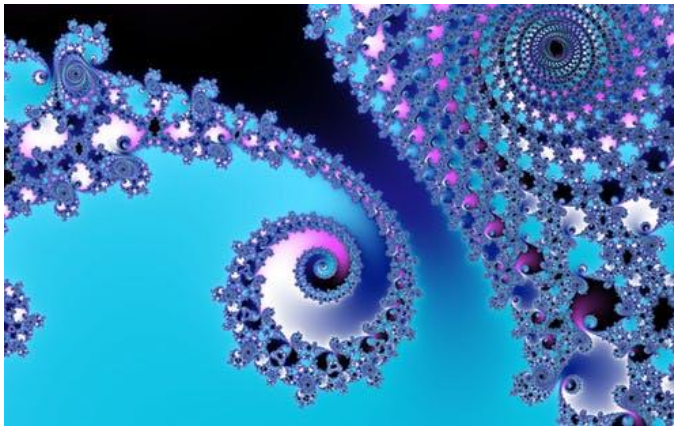


LSTM

Prof. Esp. Victor Venites



*SCHOOL OF AI – SÃO PAULO – AULA 19
– DEEPLARNING – LONG SHORT-TERM
MEMORY – RECURRENT NEURAL
NETWORK*

Até Aqui

Regressão Linear –

- Álgebra Linear
- Derivadas
- Vetores
- Matrizes

Introdução ao Machine Learning –

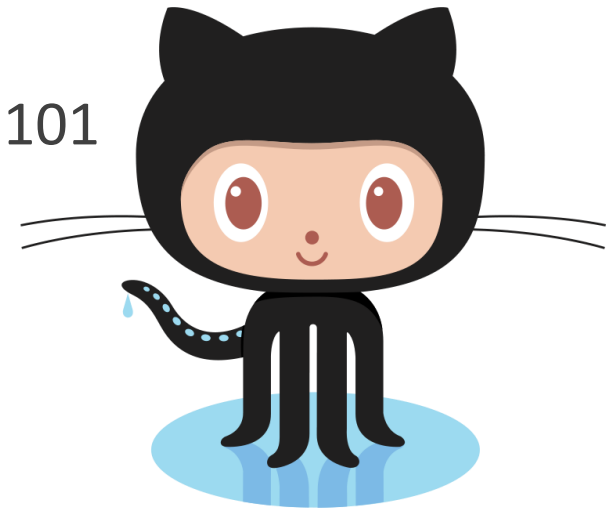
- Árvores de Decisão
- Naive Bayes
- Support Vector Machine
- KNN
- K-means

Estatística –

- Análise Descritiva
- Exploração de Dados
- Séries Temporais

Exemplos –

- Hands-On – 101
- Slides
- Python



LSTM – Long Short-Term Memory



Roteiro –

- Perceptron
- Radial Bases Function
- DeepLearning
- AutoEncoder
- Convolutional Neural Network
- **LSTM**

Objetivo

- Entender a LSTM com mais do que já se vê por ai
- Trazer uma analogia profunda
- Mostrar o que senti falta de se explicar
- Dar as ferramentas para você virar um estudante melhor
- Passar um pouco da minha experiência
- Deixar o aluno apto para aplicar
- Levantar questões... E responder a maioria!



Material: GitHub / Slides e Código

Vídeo: YouTube - Live

LSTM – Artigos Históricos

Long Short-Term Memory

Shmdhuber e Hochreiter

(1997)

Long Short-Term Memory

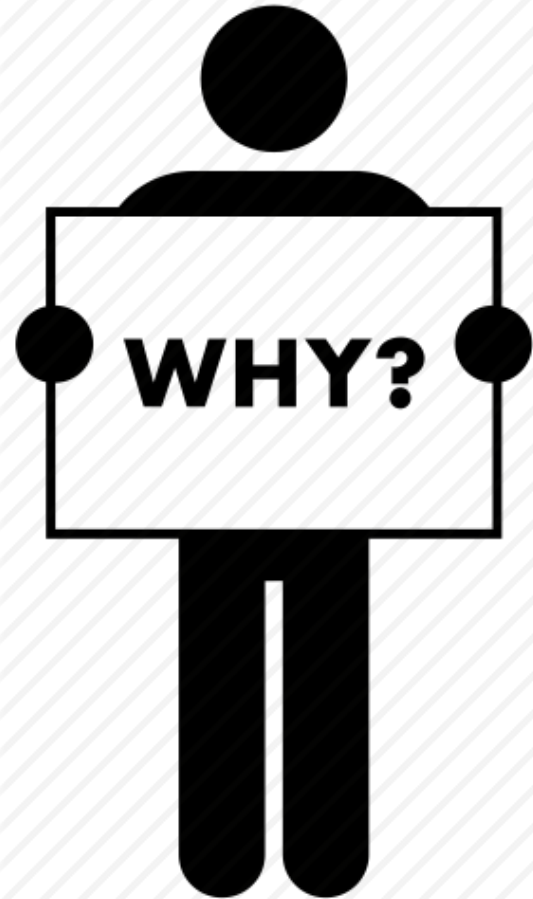
Shmdhuber e Hochreiter

(1995 e 1996)

Nuvem de Palavras – Artigos Referência

- *Maquinas de Estado Finito*
- *Learning Sequential Tasks*
- *Alternatives to BackPropagation*

Por quê LSTM?



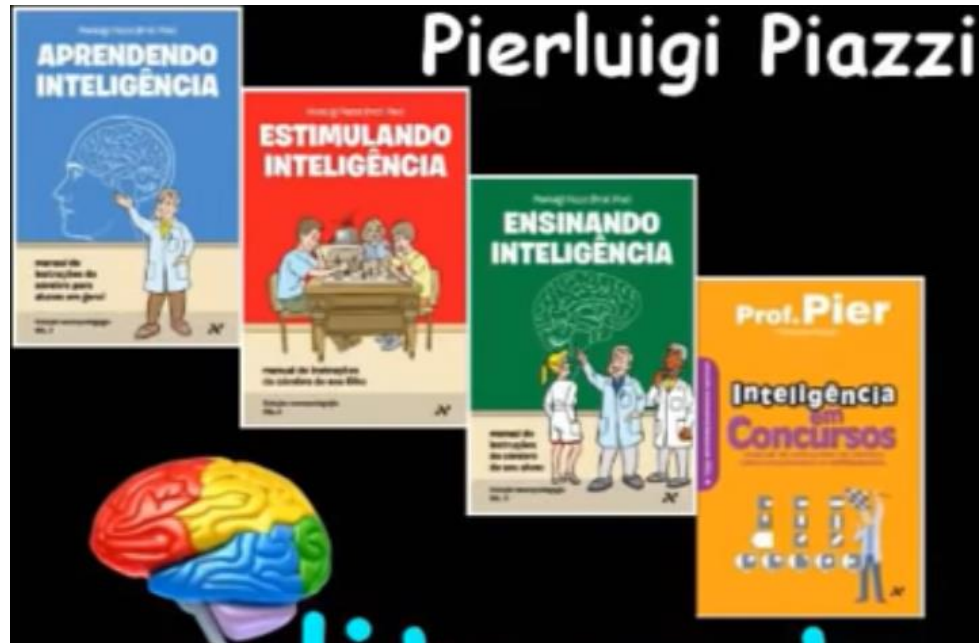
- Nós não aprendemos de forma estática
- “Matemática Dinâmica” -> Decodificação Contínua
- Ensinar uma Máquina (Computador) a reconhecer padrões diferentes e responder mais rápido á eles
- BigData (Transformação Digital)
- Aumento de Complexidade pelo grande tamanho dos dados
- Digitalização de Áudio, Vídeo e Texto
- Séries Temporais com mudanças constantes
- Vanish Problem – Problema do Decaimento ao longo do tempo

