

A primeira Escola presencial gratuita de Inteligência Artificial do Brasil



LSTM – RNN – DeepLearning

Prof. Esp. Victor Venites

AULA 15 - SÃO PAULO - 01/09/2020

Victor Venites

Formação –

- Bacharel Ciência da Computação
- Pós Engenharia Web
- Aluno Especial de Mestrado

Cursos –

Matemática Aplicada

Grupos –



Trabalhos –



Consultorias –



Até Aqui

Matemática, Python e Estatística –

- Exploração de Dados
- Introdução a Python
- Estatística Básica
- Regressão Linear
- Visualização de Dados
- Modelos de Machine Learning
- Transição de Carreira
- Perceptron

Exemplos –

- Hands-On – 101
- Slides
- Python



Inteligência Artificial



Roteiro –

- Covid-19 – Séries Temporais
- Support Vector Machine
- Séries Temporais
- Engenharia de Dados
- **LSTM – RNN - DeepLearning**

Objetivo

- Entender a LSTM com mais do que já se vê por ai
- Trazer uma analogia profunda
- Como comparar?
- Mostrar os complementos do meio acadêmico
- Passar um pouco da minha experiência em Consultorias
- Dar um caminho para o aluno estudar melhor
- Deixar o aluno apto para discutir
- Levantar questões... E responder a maioria!

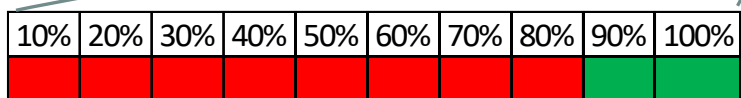
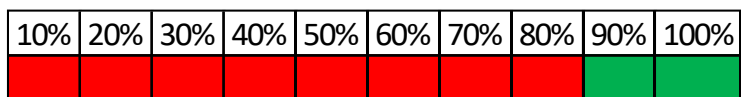


Material: GitHub / Slides e Código

Vídeo: YouTube - Live

Disclaimer

- Aos Psicólogos/Neurocientistas/Biólogos:
 - Os estudos desses modelos de Inteligência Artificial, entre outros, procuram imitar conceitos e atividades de animais da biologia do mundo para tentar resolver problemas do cotidiano.
 - Sabemos que o cérebro humano é muito mais complexo do que é falado aqui, e por isso dentro da capacidade humana procuramos criar ML com tarefas específicas e regras simplificadas.



- Aos Profissionais das Diversas Áreas:
 - Aos profissionais que enfrentam tecnologias que automatizam suas tarefas, vejam isso como uma oportunidade de explorar novas possibilidades de atuação/criatividade.
 - Estudem para criar e não apenas operar tecnologias e tarefas simples.

LSTM – Artigos Históricos

Long Short-Term Memory

Shmdhuber e Hochreiter

(1995 e 1996)

Long Short-Term Memory

Shmdhuber e Hochreiter

(1997)

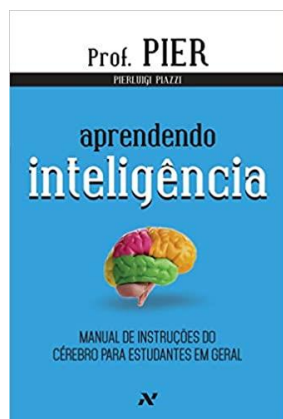
Termos Relacionados:

- Máquinas de Estado Finito
- Learning Sequential Tasks
- Alternatives to BackPropagation

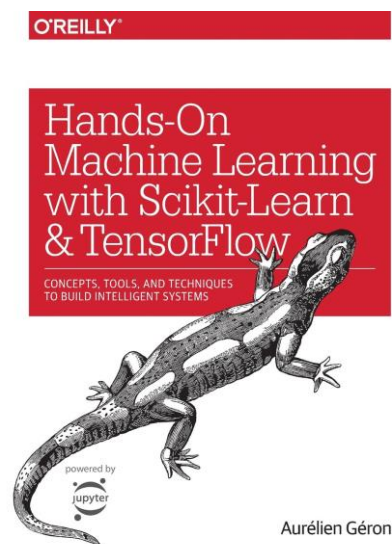
Voltando ao Básico

Nesta Aula:

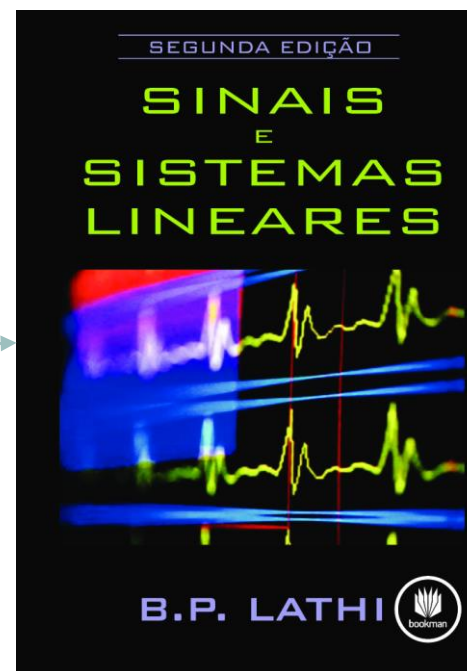
- Inteligência Natural
- Séries Temporais
- Perceptron
- RNN



