

Aprendizagem de Máquina

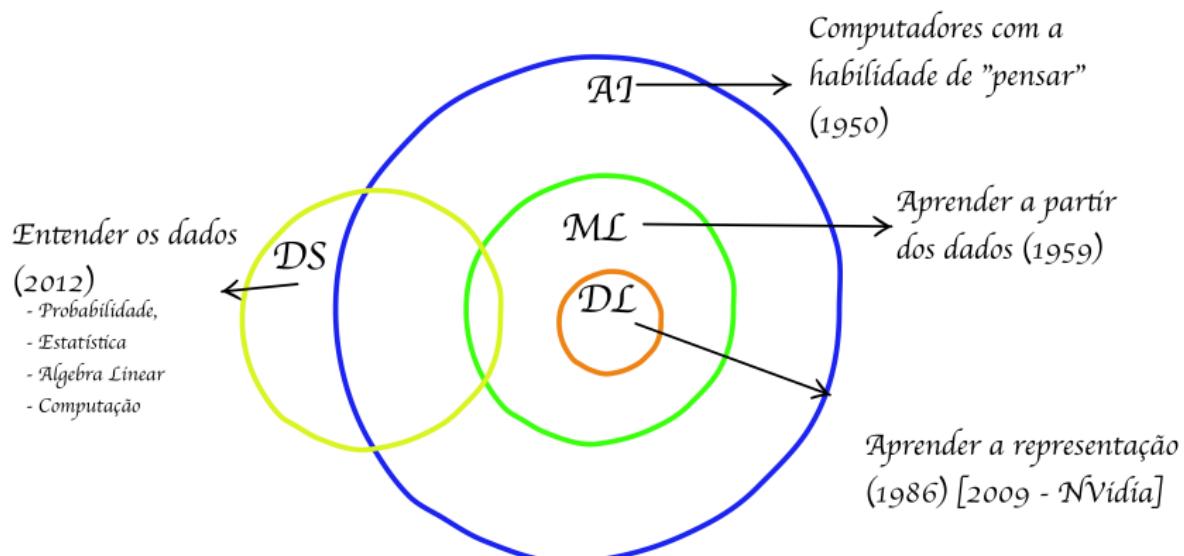
Uma visão geral

Prof. Luiz Eduardo S. Oliveira

Universidade Federal do Paraná
Departamento de Informática
<http://www.inf.ufpr.br/lesoliveira>

Qual é a diferença?

- IA (Artificial Intelligence)
- ML (Machine Learning)
- DL (Deep Learning)
- DS (Data Science)



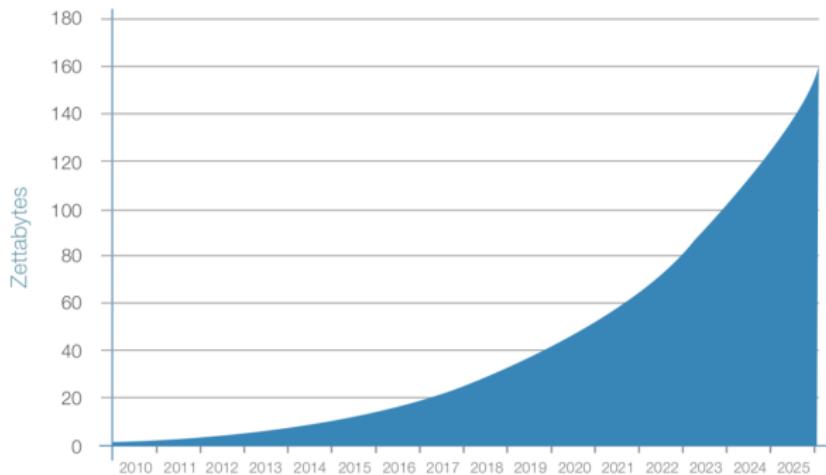
Porque agora??

Inteligência Artificial existe desde os anos 50.

- Disponibilidade de dados
- Poder computacional.
- Ferramentas.

Porque é importante?

Dados criados



Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

- Impossível para o ser humano lidar com todo esse volume de dados.
- Imagine por exemplo a quantidade de recursos necessários para olhar e rotular cada imagem do Flickr ou Google Photos.
 - ▶ Aplicação real: Perícia em computadores e telefones.



“To create the image and speech recognition algorithms designed by AutoML, Google reportedly let a cluster of **800** GPUs iterate and crunch numbers for weeks.”

<http://www.tomshardware.com/news/google-automl-artificial-intelligence-ai.34533.htm>

Machine Learning Frameworks



O que é Aprendizagem de Máquina?

- Campo de estudo que dá aos computadores a habilidade de aprender sem uma programação explícita [Arthur Samuel, 1959]

O que é Aprendizagem de Máquina?

- Campo de estudo que dá aos computadores a habilidade de aprender sem uma programação explícita [Arthur Samuel, 1959]
- Pode-se dizer que um programa de computador aprende uma experiência E a respeito de uma tarefa T e alguma medida de desempenho P, se seu desempenho em T, medido por P, melhora com a experiência E. [Tom Mitchel, 1997]

O que é Aprendizagem de Máquina?

- Campo de estudo que dá aos computadores a habilidade de aprender sem uma programação explícita [Arthur Samuel, 1959]
- Pode-se dizer que um programa de computador aprende uma experiência E a respeito de uma tarefa T e alguma medida de desempenho P, se seu desempenho em T, medido por P, melhora com a experiência E. [Tom Mitchel, 1997]
- É a ciência (e arte) de programar computadores de tal forma que eles aprendam a partir de dados [Aurélien Géron, 2017].

Desafios

- O aspecto chave da inteligência é a versatilidade, ou seja, a capacidade de fazer diversas coisas.
- Os sistemas de IA atualmente são muito bons e superam os seres humanos em algumas atividades
 - ▶ Jogos (xadrez, go, etc)
 - ▶ Diagnóstico médico (detecção de câncer de pele)
 - ▶ Geração de texto (LLM)



Quando é necessário utilizar aprendizagem de máquina?

