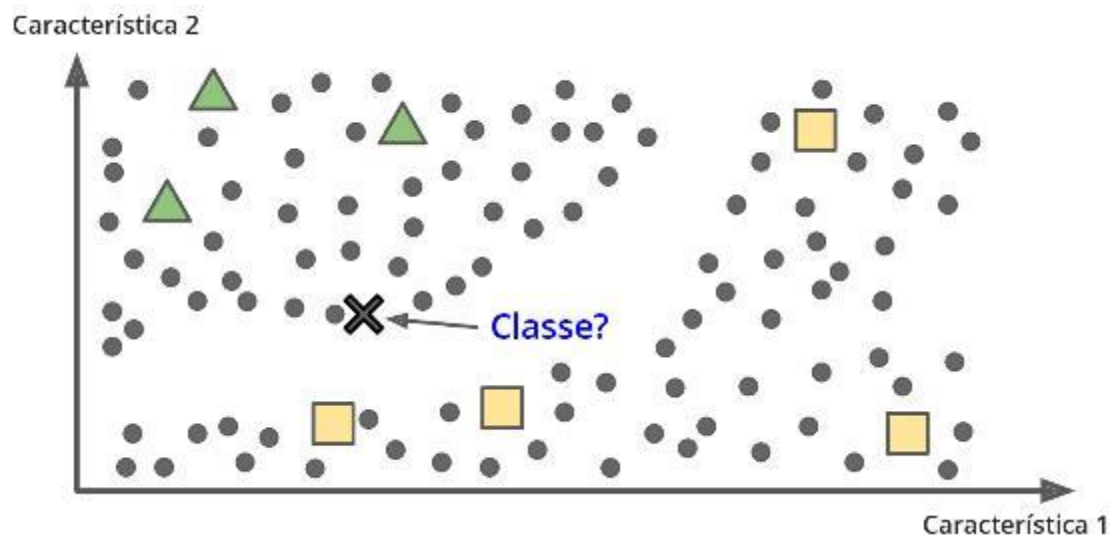


Aprendizado Semi-Supervisionado e Por Reforço

Professor
Jefferson Moraes

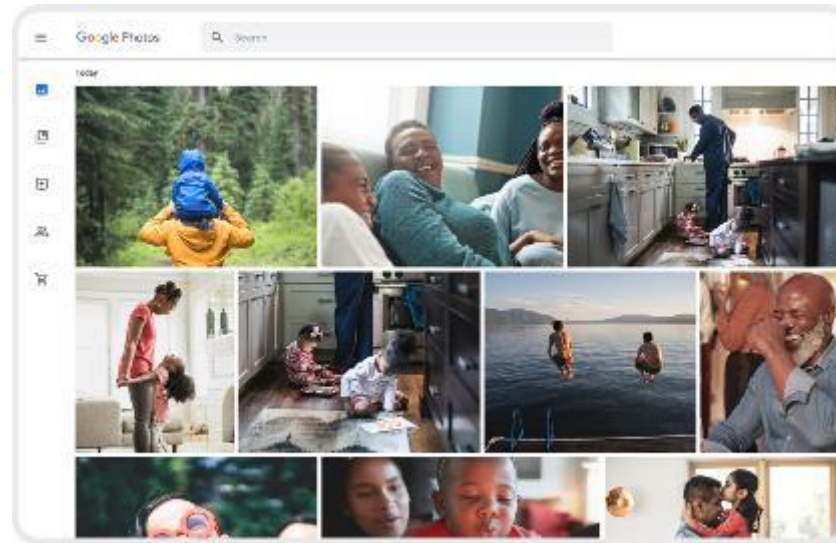
Aprendizado Semi-Supervisionado

- Rotular dados geralmente consome tempo e dinheiro
- Grande quantidade de instâncias não rotuladas e poucas instâncias rotuladas
- Alguns algoritmos conseguem lidar com dados parcialmente rotulados
- Este processo de aprendizado é chamado de semi-supervisionado



Aprendizado Semi-Supervisionado

- Serviços de hospedagem de fotos, Google Fotos
 - Aprendizado não supervisionado (clustering)
 - Ao fazer o upload de todas as fotos da família, o aplicativo reconhecerá automaticamente que a mesma pessoa (A) aparece nas fotos 1, 5 e 11, enquanto outra pessoa (B) aparece nas fotos 2, 5 e 7.
 - Aprendizado supervisionado
 - Informar quem são essas pessoas. Acrescente somente um rótulo por pessoa e o sistema será capaz de nomear todas, o que é útil para pesquisar fotos

