



Aplicação de Deep Q-Learning para Estratégias de Negociação com Ações

Uma abordagem baseada em aprendizado por reforço com dados históricos da Vale, Petrobras e BR Foods.

Integrantes:

- 1) RM357067 - Edinaldo Rodrigues de Oliveira Junior
- 2) RM358158 - Henrique Cardoso
- 3) RM358067 - Jefferson de Souza Santos
- 4) RM357344 - Wallace Vinicius Silva dos Santos



Objetivo:

Desenvolver um agente inteligente baseado em Deep Q-Learning capaz de aprender estratégias de compra e venda de ações, superando a abordagem Buy & Hold.

Objetivos específicos:

- Construir um ambiente de simulação de portfólio realista.
- Avaliar desempenho em múltiplos ativos.
- Comparar resultados com estratégias passivas.

Fluxo de Trabalho

1. Coleta e tratamento dos dados.
2. Criação do ambiente de simulação personalizado.
3. Definição do estado e espaço de ações.
4. Implementação do agente DQN.
5. Treinamento do agente com histórico de preços.
6. Armazenamento e visualização de métricas.
7. Análise e comparação com Buy & Hold.

Fonte de dados

Ativos utilizados:

- VALE3 (Vale)
- PETR4 (Petrobrás)
- BRFS3 (BR Foods)

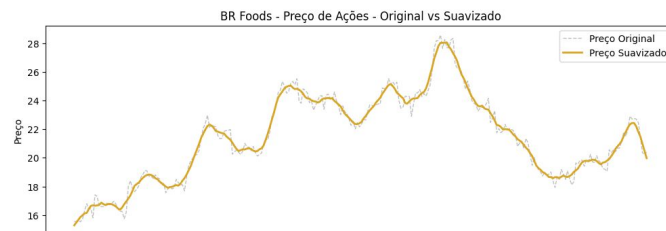
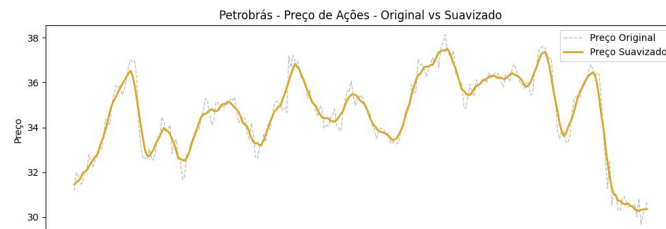
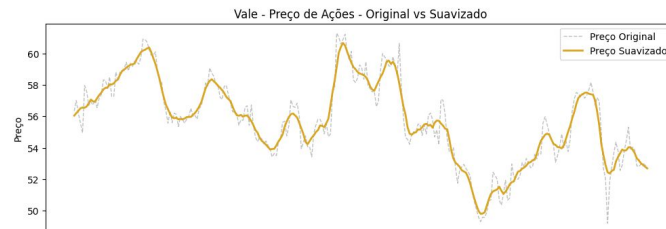
Período analisado: 2024-04-01 a 2025-05-08

Fonte: Yahoo Finance

Pré-processamento:

- Suavização de preços com média móvel exponencial.
- Conversão de datas e sincronização de tamanhos para gráficos.
- Normalização do estado para entrada na rede neural.

Suavização



Fluxo do modelo

Estrutura do Ambiente

Custom `TradingEnvironment`:

- **Estado:** [preço atual, saldo, quantidade de ações, média 3 e 5 dias, volatilidade, valores das ações nos últimos 5 dias]
- **Ações possíveis:** [0] Vender, [1] Manter, [2] Comprar
- **Recompensa:** Mudança no valor do portfólio. Favorecendo a variação com o saldo inicial.

Parâmetros definidos:

- `initial_balance` = R\$10.000
- Portfólio reiniciado a cada episódio.

