

Machine Learning

Aula 02 $\{x_n\}: x_n = \frac{1}{n}; \{y_n\} =$

max, $\lim_{n \rightarrow \infty} \sqrt[n]{D+O+O} \leq \sqrt[n]{13^n + 13^n} = \sqrt[n]{2 \cdot 13^n} = \sqrt[n]{2} \cdot \sqrt[n]{13^n} = \sqrt[n]{2} \cdot 13$

Evandro J.R. Silva

Uninassau Teresina



Sumário

1 Tipos de Sistemas do Aprendizado de Máquina

Aprendizado Supervisionado

Aprendizado não Supervisionado

Aprendizado Semissupervisionado

Aprendizado por Reforço

Aprendizado em Batch

Aprendizado online

Aprendizado baseado em instâncias

Aprendizado baseado em modelo

2 FIM

Seção Atual

1

Tipos de Sistemas do Aprendizado de Máquina

Aprendizado Supervisionado

Aprendizado não Supervisionado

Aprendizado Semissupervisionado

Aprendizado por Reforço

Aprendizado em Batch

Aprendizado online

Aprendizado baseado em instâncias

Aprendizado baseado em modelo

2 FIM

Tipos de Sistemas do Aprendizado de Máquina

- Os tipos de aprendizado são comumente categorizados da seguinte forma:

- Quanto a supervisão

Tipos de Sistemas do Aprendizado de Máquina

Os tipos de aprendizado são comumente categorizados da seguinte forma:

- Quanto a supervisão
 - Aprendizado supervisionado;
 - Aprendizado não-supervisionado;
 - Aprendizado semissupervisionado;
 - Aprendizado por reforço;

Tipos de Sistemas do Aprendizado de Máquina

Os tipos de aprendizado são comumente categorizados da seguinte forma:

- Quanto a supervisão
 - Aprendizado supervisionado;
 - Aprendizado não-supervisionado;
 - Aprendizado semissupervisionado;
 - Aprendizado por reforço;
- Quanto a apresentação de instâncias

Tipos de Sistemas do Aprendizado de Máquina

Os tipos de aprendizado são comumente categorizados da seguinte forma:

- Quanto a supervisão
 - Aprendizado supervisionado;
 - Aprendizado não-supervisionado;
 - Aprendizado semissupervisionado;
 - Aprendizado por reforço;
- Quanto a apresentação de instâncias
 - Aprendizado online;
 - Aprendizado em batch (ou lote);

Tipos de Sistemas do Aprendizado de Máquina

Os tipos de aprendizado são comumente categorizados da seguinte forma:

- Quanto a supervisão
 - Aprendizado supervisionado;
 - Aprendizado não-supervisionado;
 - Aprendizado semissupervisionado;
 - Aprendizado por reforço;
- Quanto a apresentação de instâncias
 - Aprendizado online;
 - Aprendizado em batch (ou lote);
- Quanto à forma em que funciona/opera
 - Aprendizado baseado em instâncias (também chamados de *lazy learners*);
 - Aprendizado baseado em modelo (também chamados de *eager learners*).

Aprendizado Supervisionado

- As atividades associadas a esse tipo de aprendizado são: **classificação e regressão.**
- Todas as instâncias que são passadas ao algoritmo na fase de treinamento têm sua respectiva resposta correta.

Aprendizado Supervisionado

$$\begin{aligned}
 &= 5 \cup n \sim [n] \subset \mathbb{N}^g \\
 &\{y_n\} \neq 0 \Leftrightarrow y_n \neq 0 \text{ by } \\
 &N \rightarrow R \text{ s.t. } p \\
 &\sqrt{5^n} \left\{ \frac{1}{n} \right\} A_y \\
 &x_c = \left\{ 1 + \frac{1}{n} \right\} x_n + y_n \\
 &f(x) \Leftrightarrow \exists q \in [0, 1] : \forall \\
 &(x_n - g) < \varepsilon \quad n \geq n_0 : (x_n - \\
 &\left\{ \frac{1}{n} \right\} = \left\{ \frac{1}{n+1} \right\} \quad x_n = \\
 &\left\{ 1 + \frac{1}{n} \right\} \quad x_n \leq \\
 &\varepsilon \in [0, 1] \quad \left\{ \frac{1}{n} \right\} \quad q \\
 &\dots \quad r, t, l, \dots
 \end{aligned}$$



