

Universidade Federal do Pará
Programa de Pós-Graduação em Ciência da Computação

Aprendizado de Máquina

Prof. Dr. Bianchi Serique Meiguins

Prof. Dr. Tiago Araújo

🌐 Texto

📄 Documentos

DETECTAR IDIOMA

INGLÊS

PORTUGUÊS

ESPAÑHOL

↔️

PORTUGUÊS

INGLÊS

INDONÉSIO

Tradução

🎤

0/5000

- 📧 Caixa de entrada
- 🕒 Adiados
- 📌 Importante
- Enviados
- 📄 Rascunhos 22
- ✉️ Todos os e-mails
- 🚫 Spam 31
- 🗑️ Lixeira
- 📁 Categorias
- ⌵ Mais

Emotions Legends ✕

😊

■

happy

😮

■

surprise

😐

■

neutral

😨

■

fear

😞

■

sad

😡

■

angry

😬

■

disgust

😏

■

contempt

Apply

Polarity Legends ✕

😊

■

positive

😐

■

neutral

😡

■

negative

Apply

Quantitative Legends ✕

✓

right

✗

wrong

⌚

timeout

10:37 65%

☰ AI Dungeon ↺

You rush to the hospital and find out that the man who killed you is still alive. He tells you that he didn't mean any harm when he shot you, just wanted to scare you a bit. He says that he'll be going home now though since it was such a long trip.

> Go to heaven and file a complaint about the murderer

You decide to take this opportunity to complain about the murderer to God. You head over to the church where you were baptized and file a formal complaint against him with Him. He takes pity on you and sends you straight to hell instead of sending him directly there.

⌵ What will you do?..

Help Revert Alter Remember

Tip: Remember to start your action with a verb, ex: Attack the orc



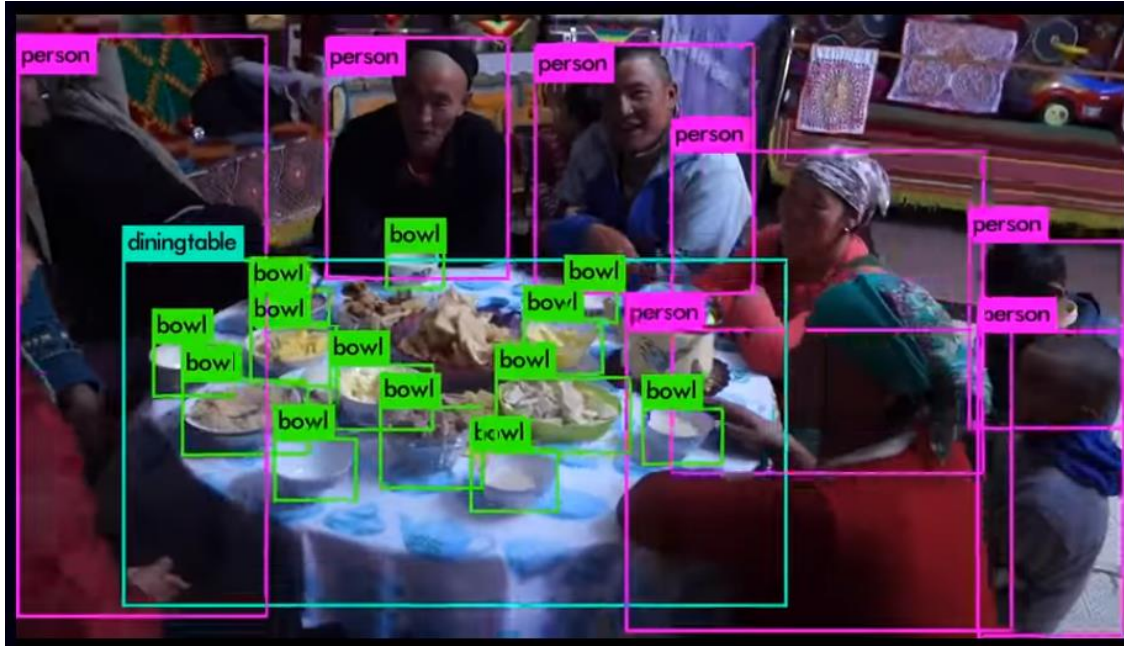
DLSS (ORIGINAL QUALITY MODE)

Ray Tracing - 1080p, Quality: High, RTX 2060



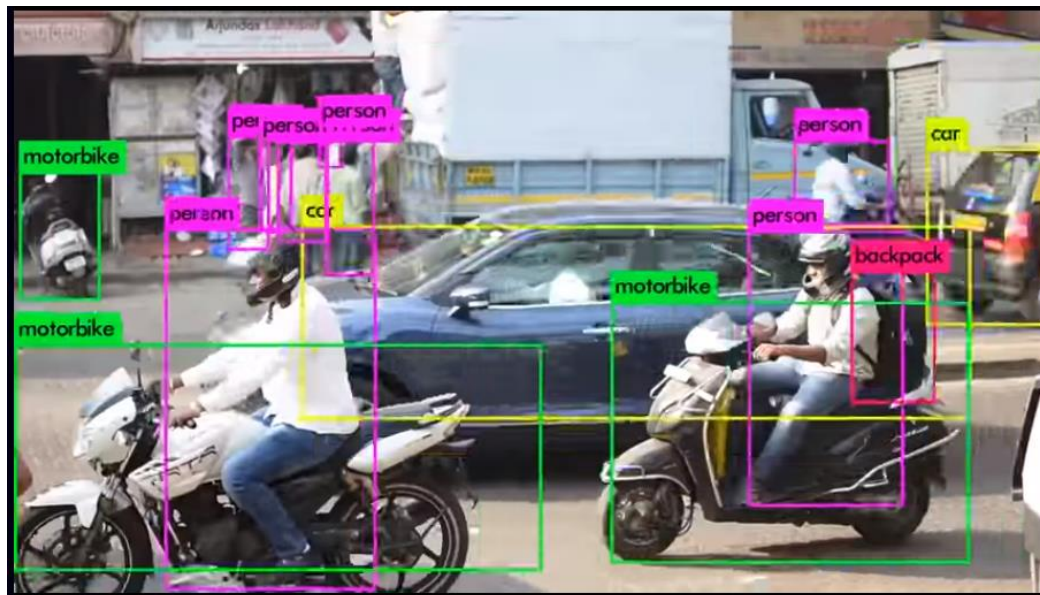
DLSS 2.0 (QUALITY MODE)

Ray Tracing - 1080p, Quality: High, RTX 2060



Tempo de inferência < 100ms

Tempo de
inferência de
20ms

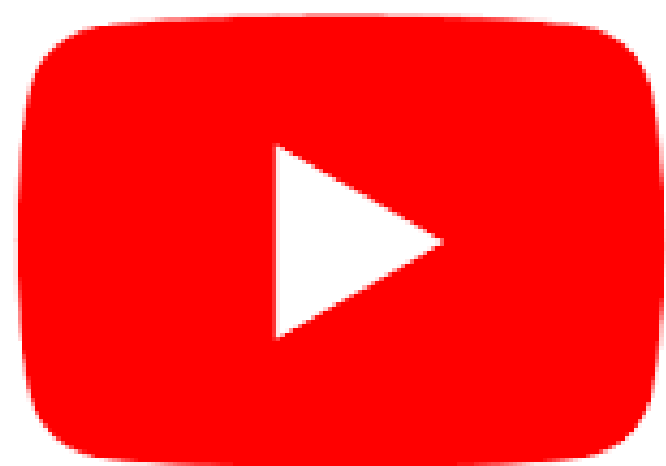


Ciência de dados

- Visualização da Informação e Aprendizado de Máquina são parte vital da Ciência de Dados (*Data Science*), gerando representações, sejam visuais ou na forma de modelos.
- Ciência de dados vincula princípios, processos e técnicas para entender um fenômeno pela análise automática de dados.
- Apoia a *Data Driven Decision Making*



NETFLIX



YouTube

Inteligência Artificial

Agentes Inteligentes

Lógica

Robótica

Otimização

Busca

Aprendizado de Máquina

Aprendizado
Supervisionado

Aprendizado
Não-Supervisionado

Aprendizado
por Reforço

Deep Learning

Definição

- Área de estudo que estuda a habilidade de computadores aprenderem sem programação explícita.
- É dito que um programa de computador aprende com a experiência **E** em respeito a uma tarefa **T** baseado na medida de desempenho **P**, se o desempenho em **T**, medido por **P**, melhora com a experiência **E**.