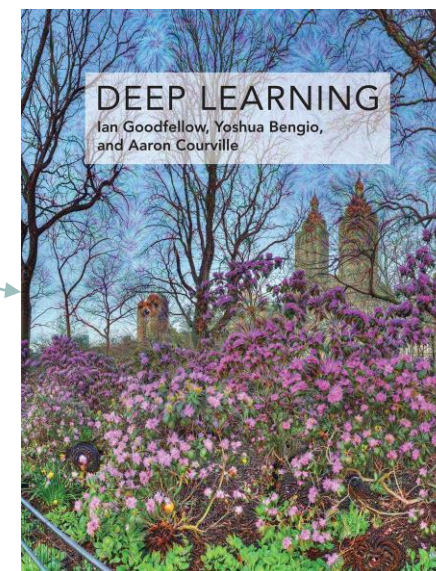
1º Inteligência Natural



2º Perceptron



3º Series Temporais

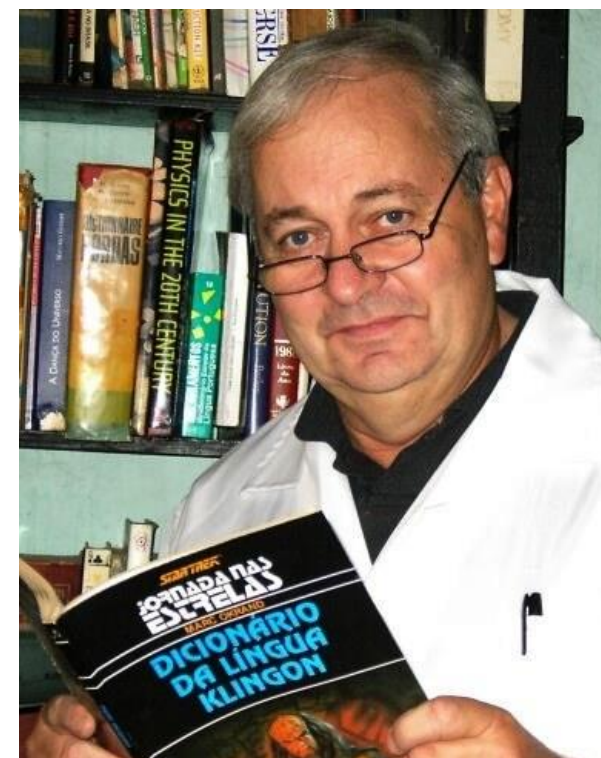


4º DeepLearning

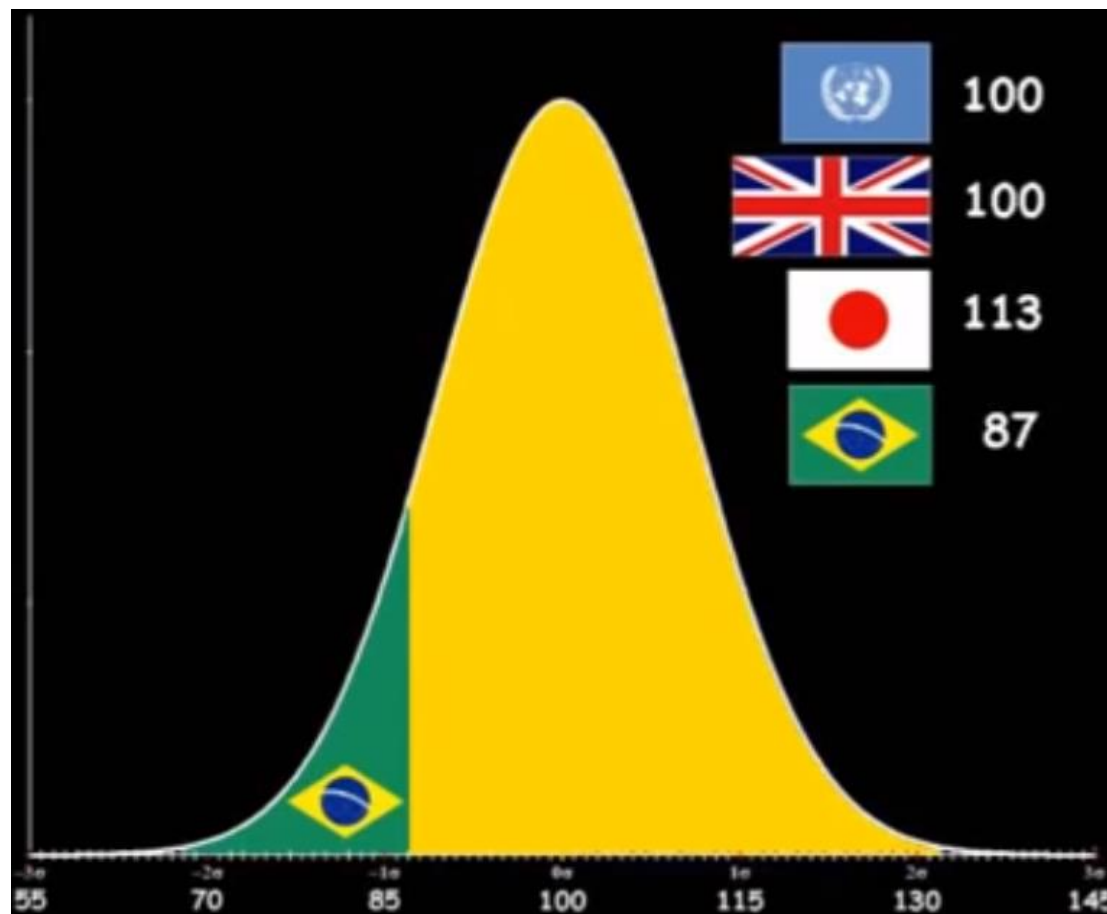
Inteligência Natural – Pierluigi Piazzi



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

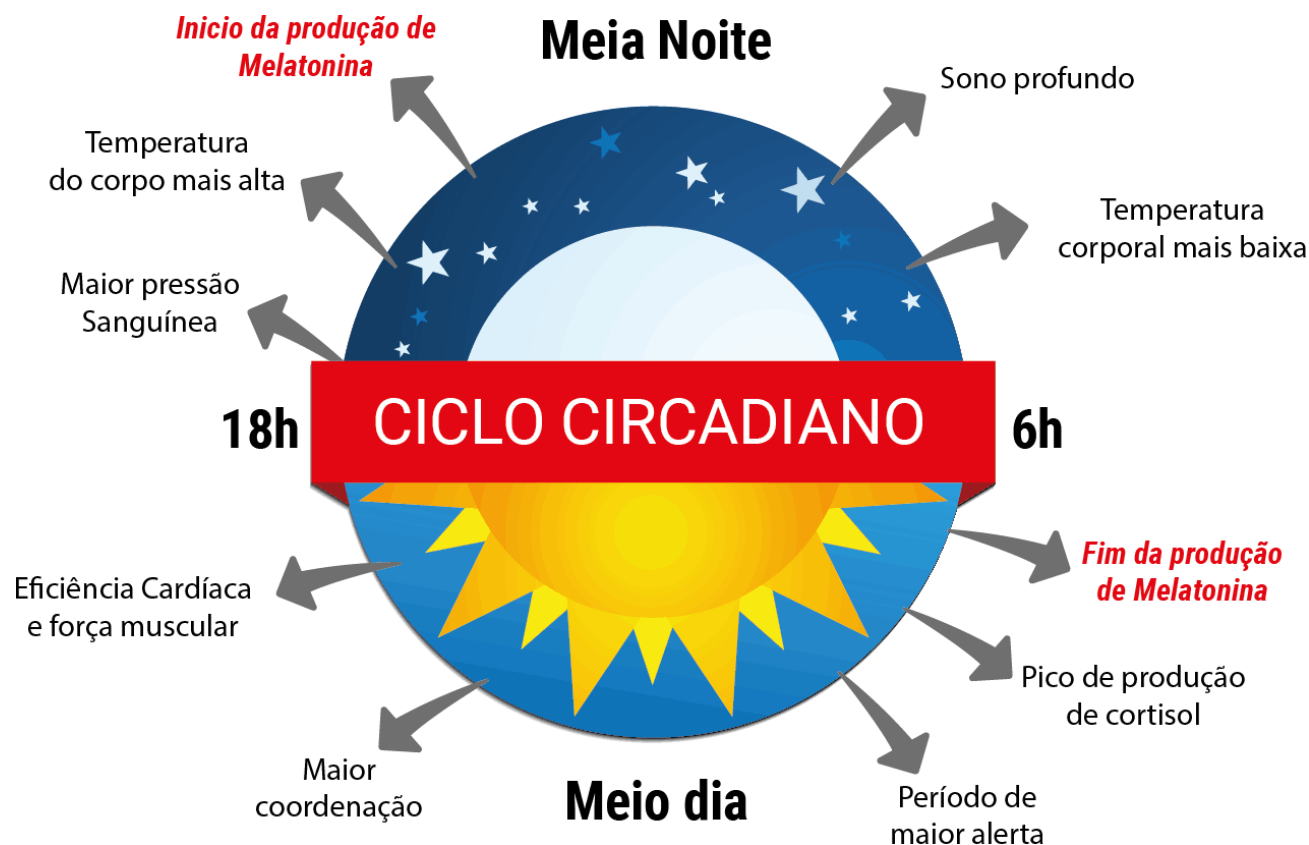


“Inteligência se Aprende”



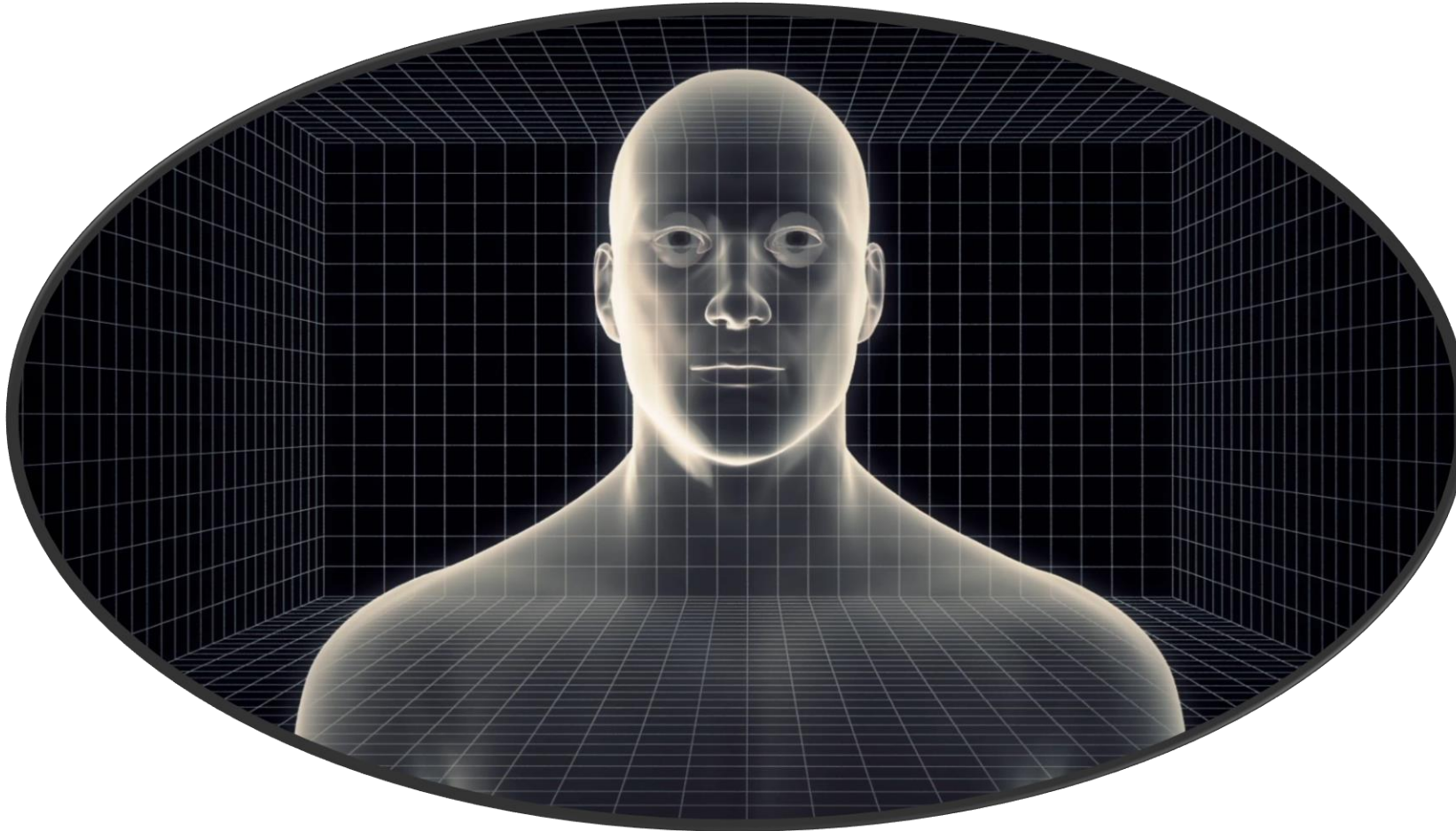
- Ao vermos a média de QI neste gráfico e vendo todo conteúdo do Professor Pier ele diz “Inteligência se Aprende”.
- Pierluigi: Começou com testes de 132 de QI e alguns anos depois tinha testes com 184 de QI.
- Mensa Brasil: Associação de pessoas com alto QI no país.

