

REINFORCEMENT LEARNING

TRABALHO FINAL

Projeto QuantumFinance: Fundo Automatizado com Reinforcement Learning

Entregantes:

BONG - RM359114

PAULA - RM359240

ROBERTO - RM359038

08/06/2025

Projeto QuantumFinance: Fundo Automatizado com Reinforcement Learning

O mercado financeiro é um ambiente dinâmico, complexo e de alta incerteza. Decisões de investimento, tradicionalmente tomadas por analistas humanos, são suscetíveis a vieses emocionais e limitações na capacidade de processar grandes volumes de dados em tempo real.

A **QuantumFinance** busca inovar ao criar um fundo de investimentos automatizado para operar ativos da bolsa brasileira (B3), utilizando Aprendizagem por Reforço (RL) para desenvolver um sistema autônomo capaz de aprender estratégias de negociação lucrativas para os ativos Vale (VALE3), Petrobrás (PETR4) e Brasil Foods (BRFS3).



Problemática e Desafio

A problemática central do projeto é: **como desenvolver um sistema autônomo capaz de aprender uma estratégia de negociação lucrativa para ativos da B3, adaptando-se às variações do mercado sem intervenção humana?**

A Aprendizagem por Reforço surge como solução promissora, permitindo que um "agente" aprenda a tomar decisões ótimas através de tentativa e erro, recebendo recompensas ou penalidades por suas ações, de forma análoga a um operador de mercado que aprende com seus lucros e prejuízos.



Ambiente Complexo

Mercado financeiro com alta volatilidade e múltiplas variáveis



Decisão Autônoma

Sistema capaz de operar sem intervenção humana



Aprendizado Contínuo

Adaptação constante às novas condições de mercado



Motivação e Objetivos

A principal **motivação** é criar um diferencial competitivo para a QuantumFinance, utilizando Inteligência Artificial para desenvolver estratégias de trading que superem as limitações humanas e se adaptem rapidamente às mudanças do mercado.

Estratégias Adaptativas

Capazes de ajustar suas táticas às novas condições de mercado, identificando padrões e tendências em tempo real.

Decisões Não-emocionais

Livres de vieses comportamentais como medo e euforia, mantendo consistência mesmo em momentos de alta volatilidade.

Sistemas Escaláveis

Com capacidade de analisar múltiplos ativos e indicadores simultaneamente, processando volumes de dados impossíveis para analistas humanos.

O **objetivo** deste projeto é planejar, desenhar e simular um agente de Reinforcement Learning que aprenda a gerenciar um portfólio com os ativos VALE3, PETR4 e BRFS3, maximizando o retorno financeiro através de decisões autônomas de compra, venda ou manutenção.



Estrutura do Agente de RL

Para resolver este problema, propomos a utilização de um agente baseado em **Deep Q-Network (DQN)**, um algoritmo que combina Q-Learning com redes neurais profundas, ideal para lidar com os espaços de estados contínuos do mercado financeiro.



Ingestão de Dados

Captura de séries temporais históricas de preços dos ativos VALE3, PETR4 e BRFS3 via Yahoo Finance, incluindo preços de Abertura, Máxima, Mínima, Fechamento e Volume.



Processamento

Modelagem como Processo de Decisão de Markov (MDP), com estados representando a situação do mercado e do portfólio, e ações de compra, venda ou manutenção.



Avaliação

Análise da performance do agente através de métricas como Lucro/Prejuízo Total, Evolução do Portfólio e Sharpe Ratio.

Modelagem do Problema

O problema será modelado como um **Processo de Decisão de Markov (MDP)**, com estados, ações e recompensas claramente definidos para permitir o aprendizado do agente.

Estados (State)

Representação do mercado e portfólio incluindo posse de ativos, saldo em dinheiro e histórico de preços dos últimos dias.

