Deep Learning

Redes Neurais Recorrentes

Aula 7

Bom dia!

Eu sou Diego Dorgam

Alguma pergunta que você quer fazer?!

http://bit.ly/dl-unb7
https://t.me/DeepLearningUnB
@diegodorgam

Instruções Gerais

- ✓ Traga seu laptop
- ✓ Use Software Livre
- ✓ Não converse por voz
- ✓ Se não entender, pergunte!
- ✓ Se entender, explique!
- ✓ DIVIRTA-SE

O que Aprendemos na aula passada?

1. Intuição
    - O que são CNNs
    - Convolution Operation
    - Pooling
    - Flattening
    - Full Connection
2. Prática
    - Reconhecimento de imagens

O que Vamos Aprender HOJE?

1. Intuição
    - Conceito de RNNs
    - Vanishing Gradient Problem
    - LSTMs and Variations
2. Prática
    - Building RNN

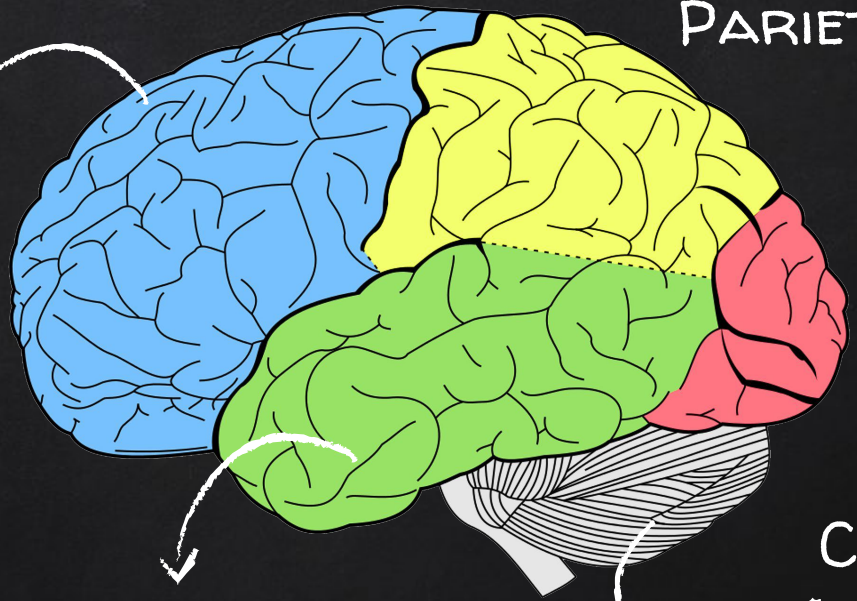
O que são RNNs

Redes Neurais Recorrentes

Onde Estamos

Supervisionado	Artificial Neural Networks	Regression & Classification
	Convolutional Neural Networks	Computer Vision
	Recurrent Neural Networks	Time Series Analysis
Não Supervisionado	Self-Organizing Maps	Feature Detection
	Deep Boltzmann Machines	Recommendation Systems
	AutoEncoders	Recommendation Systems