Pipeline



Pipeline

Aprendizado de máquina



Visualização da informação



Representação

- A percepção do computador sobre o mundo é limitada.
- Para tarefas de aprendizado de máquina, devemos escolher quais dimensões do nosso problema podem ser utilizadas.

Estudo com antibióticos



Medição do experimento

```
..., {  
  "Bacteria": "Aerobacter aerogenes",  
  "Penicillin": 870,  
  "Streptomycin": 1,  
  "Neomycin": 1.6,  
  "Gram_Staining": "negative",  
  "Genus": "other"  
},  
{  
  "Bacteria": "Bacillus anthracis",  
  "Penicillin": 0.001,  
  "Streptomycin": 0.01,  
  "Neomycin": 0.007,  
  "Gram_Staining": "positive",  
  "Genus": "other"  
}, ...
```

Vetor de características

```
[ [870, 1, 1.6, 0], [0] ]  
[ [0.001, 0.01, 0.007, 0], [1] ]
```

Tarefas

- Aprendizado Supervisionado
 - **Entrada**: Conjunto de dados com rótulo
 - **Objetivo**: Descobrir o rótulo de amostras não vistas
 - Tipos:
 - Classificação – Rótulo discreto
 - Regressão – Rótulo contínuo
- Aprendizado Não-Supervisionado
 - **Entrada**: Conjunto de dados sem resposta
 - **Objetivo**: Encontrar estrutura nos dados

Processo

Dados

- Normalização
- Transformações
- Validação
- Seleção de atributos

Treino

- Seleção de hiperparâmetros
- Seleção de métodos
- Teste e validação

Produção

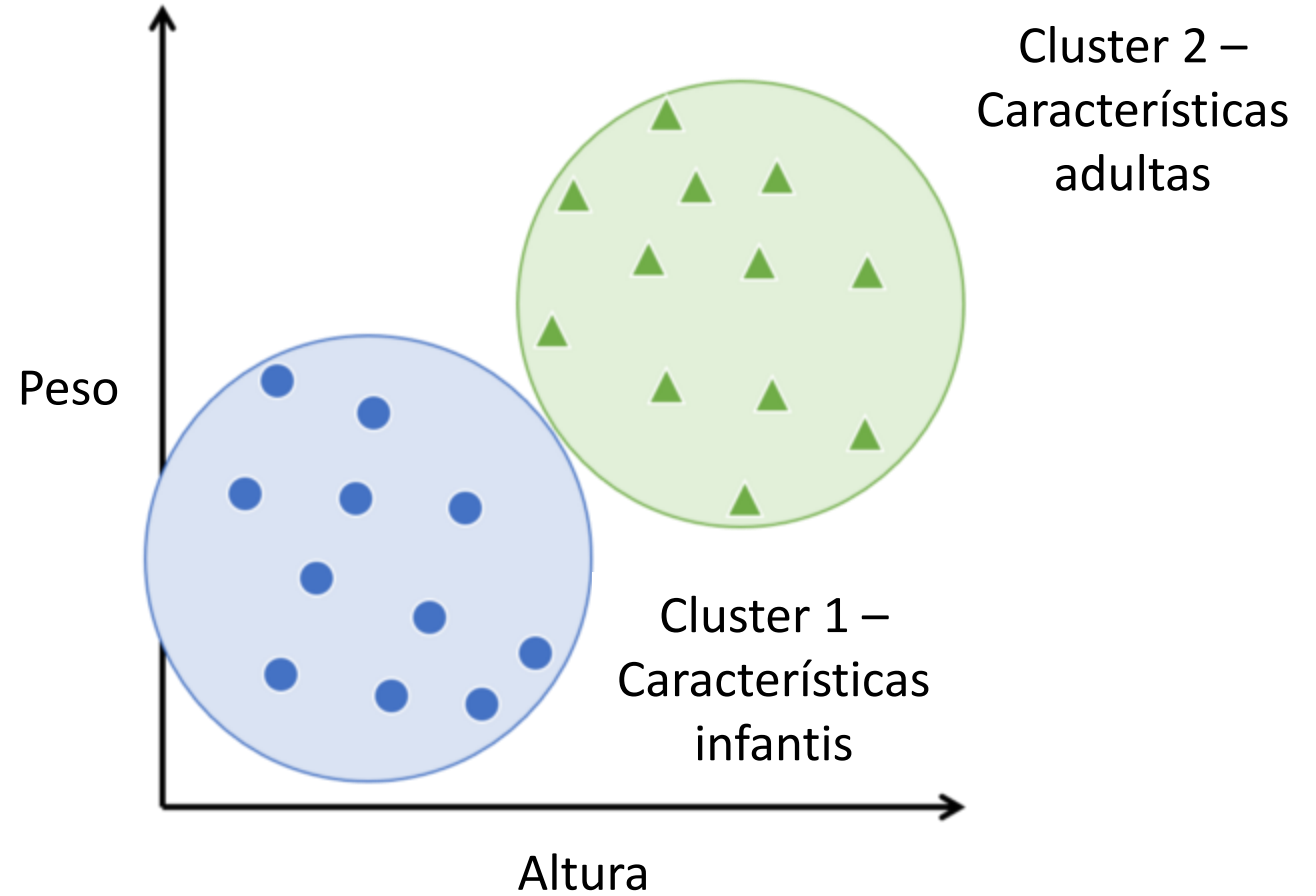
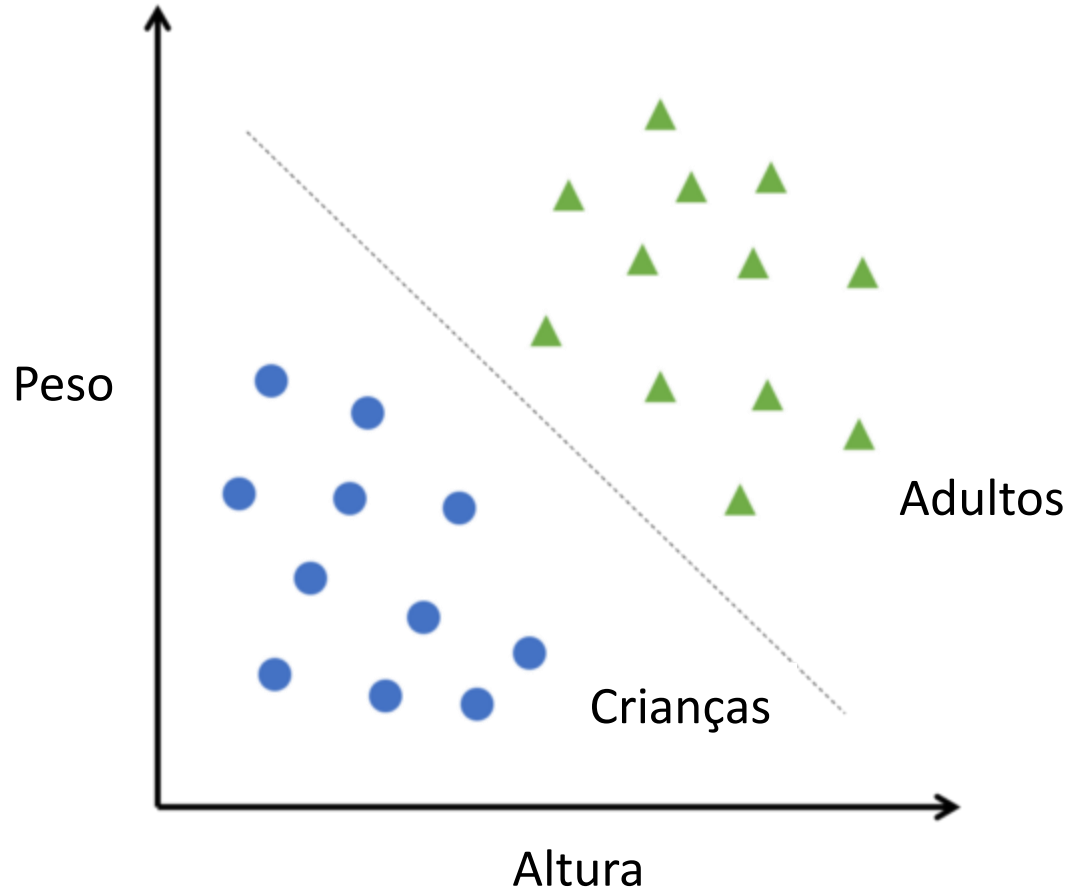
- Infraestrutura
- Monitoramento
- Refinar modelo para escalar

