



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

MATHEUS DO NASCIMENTO ARAÚJO

INTELIGÊNCIA ARTIFICIAL APLICADA AO MERCADO FINANCEIRO:
Uma pesquisa exploratória literária acerca do uso do *Financial Reinforcement Learning (FinRL)* no mercado de ações

Recife

2023

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
SISTEMAS DE INFORMAÇÃO

MATHEUS DO NASCIMENTO ARAÚJO

**INTELIGÊNCIA ARTIFICIAL APLICADA AO MERCADO FINANCEIRO:
Uma pesquisa exploratória literária acerca do uso do *Financial Reinforcement Learning (FinRL)* no mercado de ações**

TCC apresentado ao Curso de Sistemas de Informação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de Bacharel em Sistemas de Informação.

Orientador(a): Prof^o. Dr. Sergio Ricardo de Melo Queiroz

Recife

2023

FICHA CATALOGRÁFICA

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Araujo, Matheus do Nascimento.

Inteligência artificial aplicada ao mercado financeiro: uma pesquisa exploratória literária acerca do uso do Financial Reinforcement Learning (FinRL) no mercado de ações / Matheus do Nascimento Araujo. - Recife, 2023.

43 p

Orientador(a): Sergio Ricardo de Melo Queiroz

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado, 2023.

8.7.

Inclui referências.

1. Inteligência artificial. 2. Aprendizado por reforço profundo (DRL). 3. Automação. 4. Tomada de decisão. 5. Mercado financeiro. I. Queiroz, Sergio Ricardo de Melo. (Orientação). II. Título.

000 CDD (22.ed.)

MATHEUS DO NASCIMENTO ARAÚJO

**INTELIGÊNCIA ARTIFICIAL APLICADA AO MERCADO FINANCEIRO:
Uma pesquisa exploratória literária acerca do uso do *Financial Reinforcement Learning (FinRL)* no mercado de ações**

TCC apresentado ao Curso de Sistemas de Informação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de Bacharel em Sistemas de Informação.

Aprovado em: 18/09/2023.

BANCA EXAMINADORA

Profº. Dr. Sergio Ricardo de Melo Queiroz (Orientador)
Universidade Federal de Pernambuco

Profº. Dr. Francisco de Assis Tenório de Carvalho (Examinador Interno)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Ao meu pai e à minha mãe, por sempre terem tido paciência e amor comigo;

Aos meus amigos de infância, pelas décadas de amizade inabalável;

Ao meu irmão, por me ensinar a humildade de não ser egoísta;

Aos meus amigos de escola, pela companhia de sempre;

Ao Matt, por mostrar que a TIC era o meu caminho;

Ao Prof. Sérgio, pelo suporte nesta caminhada;

Ao meu Centro-Oeste, o coração do Brasil;

A Pernambuco, minha enorme paixão;

À vovó e aos meus antepassados;

Ao Wilson, ao Gui e ao Pedro,

bem como a todos os de-

mais colegas que tor-

naram a gradua-

ção mais fe-

liz e leve.

Axé!

.

EPÍGRAFE

"A primeira regra de qualquer tecnologia utilizada nos negócios é que a automação aplicada a uma operação eficiente aumentará a eficiência. A segunda é que a automação aplicada a uma operação ineficiente aumentará a ineficiência."

Bill Gates

RESUMO

No contexto atual das estratégias de investimento, inseridas no dinâmico cenário de negociação de ações, a importância do uso da Inteligência Artificial (IA) e os desafios subjacentes à sua aplicação tornam-se evidentes. Dentre os métodos mais utilizados, a Aprendizagem Profunda por Reforço (DRL) emerge como uma abordagem promissora para enfrentar esses desafios. Os estudos selecionados para esta pesquisa exploratória propõem uma estratégia de alocação de portfólio composta, que integra, entre outros, os algoritmos de Otimização de Política Próxima (PPO), Ator Crítico com Vantagem (A2C) e Gradiente de Política Determinística Profunda (DDPG), os quais se destacam por sua adaptabilidade a diversas situações de mercado. Os algoritmos previamente mencionados são empregados no *framework* aberto conhecido como *Financial Reinforcement Learning (FinRL)*, o qual os combina e introduz novas metodologias, surgindo como uma solução para simplificar a aplicação da Aprendizagem Profunda por Reforço no mercado de ações. Além disso, a ferramenta é capaz de simular ambientes de treinamento e é customizável para atender às restrições legais em diferentes jurisdições financeiras. A contribuição prática do *FinRL* é evidenciada pela disponibilidade de tutoriais que facilitam e propõem uma aplicação tangível da Aprendizagem Profunda por Reforço em cenários de negociação. Em resumo, todos os algoritmos e o *framework* mencionados oferecem melhorias substanciais nas estratégias de negociação em mercados financeiros complexos. Visto isso, este estudo, por meio de pesquisa exploratória literária, visa analisar a aplicação da aprendizagem de máquina no mercado de títulos, com o propósito de avaliar a estratégia previamente mencionada.

Palavras-chave: inteligência artificial; aprendizado por reforço profundo (DRL); automação; tomada de decisão; mercado financeiro.

ABSTRACT

In the current scenario of investment strategies, within the dynamic landscape of stock trading, the significance of utilizing Artificial Intelligence (AI) and the underlying challenges of its application become evident. Among the most employed methods, Deep Reinforcement Learning (DRL) emerges as a promising approach to address these challenges. The selected studies for this exploratory research propose a composite portfolio allocation strategy that integrates algorithms such as Proximal Policy Optimization (PPO), Actor-Critic with Advantage (A2C), and Deep Deterministic Policy Gradient (DDPG), known for their adaptability to a multitude of market situations. These aforementioned algorithms are employed within the open framework known as Financial Reinforcement Learning (FinRL), which combines and introduces novel methodologies, serving as a solution to simplify the application of Deep Reinforcement Learning in the stock market. Moreover, the tool can simulate training environments and is customizable to adhere to legal constraints in different financial jurisdictions. The practical contribution of FinRL is demonstrated through the availability of tutorials that facilitate and propose a tangible application of Deep Reinforcement Learning in trading scenarios. In summary, all the mentioned algorithms and the framework offer substantial enhancements in trading strategies within complex financial markets. Considering the context presented above, this study, through literary research, aims to analyze the application of machine learning in the securities market, with the purpose of evaluating the aforesaid strategy.

Keywords: artificial intelligence; deep reinforcement learning (DRL); automation; decision-making; financial market.

SUMÁRIO

1 INTRODUÇÃO

1.1 Contexto

1.2 Perguntas de Pesquisa

1.3 Objetivos gerais e específicos

1.3.1 Objetivos gerais

1.3.2 Objetivos específicos

1.4 Organização do Trabalho

2 REFERENCIAL TEÓRICO

2.1 Inteligência artificial (IA)

2.1.1 Definição, tipos e aplicabilidade

2.1.2 Aprendizado por reforço profundo (DRL)

2.1.2.1 Processos de decisão de Markov (MDPs)

2.1.3 Automação e tomada de decisão

2.1.4 OpenAI

2.2 Inteligência artificial no mercado financeiro

2.2.1 Otimização de Política Próxima (PPO)

2.2.2 Ator Crítico de Vantagem (A2C)

2.2.3 Gradiente de Política Determinística Profunda (DDPG)

2.2.4 *Financial Reinforcement Learning (FinRL)*

2.2.5 Padrão Gym

2.2.6 *Deep Q-Learning (DQN)*

3 METODOLOGIA

3.1 Bibliografia relacionada ao tema

3.2 Análise da documentação do *FinRL*

3.3 Análise gráfica do uso de algoritmos combinados no mercado de ações

3.4 Análise amostral do código-fonte do *FinRL*

3.5 Considerações divergentes sobre o uso do *FinRL* e da IA no mercado de ações

4 RESULTADOS E DISCUSSÕES

4.1 Questões da pesquisa

4.2 Discussão dos objetivos

4.3 Trabalhos futuros

5 CONCLUSÃO

6 REFERÊNCIAS

1 INTRODUÇÃO

1.1 Contexto

No conjunto de fatores que influenciam os investimentos, as estratégias de negociação de ações desempenham um papel fundamental na busca por retornos lucrativos. Entretanto, essa tarefa se torna um desafio complexo em um mercado em constante mutação. Para enfrentar esse cenário, a abordagem da aprendizagem profunda por reforço (DRL) tem ganhado destaque como uma promissora ferramenta nas finanças quantitativas, oferecendo a perspectiva de aprimorar consideravelmente as estratégias de negociação. Vale ressaltar, também, que o setor financeiro é o ambiente quase ideal para o uso de aprendizagem profunda, pois o mesmo tem regras definidas, procedimentos regrados e recompensas (YANG *et al.*, 2020).