Ações (Action)

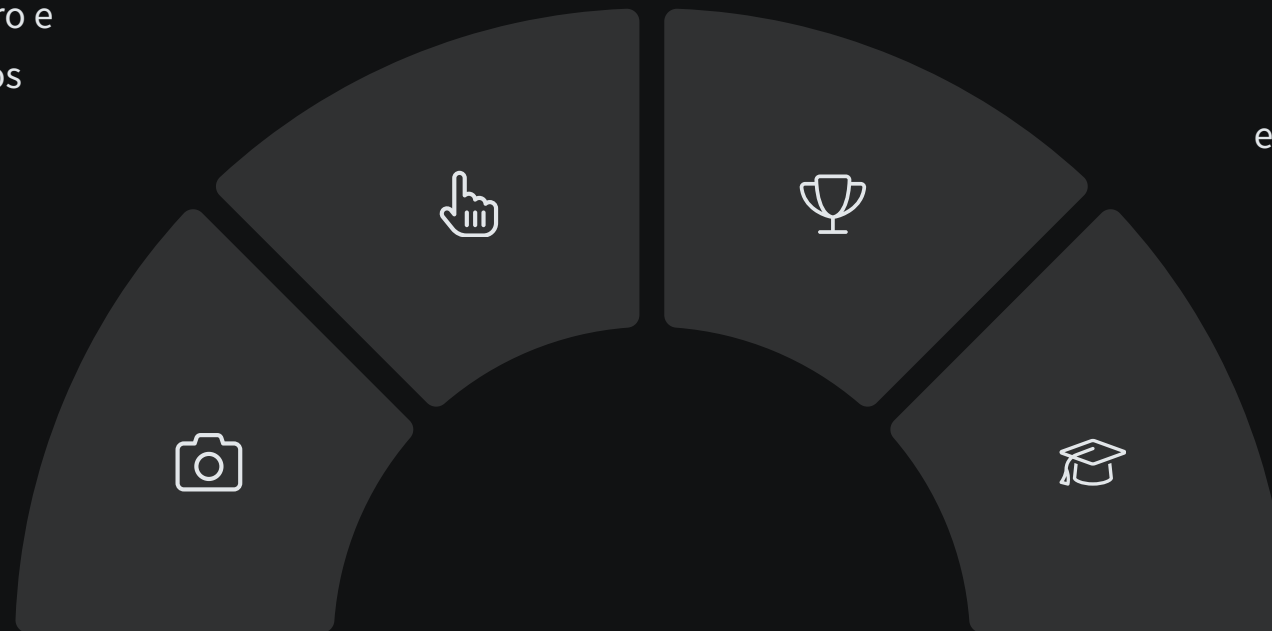
Decisões para cada ativo: Manter (0), Comprar (1) ou Vender (2), resultando em 27 combinações possíveis.

Recompensa (Reward)

Variação do valor total do portfólio de um dia para o outro, incentivando ações que geram lucro.

Algoritmo DQN

Rede neural que aproxima a função $Q(s,a)$, com Experience Replay e estratégia Epsilon-Greedy.



Algoritmo DQN em Detalhes

O **Deep Q-Network (DQN)** é o coração do sistema, combinando redes neurais com Q-Learning para aprender estratégias ótimas de negociação através da experiência acumulada.

Rede Neural

Utilizamos um Perceptron de Múltiplas Camadas para aproximar a função $Q(s, a)$. A rede recebe o estado como entrada e retorna o valor Q esperado para cada ação possível, permitindo a escolha da melhor decisão.