O CÉREBRO



FRONTAL LOBE

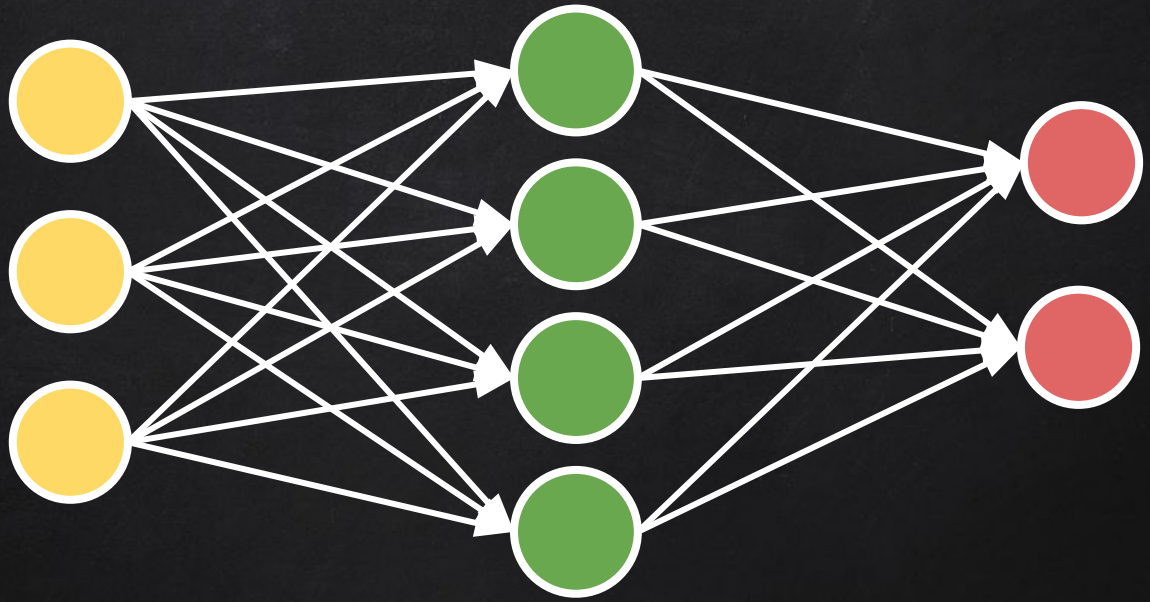
PARIETAL LOBE

OCCIPITAL LOBE

TEMPORAL LOBE

CEREBELLUM

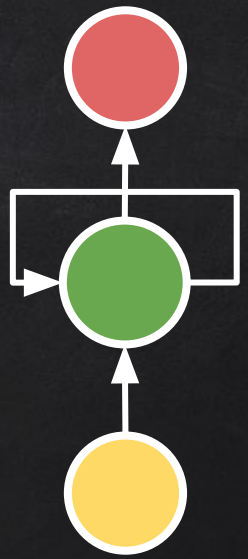
RNN Representation



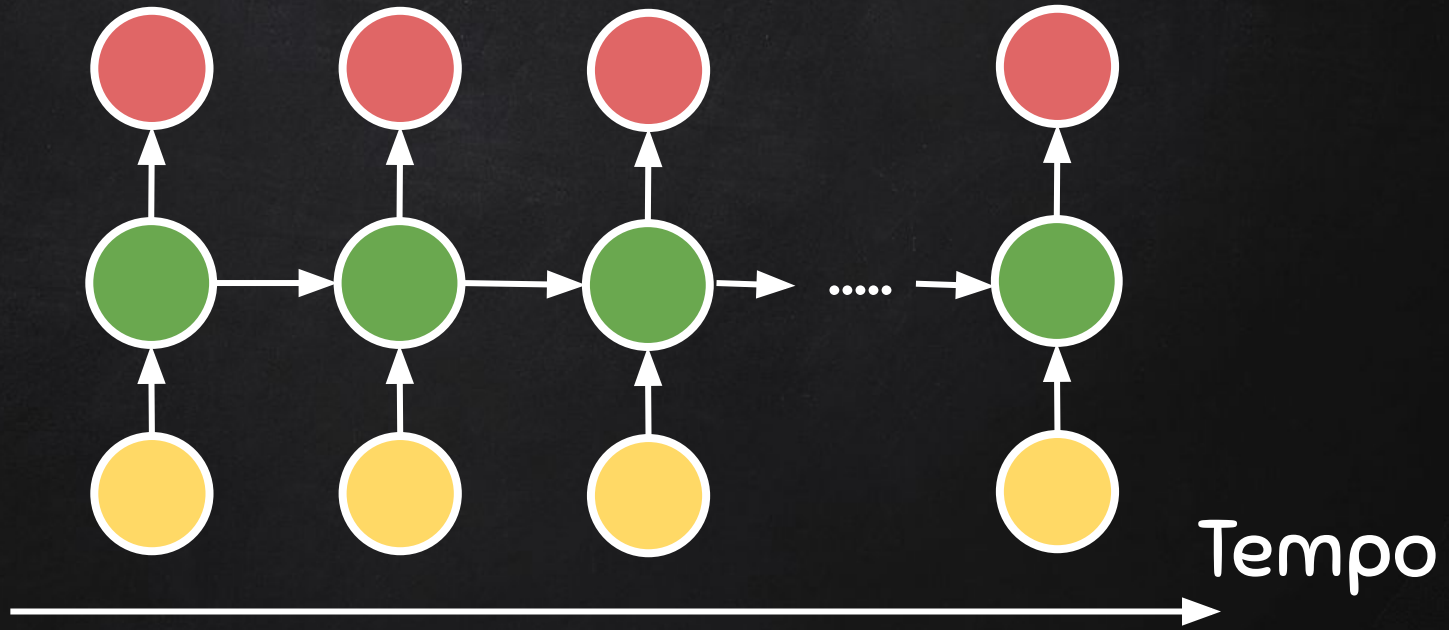
RNN Representation

RNN Representation

RNN Representation



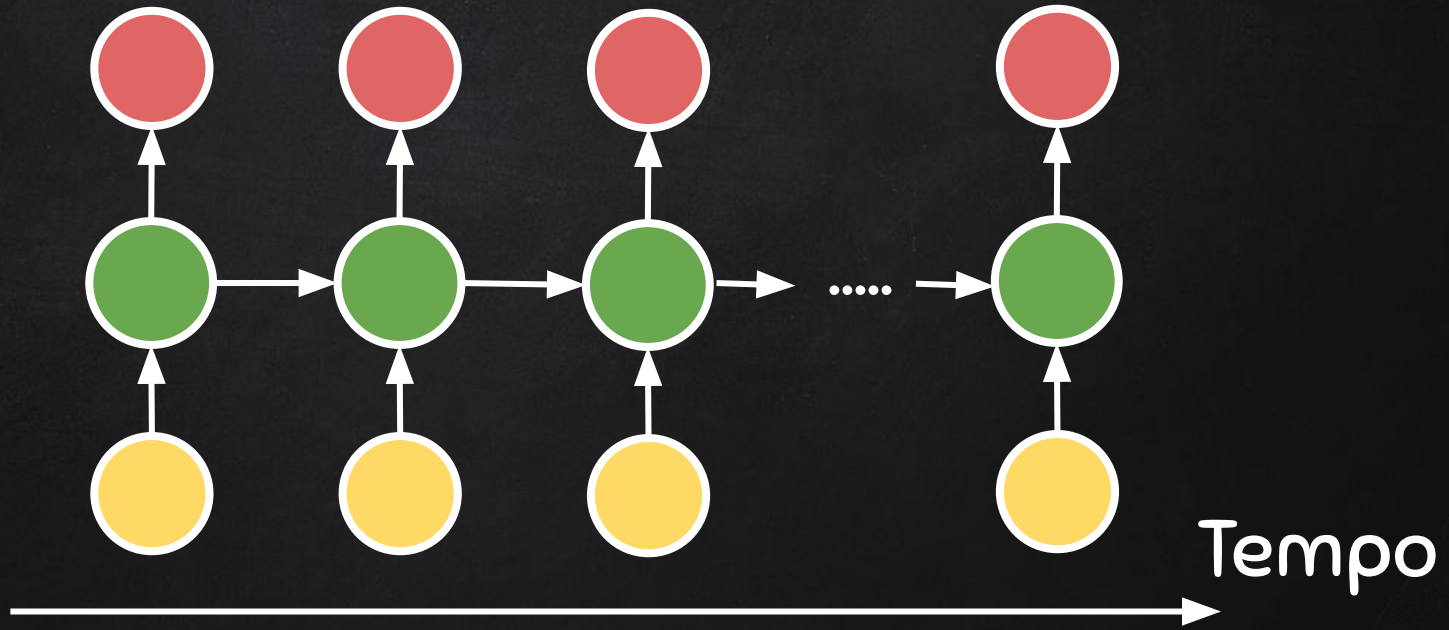
RNN Representation



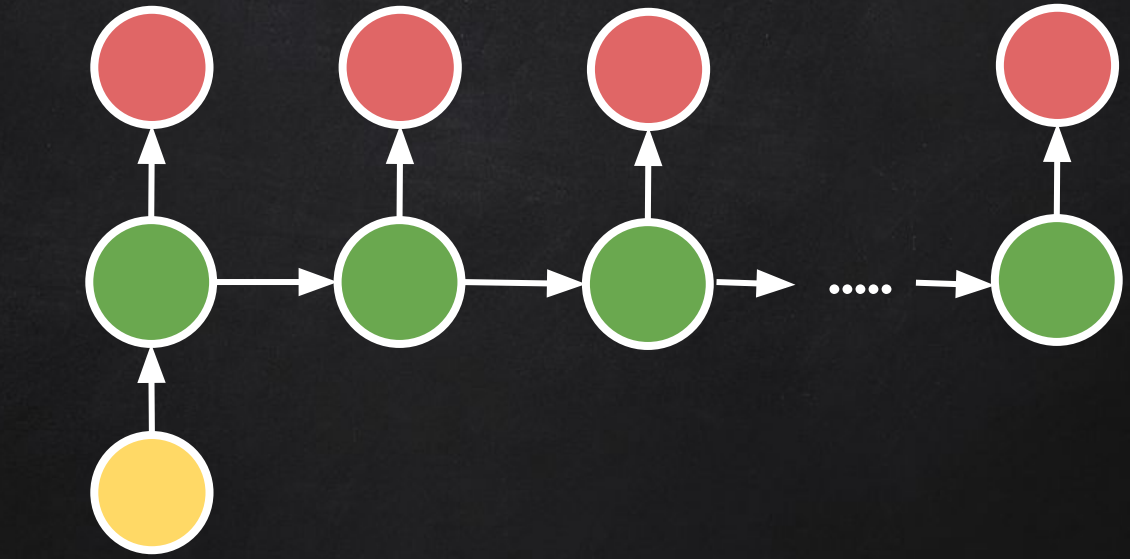
RNN Representation

Tempo

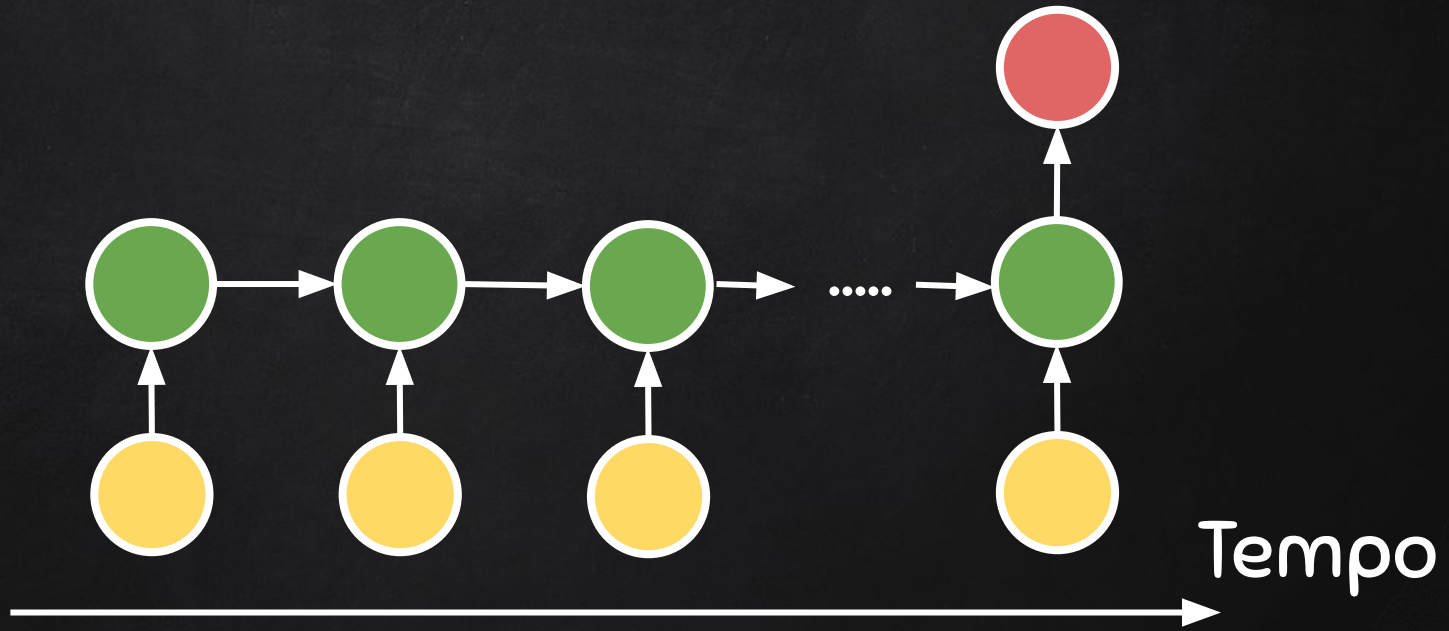
RNN Representation



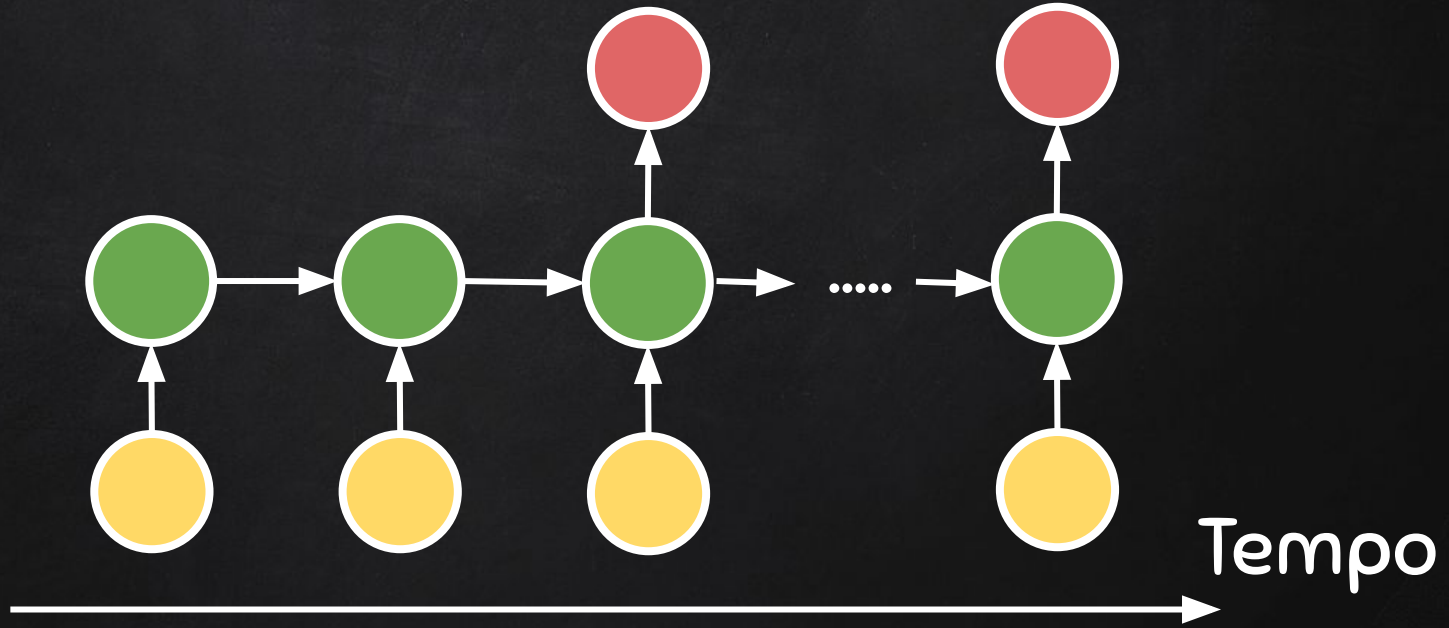