Uma proposta inovadora emerge nesse contexto: a criação de um método composto por algoritmos trabalhando em conjunto, aproveitando e potencializando a DRL para aprender estratégias de negociação de ações. Tal abordagem, proposta pela equipe do *framework* (YANG *et al.*, 2020), combina três algoritmos principais (dentre outros): a Otimização de Política Próxima (PPO), o Ator Crítico de Vantagem (A2C) e o Gradiente de Política Determinística Profunda (DDPG). A ideia central é amalgamar as virtudes de cada um desses algoritmos, bem como outros, em uma estratégia híbrida que possa se adaptar flexivelmente às diferentes condições do mercado. O resultado é um *framework* revolucionário chamado *FinRL (Financial Reinforcement Learning)*, que surge como um recurso para tornar a aprendizagem profunda por reforço mais acessível e aplicável no cenário do mercado de ações (LIU *et al.*, 2021).

O *FinRL* é projetado com base em princípios fundamentais que incluem a simplicidade, aplicabilidade, extensibilidade, personalização e reprodutibilidade. Estruturado em três camadas modulares (incluindo os algoritmos citados anteriormente), o *framework* incorpora códigos refinados para o contexto financeiro. Essa abordagem reduz consideravelmente a complexidade da implementação e minimiza a necessidade de depuração extensiva, acelerando significativamente o processo de criação e teste de estratégias (LIU *et al.*, 2021).

Uma característica distintiva do *FinRL* é a capacidade de simular ambientes de treinamento, utilizando tanto dados históricos, quanto *APIs* de negociação em tempo real. Isso confere aos *traders* a flexibilidade de treinar seus agentes em diversos cenários de mercado, ajustando a granularidade do tempo conforme necessário. Além disso, o *FinRL* é altamente customizável, permitindo que os usuários adaptem o *framework* às suas próprias restrições e preferências de negociação (LIU *et al.*, 2021).

1.2 Questões de Pesquisa

Dentro do contexto supracitado, o presente trabalho almeja, em consonância às possibilidades atuais, responder às seguintes perguntas:

Q1. Qual o impacto que o uso amplo da DRL pode vir a trazer no futuro?

Q2. Há diminuição de riscos reais e aumento do retorno financeiro com o uso de algoritmos?

Q3. O uso do *FinRL* para o mercado de ações é confiável e robusto?

Q4. Qual a performance do *framework* em comparação aos índices do mundo real?

1.3 Objetivos gerais e específicos

1.3.1 Objetivos Gerais:

Este estudo, por meio de pesquisa exploratória empírica e literária, tem como objetivo avaliar a aplicação da aprendizagem de máquina no mercado de ações, por intermédio do *FinRL*.

1.3.2 Objetivos Específicos:

- Avaliar o uso da DRL no mercado de ações;
- Avaliar os riscos relacionados a esta utilização;
- Mensurar a confiabilidade e a robustez do *FinRL*, através de pesquisa empírica e literária;
- Investigar brevemente o código-fonte.

1.4 Organização do trabalho

Este documento está organizado em cinco capítulos. O primeiro capítulo é a introdução, cujo objetivo é contextualizar o presente trabalho de conclusão de curso, bem como suas questões e objetivos. No segundo capítulo, são apresentados alguns dos fundamentos teóricos do estudo. O terceiro capítulo detalha a metodologia de pesquisa, com enfoque na abordagem de pesquisa exploratória literária, sendo esta dividida em etapas, bem como explicitar as limitações do estudo. No quarto capítulo, são revelados os resultados da pesquisa e discussões. Por fim, no quinto e último capítulo, é discutida a conclusão do texto, bem como possíveis caminhos para trabalhos futuros.

2 REFERENCIAL TEÓRICO

2.1 Inteligência Artificial (IA)

2.1.1 Definição, Tipos e Aplicabilidade

A inteligência artificial (IA) refere-se à simulação de processos de inteligência humana por meio de sistemas de computador (RUSSELL; NORVIG, 2016). Esses sistemas são projetados para realizar tarefas que normalmente requerem inteligência humana, como reconhecimento de padrões, tomada de decisões, resolução de problemas e aprendizado. Existem diferentes tipos de IA, incluindo IA fraca e forte. A IA fraca, também conhecida como IA estreita, é projetada para executar tarefas específicas e é amplamente utilizada em aplicativos como assistentes virtuais e chatbots (IBM, 2020). Já a IA forte é capaz de realizar tarefas gerais de inteligência humana e ainda é um campo em desenvolvimento (IBM, 2020).

2.1.2 Aprendizado por Reforço Profundo (DRL)

O Aprendizado por Reforço Profundo (DRL) é uma abordagem interdisciplinar que combina técnicas do aprendizado profundo e do aprendizado por reforço. Por meio do uso da DRL, os agentes aprendem a tomar decisões (também entendido como realizar ações) em ambientes complexos por meio da interação direta com eles (SUTTON; BARTO, 2018). Essa metodologia se baseia no conceito de recompensas, onde o agente busca maximizar as recompensas acumuladas ao longo do tempo, adaptando suas ações para otimizar o desempenho em tarefas específicas. Uma característica central da DRL é a modelagem das interações entre os agentes e os ambientes por meio dos Processos de Decisão de Markov (MDPs), que são estruturas matemáticas que formalizam a sequência de estados e ações em um ambiente. A ilustração abaixo (Imagem 2.1) mostra como, esquematicamente, os MDPs e o processo de DRL são semelhantes. A diferenciação ocorre, como a figura mostra, no agente, que é complexo e sem intervenção humana direta, contendo múltiplas camadas de processamento, e

o ambiente, que é de larga escala e complexo - no que tange aos dados a serem observados (WU *et al.*, 2019).

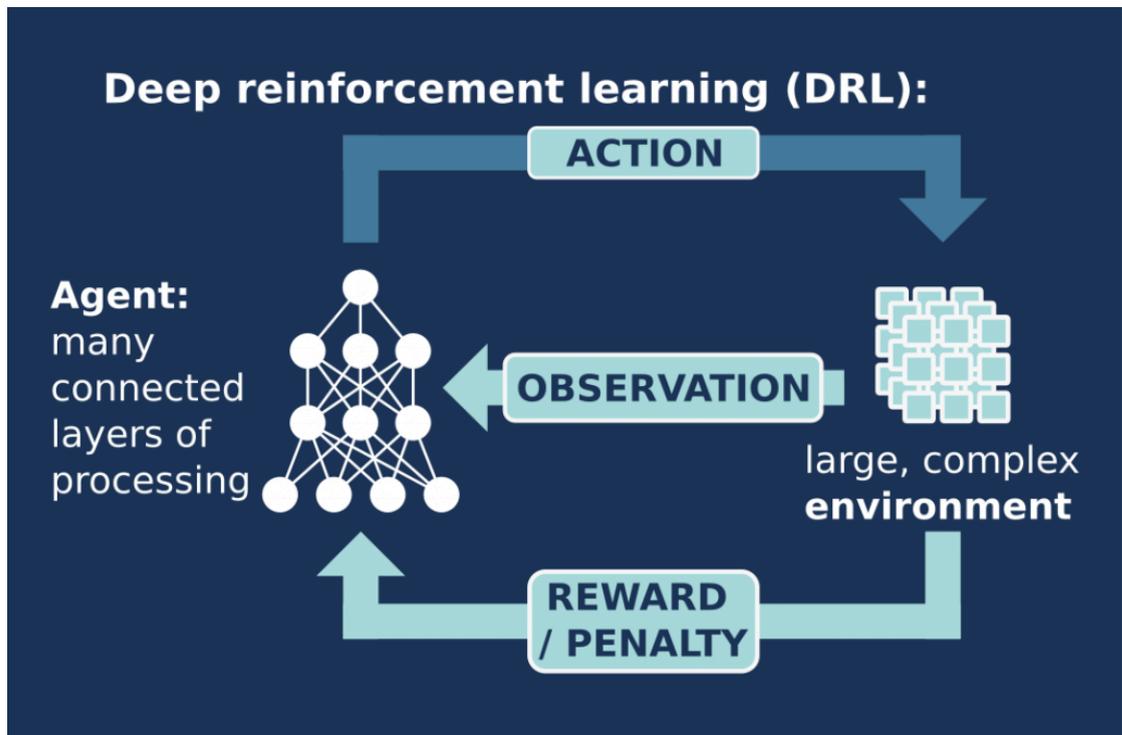


Imagem 2.1: Uma esquematização do DRL. (WU *et al.*, 2019)

2.1.2.1 Processos de decisão de Markov

Os Processos de Decisão de Markov (MDPs) são uma estrutura fundamental na modelagem de ambientes de aprendizado por reforço, incluindo o Aprendizado por Reforço Profundo (DRL) (SUTTON; BARTO, 2018). Eles descrevem uma sequência de estados e ações em um ambiente, em que a transição de um estado para outro é governada pela propriedade de Markov, ou seja, o futuro é condicionalmente independente do passado, dado o estado atual. Os MDPs incluem estados, ações, recompensas, uma função de transição que descreve a probabilidade de transição entre estados após uma ação e uma função de recompensa que atribui valores às ações realizadas nos estados. Essa estrutura permite que agentes de IA aprendam a tomar decisões sequenciais para maximizar recompensas ao longo do tempo (PUTERMAN, 2014). A Imagem 2.2 abaixo, adaptada do livro *Foundations of Reinforcement Learning with applications in Finance* (RAO; JELVIS, 2023), demonstra como as entidades descritas anteriormente

interagem entre si neste processo de iterações ao longo do tempo.

