

Aprendizagem por Reforço e IA em Jogos

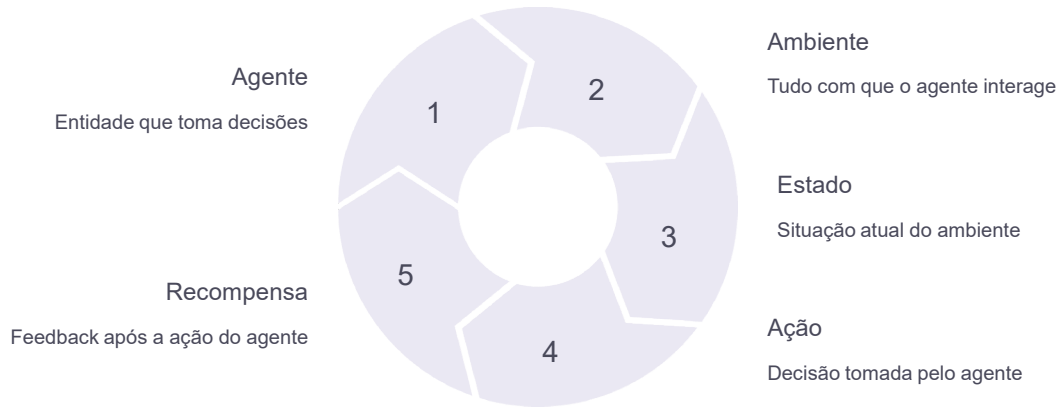
Prof. Edson Pedro Ferlin



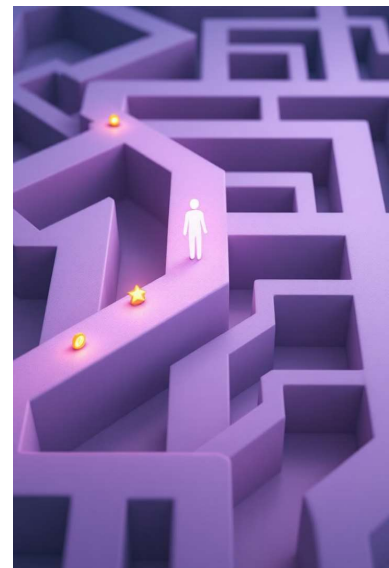
O que é Aprendizado por Reforço?

- 1 **Aprendizado de Máquina**
Agente aprende a tomar decisões sequenciais
- 2 **Maximização de Recompensa**
Busca maximizar recompensa cumulativa
- 3 **Interação com Ambiente**
Realiza ações e recebe recompensas ou penalidades

Componentes do Aprendizado por Reforço



Como o Aprendizado por Reforço Funciona?



Aplicações do Aprendizado por Reforço

Jogos

Agentes aprendendo a jogar
Xadrez, Go, Dota 2

Robótica

Robôs realizando tarefas
complexas

Controle de Sistemas

Otimização de tráfego, gerenciamento de energia



Exemplo de Aplicação em Jogos

AlphaGo

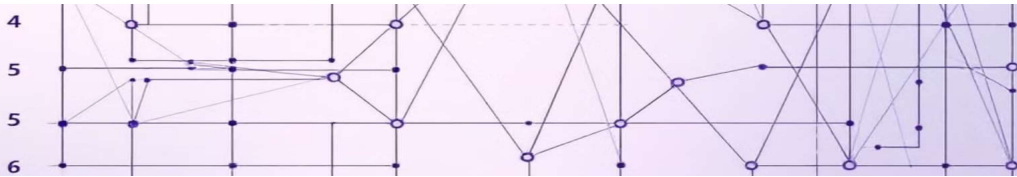
Agente da DeepMind que superou campeões mundiais de Go



Dota 2

OpenAI Five jogou contra profissionais humanos





Algoritmos de Aprendizagem por Reforço

Q-Learning

Algoritmo baseado em tabela para estimar qualidade das ações

Deep Q-Networks (DQN)

Usa redes neurais para aproximar função de valor Q

Policy Gradient

Aprende diretamente a política, não o valor das ações

Biblioteca OpenAI Gym



Variedade de Ambientes

Fornece ambientes para treinar agentes



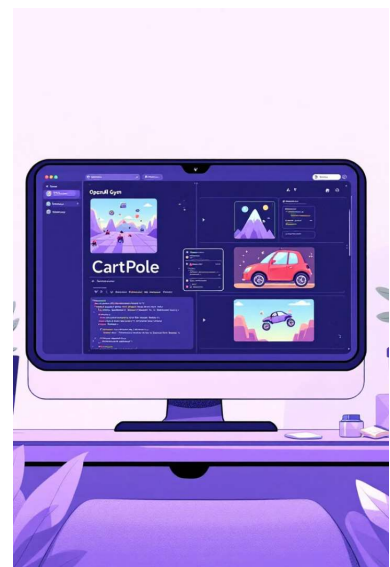
Testes Padronizados

Permite testar diferentes algoritmos



Comparação Simples

Facilita comparação entre algoritmos



Como o Gym Funciona?



Exemplo de Criação de um Agente Simples com Gym

Criar o Ambiente: Usamos `gym.make("CartPole-v1")` para criar o ambiente "CartPole", onde o agente deve equilibrar um pêndulo.

Inicializar o Ambiente: `env.reset()` reinicia o ambiente e retorna o estado inicial.

Ação Aleatória: `env.action_space.sample()` retorna uma ação aleatória que o agente toma.

Interação com o Ambiente: `env.step(action)` aplica a ação e retorna o novo estado, recompensa, se o episódio acabou, e outras informações.

Renderização: `env.render()` exibe o ambiente durante a execução.

```
import gym

# Criar o ambiente
env = gym.make("CartPole-v1")

# Inicializar o ambiente
state = env.reset()

# Realizar uma ação aleatória no ambiente
for _ in range(1000):
    env.render() # Exibir o ambiente
    action = env.action_space.sample() # Ação aleatória
    state, reward, done, info = env.step(action) # Realizar a ação

    if done:
        break

env.close()
```


Contato



eferlin@live.com



(BLOG) professorferlin.blogspot.com

(SITE) professorferlin.webnode.com.br

(YOUTUBE) [ProfEdsonPedroFerlin](https://www.youtube.com/ProfEdsonPedroFerlin)