Voltando ao Básico

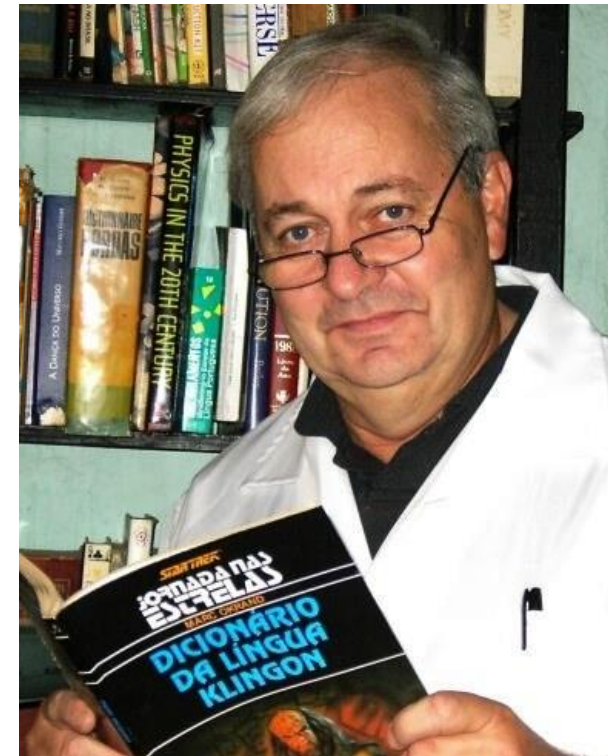
Nesta Aula:

- Inteligência Natural
- Séries Temporais
- Perceptron
- RNN

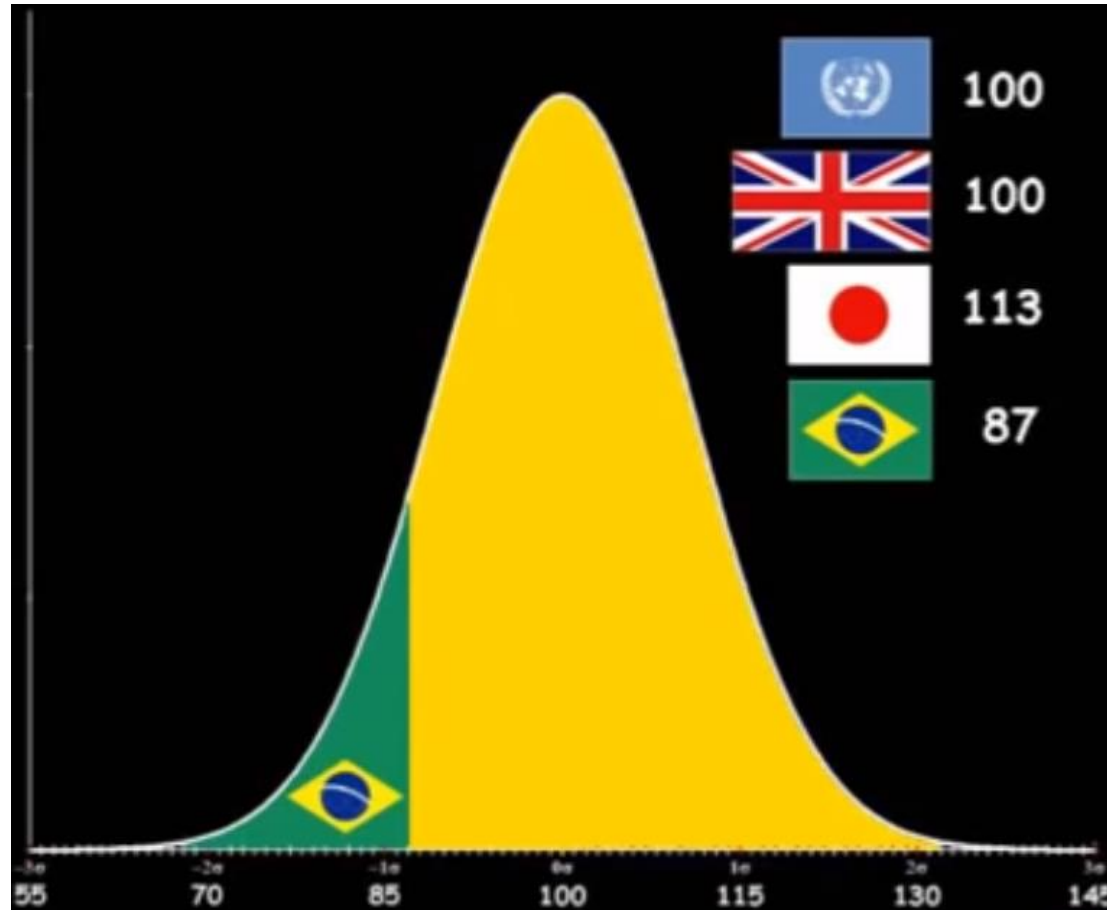
Inteligência Natural – Pierluigi Piazzi



1. As estudantes
2. Aos pais
3. Aos professores
4. Aos autodidatas

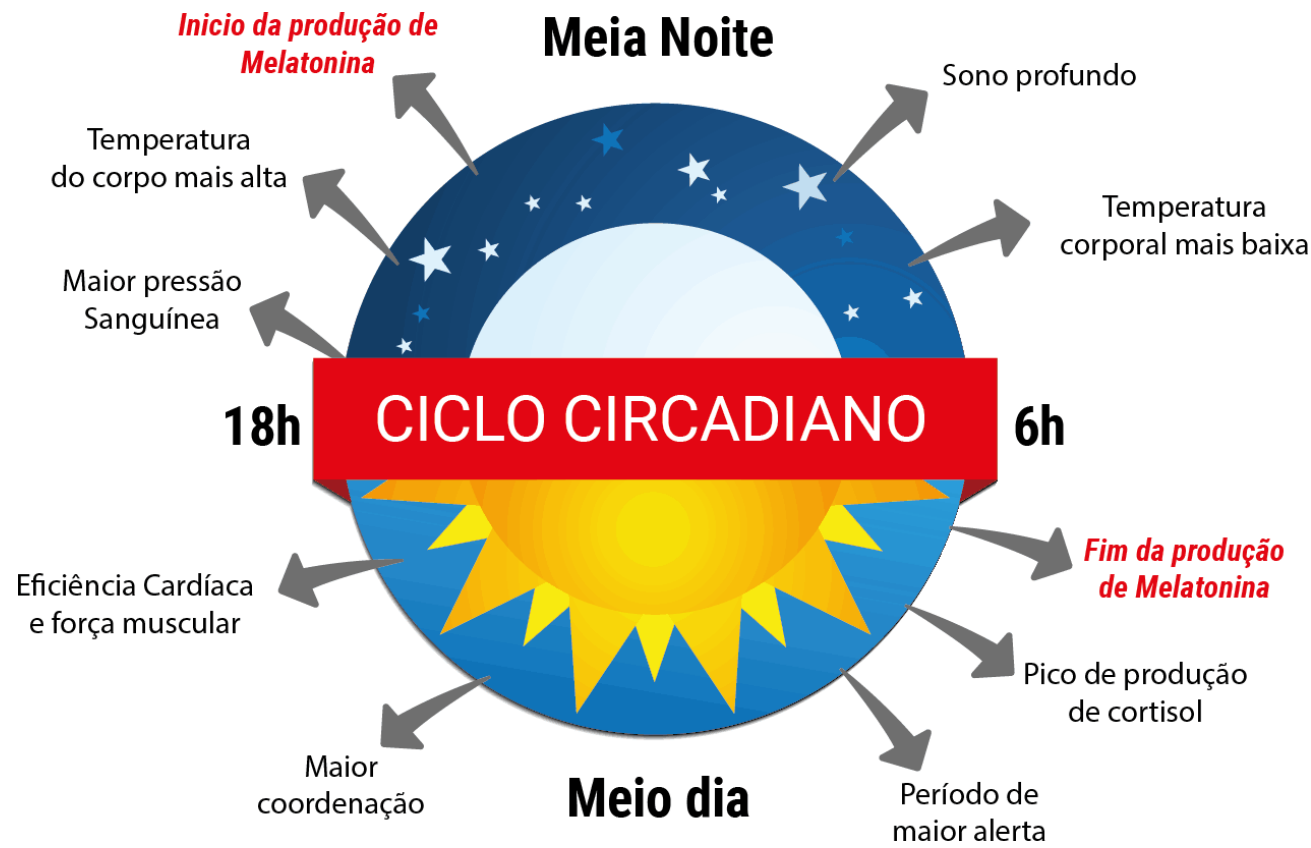


“Inteligência se Aprende”



