O que é IA?

- Aula 1
  - O que é IA?
  - Mapa da IA
  - Matemática?
  - Exemplos
- Aula 2 (em aberto)
  - LLMs
  - Transformers & Prompts
  - Exemplos (BERT?)
- Aula 3 (em aberto)
  - CNNs
  - Visão computacional
  - Exemplos (ResNet ou Yolo?)

Mapa da IA

Inteligência Forte

O que é:
Sistema que não apenas **simula** inteligência, mas realmente "entende".

Teria **consciência**, **intencionalidade**, **autoconsciência**, **experiência subjetiva**.
Capacidade de aprender qualquer coisa que um humano aprende.

É um conceito **hipotético**, não existe hoje.

Seria um "ser" (artificial?) com mente própria.

Inteligência Fraca

O que é:
Sistemas criados para executar **tarefas específicas**.
Não possuem consciência, entendimento profundo ou intenção própria.

Exemplos:
Chatbots
Sistemas de recomendação
Reconhecimento de fala
Modelos de visão (YOLO, ResNet)
LLMs atuais (GPT, LLaMA)

É competente somente no domínio para o qual foi treinada. Também chamada de "inteligência estreita".

Inteligência Forte

Você!

Inteligência Fraca

Nossas aulas!

Inteligência Forte **Inteligência Fraca**

Você! AGI? Nossas aulas!

O que é AGI:
Uma IA capaz de **realizar qualquer tarefa intelectual** que um ser humano pode realizar.
Tem **generalização**, **razão**, **planejamento**, **aprendizado universal**.

Não necessariamente precisa ter “consciência” - isso pertence ao debate da *inteligência forte*.

AGI ainda não existe, embora alguns pesquisadores argumentem que estamos chegando perto.

Inteligência Forte

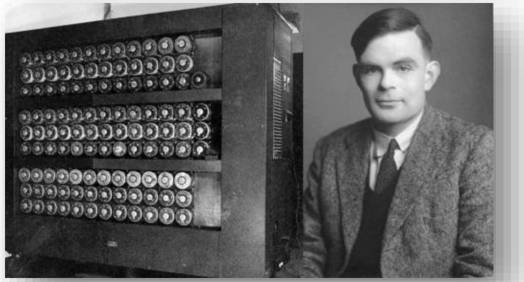
Inteligência Fraca

Você!

Nossas aulas!

Alan Turing propôs uma maneira chamada Teste de Turing, que também funciona como uma definição de inteligência.

Se um ser humano não conseguir distinguir entre uma pessoa real e um sistema de computador em um diálogo baseado em texto, o sistema é considerado inteligente.



https://github.com/microsoft/AI-For-Beginners/blob/main/translations/br/README.md

Inteligência Forte

Você!

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

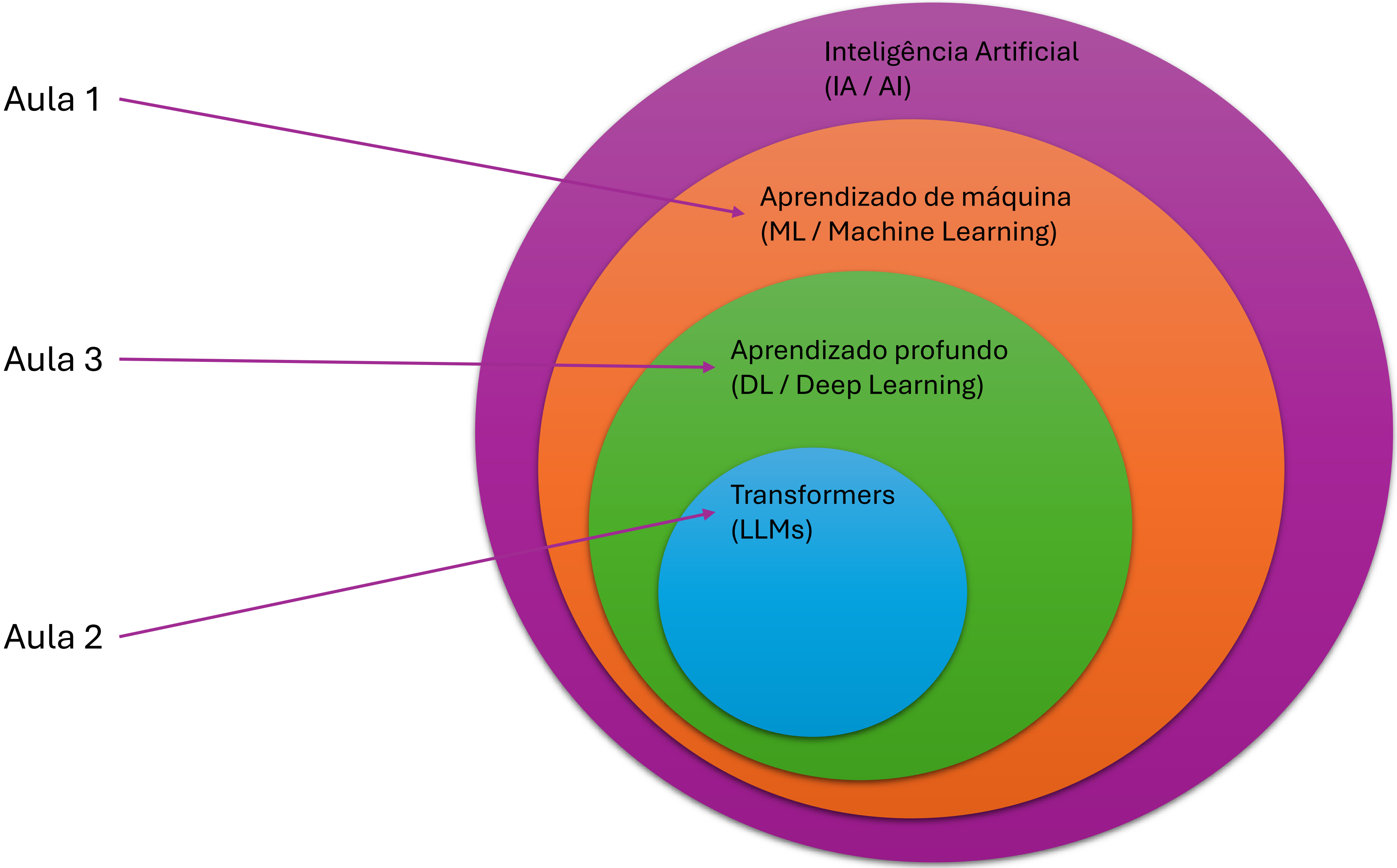
I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

Inteligência Fraca

Nossas aulas!

https://github.com/microsoft/AI-For-Beginners/blob/main/translations/br/README.md

Inteligência Fraca

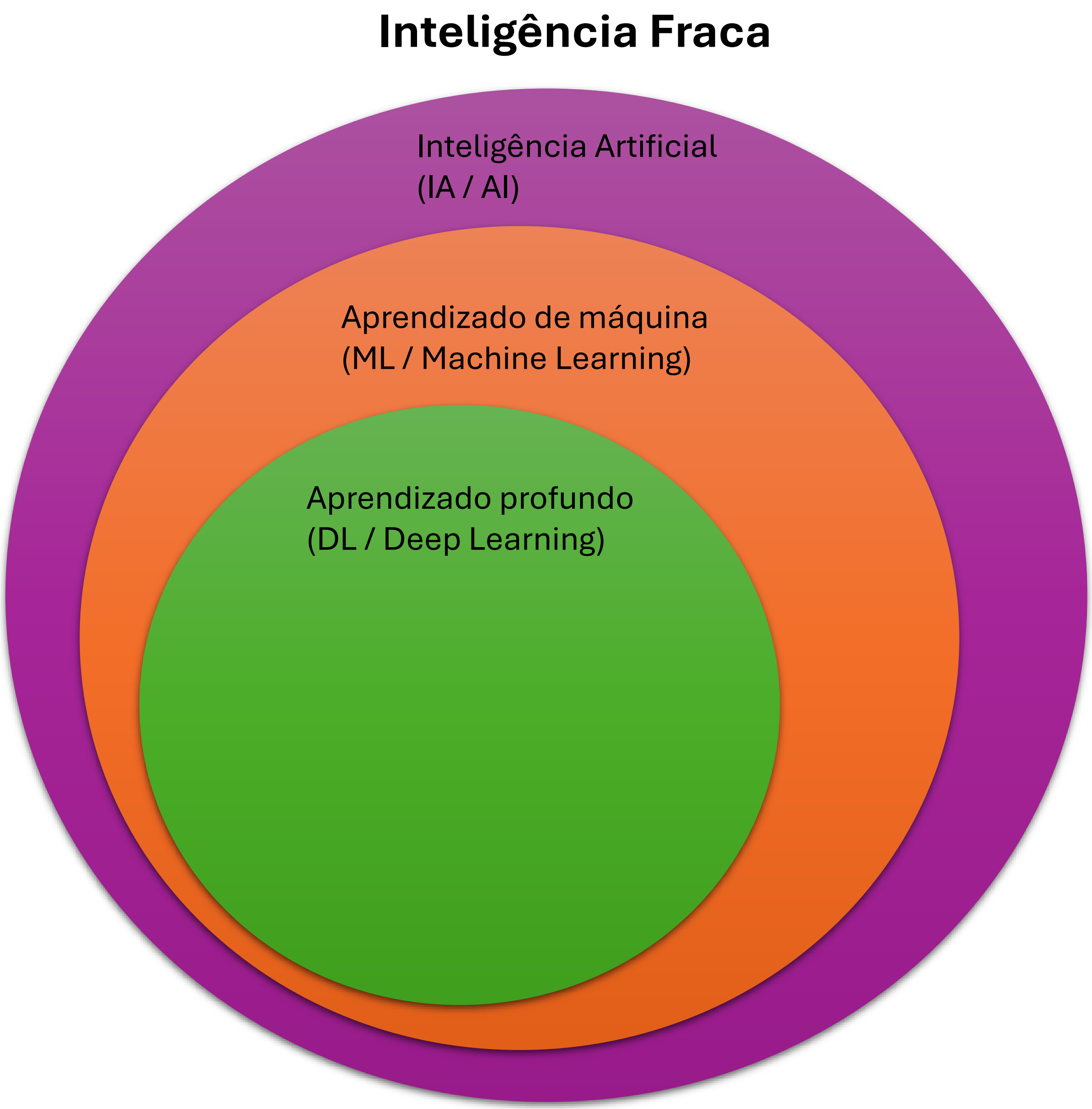


- Aprendizado de máquina
  - Regressão (Linear)
  - Classificação (SVM)
  - Clusterização (K-means)
  - Random Forest
  - GBDT / XGBoost
  - Reinforcement Learning (parte em ML, parte fora)

Inteligência Fraca

- Aprendizado de máquina
  - Regressão (Linear)
  - Classificação (SVM)
  - Clusterização (K-means)
  - Random Forest
  - GBDT / XGBoost
  - Reinforcement Learning (parte em ML, parte fora)

- Aprendizado profundo
  - CNNs (visão computacional)
  - RNNs, LSTMs, GRU
  - Transformers
  - Autoencoders
  - GANs

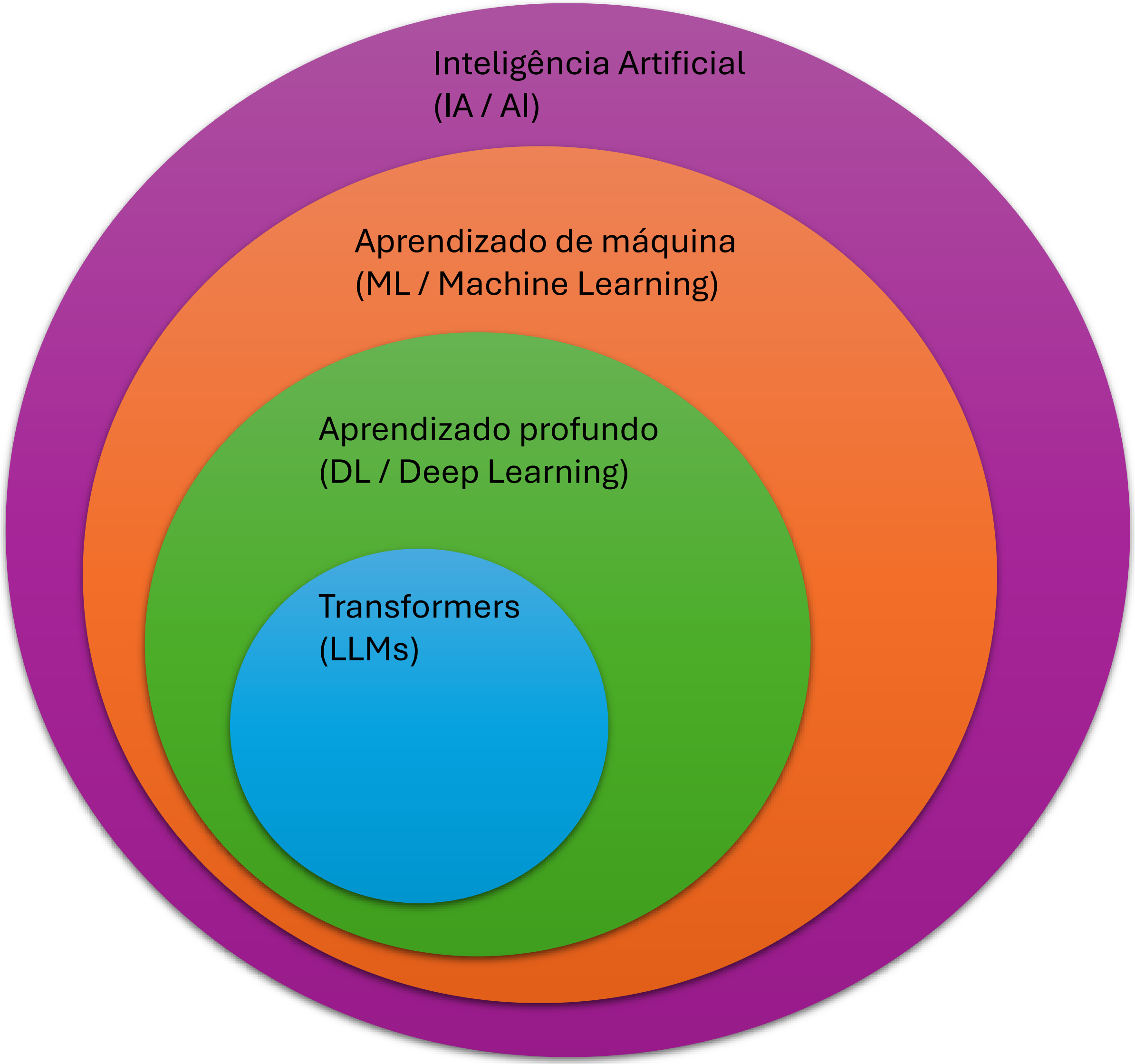


- Aprendizado de máquina
  - Regressão (Linear)
  - Classificação (SVM)
  - Clusterização (K-means)
  - Random Forest
  - GBDT / XGBoost
  - Reinforcement Learning (parte em ML, parte fora)