Quando é necessário utilizar aprendizagem de máquina?

Sempre que não for possível escrever um algoritmo determinístico para resolver o problema. Nesses casos o conhecimento deve ser extraído a partir de exemplos.

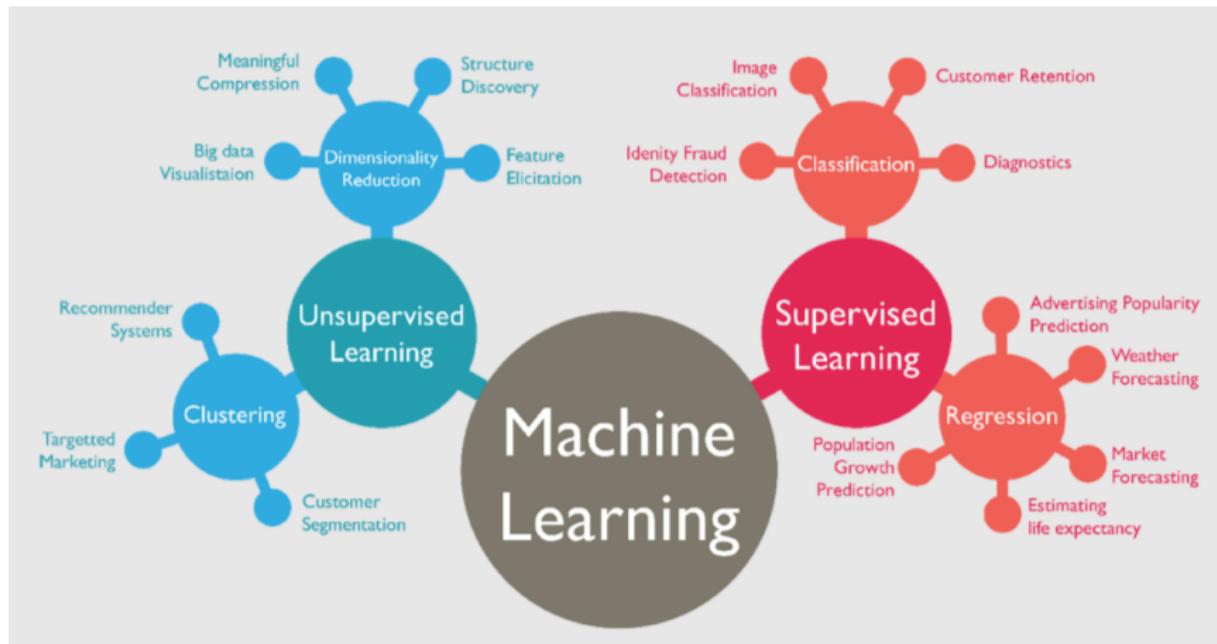
Amazon Reviews:

- Negativos:
 - ▶ The Worst!: A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.
 - ▶ Awful beyond belief!: I feel I have to write to keep others from wasting their money. This book seems to have been written by a 7th grader with poor grammatical skills for her age!
 - ▶ Oh dear: I was excited to find a book ostensibly about Muslim feminism, but this volume did not live up to the expectations
- Positivos:
 - ▶ Even Mommy has fun with this one!: My four year old daughter loves everything Barbie and loves the Rapunzel movie. This game is tons of fun, even for a 42 year old.
 - ▶ Rapunzel: This is such a great game both my 3 year old son and 7 year old daughter love it. I like to play to if they would let me!

Exemplos de Aplicações

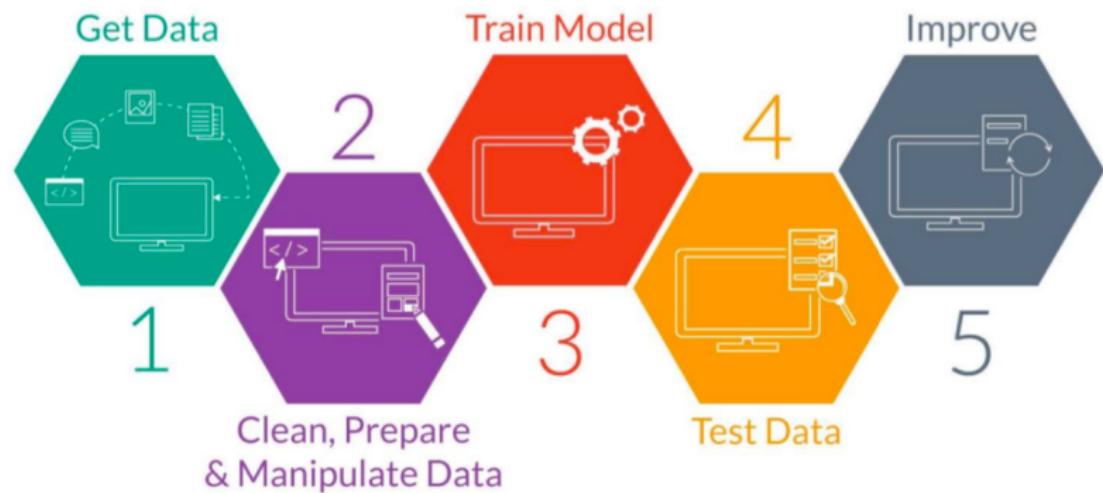


Tipos de Aprendizagem

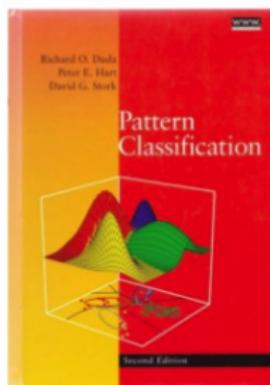


- Supervisionada: dados de treinamento estão rotulados.
- Não supervisionada: Sem rótulos; descobrir estruturas.
- Semi-supervisionada: Uso de dados não rotulados para melhorar a aprendizagem supervisionada.