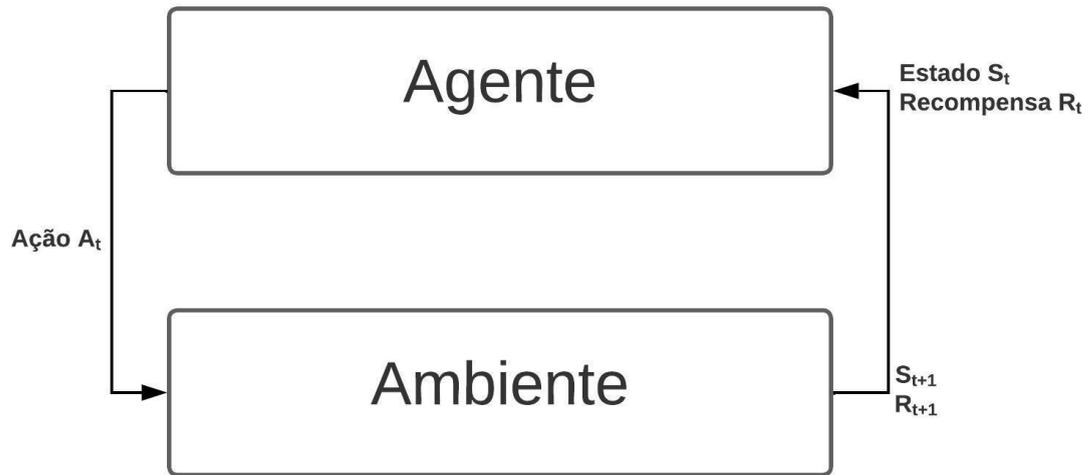


Imagem 2.2: O *framework* dos MDPs. (Adaptado de RAO; JELVIS, 2023)

2.1.3 Automação e Tomada de Decisão

A automação no contexto da inteligência artificial se refere à capacidade de sistemas autônomos executarem tarefas ou processos sem intervenção humana direta. Na tomada de decisão automatizada, algoritmos e modelos de IA são empregados para analisar dados, identificar padrões e tomar decisões com base em regras predefinidas ou aprendidas. Isso é especialmente relevante em finanças, onde algoritmos podem tomar decisões de compra ou venda de ativos com base em análises de mercado em tempo real ou histórico. (LUGER, 2019)

2.1.4 OpenAI

OpenAI é uma organização de pesquisa em inteligência artificial com a intenção declarada de desenvolver inteligência artificial geral segura e benéfica. Ela define inteligência artificial geral como “sistemas altamente autônomos que superam os humanos na maioria dos trabalhos economicamente valiosos”. Ela é proprietária da *OpenAI Gym*. (THE TIMES OF INDIA, 2023).

2.2 Inteligência Artificial no Mercado Financeiro

2.2.1 Otimização de Política Próxima (PPO)

O PPO é um algoritmo de otimização de política proximal que pertence à família dos métodos de aprendizado por reforço. Ele é projetado para aprimorar a política de um agente (ou seja, suas ações em diferentes estados) enquanto equilibra a exploração de novas ações com a maximização das recompensas obtidas. O PPO se destaca por seu desempenho estável e capacidade de lidar com grandes conjuntos de dados, tornando-o uma escolha popular em aplicações financeiras. (SCHULMAN *et al.*, 2017)

2.2.2 Ator Crítico de Vantagem (A2C)

O A2C é um algoritmo de aprendizado por reforço que combina as abordagens do Ator Crítico e do Aprendizado Profundo. Ele divide o processo de aprendizado em dois componentes principais: o "Ator", que aprende ações ótimas com base na política atual, e o "Crítico", que avalia o valor esperado das ações em um determinado estado. Essa combinação permite um aprendizado mais eficiente e estável. (MNIH *et al.*, 2016)

2.2.3 Gradiente de Política Determinística Profunda (DDPG)

O DDPG é um algoritmo que lida com espaços de ação contínuos em problemas de aprendizado por reforço. Ele utiliza redes neurais profundas para aproximar a função de valor e a política determinística, permitindo que o agente tome ações em um espaço contínuo. O DDPG é particularmente útil em aplicações financeiras, onde as ações podem variar em uma escala contínua. (LILLICRAP *et al.*, 2016)

2.2.4 Financial Reinforcement Learning (FinRL)

O *FinRL* é um *framework* de aprendizado por reforço especialmente projetado para aplicações financeiras. Ele fornece ferramentas e algoritmos adaptados ao contexto do mercado financeiro, facilitando a criação de estratégias automatizadas de negociação. O *FinRL* abrange tanto a modelagem do ambiente financeiro quanto a escolha de ações, permitindo que os agentes aprendam a otimizar suas decisões em cenários reais. (AI4FINANCE FOUNDATION, 2021)

2.2.5 Padrão Gym

O “Padrão Gym” refere-se à *OpenAI Gym*, que é uma biblioteca de código aberto em Python para desenvolver e comparar algoritmos de aprendizado por reforço. Ela fornece uma API padrão para comunicação entre algoritmos de aprendizado e ambientes, bem como um conjunto padrão de ambientes compatíveis com essa API. Desde o seu lançamento, a API do Gym tornou-se o padrão do campo para fazer tais feitos. (BROCKMAN *et al.*, 2016)

2.2.6 Deep Q-Learning (DQN)

O Q-learning é um algoritmo de aprendizado por reforço que, diferentemente do *framework Q-Learning* original (que utiliza tabelas, ao invés de redes neurais), não depende de um modelo para aprender o valor de uma ação em um estado específico (MATIISEN, 2015). Ele é capaz de lidar com situações em que não há necessidade de um modelo do ambiente e pode enfrentar desafios que envolvem transições e recompensas estocásticas sem precisar de ajustes adicionais, podendo produzir diversas ações partindo apenas de um estado, como podemos ver abaixo na Imagem 2.3. (SINGH, 2022)

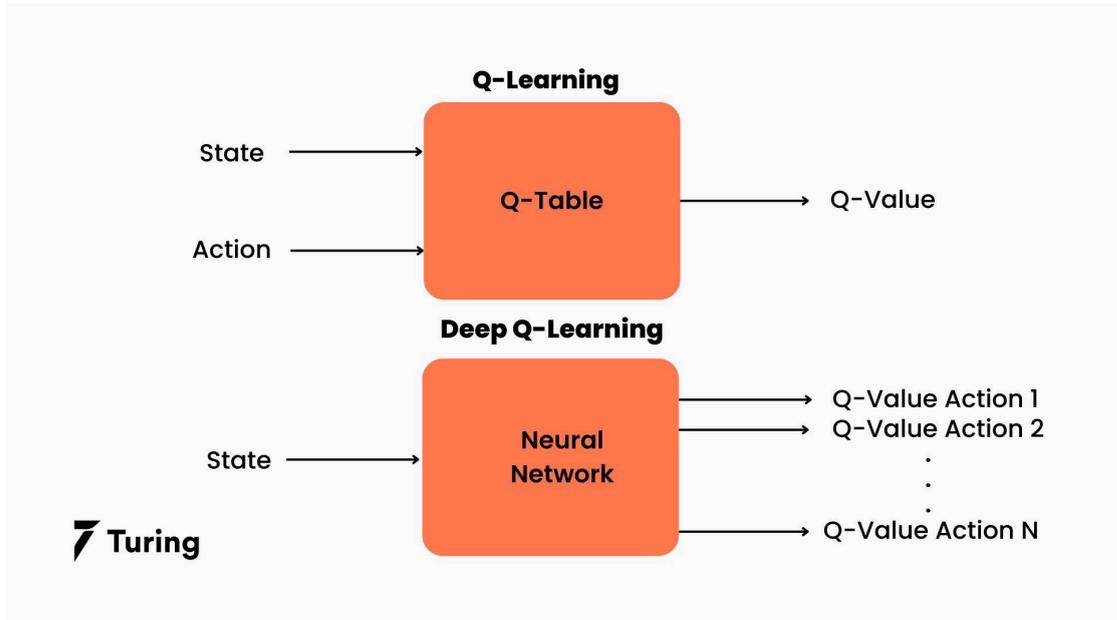


Imagem 2.3: *Q-Learning vs. Deep Q-Learning*. (Singh, *A comprehensive guide to neural networks in Deep Q-Learning*, 2022)

3 METODOLOGIA

A metodologia empregada neste estudo envolve a abordagem de pesquisa exploratória empírica e literária. Nesse sentido, busca-se compreender e mapear de forma abrangente o contexto em estudo, sem a intenção de oferecer respostas definitivas ou conclusivas (THEODORSON; THEODORSON, 1970). Este estudo permitirá identificar elementos e variáveis relevantes, bem como estabelecer diretrizes para pesquisas mais detalhadas no futuro. É importante frisar que, devido ao fato de ser um estudo teórico literário, existem limitações, sendo parte delas a dependência de fontes secundárias, a ausência de *surveys* e o foco em informações já existentes.