Parachoque
Farol
Parabrisa/teto
Rodas
Aerofólio



Aprendizado Supervisionado

$$\begin{aligned} & \{x_n\}_{n=1}^{\infty} \subset \mathbb{R}^d \\ & y_n \neq 0 \Leftrightarrow y_n = 1 \text{ ou } -1 \\ & N \rightarrow R \text{ s.t. } p \\ & \sqrt{N} \left\{ \frac{1}{n} \right\} A_y \\ & x_c = \left\{ 1 + \frac{1}{n} \right\} x_n + y_n \\ & f(x) \Leftrightarrow \exists q \in [0, 1] : \forall \\ & (x_n - g) < \epsilon \quad n \geq n_0 : (x_n - g) \in \left[\frac{1}{n}, \frac{1}{n+1} \right] \\ & \left(\frac{1}{n}, \frac{1}{n+1} \right] = \left[\frac{1}{n}, \frac{n+1}{n} \right] \quad x_n \leq q \\ & \epsilon \in [0, 1] \quad \left\{ \frac{1}{n} \right\} \quad \left(\frac{1}{n}, \frac{1}{n+1} \right] \end{aligned}$$

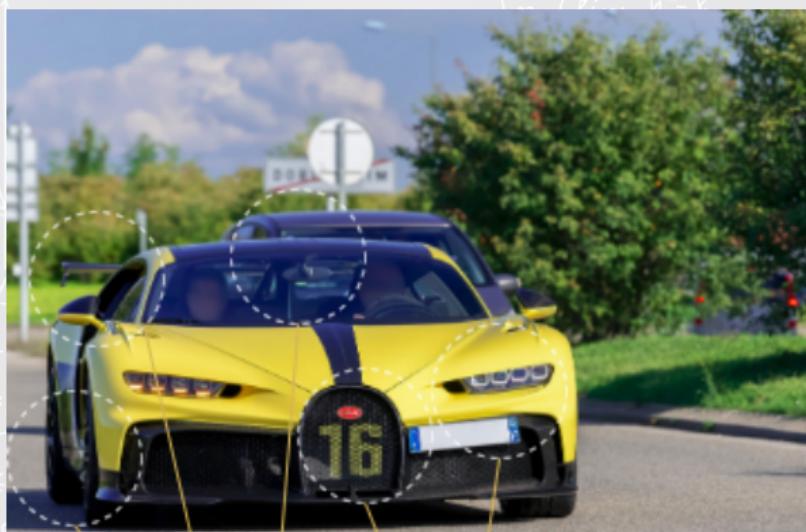


Parachoque
Farol
Parabrisa/teto
Rodas
Aerofólio



Aprendizado Supervisionado

$$\begin{aligned}
 & \{x_n\} \subset \mathbb{R}^n \quad \{y_n\} \subset \mathbb{R}^m \\
 & \{y_n\} \neq 0 \Leftrightarrow y_n \neq 0_{B_y} \\
 & N \rightarrow R \quad x: p \\
 & \sqrt[3]{5^n} \left\{ \frac{1}{n} \right\} A_y \\
 & x_c = \left\{ 1 + \frac{1}{n} \right\} x_n + y_n \\
 & f(x) \Leftrightarrow \exists q \in [0, 1] : \forall \\
 & (x_n - g) < \epsilon \quad n \geq n_0 : (x_n - \\
 & \left\{ \frac{1}{n} \right\} = \left\{ \frac{1}{n} \right\} \quad x_n = \\
 & \left\{ \frac{1}{n+1} \right\} \quad x_n \leq \\
 & \epsilon \in [0, 1] \quad \left\{ \frac{1}{n} \right\} \quad q \\
 & \dots \quad r, t, \dots
 \end{aligned}$$



Parachoque
Farol
Parabrisa/teto
Rodas
Aerofólio

$$\begin{aligned}
 & \left(x, \frac{\pi}{n} \right) \quad \{x_n\} \subset \mathbb{R}^n \quad n \geq \infty \\
 & \sqrt[n]{4^{n+1}} \quad \sum_{n=1}^{\infty} x_n \\
 & f(x), f'(x') \leq c \\
 & \sqrt[3]{3^n} \quad \left\{ \frac{1}{n} \right\} \\
 & \dots \quad \sqrt[3]{5^n} \quad \left\{ \frac{1}{n} \right\} \\
 & \dots \quad x - x
 \end{aligned}$$

Aprendizado Supervisionado



Aprendizado Supervisionado



Buggati ou McLaren?

Aprendizado Supervisionado

Buggati ou McLaren?