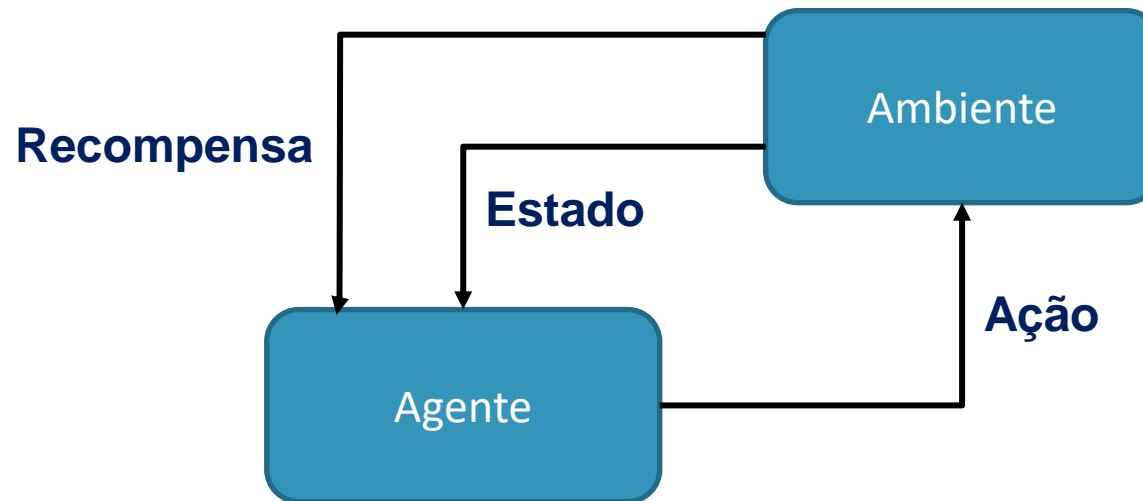


Aprendizado Semi-Supervisionado

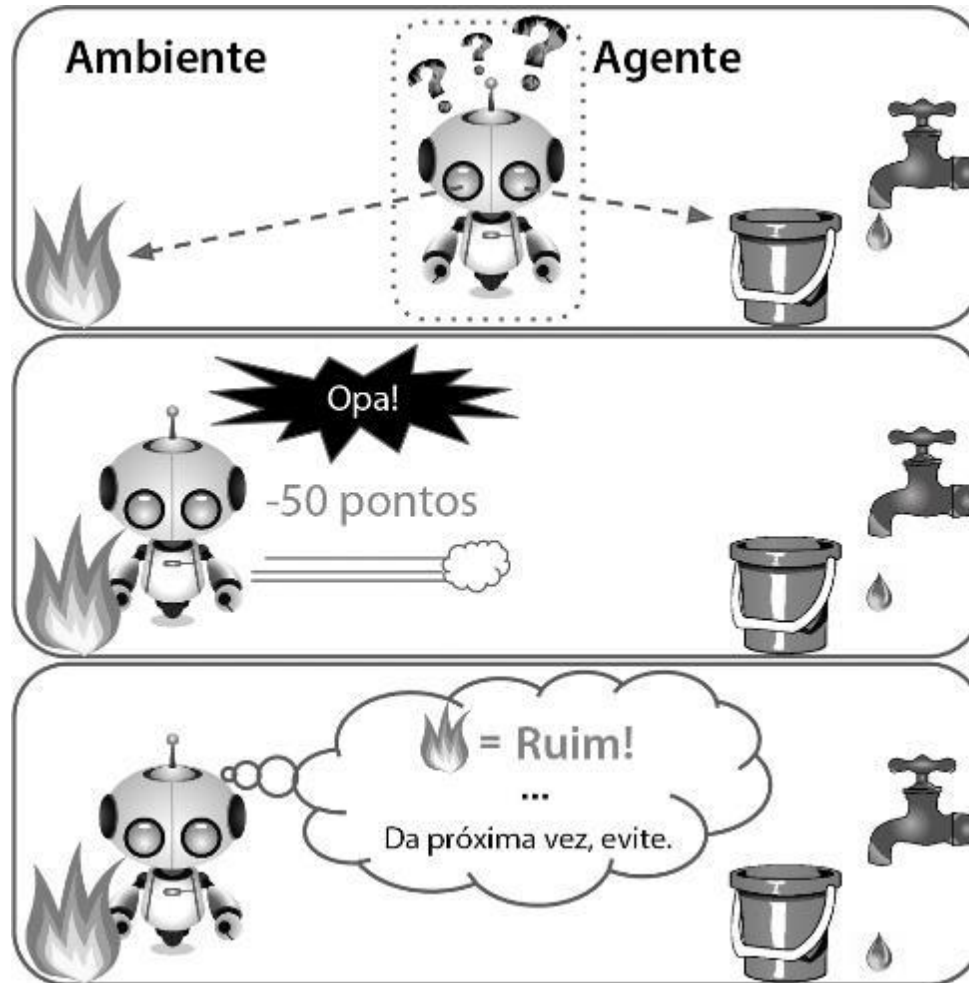
- A maioria dos algoritmos de aprendizado semi-supervisionado são combinações de algoritmos supervisionados e não supervisionados
 - Ex: Redes Neurais de Crenças Profundas (*DBNs – Deep Belief Networks*)
 - Baseadas em componentes não supervisionados, as chamadas máquinas restritas de Boltzman (RBMs), empilhas uma por cima da outra
 - RBMs são treinadas sequencialmente de forma não supervisionada, assim o sistema otimiza usando técnicas de aprendizado supervisionado