Como Aprendemos?

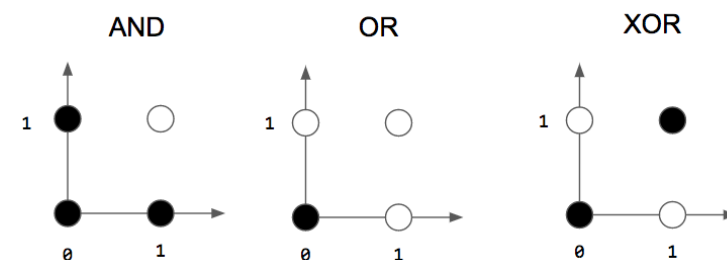
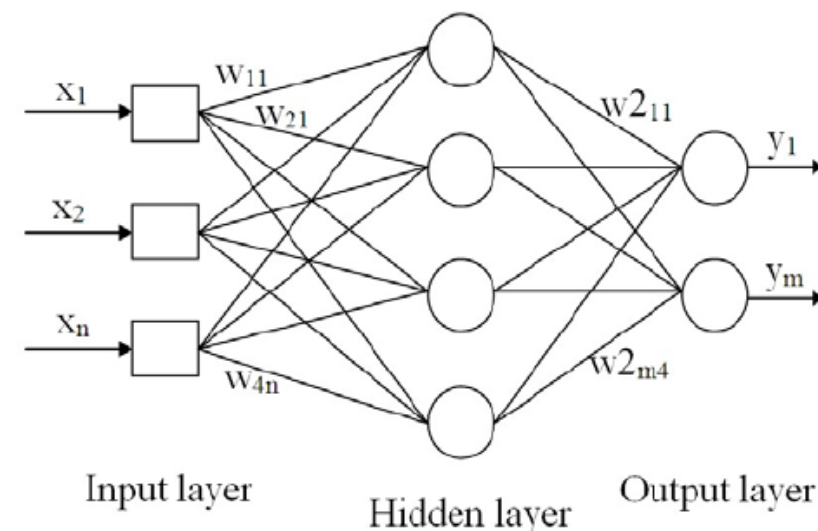
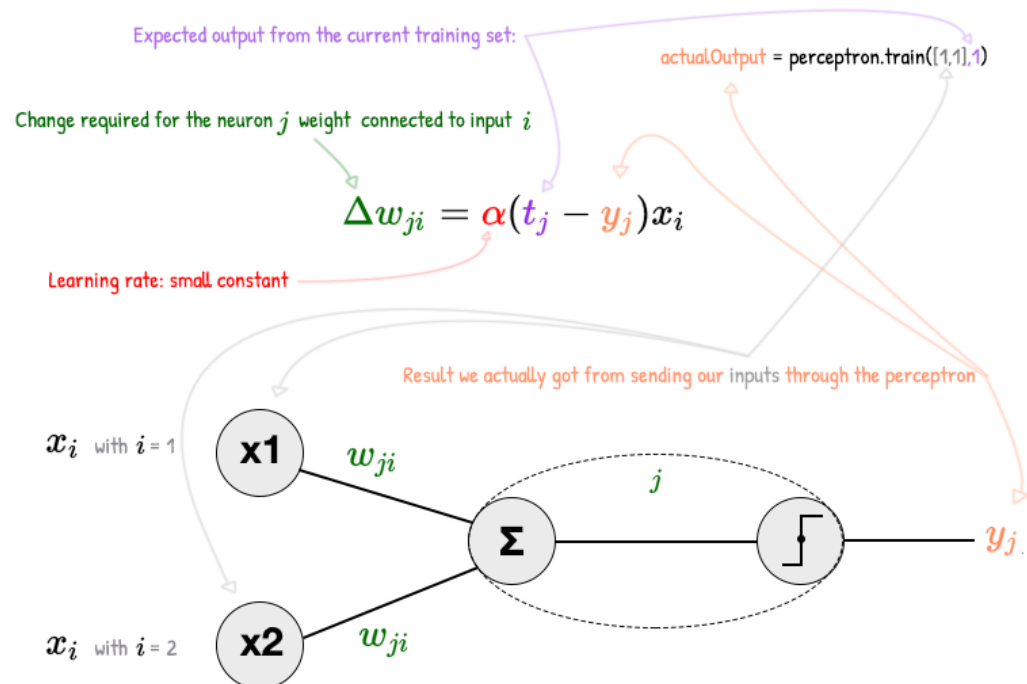


- Analogia do Ciclo com a LSTM
- “Aula dada, aula estudada hoje”
- Inteligência Natural
- Ciclo de aprendizado do cérebro humano no mesmo dia
- Nosso LOOP de aprendizado

O que Acontece no Sono?



Neurônio - Perceptron



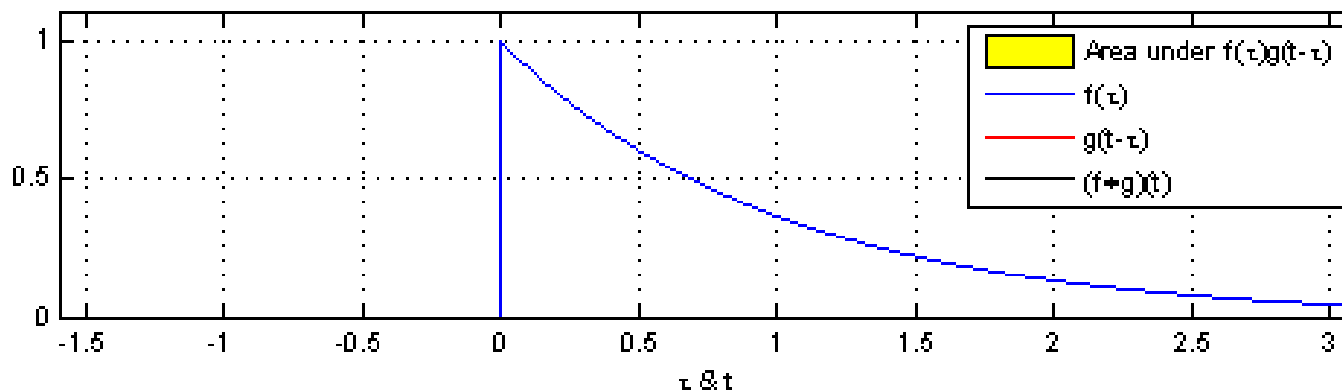
- Aula 13/2019 - 18/06 - Parte 2 - Introdução a redes Neurais e Perceptrons

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

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 |

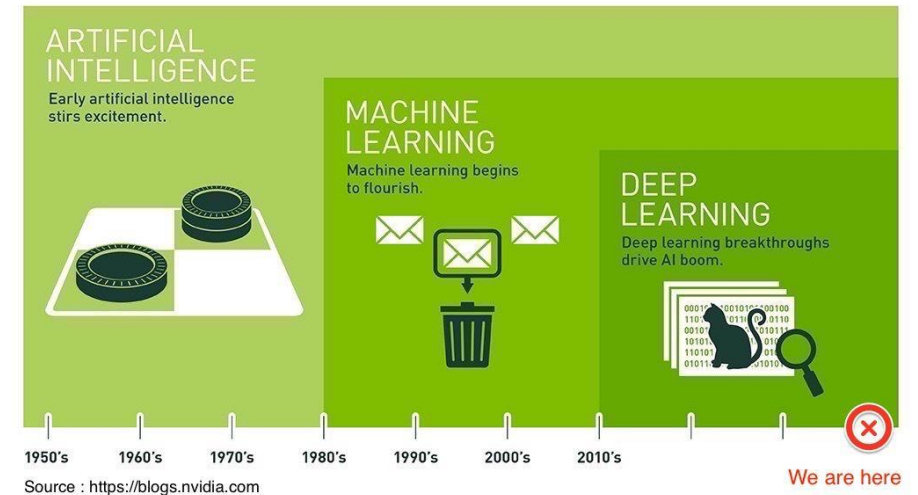


- Aula 09/2020 - Escola Livre de IA - Séries Temporais

<https://www.youtube.com/watch?v=tYmvx7HvqKg&list=PLxalRiHfWZGfcZAr1353ZkxEByg-fS3bh&index=4>

DeepLearning

- Modelos de Machine Learning com foco em problemas mais complexos e específicos;
- Onde uma Rede Neural mais generalista perde em performance, tempo, acurácia, etc..
- Exemplos: RNN; AutoEncoders; Reinforcement Learning; CNN.