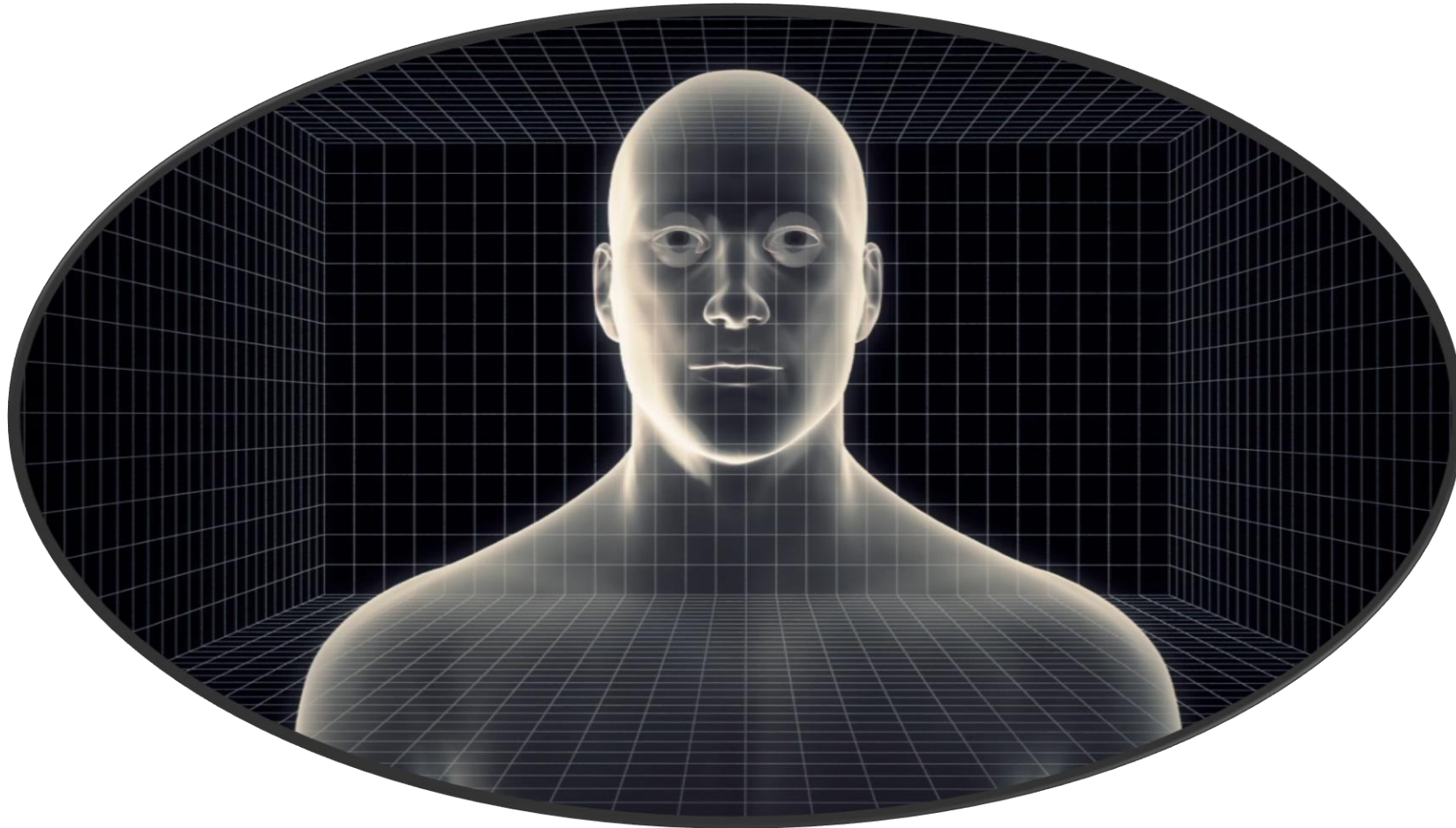
- Ao vermos a média de QI neste gráfico e vendo todo conteúdo do Professor Pier ele diz “Inteligência se Aprende”

Como Aprender?

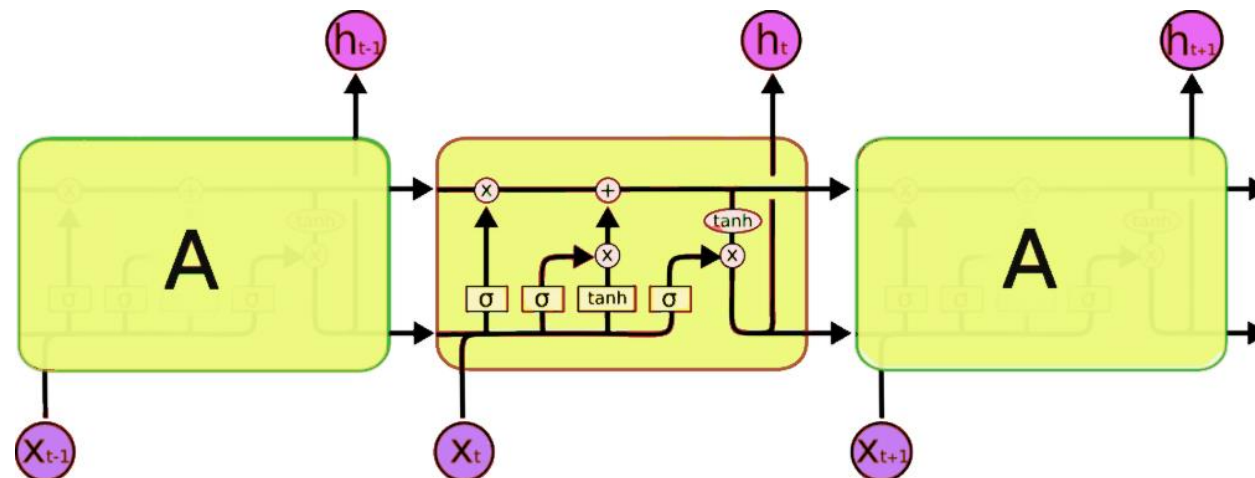


- Analogia do Ciclo com a LSTM
- “Aula dada, aula estudada hoje”
- Inteligência Natural
- Ciclo de aprendizado do cérebro humano no mesmo dia
- Nosso LOOP de aprendizado

O que Acontece no Sono?

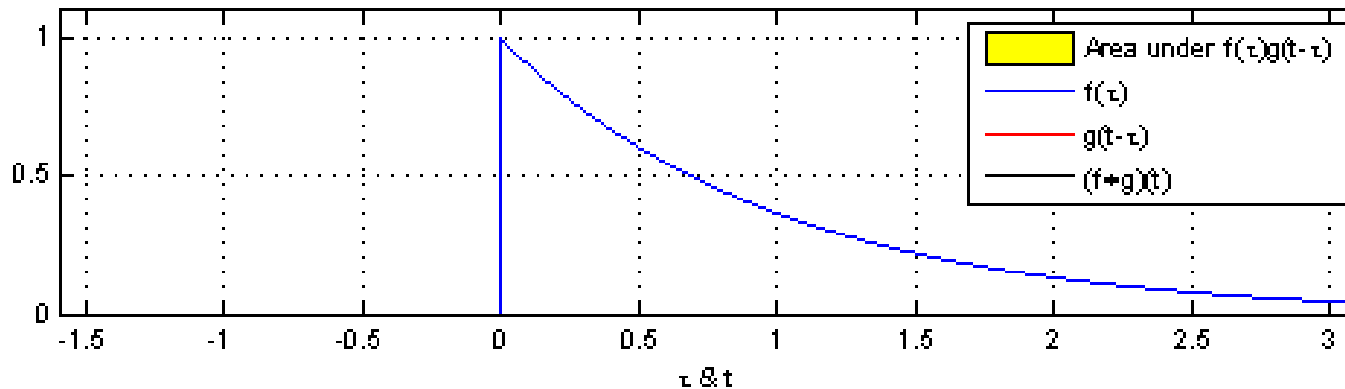


E na LSTM?



