



Universidade Federal do
Recôncavo da Bahia



Universidade Federal
de São João del-Rei



Introdução ao Aprendizado por Reforço

Prof. Me. André Luiz Carvalho Ottoni

Aprendizado por Reforço: Engenharia e Estatística impulsionando a sociedade
Projeto financiado pelo Edital 001/2019/UFSJ/Reitoria

Centro de Ciências Exatas e Tecnológicas
UFRB - Universidade Federal do Reconcávo da Bahia

São João del-Rei (MG), outubro de 2019

Plano de Aula

1 Introdução

2 Agentes com Aprendizagem

3 Aplicações

Aplicações do Aprendizado de Máquina:

- Aproximador Universal de Funções
 - Objetivo: mapear o relacionamento entre variáveis.
 - Aplicação: processos cuja a modelagem é complexa.
- Controle de Processos:
 - Objetivo: identificar ações de controle que permitam alcançar os requisitos de qualidade.
 - Aplicações: Automação, Robótica, Aeronaves.

Aplicações do Aprendizado de Máquina:

- Reconhecimento/classificação de padrões:
 - Objetivo: associar um padrão de entrada para uma das classes previamente definidas.
 - Aplicações: Reconhecimento de imagens, voz.
- Agrupamento de dados (clusterização):
 - Objetivo: detecção de similaridades entre os diversos padrões de entrada.
 - Aplicações: identificação de classes de dados.

Aplicações do Aprendizado de Máquina:

- Sistemas de Previsão:
 - Objetivo: estimar valores futuros. Leva-se em consideração diversas medidas prévias.
 - Aplicações: previsão de mercados financeiros, previsões climáticas.
- Otimização de Sistemas:
 - Objetivo: minimizar/Maximizar uma função custo (objetivo).
 - Aplicações: otimização combinatória.

Relevância do (*Machine Learning*) no Mundo

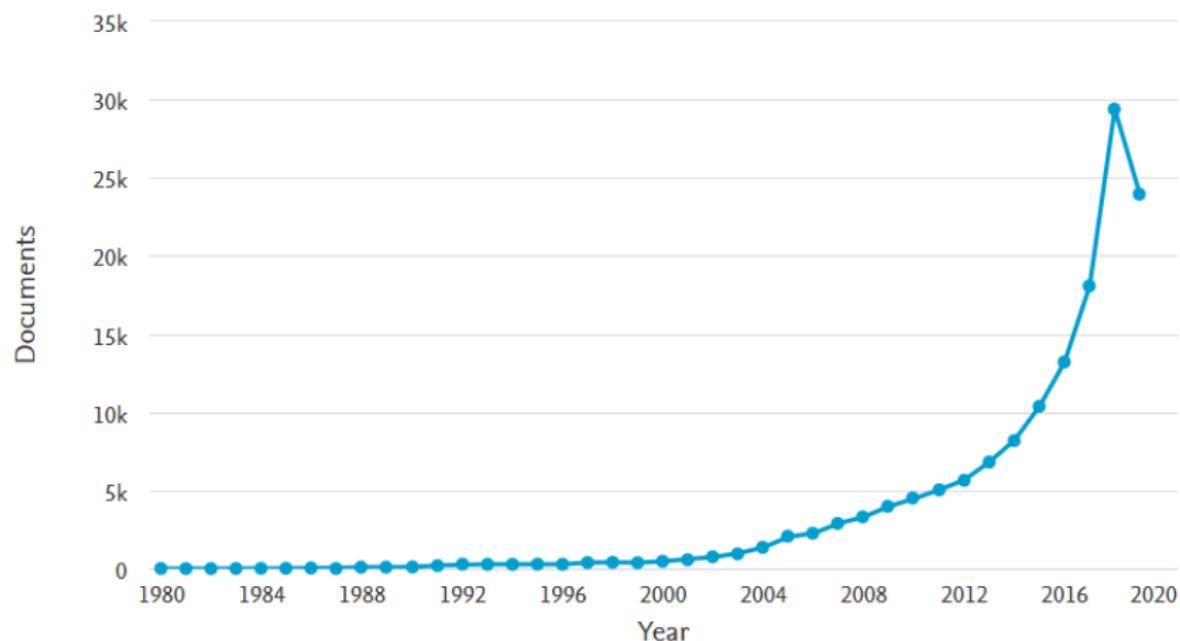


Figura 1: *Machine Learning* no Mundo. Fonte: Scopus (Agosto - 2019).