- Aula 16 - Parte 1 - DeepLearning – AutoEncoder

https://youtu.be/ep_j7pvFg9c?list=PLxaLRiHfWZGc1InNs_oU2av-z1SttRzOic

- Aula 16 - Parte2 - DeepLearning – AutoEncoder

https://youtu.be/v-7vLPR2Azc?list=PLxaLRiHfWZGc1InNs_oU2av-z1SttRzOic



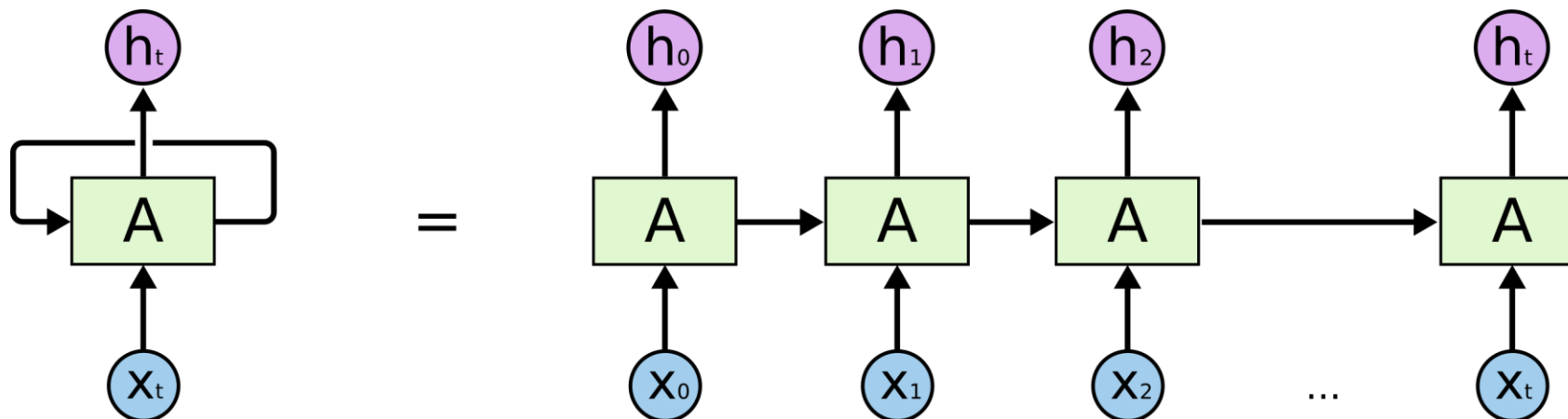
- Aula 21 - Parte1 - Reinforcement Learning

https://youtu.be/nY22xKsa_xU?list=PLxaLRiHfWZGc1InNs_oU2av-z1SttRzOic

- Aula 21 - Parte 2 - Reinforcement Learning

https://youtu.be/vNI6SgWLnOQ?list=PLxaLRiHfWZGc1InNs_oU2av-z1SttRzOic

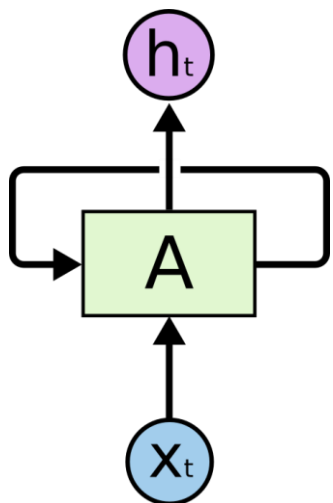
Rede Neural Recorrente



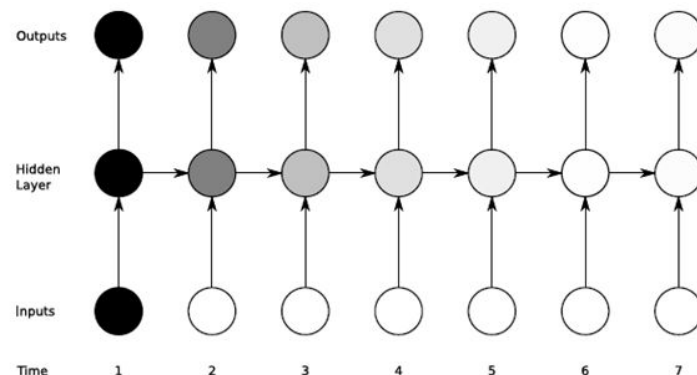
- Os Neurônios agora funcionam de forma Sequencial, como um loop de atualização contínua dos pesos e previsões. Assim, pode atender a demandas de stream de dados ininterruptos;
- BackPropagation muito pesado para rodar em toda base;
- Dados muito antigos podem influenciar em uma cultura atual de dados;
- Variações grandes e imprevistas nos novos dados podem fazer o modelo de perder um pouco até se adaptar.

RNN – Vanishing e Exploding

Recurrent Neural Networks



- Vanishing gradient problem



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

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

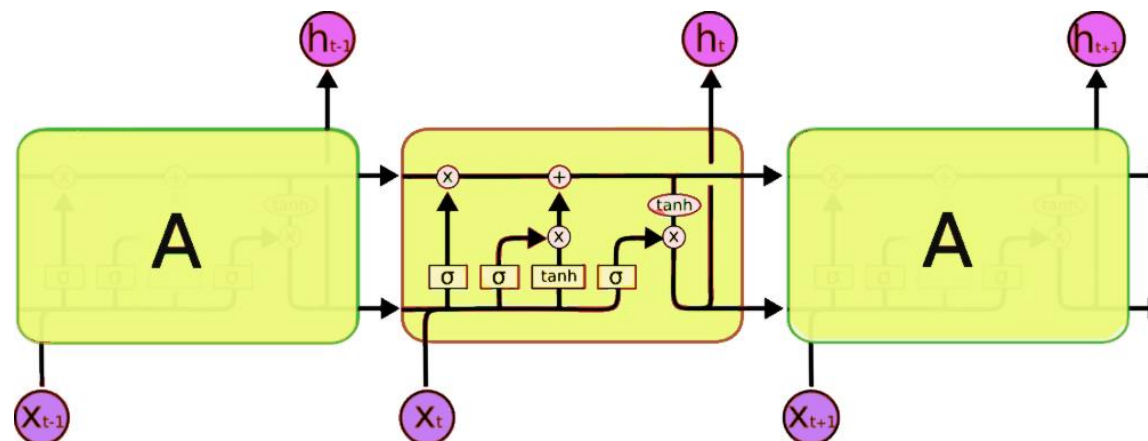
$$b_h^t = \theta_h(a_h^t)$$

Sensitivity decay exponentially over the time

E na LSTM?

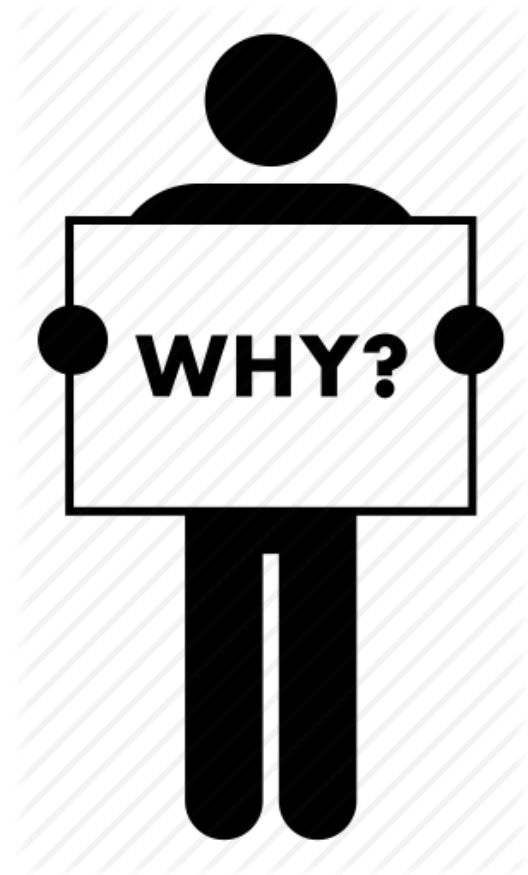


- Adaptamos a ideia de ciclo de aprendizado humano criando uma alegoria em versão artificial;
- A Rede Neural do LSTM, parecido com um sonho, decide se vale a pena lembrar de algumas experiencias e decisões tomadas.
- Além de ponderar uma previsão baseada na memória de Curto e Longo Prazo.