Rede Neural

Estrutura da rede neural (Q-Network):

Entrada: vetor de estado com 9 variáveis

Camadas totalmente conectadas:

- `fc1`: 128 neurônios → ReLU
- `fc2`: 256 neurônios → ReLU
- `fc3`: 128 neurônios → ReLU
- `fc4`: 64 neurônios → ReLU
- `fc5`: saída com `action_size` neurônios (representando as ações possíveis)

Ativação: ReLU em todas as camadas ocultas

Saída: valores Q estimados para cada ação possível (`vender`, `manter`, `comprar`)

Rede Neural

Configurações principais:

- `gamma` = 0.99 (desconto de recompensas futuras)
- `epsilon` decrescente (exploração → exploração)
- `batch_size` = 128, `memory_size` = 300.000
- `target_update_freq` = 100

Funcionalidades:

- `act()`: política ϵ -greedy (exploração controlada)
- `store_experience()`: buffer com replicação de bons episódios
- `train()`: aprendizado via Bellman + minibatch
- `update_target_network()`: estabilidade
- `update_epsilon()`: reduz exploração com o tempo

Fluxo do modelo

Treinamento

Número de episódios: 400

Métricas monitoradas:

- Recompensa total por episódio.
- Valor final do portfólio.
- Episódio com melhor desempenho.

Decisões adicionais:

- Treinamento contínuo a cada passo.
- Armazenamento de valores diários do portfólio para análise.

```
Episode 10/400 completed with portfolio of 9253. Com sugestão de ação 1 e recompensa acumulada de 92.4.
Episode 20/400 completed with portfolio of 10322. Com sugestão de ação 1 e recompensa acumulada de 113.7.
Episode 30/400 completed with portfolio of 9547. Com sugestão de ação 2 e recompensa acumulada de 104.6.
Episode 40/400 completed with portfolio of 8918. Com sugestão de ação 2 e recompensa acumulada de 90.9.
Episode 50/400 completed with portfolio of 9373. Com sugestão de ação 0 e recompensa acumulada de 107.3.
Episode 60/400 completed with portfolio of 9453. Com sugestão de ação 0 e recompensa acumulada de 95.5.
Episode 70/400 completed with portfolio of 9708. Com sugestão de ação 1 e recompensa acumulada de 144.5.
Episode 80/400 completed with portfolio of 8869. Com sugestão de ação 0 e recompensa acumulada de 104.8.
Episode 90/400 completed with portfolio of 10315. Com sugestão de ação 2 e recompensa acumulada de 172.8.
Episode 100/400 completed with portfolio of 9936. Com sugestão de ação 0 e recompensa acumulada de 158.9.
Episode 110/400 completed with portfolio of 9334. Com sugestão de ação 1 e recompensa acumulada de 142.2.
Episode 120/400 completed with portfolio of 9085. Com sugestão de ação 1 e recompensa acumulada de 130.2.
Episode 130/400 completed with portfolio of 9081. Com sugestão de ação 0 e recompensa acumulada de 145.6.
Episode 140/400 completed with portfolio of 11259. Com sugestão de ação 1 e recompensa acumulada de 202.2.
Episode 150/400 completed with portfolio of 10269. Com sugestão de ação 0 e recompensa acumulada de 186.0.
Episode 160/400 completed with portfolio of 9072. Com sugestão de ação 0 e recompensa acumulada de 156.2.
Episode 170/400 completed with portfolio of 10663. Com sugestão de ação 0 e recompensa acumulada de 197.8.
Episode 180/400 completed with portfolio of 10682. Com sugestão de ação 2 e recompensa acumulada de 208.9.
Episode 190/400 completed with portfolio of 10004. Com sugestão de ação 0 e recompensa acumulada de 195.4.
Episode 200/400 completed with portfolio of 10802. Com sugestão de ação 1 e recompensa acumulada de 219.5.
Episode 210/400 completed with portfolio of 10904. Com sugestão de ação 2 e recompensa acumulada de 222.9.
Episode 220/400 completed with portfolio of 11077. Com sugestão de ação 0 e recompensa acumulada de 242.5.
Episode 230/400 completed with portfolio of 12855. Com sugestão de ação 1 e recompensa acumulada de 310.7.
Episode 240/400 completed with portfolio of 10652. Com sugestão de ação 0 e recompensa acumulada de 228.4.
Episode 250/400 completed with portfolio of 11290. Com sugestão de ação 2 e recompensa acumulada de 241.0.
```