One to Many



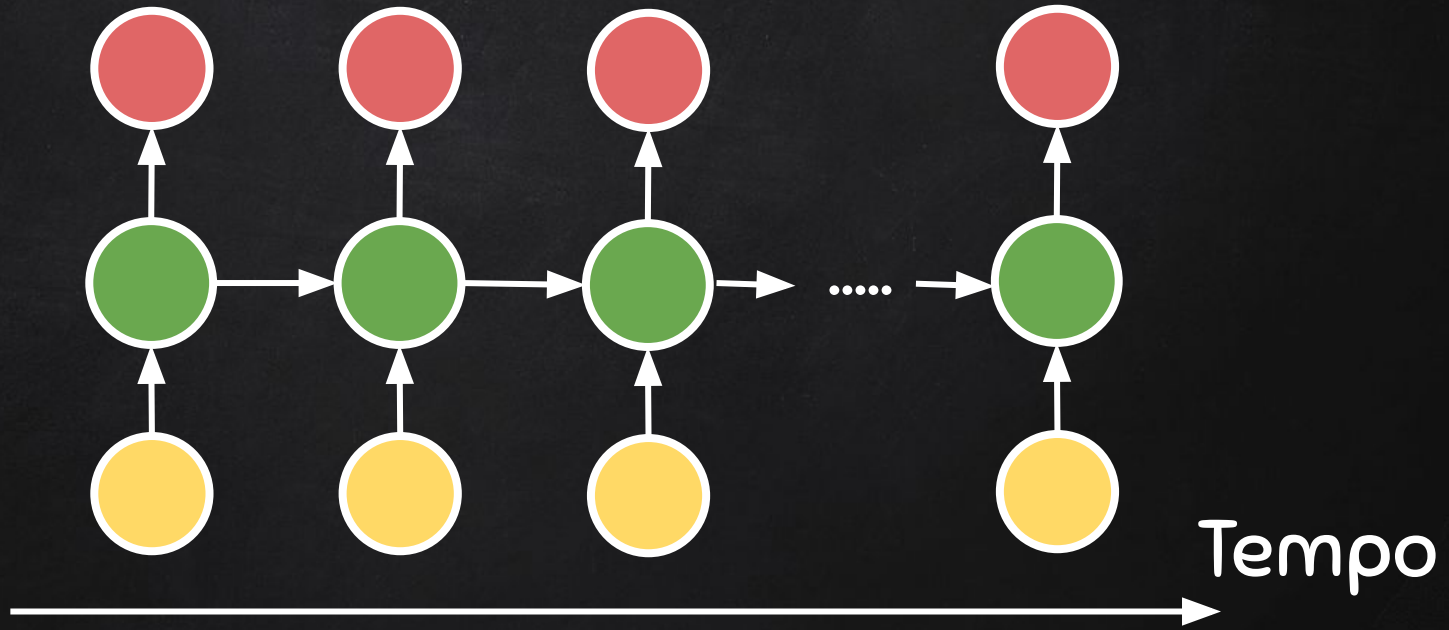
Many to One



Many to Many



Many to Many



Sunspring

A Sci-Fi Short Film Starring Thomas Middleditch

https://youtu.be/LY7x2Ihqjmc

It's No Game

A Sci-fi Short Film Starring David Hasselhoff

https://youtu.be/5qPgG98_CQ8

Vanishing Gradient Problem

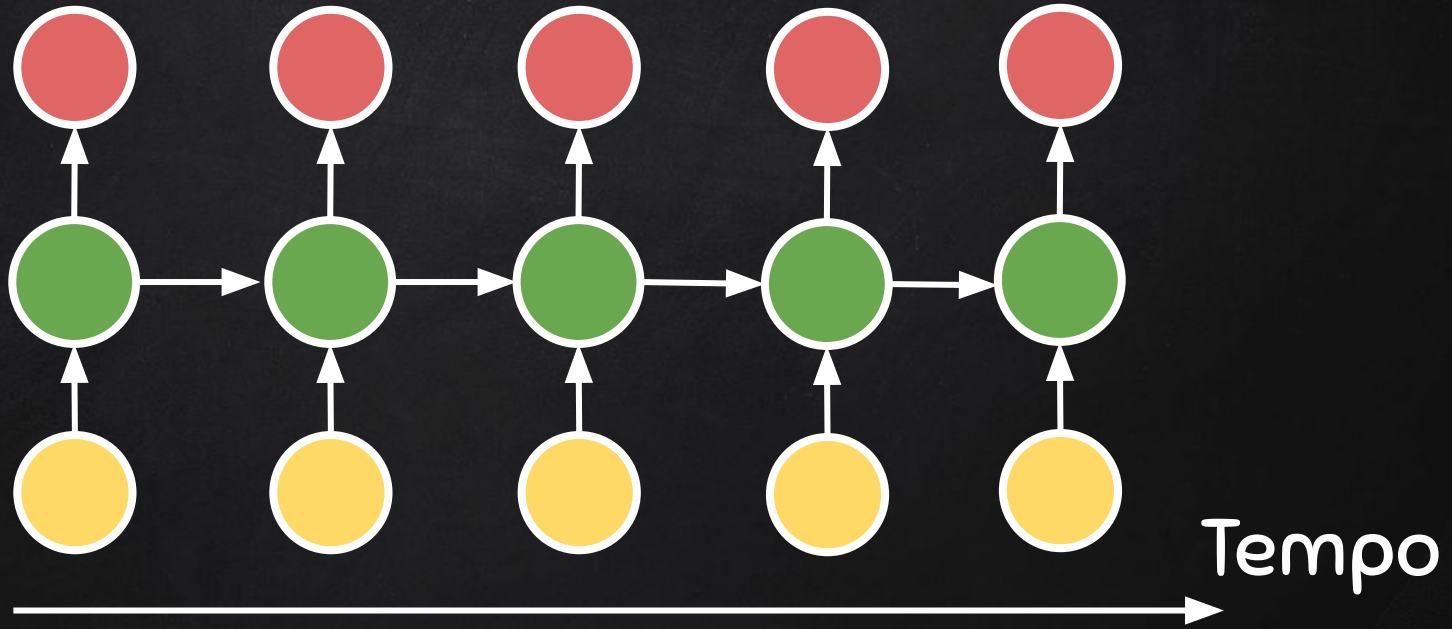
Redes Neurais Recorrentes

Sepp Hochreiter

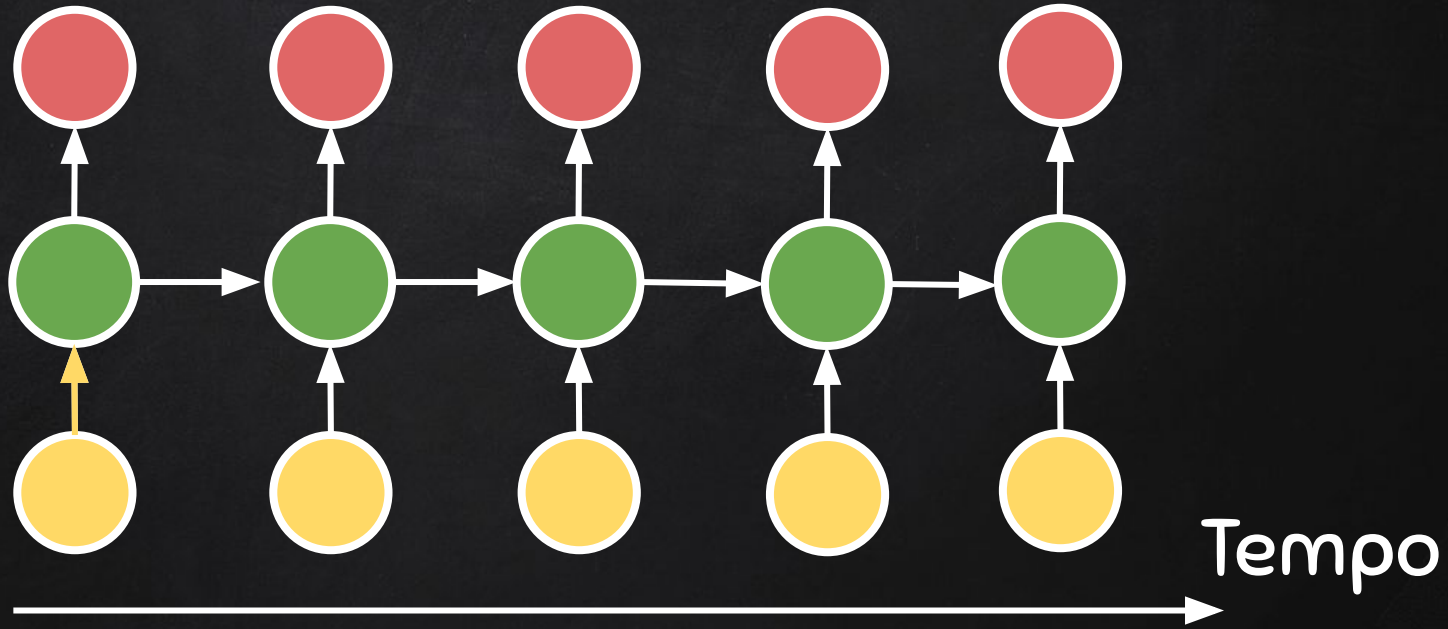
1991 Sepp Hochreiter's Analysis of fundamental problem of deep learning