Aprendizado Supervisionado



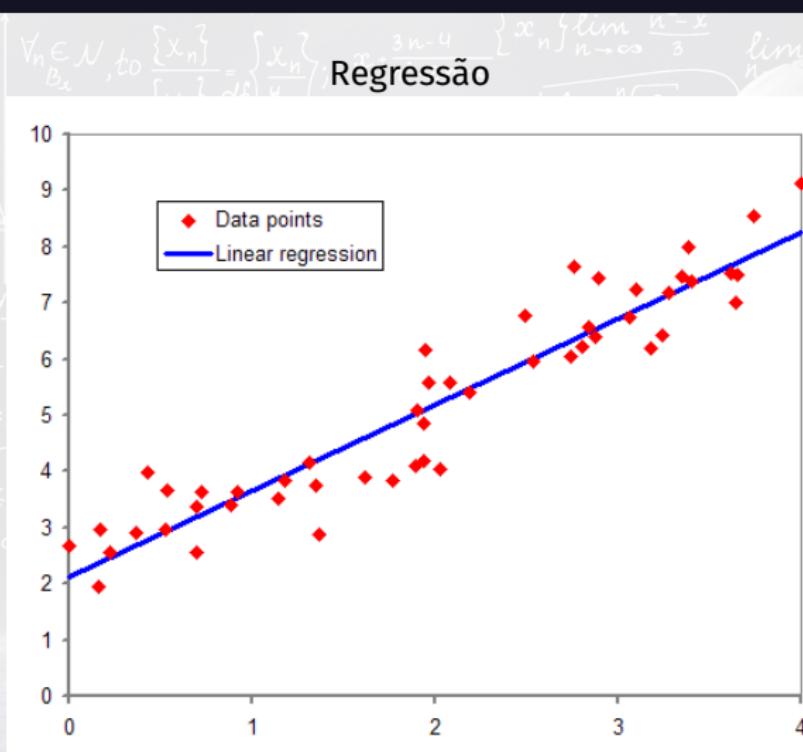
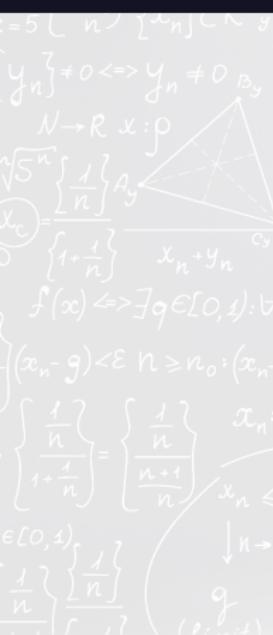
Aprendizado Supervisionado

Buggati ou McLaren?