Este capítulo está dividido em cinco partes: levantamento bibliográfico inicial relacionado ao tema, análise da documentação do *FinRL* e a análise do gráfico de performances dos mercados de ações projetados pelo *framework* em comparação à performance do índice *Dow Jones Industrial Average* (DJIA), avaliação amostral dos códigos do projeto (disponíveis em seu *GitHub*) e investigação sobre as divergências que existem em relação ao uso do *FinRL* e da IA no mercado de ações

3.1 Bibliografia relacionada ao tema

Diversos estudos vêm abordando o tema do uso da DRL e do *deep learning* no mercado de ações. Enquanto pesquisadores em Oxford exploraram a aplicação da aprendizagem profunda na gestão de portfólio (ZHANG; ZOHREN; ROBERTS, 2020), em Chongqing realizou-se uma abrangente pesquisa sobre o uso de DRL para o gerenciamento do mesmo (HUANG; ZHOU; SONG, 2020); já em Istambul, dois engenheiros filiados ao IEEE (Instituto de Engenheiros Elétricos e Eletrônicos), buscaram discutir o uso intenso da DRL no mercado de ações, saindo de técnicas de aprendizado supervisionado e introduzindo o uso de automação total da compra e venda de ações (KABBANI; DUMAN, 2022). Esses são apenas alguns dos artigos que destacam a importância do uso de *deep learning* e de aprendizado por reforço profundo ao lidar com a complexidade das decisões de investimento, melhorando a eficiência e a eficácia das estratégias de negociação.

Na implementação de estratégias da DRL no contexto do mercado de ações, as ferramentas disponíveis desempenham um papel crucial. Duas das bibliotecas mais amplamente utilizadas nesse campo são *TensorFlow* e *PyTorch*. O *TensorFlow*, desenvolvido pelo Google, é uma plataforma de código aberto amplamente reconhecida que oferece uma ampla gama de ferramentas para desenvolver e treinar modelos de aprendizado de máquina, incluindo a DRL (ABADI *et al.*, 2016). Por outro lado, o *PyTorch* é uma estrutura de aprendizado profundo que se destacou por sua flexibilidade e abordagem mais intuitiva, permitindo que os pesquisadores experimentem de forma rápida e eficaz com ideias de modelagem complexas (PASZKE *et al.*, 2019).

No entanto, para abordagens específicas na área financeira, como a negociação de ações, surgiram bibliotecas direcionadas que simplificam ainda mais a implementação de algoritmos da DRL. Duas delas são o *Stable Baselines 3* e o *FinRL*. O *Stable Baselines 3* é uma extensão da biblioteca *OpenAI Baselines*, oferecendo implementações pré-ajustadas de diversos algoritmos da DRL, como *DQN*, *A2C* e *PPO*, que podem ser diretamente aplicados a cenários de negociação. Isso reduz a curva de aprendizado e facilita a prototipagem de estratégias de negociação baseadas em DRL (RAFFIN *et al.*, 2020). Por sua vez, o *FinRL* é uma biblioteca específica para finanças, construída sobre o *TensorFlow*. Ela foi projetada para envolver os desafios únicos da negociação de ações, oferecendo ferramentas que simplificam a integração de dados históricos e de mercado, bem como a implementação de estratégias da DRL. Ela será explicitada mais especificamente a seguir (LIU *et al.*, 2020).

Existem ainda três outras bibliotecas mencionadas pelos autores da *FinRL*: a *Dopamine*, a *Rlib* e a *TensorLayer*. A *Google Dopamine* é uma biblioteca *open source* que busca facilitar o desenvolvimento rápido de algoritmos de Aprendizado por Reforço Profundo (DRL), enfatizando a facilidade de uso e adaptabilidade (CASTRO; BELLEMARE, 2019). O *Rlib*, que faz parte do projeto *Ray*, é um *framework* de código aberto para dimensionar aplicativos de IA e Python, como aprendizado de máquina. Ele permite a implementação de uma ampla variedade de algoritmos com alto desempenho e escalabilidade (LIANG, 2018). Por fim, a *TensorLayer*, baseada no *TensorFlow*, é destinada a pesquisadores que buscam redes neurais personalizáveis, sendo projetada com simplicidade, flexibilidade e alto

desempenho em mente. A biblioteca possui uma abstração de camada de alto nível que é de fácil aprendizado e APIs transparentes e flexíveis. (DONG, 2017).

3.2 Análise da documentação do *FinRL*

A documentação, per se, do *framework* não é tão extensa em seu GitHub, o que desencadeou uma análise mais detalhada dos cinco artigos publicados pelos autores. O primeiro artigo analisado, no caso, propõe o *FinRL* como um projeto de código aberto em criação (LIU *et al.*, 2018). Já o último (LIU *et al.*, 2022), demonstra que a biblioteca já operava com três camadas distintas: camada de aplicação, camada de agente e camada de ambiente, de forma refinada e com foco total no mercado, sem a presença das criptomoedas. A Imagem 3.1, que será utilizada para ilustrar, será a de um dos artigos de 2020 (YANG *et al.*, 2020).

A camada de aplicação visa disponibilizar múltiplas tarefas de negociação de demonstração, atuando como base para os usuários desenvolverem suas próprias estratégias (LIU *et al.*, 2021). Neste trabalho, tendo como base o trabalho intermediário de 2020 (YANG *et al.*, 2020), exploraremos a utilização da parte de mercado de ações e alocação de portfólio desta camada, que também envolve *High Frequency Trading* - um método financeiro rápido e automatizado com velocidades, taxas de rotatividade e proporções de ordens elevadas, baseado em dados financeiros de alta frequência e ferramentas de negociação eletrônicas - (RUNDO *et al.*, 2019), negociações em criptomoedas, impacto de regulações estatais sobre o mercado e outras ações definidas pelo usuário.

A camada de agente, por sua vez, facilita a integração de algoritmos DRL otimizados provenientes de bibliotecas existentes, seguindo um fluxo unificado (Imagem 3.2). A imagem, que é uma adaptação da ilustração clássica de um MDP (presente no referencial teórico deste trabalho), traz o estado da carteira - "*State*" - e a recompensa da atividade algorítmica - "*Reward*" - como entes influenciados pelo ambiente e que influenciam os agentes de negociação - "*Trading Agents*" -, que são um amálgama dos mais diversos algoritmos utilizados no modelo. Já as ações - de compra, venda e manutenção da posse do título - a serem tomadas - "*Actions*" - seguem um fluxo inverso, onde os negociadores influenciam o ambiente através de suas escolhas, de forma mimética ao mundo real (YANG *et al.*, 2020). É importante

ressaltar que, além dos algoritmos previamente mencionados neste trabalho, existem derivados do DDPG, do *Deep Q Learning (DQN)* e outras abordagens algorítmicas presentes, todos inseridos no contexto de bibliotecas DRL como *ElegantRL*, *RLlib*, *Stable Baselines3* (LIU *et al.*, 2021).

Já a camada de ambiente tem a finalidade de abranger dados históricos e APIs de negociação em tempo real de diversos mercados em ambientes de treinamento, aderindo ao *padrão Gym*. Nas tarefas de negociação da camada superior, é possível acionar diretamente os algoritmos DRL presentes na camada de agente e os ambientes de mercado da camada de ambiente (LIU *et al.*, 2020). O design do ambiente é crucial na aprendizagem profunda por reforço, uma vez que o agente aprende por meio de interações de tentativa e erro com o mesmo. Um ambiente bem projetado, que emula mercados do mundo real, aprimora a aprendizagem. Ressalta-se que, devido à natureza estocástica e interativa das atividades financeiras de mercado, elas são modeladas como Processos de Decisão de Markov (MDPs) (RUNDO *et al.*, 2019).

No *FinRL*, a camada de ambiente é responsável por observar os dados de mercado atuais e traduzir essas informações em estados para o problema do MDP. As variáveis de estado são categorizadas em aquelas que representam o estado do agente e o estado do mercado (LIU *et al.*, 2020). Por exemplo, no caso da negociação de ações, o estado do mercado engloba os preços de abertura, alta, baixa e fechamento, volume e indicadores técnicos. O estado do agente inclui o saldo da conta e as posições de ações. Todos esses passos não divergem do que é feito por *traders* no cotidiano (INFOMONEY, 2022).

