

Inteligência Computacional em Hardware

Aprendizado de Máquina

Noções Iniciais

Tiago Oliveira Weber
2024

- ▶ Contextualização
- ▶ Conceitos
- ▶ Classificações
- ▶ Aprendizado Supervisionado
- ▶ Árvores de Decisão

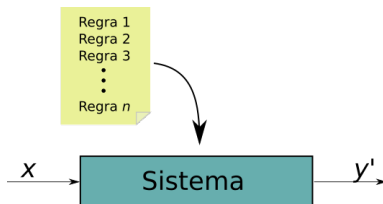
- ▶ **Aprender**: processo de obter novo *conhecimento*.
- ▶ respostas melhores para um mesmo conjunto de entradas em um próximo momento / adaptação

- ▶ Aprendizado através de exemplos
- ▶ Aprendizado por informações fornecidas (ser "contado")
- ▶ Aprendizado através de tentativas ("fazendo")

- ▶ conjunto treinamento $(\mathbf{x}_i, \mathbf{y}_i)$
- ▶ é usado para obter a função $f:D \rightarrow S$
 - ▶ sendo \mathbf{x}_i um vetor do espaço de domínio D
 - ▶ e \mathbf{y}_i um vetor do espaço de solução S .
 - ▶ $i = 1, 2, \dots, n$ é o índice dos dados de treinamento
- ▶ Exemplo: redes neurais

amostra	x_1	x_2	\dots	x_k	y (classe)
1	10	4	\dots	4	0
2	4	3	\dots	1	1
3	9	3	\dots	3	0
\vdots	\vdots	\vdots	\dots	\vdots	\vdots
n	10	4	\dots	0	1

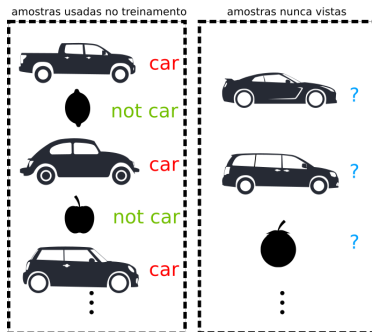
- ▶ Criação de regras heurísticas dados por um **instrutor**:
 - ▶ pode passar informação diretamente na forma de regras/heurísticas ou
 - ▶ pode passar na forma de linguagem natural, que posteriormente são convertidas em regras
- ▶ Exemplo:
 - ▶ sistemas de IA simbólicas
 - ▶ sistemas Fuzzy



- ▶ Sistema começa a operar com pouca informação e vai acumulando experiências com base nos resultados.
- ▶ Exemplo:
 - ▶ uso de algoritmos genéticos, metaheurísticas de otimização em geral,...

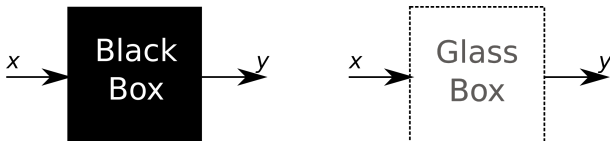
► Generalização

- é o processo de prever soluções (y) para valores de entrada novos (não usados no treinamento)
- boa generalização é característica essencial para técnicas de aprendizado de máquina
- Exemplos:
 - reconhecer objetos em novas imagens, classificar texturas de novas amostras de tecido, classificar novas amostras de áudio...



► Explicação

- possibilidade de rastrear (de forma compreensível) o processo de alcançar a resposta a partir das entradas
- altamente desejável para sistemas que envolvam segurança e responsabilidade (*accountability*)
- fácil para sistemas simbólicos
- difícil para sistemas conexionistas
- *blackbox vs glassbox*
- *explainable AI - XAI*



▶ Validação

- ▶ testar quão boa as saídas do sistema são através de comparação com especialistas ou outros sistemas.
- ▶ aqui, descrito de forma genérica. posteriormente, em mais detalhe.