GMG

Aprendizado Supervisionado



Regressão

Aprendizado Supervisionado

- Classificadores e preditores mais comuns

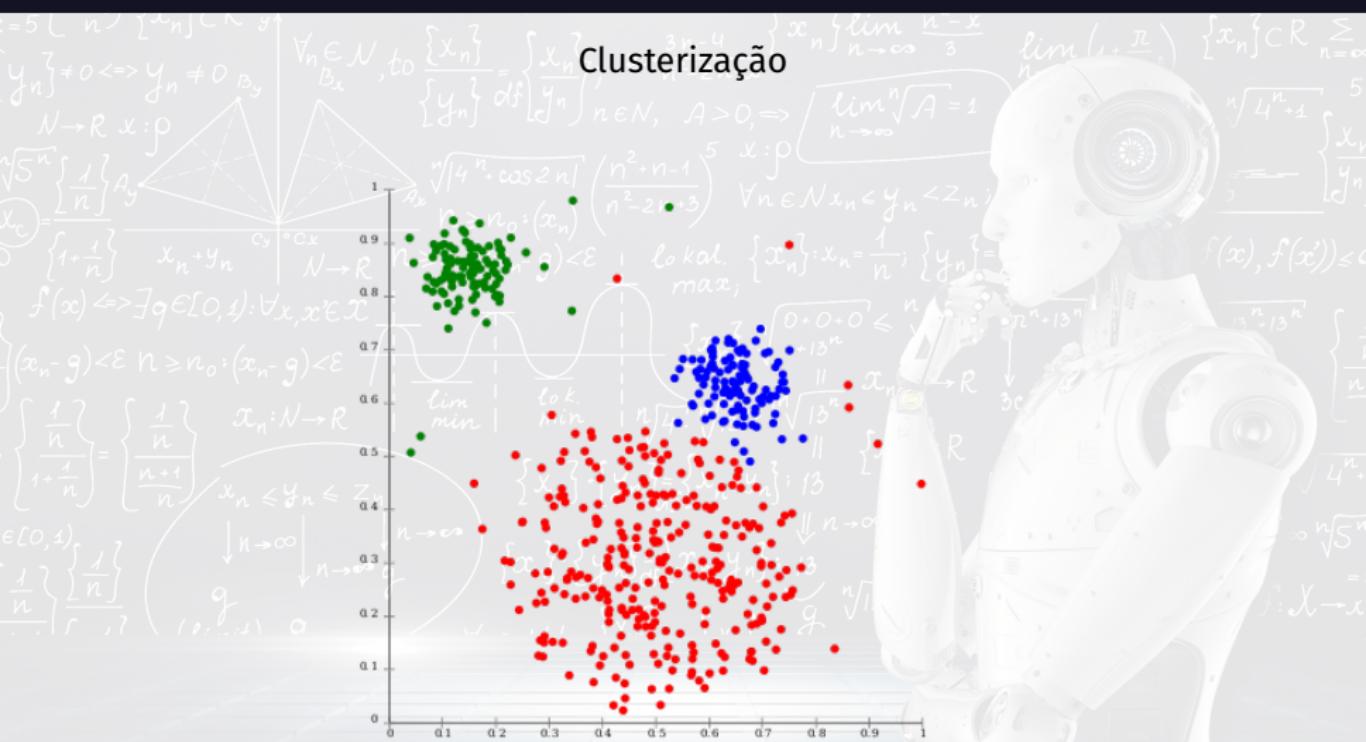
- Árvore de Decisão
- K-Vizinhos mais próximos (k-NN)
- Máquinas de Vetores de Suporte (SVM)
- Naive Bayes
- Redes Neurais Artificiais
- Regressão [Linear/Logística/Polinomial]

Aprendizado não Supervisionado

- As atividades associadas a esse tipo de aprendizado são: **clusterização, detecção de anomalias ou de novidades, visualização e redução de dimensionalidade e aprendizado de regras por associação.**

- Nenhuma instância possui categoria, ou seja, não há certo ou errado. O objetivo é que através de algum algoritmo a máquina nos dê *insights* sobre observações que são “parecidas” entre si.

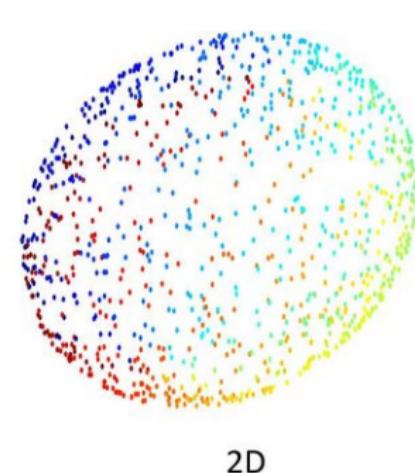
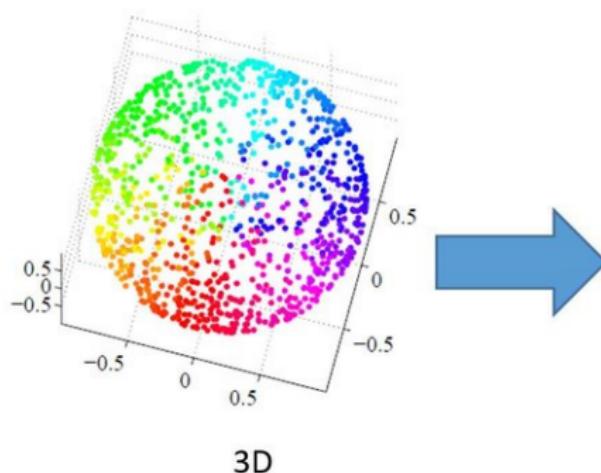
Aprendizado não Supervisionado