O processo de treinamento de um *RL* envolve a observação de mudanças nos preços, tomada de ações e cálculo de recompensas. Ao interagir com o ambiente, o agente se atualiza iterativamente e, eventualmente, adquire uma estratégia de negociação para maximizar os retornos esperados (JONES, 2017). Aqui, o modelo segue essa cartilha basilar. O *FinRL* transforma dados de mercado reais em ambientes de treinamento no estilo *Gym*, seguindo o princípio da simulação baseada no tempo. Inspirado pelo *OpenAI Gym*, o *framework*, de forma mimética ao seu correspondente, oferece uma coleção de ambientes de treinamento versáteis para várias tarefas de negociação (BROCKMAN *et al.*, 2016).

Para além da sua capacidade funcional, a eficácia do *FinRL* também depende da qualidade dos dados e da robustez dos algoritmos utilizados (LIU *et al.*, 2020). Para que os utilizadores consigam otimizar a sua estratégia de negociação, é essencial que tenham uma compreensão sólida do ambiente de negociação, das variáveis do mercado e de como estas interagem entre si. Dada a volatilidade inerente aos mercados financeiros, a tomada de decisões rápidas e bem informadas é crucial (ASHFORD, 2023).

Um dos maiores desafios na DRL é garantir que os modelos sejam generalizáveis a diferentes condições de mercado. Isso é particularmente verdadeiro no contexto financeiro, onde as condições de mercado podem mudar rapidamente devido a uma miríade de fatores externos, como mudanças políticas, desastres naturais, entre outros (AN; SUN; WANG, 2022). A capacidade do *FinRL* de simular essas condições, integrando múltiplos conjuntos de dados e algoritmos, oferece uma vantagem considerável na preparação de agentes de negociação para o mundo real (LIU *et al.*, 2021).

Além disso, a modularidade do *FinRL*, com suas camadas distintas, permite aos utilizadores adaptarem ou expandirem facilmente as funcionalidades existentes. Por exemplo, enquanto o ambiente atual fornece dados históricos e *APIs* de negociação em tempo real, futuros desenvolvedores podem querer integrar fontes de dados adicionais ou implementar técnicas de aprendizado mais avançadas para refinar a tomada de decisão do agente (LIU *et al.*, 2021).

Concluindo, o *FinRL* emerge como uma solução inovadora no campo da DRL para negociação financeira. A sua estrutura modular e a integração de algoritmos de ponta fornecem um quadro robusto para pesquisa e desenvolvimento. Contudo, como em qualquer ferramenta, a eficácia final do *framework* depende em grande parte da capacidade dos seus usuários de interpretar corretamente os dados, entender o ambiente de negociação e adaptar-se continuamente às mudanças no mercado (DENG *et al.*, 2017).

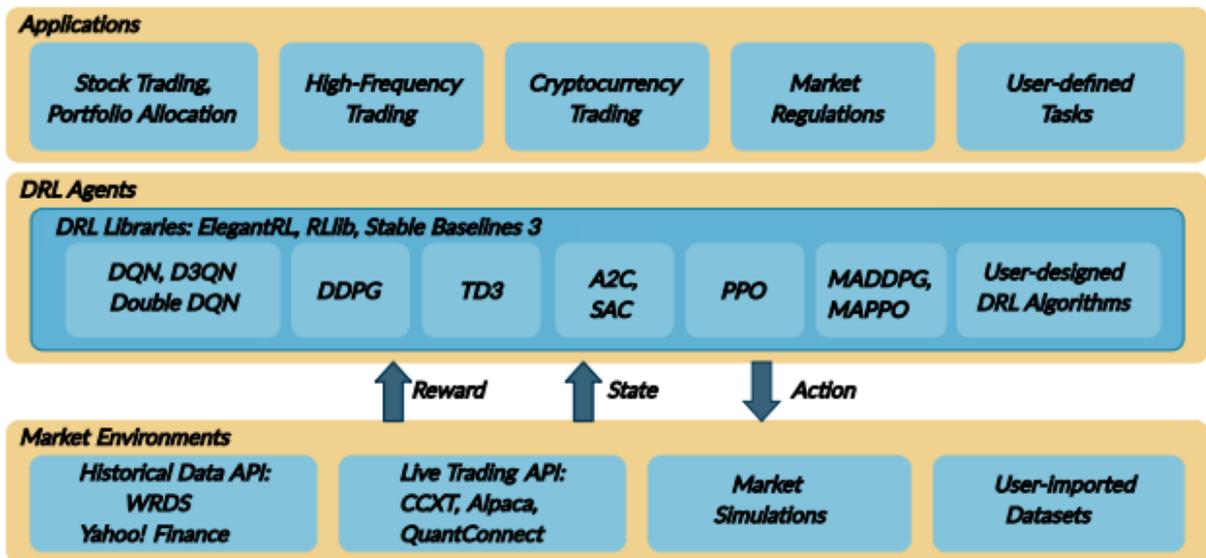


Imagem 3.1 - (YANG *et al.*, 2020)

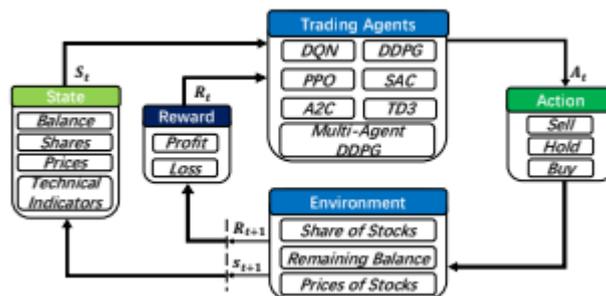


Imagem 3.2 - (YANG *et al.*, 2020)

3.3 Análise gráfica do uso de algoritmos combinados no mercado de ações

No artigo nomeado "FinRL-Meta: Ambientes de Mercado e Referências para Aprendizado por Reforço Financeiro Orientado a Dados", em tradução livre, de Yang *et al* (2022), apresentado na *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, os autores amalgamam experimentos utilizando o *framework* criado pelos próprios com a intenção de apresentá-los ao público (LIU *et al.*, 2022). Este corrente trabalho se atentará a um dos gráficos de performance comparada presentes no artigo. No gráfico abaixo (Imagem 3.3), os autores avaliam o retorno acumulado no mercado de ações no mundo real (DJIA) em comparação com o uso dos algoritmos usados de forma isolada (A2C, PPO e DDPG). Por fim, temos a estratégia *Ensemble*, proposta inicialmente em um dos artigos de 2020 (YANG *et al.*, 2020), que seria uma abordagem utilizando múltiplos algoritmos para potencializar o retorno da carteira formada pelos papéis do *industrial average* do índice americano Dow Jones, calculado pelo *Wall Street Journal* (LIU *et al.*, 2022). Os dados coletados se referem à biblioteca do *Yahoo! Finance* no intervalo de tempo abaixo.

O *Ensemble* reflete, basicamente, o modo como o *FinRL* trabalha. Nota-se no gráfico que, a estratégia composta supera, virtualmente, todos os demais algoritmos usados de maneira individualizada, bem como o comportamento da carteira de ações *per se* (LIU *et al.*, 2022). No experimento, o agente conjunto tem a maior relação *Sharpe* de 1.53, o que significa que ele tem o melhor desempenho no equilíbrio entre riscos e lucros. O índice *Sharpe* é uma métrica usada nas finanças para avaliar o retorno ajustado ao risco de um investimento (PINTO, 2021). Ele é calculado dividindo o excesso de retorno do investimento pelo seu desvio padrão. Quanto maior o índice *Sharpe*, melhor o desempenho do investimento em relação ao risco assumido (PINTO, 2021).



Imagem 3.3 - Performance do *Ensemble* vs. Performance “humana” (DJIA) vs. Performance dos algoritmos individualizados. (LIU et al., 2022)

3.4 Análise amostral do código-fonte do *FinRL*

Por incentivo do canal Hudson & Thames do Youtube (HUDSON & THAMES, 2023), este trabalho também optou por analisar brevemente uma parte dos códigos dos algoritmos do *FinRL*. Para evitar uma análise longa, o foco aqui será no arquivo *main.py* do repositório principal do projeto (AI4FINANCE FOUNDATION, 2023). Porém, antes de investigarmos o mesmo, na Imagem 3.4, veremos a seção de testes unitários. São apenas seis arquivos, dos quais apenas três estão minimamente completos. Dois encontram-se em branco e um encontra-se incompleto. Testes são essenciais em projetos dessa magnitude. Indicam uma boa prática de engenharia de software.