Relevância do (*Machine Learning*) no Brasil

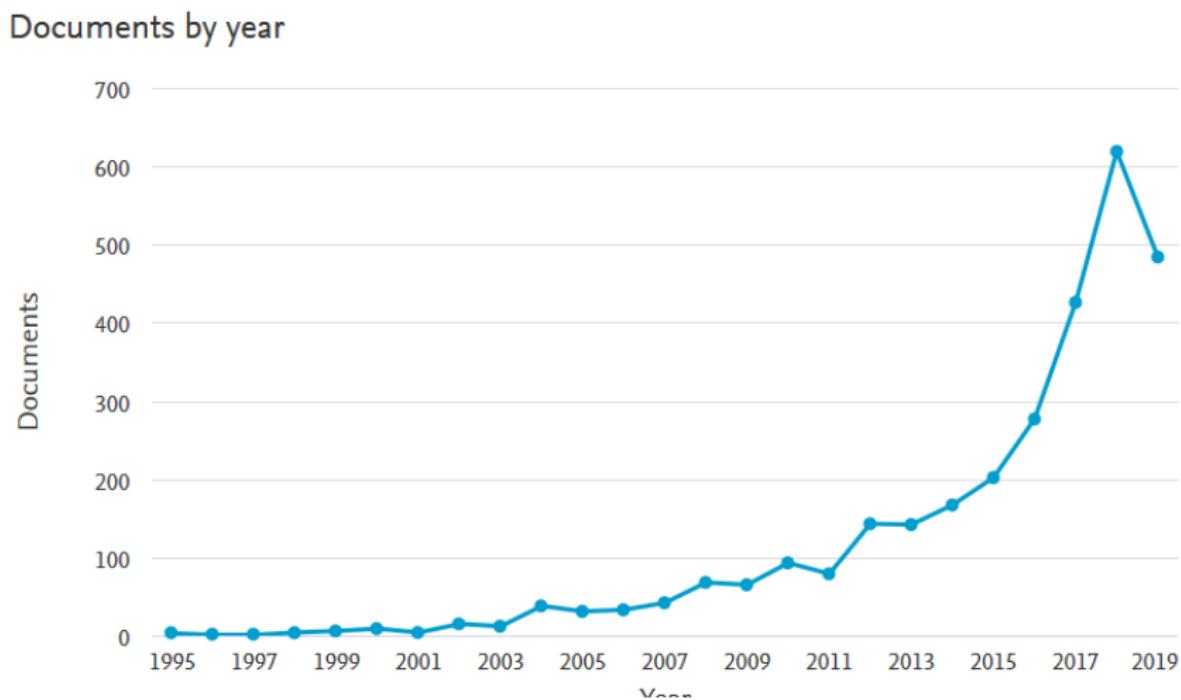


Figura 2: *Machine Learning* no Brasil. Fonte: Scopus (Agosto - 2019).