- Aprendizado profundo
  - CNNs (visão computacional)
  - RNNs, LSTMs, GRU
  - Transformers
  - Autoencoders
  - GANs

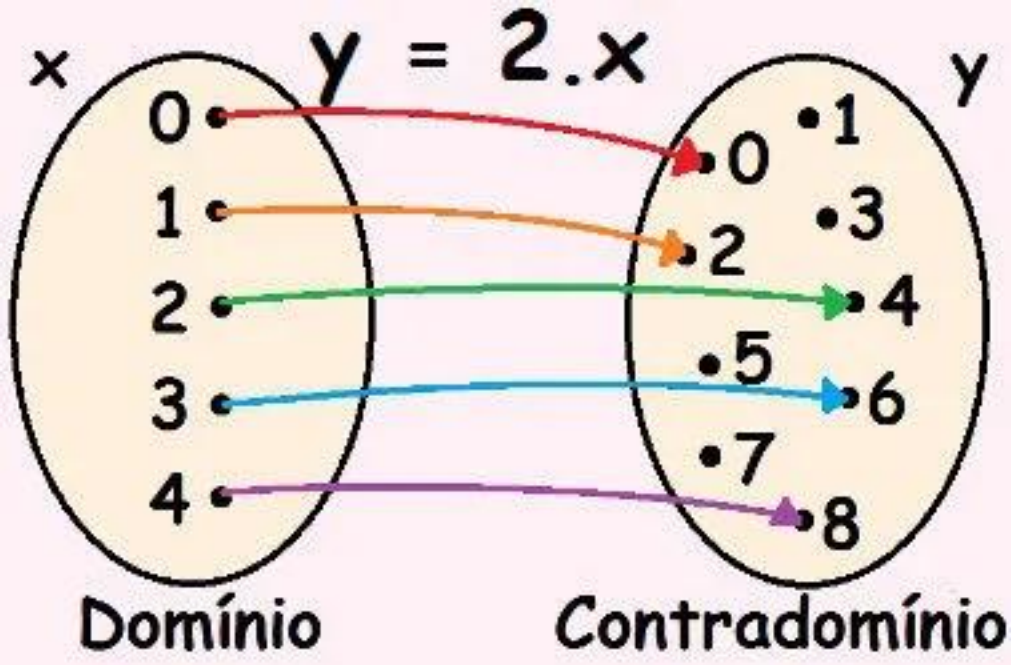
- Transformers
  - GPT
  - BERT
  - LLaMA
  - Mistral

Inteligência Fraca



Matemática?