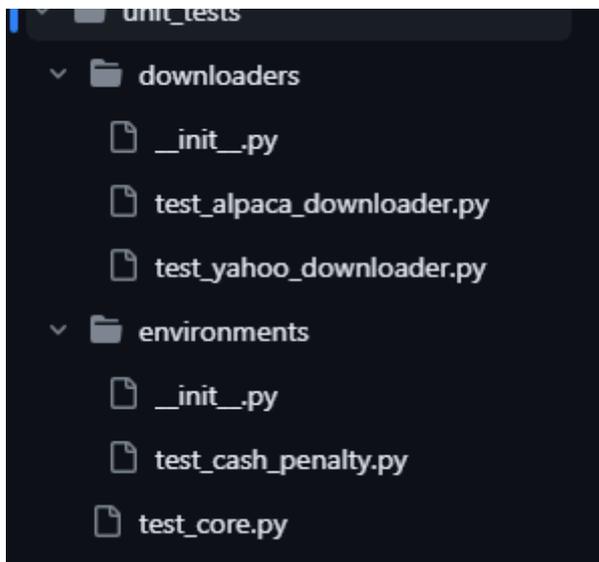


Imagem 3.4: Testes unitários do *FinRL*.

(AI4FINANCE FOUNDATION, 2023)

Na Imagem 3.5, nota-se uma repetição da chamada dos construtores. O código chama treze vezes a mesma biblioteca, de modo a importar os seus módulos. Tal processo poderia ser simplificado, visto que o arquivo de origem chamado é bem pequeno e interno à *branch* do projeto. Uma linha de código com “*” bastaria.

```
1  from __future__ import annotations
2
3  import os
4  from argparse import ArgumentParser
5  from typing import List
6
7  from finrl.config import ALPACA_API_BASE_URL
8  from finrl.config import DATA_SAVE_DIR
9  from finrl.config import ERL_PARAMS
10 from finrl.config import INDICATORS
11 from finrl.config import RESULTS_DIR
12 from finrl.config import TENSORBOARD_LOG_DIR
13 from finrl.config import TEST_END_DATE
14 from finrl.config import TEST_START_DATE
15 from finrl.config import TRADE_END_DATE
16 from finrl.config import TRADE_START_DATE
17 from finrl.config import TRAIN_END_DATE
18 from finrl.config import TRAIN_START_DATE
19 from finrl.config import TRAINED_MODEL_DIR
20 from finrl.config_tickers import DOW_30_TICKER
21 from finrl.meta.env_stock_trading.env_stocktrading_np import StockTradingEnv
22
23 # construct environment
24
25 # try:
26 #     from finrl.config_private import ALPACA_API_KEY, ALPACA_API_SECRET
27 # except ImportError:
28 #     raise FileNotFoundError(
29 #         "Please set your own ALPACA_API_KEY and ALPACA_API_SECRET in config_private.py"
30 #     )
31
32
33 def build_parser():
34     parser = ArgumentParser()
35     parser.add_argument(
36         "--mode",
37         dest="mode",
38         help="start mode, train, download_data " " backtest",
39         metavar="MODE",
40         default="train",
41     )
42     return parser
43
```

Imagem 3.5: Código verboso, que poderia ser simplificado. (AI4FINANCE FOUNDATION, 2023)

Na última figura (Imagem 3.6) notamos a falta quase completa de comentários em trechos importantes do código. Para uma ferramenta que se propõe ser inclusiva e de fácil entendimento para seus usuários, seu código deveria, ao menos, seguir uma diretriz similar.

```
def main() -> int:
    end_date=TEST_END_DATE,
    ticker_list=DOW_30_TICKER,
    data_source="yahoofinance",
    time_interval="1D",
    technical_indicator_list=INDICATORS,
    drl_lib="elegantrl",
    env=env,
    model_name="ppo",
    cwd="./test_ppo",
    net_dimension=512,
    kwargs=kwargs,
)
elif options.mode == "trade":
    from finrl import trade

    try:
        from finrl.config_private import ALPACA_API_KEY, ALPACA_API_SECRET
    except ImportError:
        raise FileNotFoundError(
            "Please set your own ALPACA_API_KEY and ALPACA_API_SECRET in config_private.py"
        )
    env = StockTradingEnv
    kwargs = {}
    trade(
        start_date=TRADE_START_DATE,
        end_date=TRADE_END_DATE,
        ticker_list=DOW_30_TICKER,
        data_source="yahoofinance",
        time_interval="1D",
        technical_indicator_list=INDICATORS,
        drl_lib="elegantrl",
        env=env,
        model_name="ppo",
        API_KEY=ALPACA_API_KEY,
        API_SECRET=ALPACA_API_SECRET,
        API_BASE_URL=ALPACA_API_BASE_URL,
        trade_mode="paper_trading",
        if_vix=True,
        kwargs=kwargs,
        state_dim=len(DOW_30_TICKER) * (len(INDICATORS) + 3)
        + 3, # bug fix: for ppo add dimension of state/observations space = len(stocks)* len(INDICATORS) + 3+ 3*len(stocks)
        action_dim=len(
            DOW_30_TICKER
        ), # bug fix: for ppo add dimension of action space = len(stocks)
    )
else:
    raise ValueError("Wrong mode.")
return 0
```

Imagem 3.6: Código com poucos comentários. (AI4FINANCE FOUNDATION, 2023)

3.5 Considerações divergentes sobre o uso do FinRL e da IA no mercado de ações

Alguns artigos anteriores avaliaram o uso do FinRL sob diferentes pontos de vista, sendo um destes o mais proeminente. O estudo conduzido por Rainer Jager (JAGER, 2021) da Universidade de Waikato, Nova Zelândia, encontrou inconsistências estatísticas no modelo proposto, sugerindo a substituição do índice de turbulência fixo por algo flexível, que reflita a volatilidade do mercado financeiro. O autor normalizou os resultados dos índices de Sharpe obtidos, através do modelo de D'Agostino e Pearson (D'AGOSTINO; BELANGER; D'AGOSTINO, 1990), de modo a avaliar os resultados da atividade do *framework FinRL*.

Por meio da análise dos achados, o autor do artigo concluiu que o *Ensemble* proposto pelos criadores do *FinRL* (YANG *et al*, 2020) não apresenta ganhos significativos de performance em relação ao algoritmo DDPG usado de forma isolada. Outra conclusão é que o uso do *CBOE Volatility Index (VIX)*, ao invés de um valor fixo para avaliar a turbulência, seria mais apropriado, pois este é um índice de mercado atualizado em tempo real de acordo com as expectativas do mercado para a volatilidade nos próximos 30 dias (KUEPPER, 2023).

Adicionalmente às divergências em relação à ferramenta, assomam-se preocupações em relação ao uso da inteligência artificial no mercado de ações. É notório que em mercados fortemente digitalizados, como os de ações e câmbio, algoritmos de IA podem melhorar a administração de liquidez e a realização de ordens volumosas com um impacto mínimo, otimizando a performance sem estar em desconformidade às condições e às regulações do mercado (OECD, 2021).

Porém, também de acordo com a OCDE:

“Semelhante aos modelos e algoritmos não baseados em IA, o uso dos mesmos modelos de machine learning por um grande número de profissionais de finanças pode potencialmente levar a um comportamento de manada e mercados unidirecionais, o que, por sua vez, pode aumentar os riscos para a liquidez e estabilidade do sistema, especialmente em momentos de estresse. Embora o algoritmo de negociação com IA possa aumentar a liquidez em tempos normais, também pode levar à convergência e,

consequentemente, a momentos de iliquidez em tempos de estresse e a quedas rápidas no mercado. A volatilidade do mercado pode aumentar através de grandes vendas ou compras executadas simultaneamente, dando origem a novas fontes de vulnerabilidades. A convergência de estratégias de negociação cria o risco de loops de feedback auto-reforçados que podem, por sua vez, desencadear movimentos bruscos de preços. Tal convergência também aumenta o risco de ataques cibernéticos, pois fica mais fácil para os cibercriminosos influenciar agentes que agem da mesma maneira.” (OECD, 2021)

4 RESULTADOS E DISCUSSÕES

Antes de iniciarmos a discussão, é importante ressaltar que o FinRL possui as seguintes características: uma arquitetura em camadas que garante transparência e permite interações exploratórias entre os níveis de agente e ambiente, modulação e extensibilidade proporcionadas por módulos independentes em cada camada, além de uma abordagem simples e aplicável que oferece tarefas de negociação referencial, customizáveis por meio de configurações intuitivas, e tutoriais acessíveis para iniciantes (LIU *et al.*, 2022). Com isso em mente, vamos ao desenvolvimento deste capítulo.

4.1 Questões de pesquisa

Com base nas investigações feitas, podemos, enfim, responder às questões previamente colocadas neste trabalho. Primeiramente, o uso amplo da DRL pode ter um impacto significativo no futuro, especialmente no campo financeiro e de negociação de ações. Com a crescente disponibilidade de dados financeiros em tempo real e o avanço contínuo das técnicas de aprendizado de máquina, espera-se que o uso da aprendizagem profunda por reforço possa levar a uma maior automatização e otimização das estratégias de negociação. Isso pode resultar em maior eficiência na tomada de decisões, bem como na identificação e exploração de oportunidades de mercado mais complexas (YANG *et al.*, 2020). No entanto, também é importante considerar os desafios éticos e regulatórios que podem surgir com o uso amplo dessas tecnologias.