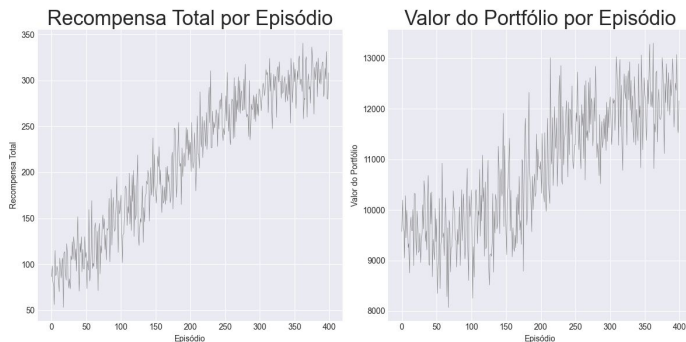
Resultado Vale

Portfólio (Buy & Hold): R\$9,556.10

Portfólio Episódio Final (400): R\$12,157.95

Portfólio Melhor Episódio (362): R\$13,298.30

Ganho Percentual Máximo: Vale: +39,18%



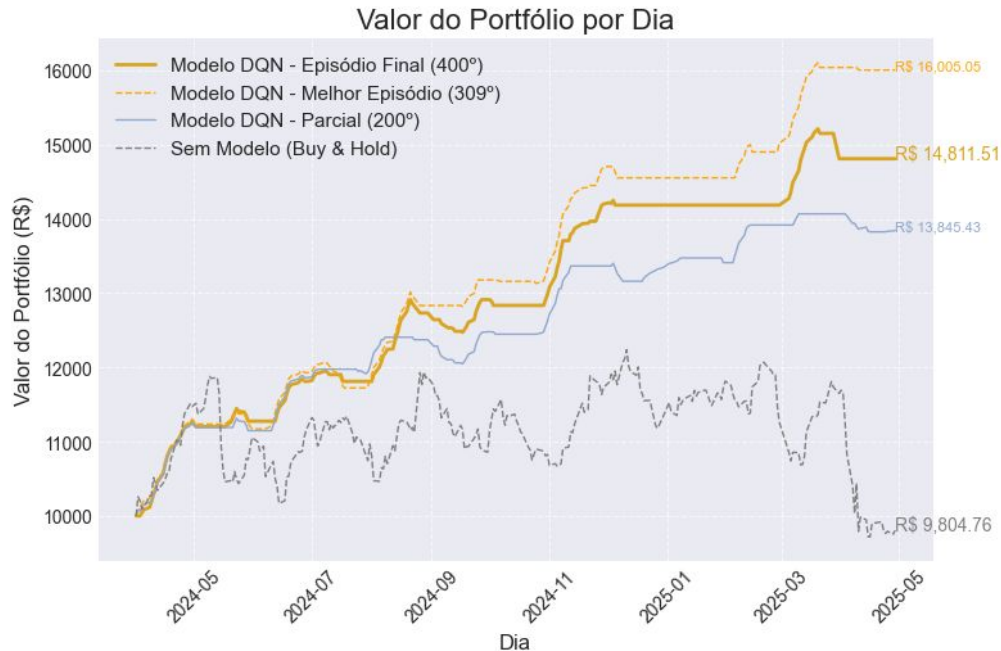
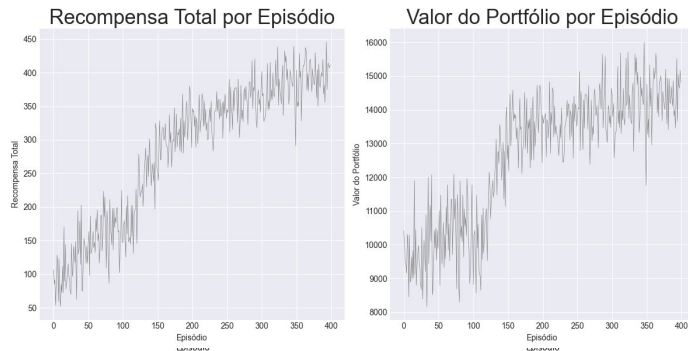
Resultado Petrobras