Pipeline



Aprendizagem Supervisionada - Classificação



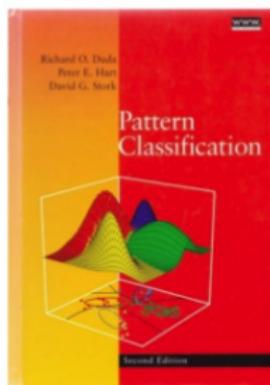
Robalo



Salmão

- Construir um sistema que classifique automaticamente dois tipos de peixe: Salmão e Robalo.
 - ▶ Precisamos de aprendizagem de máquina?

Aprendizagem Supervisionada - Classificação



Robalo



Salmão

- Construir um sistema que classifique automaticamente dois tipos de peixe: Salmão e Robalo.
 - ▶ Precisamos de aprendizagem de máquina?
 - ▶ Sim! O que precisamos para construir o sistema?

Aprendizagem Supervisionada - Classificação

1. Dados Rotulados



Salmão



- Esse conjunto de dados rotulados é chamado de base de treinamento.

Aprendizagem Supervisionada - Classificação

2. Representação

- O que diferencia um salmão de um robalo?
- Extrair atributos, os quais sejam similares aos objetos de uma mesma classe e diferentes para objetos das outras classes.
- Processo conhecido com **Extração de Características**



Aprendizagem Supervisionada - Classificação

2. Representação

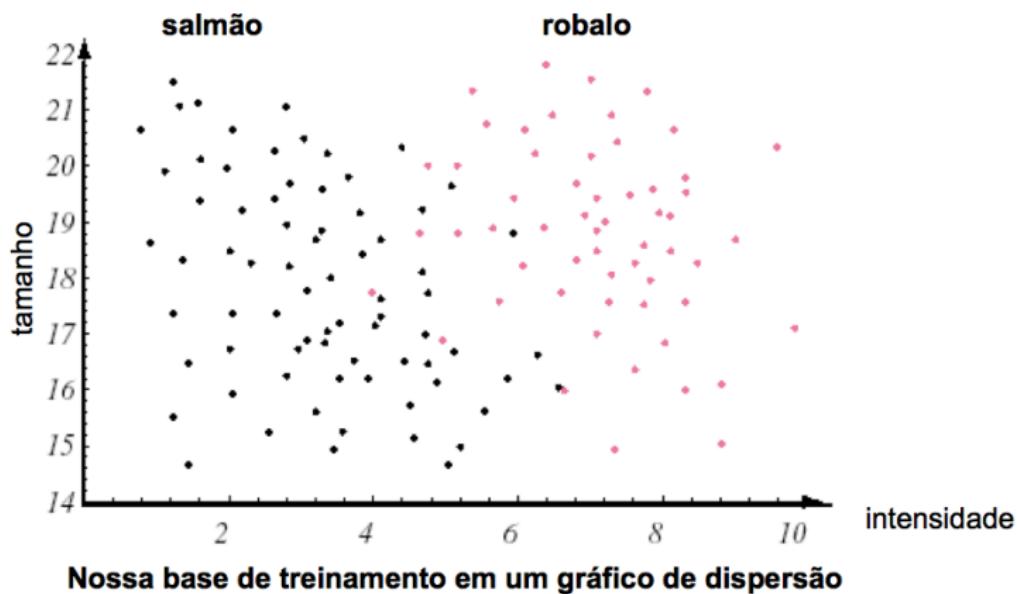
- O que diferencia um salmão de um robalo?
- Extrair atributos, os quais sejam similares aos objetos de uma mesma classe e diferentes para objetos das outras classes.
- Processo conhecido com **Extração de Características**



- Suponha que tamanho e cor (intensidade) sejam características discriminantes.

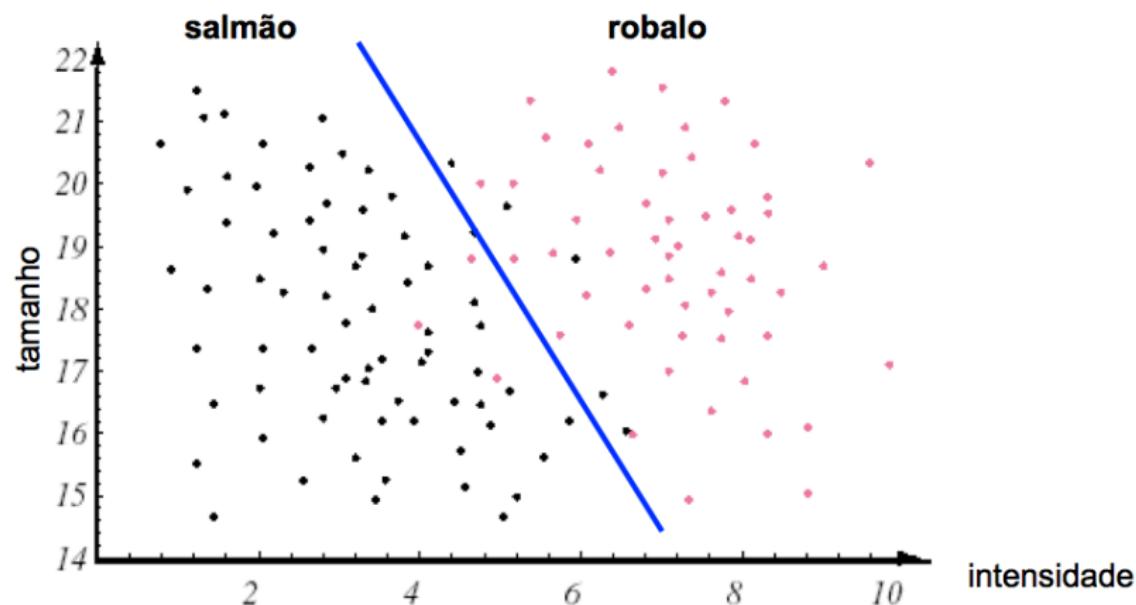
Aprendizagem Supervisionada - Classificação