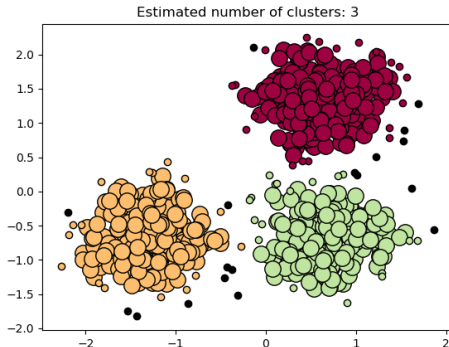
▶ Adaptação

- ▶ capacidade de mudar o sistema durante operação em um ambiente dinâmico.

- ▶ **Aprendizado Indutivo:** criar função geral a partir de conjunto de pares entrada-saída.
- ▶ **Aprendizado Analítico ou Dedutivo:** partir de uma regra geral conhecida para uma nova regra que é consequência lógica das anteriores.

- ▶ Aprendizado não-supervisionado
- ▶ Aprendizado por reforço
- ▶ Aprendizado supervisionado
- ▶ Aprendizado semi-supervisionado

- ▶ **Aprendizado não-supervisionado:**
 - ▶ Padrões são aprendidos mesmo que não seja fornecido feedback explícito.
 - ▶ Ex.: clustering

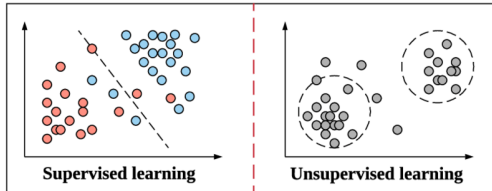


- ▶ **Aprendizado por Reforço** (*Reinforcement Learning*):
 - ▶ Padrões são aprendidos através de uma série de reforços (recompensas ou punições)
 - ▶ Ex.: número de colisões durante uma corrida simulada de carros



- ▶ **Aprendizado Supervisionado:**
 - ▶ Padrões são aprendidos através de pares entrada-saída.
 - ▶ a função escondida que mapeia as entradas para as saídas é buscada
 - ▶ Ex.: aprender a detectar imagens que representem células cancerígenas a partir de uma série de imagens médicas rotuladas por médicos.

Aprendizado Supervisionado e Aprendizado Não Supervisionado

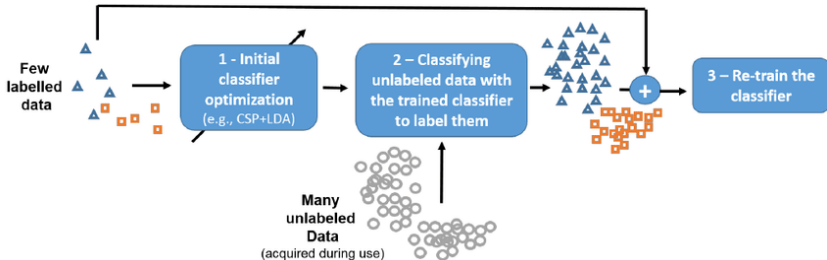


Fonte: extraído de QIAN et al. "Orchestrating Development Lifecycle of Machine Learning Based IoT Applications: A Survey", arXiv:1910.05433 [cs.DC], 2019

► **Aprendizado Semi-supervisionado:**

- apenas uma pequena parcela dos exemplos fornecidos são rotulados
- o restante dos exemplos fornecido não tem rótulo
- podem haver erros nos rótulos fornecidos
- intuitivamente, pode ser visto como exemplos dado pelo professor para preparar a turma para o exame

Princípio de aprendizado semi-supervisionado.



Fonte: extraído de Fabien LOTTE. "Signal Processing Approaches to Minimize or Suppress Calibration Time in Oscillatory Activity-Based Brain-Computer Interfaces" Proceedings of the IEEE, Institute of Electrical and Electronics Engineers (IEEE), 2015, 103 (6), pp.871-890.

