

**ESTRATÉGIAS DE INVESTIMENTO BASEADAS
EM MICROESTRUTURA DE MERCADO**

ÁLEF WILLIS MAGNO MIRANDA

**ESTRATÉGIAS DE INVESTIMENTO BASEADAS
EM MICROESTRUTURA DE MERCADO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais – Departamento de Ciência da Computação como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADOR: ADRIANO CÉSAR MACHADO PEREIRA

Belo Horizonte

Junho de 2019

© 2019, Álef Willis Magno Miranda.
Todos os direitos reservados.

Miranda, Álef Willis Magno

M672e Estratégias de investimento baseadas em
microestrutura de mercado / Álef Willis Magno
Miranda. — Belo Horizonte, 2019
xx, 55 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de
Minas Gerais – Departamento de Ciência da
Computação

Orientador: Adriano César Machado Pereira

1. Computação – Teses. 2. Inteligência Artificial.
3. Mercado Financeiro. 4. Bolsa de valores.
5. Modelagem de Dados. I. Orientador. II. Título.

CDU 519.6*82(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

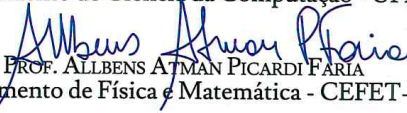
Estratégias de investimento baseadas em microestrutura de mercado

ALEF WILLIS MAGNO MIRANDA

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:


PROF. ADRIANO CÉSAR MACHADO PEREIRA - Orientador
Departamento de Ciência da Computação - UFMG


PROF. ANÍSIO MENDES LACERDA
Departamento de Ciência da Computação - UFMG


PROF. ALBENS ATMAN PICARDI FÁRIA
Departamento de Física e Matemática - CEFET-MG

Belo Horizonte, 3 de Julho de 2019.

Dedico esse trabalho a meus pais, por sempre estarem ao meu lado durante todos os desafios de minha vida, ao meu grande amor, por me dar energia e força de vontade nos momentos mais difíceis e a meus professores, que trilharam todo o caminho do conhecimento junto a mim...

“Nada é como deveria ser.”
(Autor desconhecido)

Resumo

O uso de agentes autônomos de negociação no mercado financeiro se torna cada vez mais comum, apesar de no Brasil ainda ser muito pouco explorado. Com um grande volume de dados financeiros disponíveis, torna-se viável a construção de novos modelos de negociação. O objetivo deste trabalho é propor novos indicadores baseados em microestrutura de mercado para criar novas estratégias de investimento automatizadas. Para alcançar esse objetivo, constrói-se um simulador de bolsa de valores, um conjunto de indicadores baseados em microestrutura de mercado, uma técnica de rotulação de séries de preço e, por fim, um agente autônomo de negociação. A princípio, apresenta-se um simulador capaz de reproduzir as ordens e negociações enviadas à bolsa de valores dos dias de negociação passados. Em seguida, derivam-se indicadores financeiros baseados em aspectos de microestrutura de mercado e os princípios necessários para produzir rotulação de séries de preço. Por fim, construiu-se um agente autônomo que toma decisões baseando-se em tais indicadores e realiza-se a validação experimental, na qual avaliam-se métricas financeiras do agente. A base de dados escolhida para a validação experimental consiste dos dados de negociação de 2018 do minicontrato futuro de dólar da B3. Os resultados financeiros obtidos pelo agente autônomo são avaliados em diferentes cenários considerando-se custos de operação e latências de rede. Tais resultados financeiros revelam o potencial de se utilizar dados de microestrutura de mercado para construção de estratégias automatizadas de investimento.

Palavras-chave: Mercado financeiro, Microestrutura de mercado, Livro de ofertas, Simulação realística, Agentes autônomos, Estratégias de Investimento, Modelagem de indicadores

Abstract

The use of autonomous trading agents in stock markets is becoming more common over time, although it is yet little explored in Brazil. With the large amount of available financial data, the construction of new trading models is feasible. The main objective of this work is to propose new financial indicators based on market micro-structure in order to create new automated investment strategies. To achieve this goal, a stock exchange simulator is built, a set of financial indicators based on market micro-structure and the principles for labelling prices series are defined and, finally, an autonomous trading agent is constructed. Firstly, the simulator, capable of reproduction of the orders and trades sent to the stock exchange of past trading days, is presented. Secondly, the set of financial indicators based on market micro-structure aspects and the necessary principles to label price series are defined. Lastly, the autonomous agent based on such indicators is built and, then, experimental validation is performed analysing financial metrics of the agent. In order to perform the experimental validation, data from future contracts of dollar of 2018 from B3 are chosen. The financial results obtained from the agent are evaluated in multiple scenarios varying parameters of network latency and operational costs. Such financial results show the potential of using micro-structure market data to construct automated investment strategies.