Portfólio (Buy & Hold): R\$9,804.76

Portfólio Episódio Final (400): R\$14,811.51

Portfólio Melhor Episódio (309): R\$16,005.05

Ganho Percentual Máximo: Vale: +63,25%



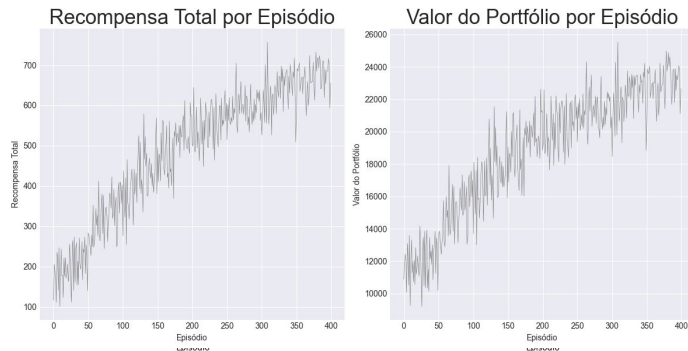
Resultado BR Foods

Portfólio (Buy & Hold): R\$14,789,29

Portfólio Episódio Final (400): R\$22,665.64

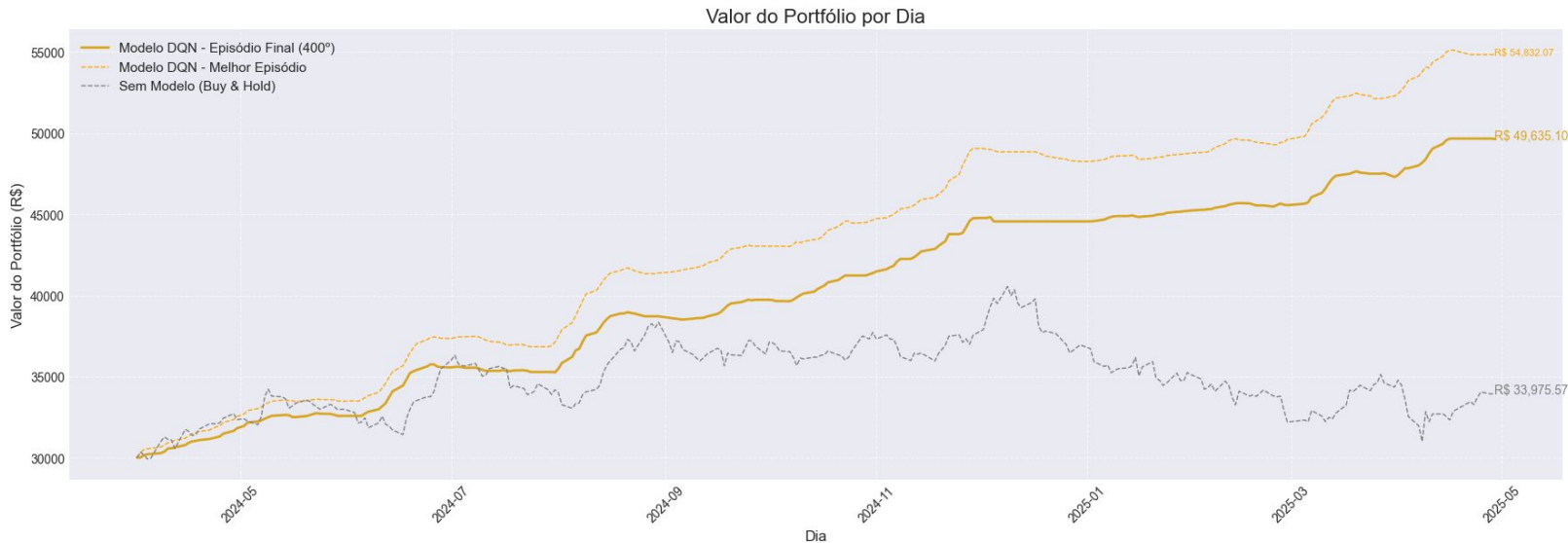
Portfólio Melhor Episódio (309): R\$25,528.72

Ganho Percentual Máximo: Vale: +72,59%



Resultado Carteira

Carteira (Todas as Ações)	Buy & Hold (R\$)	Episódio Final (R\$)	Crescimento (%)	Melhor Episódio (R\$)	Crescimento (%)
	33.975,57	49.635,10	+46,09%	54.832,07	+61,44%





Conclusões

- O agente DQN superou a estratégia Buy & Hold em todos os testes.
- Os resultados indicam aprendizado efetivo de padrões de compra/venda.
- Algoritmos de RL têm alto potencial para trading automatizado.

Obrigado