- ▶ Dado um **conjunto de treinamento** com N exemplos de pares entrada-saída:

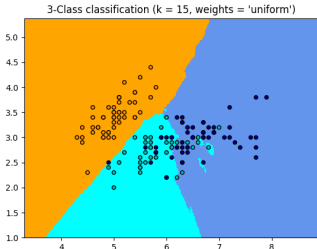
$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), (\mathbf{x}_3, y_3), \dots, (\mathbf{x}_N, y_N),$$

onde cada y_j é gerado por uma função desconhecida $y = f(\mathbf{x})$

- ▶ objetivo é descobrir a função h (hipótese) que aproxima a função f .
- ▶ a tarefa do aprendizado é buscar no **espaço das possíveis hipóteses H** por uma hipótese h que se desempenhe bem, mesmo para exemplos não fornecidos no conjunto treinamento

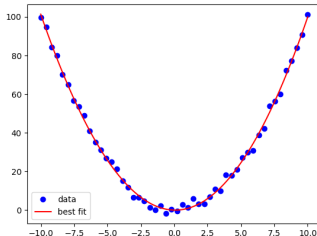
- ▶ Classificação:
 - ▶ quando a saída y é um conjunto de finito de valores
 - ▶ ex.: ensolarado/nublado/chuvoso, comprar/vender ação

Classificador com 3 classes



- ▶ Regressão:
 - ▶ quando a saída y é um valor real
 - ▶ a missão é encontrar a função escondida que mapeia entradas até a saída
 - ▶ ex.: temperatura, valor de uma ação

Regressão de uma função



- ▶ **espaço das possíveis hipóteses H** : todo o conjunto de possíveis hipóteses que podem ser construídas a partir do modelo utilizado.
 - ▶ Ex.: caso estejamos utilizando como modelo um polinômio de ordem até 5, o espaço H são todos os possíveis polinômios que podem ser construídos a partir das variações dos coeficientes disponíveis.

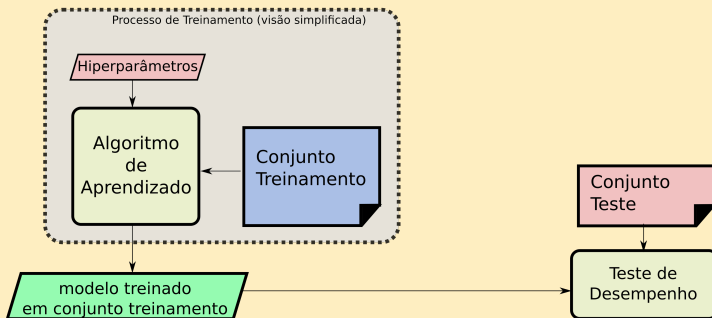


- ▶ **hipótese consistente**: hipóteses que concordam com os dados
- ▶ entre hipóteses consistentes, qual escolher?
 - ▶ navalha de Occam (ou Ockham): entre duas explicações, a que precisa menos suposições geralmente é a correta

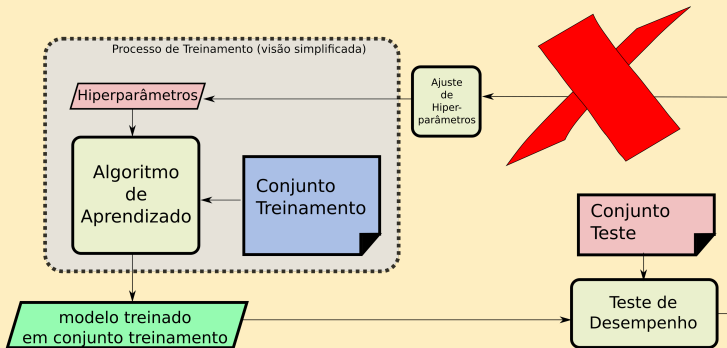
- ▶ Curve Fitting - Exemplo

- ▶ **Taxa de erro** de uma hipótese: proporção de erros que comete
- ▶ erro: $h(\mathbf{x}) \neq y$ para um exemplo (\mathbf{x}, y)
- ▶ baixa taxa de erros no **conjunto treinamento** não significa que hipótese terá boa generalização
- ▶ tipos de conjunto:
 - ▶ **conjunto treinamento**
 - ▶ **conjunto validação**
 - ▶ **conjunto teste**