Tipos de Aprendizado

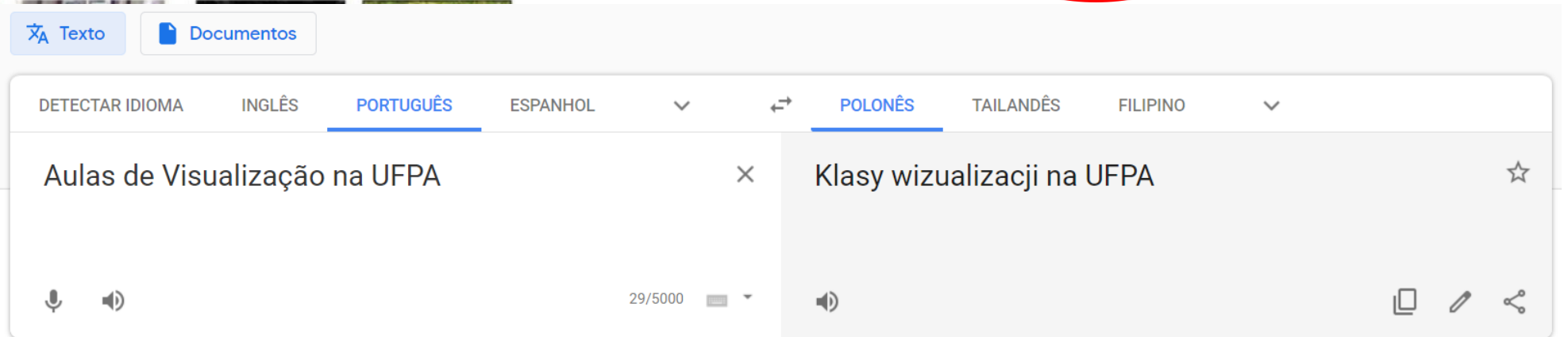
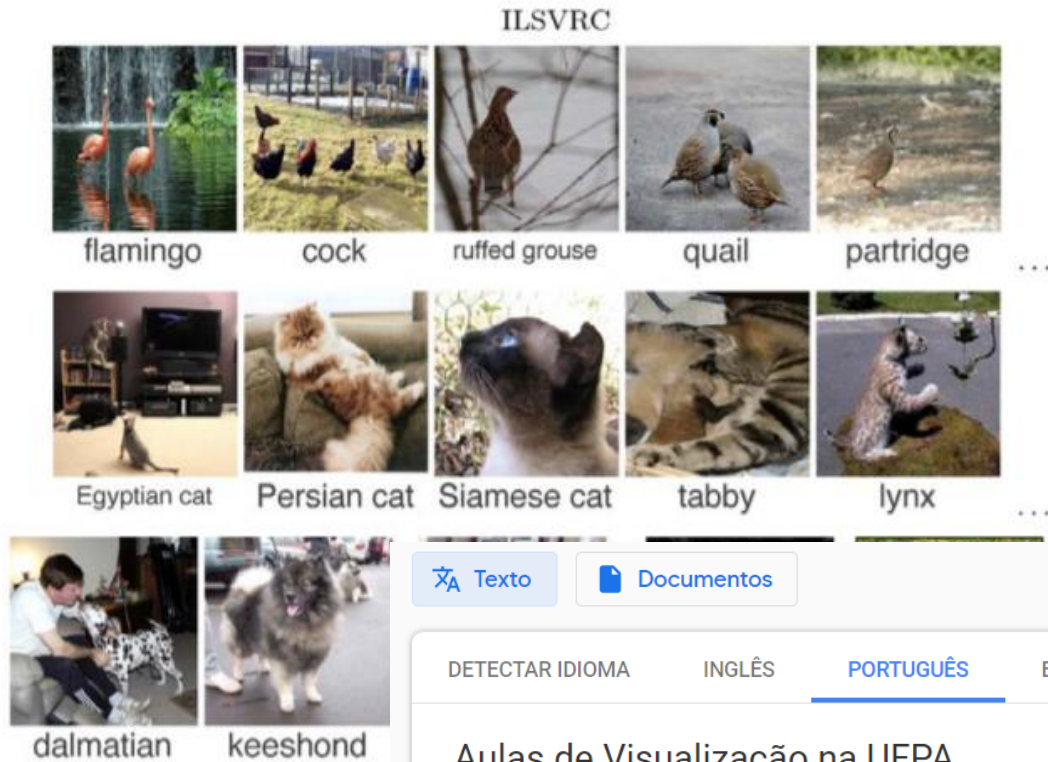
Supervisionado