Aprendizado não Supervisionado

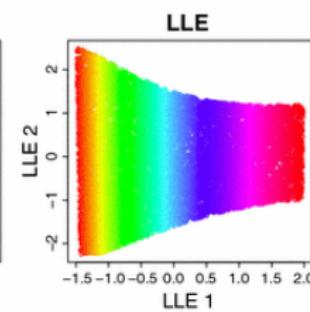
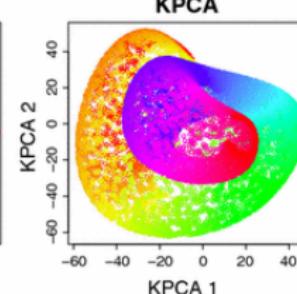
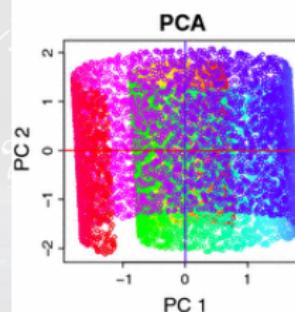
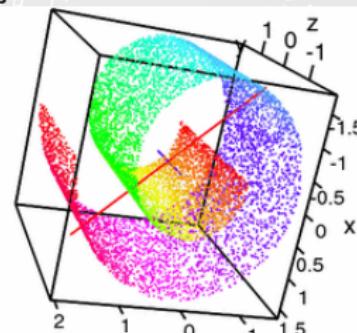
Aprendizado não Supervisionado

Redução de Dimensionalidade



Aprendizado não Supervisionado

Redução de Dimensionalidade



Aprendizado não Supervisionado

- Algoritmos mais comuns
 - K-Médias (*k-means*)
 - DBSCAN (Clusterização Espacial Baseada em Densidade de Aplicações com Ruído)
 - HCA (Análise de Cluster Hierárquico)
 - SVM com Uma Classe (*One-class SVM*)
 - Floresta de Isolamento
 - Análise de Componentes Principais (PCA)
 - LLE (*Locally Linear Embedding*)
 - t-SNE (*Distributed Stochastic Neighbor Embedding*)

Aprendizado Semissupervisionado

- As bases de dados não aparecem do nada!
- Por detrás de todas as bases de dados disponíveis houve um grande esforço pessoal, logístico e financeiro.
- Além disso, existem bases de dados que são alimentadas com novas instâncias à medida em que passa o tempo.
- Por causa disso, pode ser possível o caso de bases de dados em que nem todas as instâncias foram rotuladas.
- O **Aprendizado Semissupervisionado** ocorre em bases com dados **parcialmente rotulados**, tendo atividades do aprendizado supervisionado e também do não-supervisionado.