2. Representação



Aprendizagem Supervisionada - Classificação

3. Fronteira de decisão



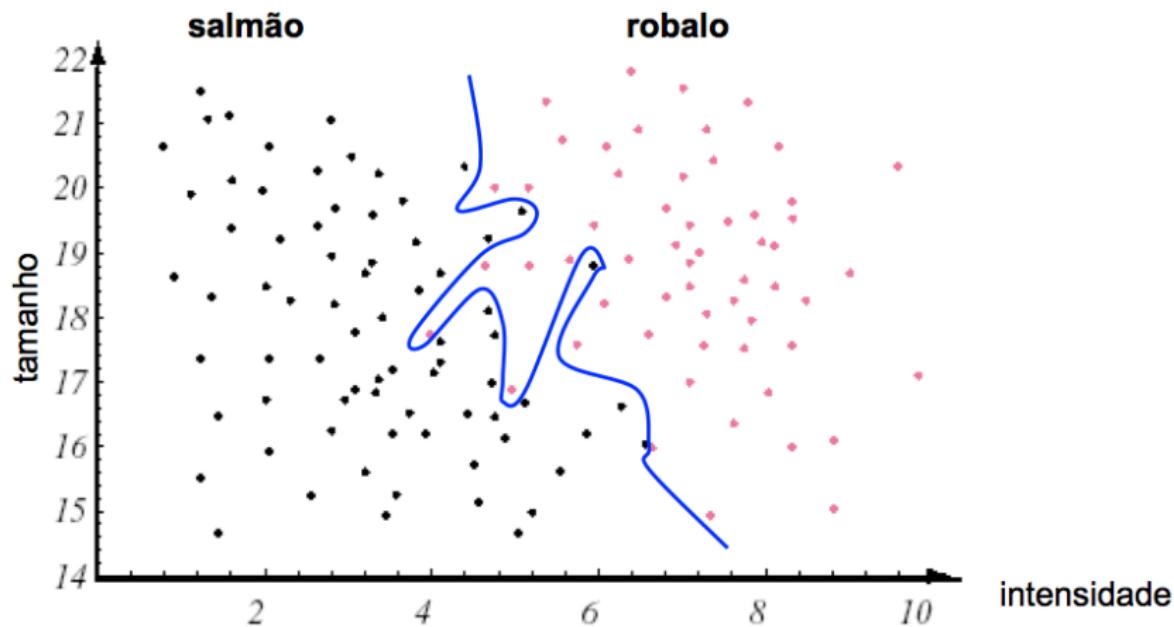
- Separação Linear: $y = ax + b$
- Funções Discriminantes Lineares (Perceptron, SVM).

3. Fronteira de decisão

- Essa regra fornece uma boa fronteira de decisão, entretanto pode ser melhorada.
- Adicionar mais características, como por exemplo, dimensões do peixe, posição relativa dos olhos, etc..., poderiam ajudar.
- Quanto mais características, mais base de treinamento será necessária
- Maldição da dimensionalidade

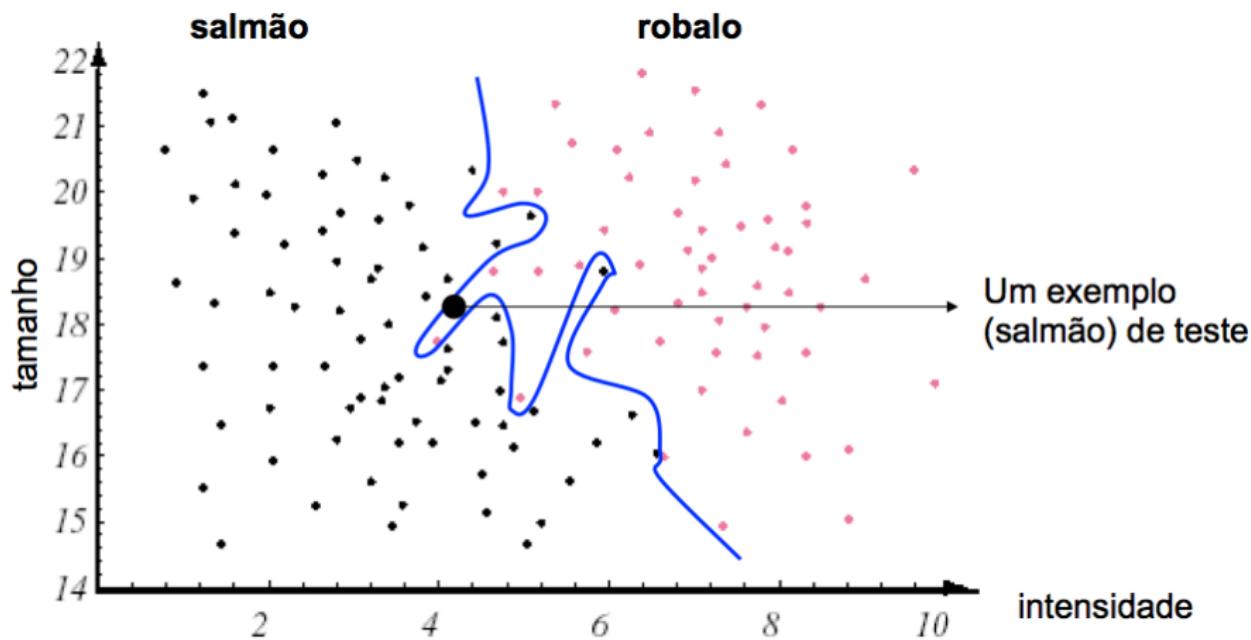
Aprendizagem Supervisionada - Classificação