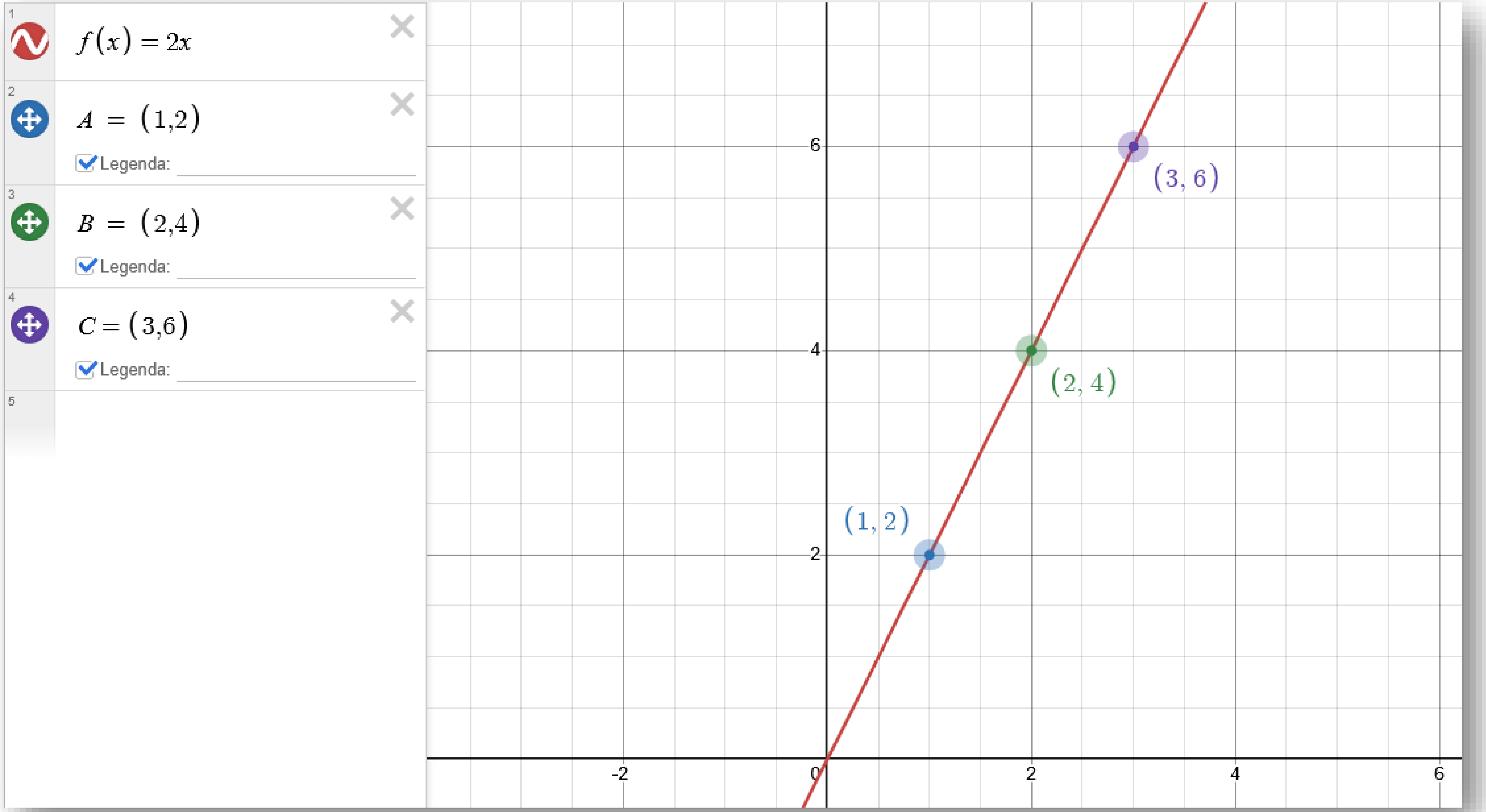
$f(x) = 2x$



https://brasilescola.uol.com.br/o-que-e/matematica/o-que-e-funcao.htm

X	Y
0	0
1	2
2	4
3	6
4	8

f(x) = 2x

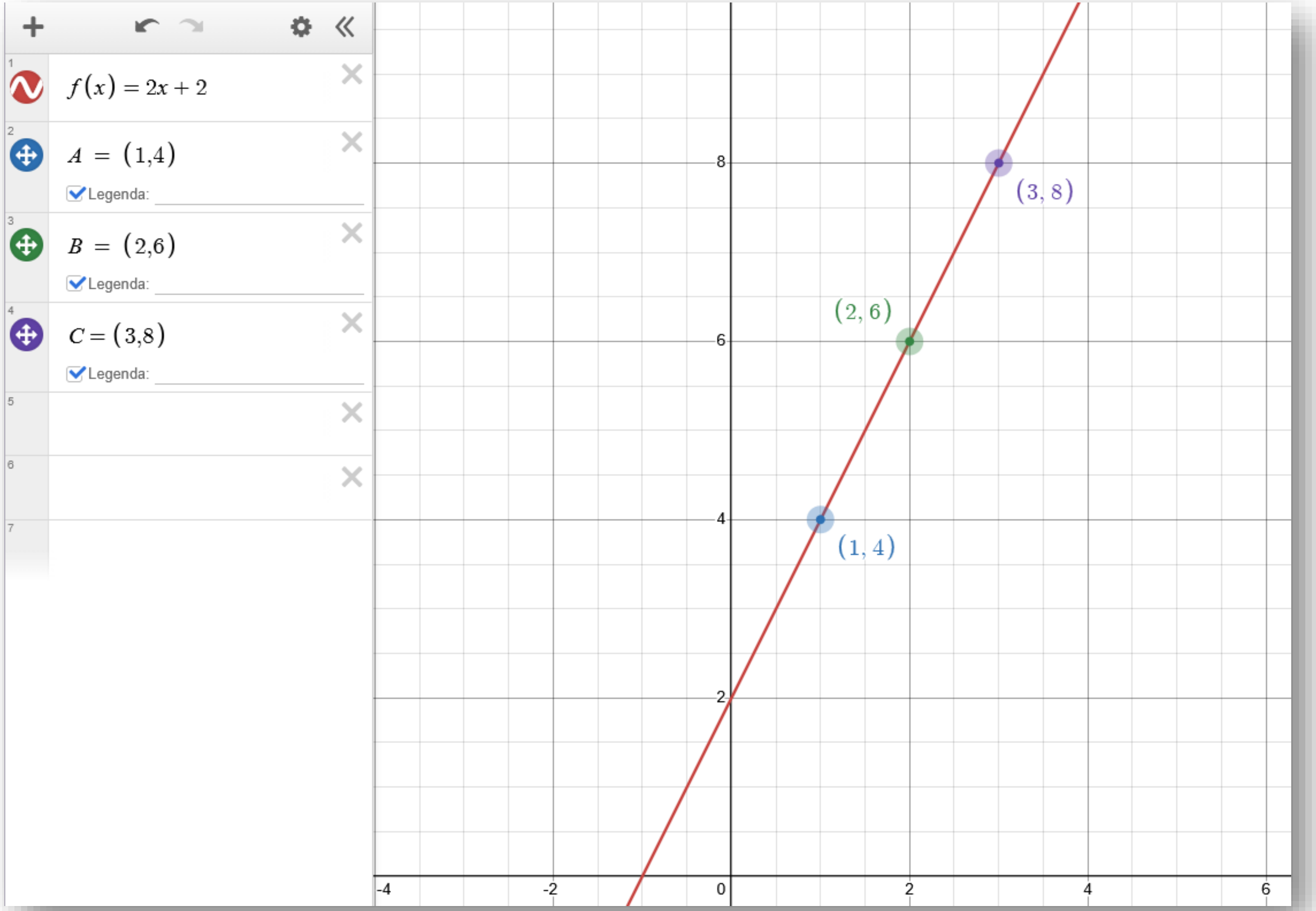


X	Y
0	0
1	2
2	4
3	6
4	8

$$f(x) = ?$$

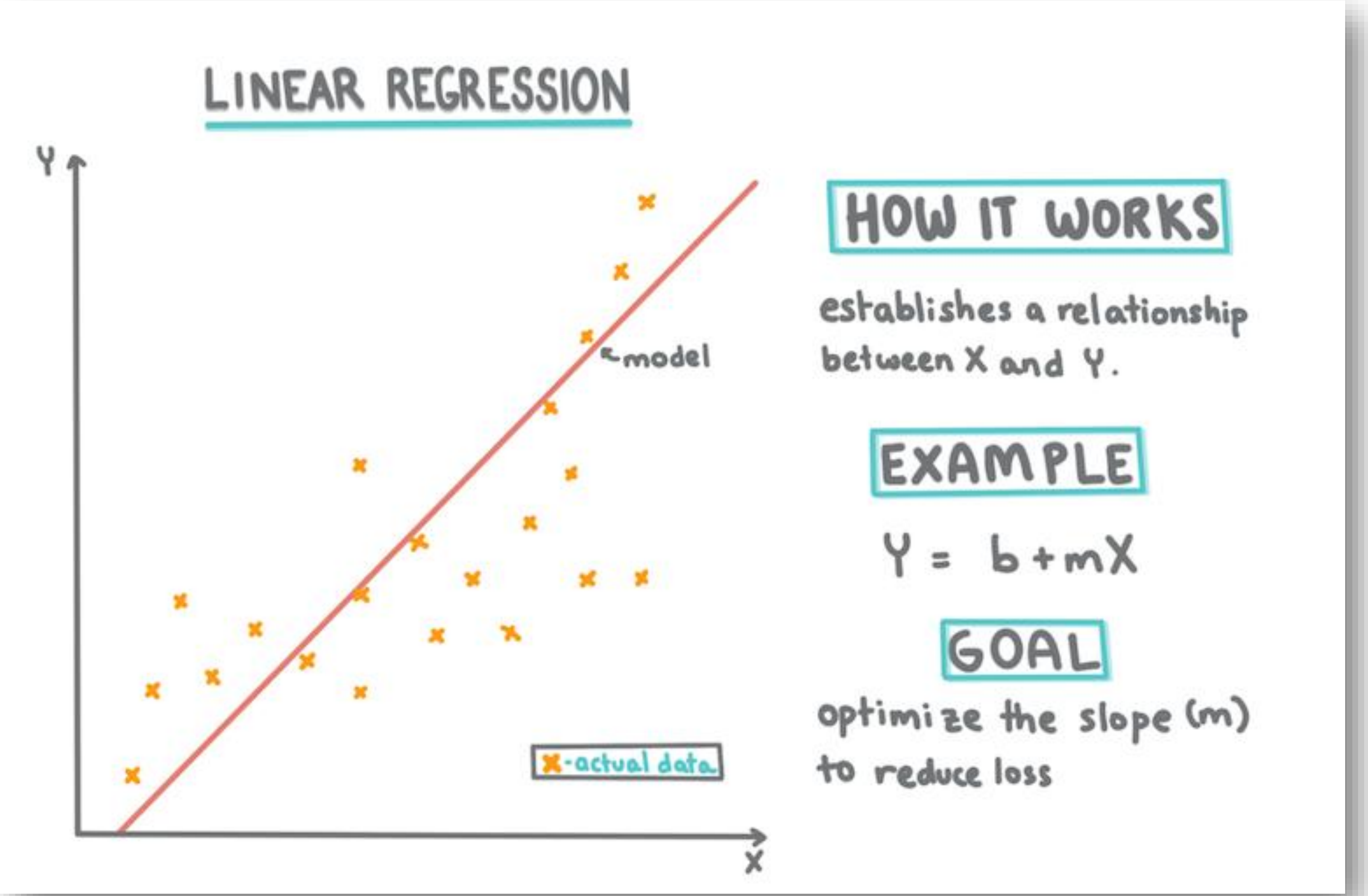
X	f(x)	Y
0	?	2
1	?	4
2	?	6
3	?	8
4	?	10

$f(x) = 2x+2$



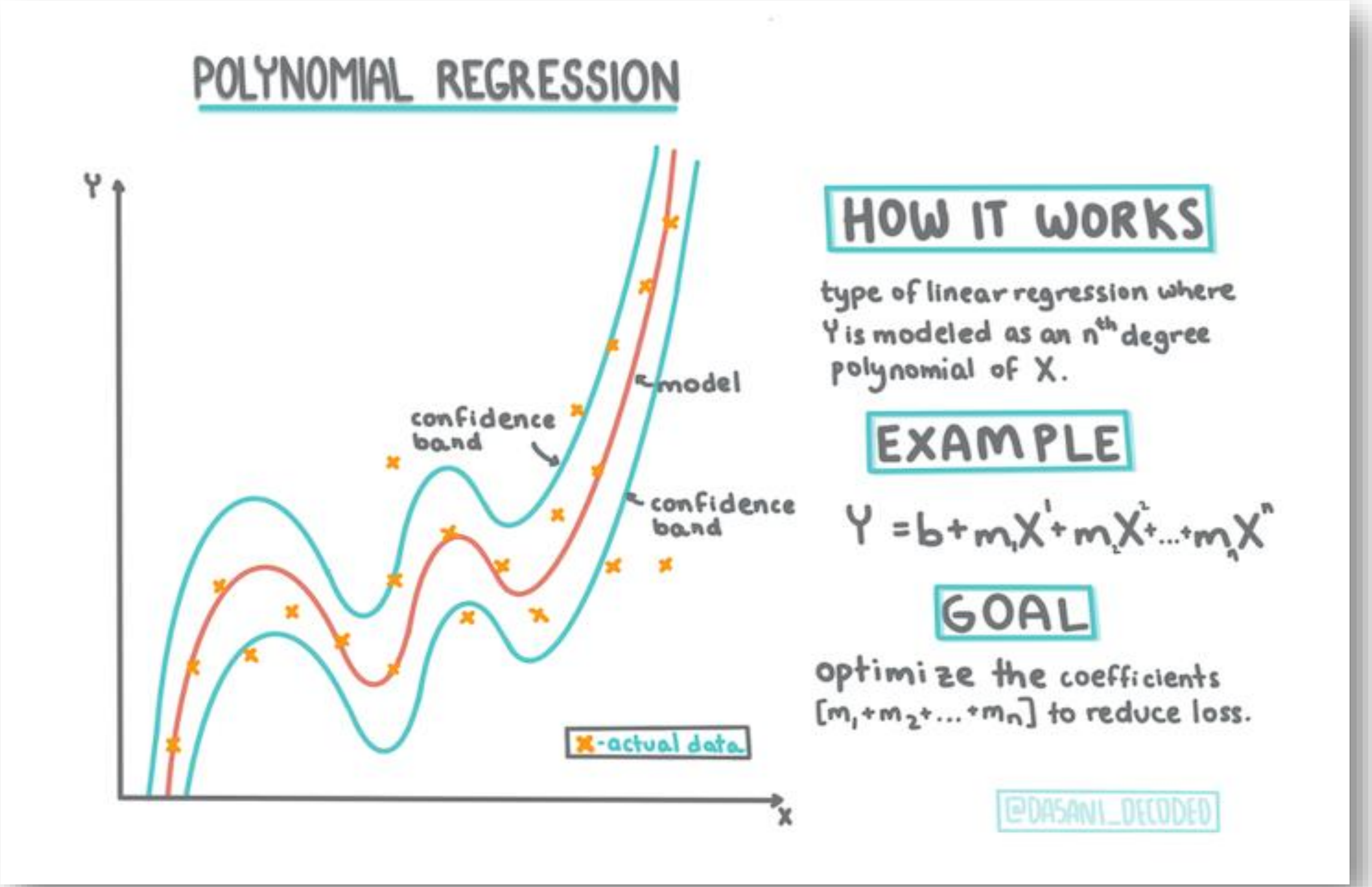
X	f(x)	Y
0	2x+2	2
1	2x+2	4
2	2x+2	6
3	2x+2	8
4	2x+2	10

Regressão



https://github.com/microsoft/ML-For-Beginners/blob/main/translations/br/2-Regression/3-Linear/README.md

https://x.com/dasani_decoded

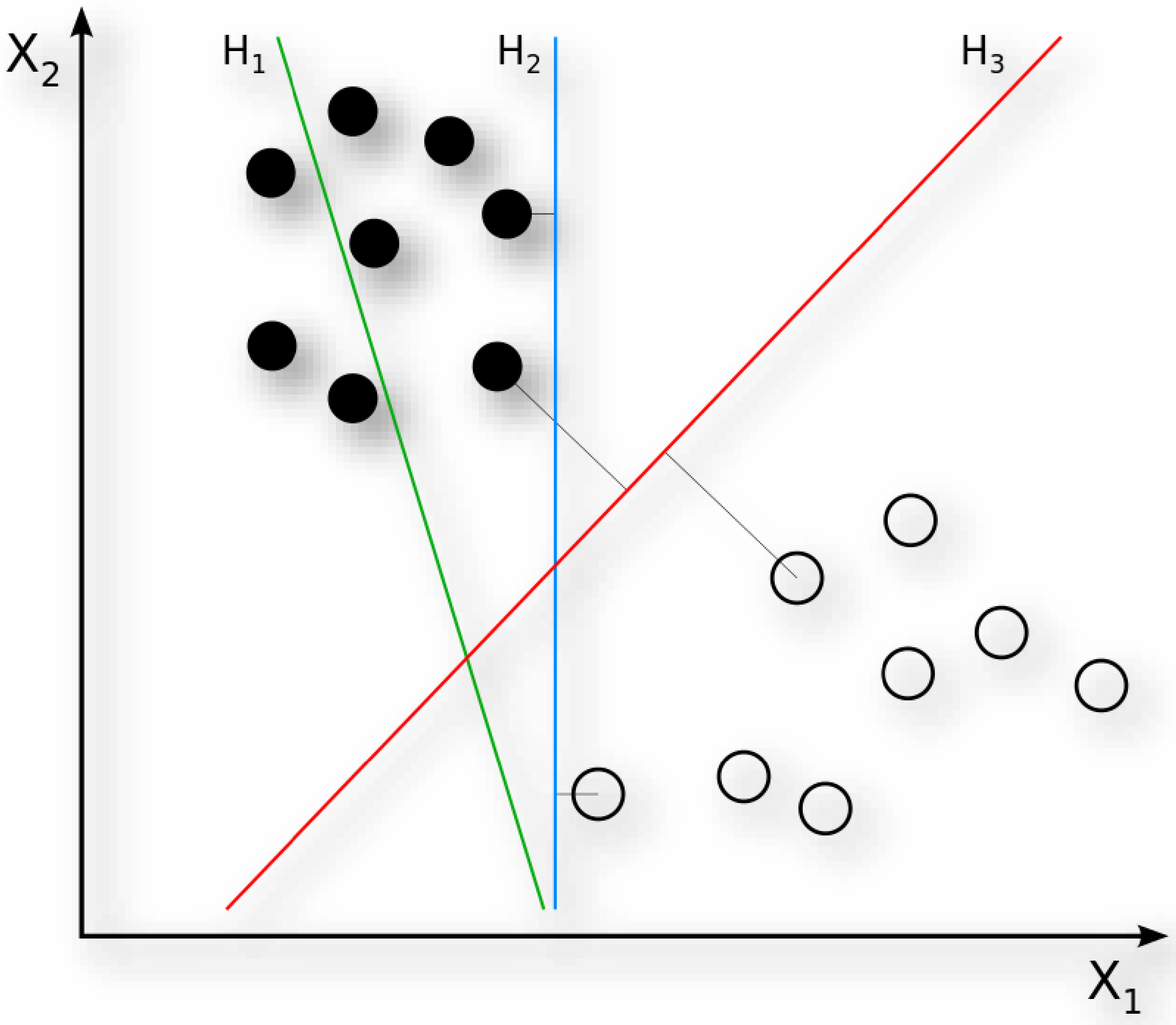


https://github.com/microsoft/ML-For-Beginners/blob/main/translations/br/2-Regression/3-Linear/README.md

https://x.com/dasani_decoded

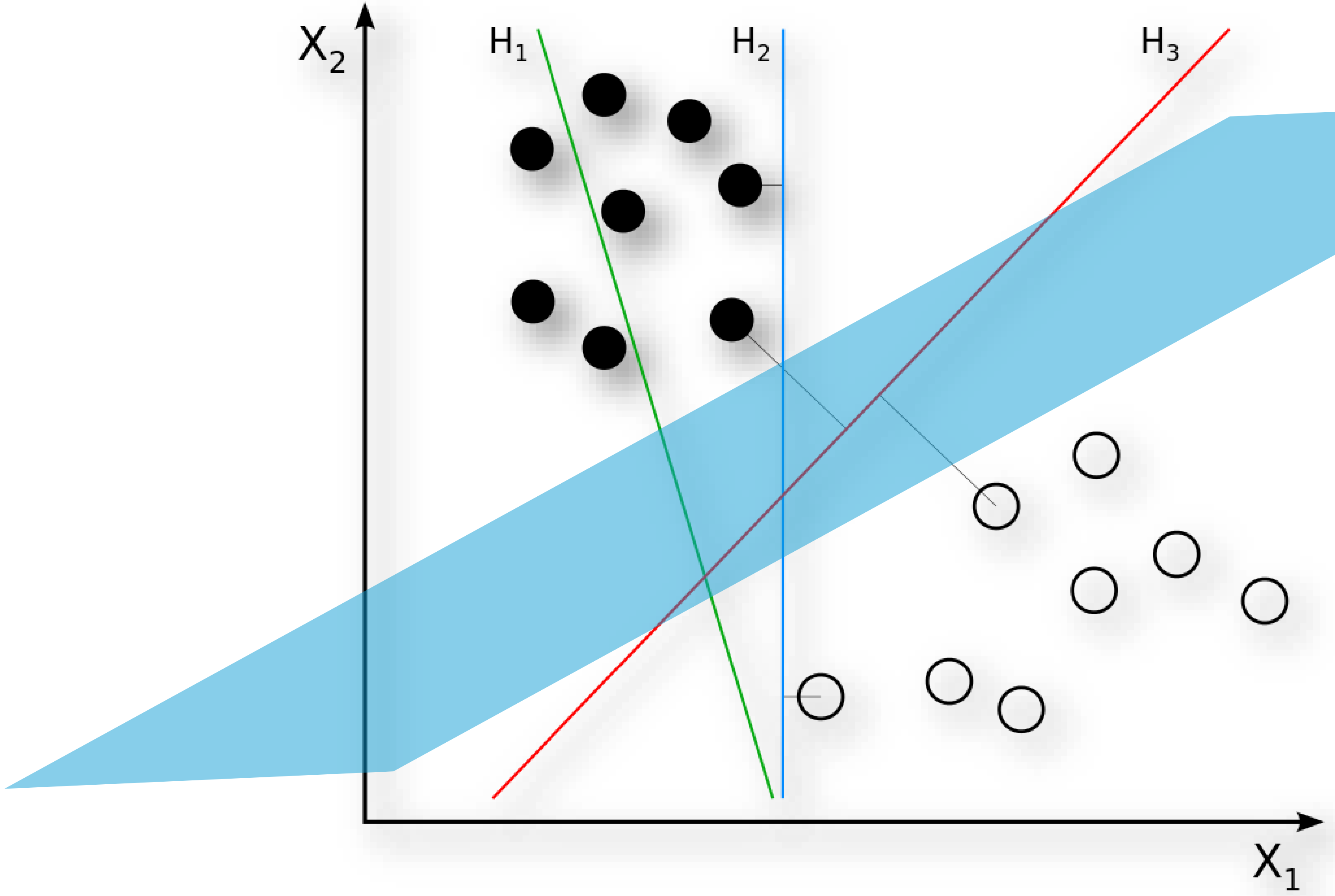
Classificação

SVM



https://github.com/microsoft/ML-For-Beginners/blob/main/4-Classification/3-Classifiers-2/images/svm.png

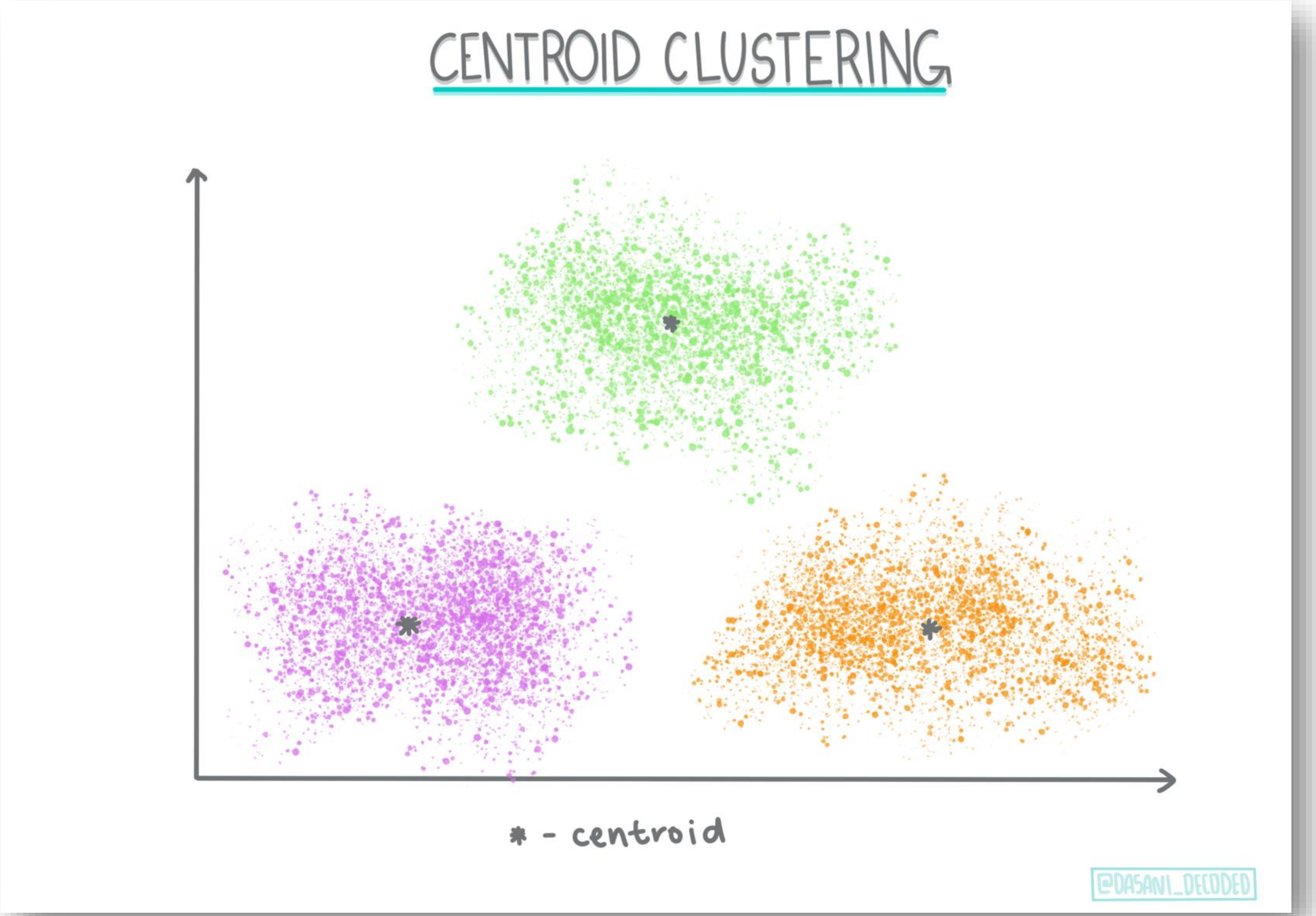
SVM



https://github.com/microsoft/ML-For-Beginners/blob/main/4-Classification/3-Classifiers-2/images/svm.png