Em segundo lugar, o uso de algoritmos, incluindo aqueles baseados na DRL, pode potencialmente levar à diminuição de riscos e aumento do retorno no mercado de ações. Algoritmos de negociação podem analisar rapidamente grandes volumes de dados, identificar padrões sutis e reagir a mudanças de mercado em tempo real. Isso pode ajudar a reduzir a influência de fatores emocionais nas decisões de investimento e permitir uma abordagem mais disciplinada (CAPITAL VIA, 2019). No entanto, é importante destacar que o sucesso dessas estratégias depende da qualidade dos dados, da robustez dos algoritmos e da compreensão do mercado financeiro (LIU *et al.*, 2020).

Com relação à terceira questão, como em qualquer ferramenta, a confiabilidade e robustez do *FinRL* dependem de vários fatores, incluindo a qualidade da implementação, a adequação dos algoritmos às situações do mercado e a manutenção contínua da biblioteca. É recomendável avaliar a biblioteca em relação aos requisitos específicos antes de adotá-la em cenários de negociação real (HUDSON & THAMES, 2023).

Finalmente, de acordo com a comunidade do *framework*, e olhando os demais dados postos, os resultados mostram que a estratégia de conjunto (*Ensemble*) do *FinRL* supera os algoritmos individuais, o Índice Dow Jones Industrial e o método de alocação de portfólio de variância mínima em termos de índice de Sharpe, equilibrando risco e retorno considerando custos de transação (LIU et al., 2022).

4.2 Discussão dos objetivos

Na extensa literatura sobre o assunto, percebemos que há uma profunda exploração do uso da DRL no mundo das ações. Esse estudo vai desde o gerenciamento de portfólios até o desenvolvimento de técnicas de negociação automatizada. Dentro desse universo, não podemos deixar de mencionar ferramentas e bibliotecas como *TensorFlow*, *PyTorch*, *Stable Baselines3* e, especialmente, o *FinRL*, que é parte integrante do tema deste estudo. Elas têm desempenhado um papel fundamental na otimização das estratégias de DRL no setor.

É notório que o texto não mergulha profundamente nos riscos, mas os não-leigos em relação ao assunto reconhecem o enorme desafio que a volatilidade dos mercados financeiros apresenta. Vale destacar como fatores externos podem alterar rapidamente as condições de mercado, mostrando a importância de se ter modelos de DRL que possam se adaptar a diversos cenários. Assim, é crucial reconhecer e avaliar esses riscos para garantir uma aplicação bem-sucedida da DRL no ambiente acionário.

Ao nos debruçarmos sobre a documentação em artigos do *FinRL*, é possível identificar uma descrição detalhada de sua estrutura, design e aplicabilidade - componentes-chave para avaliar sua confiabilidade. A eficácia do *FinRL* não se limita apenas ao seu funcionamento, mas também está atrelada à qualidade dos dados e à solidez dos algoritmos utilizados. A capacidade do *FinRL* de simular ambientes de mercado, aproveitando uma variedade de dados e técnicas, mostra sua competência em preparar agentes de negociação para o mercado real. Os fatos expostos anteriormente mostram que o *framework* aberto tem um futuro brilhante pela frente.

De outro modo, porém, a documentação não-acadêmica é ausente em certas ocasiões, forçando o usuário a ler os códigos disponíveis no GitHub para obter um conhecimento mais abrangente do que está acontecendo ao rodar o algoritmo. Adicionalmente, o programa não é tão bem comentado, pois notam-se várias linhas sem notas, e a engenharia de software, no geral, é apenas intermediária, talvez por ser um projeto aberto e que está em crescimento. Apesar de tudo isto, o funcionamento não é prejudicado.

Por fim, a metodologia apresentada está em total sintonia com os objetivos do estudo. Esta pesquisa, tanto através da análise empírica quanto da revisão literária, oferece uma perspectiva completa sobre a intersecção da aprendizagem de máquina e, particularmente, da DRL no mundo das ações. O enfoque que o presente trabalho opta por dar ao *FinRL*, bem como seu exame detalhado, reforça nossa compreensão sobre a qualidade desta biblioteca e confiabilidade do seu uso no mercado financeiro. Algumas lacunas, porém, precisam ser preenchidas.

4.3 Trabalhos futuros

O *FinRL* é apenas um *framework* num vasto conjunto de bibliotecas a serem exploradas, logo novas pesquisas exploratórias podem ser realizadas com cada uma das bibliotecas citadas neste trabalho. Todas elas possuem documentação acadêmica extensa, e, em sua maioria, são abertas, de modo que o código também se encontra em seus respectivos repositórios *Git*.

Suplementarmente, o *FinRL* encontra-se em crescimento e ainda tem algumas questões táticas pendentes, logo trabalhos práticos, envolvendo criação de códigos, documentação de programas e aprimoramento da engenharia de software são, sob o ponto de vista deste trabalho, bem vindos. A comunidade em torno desta biblioteca é ativa, empenhada e competente.

Por fim, uma temática a ser explorada seria a determinação (ou não) de limites em relação ao uso ostensivo de inteligência artificial nos mercados financeiros, visto que isto poderia ocasionar o surgimento de novos riscos.

5 CONCLUSÃO

O *FinRL* é uma biblioteca destacada pelo seu design em camadas, módulos independentes em cada estrato e por sua abordagem aplicável, provida de tutoriais e configurações intuitivas. Esta ferramenta se alinha ao crescente foco no uso de aprendizagem por reforço profundo no setor financeiro. Observa-se que, à medida que os dados financeiros em tempo real se tornam mais acessíveis e as técnicas de IA avançam, a DRL é posicionada como uma força potencial na automatização e otimização das estratégias de negociação. No entanto, enquanto ferramentas como o *FinRL* podem ajudar a identificar e explorar oportunidades de mercado, sua eficácia depende da qualidade dos dados, robustez dos algoritmos e compreensão profunda do mercado. Além disso, é imperativo considerar os desafios éticos e regulatórios associados à ampla adoção dessas tecnologias.

Quando se analisa o *FinRL* especificamente, nota-se a meticulosidade da documentação acadêmica, descrevendo sua estrutura, design e aplicabilidade, o que é crucial para avaliar sua confiabilidade. Por outro lado, a documentação não-acadêmica, em algumas instâncias, carece de informações detalhadas, o que pode levar os usuários a consultar diretamente os códigos no GitHub. Embora o fato de que o projeto apresenta certas limitações, como a falta de comentários em algumas linhas de código, a funcionalidade geral do *FinRL* permanece intacta. Há, ainda, questionamentos em relação à própria performance do *framework*, quando são realizados certos experimentos, e ao suposto uso desenfreado de *machine learning* no mercado financeiro, que poderia ocasionar episódios de “comportamento de manada” e riscos de segurança cibernética. O *FinRL* seria, em conjunto com tantos outros *frameworks*, mais um ator inserido neste contexto.

Concluindo, o estudo enfatiza a relevância do *FinRL* no panorama atual da DRL e suas aplicações no setor acionário. A abordagem investigativa adotada oferece uma visão abrangente da integração da aprendizagem de máquina no mercado financeiro, pontuando tanto as potencialidades quanto os desafios da ferramenta. Entretanto, com o crescimento contínuo do *FinRL* e a ativa comunidade por trás dele, há espaço para aprimoramentos e investigações futuras, tanto em aspectos experimentais, como exploratórios.

6 REFERÊNCIAS

ABADI, M. et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2016. Disponível em: <<https://arxiv.org/abs/1603.04467>>. Acesso em: 05 set. 2023.

AI4FINANCE FOUNDATION. FINRL: Financial reinforcement learning. Disponível em: <<https://github.com/AI4Finance-Foundation/FinRL>>. 2021. Acesso em: 26 ago. 2023.

AI4FINANCE FOUNDATION. FINRL: Master branch. Disponível em: <<https://github.com/AI4Finance-Foundation/FinRL/tree/master>>. 2023. Acesso em: 06 set. 2023.

ALDRIDGE, Irene. High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems. 2nd ed. Wiley, 2013.

AN, B.; SUN, S.; WANG, R. Deep Reinforcement Learning for Quantitative Trading: Challenges and opportunities. IEEE Intelligent Systems, v. 37, n. 2, p. 23–26, 15 set. 2022.

ASHFORD, K. What is market volatility - and how should you manage it? Disponível em: <<https://www.forbes.com/advisor/investing/what-is-volatility/>>. Acesso em: 05 set. 2023.

BROCKMAN, Greg et al. OpenAI Gym. OpenAI, 2016. Disponível em: <<https://arxiv.org/abs/1606.01540>>. Acesso em: 05 set. 2023.

