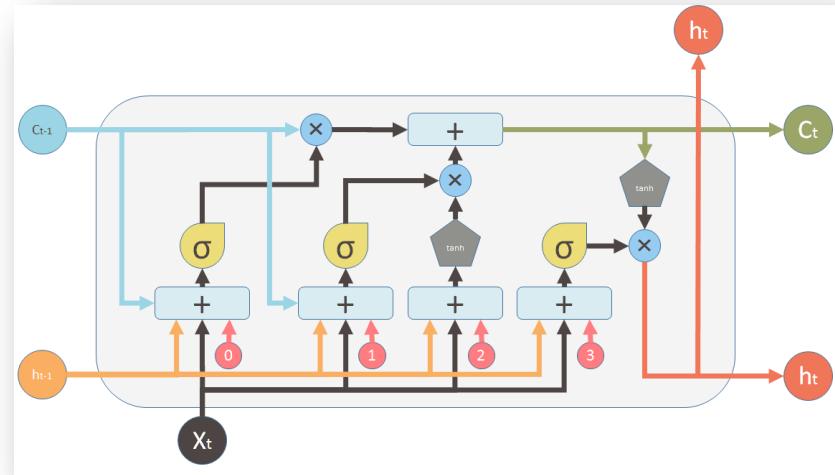
Leitura adicional:

Understanding LSTM and its diagrams

Shi Yan (2016)

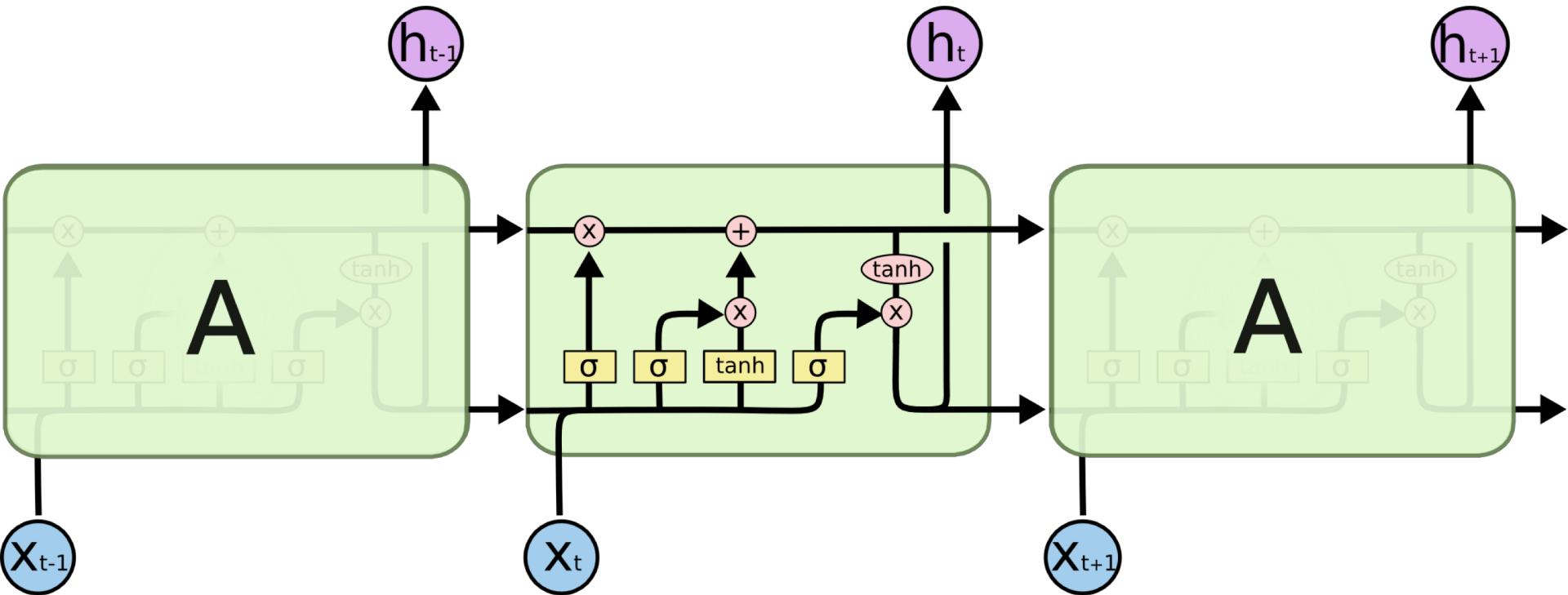
Link:

<https://medium.com/@shiyans/understanding-lstm-and-its-diagrams-37e2f46f1714>



LSTM – Intuição Prática

Long Short-Term Memory



Fonte da imagem: colah.github.io

Long Short-Term Memory

Cell sensitive to position in line:

The sole importance of the crossing of the Berezina lies in the fact that it plainly and indubitably proved the fallacy of all the plans for cutting off the enemy's retreat and the soundness of the only possible line of action--the one Kutuzov and the general mass of the army demanded--namely, simply to follow the enemy up. The French crowd fled at a continually increasing speed and all its energy was directed to reaching its goal. It fled like a wounded animal and it was impossible to block its path. This was shown not so much by the arrangements it made for crossing as by what took place at the bridges. When the bridges broke down, unarmed soldiers, people from Moscow and women with children who were with the French transport, all--carried on by vis inertiae--pressed forward into boats and into the ice-covered water and did not, surrender.

Cell that turns on inside quotes:

"You mean to imply that I have nothing to eat out of.... On the contrary, I can supply you with everything even if you want to give dinner parties," warmly replied Chichagov, who tried by every word he spoke to prove his own rectitude and therefore imagined Kutuzov to be animated by the same desire.

Kutuzov, shrugging his shoulders, replied with his subtle penetrating smile: "I meant merely to say what I said."

Fonte da imagem: karpathy.github.io

Long Short-Term Memory

Cell that robustly activates inside if statements:

```
static int __dequeue_signal(struct sigpending *pending, sigset_t *mask,
                           siginfo_t *info)
{
    int sig = next_signal(pending, mask);
    if (sig) {
        if (current->notifier) {
            if (sigismember(current->notifier_mask, sig)) {
                if (!(current->notifier)(current->notifier_data)) {
                    clear_thread_flag(TIF_SIGPENDING);
                    return 0;
                }
            }
        }
        collect_signal(sig, pending, info);
    }
    return sig;
}
```

Cell that is sensitive to the depth of an expression:

```
#ifdef CONFIG_AUDITSYSCALL
static inline int audit_match_class_bits(int class, u32 *mask)
{
    int i;
    if (classes[class]) {
        for (i = 0; i < AUDIT_BITMASK_SIZE; i++)
            if (mask[i] & classes[class][i])
                return 0;
    }
    return 1;
}
```