Clusterização

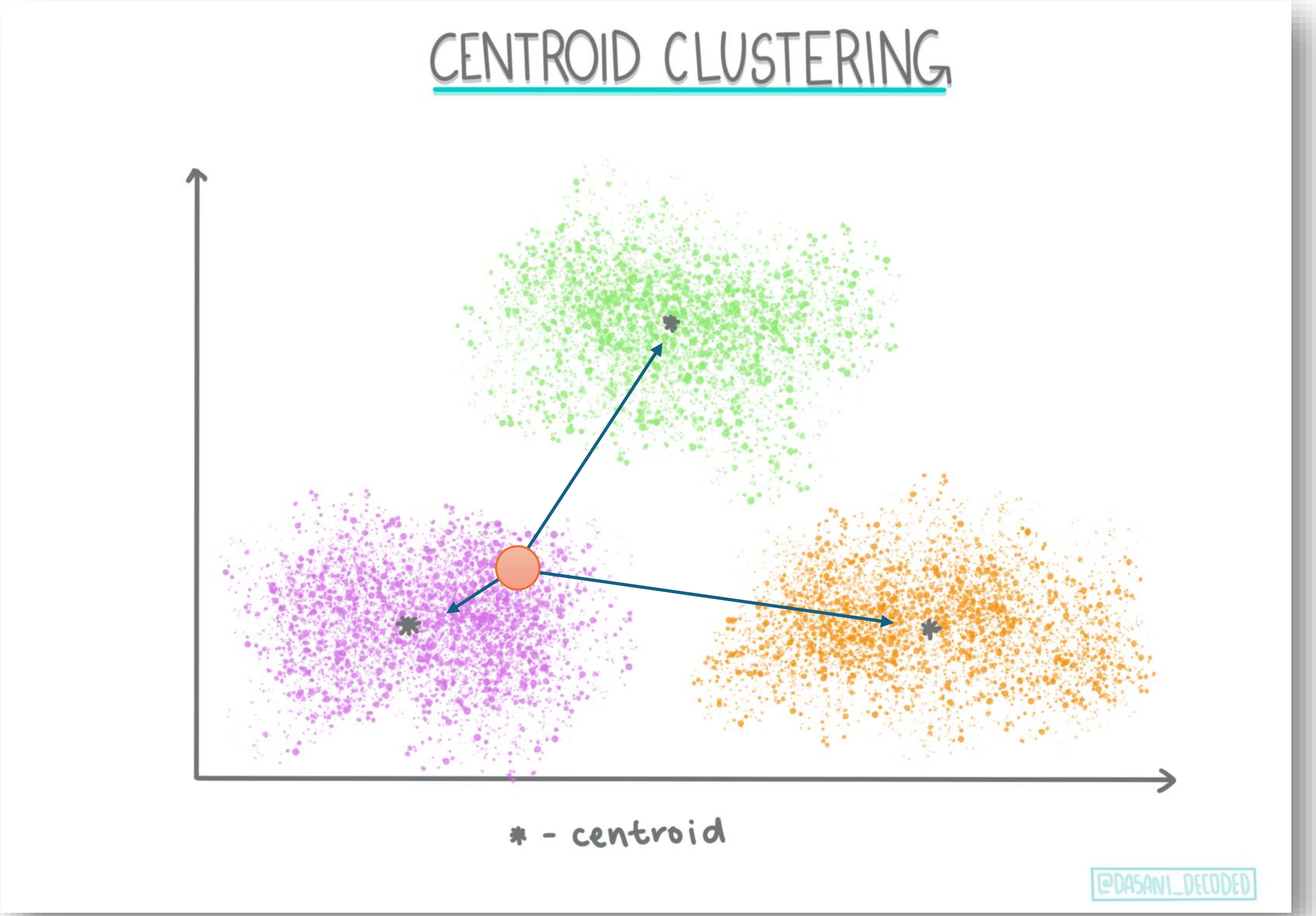
K-means



https://github.com/microsoft/ML-For-Beginners/blob/main/translations/br/5-Clustering/1-Visualize/README.md

https://x.com/dasani_decoded

K-means



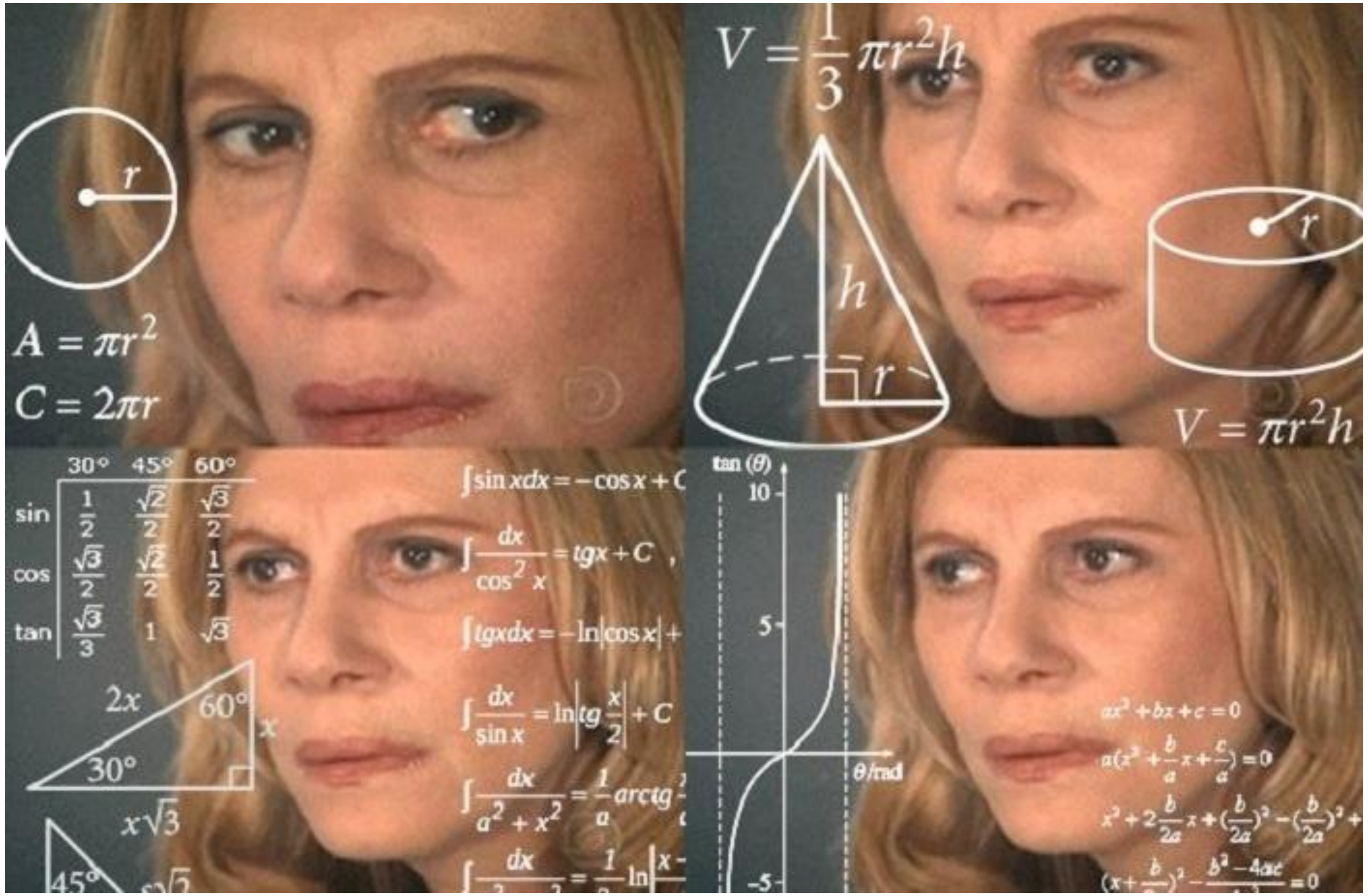
https://github.com/microsoft/ML-For-Beginners/blob/main/translations/br/5-Clustering/1-Visualize/README.md

https://x.com/dasani_decoded

É tudo número?

Textos ~~que usei~~ na escola?

Meu aspirador robô que me reconhece?



E isso aqui?

E meu GPT que disse que quer namorar comigo?

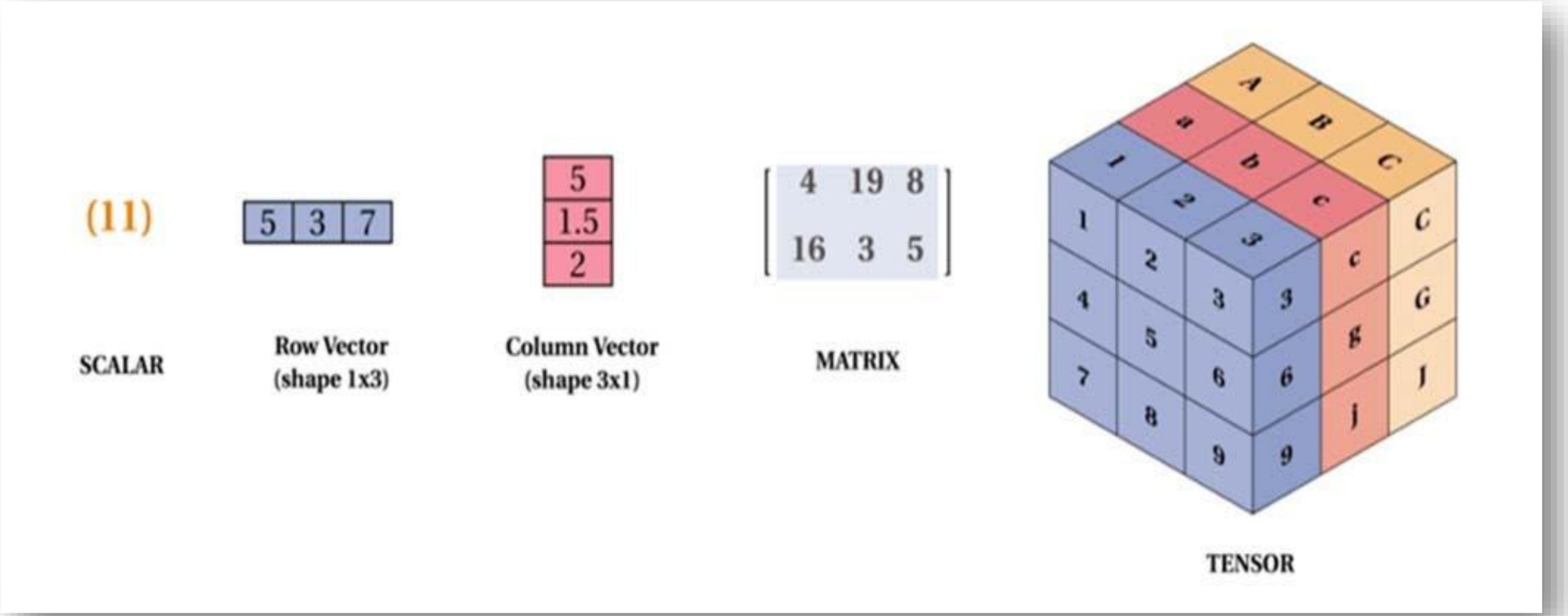
Ou isso?

É tudo número?