http://people.idsia.ch/~juergen/fundamentaldeeplearningproblem.html

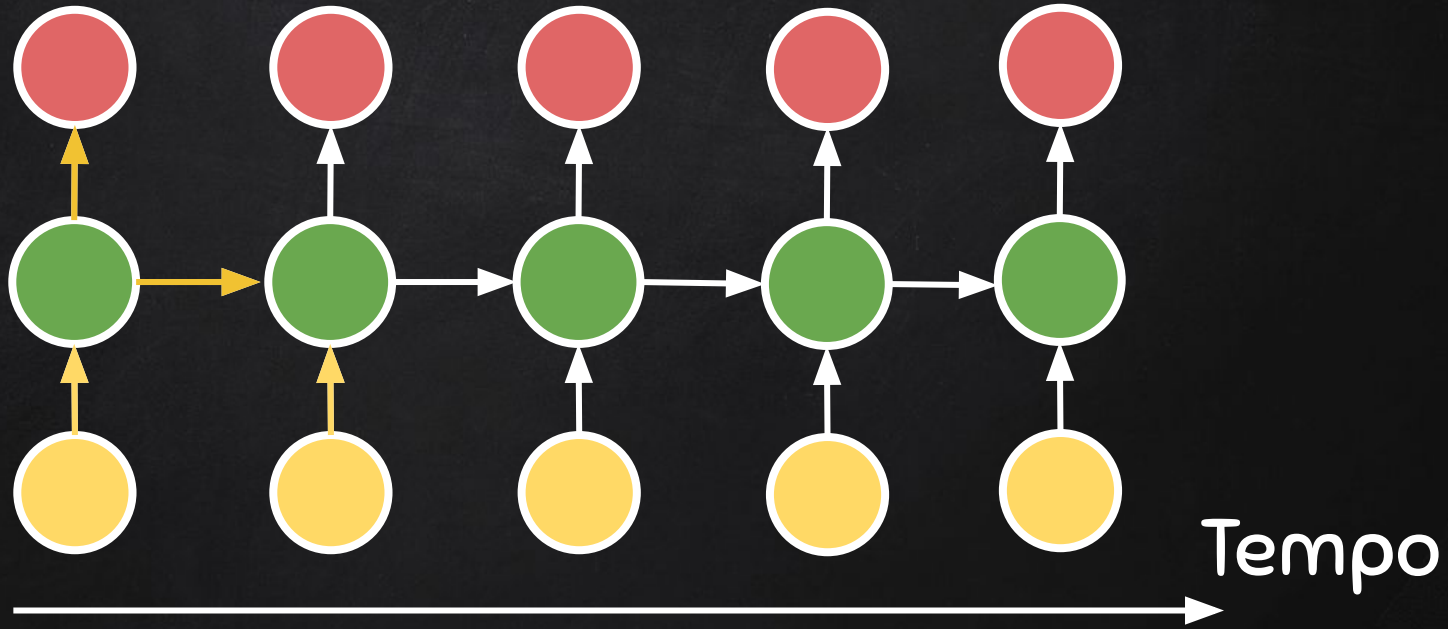
Vanish Gradient Problem



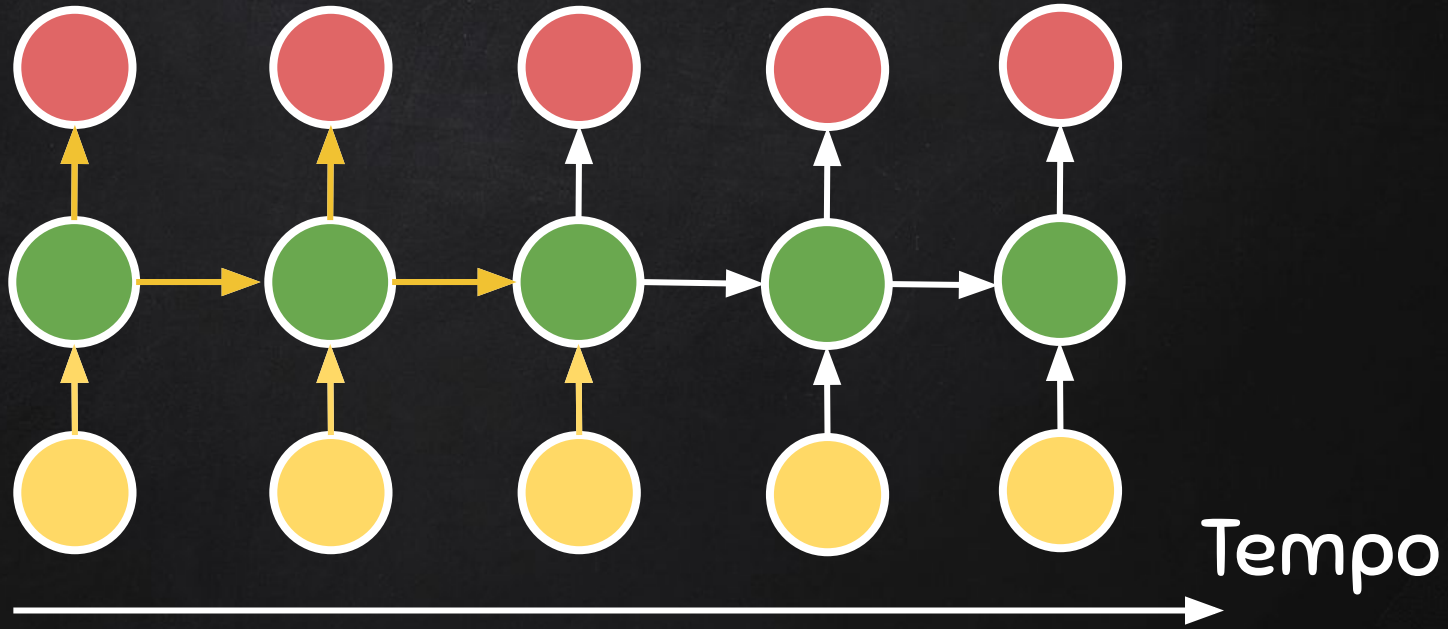
Vanish Gradient Problem



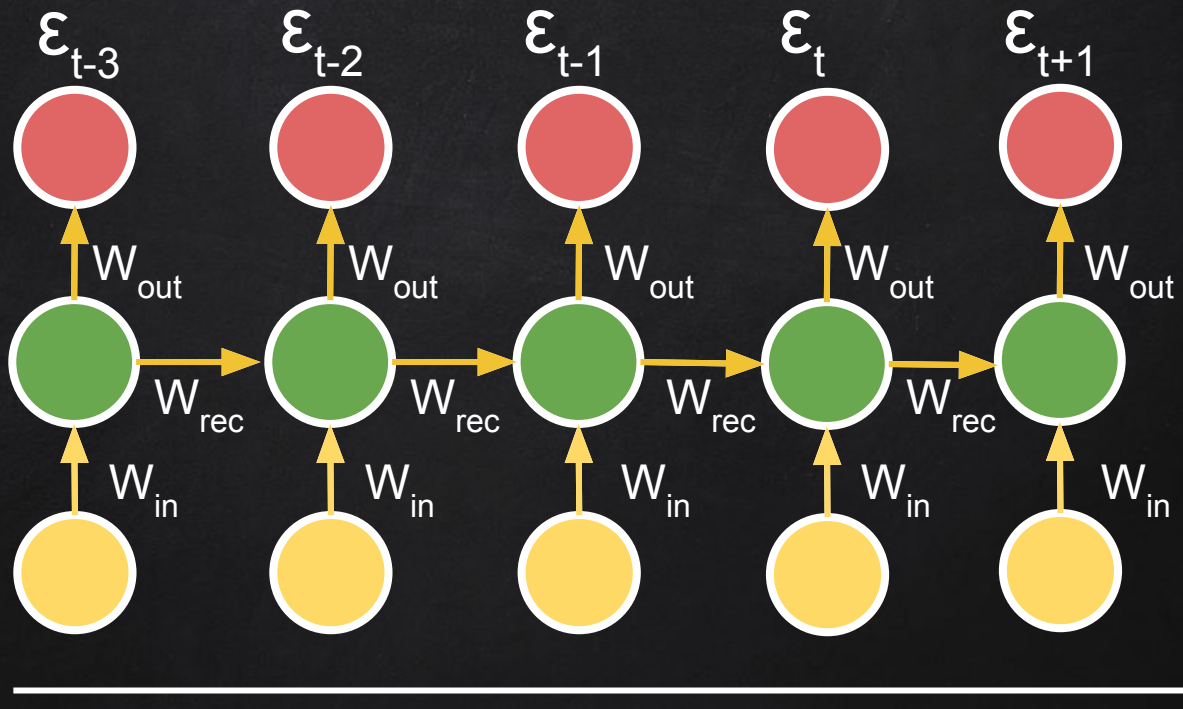
Vanish Gradient Problem



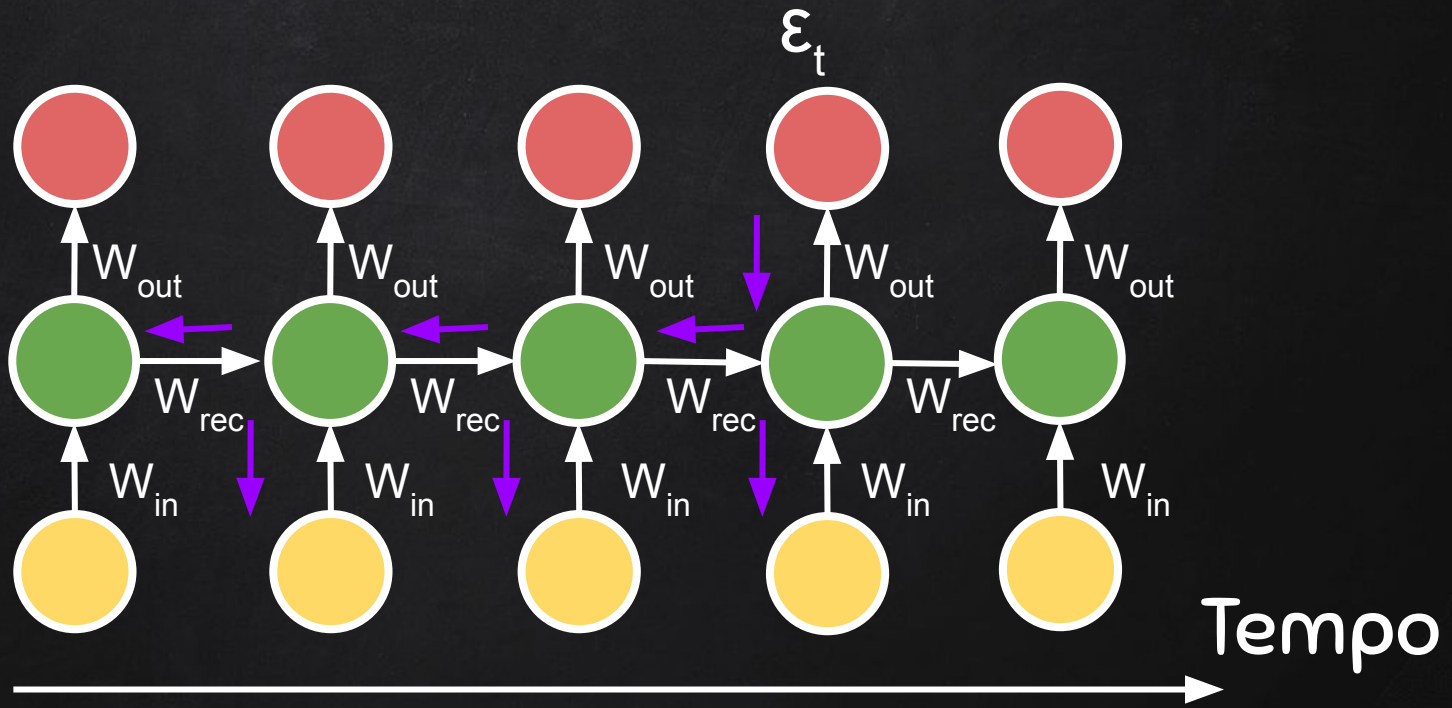
Vanish Gradient Problem



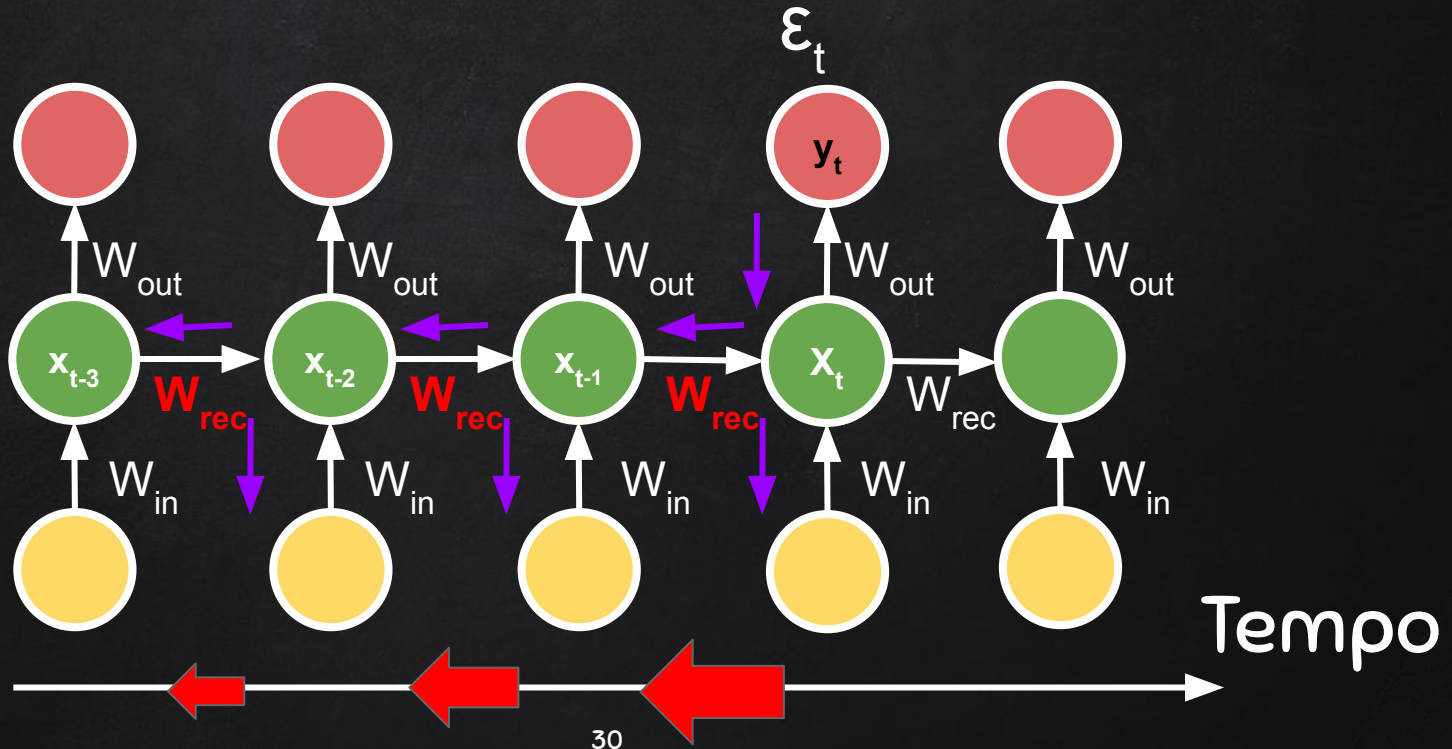