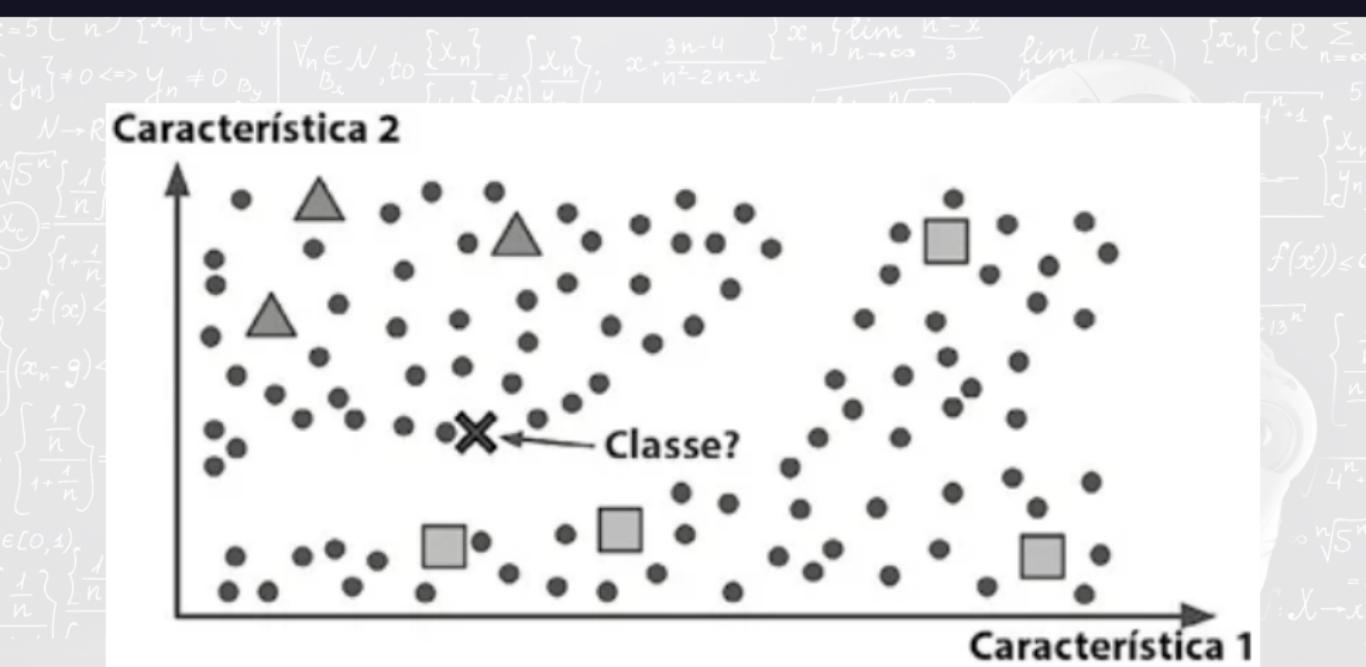
Aprendizado Semissupervisionado

Exemplo

- **Google Fotos** - suponha que você fez o upload das fotos do almoço em família do último domingo.

- A parte não-supervisionada do algoritmo vai conseguir detectar uma determinada pessoa **A** em todas as fotos em que ela aparece.
- Então você pode fornecer o rótulo daquela pessoa e informar, por exemplo, que é o seu pai. A partir de agora, é possível utilizar do aprendizado supervisionado para reconhecer o seu pai em todas as outras fotos que você tenha upado.

Aprendizado Semissupervisionado



Aprendizado Semissupervisionado

Algoritmos

- DBNs - Redes Neurais de Crenças Profundas
- RBMs - Máquinas Restritas de Boltzmann

Aprendizado por Reforço

- Lembra, bem pouco, o aprendizado supervisionado.
- Sua característica principal é **premiar ou punir** a máquina dependendo do que ela tenha decidido fazer.
- O objetivo da máquina é ser premiada o máximo possível.
- No **aprendizado supervisionado**, a máquina sabe o quanto acertou ou o quanto errou. No **aprendizado por reforço** ela só sabe se acertou ou errou.
- Exemplo real: **AlphaGo** da DeepMind, que venceu o campeão mundial de Go em 2017.

Aprendizado em *batch* (ou lote)

- O lote se refere aos dados.
- Neste tipo de aprendizado os dados são parte de um grande conjunto, o qual é apresentado inteiro para a máquina.
- Uma vez que termina o treinamento, o algoritmo estará “finalizado”, no sentido de pronto para ser usado.
- Caso surjam novos dados, o algoritmo deverá ser treinado novamente, mas do zero. Ao fim do novo treinamento, teremos uma nova versão do algoritmo.
- É também referido como aprendizado **offline**.

Aprendizado em *batch* (ou lote)

- Esse tipo de aprendizado vai requerendo mais **recursos computacionais** (HD, memória, poder de processamento) à medida da quantidade de dados a serem utilizados.
- Uma IA sobre algum problema simples, e com poucos dados, pode ser treinada em computadores comuns, incluindo tablet e smartphone.
- Uma IA sobre problemas do mundo real, extremamente complexos, com quantidades avassaladoras de dados só podem ser treinadas por grandes empresas ou governos, em super computadores, ou clusters de computadores.

Aprendizado online ou incremental

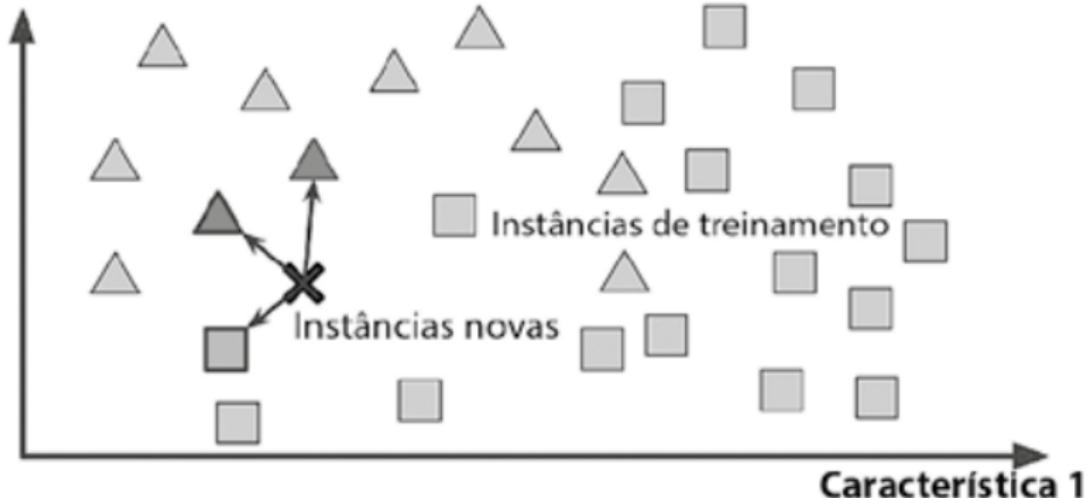
- Neste caso a IA pode ser treinada de forma incremental, com novas instâncias sendo fornecidas de forma sequencial, individual ou em pequenos grupos (ou *mini-batches*).
- A atualização da IA acontece de forma rápida e com baixo custo computacional.
- É boa para problemas que **geram dados em um fluxo contínuo**. Por exemplo, no **mercado financeiro**, onde a todo momento os preços das ações estão mudando. Outra situação é no **monitoramento de saúde**, onde dados sobre um paciente são coletados a todo instante.
- Também pode ser usada em caso de limitações de recursos computacionais.