Experience Replay

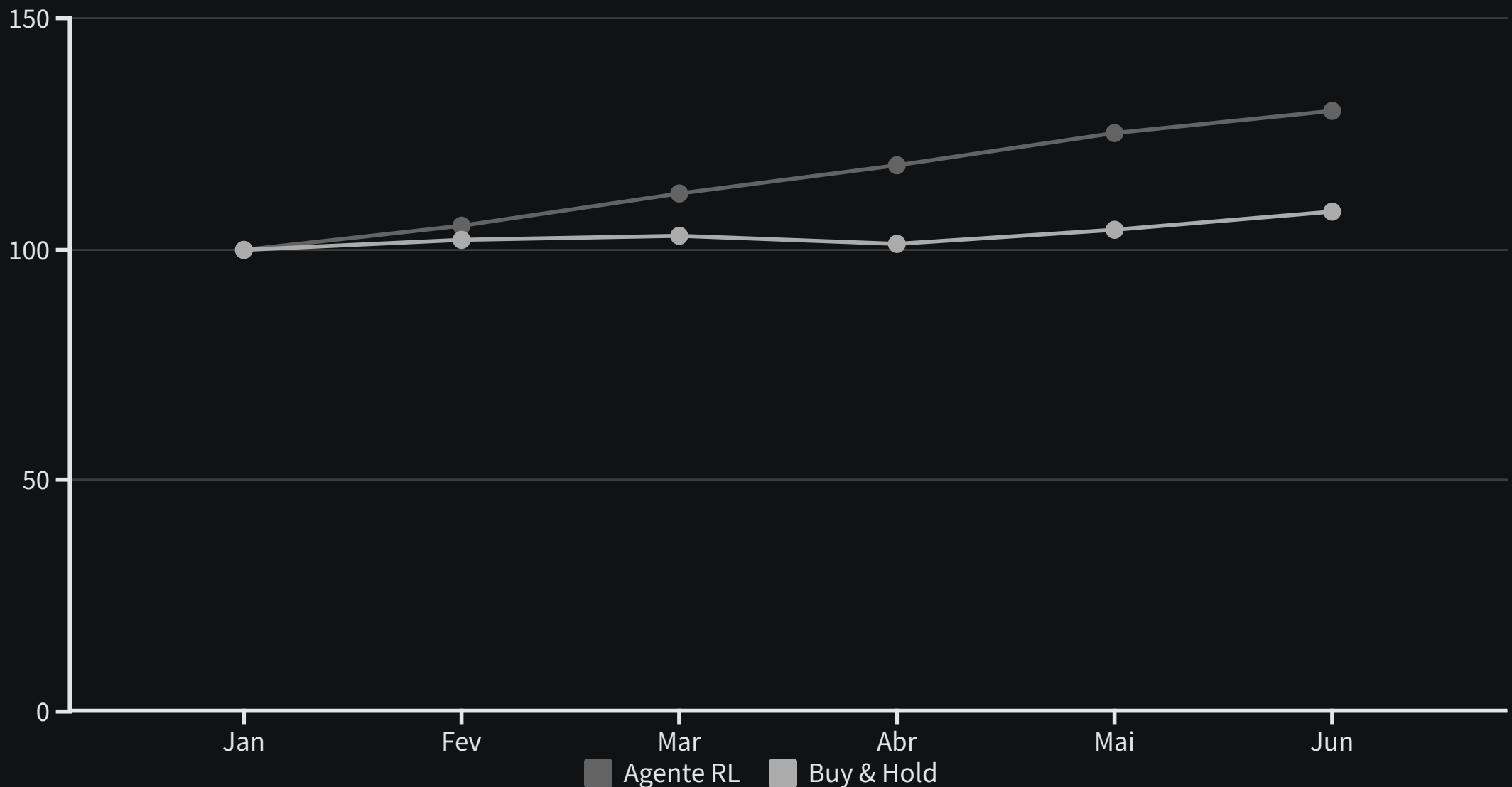
O agente armazena suas experiências (estado, ação, recompensa, próximo_estado) em uma memória. Durante o treino, ele amostra lotes aleatórios dessa memória para treinar a rede, quebrando a correlação entre experiências sequenciais.

Exploration vs. Exploitation

Usando a estratégia Epsilon-Greedy, o agente inicialmente explora o ambiente com ações aleatórias (alto epsilon) e gradualmente passa a confiar mais na sua política aprendida (baixo epsilon).

Avaliação de Performance

A saída final do processo é uma **política de negociação** aprendida pela rede neural. Em um ambiente de simulação (backtesting), avaliamos a performance do agente em um período de teste com dados que ele nunca viu antes.



As principais **métricas de avaliação** incluem o Lucro/Prejuízo Total (valor final do portfólio menos capital inicial), a Evolução do Portfólio comparada com estratégias passivas, e o Sharpe Ratio, que mede o retorno ajustado ao risco.

Considerações e Próximos Passos

Embora promissor, o modelo apresenta limitações que precisam ser consideradas para uma implementação realista, como custos de transação, liquidez e impacto de mercado, além do risco de overfitting nos dados de treinamento.



O potencial futuro do projeto inclui o enriquecimento dos estados com indicadores técnicos como Médias Móveis e RSI, a incorporação de análise de sentimento de notícias, a exploração de algoritmos mais avançados como A2C e PPO, e a implementação de estratégias sofisticadas de gerenciamento de risco.