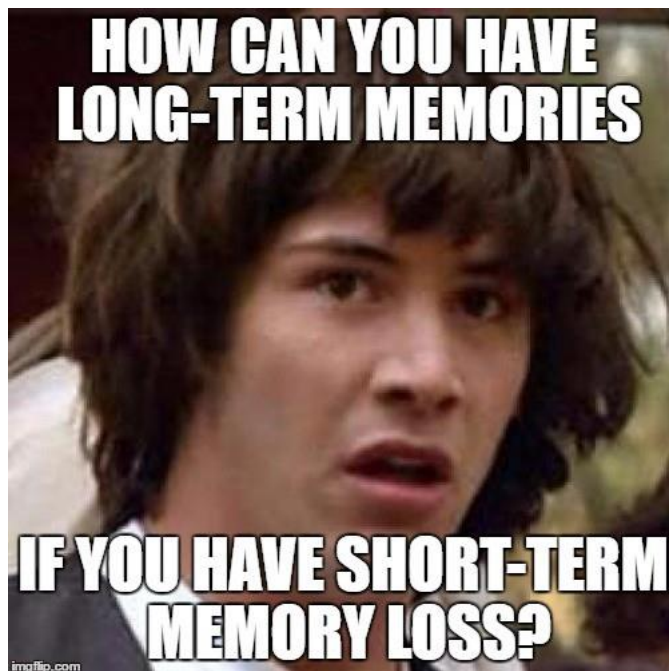


Por quê LSTM?

- Para resolver problemas diversos com soluções diversas
- 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



O que é LSTM?



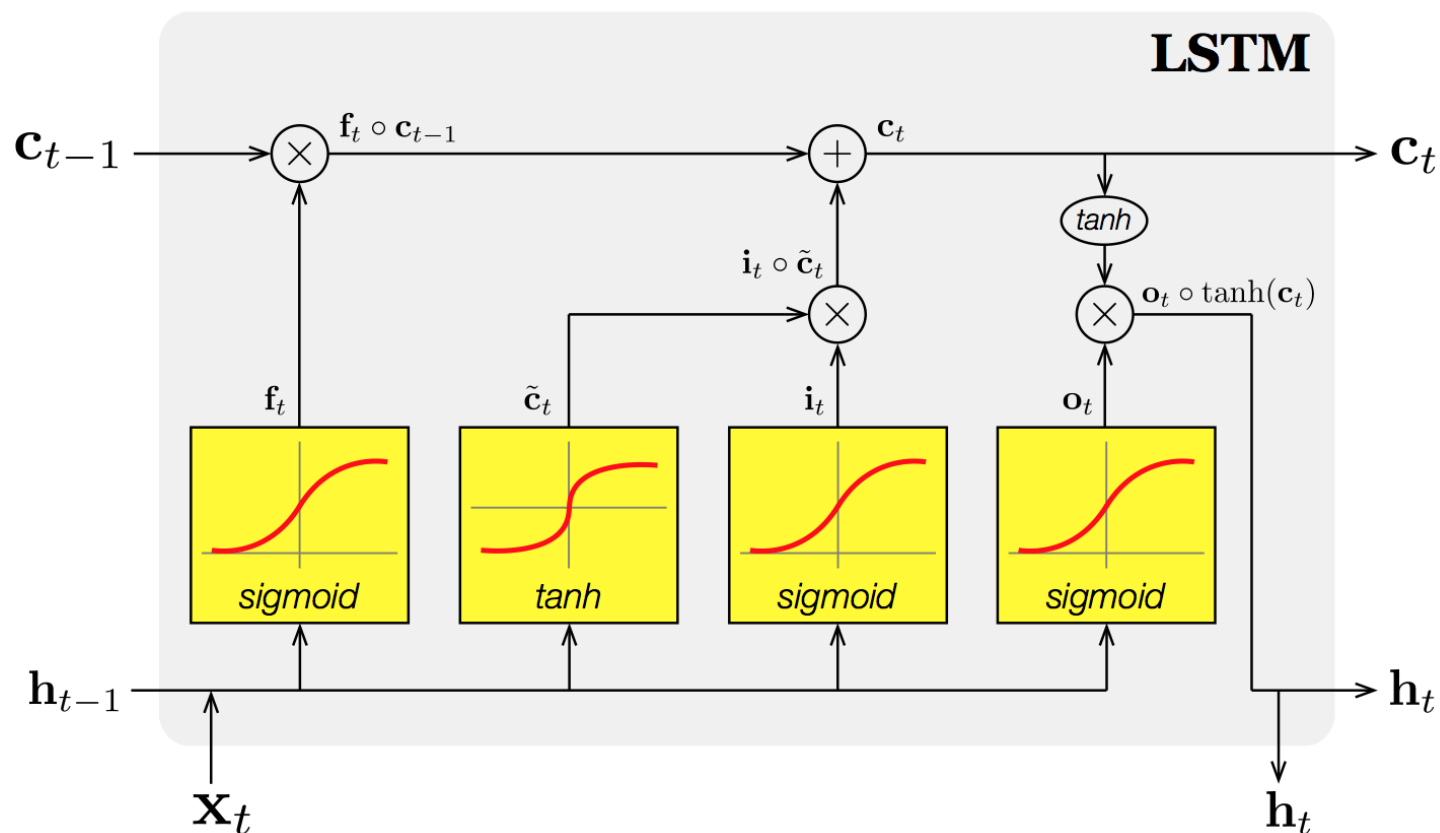
Uma Rede Neural Recorrente com Neuronios Perceptron rodando série sequencial.

Conceitos:

- Memória
- Estado
- Decisão
- Pesos e Bias



Como a Célula LSTM Funciona?



Gating variables

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

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

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

Candidate (memory) cell state

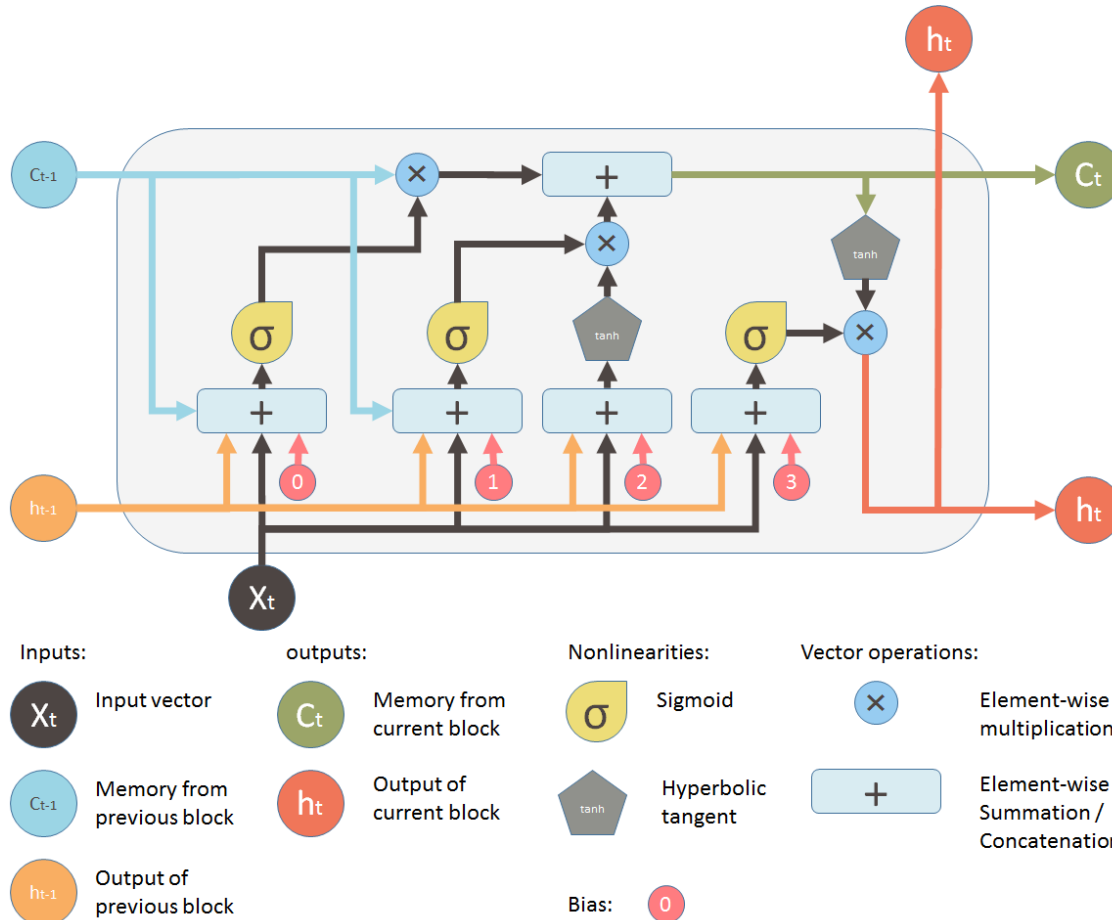
$$\tilde{c}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c)$$

Cell & Hidden state

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t$$

$$h_t = o_t \circ \tanh(c_t)$$

Arquitetura LSTM?