Série Temporal

X1	X2	X3	...	Xn	Y
1	2	3	...	100	73
2	3	4	...	200	98
3	4	5	...	300	10
6	7	8	...	400	13
9	10	11	...	500	25



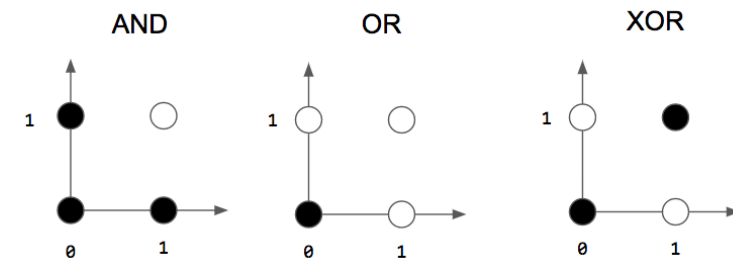
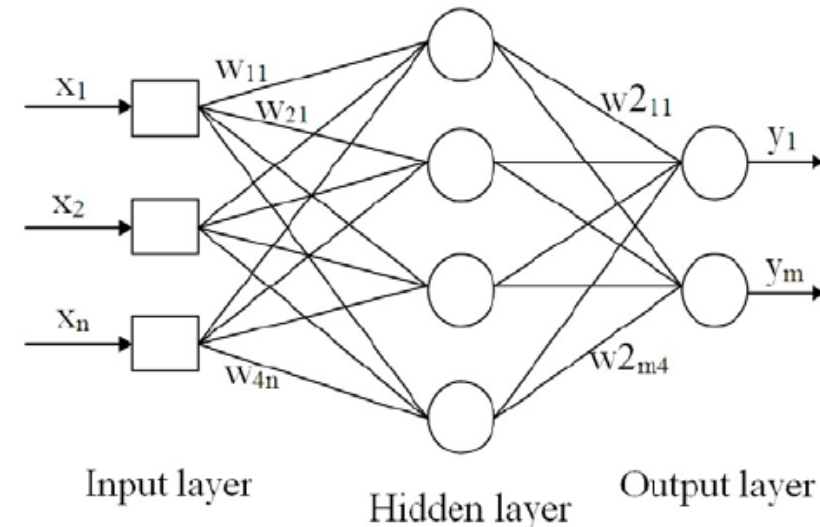
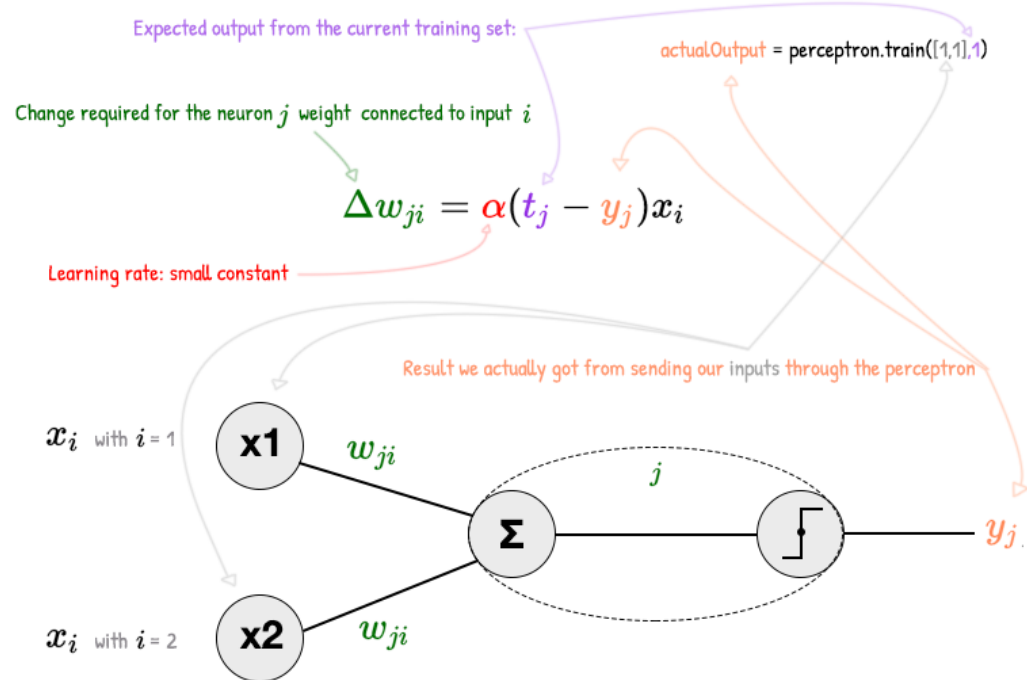
X	Y
1	73
2	98
3	10
4	42
5	35
...	...
n	25



- Aula 11/2019 - 21/05 - Parte 2 - Série Temporais aplicada a Investimento e Bitcoins

<https://www.youtube.com/watch?v=lbQ5UEOWi1o>

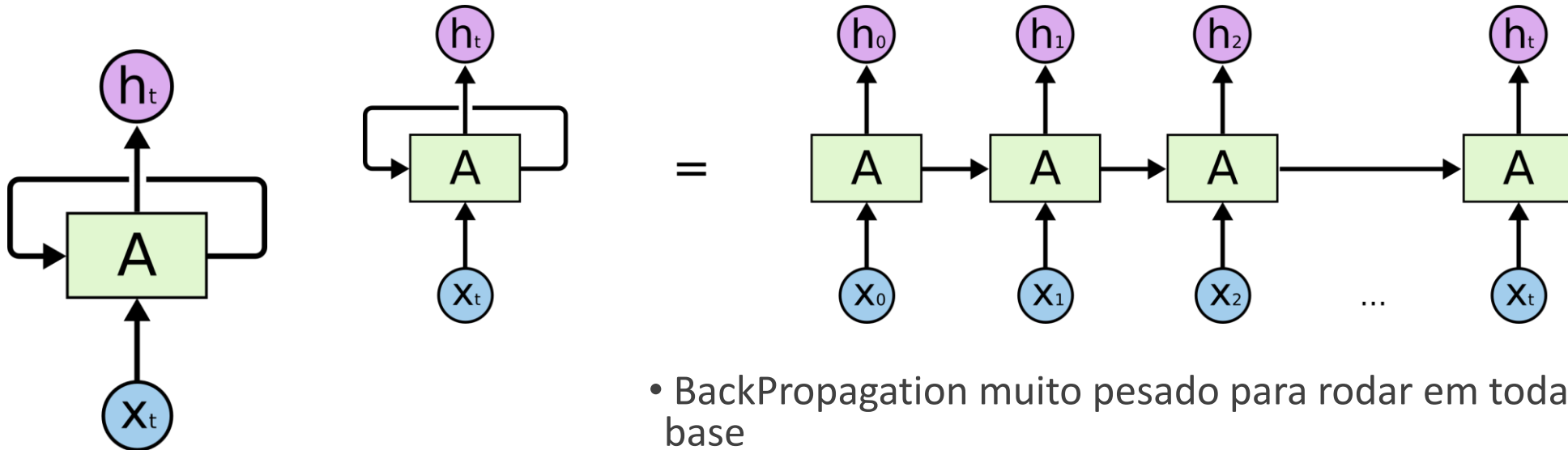
Rede Neural Perceptron



- Aula 13/2019 - 18/06 - Parte 2 - Introdução a redes Neurais e Perceptrons

- <https://www.youtube.com/watch?v=TOcBk3BzBXU>

Rede Neural Recorrente



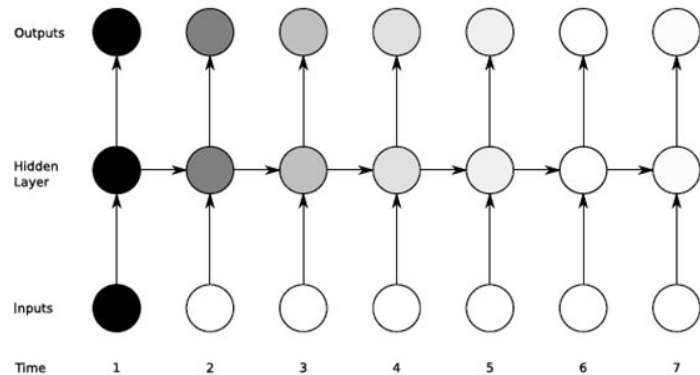
$W_{rec} \sim \text{small}$ \Rightarrow Vanishing
 $W_{rec} \sim \text{large}$ \Rightarrow Exploding