Vanish Gradient Problem



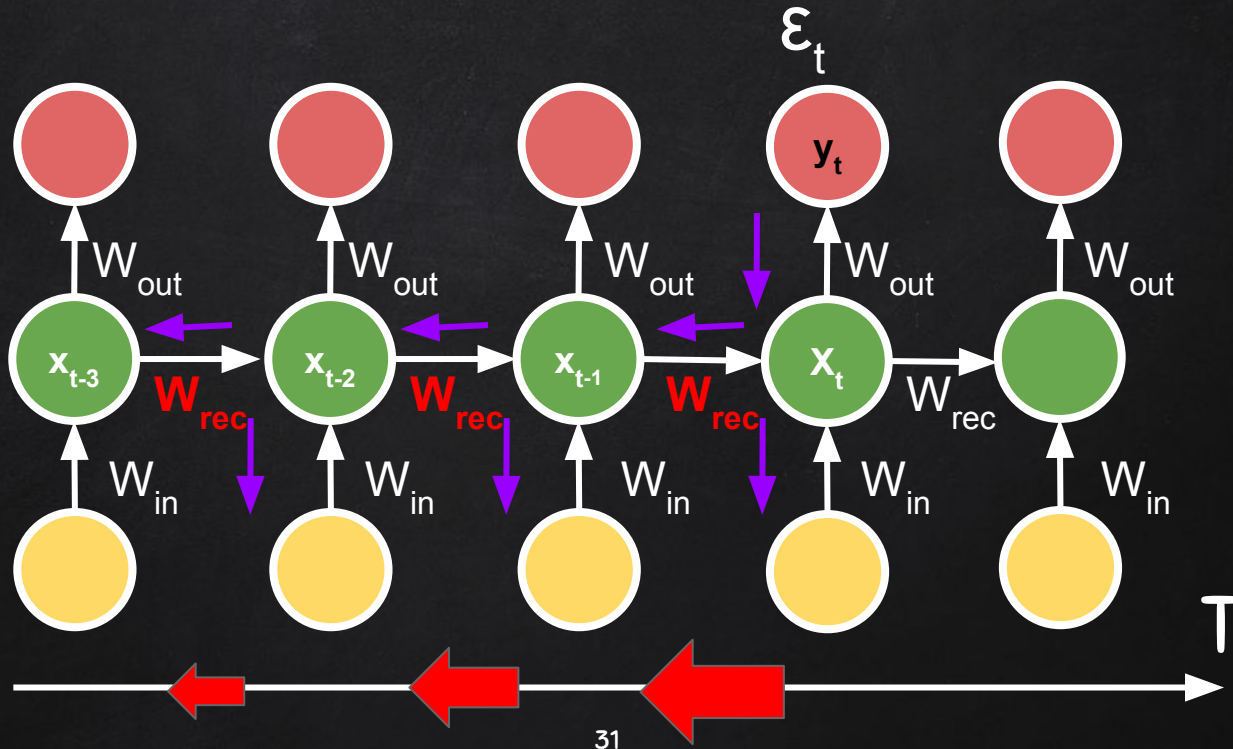
Vanish Gradient Problem



Vanish Gradient Problem



Vanish Gradient Problem



W_{rec} ~ peq = Vanishing

W_{rec} ~ grande = Exploding

Tranquilo?

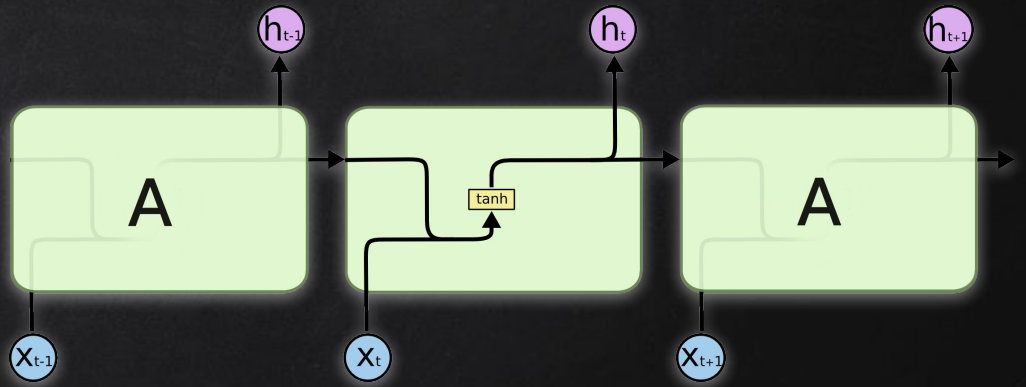
LSTMs

Long Short Term Memory

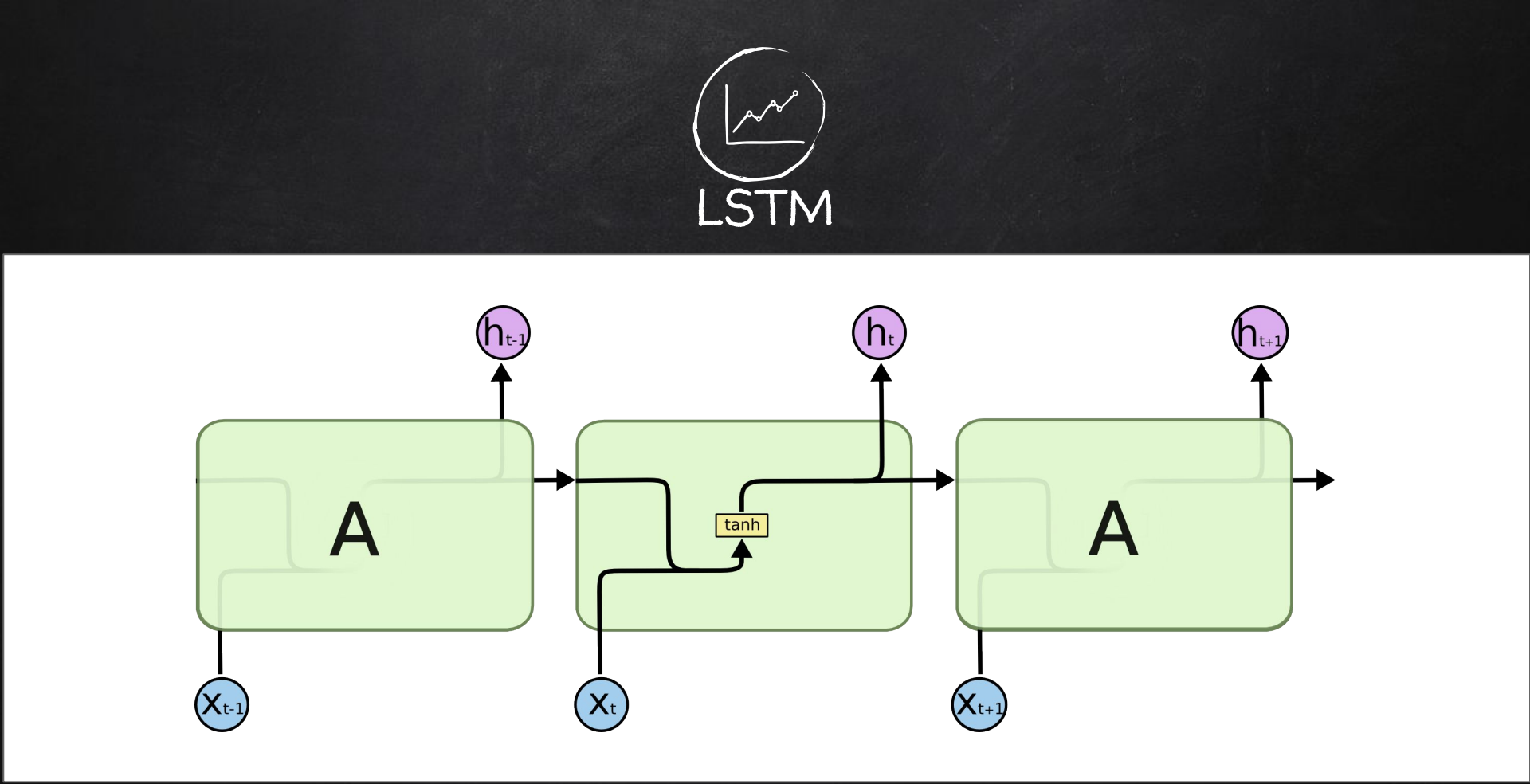
Chegou o momento!