VS

Não-Supervisionado



Aprendizado Supervisionado

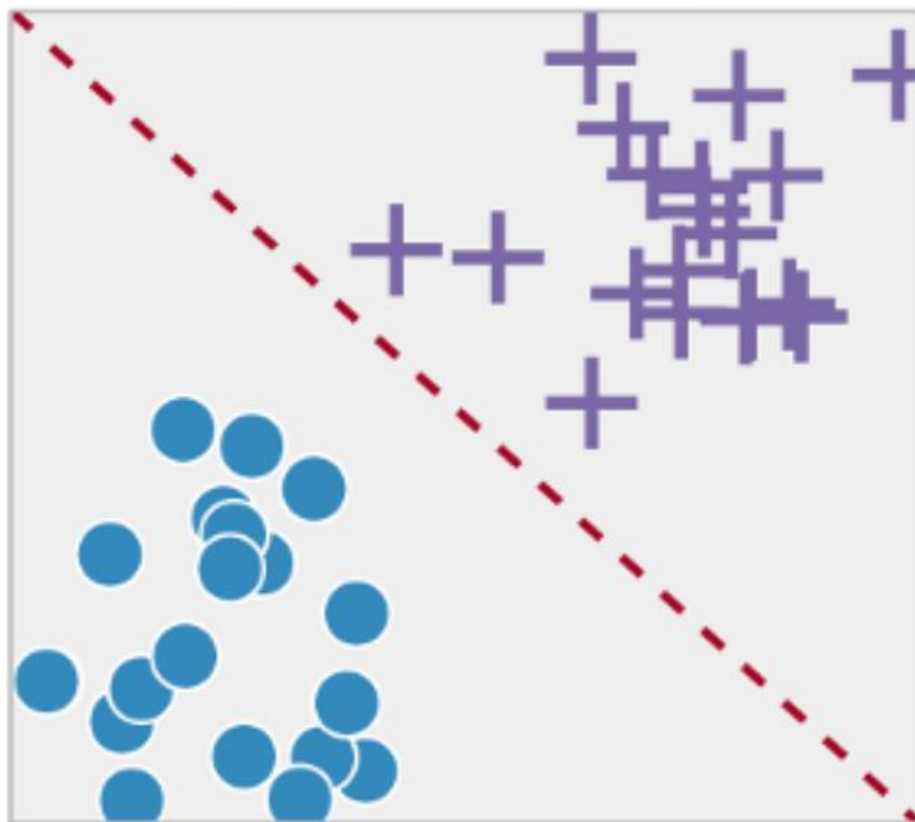


Aprendizado Supervisionado

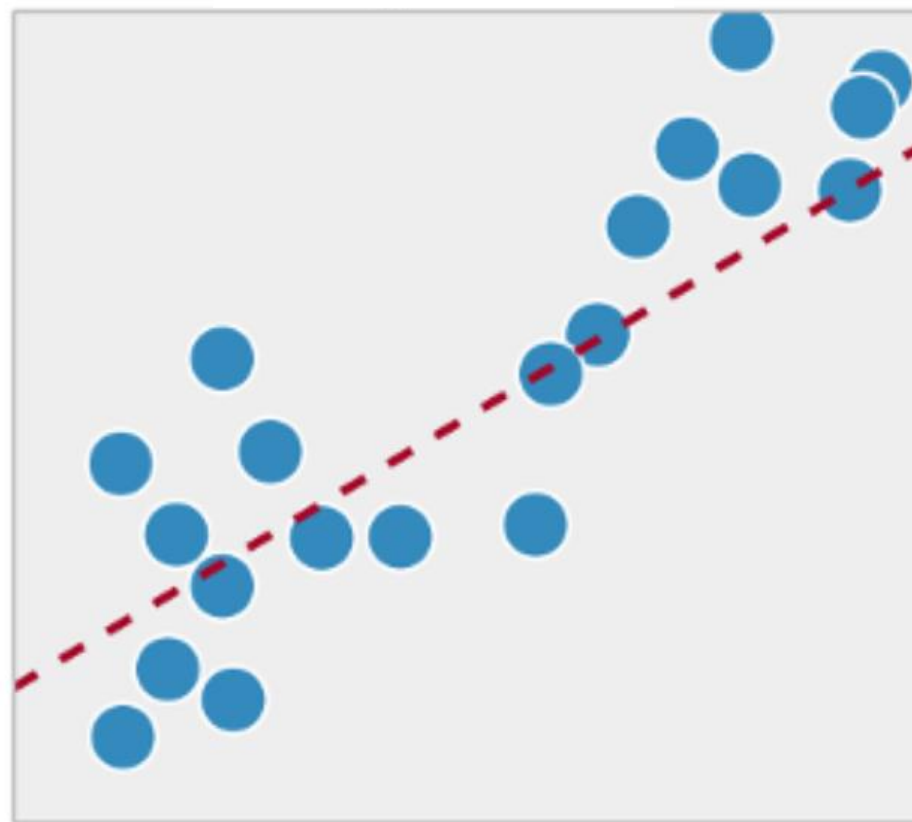
- Dado um conjunto de dados $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ um algoritmo de aprendizado procura uma função $f(X) = Y$
- X = vetores de características (x_i representa o i -ésimo vetor)
- Y = rótulos (y_i representa o i -ésimo rótulo)
- Dois tipos, depende de Y
 - Regressão quando é Y contínuo
 - Classificação quando é Y categórico

Aprendizado Supervisionado

Classificação



Regressão



Otimização

- Sem treino
- Mínimos Quadrados
- Método do Gradiente

K-Nearest Neighbor (KNN)

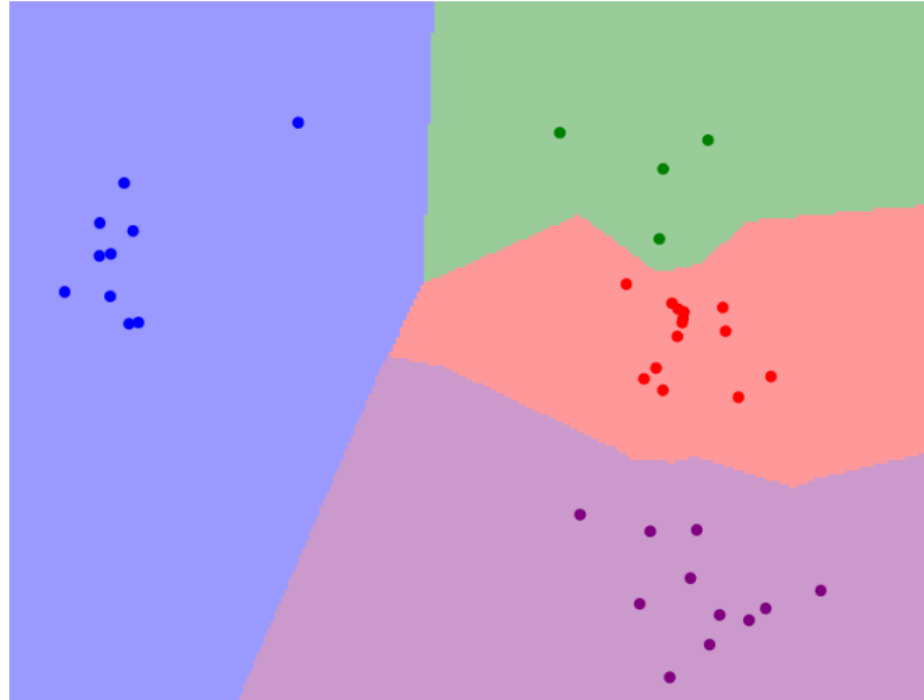
- Algoritmo simples, sem otimização
- Pontos novos procuram os vizinhos mais próximos, calculando distância
- As classes dos pontos mais próximos define a classe dos pontos novos

K-Nearest Neighbors Demo

This interactive demo lets you explore the K-Nearest Neighbors algorithm for classification.

Each point in the plane is colored with the class that would be assigned to it using the K-Nearest Neighbors algorithm. Points for which the K-Nearest Neighbor algorithm results in a tie are colored white.

You can move points around by clicking and dragging!



Metric

☒ L1 ☐ L2

Num classes

☐ 2 ☐ 3 ☒ 4 ☐ 5

Num Neighbors (K)

☒ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7

Num points

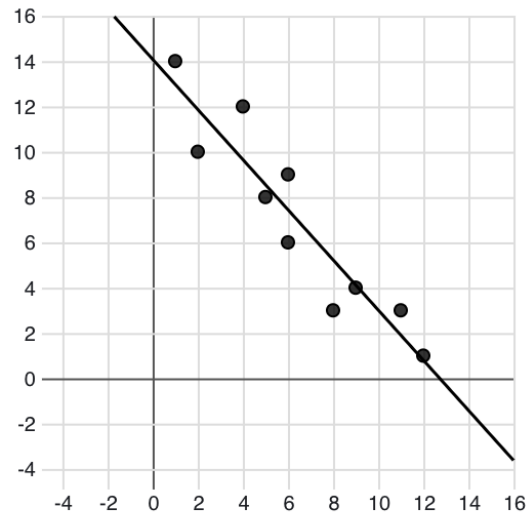
☐ 20 ☐ 30 ☒ 40 ☐ 50 ☐ 60

[K-Nearest Neighbors Demo](#)