3. Fronteira de decisão



Aprendizagem Supervisionada - Classificação

3. Fronteira de decisão



3. Fronteira de decisão

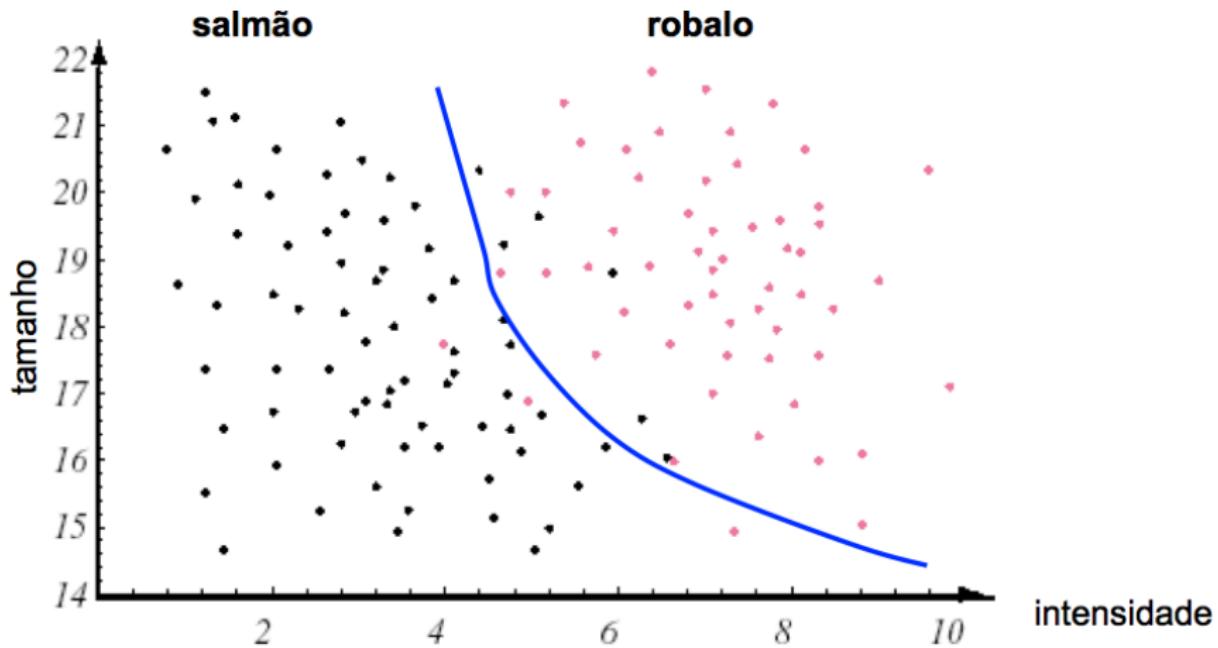
- Apesar de complexo, esse modelo é pior do que o modelo linear.
- A principal característica de um modelo deve ser a sua capacidade de generalizar.
- Modelos muito complexos geralmente não generalizam bem, pois decoram a base de treinamento (over-fitting).

3. Fronteira de decisão

- Por outro lado, modelos muito simples tendem a generalizar demais (under-fitting).
- Na aprendizagem de máquina, veremos técnicas para evitar over- e under-fitting, ou seja, construir fronteiras de decisão nem tão simples e nem tão complexas

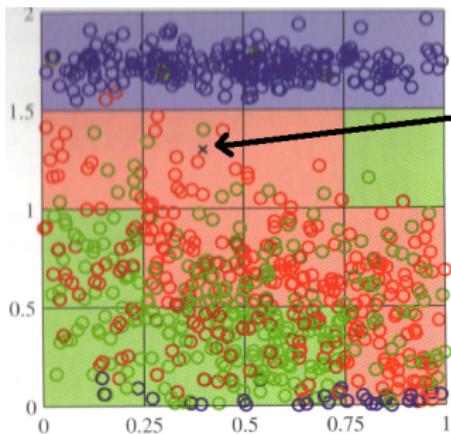
Aprendizagem Supervisionada - Classificação

3. Fronteira de decisão



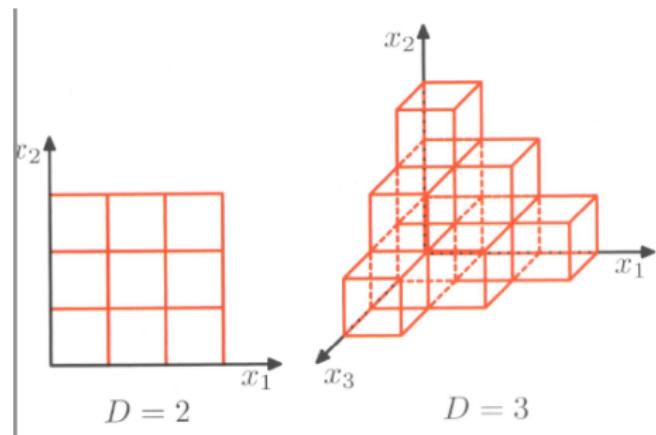
Maldição da Dimensionalidade

- Em geral, na busca por melhores resultados na aprendizagem, criam-se vetores de características com mais informações.
- Para ilustrar o problema, vamos considerar o exemplo da Figura abaixo, com duas características x_6 e x_7 e três classes.
- dim1



Maldição da Dimensionalidade

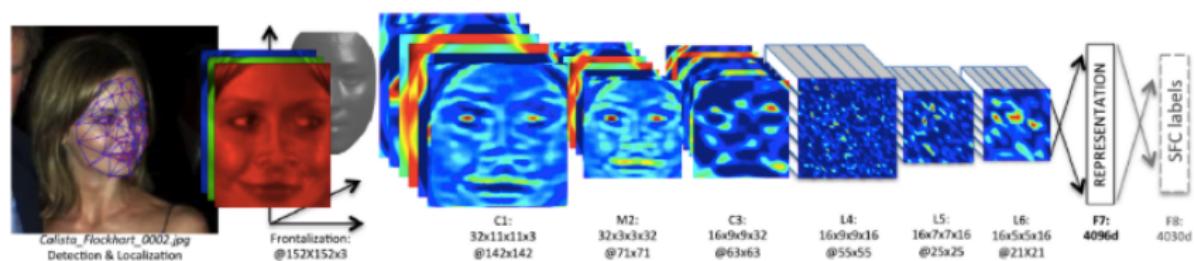
- Conforme adicionamos características ao vetor, o número de células cresce de forma exponencial.
- Em função disso, quanto mais características temos, mais base de dados temos que ter para preencher todas as células.
- Na maioria dos problemas, entretanto, a quantidade de dados disponível para a aprendizagem é limitada.