Relevância do (*Machine Learning*) nos EUA

Documents by year

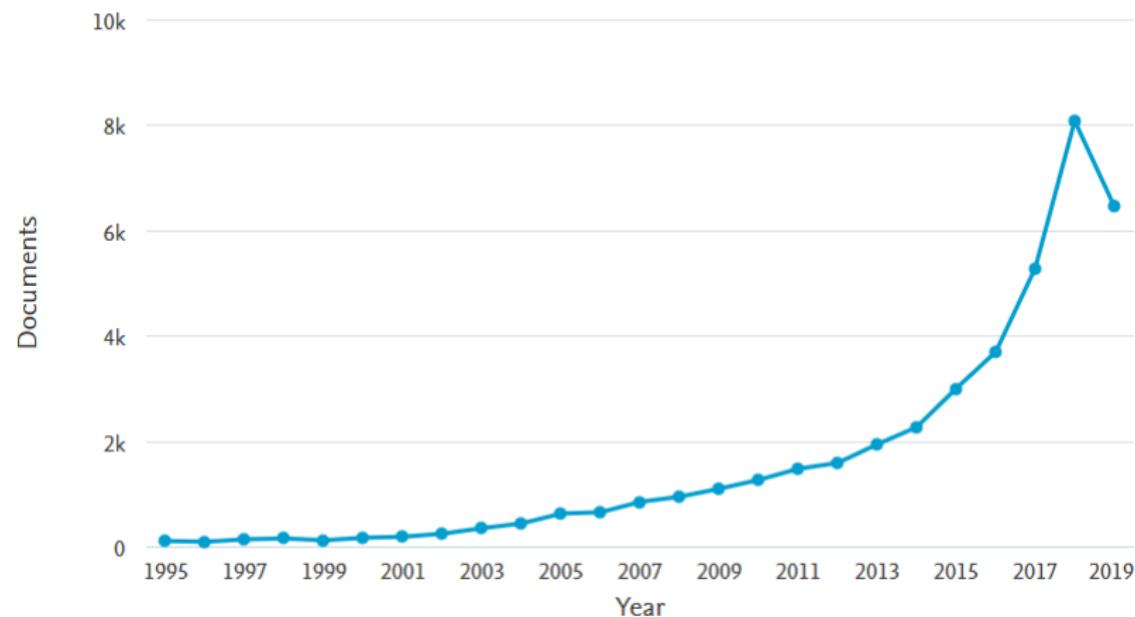


Figura 3: *Machine Learning* nos EUA. Fonte: Scopus (Agosto - 2019).

Fatores para o Aprendizado de Máquina:

- Que *componente* deve ser melhorado.
- O *conhecimento prévio* que o agente já tem.
- Que *representação* é usada para os dados e para o componente.
- Que *feedback* está disponível para aprendizagem.

Tipos de Aprendizado de Máquina:

- Aprendizado Supervisionado.
- Aprendizado por Reforço.
- Aprendizado não-supervisionado.

Aprendizado Supervisionado:

- O agente observa alguns exemplos de pares de entrada e saída.
- Aprendizado com “professor”.
- “Professor”: ensina qual a resposta correta para cada entrada.
- Resposta correta: saída desejada.
- Aprendizado de uma função que faz o mapeamento da entrada para a saída.
- Ex. de Métodos:
 - Regressão Linear.
 - Regressão Logística.
 - Redes Neurais Artificiais: Perceptron.

Aprendizado Supervisionado

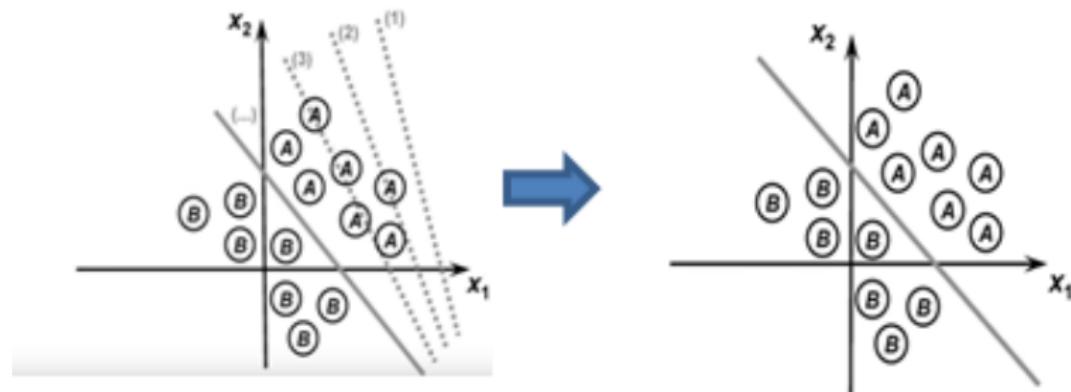


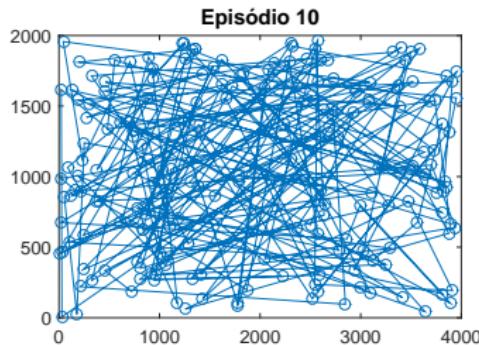
Figura 4: Exemplo de Aprendizado Supervisionado (Perceptron). Fonte: Nunes et al (2016).