Sim

Matriz?

$$f(x) = y$$

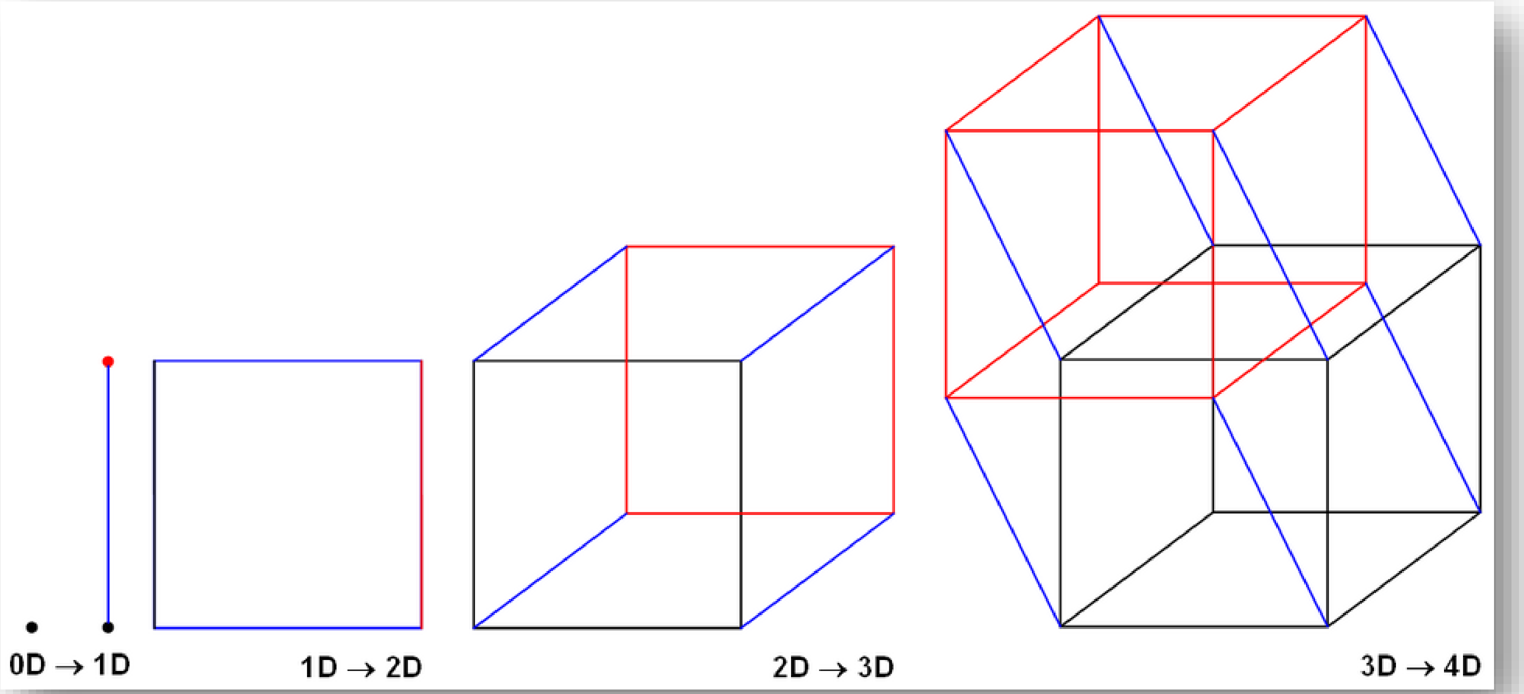


https://medium.com/@fathahka/scalar-vector-matrix-tensor-in-linear-algr-f1bc673fa4eb

Espaço?

f(x) = y

?



https://commons.wikimedia.org/wiki/File:Hypercube-construction-4d.png

$$v = (x_1, x_2, x_3, x_4)$$

$$\downarrow$$

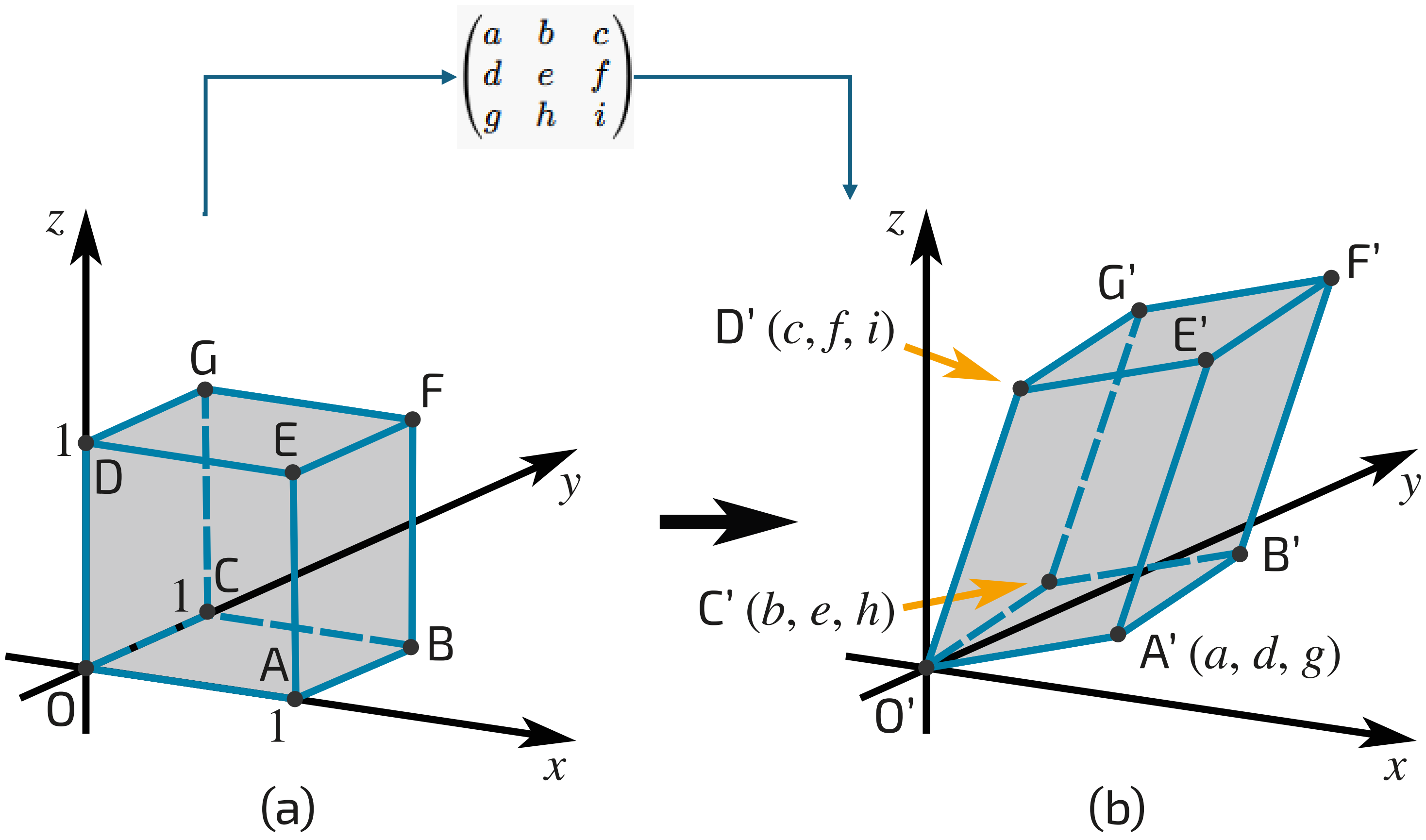
$$\mathbb{R}^4 = \{(x_1, x_2, x_3, x_4) \mid x_i \in \mathbb{R}\}$$

$$\downarrow$$

$$\begin{bmatrix} 5 \\ -2 \\ 0.7 \\ 11 \end{bmatrix}$$

Transformação?

f(x) = y



https://isaacscience.org/concepts/matrices_linear_transformations3d

$$v = (x_1, x_2, x_3, x_4)$$

$$\downarrow$$

$$\mathbb{R}^4 = \{(x_1, x_2, x_3, x_4) \mid x_i \in \mathbb{R}\}$$

$$\downarrow$$

$$\begin{bmatrix} 5 \\ -2 \\ 0.7 \\ 11 \end{bmatrix}$$

Rede neural?

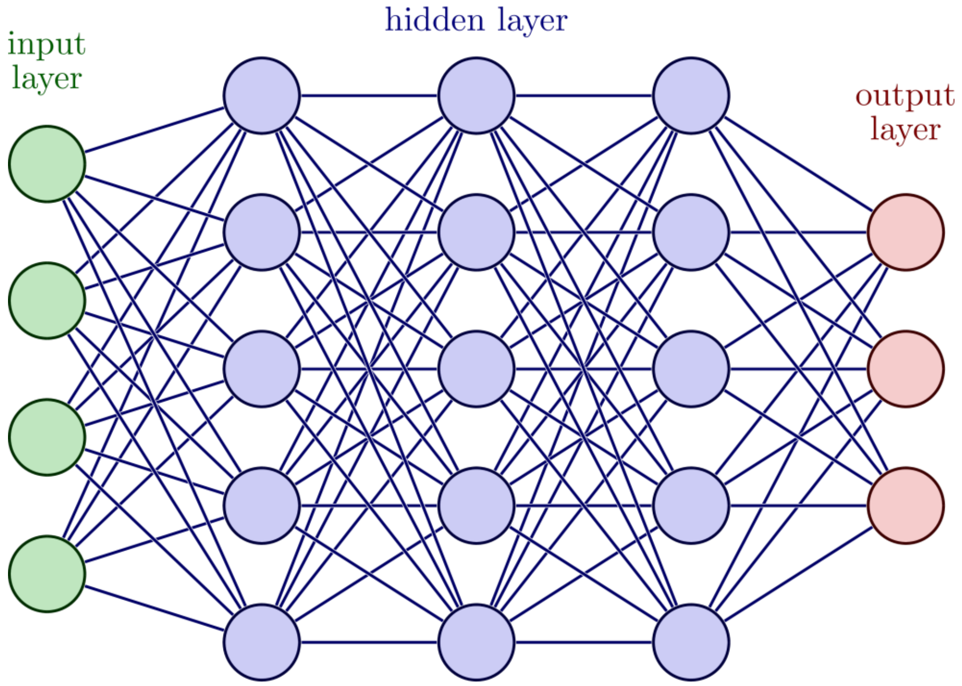
$$f(x) = y$$

$$T : V \to W$$

$$T(u+v) = T(u) + T(v)$$

$$T(\alpha v) = \alpha T(v)$$

Entrada



Saída

Uma rede neural é uma composição de várias transformações afins (Álgebra Linear) combinadas com funções não lineares

https://tikz.net/neural_networks/

$$f(x) = Wx + b \qquad y = \sigma(Wx + b)$$

Textinho?

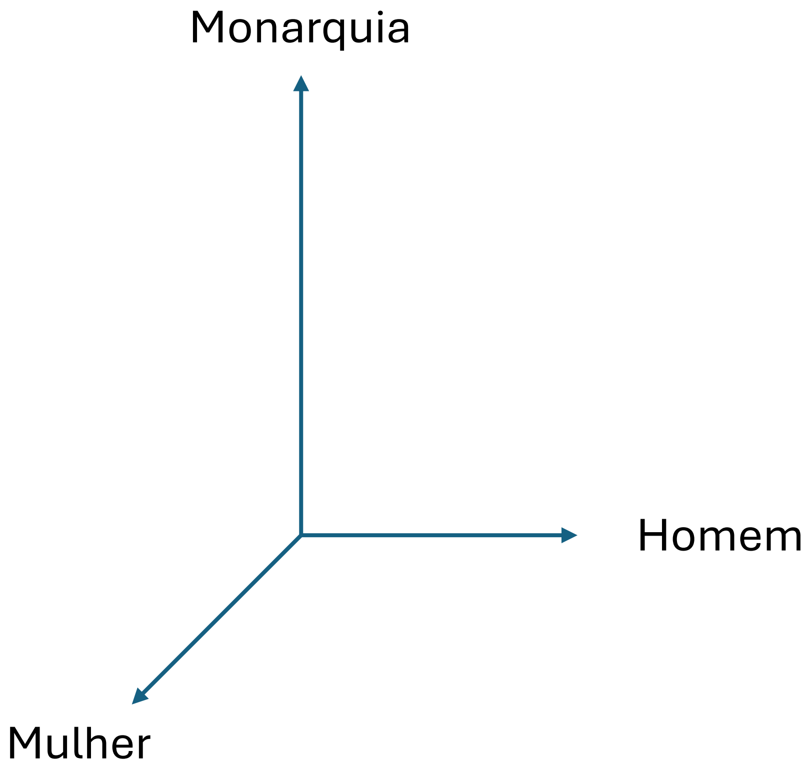
E meus textinhos?

A **rainha** é bonita

Tokens “Pedaços de frase”	**Embedding** “significado numérico dela”
A	vetor(4096)
rainha	vetor(4096)
é	vetor(4096)
bonita	vetor(4096)

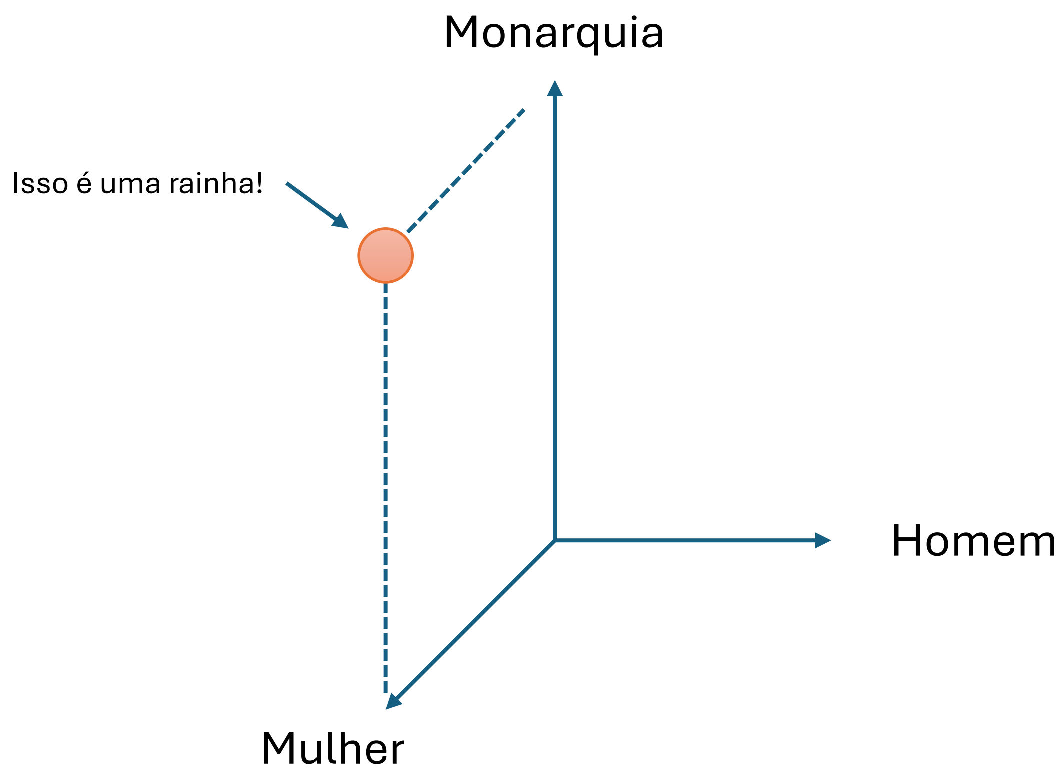
E meus textinhos?