Aprendizagem Supervisionada - Classificação

Deep Learning (Representation Learning)

- Extrair a representação a partir dos dados.
- Qualquer algoritmos de aprendizagem de máquina depende da representação usada.
- Definição de características não é uma tarefa trivial.
- Porque não deixa-la então por conta do algoritmo de aprendizagem?

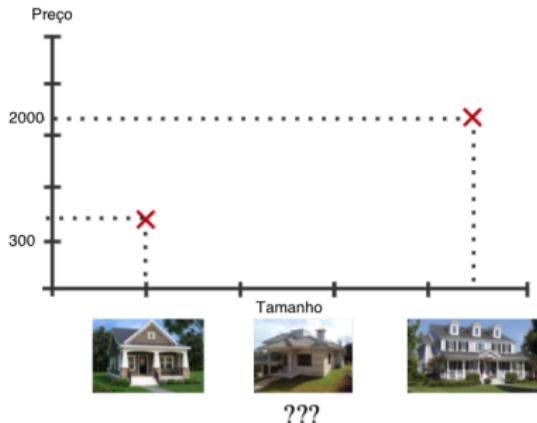


Aprendizagem Supervisionada - Regressão



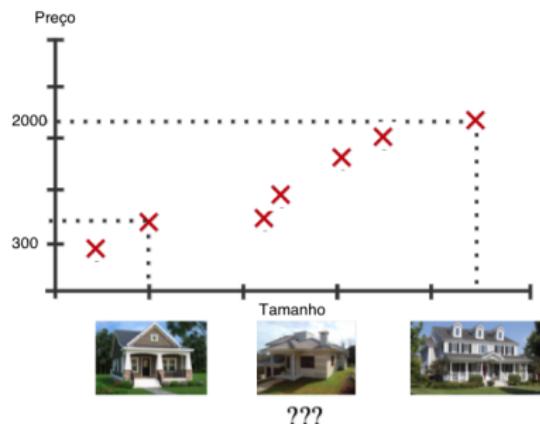
- Construir um sistema para estimar o preço de um imóvel.

Aprendizagem Supervisionada - Regressão



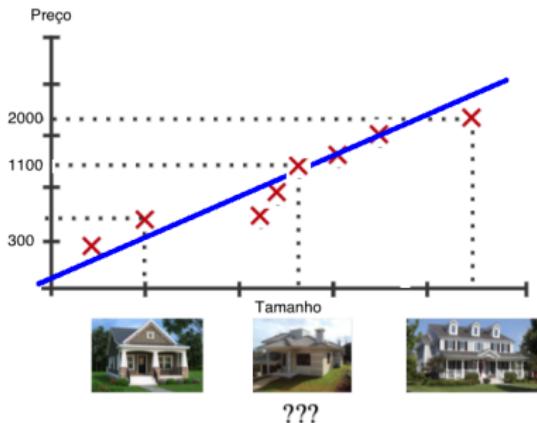
- Com base em informações de preço e tamanho de diversas casas, a tarefa consiste em estimar o preço de uma dada casa.
- Diferentemente da classificação, onde a saída do classificador é categórica, na regressão a saída é um valor real.

Aprendizagem Supervisionada - Regressão



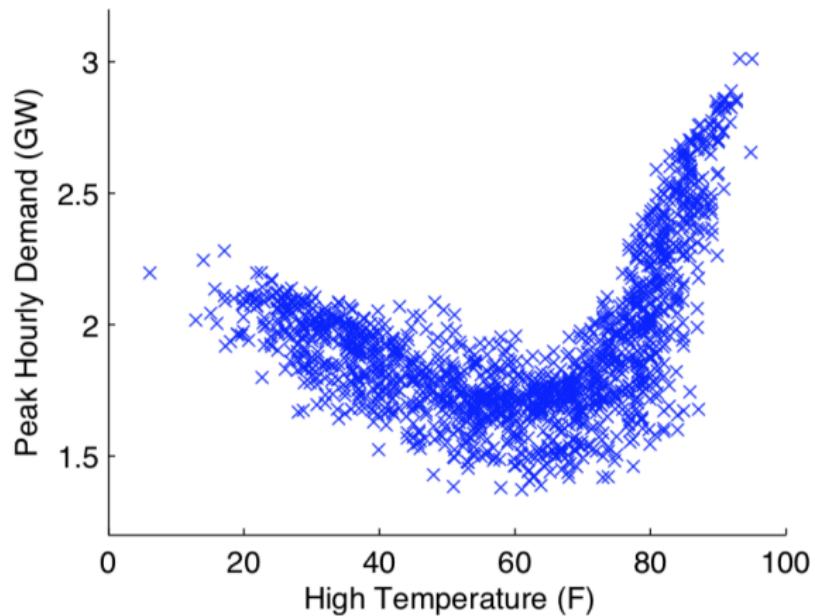
- Qual seria o preço estimado para a casa em questão?