Aprendizado por Reforço

- O sistema de aprendizado, denominado de **agente** nesse contexto, pode assistir o ambiente, selecionar e executar ações e obter **recompensas** em troca (ou **penalidades** na forma de recompensas negativas)
- Deve aprender por si só qual é a melhor estratégia (**política**)
- Uma política define qual **ação** o agente deve escolher quando está em determinada situação



Aprendizado por Reforço



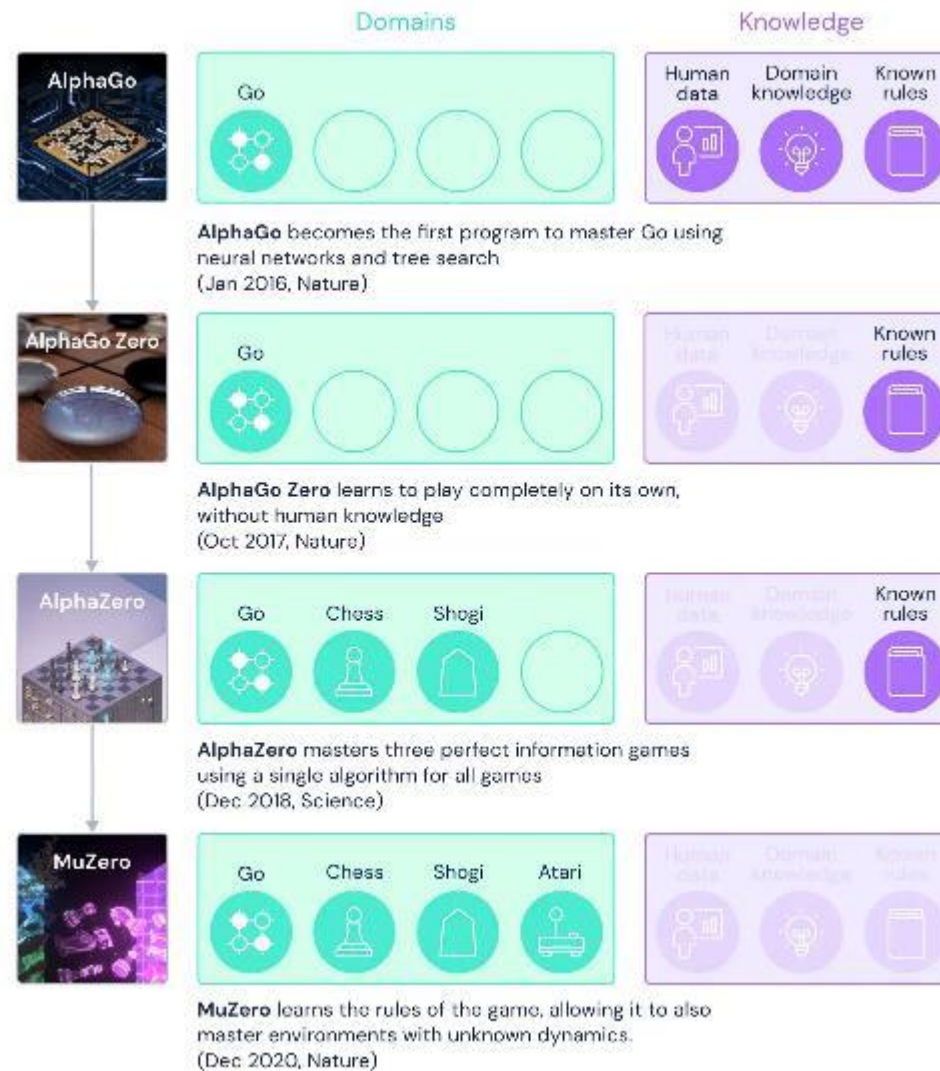
- 1 Observe
- 2 Selecione a ação utilizando a política
- 3 Ação!
- 4 Ganhe a recompensa ou a penalidade
- 5 Atualize a política (etapa de aprendizado)
- 6 Itere até encontrar uma boa política

Aprendizado por Reforço

- Programa AlphaGo da Deep Mind é um bom exemplo de aprendizado por reforço
- Venceu o campeão mundial Ke Jie no Jogo Go (maio de
 - Aprendeu sua política de vitória analisando milhões de jogos e depois jogando muitos jogos contra si mesmo
 - O aprendizado foi desativado durante os jogos contra o campeão
 - AlphaGo estava usando a política que havia aprendido



Aprendizado por Reforço Profundo



Próxima Aula

Aprendizado em batch e online