A **rainha** é bonita

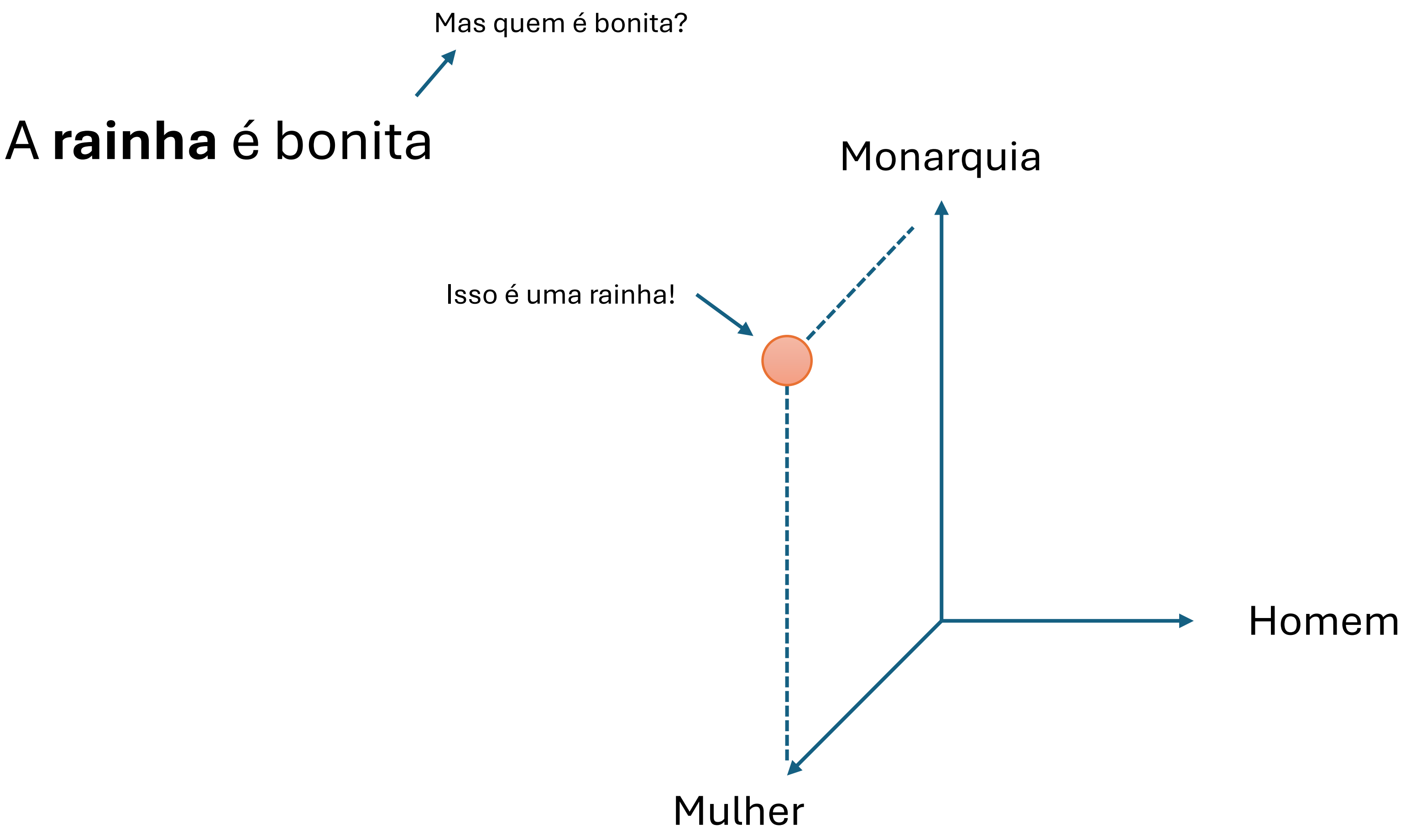


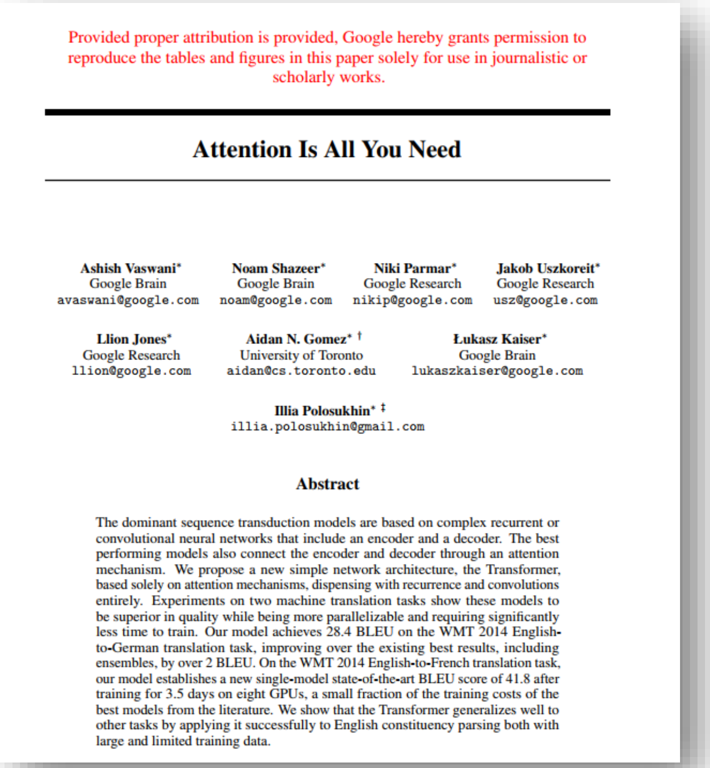
E meus textinhos?

A **rainha** é bonita



E meus textinhos?





Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

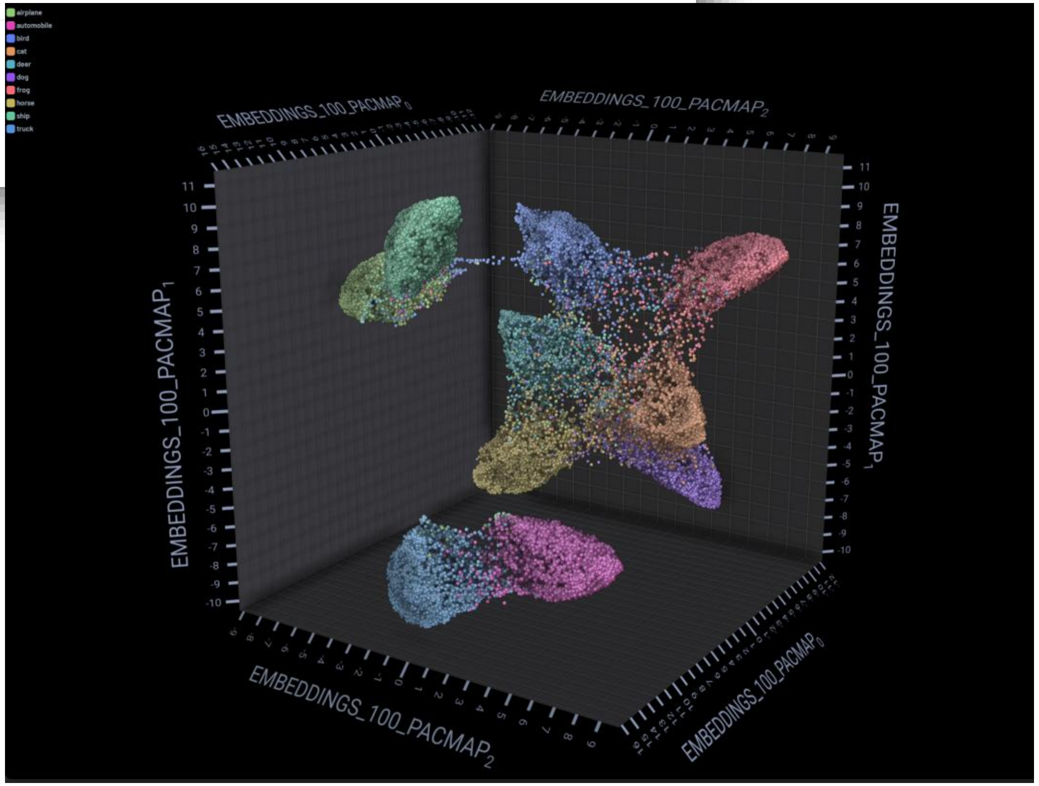
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

E meus textinhos?

GPT-3.5 Turbo	2022	**4096**	100k	**~6–20B (estimado)**
GPT-4 (base)	2023	**~12k** (estimado)	>100k	**~200B+ (estimado)**
GPT-4 Turbo / GPT-4.1	2024	**não divulgado**, estimado **8k–10k**	>100k	**não divulgado**, estimado **>100B**

E meus textinhos?

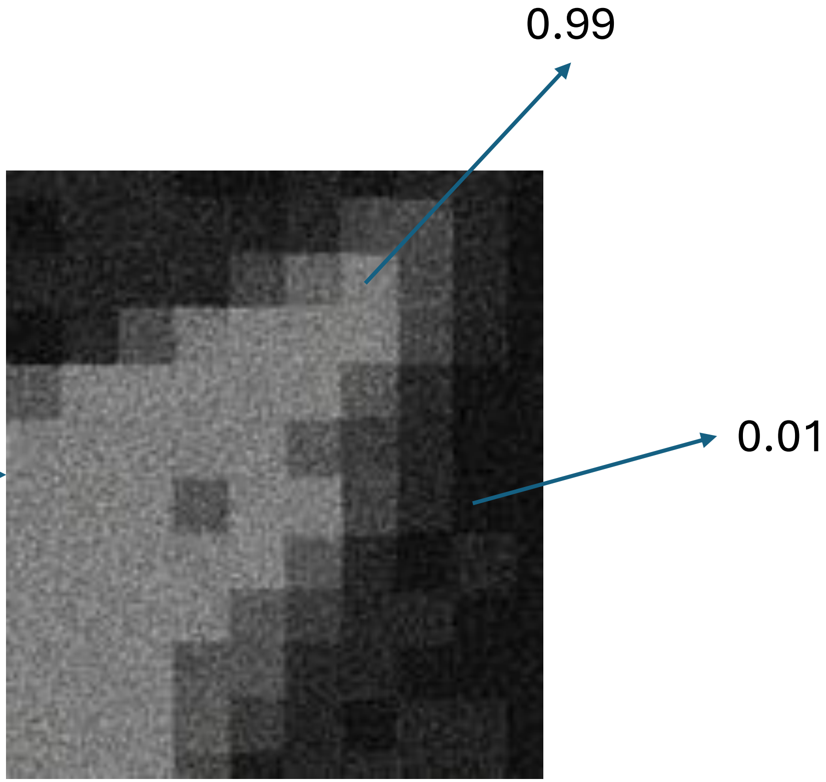
GPT-3.5 Turbo	2022	**4096**	100k	**~6–20B (estimado)**
GPT-4 (base)	2023	**~12k** (estimado)	>100k	**~200B+ (estimado)**
GPT-4 Turbo / GPT-4.1	2024	**não divulgado**, estimado **8k–10k**	>100k	



https://docs.3lc.ai/3lc/latest/user-guide/python-package/embeddings.html

Gatinho?

E meus gatinhos?

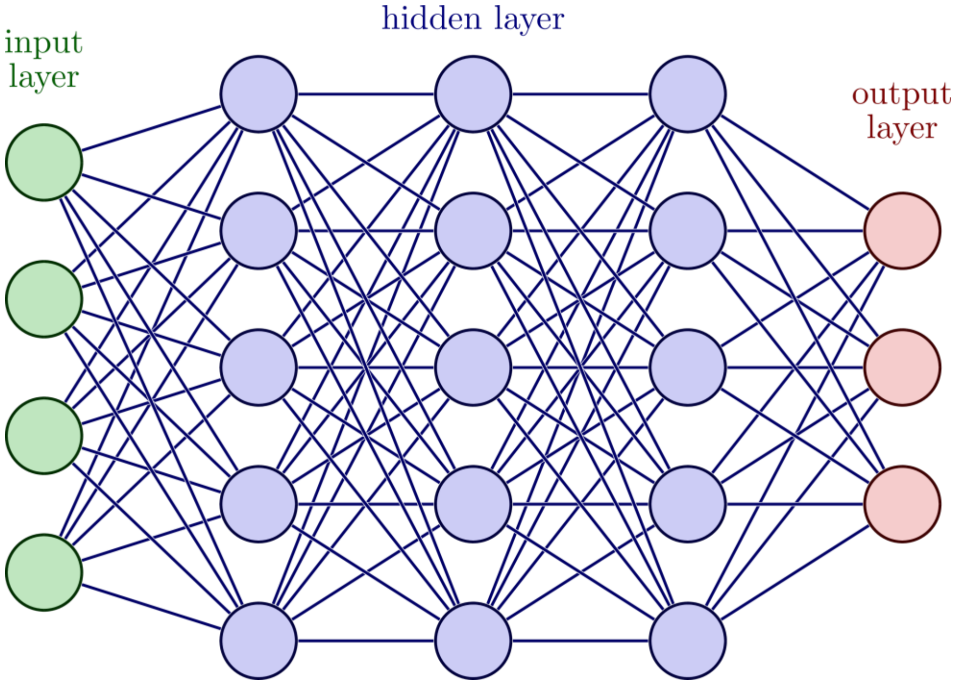


ResNet50
3 canais (Red, Green, Blue)
224 × 224 pixels

[0.73, 0.41, 0.88, 0.56, 0.92, 0.37, 0.64, 0.29, 0.77, 0.53, 0.18, 0.94,...

E meus gatinhos?

[0.73, 0.41, 0.88, 0.56, 0.92, 0.37, 0.64, 0.29, 0.77, 0.53, 0.18, 0.94,...



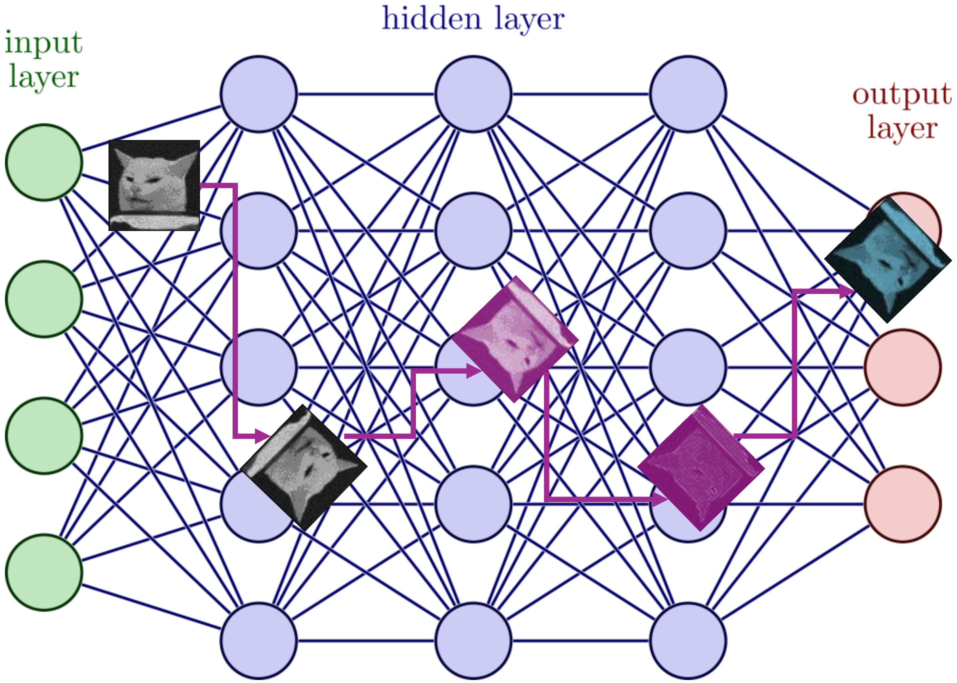
https://tikz.net/neural_networks/

E meus gatinhos?