- BackPropagation muito pesado para rodar em toda base
- Dados muito antigos podem influenciar em uma cultura atual de dados
- Variações grandes e imprevistas nos novos dados podem fazer o modelo de perder um pouco até se adaptar

Rede Neural Recorrente

Recurrent Neural Networks

- Vanishing gradient problem



$$a_h^t = \sum_{i=1}^I w_{ih} x_i^t + \sum_{h'=1}^H w_{h'h} b_{h'}^{t-1}$$

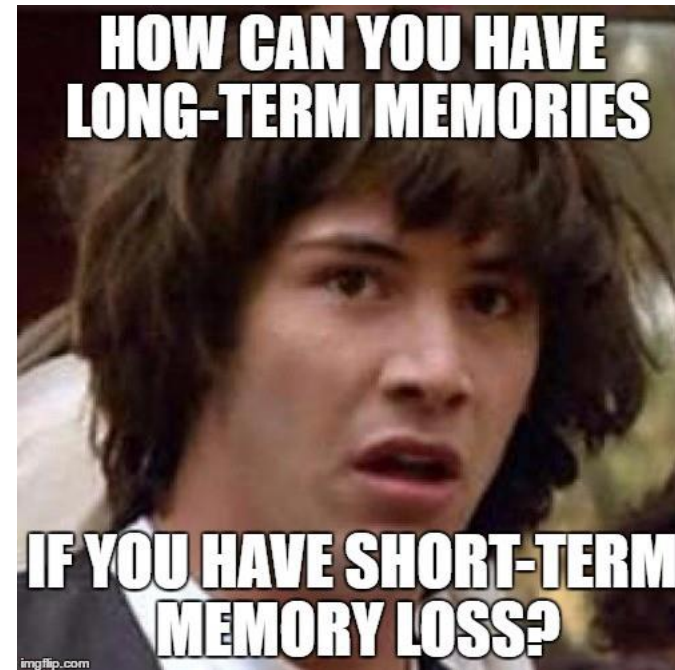
$$b_h^t = \theta_h(a_h^t)$$

Sensitivity decay exponentially over the time

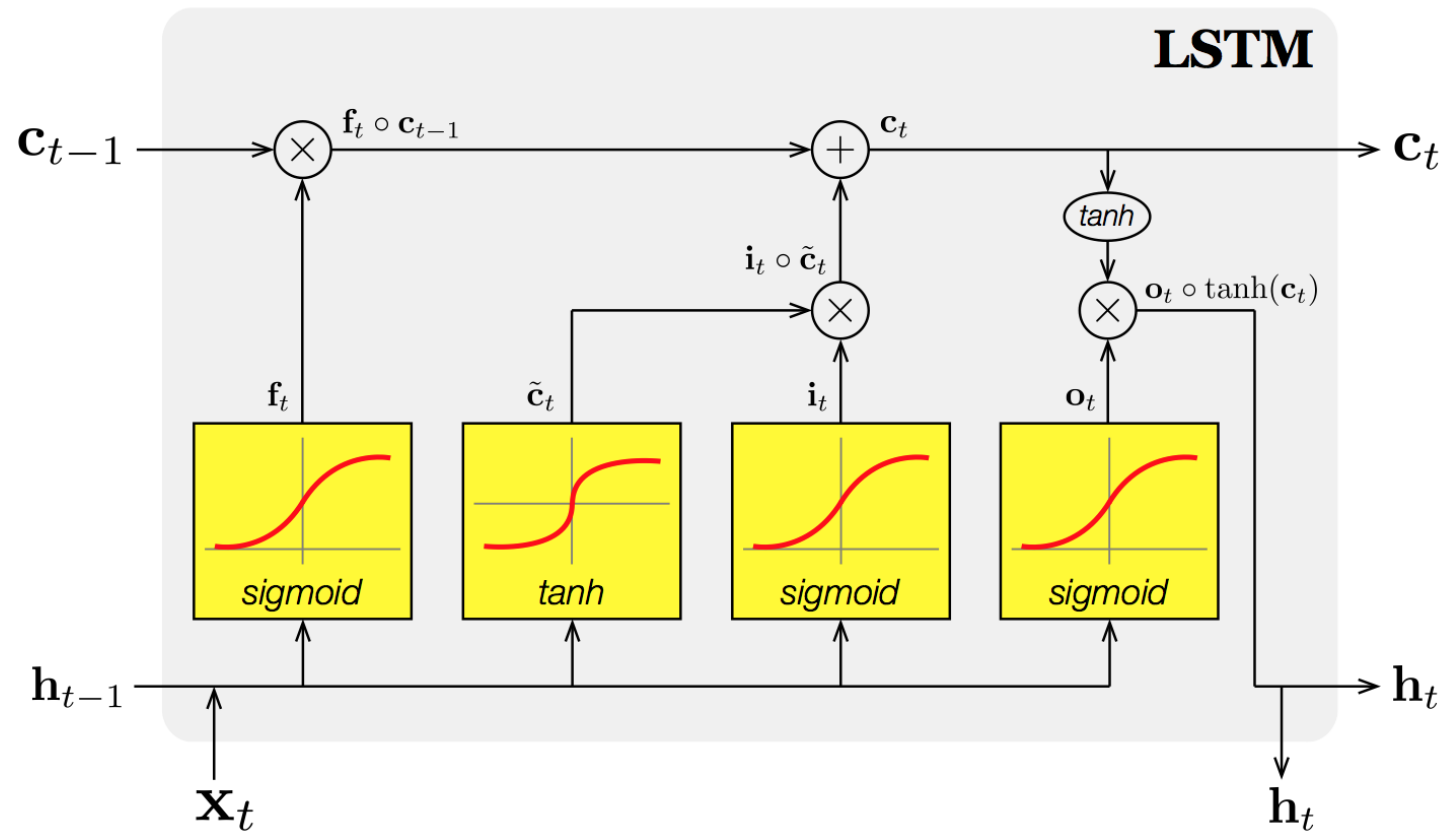
O que é LSTM?

Conceitos:

- Memória
- Estado
- Pesos e Bias



Como a Célula LSTM funciona?