Understanding LSTM Networks

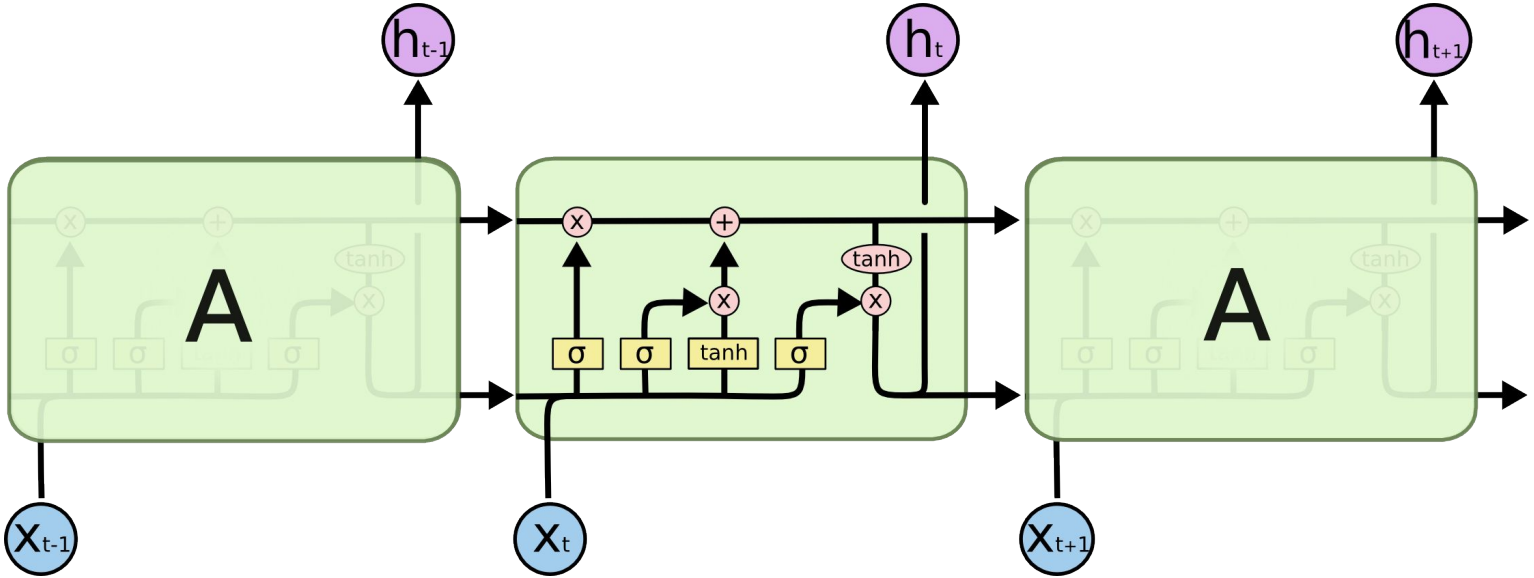
Cristopher Olah (2015)



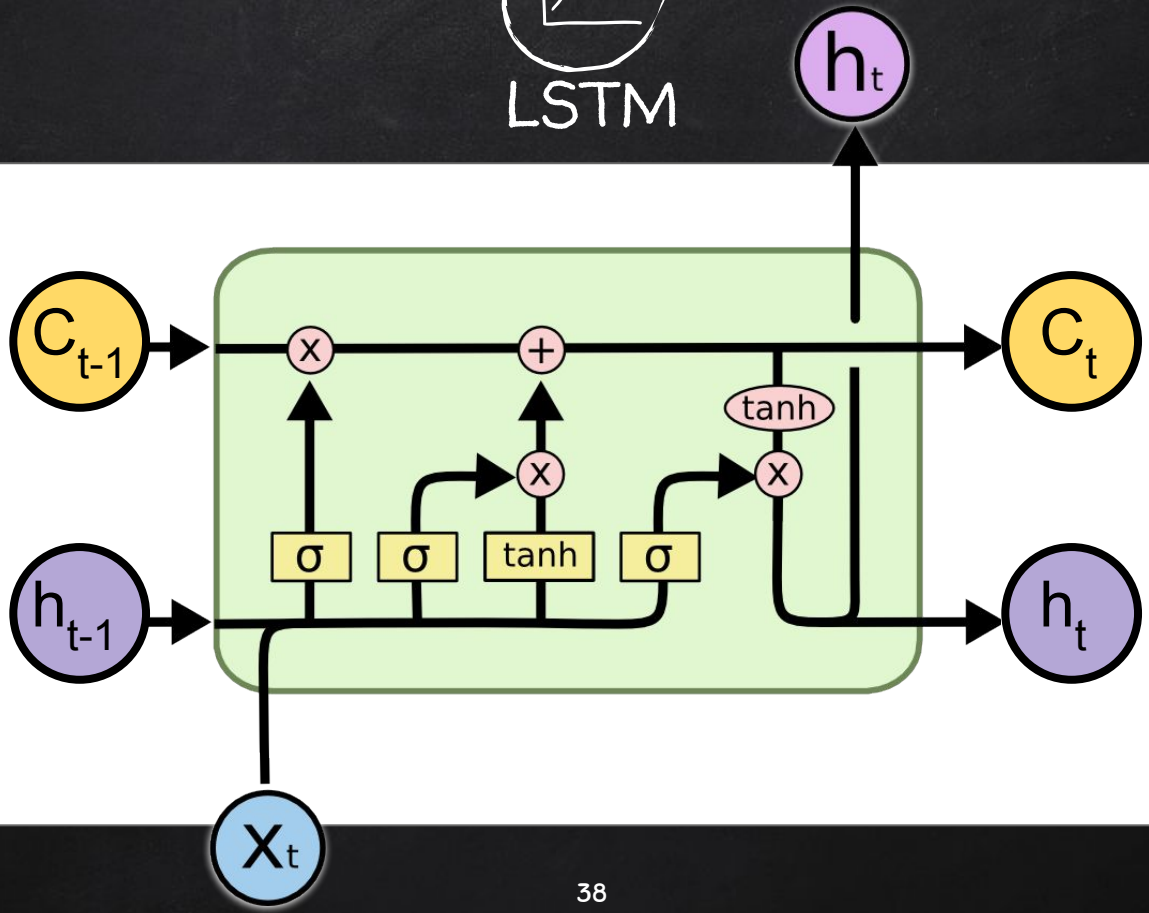
http://colah.github.io/posts/2015-08-Understanding-LSTMs/