Aprendizado baseado em instâncias

- Neste tipo de aprendizado, levamos em consideração o pressuposto de que observações de uma mesma classe são similares.
- A partir disso podemos verificar um **nível de similaridade** entre instâncias diferentes.

Aprendizado baseado em instâncias

Característica 2

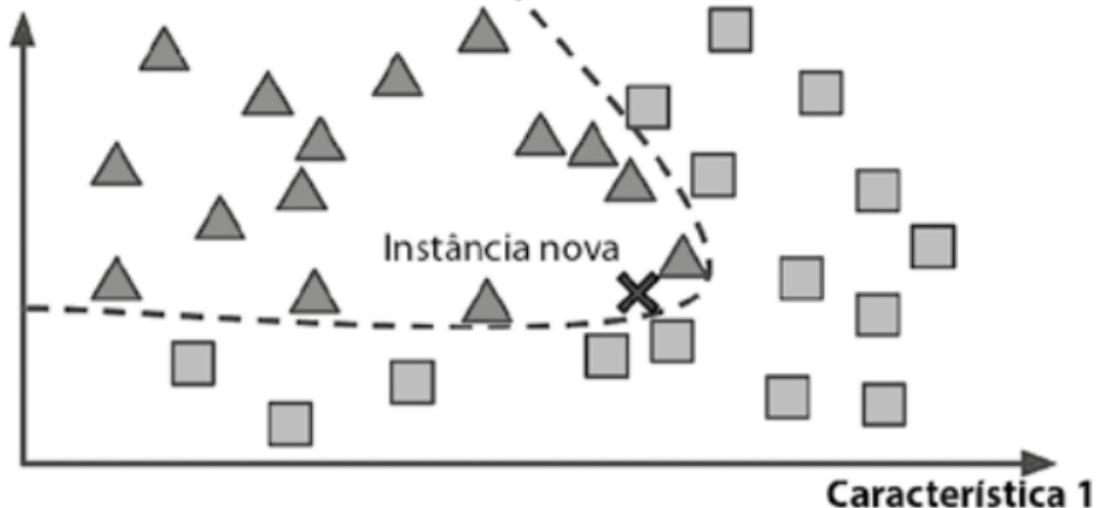


Aprendizado baseado em modelos

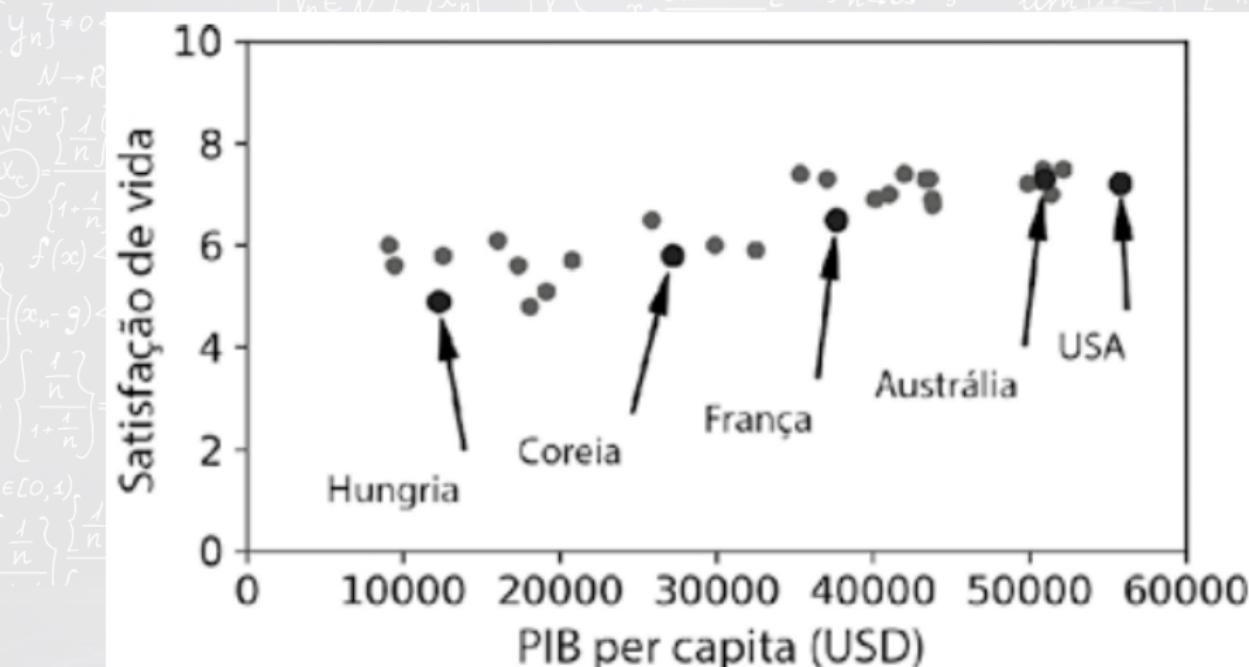