Aprendizado por Reforço:

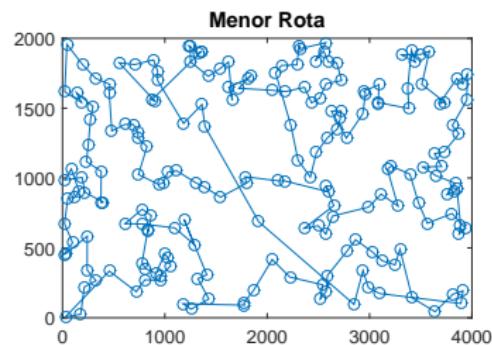
- O agente aprende a partir do sucesso e fracasso no ambiente.
- Recebe sinais de reforços e penalidades.
- Aprendizado por interação no ambiente.
- Objetivo: qual a melhor ação para cada estado do ambiente?
- Ex. de Métodos:
 - Q-learning.
 - SARSA.
 - R-learning.

Aprendizado por Reforço:

- Exemplos de rotas para kroA200, com SARSA , $\alpha = 0,7894$, $\gamma = 0,0894$ e $\varepsilon = 0,01$:



(a) Episódio 10: 331.527



(b) Episódio 2206: 34.795

Figura 5: Representação de rotas encontrados para kroA200.

Aprendizado não-supervisionado:

- O agente aprende padrões de entrada.
- Tarefa comum: agrupamento de dados.
- Não existem saídas desejadas.
- Verifica as particularidades existentes existentes em comum entre os elementos da amostra
- Ex. de Métodos:
 - K-médias.
 - Algoritmo Hierárquico.

Aprendizado Não-Supervisionado

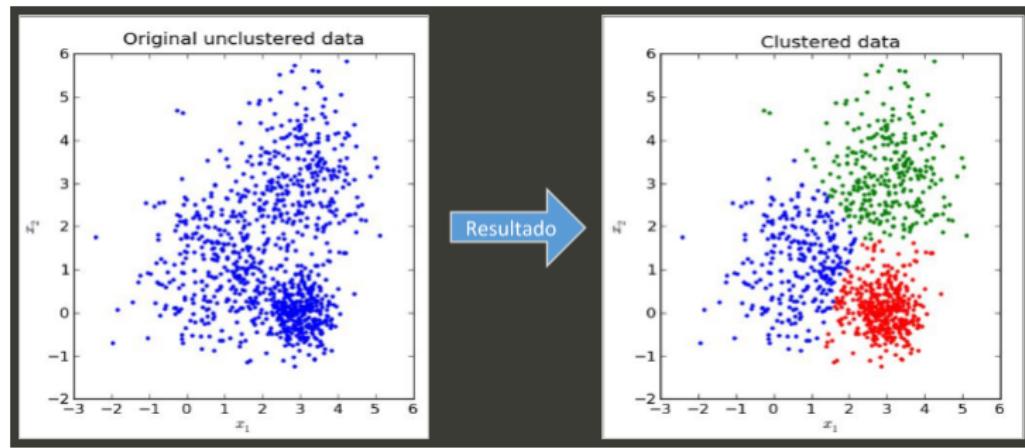


Figura 6: Exemplo de Aprendizado Não-Supervisionado. Fonte:
<http://carlosbaia.com/2016/07/17/introducao-ao-machine-learning/>.

Estrutura de Agentes:

- **Programa do agente:**

- Implementa a função do agente.
- Mapeia percepções em ações.

- **Arquitetura:** possibilita agir no ambiente.

- Dispositivo com sensores e atuadores físicos.

$$agente = programa + arquitetura$$

(1)

Tipos de agentes:

- Agentes Reativos Simples.
- Agentes Reativos Baseados em Modelos.
- Agentes Baseados em Objetivos.
- Agentes Baseados na Utilidade.
- Agentes com Aprendizagem.

- **Agente Reativo Simples**

- Seleciona ações com base em percepções atuais.
- Ignora o restante do histórico de percepções.
- Regra condição-ação.
- Comportamento limitado.
- Ex.: Frear o carro bruscamente diante de uma adversidade.

- **Agente Reativo Baseados em Modelos**

- Forma de lidar com observação parcial.
- Monitora parte do mundo que não pode ver agora.
- Mantém algum tipo de estado interno.
- Depende do histórico de percepções.
- Ex. de modelos.: Mapas do ambiente.