LSTM
h_{t-1}
h_t
h_{t+1}
A
tanh
A
X_{t-1}
X_t
X_{t+1}

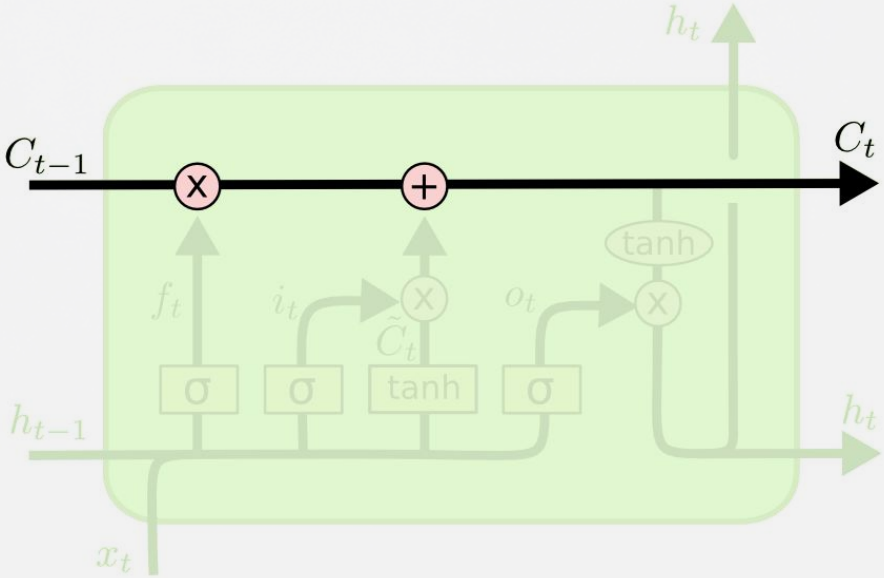
LSTM



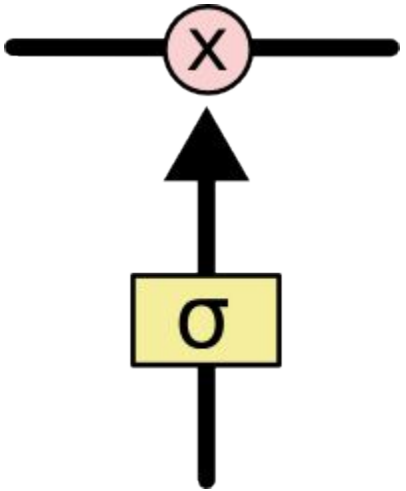
LSTM



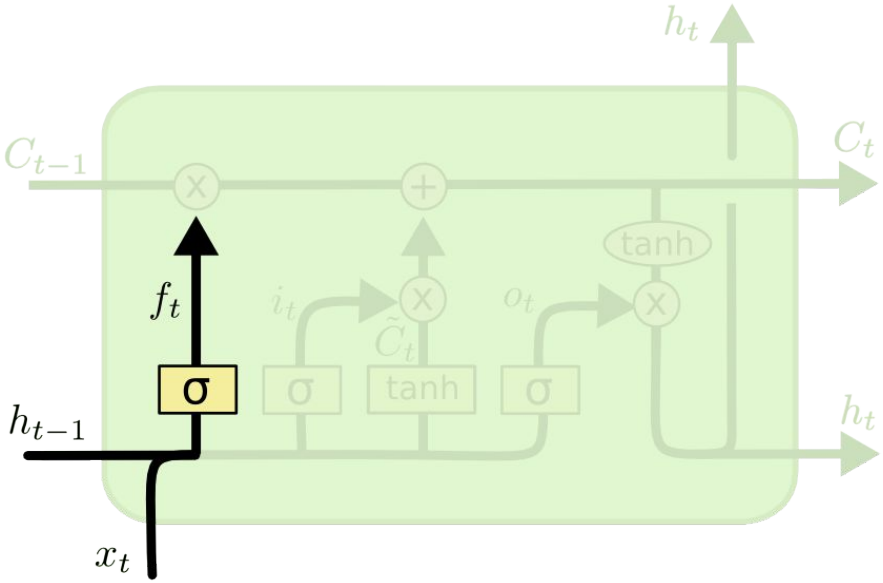

LSTM


h_t
C_{t-1}
C_t
tanh
f_t
i_t
o_t
$\tilde{C}_t$
σ
σ
tanh
σ
h_{t-1}
h_t
x_t


LSTM

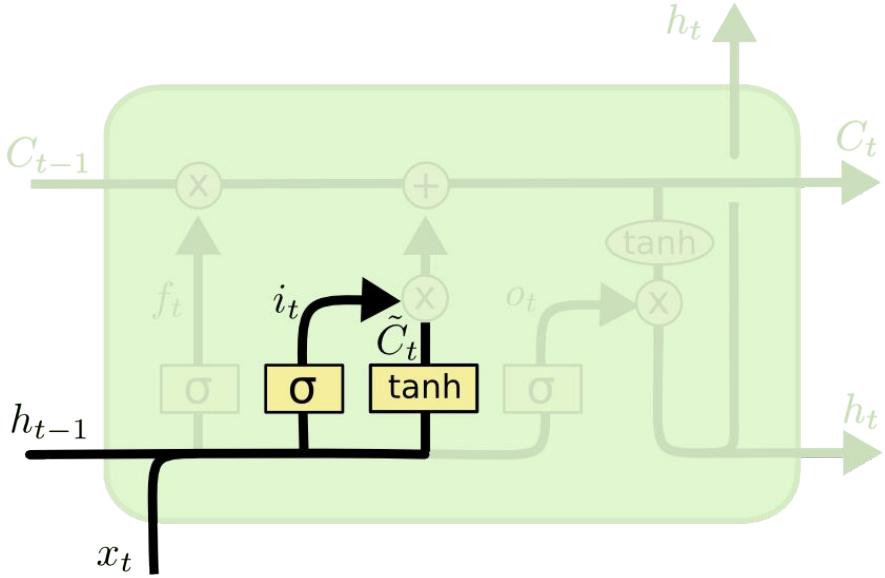

X
σ

Forget Gate



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

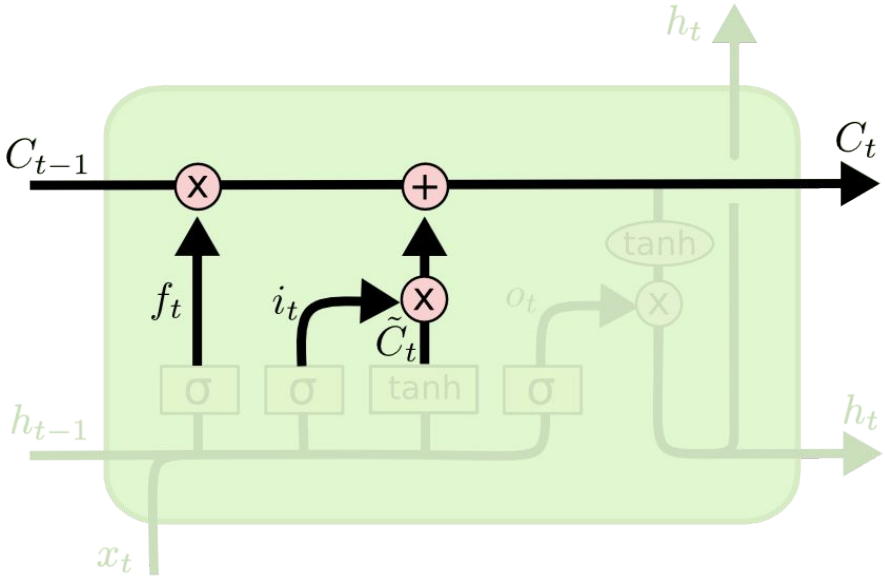
Input Gate



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

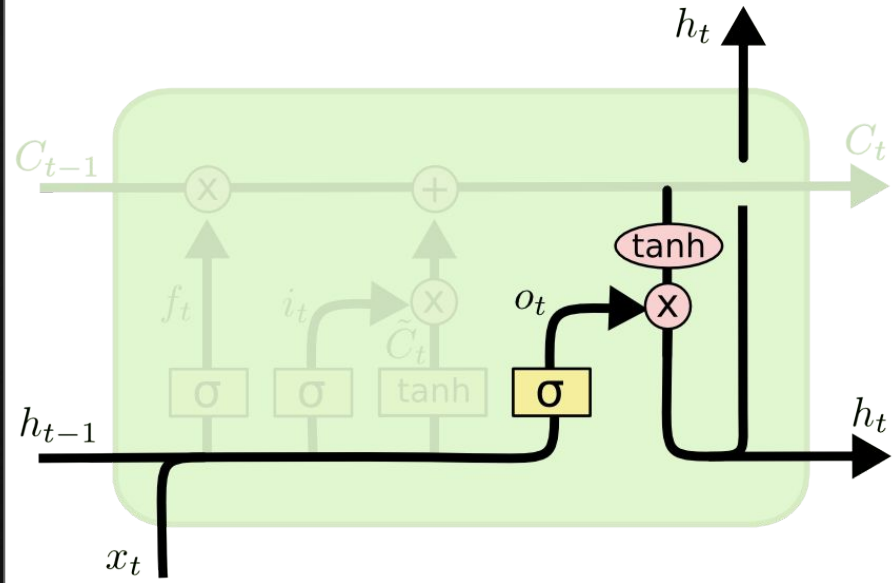
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Update Gate



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Output Gate

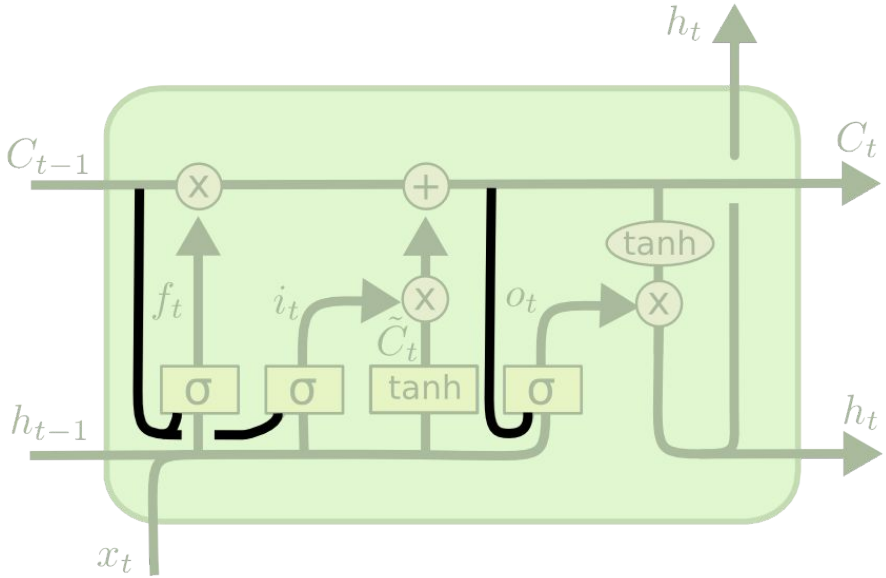


$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right)$$

$$h_t = o_t * \tanh\left(C_t\right)$$

Agora ficou claro...