Fonte da imagem: karpathy.github.io

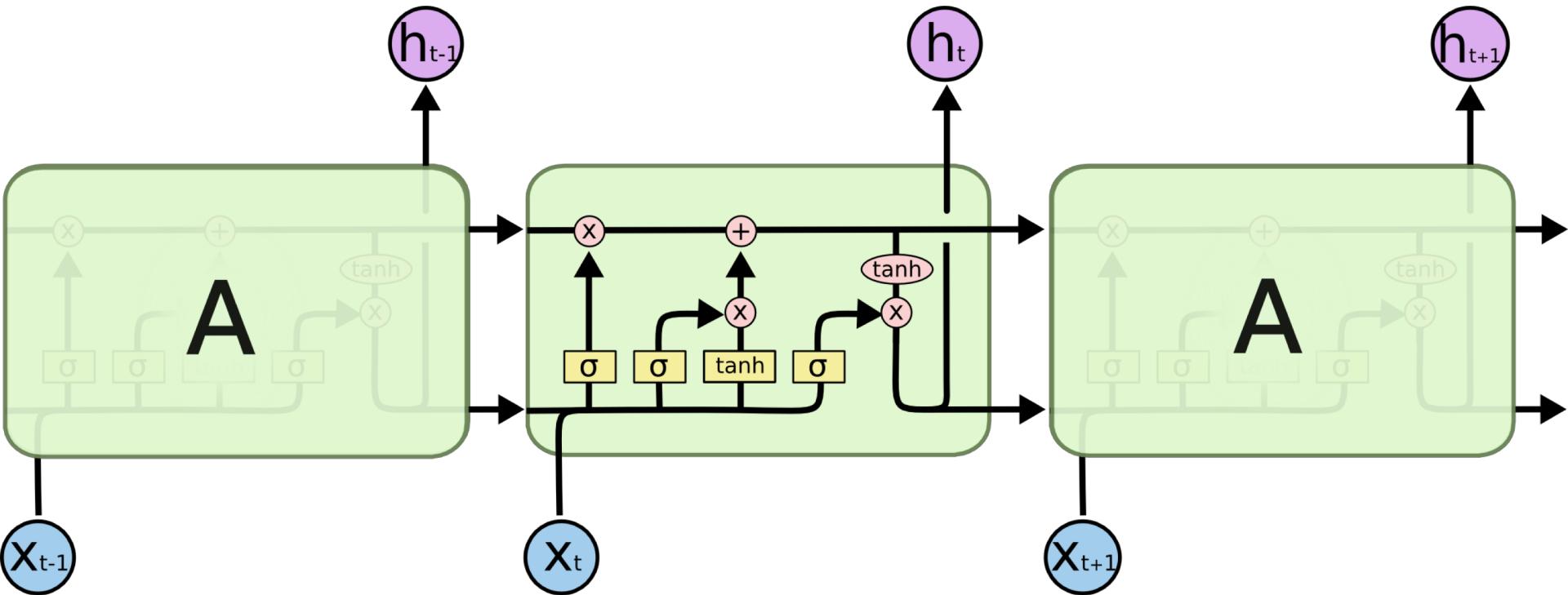
Long Short-Term Memory

A large portion of cells are not easily interpretable. Here is a typical example:

```
/* Unpack a filter field's string representation from user-space
 * buffer. */
char *audit_unpack_string(void **bufp, size_t *remain, size_t len)
{
    char *str;
    if (!*bufp || (len == 0) || (len > *remain))
        return ERR_PTR(-EINVAL);
    /* Of the currently implemented string fields, PATH_MAX
     * defines the longest valid length.
    */
```

Fonte da imagem: karpathy.github.io

Long Short-Term Memory



Fonte da imagem: colah.github.io

t t p : / / w w w . y n e t n e w s . c o m /] E n g l i s h - l a n g u a g e w e b s i t e o f I s r a e l
t p : / / w w w . b a c a h e t s . c o m / - x g l i s h l i n g u a g e s a i r s i t e o f t s l a e l
d : x n e . w a e a . a w a t o a . s & n t i a c a - s a r d e e l h o a n t b i s a n f a n r e i f
m w - 2 ♦ p i i i s o e s s i s . / e r n . c] (d c e e n e p e s a a i k i i e e l e d h , i r t h r a o n s
d r . < : a h b - n p t w t . x i g h / m a) T v d r y z i c o u e d l s u : t h a - o o t u , s t u i f
s t p , t c o a 2 d r u l w o c l e n s r] p . l l v a o d , , e y t c - n d m - o i b u v s] b b i m s u l t

g e s t n e w s p a p e r ' ' [[Y e d i o t h A h r o n o t h]] ' ' ' ' H e b r e w - l a n g u a g e
e t a a w s p a p e r s o ' [[T e l t i (f e a n e m t i) ' ' * ' ' [e r r e w s l e n g u a g e :
i r s c o e e n a i T T h A o a i n h S r m u w] e y s [' i n e i a ' s i w d d e ' h s
u s . s e t l g o r s . a s a t C a r e e g ' a C l r i s z] i e ' : : , # : T A a a a a t B a s e e i l o '
- t u a e v r t i d , t B A m S u s y u t]] A s a o i g s] , . . : s M B o l o u s : T o u a - n : d
a , d , i i u i t i c p .] (I S v H v t u s u i e D n o e g a n o . ,] : { C C u i b o h e C y b k s l s : r