CAPITAL VIA. “Algo trading” – taking emotions out of the decision making process. Disponível em: <<https://www.capitalvia.com/blog/algo-trading-taking-emotions-out-of-the-decision-making-process>>. Acesso em: 06 set. 2023.

CASTRO, P. S.; BELLEMARE, M. G. Dopamine 2.0: Providing more flexibility in reinforcement learning research. Disponível em:

<<https://opensource.googleblog.com/2019/02/dopamine-2.0.html>> ;. Acesso em: 06 set. 2023.

D'AGOSTINO, R. B.; BELANGER, A.; D'AGOSTINO, R. B. A suggestion for using powerful and informative tests of normality. *The American Statistician*, v. 44, n. 4, p. 316, 1990.

DENG, Y. et al. Deep direct reinforcement learning for financial signal representation and trading. *IEEE transactions on neural networks and learning systems*, [S.l.], v. 28, n. 3, p. 653–664, mar. 2017. ISSN 2162-237X. Disponível em: <<https://ieeexplore.ieee.org/document/7407387>>. Acesso em: 05 set. 2023.

DONG, Hao et al. TensorLayer: A Versatile Library for Efficient Deep Learning Development. *ACM Multimedia*, 2017. Disponível em: <<https://arxiv.org/abs/1707.08551>>. Acesso em: 05 set. 2023.

HUDSON & THAMES. FinRL - Reinforcement Learning in Finance; Hudson & Thames, 05 jun. 2023. Disponível em: <<https://www.youtube.com/watch?v=8464ezZWJrE>>. Acesso em: 26 ago. 2023.

HUANG, G.; ZHOU, X.; SONG, Q. Deep Reinforcement Learning for Portfolio Management. Disponível em: <<https://arxiv.org/ftp/arxiv/papers/2012/2012.13773.pdf>>. Acesso em: 05 set. 2023.

IBM, I. O que é a ia forte? Disponível em: <<https://www.ibm.com/br-pt/topics/strong-ai>>. Acesso em: 03 set. 2023.

INFOMONEY. Análise Técnica de ações: aprenda a interpretar gráficos e índices. Disponível em: <<https://www.infomoney.com.br/guias/analise-tecnica/>>. Acesso em: 05 set. 2023.

JAGER, R. A. Evaluation and potential improvements of a deep reinforcement learning model for Automated Stock Trading. *SSRN Electronic Journal*, 2021.

JIANG, Z. et al. Financial Reinforcement Learning: Application, Comparison and Tools. Disponível em: <<https://arxiv.org/abs/2011.09607>>. Acesso em 03 set. 2023.

JONES, M. Train a software agent to behave rationally with reinforcement learning. Disponível em: <<https://developer.ibm.com/articles/cc-reinforcement-learning-train-software-agent/>>. Acesso em: 05 set. 2023.

KABBANI, T.; DUMAN, E. Deep Reinforcement Learning Approach for trading automation in the stock market. IEEE Access, v. 10, p. 93564–93574, 05 jul. 2022.

KUEPPER, J. *CBOE Volatility Index (VIX): What does it measure in investing?* Disponível em: <<https://www.investopedia.com/terms/v/vix.asp#:~:text=Key%20Takeaways,market%20when%20making%20investment%20decisions.>>. Acesso em: 08 set. 2023.

LIANG, E. et al. RLib: Abstractions for Distributed Reinforcement Learning. In: International Conference on Machine Learning (ICML), 2018. p. 10. Disponível em: <<https://doi.org/10.48550/arXiv.1712.09381>>. Acesso em: 05 set. 2023.

LILLICRAP, T. P. et al. Continuous control with deep reinforcement learning. In: International conference on learning representations, 4., 2016, San Juan. Proceedings. San Juan: ICLR, 2016. Disponível em: <<https://arxiv.org/abs/1509.02971>>. Acesso em 03 set. 2023.

LIU, X. Y. et al. FinRL-Meta: Market Environments and Benchmarks for Data-Driven Financial Reinforcement Learning. NeurIPS, 2022. Disponível em: <<https://arxiv.org/abs/2211.03107>>. Acesso em: 05 set. 2023.

LIU, X. Y. et al. FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. Deep RL Workshop, NeurIPS 2020, 2020. Disponível em: <<https://arxiv.org/abs/2011.09607>>. Acesso em: 05 set. 2023.

LIU, X. Y. et al. FinRL: Deep reinforcement learning framework to automate trading in quantitative finance. ACM International Conference on AI in Finance (ICAIF), 2021. Disponível em: <<https://arxiv.org/abs/2111.09395>>. Acesso em: 05 set. 2023.

LIU, X. Y. et al. Practical deep reinforcement learning approach for stock trading. NeurIPS Workshop on Deep Reinforcement Learning, 2018. Disponível em: <<https://arxiv.org/abs/1811.07522>>. Acesso em: 05 set. 2023.

LUGER, G. F. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Pearson, 2019.

MATIISEN, T. Demystifying Deep Reinforcement Learning. neuro.cs.ut.ee. Computational Neuroscience Lab, 19 dez. 2015. Disponível em: <<https://neuro.cs.ut.ee/demystifying-deep-reinforcement-learning/>>. Acesso em: 06 set. 2023.

MENKVELD, A. J.; QI, T. Dynamic Limit Order Market with Fast and Slow Traders. Review of Financial Studies, v. 28, n. 12, p. 3407-3440, 2015.

MNIH, V. et al. Asynchronous methods for deep reinforcement learning. In: International conference on machine learning, 33., 2016, New York. Proceedings. New York: PMLR, 2016. p. 1928-1937. Disponível em: <<https://arxiv.org/abs/1602.01783>>. Acesso em: 03 set. 2023.

OECD. Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers. 2021. Disponível em <<https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf>>. Acesso em: 08 set. 2023.

PASZKE, A. et al. PyTorch: An imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems, v. 32, 2019. Disponível em: <<https://arxiv.org/abs/1912.01703>>. Acesso em: 05 set. 2023.

PINTO, L. O que É índice Sharpe? Saiba Como Comparar Risco e retorno nos fundos de Investimento. Disponível em: <<https://conteudos.xpi.com.br/aprenda-a-investir/relatorios/o-que-e-indice-sharpe-sai-ba-como-comparar-risco-e-retorno-nos-fundos-de-investimento/>>. Acesso em: 05 set. 2023.

PUTERMAN, M. L. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, 2014.

RAFFIN, A. et al. Stable Baselines3. GitHub, 2020. Disponível em: <<https://github.com/DLR-RM/stable-baselines3>>. Acesso em: 05 set. 2023.

RAO, A.; JELVIS, T. Foundations of Reinforcement Learning with applications in Finance. Boca Raton: Chapman & Hall, CRC Press, Taylor & Francis Group, 2023.

RUNDO, F. et al. Advanced markov-based machine learning framework for making Adaptive Trading System. Computation, v. 7, n. 1, p. 4, 2019.

RUSSELL, S.; NORVIG, P. Inteligência Artificial: Tradução da 3ª edição. São Paulo: Pearson, 2016.

SCHULMAN, J. et al. Proximal Policy Optimization Algorithms, 28 ago. 2017. OpenAI. Disponível em: <<https://arxiv.org/pdf/1707.06347.pdf>>. Acesso em: 03 set. 2023.

SILVER, D. Introduction to reinforcement learning with David Silver. Disponível em: <<https://www.deepmind.com/learning-resources/introduction-to-reinforcement-learning-with-david-silver>>. Acesso em: 02 set. 2023.

SINGH, S. A comprehensive guide to neural networks in Deep Q-Learning. Disponível em: <<https://www.turing.com/kb/how-are-neural-networks-used-in-deep-q-learning>>. Acesso em: 03 set. 2023.

SUTTON, R. S.; BARTO, A. G. Reinforcement Learning: An Introduction. The MIT Press, 2018.

THE TIMES OF INDIA. OpenAI, the company behind ChatGPT: What all it does, how it started and more. The Times of India. Disponível em: <<https://timesofindia.indiatimes.com/gadgets-news/openai-the-company-behind-chat-gpt-what-all-it-does-how-it-started-and-more/articleshow/97297027>>. Acesso em: 28 de agosto de 2023.

THEODORSON, G. A. & THEODORSON, A. G. A modern dictionary of sociology. London, Methuen, 1970.

WU, Y. et al. Using deep reinforcement learning to allocate resources in a system of network slices. Disponível em: <<https://wu-lab.exeter.ac.uk/summaries/2019-drl/>>. Acesso em: 03 set. 2023.

YANG, H. et al. Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy. SSRN, 11 set. 2020. Disponível em: <<https://dx.doi.org/10.2139/ssrn.3690996>>. Acesso em: 05 set. 2023.

ZHANG, Z.; ZOHREN, S.; ROBERTS, S. Deep learning for portfolio optimization. The Journal of Financial Data Science, v. 2, n. 4, p. 8–20, 2020.