- **Agente Reativo Baseados em Objetivos**

- Conhecer o estado atual nem sempre é suficiente para decidir o que fazer.
- Objetivos: situações desejáveis.
- Ex.: O táxi saber onde quer chegar.
- Ex.: O time de futebol saber que precisa fazer gols.

• Agente Reativo Baseados na Utilidade

- Objetivos pode não ser suficientes para gerar um comportamento de alta qualidade.
- Ex.: Existem várias rotas para um táxi chegar ao seu destino.
- Algumas rotas são mais rápidas, seguras, econômicas.
- Função de Utilidade: internalização da medida de desempenho.
- Um agente que escolhe as ações que maximizem sua utilidade:
 - Racional de acordo com a medida de desempenho externa.

Como os agentes podem aprender?

- **Elemento de aprendizado:**
 - Responsável pela execução de aperfeiçoamentos.
- **Elemento de desempenho:**
 - Responsável pela seleção de ações.
- **Crítico:**
 - Realimenta o elemento de aprendizado.
 - Determina a maneira que o elemento de desempenho deve ser modificado.
- **Gerador de Problemas:**
 - Sugere ações que levarão a novas experiências.
 - Ações exploratórias.

Exemplo: Táxi Automatizado

- **Elemento de desempenho:**

- Coleção de conhecimentos/procedimentos que o táxi tem para selecionar suas ações.

- **Crítico:**

- Observa o mundo e repassa as informações para o elemento de aprendizado.
- Ex.: O táxi faz uma manobra arriscada e os outros motoristas de irritam.

- **Elemento de aprendizado:**

- Verifique se a ação tomada foi ruim.
- Reformula a regra e repassa para o elemento de desempenho

- **Gerador de Problemas:**

- Identifica ações que necessitam de melhorias.
- Ex.: Testar os freios em diferentes superfícies.

Previsão de Mercados de Ações

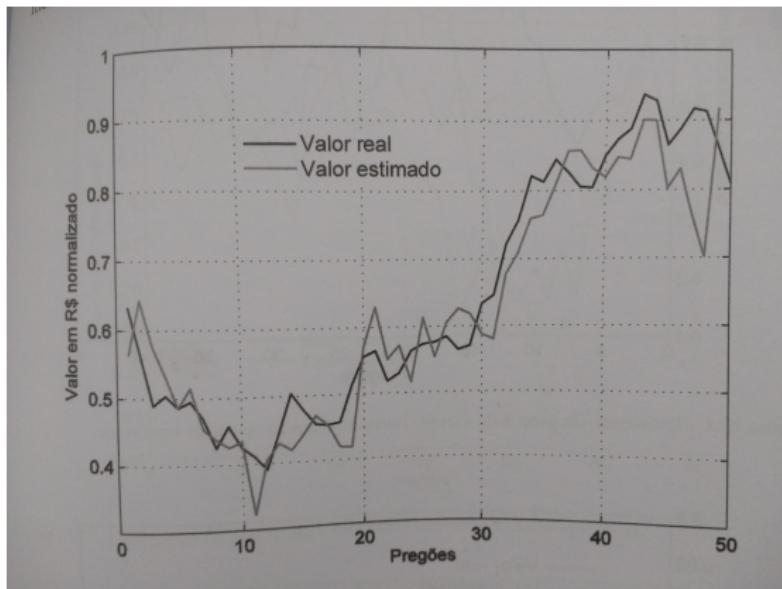


Figura 7: Exemplo de previsão de mercado de ações. Fonte: Nunes et al (2016).

Previsão de Mercados de Ações

- **Entradas:**

- Preço de abertura.
- Preço Mínimo.
- Preço de Fechamento.
- Volume Comercializado.
- Taxa de juros.
- Índice de produção industrial.
- Índice Bovespa.

- **Saída:** Preço de fechamento das ações para dias futuros.

Sistema para Diagnóstico de Doenças

Pacientes	Doenças						
	D ₁	D ₂	D ₃	D ₄	D ₅	(...)	D ₂₁
1	1	0	0	0	0	(...)	1
2	1	0	0	1	0	(...)	0
3	0	0	0	0	0	(...)	0
4	1	0	1	0	0	(...)	0
5	1	0	0	0	0	(...)	0
6	0	0	1	0	0	(...)	0
7	1	0	0	1	0	(...)	0
(...)	1	0	0	1	0	(...)	0
50	1	0	0	1	0	(...)	0

Figura 8: Classe representando presença ou ausência de doenças. Fonte: Nunes et al (2016).