Gating variables

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o)$$

Candidate (memory) cell state

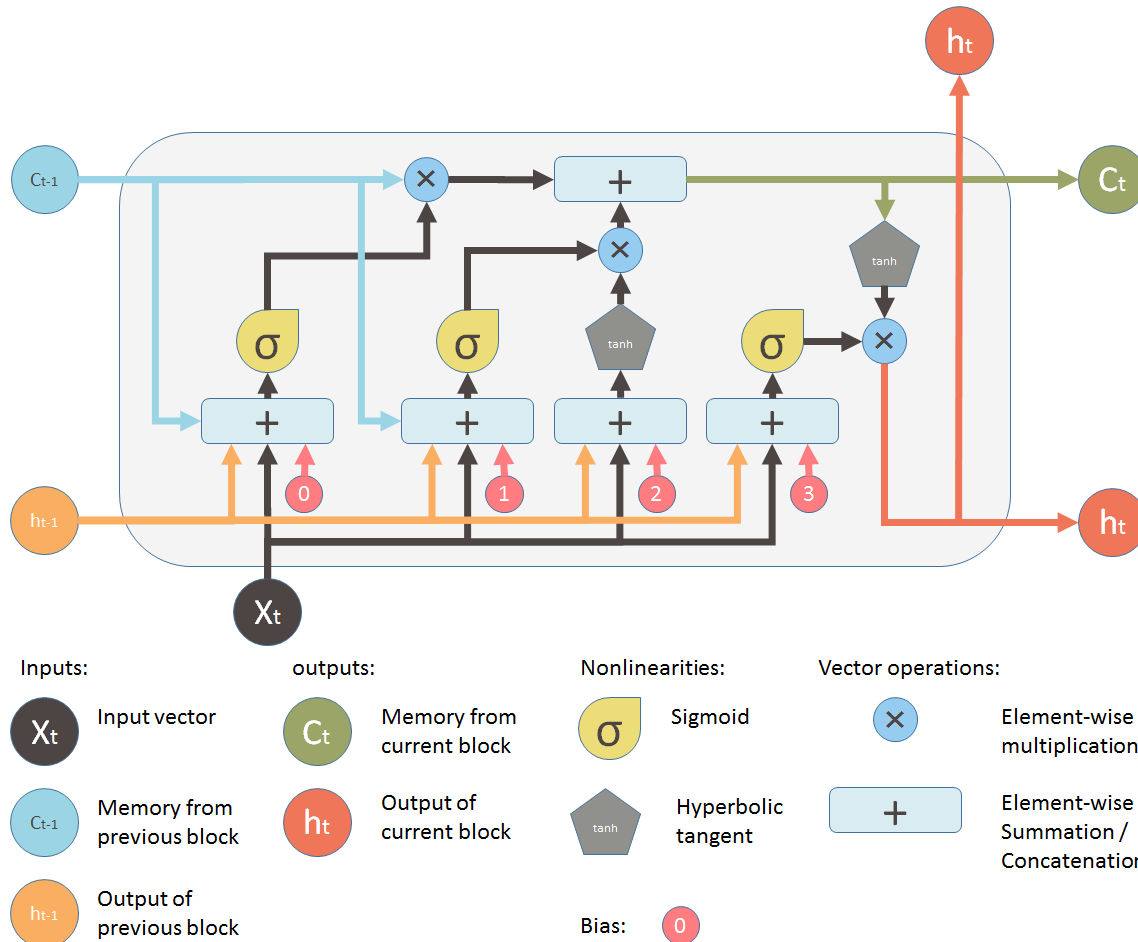
$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c)$$

Cell & Hidden state

$$\mathbf{c}_t = \mathbf{f}_t \circ \mathbf{c}_{t-1} + \mathbf{i}_t \circ \tilde{\mathbf{c}}_t$$

$$\mathbf{h}_t = \mathbf{o}_t \circ \tanh(\mathbf{c}_t)$$

Arquitetura LSTM



- Perceptron = Ativação($Wx \cdot X + Wh \cdot H + b$)

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

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

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

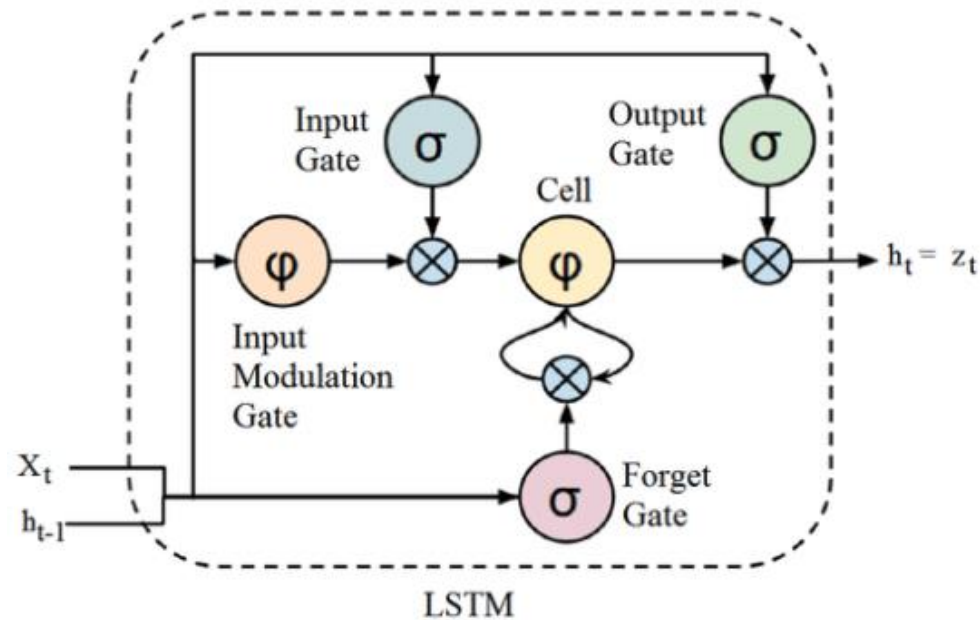
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

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

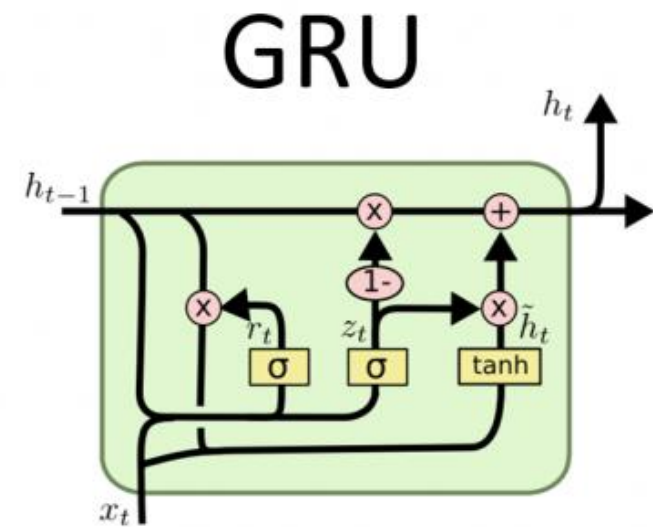
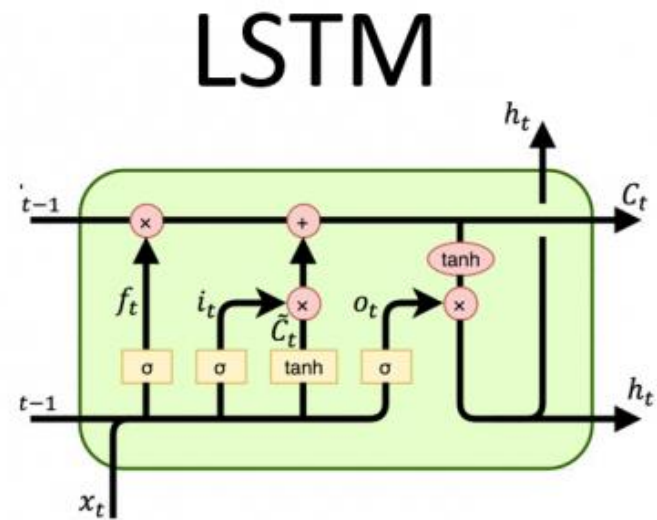
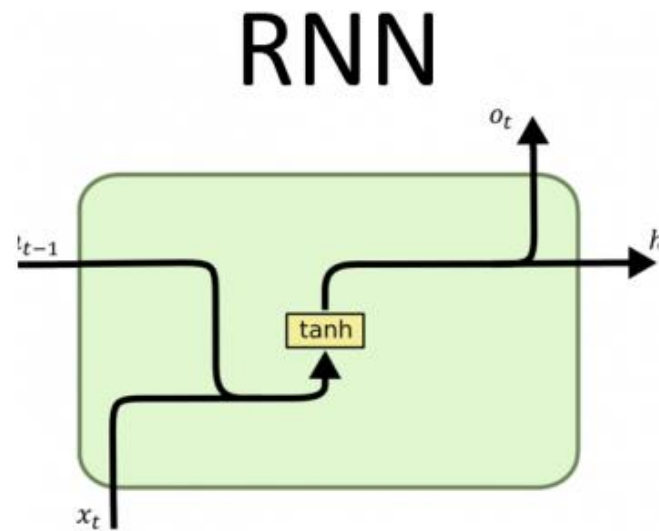
$$h_t = o_t * \tanh(C_t)$$