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

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

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

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

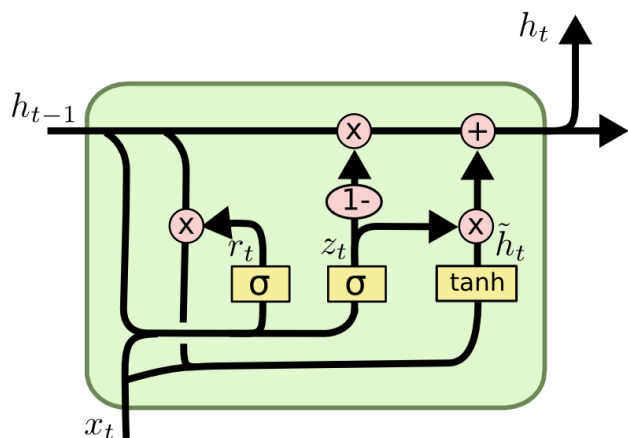
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

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

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

LSTM Simplificada = GRU

- GRU => Gated Recurrent Unit
- Por Kyunghyun Cho em 2014

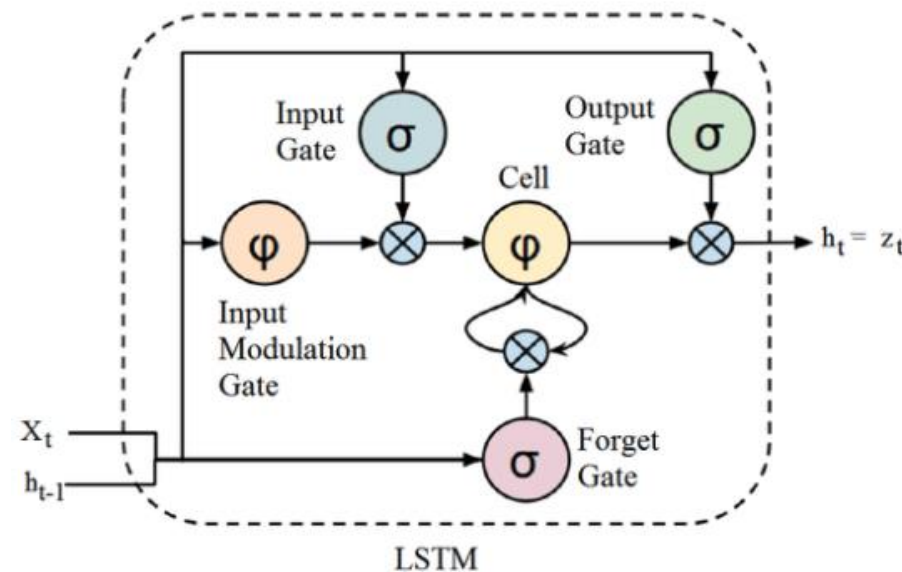


$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

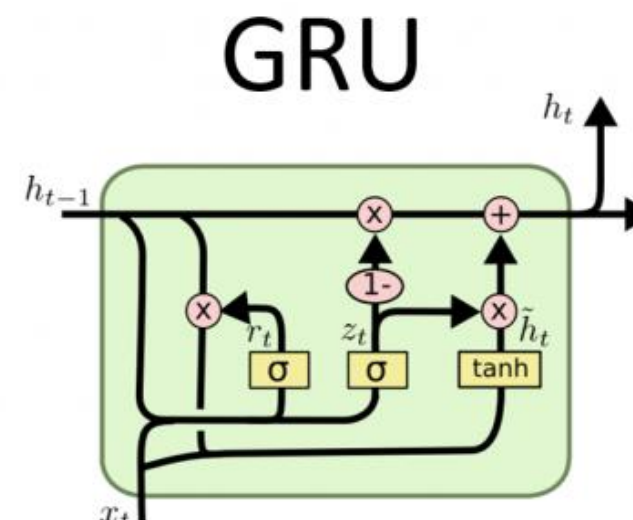
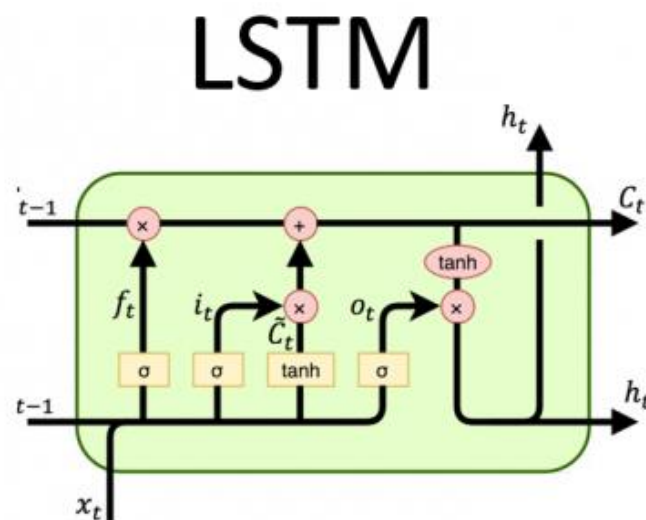
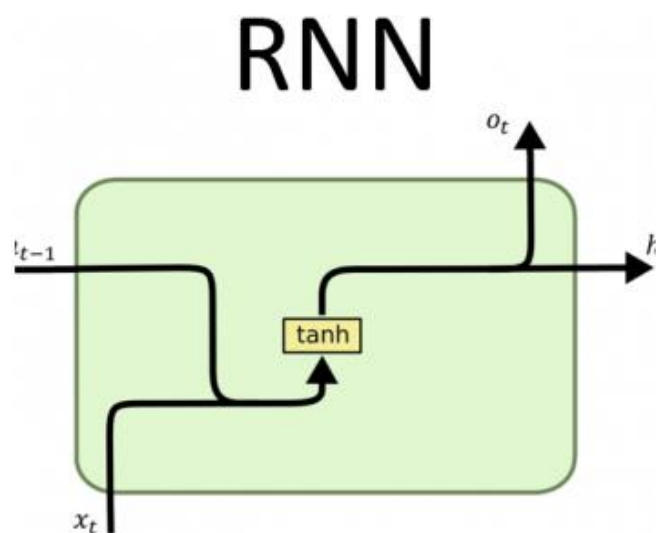
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

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

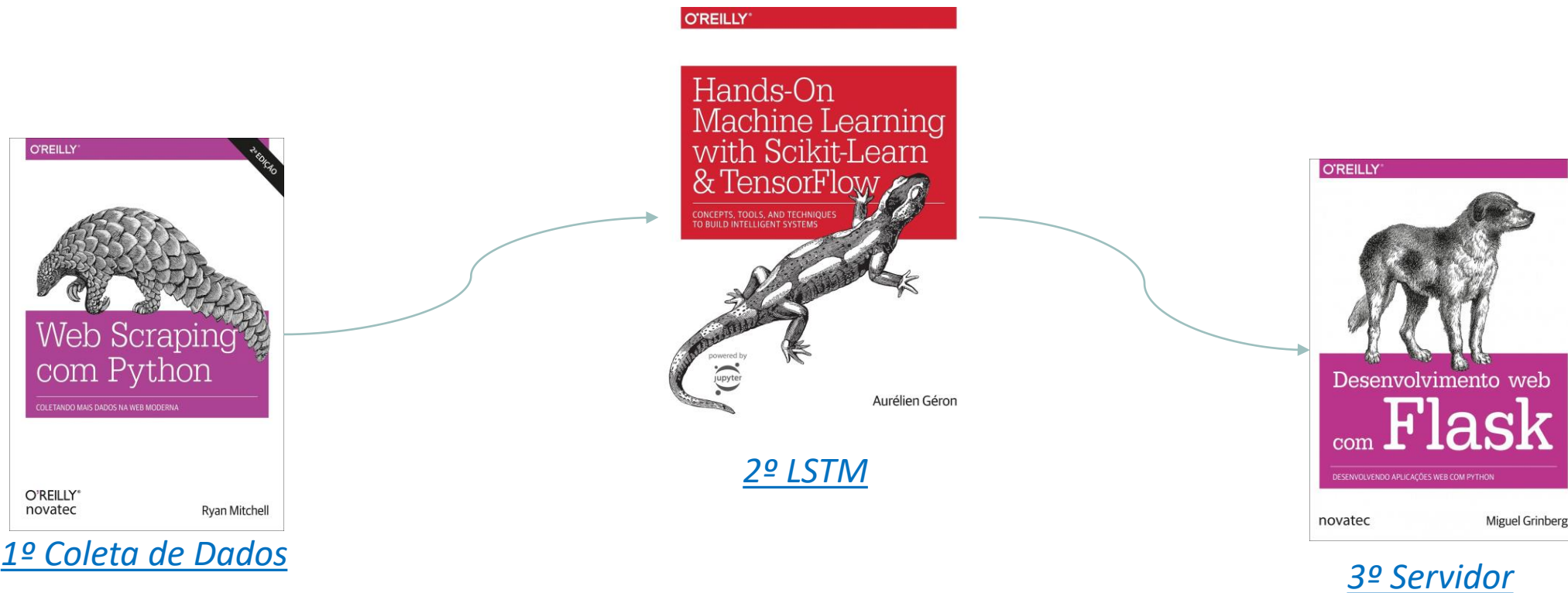
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$