Aprendizagem Supervisionada - Regressão



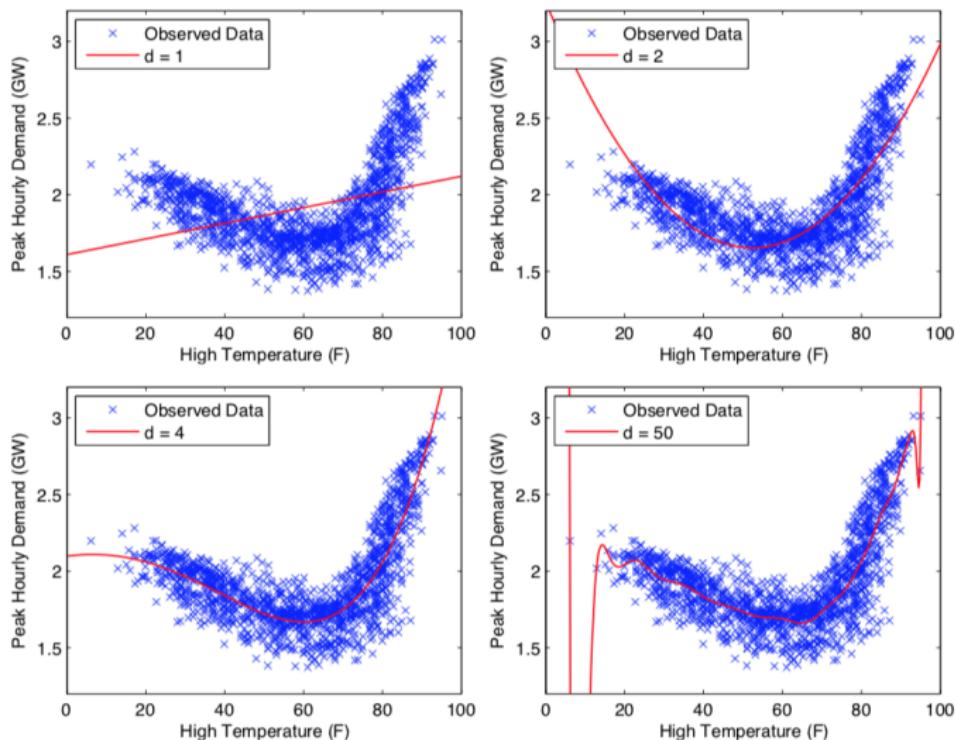
- Dado um conjunto de dados rotulados (tamanho da casa, preço) podemos encontrar um modelo para estimar o preço de outras casas.
- O modelo mais simples é a regressão Linear ($y = ax + b$)
- Cuidado para não extrapolar sua base de conhecimento.

Aprendizagem Supervisionada - Regressão



- Em alguns casos um modelo linear não é a melhor solução.

Aprendizagem Supervisionada - Regressão



- Under-fitting vs Over-fitting

Aprendizagem não supervisionada

- Em problemas de aprendizagem supervisionada, contamos como a tupla $[X, y]$, em que o objetivo é classificar y usando o vetor de características X .
- Na aprendizagem não supervisionada, temos somente o vetor X .
- Nesse caso, o objetivo é descobrir alguma coisa a respeito dos dados
 - ▶ Por exemplo, como eles estão agrupados.
- Mais subjetiva que a aprendizagem supervisionada uma vez que não existe um objetivo simples como a classificação.

Dados não rotulados

Obtenção de dados não rotulados não é custosa!

Aprendizagem não supervisionada

Exemplos de aplicações:

- Segmentação de mercado (Marketing)
- Agrupamento de pacientes|clientes|documentos, etc...
- Sistemas de Recomendação

Frequently Bought Together

Total price: \$83.09

Add both to Cart

Add both to List

This item: Structure and Interpretation of Computer Programs - 2nd Edition (MIT Electrical Engineering and... by Harold Abelson Paperback \$50.50

The Pragmatic Programmer: From Journeyman to Master by Andrew Hunt Paperback \$32.59

Customers Who Bought This Item Also Bought

The Little Schemer - 4th Edition
Daniel P. Friedman
★★★★★ 64
Paperback \$36.00 ✓Prime

Instructor's Manual Va Structure and Interpretation of Computer Programs...
Gerald Jay Sussman
★★★★★ 5
Paperback \$28.70 ✓Prime

The Pragmatic Programmer
Andrew Hunt David Thomas
★★★★★ 328
Paperback \$32.59 ✓Prime

Introduction to Algorithms, 3rd Edition (MIT Press)
Thomas H. Cormen
★★★★★ 313
Paperback \$68.32 ✓Prime

An Introduction to Functional Programming Through Lambda Calculus
Greg Michaelson
★★★★★ 23
Paperback \$20.70 ✓Prime

Purely Functional Data Structures
Chris Okasaki
★★★★★ 19
Paperback \$40.74 ✓Prime

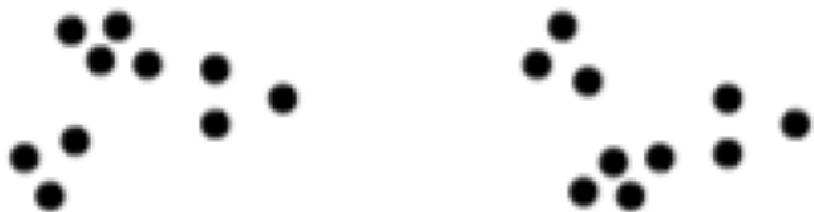
Code: The Hidden Language of Computer Hardware and Software
Chris Okasaki
★★★★★ 334
Paperback \$17.99 ✓Prime

The Little Prover (MIT Press)
Daniel P. Friedman
★★★★★ 4
Paperback \$31.78 ✓Prime



Aprendizagem não supervisionada

Quantos grupos (clusters) existem nesses dados?