i c a l s : ' ' * ' [[G l o b e s]] ' ' [h t t p : / / w w w . g l o b e s . c o . i l /] b u s i
c a l : ' ' * ' [T a a b a] ' ' ([t t p : / / w w w . b u o b a l . c o m u n / s A - y t i r
s t l ' [h A e o v e l t s a h a d : x g e . w a o i r . r t o a . e l . i t & a i e
t t ' ' ' & [& & m C o e r o n e ' : : , i ' o d w . , : n i i i s a a u e . e n i / o m l c C . (e f t g
a ' n : , C : & : # * : a f D r u s u l l , . o m e l p < , d h a ; d e u o o t / i h n c s i f S , u r h
n k i <] : & 1 1 s T G u i t r s i , : b a c m r - x t p o b - g r e s i s l e r l n a f a D] l o s p t

i l y * ' ' [[H a a r e t z | H a ' A r e t z]] ' ' [h t t p : / / w w w . h a a r e t z . c o . i l /]
l y * ' ' [[T e r r d n F e r a n t a h]] ' ' ([t t p : / / w w w . b o n m d s t . c o m u n / s -
r e ' ' h A i l n n t e H a l s r c n o l ' s a h a d : x n e . w a a m r t d h e o h o l c 8
k i . : * s C O S a n l t h i T i m ' l i] e : , i m c d w - 2 ♦ p h i F o n t e d a i m a g e : k a r p a t h y . g i t h u b . i o
d s - ! [t B T C o m m g d] W o n a a e , : . b a e r r . < t a i b - d u l c n n c / a r n e s i l
n d s # & : G l D u v c c s a o S u c l t e l l z l . : o ' o m t l . : e o a 2 n i v f s r o o e i u n a l a

Leitura Adicional

Leitura adicional:

The Unreasonable Effectiveness of Recurrent Neural Networks

Andrej Karpathy (2015)

Link:

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

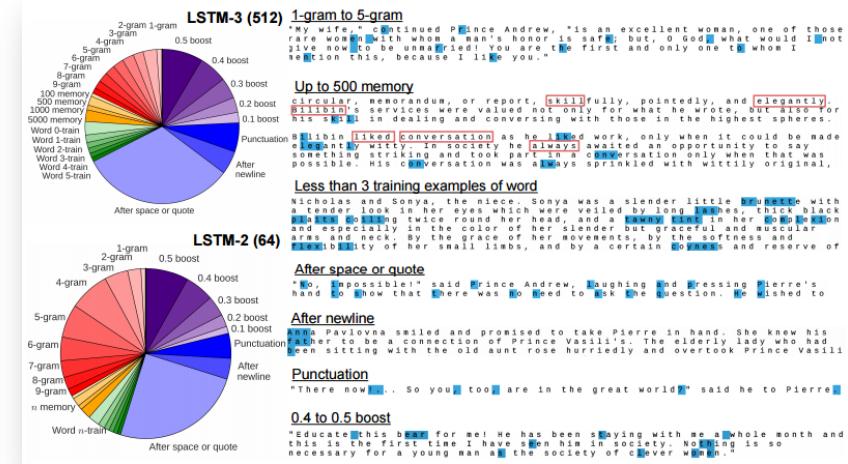
A neural network diagram illustrating the generation of a Hebrew sentence from an English input. The input "English-language websites of Israel's lar" is processed by an RNN layer, which generates the output "הוֹרְכָּדָה שֶׁבְּנֵי יִשְׂרָאֵל מִתְּבִּינָה". The diagram shows the flow of information through various layers, including embedding, hidden states, and a final layer that maps to Hebrew characters.

Leitura Adicional

Leitura adicional:

Visualizing and Understanding Recurrent Networks

Andrej Karpathy et al. (2015)