- Neste tipo de aprendizado, buscamos encontrar um modelo (literalmente uma função matemática) que melhor descreva os dados observados.

- Quanto melhor o modelo, melhor ele conseguirá realizar alguma classificação ou predição

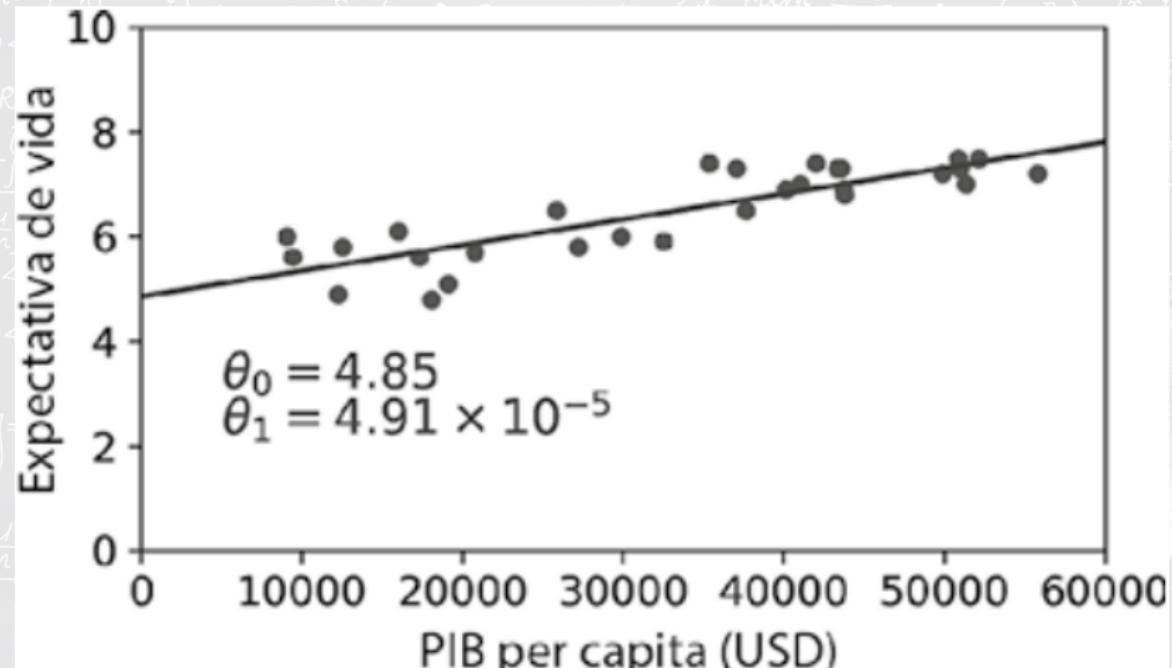
Aprendizado baseado em modelo



Aprendizado baseado em modelo



Aprendizado baseado em modelo



Exemplo de como fazer essa regressão no livro-texto!

Seção Atual