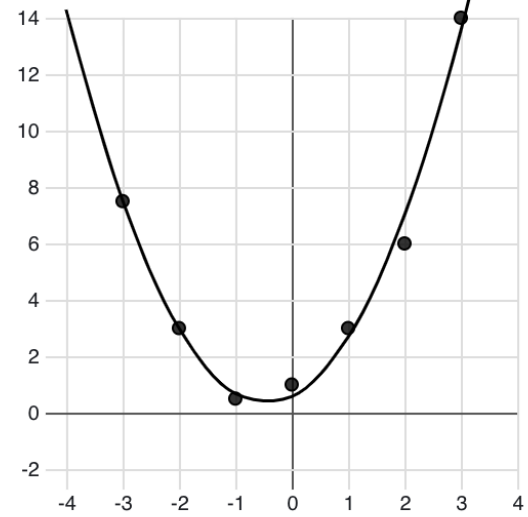
Regressão logística

- Usa uma função para definir o relacionamento entre os dados
- Pode assumir muitas formas
- Como escolher a função $f(X)$? Função de custo

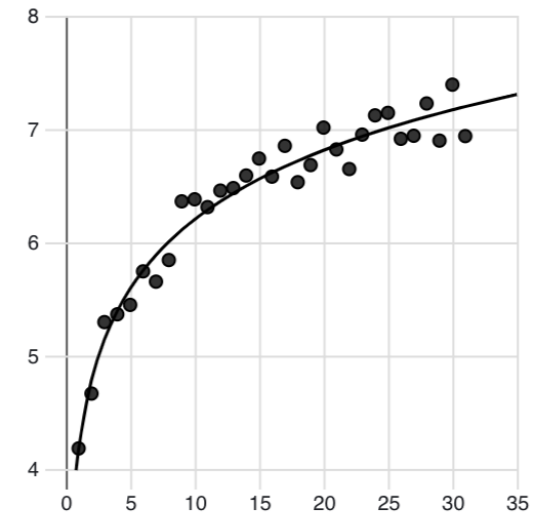
Linear `d3.regressionLinear()`



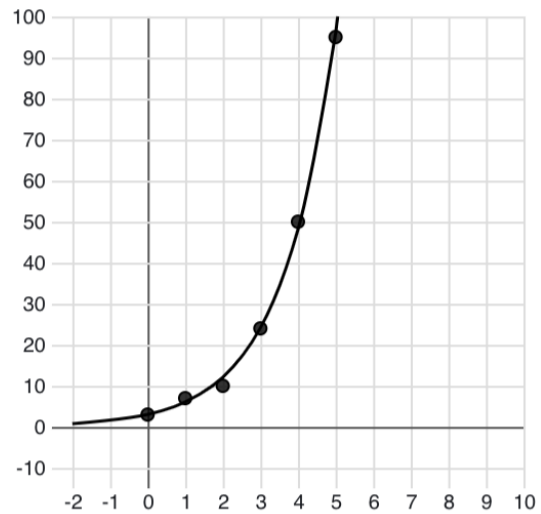
Quadratic `d3.regressionQuad()`



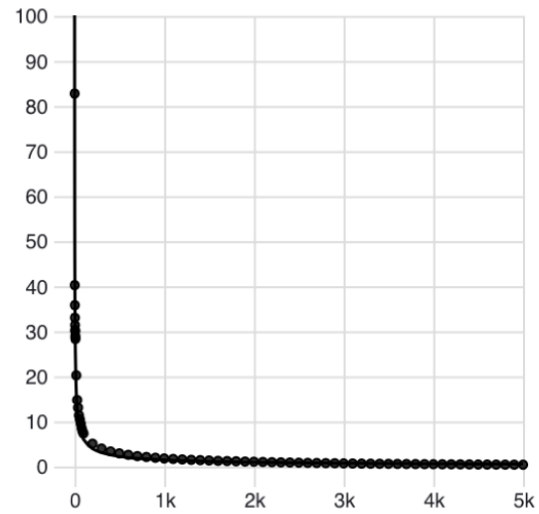
Logarithmic `d3.regressionLog()`



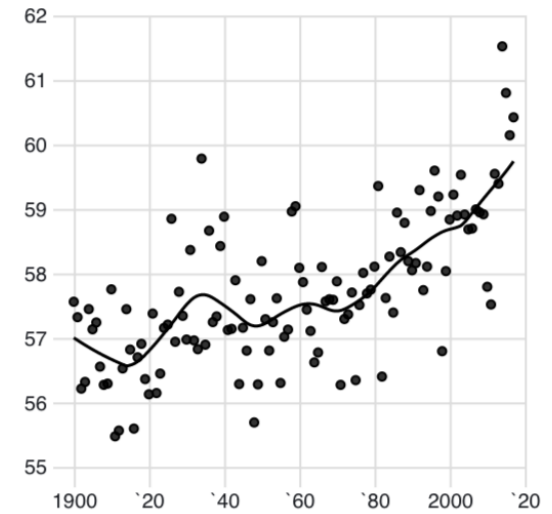
Exponential `d3.regressionExp()`



Power law `d3.regressionPow()`



LOESS `d3.regressionLoess()`



Função de Custo

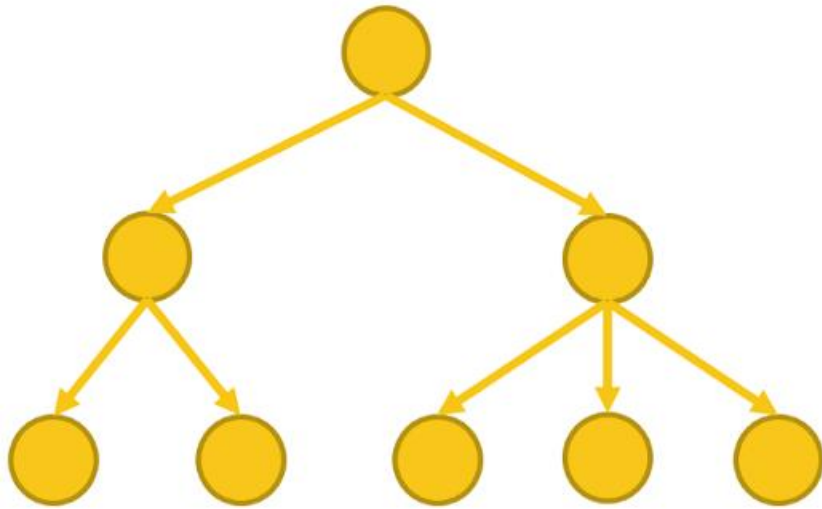
- Hipótese: $f(X) = \theta_0 + \theta_1 x$
- Paramêtros: θ_0, θ_1
- Minimizar: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (f(x_i) - y_i)^2$
- Objetivo: $\underset{\theta_0, \theta_1}{\text{minimizar}} J(\theta_0, \theta_1)$