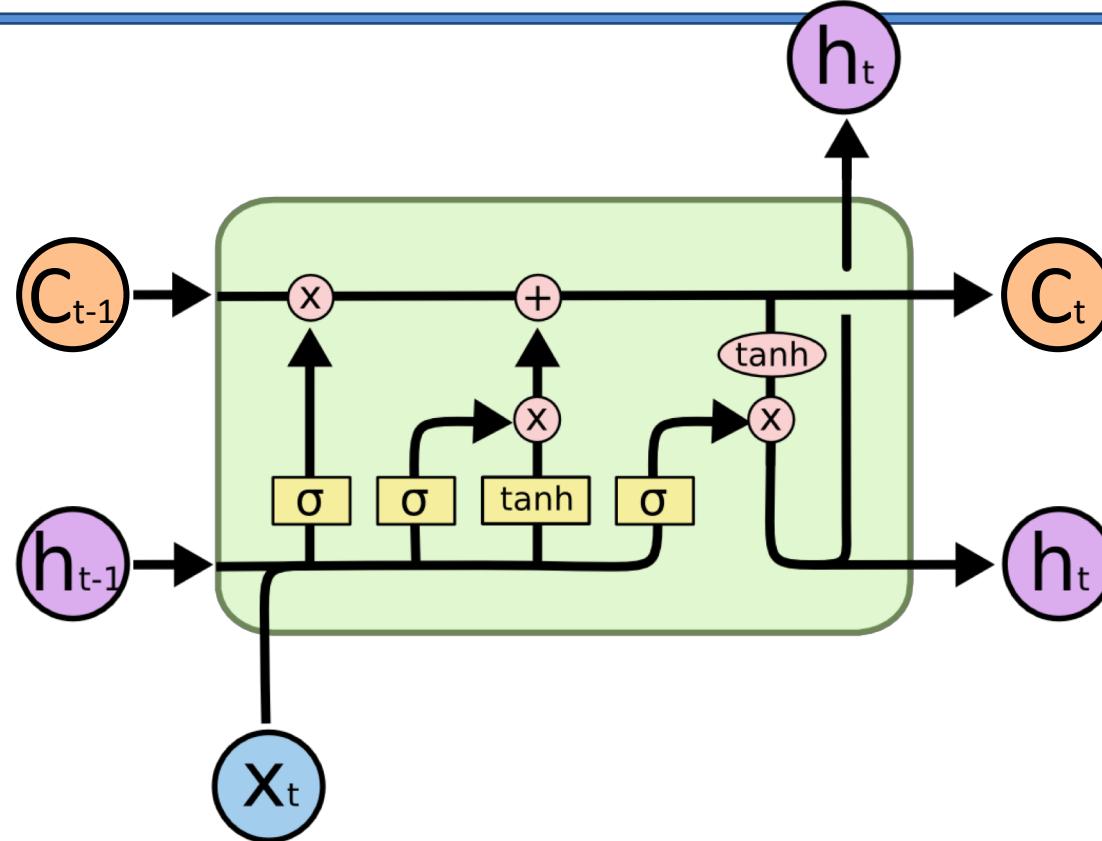


Link:

<https://arxiv.org/pdf/1506.02078.pdf>

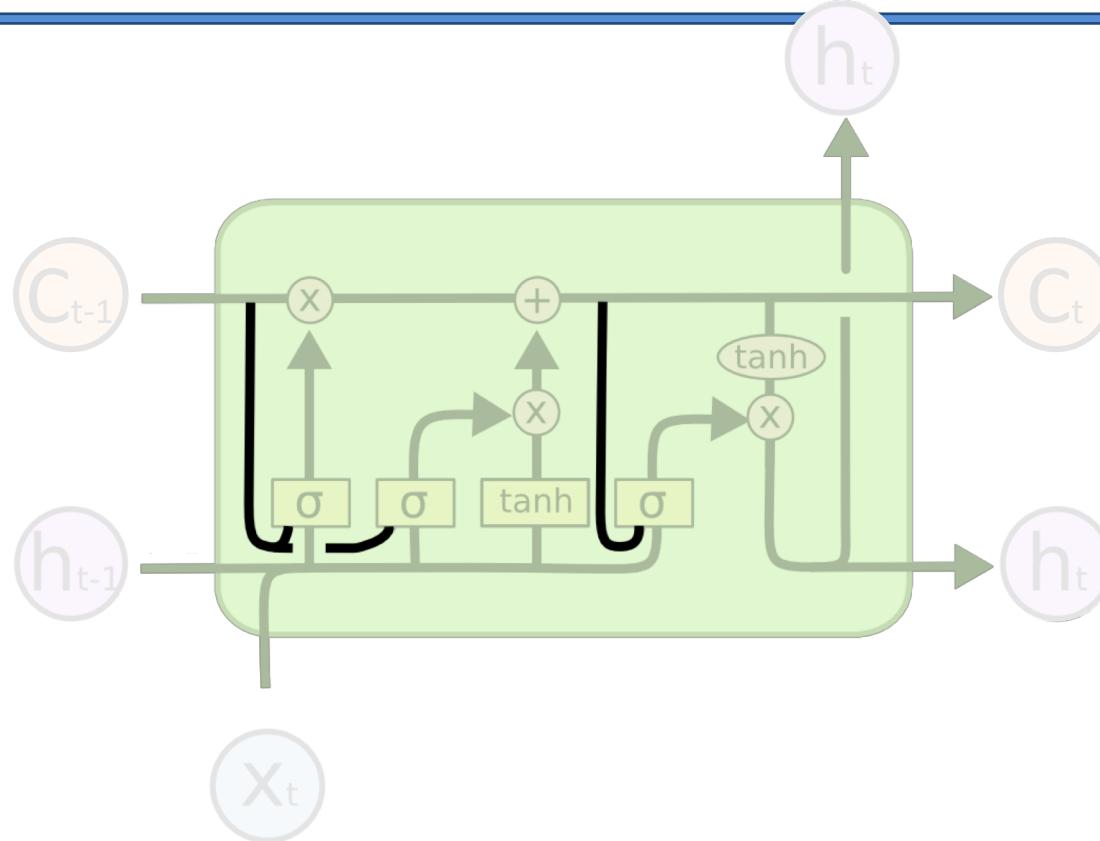
LSTM – Variações

LSTM – Variações



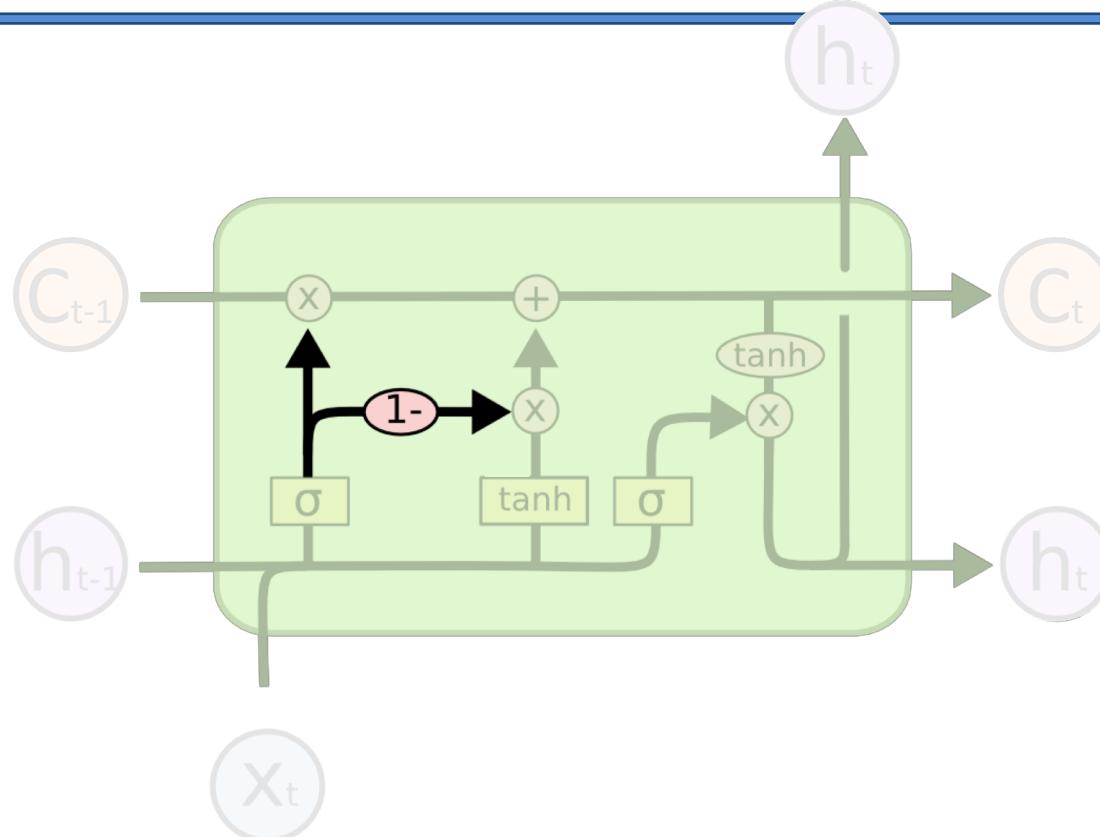
Fonte da imagem: colah.github.io

LSTM – Variações



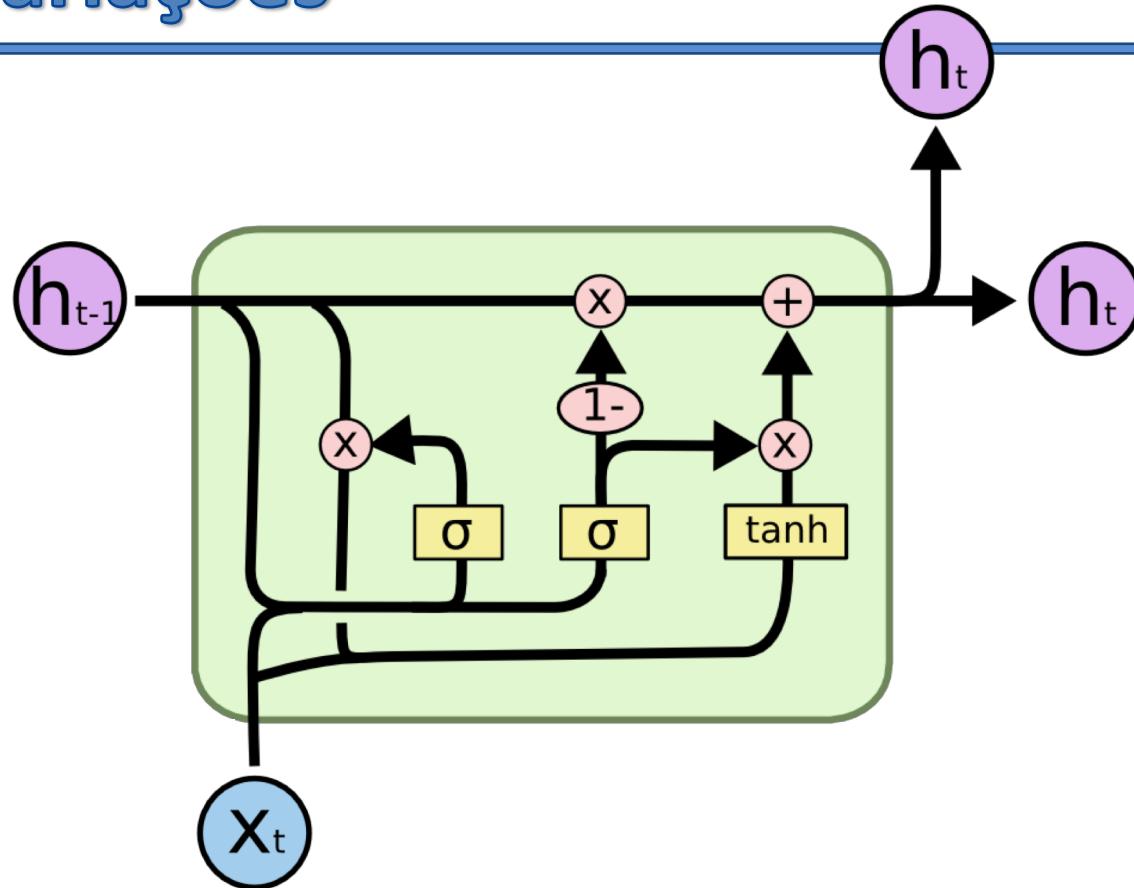
Fonte da imagem: colah.github.io

LSTM – Variações



Fonte da imagem: colah.github.io

LSTM – Variações



Long Short-Term Memory

Leitura adicional

LSTM: A Search Space Odyssey

Klaus Greff et al. (2015)

Link:

<https://arxiv.org/pdf/1503.04069.pdf>