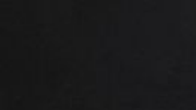
Peephole Connections



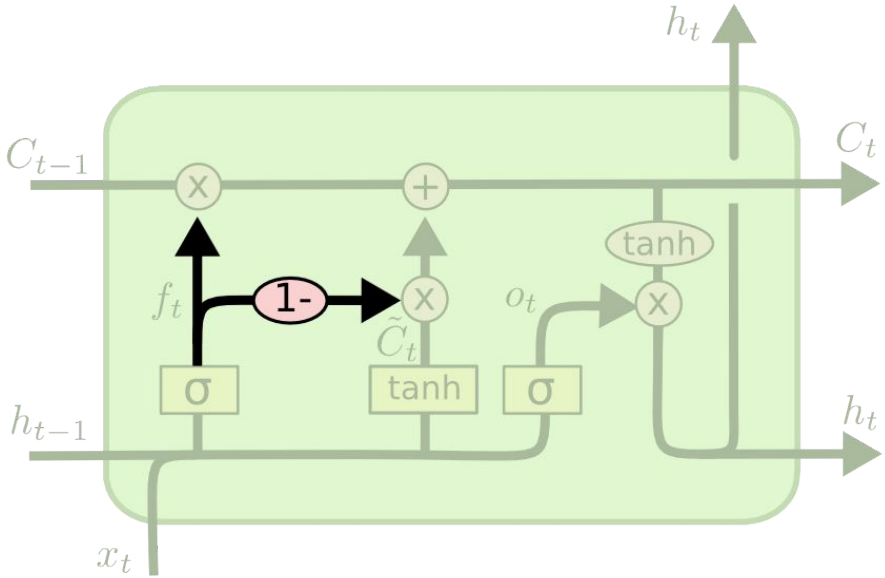
$$f_t = \sigma\left(W_f \cdot [\boldsymbol{C_{t-1}}, h_{t-1}, x_t] + b_f\right)$$

$$i_t = \sigma\left(W_i \cdot [\boldsymbol{C_{t-1}}, h_{t-1}, x_t] + b_i\right)$$

$$o_t = \sigma\left(W_o \cdot [\boldsymbol{C_t}, h_{t-1}, x_t] + b_o\right)$$

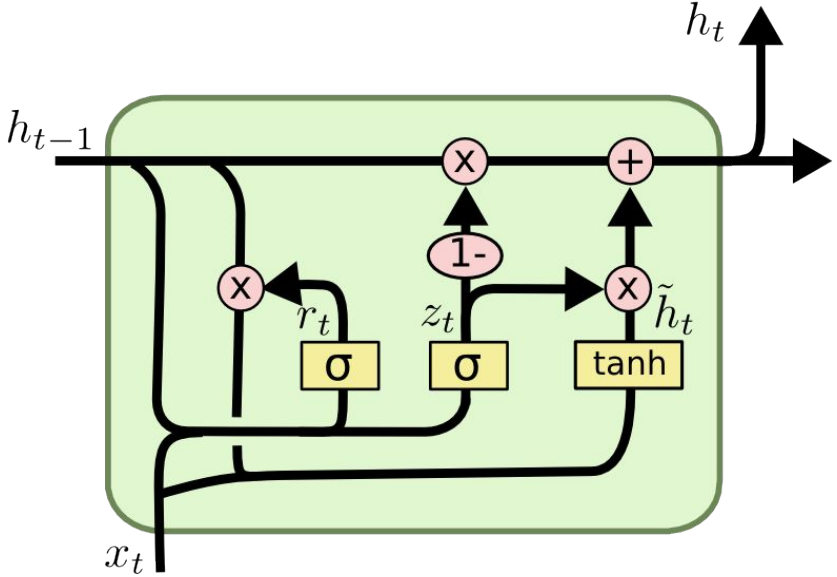


Coupled Forget and Input Gates



$$C_t = f_t * C_{t-1} + (1 - \boldsymbol{f_t}) * \tilde{C}_t$$

Gated Recurrent Unit



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

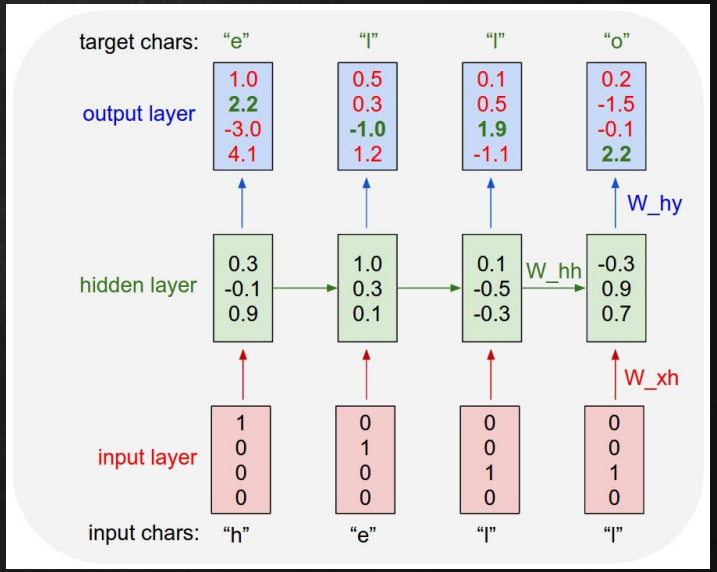
$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

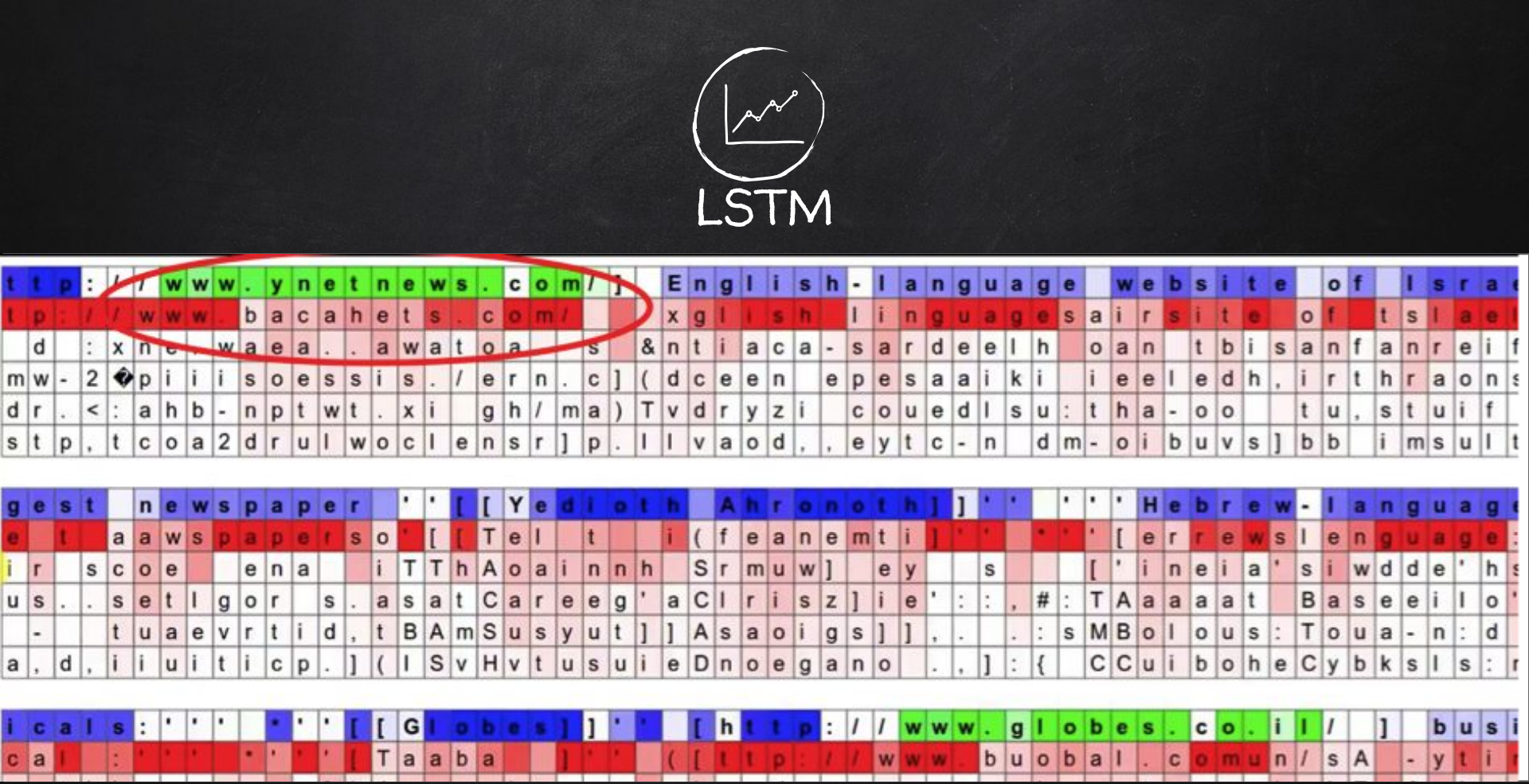
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

The Unreasonable Effectiveness of Recurrent Neural Networks

Andrej Karpathy (2015)



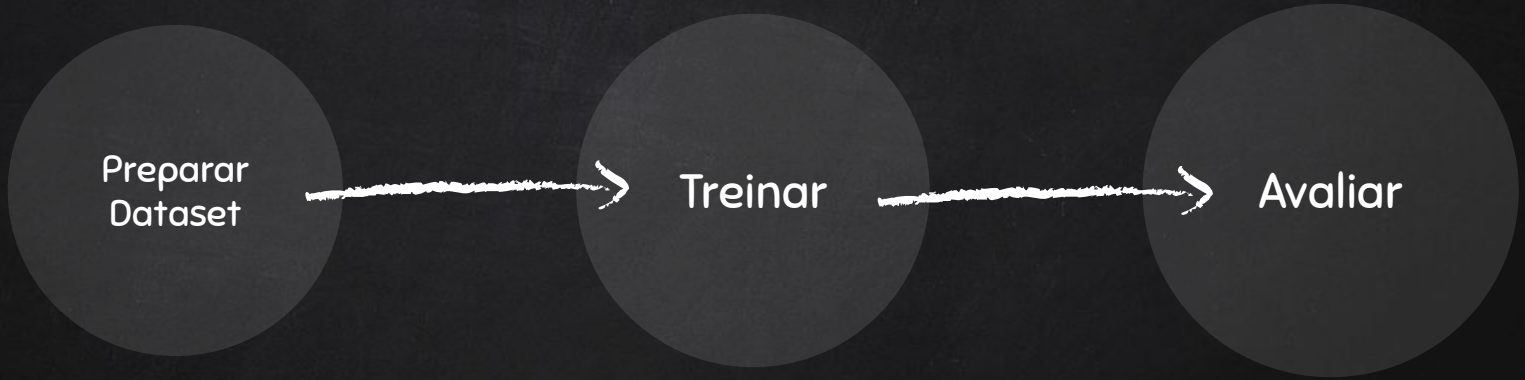
http://karpathy.github.io/2015/05/21/rnn-effectiveness/


LSTM

Prática
Building RNN

Prevendo preços de ações

Building CNN



https://github.com/deeplearningunb/building-rnn

Dever de Casa...

Valendo nota =)

1. Checkout na branch com seu nome
2. Faça ajustes nos hyper parâmetros
3. TROQUE A LSTM POR UMA GRU
4. Commit do resultado (NA SUA BRANCH)

Obrigado!

Dúvidas?

http://bit.ly/dl-unb7
https://t.me/DeepLearningUnB
@diegodorgam

Credits

Special thanks to all the people who made and released these awesome resources for free:

- ✘ Presentation template by SlidesCarnival
- ✘ Photographs by Unsplash