0

1

[0.73, 0.41, 0.88, 0.56, 0.92, 0.37, 0.64, 0.29, 0.77, 0.53, 0.18, 0.94,...



[**0** = 0.99 | **1** = 0.01]

https://tikz.net/neural_networks/

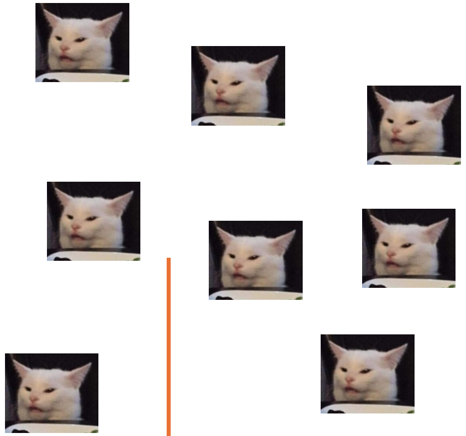
Isso é um gato!

Probabilidade de ser 0 = 0.99%

Probabilidade de ser 1 = 0.01%

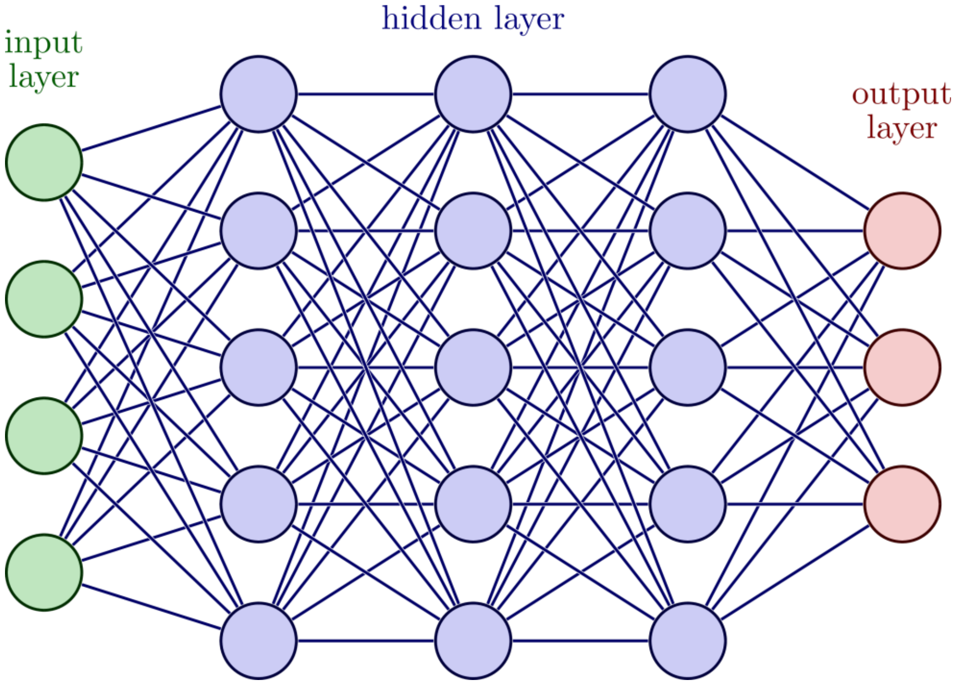
E meus gatinhos?

Aqui é tudo 0

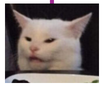


Aqui é tudo 1

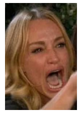
Atualiza a rede



https://tikz.net/neural_networks/

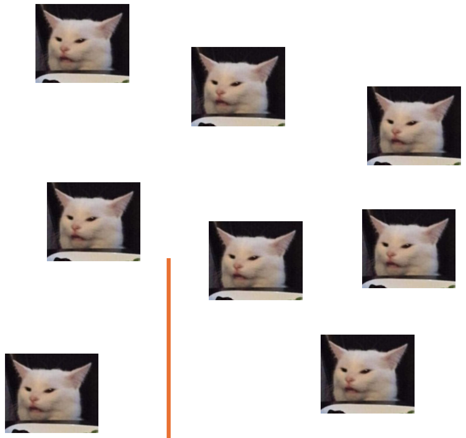


0 ou 1?



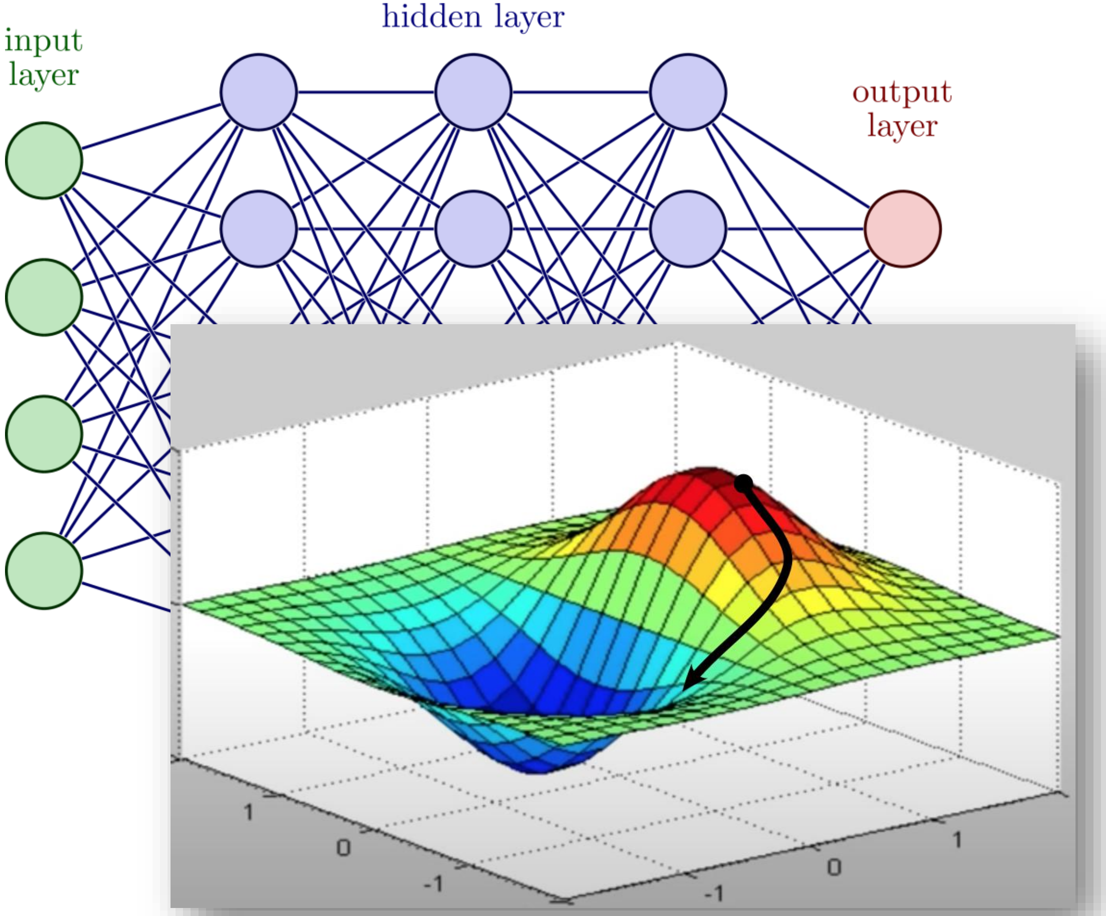
E meus gatinhos?

Aqui é tudo 0

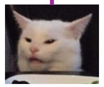


Aqui é tudo 1

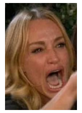
Atualiza a rede



https://mriquestions.com/back-propagation.html



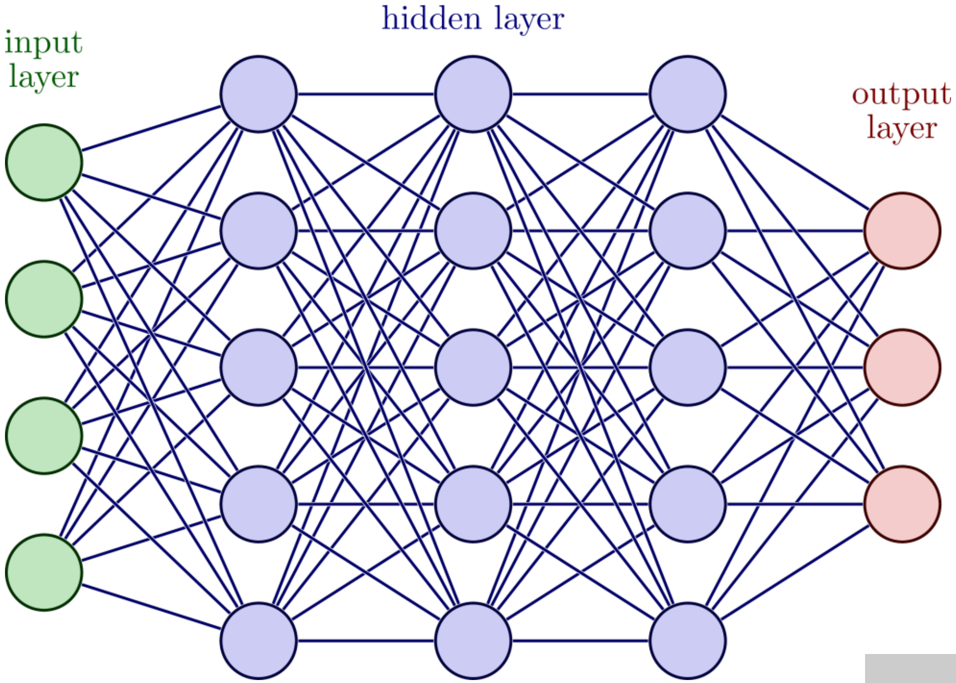
0 ou 1?



Quão longe estou da resposta?

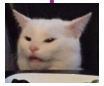
E meus gatinhos?

Aqui é tudo 0

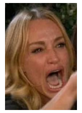


https://tikz.net/neural_networks/

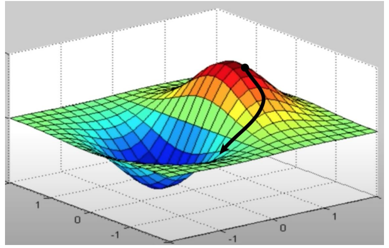
Atualiza a rede



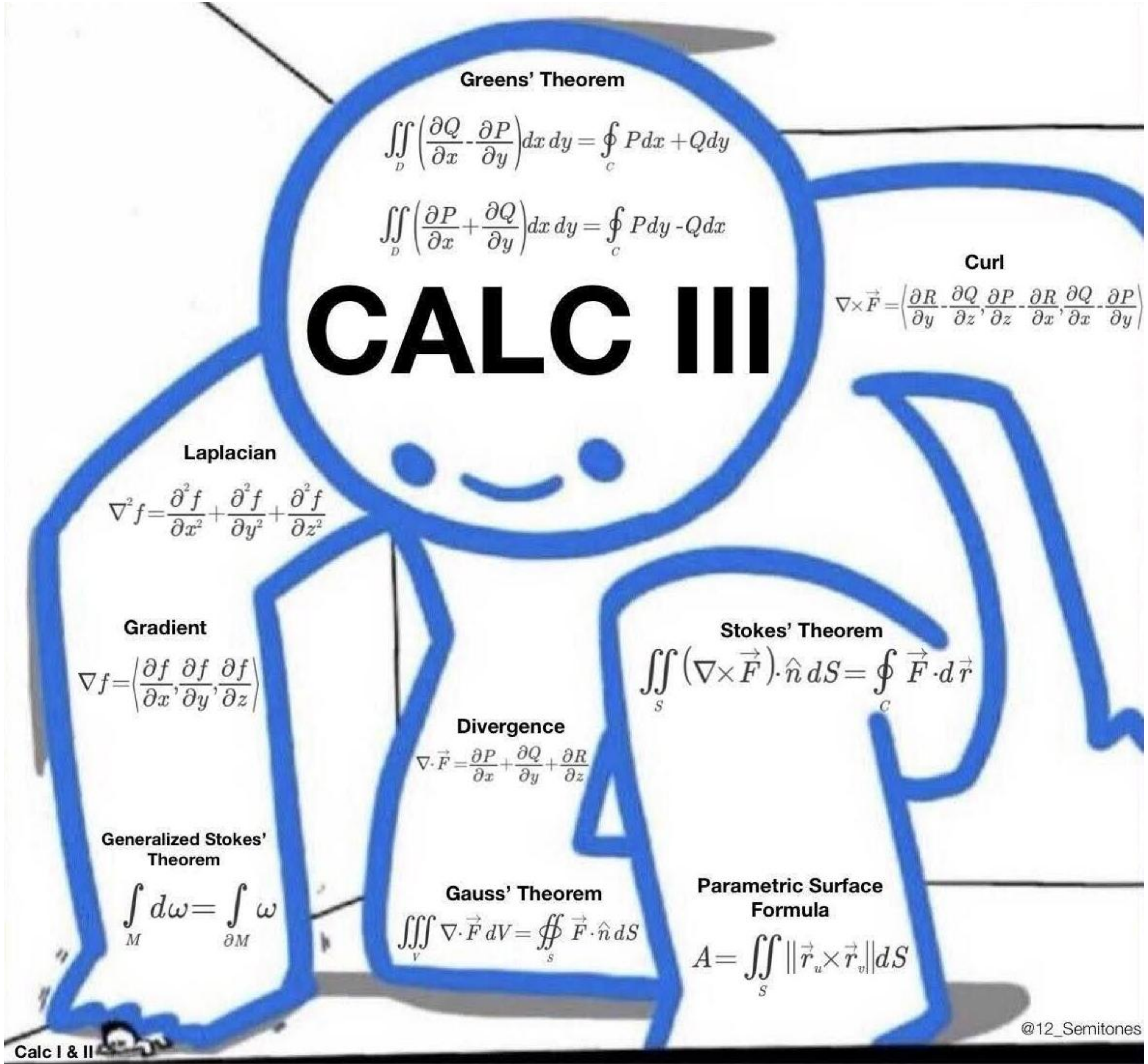
0 ou 1?