Keywords: Stock Market, Market micro-structure, order book, Realistic simulation, Autonomous agents, Investment Strategy, Indicators modelling

Lista de Figuras

2.1	Período da bolsa aberta e suas divisões	7
2.2	Representação de um livro de ofertas.	9
4.1	Metodologia e seus detalhes.	21
4.2	Etapas do tratamento de dados.	24
4.3	Representação do funcionamento do simulador.	25
4.4	Série de preço rotulada do mini-dólar futuro no dia 02 de Março de 2018. .	30
4.5	<i>Spread</i> e do <i>mid-price</i>	32
4.6	Compras e vendas no período de negociação	33
4.7	Indicadores de posição e desbalanceamento.	34
4.8	Indicador de velocidade	35
4.9	Indicadores de densidade e profundidade	36
4.10	Processo de avaliação da estratégia de investimento.	37
5.1	Capital acumulado em 01/03/2018	43
5.2	Capital acumulado em 02/03/2018	43
5.3	Capital acumulado em 05/03/2018	44
5.4	Capital diário e acumulado entre os meses de Março e Dezembro de 2018. .	44
5.5	Capital acumulado em 01/03/2018 em diferentes latências	46
5.6	Capital acumulado em 02/03/2018 em diferentes latências	46
5.7	Capital acumulado em 05/03/2018 em diferentes latências	47
5.8	Capital acumulado de Março a Dezembro de 2018 em diferentes latências. .	47

Lista de Tabelas

3.1	Relação dos trabalhos	19
5.1	Matriz de confusão	42
5.2	Resultados do cenário de custo zero e latência nula.	44
5.3	Resultados do cenário de custo positivo e latência nula.	45
5.4	Resultados sumarizados	48
5.5	Avaliação geral	48

Sumário

Resumo	xi
Abstract	xiii
Lista de Figuras	xv
Lista de Tabelas	xvii
1 Introdução	1
1.1 Visão geral	1
1.2 Motivação	2
1.3 Objetivos	2
1.4 Contribuições	3
1.5 Organização do documento	4
2 Fundamentos conceituais	5
2.1 Regras de negociação	5
2.1.1 Características das ordens	5
2.1.2 Posicionamento e execução das ordens	6
2.1.3 Definição de preço inicial	7
2.2 Livro de Ofertas	8
2.2.1 Estrutura	8
2.2.2 Formação de preço	10
3 Trabalhos relacionados	11
4 Metodologia	21
4.1 Simulador	22
4.1.1 Tratamento de dados	22
4.1.2 Simulação	25

4.1.3	Notação	27
4.2	Rotulação das séries de preço	29
4.3	Indicadores	30
4.3.1	<i>Spread</i> e <i>Mid-price</i>	30
4.3.2	Compras, vendas e volume	32
4.3.3	Posição e desbalanceamento	32
4.3.4	Velocidade	34
4.3.5	Densidade e profundidade	35
4.4	Construção e validação da estratégia	37
4.4.1	Dados	37
4.4.2	Modelo de classificação	38
4.4.3	Estratégia	39
5	Experimentos: Resultados e análise	41
5.1	Dados	41
5.2	Modelo de classificação	42
5.3	Modelo de negociação	42
5.3.1	Custo zero, latência nula	43
5.3.2	Custo positivo, latência nula	45
5.3.3	Custo zero, latência positiva	45
5.3.4	Avaliação geral	47
6	Conclusão	49
	Referências Bibliográficas	51

Capítulo 1

Introdução

Esse capítulo apresenta uma visão geral sobre mercados financeiros, uso de agentes autônomos e microestrutura de mercado. Em seguida, são apresentadas as motivações e os objetivos deste trabalho. Por fim, apresentam-se a contribuição gerada e a organização do trabalho.

1.1 Visão geral

O mercado financeiro centraliza a negociação de diversos tipos de papéis ao redor do mundo, como ações e derivativos. Ele permite que investidores comprem e vendam esses papéis de acordo com suas preferências e necessidades. O preço desses papéis é definido, genericamente, pelas leis de oferta e demanda Baumol [2007].