LSTM Simplificada = GRU

- Perceptron = Ativação($Wx \cdot X + Wh \cdot H + b$)



Resumo de Evolução



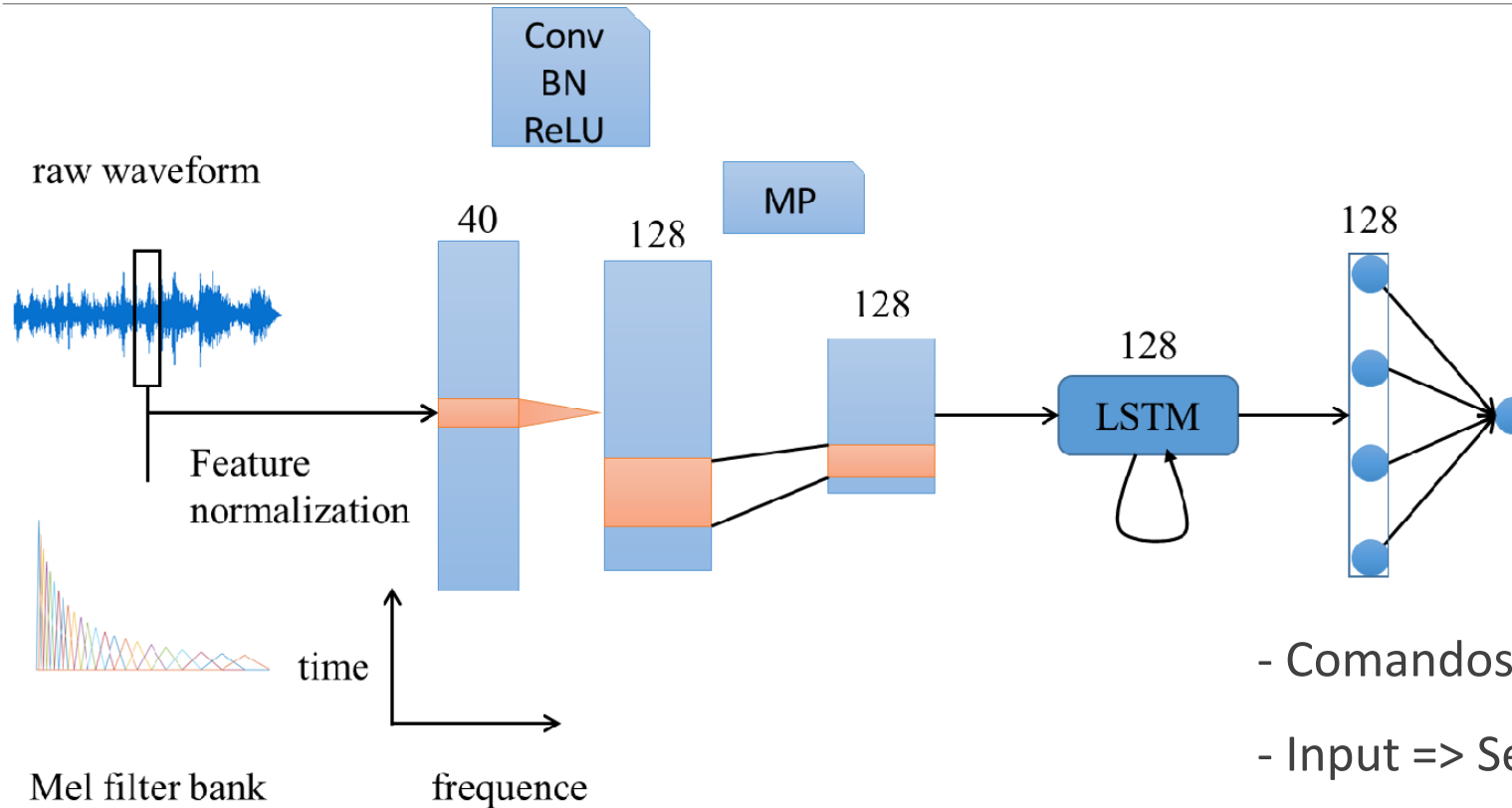
Onde adquirir bases de Dados?



Sites conhecidos:

- **Kaggle** – Bitcoin Historical Data – até Agosto/2019
<https://www.kaggle.com/mczielinski/bitcoin-historical-data/downloads/bitcoin-historical-data.zip/17>
- Base utilizada para aula de hoje
- Tem 240MB por isso não foi possível colocar no GitHub

Exemplos



- Comandos de Vós
- Input => Sequência de Dados
- Output => Comando de Voz(celular)

Exemplos



A. Read the passage. Answer the questions correctly.



THE FOX AND THE STORK

A fox invited a stork to his house for dinner. The fox served soup in shallow bowls. The fox licked up his soup very quickly. The stork was not able to drink the soup because his beak was too long. The stork went home hungry.

The next day, the stork invited the fox to his house for lunch. They were also having soup for lunch. The stork served the soup in tall jugs. The stork drank the soup easily but the fox could not reach the soup inside the tall jug. This time it was the fox that went home hungry.

1 Why did the fox invite the stork to his house?

- A** For lunch **B** For dinner **C** For breakfast

2 The stork could not drink the soup because _____.

- A** he was too hungry **B** his beak was too long **C** the soup was too hot

- Interpretação de Texto

- Corretor Ortográfico

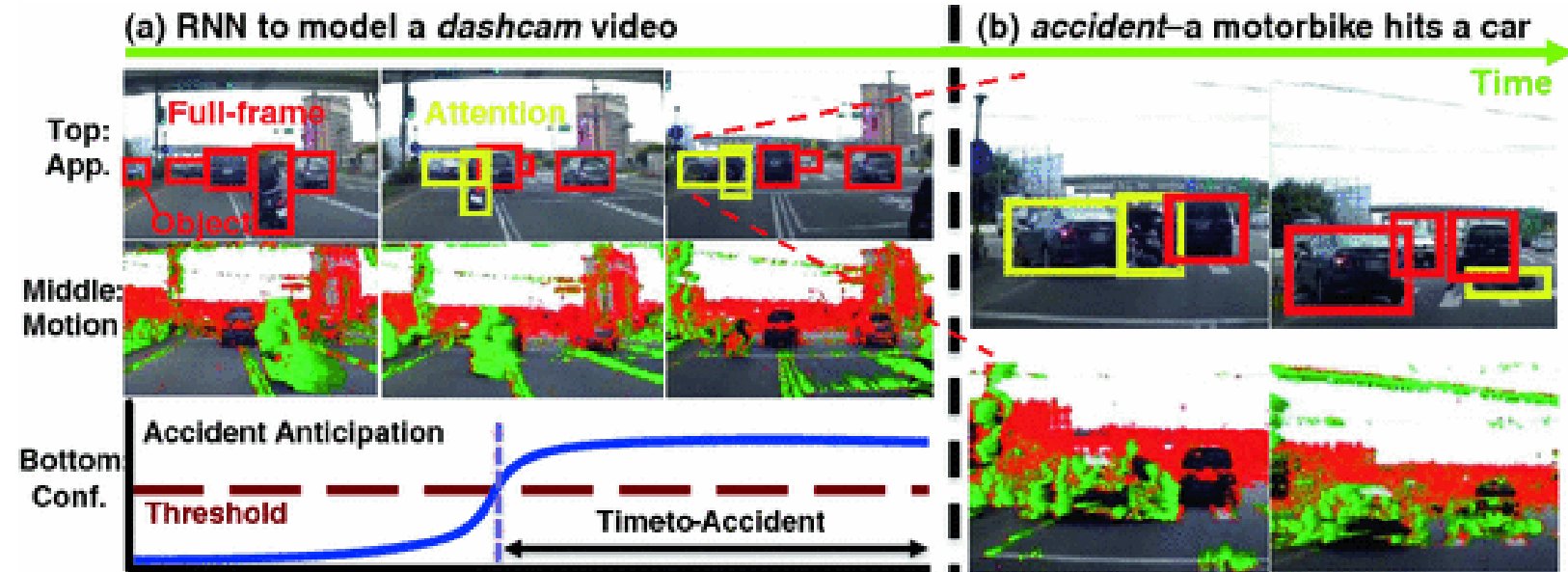
-- Qual a resposta dos exercícios?