Árvores de decisão

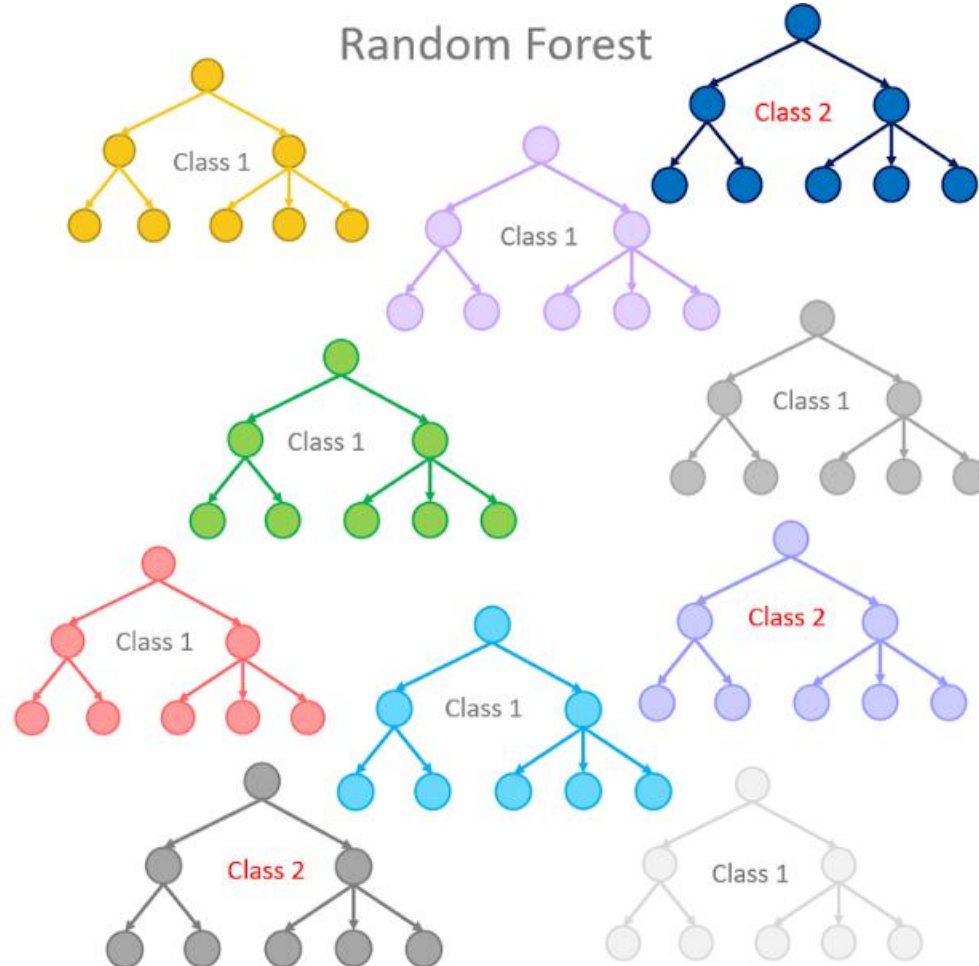
- Usa um conjunto simples de decisões hierárquicas para tomar decisões
- Simples de entender, rápido de executar
- O conjunto de árvores (floresta) perde em interpretabilidade mas ganha em acurácia
- Aprendizado pode ser baseado em entropia, que cria nós baseado no ganho de informação

Árvores de decisão e Florestas Aleatórias

Single Decision Tree

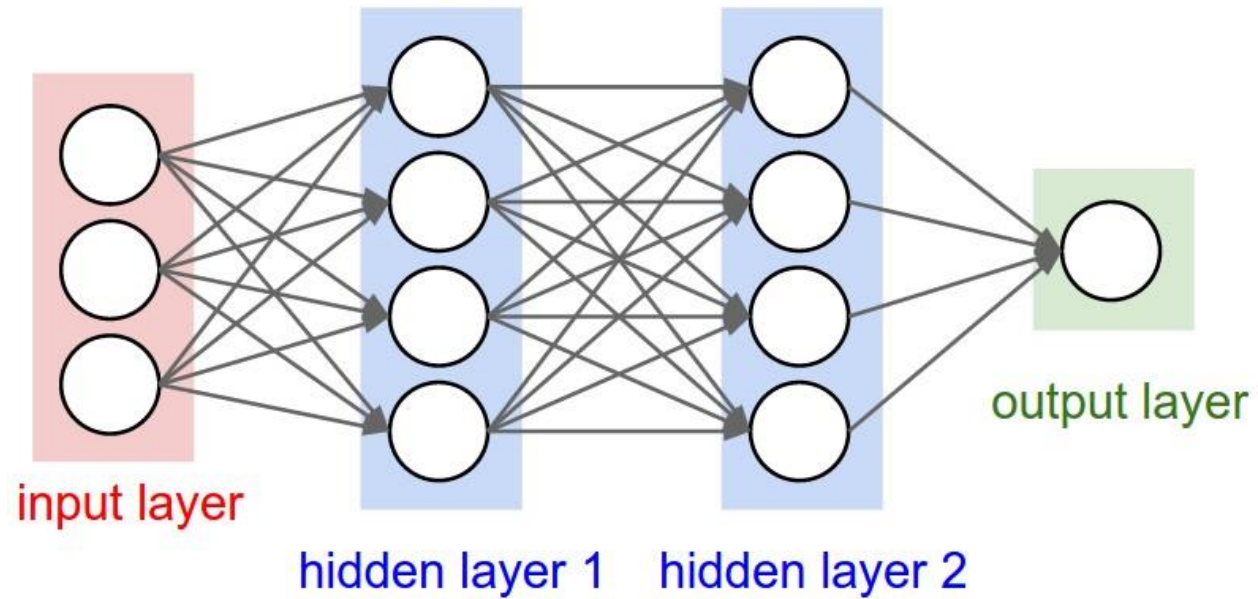
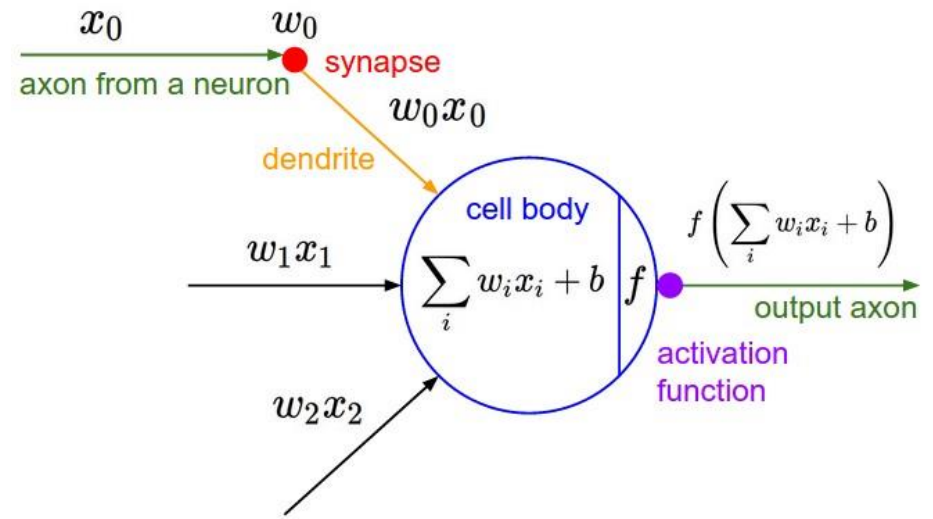
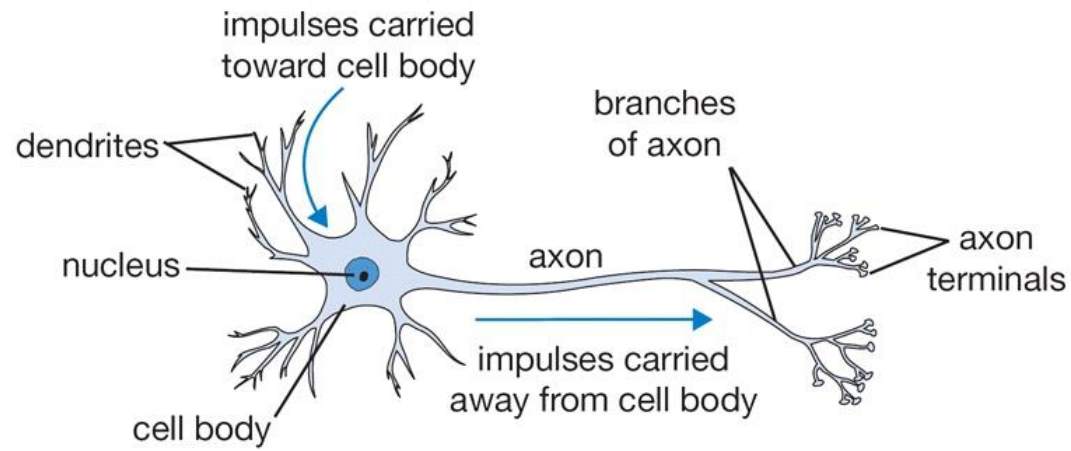