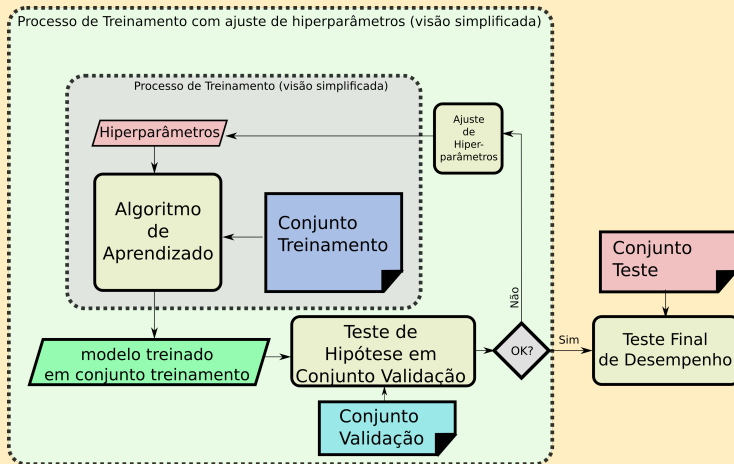
Processo sem otimização de hiperparâmetros



Processo com uso incorreto do conjunto teste



Processo com uso de conjunto validação para escolha de hiperparâmetros



Uso do dataset

total de amostras disponíveis

conjunto treinamento

conjunto teste

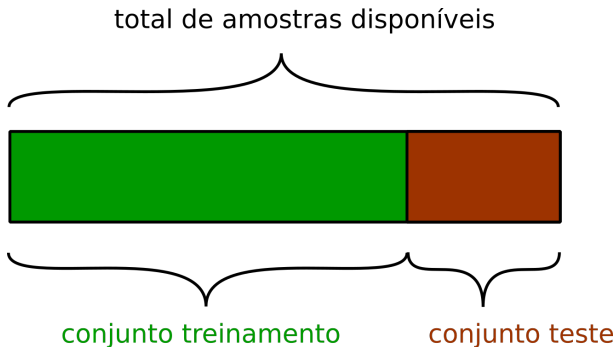
conjunto treinamento

conjunto validação

conjunto teste

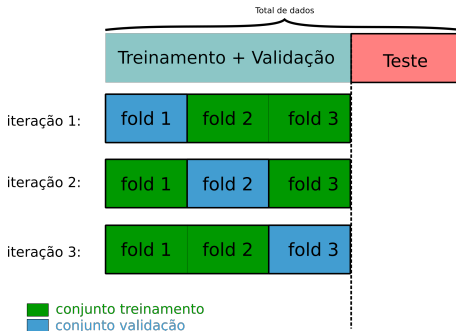
Holdout Cross-Validation

- ▶ divide-se aleatoriamente os dados disponíveis em conjunto treinamento e conjunto teste
 - ▶ ex.: 80% para treinamento, 20 % para teste
- ▶ vantagem: simplicidade
- ▶ desvantagem: não usa bem os dados



k-fold Cross-Validation

- ▶ divide-se os dados em k subconjuntos de mesmo tamanho
- ▶ rodamos k treinamentos:
 - ▶ em cada um:
 - ▶ uma das partes é usada para validação;
 - ▶ as demais são usadas para treinamento;
 - ▶ é feita a média do desempenho nos conjuntos teste das k rodadas
- ▶ vantagem: cada parte é usada tanto para treinamento quanto para validação (em momentos distintos)



Tópicos estudados

- ▶ Contextualização
- ▶ Conceitos
- ▶ Classificações
- ▶ Conjuntos Treinamento/Validação/Teste

- ▶ Stuart J. Russel, Peter Norvig. **Artificial Intelligence: A Modern Approach**. (3rd ed.). Prentice Hall Press, USA, 2009.



- ▶ Kubat, Miroslav. **An Introduction to Machine Learning**. Springer International Publishing, second edition, 2017.



- ▶ Géron, Aurélien. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. O'Reilly Media, 2019.