O mercado financeiro é composto de forma geral por bolsas de valores, corretoras e investidores, sejam eles empresas ou pessoas físicas. No Brasil, a *Brasil Bolsa Balcão* B3 [2019a], resultado da fusão entre *BMF&Bovespa* e *Cetip* B3 [2019b], é responsável pela negociação dos papéis de empresas brasileiras e de outros derivativos. A legislação do mercado financeiro brasileiro é definida pela *Comissão de Valores Mobiliários* CVM [2019], que fica responsável por punir infratores, como usuários de informação privilegiada Bhattacharya & Daouk [2002].

O uso de agentes autônomos Russell & Norvig [2016] no mercado financeiro tem se tornado bem comum. Os agentes são capazes de tomar decisões puramente analíticas de alta precisão e sem interferência externa. Tais decisões podem ser modeladas por especialistas de mercado ou mesmo serem derivadas a partir de técnicas de aprendizado de máquina. Existem diferentes tipos de agente considerando-se sua frequência de operação e quantidade de papéis distintos que operam, mas todos buscam maximizar o lucro de seus operadores.

Ao se falar em *microestrutura de mercado* O'hara [1995], refere-se aos detalhes operacionais das negociações realizadas em determinado papel. No contexto de bolsa de valores, trata-se das regras que regem o envio de ordens e realização das negociações desses papéis. Essas regras são materializadas em um livro de ofertas, responsável por controlar ofertas de compra e venda e realizar transações quando os preços estão de acordo.

Esse trabalho aplica os conceitos de microestrutura de mercado para construção de novos indicadores financeiros. Esse indicadores são utilizados para a construção de um agente autônomo de investimento utilizando-se aprendizado de máquina Bishop [2006]. Por fim, avalia-se seu desempenho financeiro em um simulador de bolsa de valores.

1.2 Motivação

As séries de preço de determinado papel são produto de um sistema complexo de diversos agentes. Dessa forma, a previsão de tendências de mercado se apresenta como um grande desafio, o que gera oportunidades para aplicações de técnicas computacionais como aprendizado de máquina e reconhecimento de padrões.

A *B3* disponibiliza abertamente todos os dados referentes a microestrutura de mercado de todos seus papéis. A partir disso, é viável a reconstrução dos cenários de cada dia de mercado e a análise precisa desses dados.

Microestrutura de mercado pode representar uma nova dimensão nesse tipo de estudo, trazendo mais informação sobre o mercado que a série de preços e volume. Dessa forma, há grande possibilidade de melhorar os resultados dos trabalhos que utilizam apenas séries de preço e suas derivações.

1.3 Objetivos

O objetivo geral deste trabalho é propor novos indicadores financeiros baseados em microestrutura de mercado para criar novas estratégias de investimento automatizadas. Dessa forma, os objetivos específicos desse trabalho podem ser desdobrados em:

- **Derivação de indicadores:** Utilizando-se o livro de ofertas de base, deriva-se novos indicadores que buscam encapsular informações sobre as tendências de mercado;

- **Rotulação de séries de preço:** Decidir os melhores momentos para se investir é um problema em aberto. Dessa forma, utiliza-se um algoritmo de rotulação para indicar se determinadas séries de preço estão em tendência de crescimento ou decréscimo. Dessa forma, apresentam-se as características necessárias para se produzir séries de preço bem rotuladas. Essa rotulação é de suma importância para a produção dos modelos de classificação de tendências de preço;
- **Construção de um modelo de classificação de tendência:** A partir dos indicadores derivados e uma série de preços bem rotulada, produz-se um modelo de classificação de tendências de preço utilizando-se técnicas de aprendizado de máquina;
- **Construção de um modelo de negociação:** Uma estratégia de investimento é construída utilizando-se de entrada o modelo de classificação de tendência. Esse estratégia é implementada por meio de um agente autônomo de investimento buscando lucrar em suas operações financeiras.