Resumo da Evolução

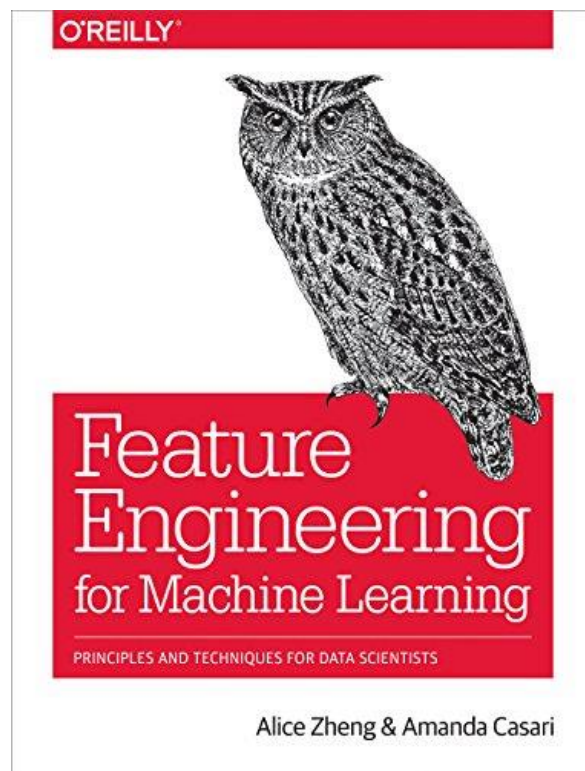


E o PipeLine?



- Junto ao PipeLine veremos as Features Engieenring que merecem Slides mais detalhados.

Feature Engineering – Imagem e Áudio



- Imagem: Convolução; Técnicas de filtro de Imagem; Blur, Recortes de Imagem.

Aula 17 - Parte 1 - Redes Neurais Convolucionais –

<https://www.youtube.com/watch?v=ogTOo6xuc04&list=PLxaLRiHfWZGc1InNsoU2av-z1SttRzOic&index=11>

Aula 17 - Parte 2 - Rede Neural Convolutonal –

<https://www.youtube.com/watch?v=TSYwzVWQVWk&list=PLxaLRiHfWZGc1InNsoU2av-z1SttRzOic&index=12>

- Áudio: FFT(Fast Fourier Transform)

Exemplos - Texto



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

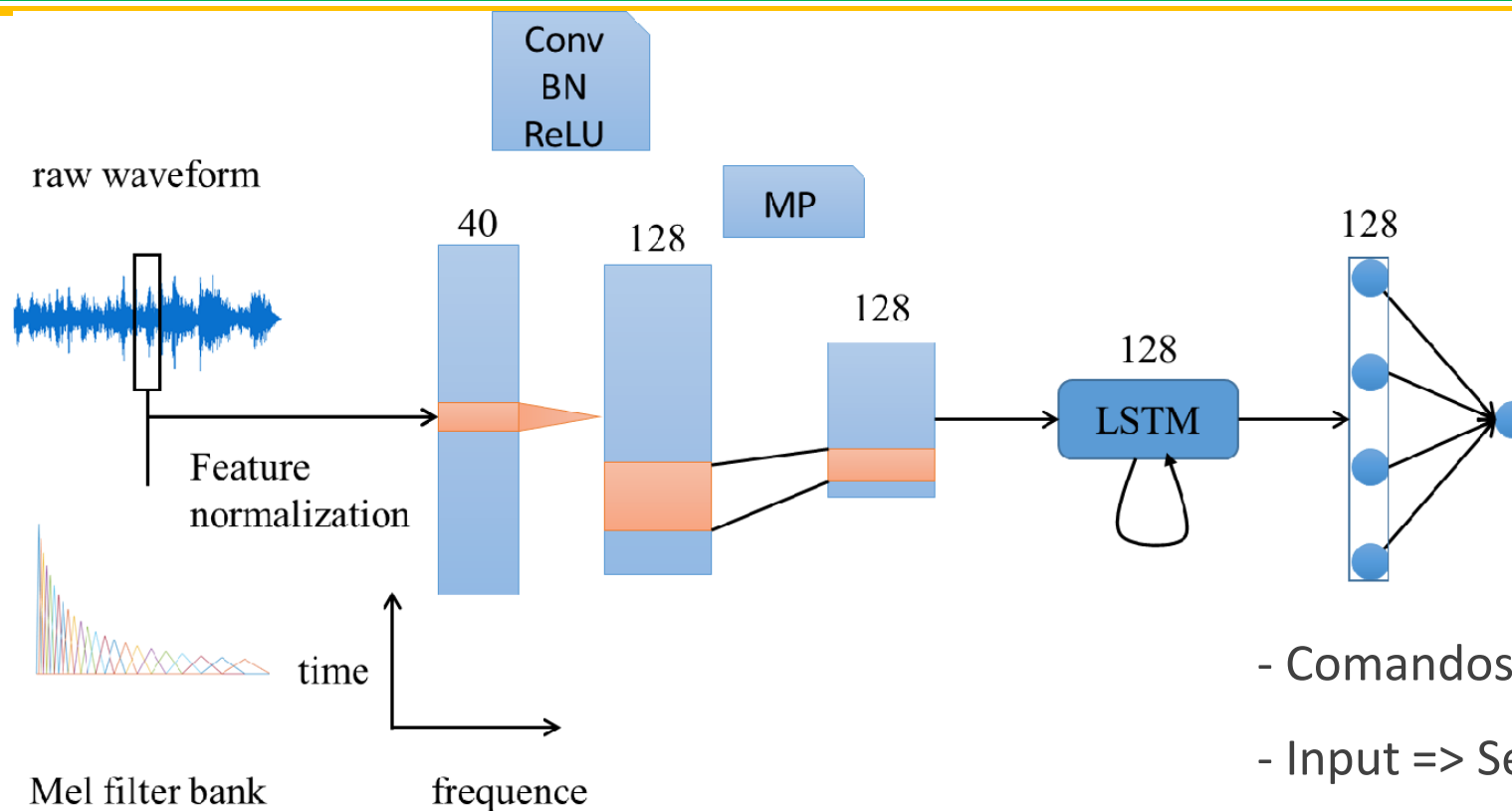
- Corretor Ortográfico

-- Qual a resposta dos exercícios?