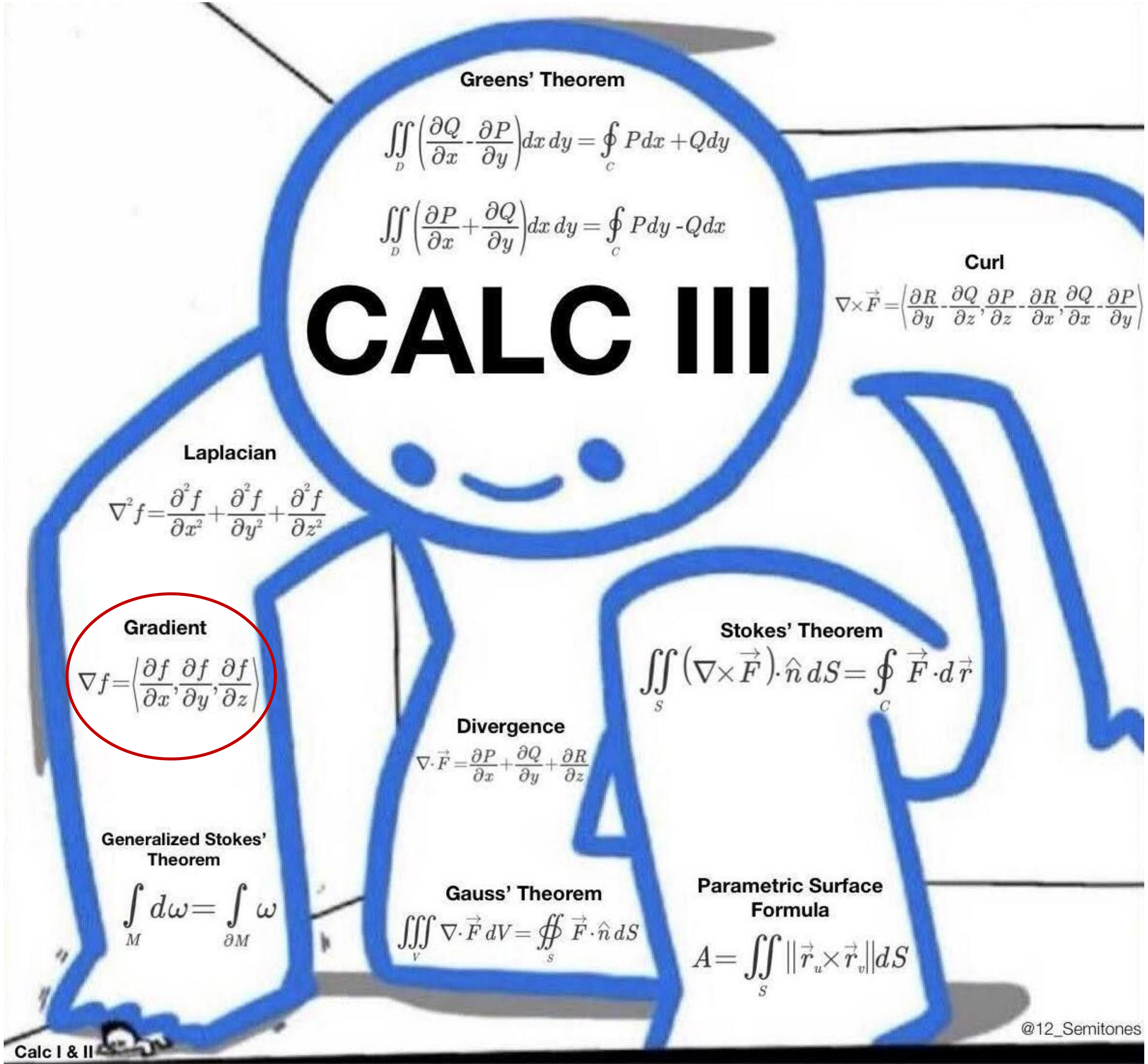


Aqui é tudo 1



https://mriquestions.com/back-propagation.html

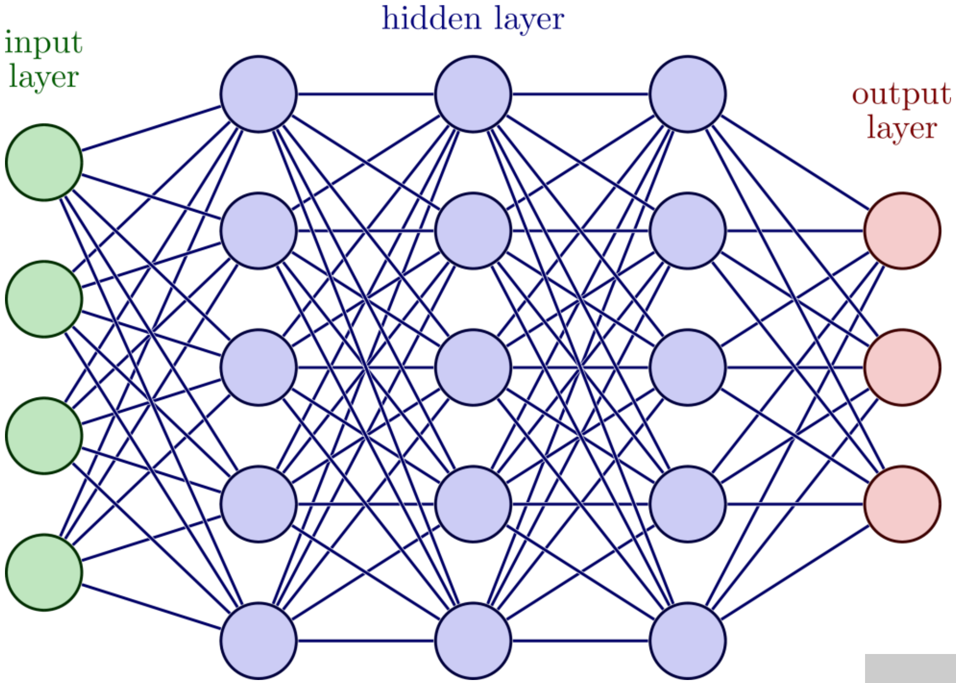

Greens' Theorem
$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_C P dx + Q dy$
$\iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dx\, dy = \oint_C P dy - Q dx$
CALC III
Curl
$\nabla \times \vec{F} = \left\langle \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right\rangle$
Laplacian
$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$
Gradient
$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$
Stokes' Theorem
$\iint_S (\nabla \times \vec{F}) \cdot \hat{n}\, dS = \oint_C \vec{F} \cdot d\vec{r}$
Divergence
$\nabla \cdot \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$
Generalized Stokes' Theorem
$\int_M d\omega = \int_{\partial M} \omega$
Gauss' Theorem
$\iiint_V \nabla \cdot \vec{F}\, dV = \oiint_S \vec{F} \cdot \hat{n}\, dS$
Parametric Surface Formula
$A = \iint_S \|\vec{r}_u \times \vec{r}_v\| dS$
@12_Semitones
Calc I & II


Greens' Theorem
$\iint_D \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\,dy=\oint_C P\,dx+Q\,dy$
$\iint_D \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}\right)dx\,dy=\oint_C P\,dy-Q\,dx$
CALC III
Curl
$\nabla\times\vec{F}=\left\langle\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z},\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x},\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right\rangle$
Laplacian
$\nabla^2 f=\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}$
Gradient
$\nabla f=\left\langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right\rangle$
Stokes' Theorem
$\iint_S (\nabla\times\vec{F})\cdot\hat{n}\,dS=\oint_C \vec{F}\cdot d\vec{r}$
Divergence
$\nabla\cdot\vec{F}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}$
Generalized Stokes' Theorem
$\int_M d\omega=\int_{\partial M}\omega$
Gauss' Theorem
$\iiint_V \nabla\cdot\vec{F}\,dV=\oiint_S \vec{F}\cdot\hat{n}\,dS$
Parametric Surface Formula
$A=\iint_S \|\vec{r}_u\times\vec{r}_v\|dS$
Calc I & II
@12_Semitones

E meus gatinhos?

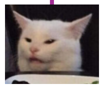
Aqui é tudo 0

Aqui é tudo 1

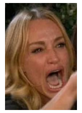


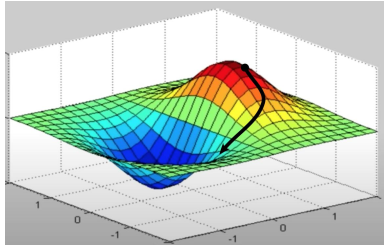
https://tikz.net/neural_networks/

Atualiza a rede



0 ou 1?

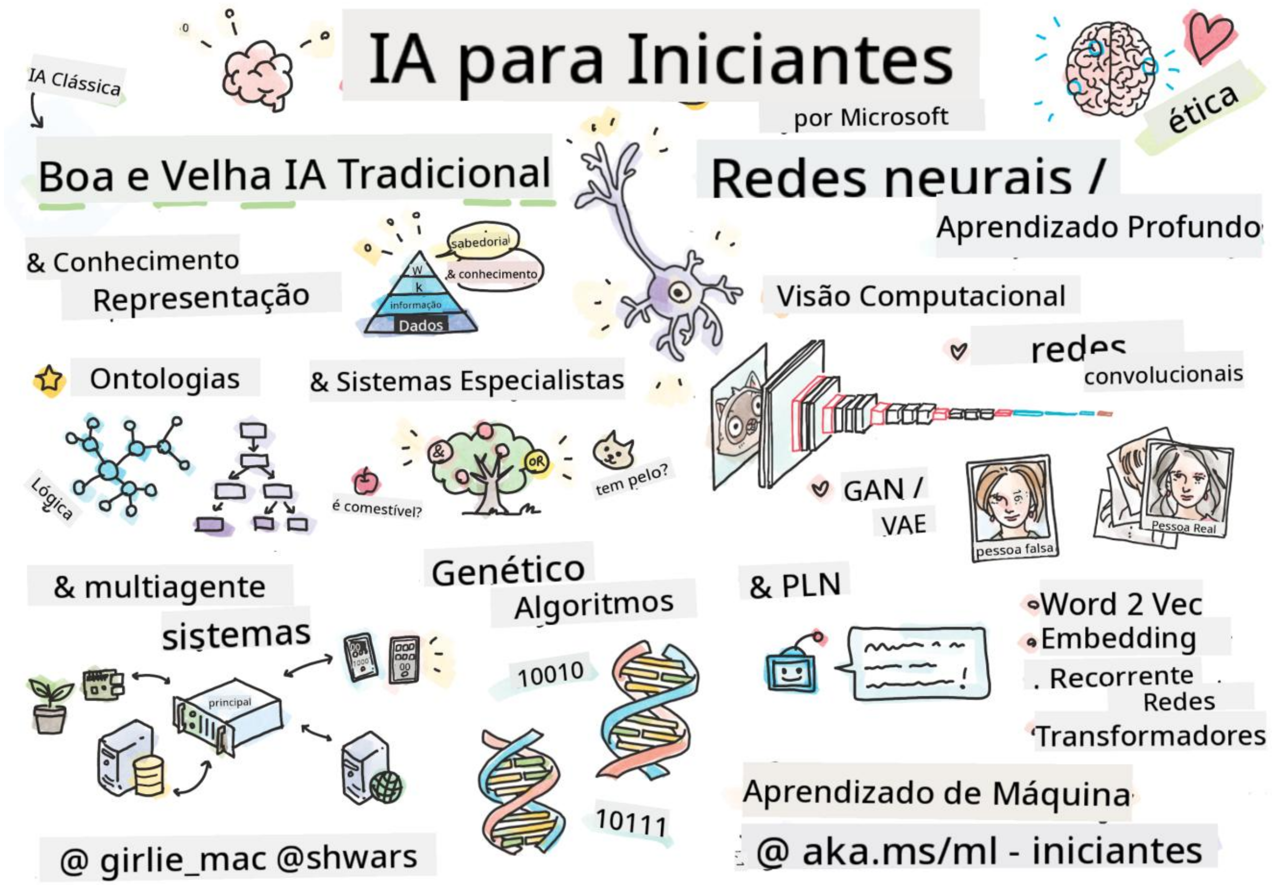




https://mriquestions.com/back-propagation.html

É tudo número?

Sim



https://github.com/microsoft/AI-For-Beginners

https://github.com/microsoft/AI-For-Beginners/blob/main/translations/br/README.md

https://github.com/microsoft/Data-Science-For-Beginners

https://github.com/microsoft/Data-Science-For-Beginners/blob/main/translations/br/README.md

https://github.com/microsoft/ai-agents-for-beginners

https://github.com/microsoft/ai-agents-for-beginners/blob/main/translations/br/README.md

Aprenda Machine Learning - Um Currículo

> 🌍 Viaje pelo mundo enquanto exploramos Machine Learning por meio de culturas globais 🌍

Os Cloud Advocates da Microsoft têm o prazer de oferecer um currículo de 12 semanas e 26 lições sobre **Machine Learning**. Neste currículo, você aprenderá sobre o que às vezes é chamado de **machine learning clássico**, utilizando principalmente a biblioteca Scikit-learn e evitando aprendizado profundo, que é abordado em nosso [currículo de IA para Iniciantes](). Combine essas lições com nosso currículo ['Ciência de Dados para Iniciantes'](), também!

Viaje conosco pelo mundo enquanto aplicamos essas técnicas clássicas a dados de várias regiões do mundo. Cada lição inclui questionários antes e depois da aula, instruções escritas para completar a lição, uma solução, uma tarefa e muito mais. Nossa pedagogia baseada em projetos permite que você aprenda enquanto constrói, uma maneira comprovada de fixar novas habilidades.

✍️ **Agradecimentos especiais aos nossos autores** Jen Looper, Stephen Howell, Francesca Lazzeri, Tomomi Imura, Cassie Breviu, Dmitry Soshnikov, Chris Noring, Anirban Mukherjee, Ornella Altunyan, Ruth Yakubu e Amy Boyd

🎨 **Agradecimentos também aos nossos ilustradores** Tomomi Imura, Dasani Madipalli e Jen Looper

🙏 **Agradecimentos especiais 🙏 aos nossos autores, revisores e colaboradores de conteúdo Microsoft Student Ambassador**, especialmente Rishit Dagli, Muhammad Sakib Khan Inan, Rohan Raj, Alexandru Petrescu, Abhishek Jaiswal, Nawrin Tabassum, Ioan Samuila e Snigdha Agarwal

🤩 **Gratidão extra aos Microsoft Student Ambassadors Eric Wanjau, Jasleen Sondhi e Vidushi Gupta pelas nossas lições em R!**

https://github.com/microsoft/ML-For-Beginners/

https://github.com/microsoft/ML-For-Beginners/blob/main/translations/br/README.md github.com/lnncrs