Sistema para Diagnóstico de Doenças

- **Entradas:** Pacientes com tipos de doenças.
 - D_1 : Doenças infeccioas.
 - D_2 : Tumores.
 - D_3 : Doenças no sangue.
 - D_4 : Doenças endócrinas.
 - D_5 : Transtornos mentais e comportamentais.
- **Saída:** Classes de tratamento.

Controle de trajetória de robôs móveis: robô seguidor de fonte de luz

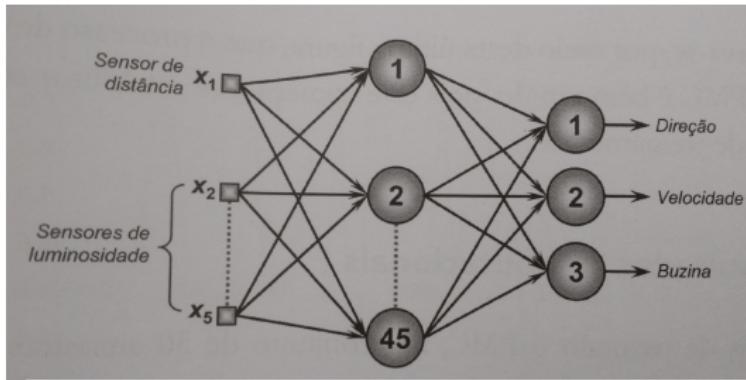


Figura 9: Topologia da rede para controle do robô. Fonte: Nunes et al (2016).

Controle de trajetória de robôs móveis: robô seguidor de fonte de luz

- **Entradas:** Pacientes com tipos de doenças.

- x_1 : Sensor de distância.
- x_2 a x_5 : Sensores de luminosidade.

- **Saída:**

- Direção.
- Velocidade.
- Buzina.

Estimação de desempenho de sistemas elétricos

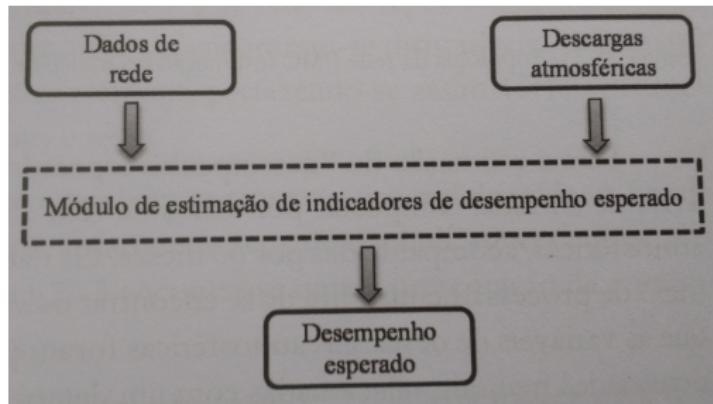


Figura 10: Sistema de estimativa de desempenho. Fonte: Nunes et al (2016).

Estimação de desempenho de sistemas elétricos

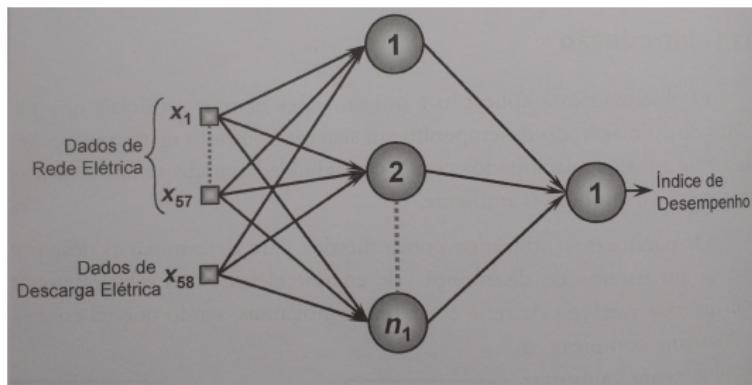


Figura 11: Topologia da rede. Fonte: Nunes et al (2016).