Aprendizagem não supervisionada

Dois



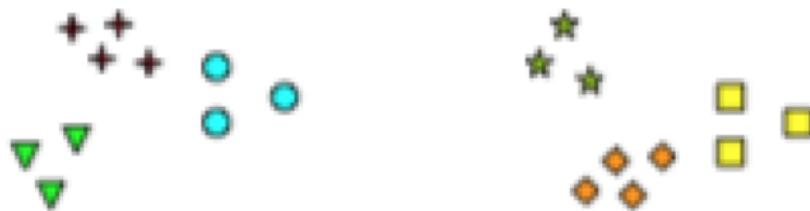
Aprendizagem não supervisionada

Quatro



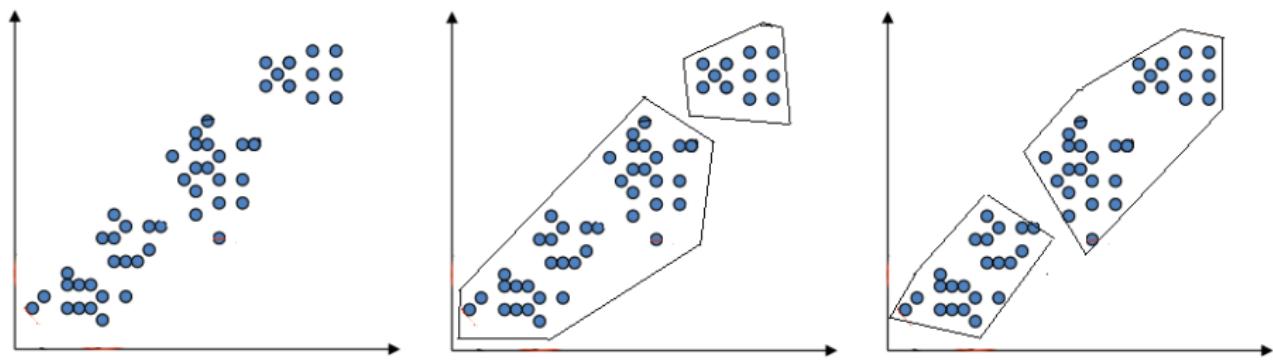
Aprendizagem não supervisionada

Seis



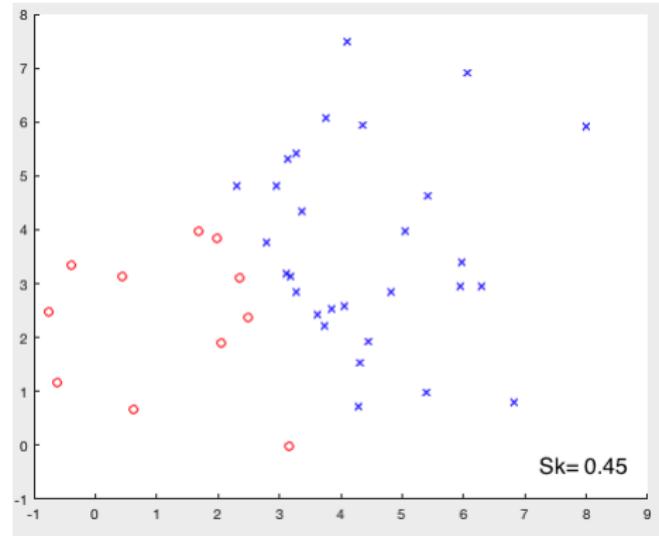
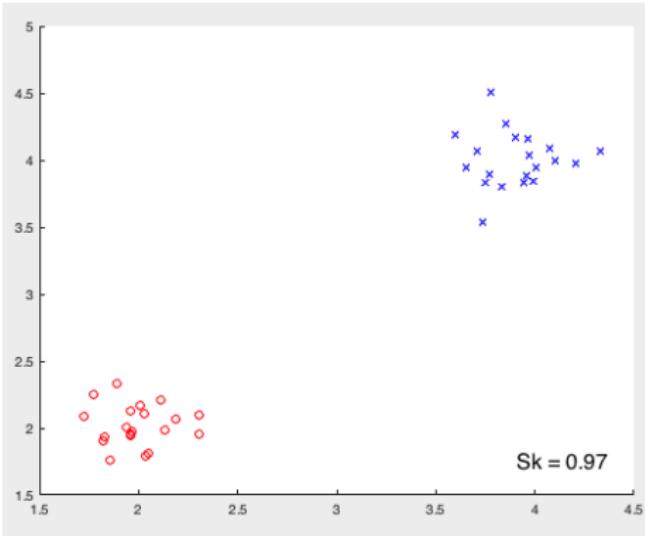
Aprendizagem não supervisionada

O que é um bom clustering?



- Compactos e longe um dos outros.
- Índices de clusters

Avaliação em Aprendizagem não Supervisionada



Aprendizagem não Supervisionada

LLM (Large Language Models)

- Mecanismos de atenção (Transformers)
- GPT, Llama, etc..
- O modelo é treinado em grandes quantidades de texto sem rótulos específicos. O objetivo do treinamento é aprender a prever a próxima palavra ou sequência de palavras em uma frase, dada a sequência anterior.

Redução de Dimensionalidade

- Independentemente do algoritmo de aprendizagem de máquina escolhido, um dos principais aspectos na construção de um bom classificador é a utilização de características discriminantes.
- A adição de uma nova característica não significa necessariamente um bom classificador
- Depois de um certo ponto, adicionar novas características pode piorar o desempenho do classificador.
- Aspectos diretamente relacionados com a escolha das características:
 - ▶ Desempenho
 - ▶ Tamanho da base de dados.

Identificar e selecionar um subconjunto de características relevantes para um determinado problema, a partir de um conjunto inicial.

Redução de Dimensionalidade

- Não é um problema trivial
- Em problemas reais, características discriminantes não são conhecidas a priori.
- Características raramente são totalmente independentes.
- Duas características irrelevantes, quando unidas pode formar uma nova característica relevante e com bom poder de discriminação
- Problema de otimização, em que o espaço de busca é 2^N (N é o número de características).

Redução de Dimensionalidade

Wrapper vs Filter

- Se o algoritmo de aprendizagem está envolvido no processo de seleção de atributos, o método é chamado de Wrapper
- Caso contrário, Filter.
- Os métodos wrapper geralmente fornecem melhores resultados, entretanto com um maior custo computacional.
 - ▶ O classificador é treinado para cada subconjunto de características que deve ser avaliado
 - ▶ Inviável para 2^N

eXplainable AI

- Com o aumento da adoção de soluções de ML e com a falta de transparência e explicações dos modelos a área de eXplainable Artificial Intelligence (XAI) passou a crescer de forma acelerada nos últimos anos.
- Este campo de pesquisa tem por objetivo tornar os resultados dos sistemas de IA mais comprehensíveis para humanos.
- Apresentando mecanismos internos mais transparentes e provendo explicações para suas decisões em alguns níveis de detalhes.
- Abordagens
 - ▶ Feature Importance (Árvores de Decisão)
 - ▶ Permutation Feature Importance (remover uma característica e verificar o impacto)
 - ▶ LIME (Local Interpretable Model-Agnostic Explanation)
 - ▶ SHAP (SHapley Additive exPlanations)