-1-) **B**

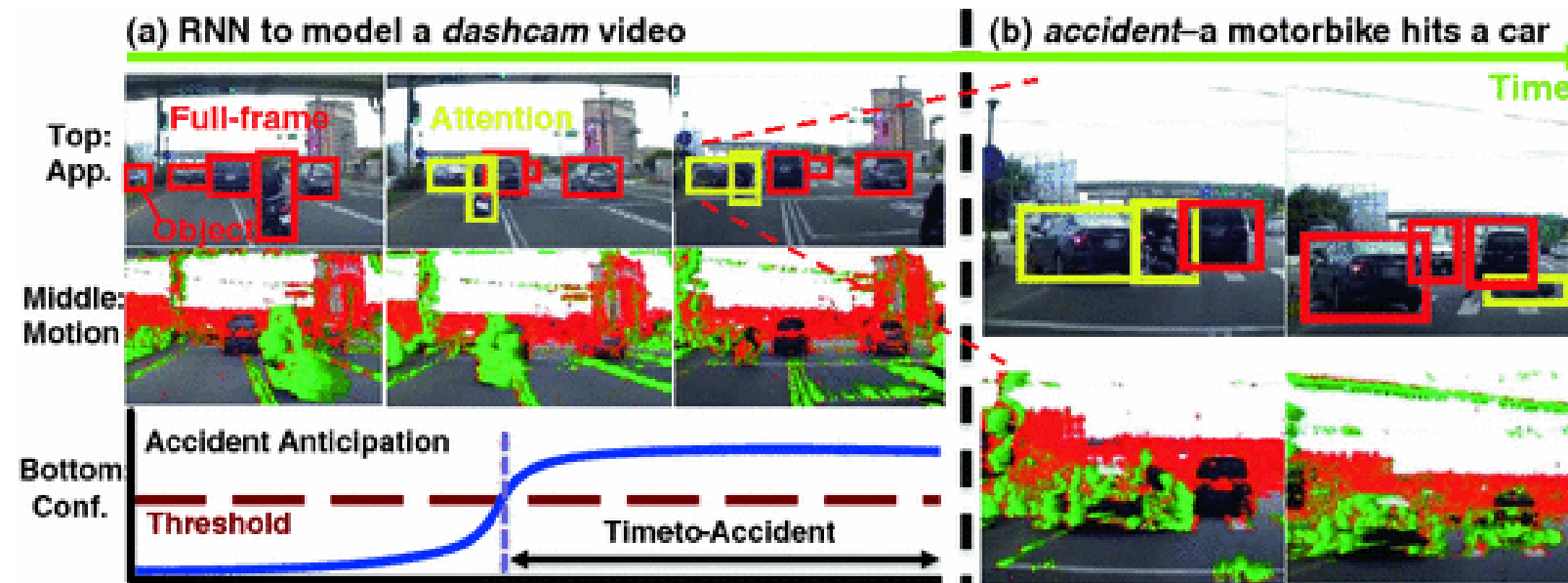
2-) **B**

Exemplos - Áudio



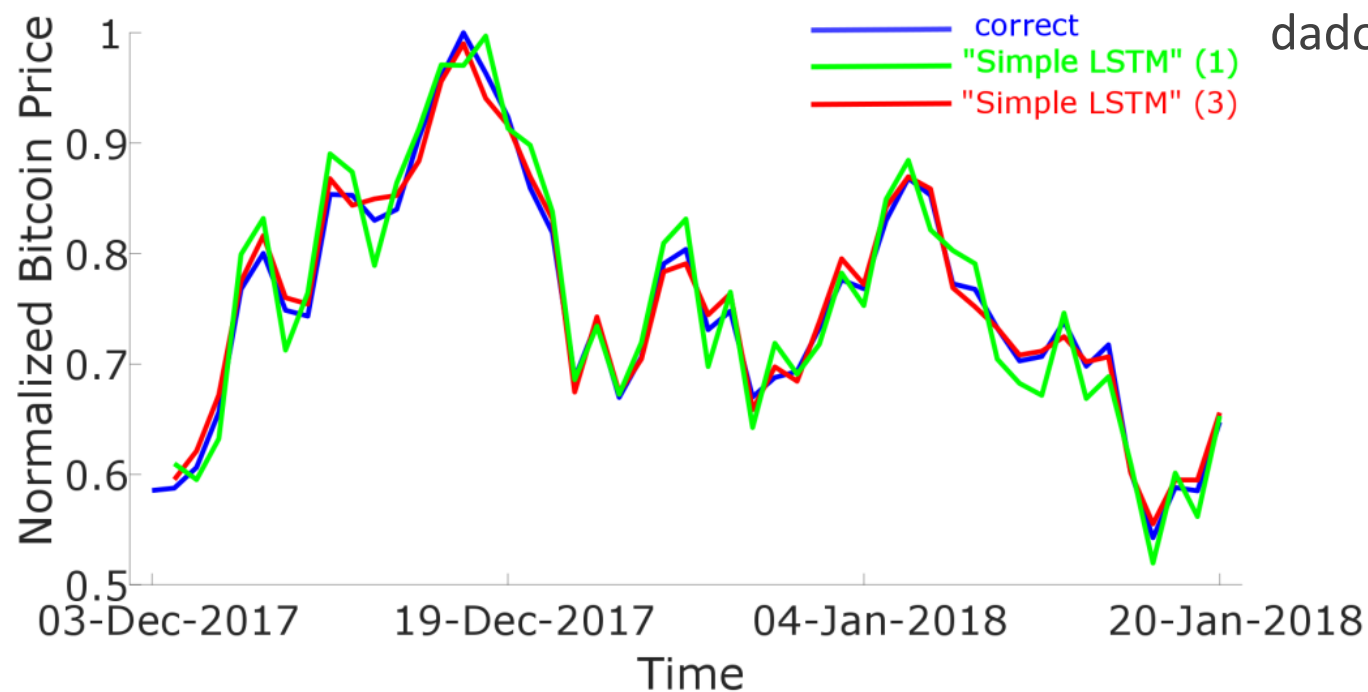
- Comandos de Vós
- Input => Sequência de Dados
- Output => Comando de Voz(celular)

Exemplos - Vídeo



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

Exemplos - BitCoin



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

Onde Adquirir Bases de Dados?



Bitcoin Historical Data - Kaggle

- <https://www.kaggle.com/mczielinski/bitcoin-historical-data>
- Base utilizada para aula de hoje
- Tem 240MB por isso não foi possível colocar no GitHub



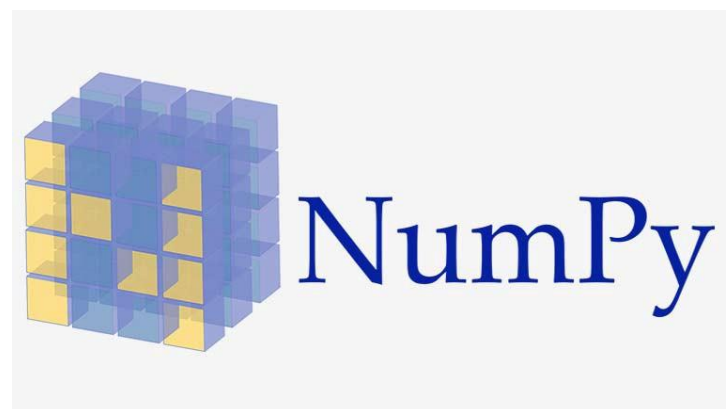
Passo-a-Passo – LSTM – Case BitCoin



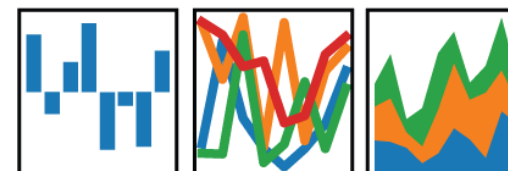
- 1 -> Baixar Bitcoins
- 2 -> Visualizar Gráficos
- 3 -> Tratar e Filtrar
- 4 -> Testar modelo LSTM
- 5 -> Discutir melhorias
- ...
- X -> Dominar MatriX



Hands-On – Python



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?

Exercícios



- Quando você usa DeepLearning?
- Por quê usar RNN comuns podem ser mais um problema do que solução?
- Onde você isso sendo aplicado na sua empresa?
- Onde você um mercado promissor para as RNNs?
- Dê uma diferença entre LSTM e GRU.

Referências Bibliográficas - YouTube

Professor Victor Venites in School of AI – SP –

<https://www.youtube.com/playlist?list=PLxaLRiHfWZGc1InNsoU2av-z1SttRzOic>

Prof. Victor Venites em Escola Livre de IA –

<https://www.youtube.com/playlist?list=PLxaLRiHfWZGfcZAr1353ZkxEByg-fS3bh>

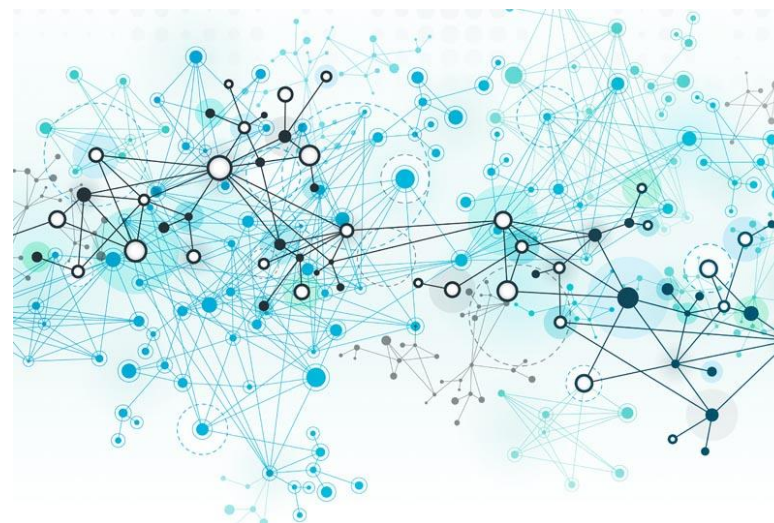


Como passar em concursos, vestibulares e Enem? - Professor Pierluigi Piazzzi –

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

Obrigado!

Att,
Victor Venites



E-mail: contato@victorvenites.com

 LinkedIn: <https://www.linkedin.com/in/victor-venites/>