-1-) **B**

2-) **B**

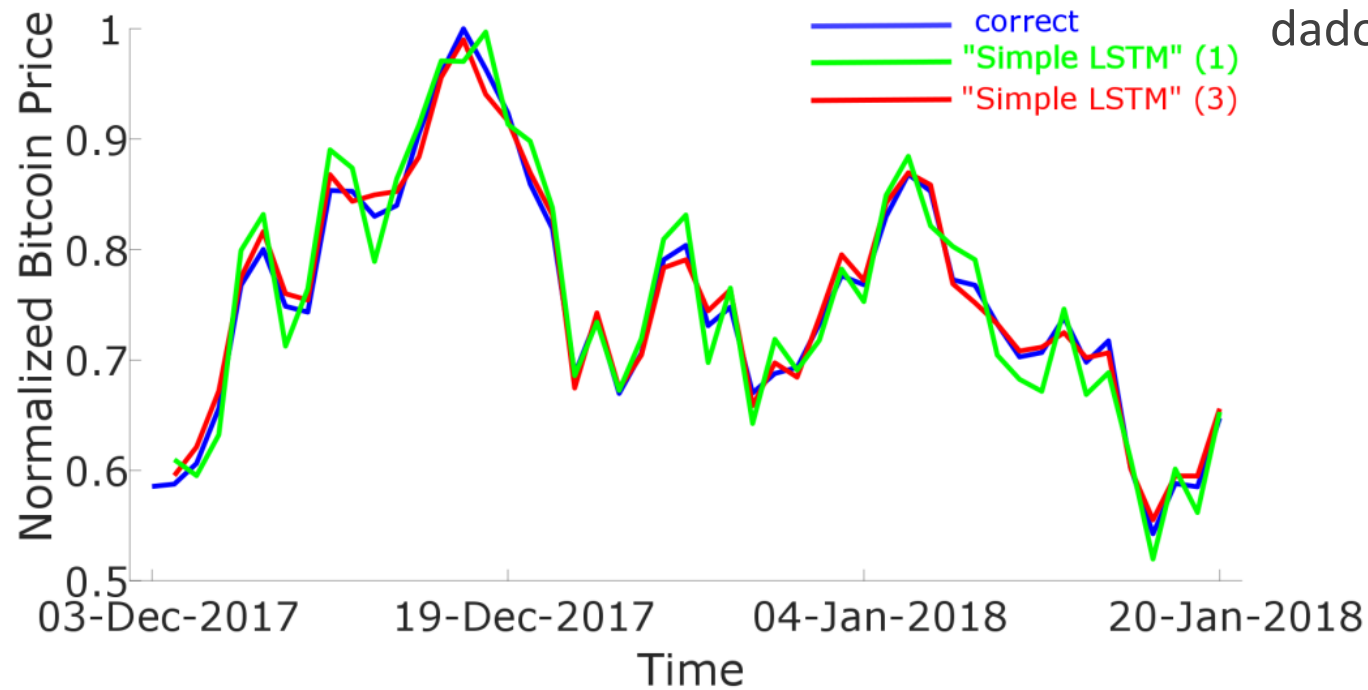
--

Exemplos



- Reconhecimento de Ciclistas
- Evitar Colisões de Carros
- Input => Sequência de Dados
- Output => Sequência de Eventos

Exemplos



- Séries Temporais -> BitCoin
- Recebendo dados 24h por dia
- Rodando em Loop, para cada no dado que chega

Passo-a-Passo – LSTM – Case BitCoin



1 -> Importar imagens

2 -> Visualizar

3 -> Variáveis

4 -> Teste

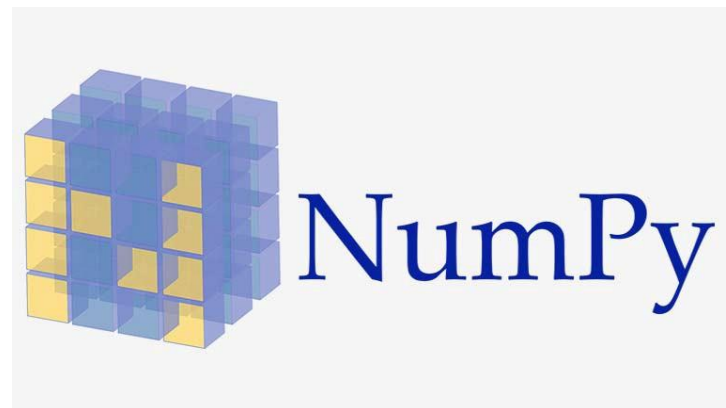
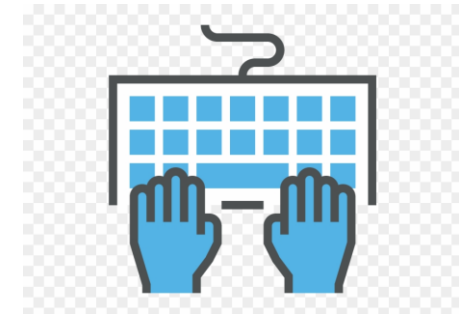
5 -> Discutir melhorias

...

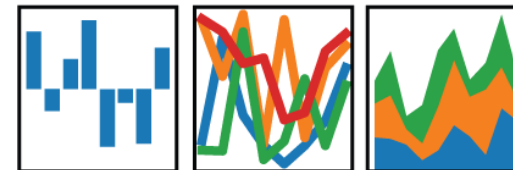
X -> Dominar MatriX



Hands-On



pandas
 $y_i t = \beta' x_{it} + \mu_i + \epsilon_{it}$



Revisão

- Aplicações
- Dúvidas
- Feedback...
 - O que achou da aula?
 - Como foi sua experiencia?
 - E os Slides? Agradáveis?




Referências Bibliográficas - Livros



Comece Pelo Porquê – Simon Sinek(2018), ISBN 978-85-431-0663-2

Introdução à Ciência de Dados – Fernando Amaral (2016), ISBN 978-85-7608-934-6

Deep Learning 

<https://books.google.com.br/books?isbn=0262035618> - Traduzir esta página



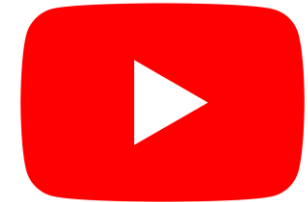
Ian Goodfellow, Yoshua Bengio, Aaron Courville - 2016 - Visualização - Mais edições

The hierarchy of concepts allows the computer to learn complicated concepts by building them out of simpler ones; a graph of these hierarchies would be many layers deep. This book introduces a broad range of topics in deep learning.

Referências Bibliográficas - YouTube

School of AI São Paulo –

<https://www.youtube.com/channel/UCcQgGC19k35ayQNsspyyBhQ>



Como passar em concursos, vestibulares e Enem? - Professor Pierluigi Piazzini –

<https://www.youtube.com/watch?v=0qxDU96J-do>

Obrigado!

Att,

Victor Venites

LinkedIn: <http://victorvenites.com/>

E-mail: contato@victorvenites.com