Random Forest



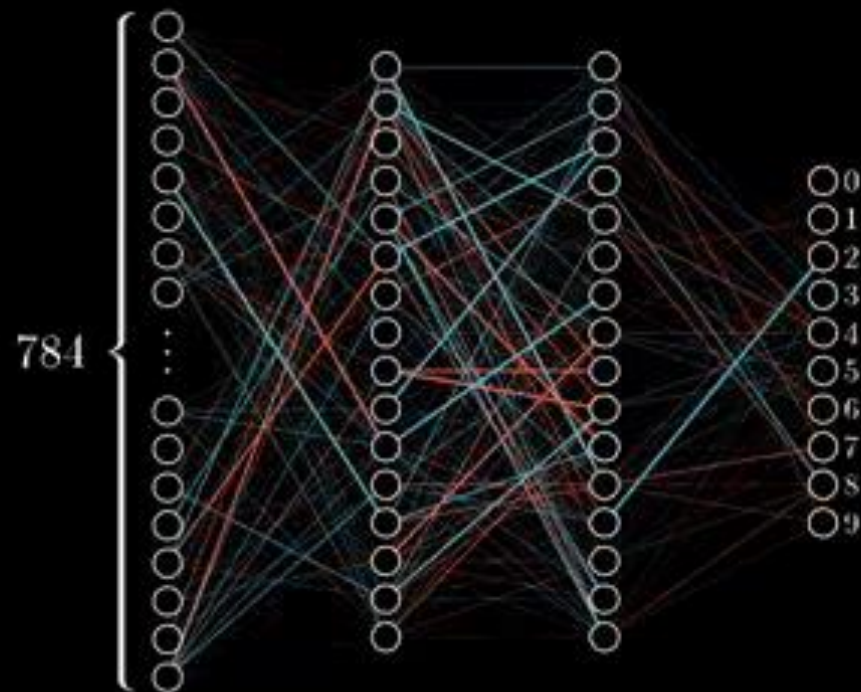
Rede Neural

- Vagamente inspirado por um modelo biológico
- Usa conexões de neurônios para calcular valores e escolher classes
- Altamente eficiente em tarefas de percepção
- O uso de Redes Neurais atende por outro nome: *Deep Learning*

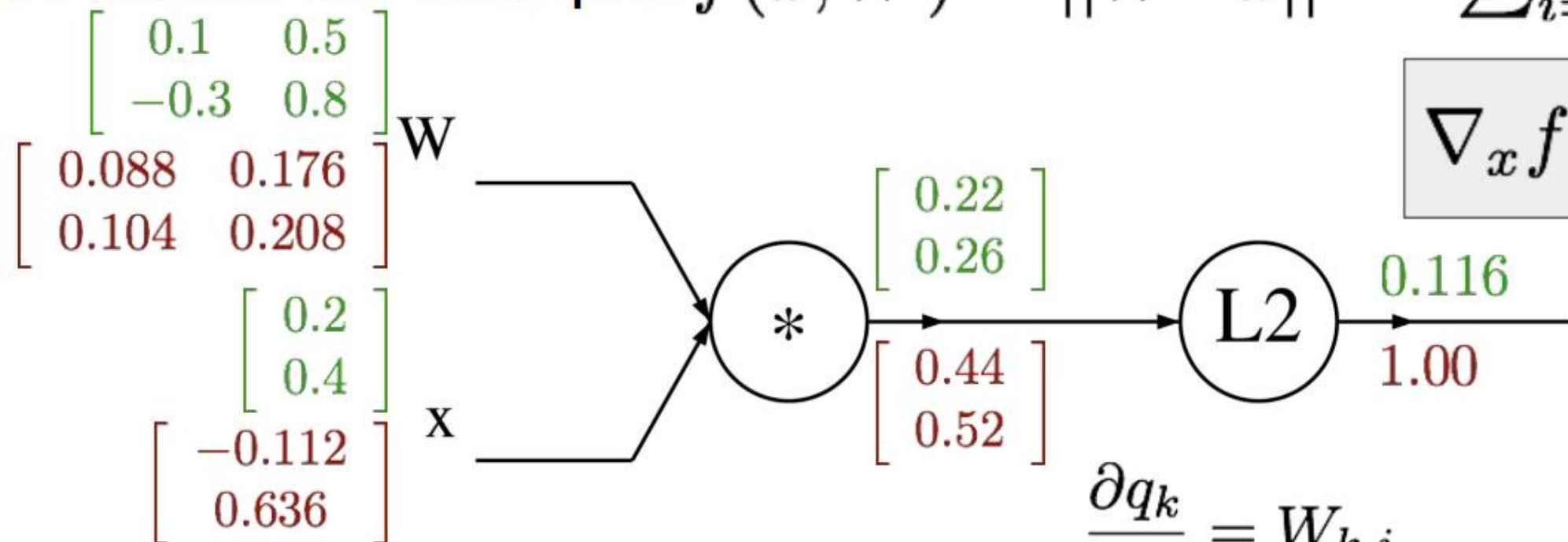


Treino

Training in
progress...



A vectorized example: $f(x, W) = ||W \cdot x||^2 = \sum_{i=1}^n (W \cdot x)_i^2$

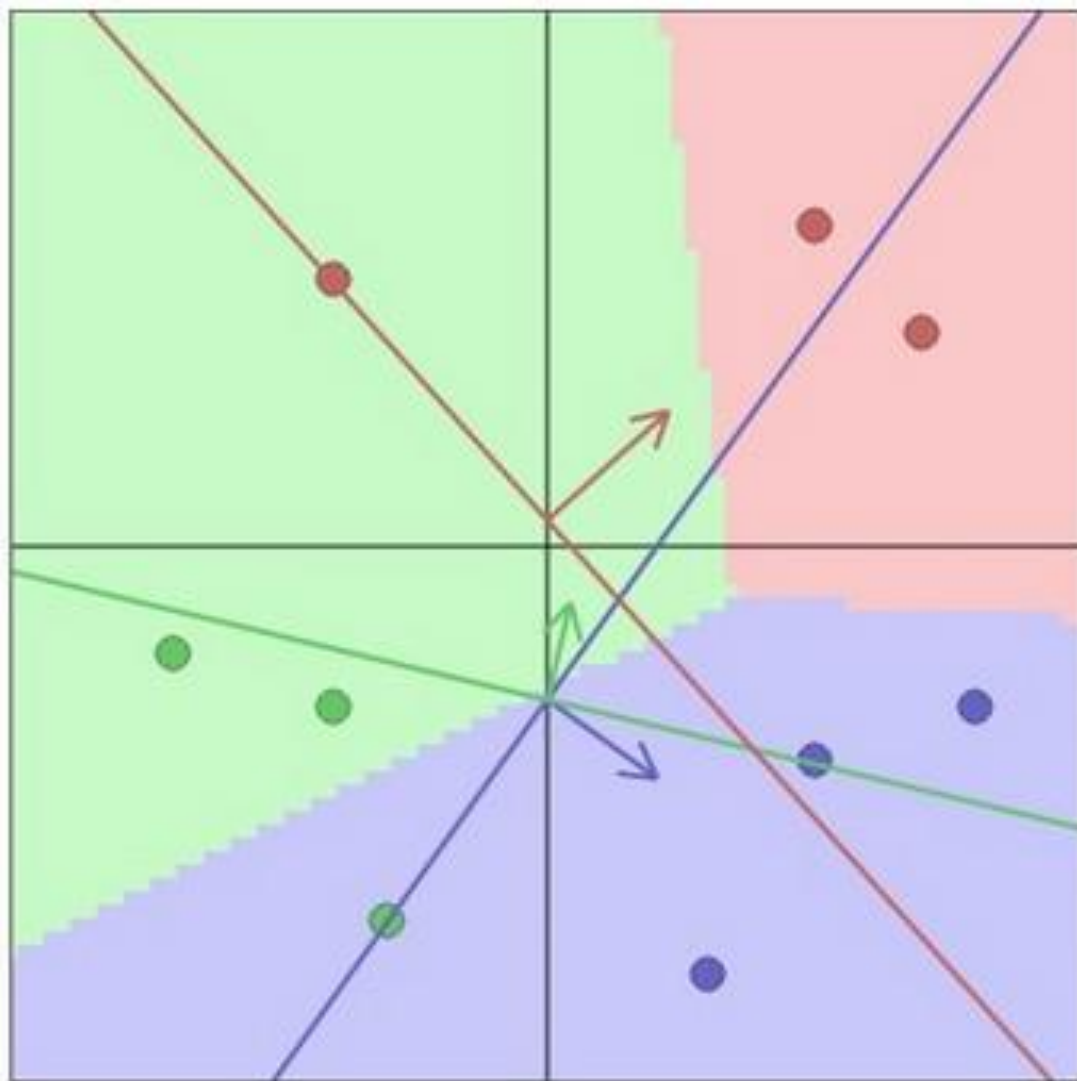


$$\nabla_x f = 2W^T \cdot q$$

$$q = W \cdot x = \begin{pmatrix} W_{1,1}x_1 + \dots + W_{1,n}x_n \\ \vdots \\ W_{n,1}x_1 + \dots + W_{n,n}x_n \end{pmatrix}$$

$$f(q) = ||q||^2 = q_1^2 + \dots + q_n^2$$

$$\begin{aligned} \frac{\partial q_k}{\partial x_i} &= W_{k,i} \\ \frac{\partial f}{\partial x_i} &= \sum_k \frac{\partial f}{\partial q_k} \frac{\partial q_k}{\partial x_i} \\ &= \sum_k 2q_k W_{k,i} \end{aligned}$$



$W[0,0]$	$W[0,1]$	$b[0]$
2.06 -0.12	1.48 0.11	-0.42 0.00
$W[1,0]$	$W[1,1]$	$b[1]$
0.44 0.19	-1.82 -0.37	0.52 0.11
$W[2,0]$	$W[2,1]$	$b[2]$
2.27 0.27	-2.04 0.14	-0.10 -0.11

Step size: 0.10000

Single parameter update

Start repeated update

Stop repeated update

Randomize parameters

$x[0]$	$x[1]$	y	$s[0]$	$s[1]$	$s[2]$	L
0.50	0.40	0	1.20	0.01	0.22	0.02
0.80	0.30	0	1.67	0.33	1.10	0.44
0.30	0.80	0	1.38	-0.80	-1.05	0.00
-0.40	0.30	1	-0.80	-0.20	-1.62	0.39
-0.30	0.70	1	-0.01	-0.88	-2.21	1.87
-0.70	0.20	1	-1.57	-0.15	-2.10	0.00
0.70	-0.40	2	0.43	1.55	2.31	0.25
0.50	-0.60	2	-0.28	1.83	2.26	0.57
-0.40	-0.50	2	-1.98	1.26	0.01	2.24

mean:
0.64

Total data loss: 0.64
Regularization loss: 1.92
Total loss: 2.57

L2 Regularization strength: 0.10000

[Linear Classification Loss Visualization](#)

Tinker With a **Neural Network** Right Here in Your Browser.

Don't Worry, You Can't Break It. We Promise.



Epoch
000,000

Learning rate
0.03

Activation
Tanh

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



FEATURES

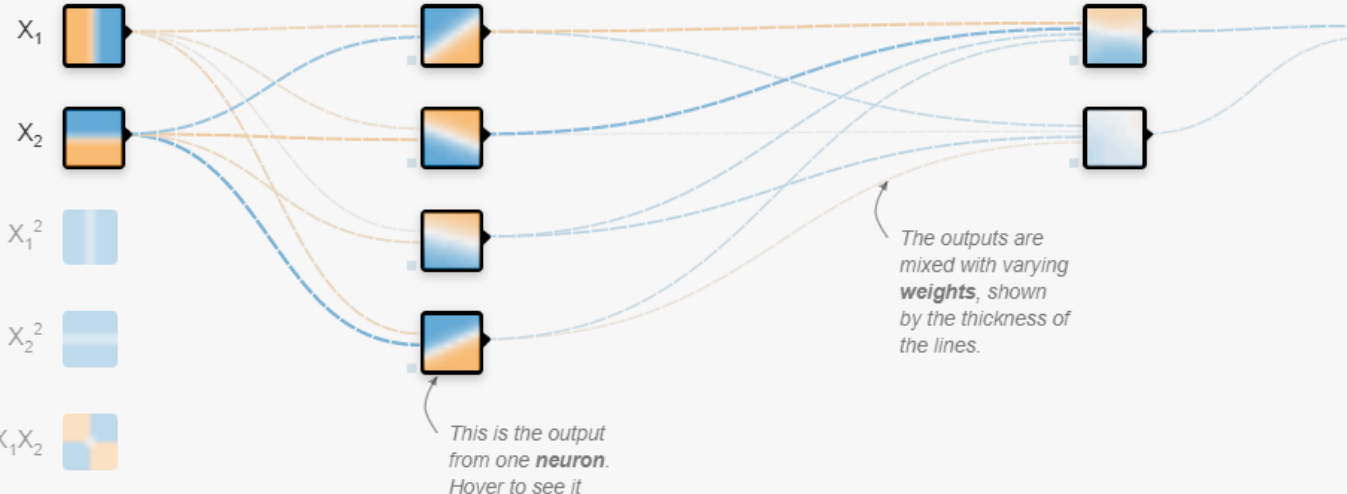
Which properties do you want to feed in?



+ - 2 HIDDEN LAYERS

+ -
4 neurons

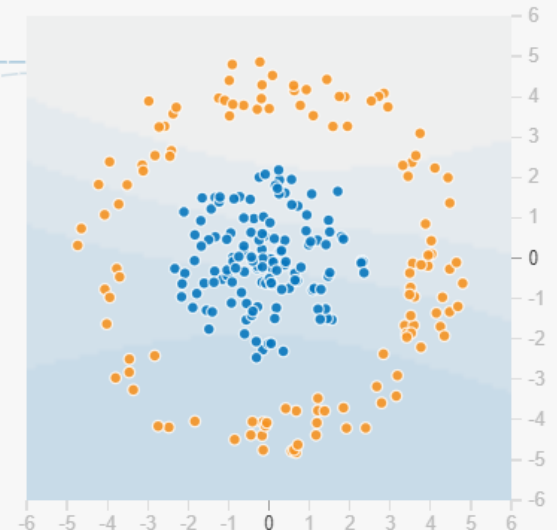
+ -
2 neurons



OUTPUT

Test loss 0.514

Training loss 0.499



[TensorFlow Playground](#)

Aprendizado não-supervisionado

- Dado um conjunto de dados $\{x_1, x_2, \dots, x_n\}$ um algoritmo de aprendizado procura estruturas latentes
- Tarefa comum: clusterização
- Maior parte da informação disponível hoje não tem rótulo





[This person does not exist](#)

Clusterização

- Mesmo sem rótulo, encontrar meios de agrupar informação tem muitas utilidades
- Exemplos:
 - Identificação de células cancerígenas
 - Detecção de fraudes e dispor linhas de crédito
 - Propaganda
 - Redes sociais
 - Análise de astros

K-Means

- Método simples de agrupamento, útil quando se sabe um pouco do domínio
- Parâmetro k define o número de clusters, que são definidos em razão de um centroide.
- Os pontos se agrupam nos clusters pela distância até os centroides, que são atualizados em cada etapa
- Muito sensível

DBSCAN

- Procura clusters pela densidade de pontos
- Parâmetros ϵ , *minPoints* : distância e densidade esperada
- Os pontos se agrupam nos clusters pela distância até os centroides, que são atualizados em cada etapa
- Um pouco mais robusto, mas depende muito da densidade