1.4 Contribuições

Esse projeto realiza um estudo detalhado do processo de negociação de papéis nas bolsas de valores. Uma abordagem baseada em microestrutura de mercado abre novas possibilidades para construção de agentes autônomos com maior performance financeira. De forma específica, esse projeto traz as seguintes contribuições:

- **Simulador de bolsa de valores:** O projeto produz um simulador capaz de reprodução das ordens e negociações da bolsa de valores de dias passados;
- **Novos indicadores:** Os indicadores extraídos a partir do livro de ofertas são úteis para qualquer outro trabalho na área, possibilitando seu uso em diferentes estratégias de investimento;
- **Agente de negociação:** Um agente autônomo de investimento é construído e avaliado em diversos cenários que aproximam a realidade. Esse agente é passível de aplicação real desde que seus requisitos sejam atendidos.

1.5 Organização do documento

As próximas seções estão divididas da seguinte forma:

- **(Cap. 2) Fundamentos conceituais:** Apresenta-se toda a base teórica de mercado financeiro necessária para a execução do projeto;
- **(Cap. 3) Trabalhos relacionados:** Apresentam-se trabalhos relacionados a mercado financeiro, aprendizado de máquina e agentes autônomos.
- **(Cap. 4) Metodologia:** Neste capítulo, mostra-se o processo de construção e aplicação do simulador da bolsa de valores, princípios para rotulação de séries de preço, a construção dos indicadores e a estratégia de investimento construída, junto de seu processo de validação experimental;
- **(Cap. 5) Experimentos: Resultados e análise:** Descrevem-se os dados utilizados, a performance dos modelos de classificação e da estratégia de investimento. Por fim, uma avaliação financeira detalhada, considerando-se diferentes cenários, é apresentada.

Capítulo 2

Fundamentos conceituais

Ao longo deste trabalho, diversos termos e conceitos relacionados ao mercado financeiro são utilizados. Desta forma, busca-se descrever cada um deles e detalha-se suas características mais relevantes. Todos eles compõem uma área de finanças denominada *microestrutura de mercado*.

O termo *Microestrutura de mercado* pode ser definido como ‘o estudo dos processos e desfechos da negociação de bens sob regras de negociação explícitas’ O’hara [1995]. O termo refere-se aos mecanismos operacionais de negociação de determinados bens e como esses mecanismos afetam o processo de formação de preço.

Considerando-se o contexto de mercado financeiro, pode-se dizer que microestrutura de mercado se refere às *regras* de negociações de *papéis*, sejam eles ações ou contratos futuros. Essas regras podem então ser materializadas na construção de um *livro de ofertas*.

2.1 Regras de negociação

O processo de negociação de papéis deve ser justo e implicitamente obedecer às *regras de oferta e demanda*. Aplicando-se tais requisitos, garante-se empiricamente que o preço de um determinado papel é seu preço *real*. São apresentadas a seguir as características das ordens e como elas são negociadas.

2.1.1 Características das ordens

As *ordens* representam a vontade dos investidores em relação a um papel. Elas possuem dimensões suficientes para descrever as condições nas quais elas devem ser executadas. As dimensões mais relevantes são descritas a seguir:

- **Sentido:** Uma ordem pode ter seu sentido descrito como *compra* ou *venda*. Ordens de compra informam o interesse de se obter determinado papel, enquanto que ordens de venda informam o interesse de se vender um papel em sua posse. Existem casos de venda sem de fato possuir o papel em questão, isso é descrito como *venda a descoberto*, onde busca-se lucro esperando-se pela desvalorização de determinado papel Harris [2003];
- **Volume:** A quantidade esperada de papéis a ser negociada. Uma ordem pode ser parcialmente negociada, isto é, apenas uma fração de seu tamanho é de fato executada;
- **Momento:** Horário em que a ordem foi recebida pela bolsa. O horário tem grande importância para determinar a prioridade das ordens entre si;
- **Classe:** Uma ordem pode ser definida em duas classes mais básicas, *ordem limite* e *ordem a mercado*.

Uma **ordem limite** pode definir o preço no qual determinado papel deve ser negociado. Dessa forma, a ordem é apenas executada no preço previamente definido. Por outro lado, não há garantias de que a ordem será executada, dependendo-se apenas da oferta e demanda do papel.

Já uma **ordem a mercado** não define preço. É uma ordem que será executada no momento de sua recepção pela bolsa. Uma ordem a mercado aceita o preço vigente do papel como seu preço de execução;

- **Preço** (opcional): Preço pelo qual o papel será negociado. Sua definição é condicionada à classe da ordem.

2.1.2 Posicionamento e execução das ordens

Com as dimensões definidas, pode-se descrever como o posicionamento das ordens é realizado. Esse posicionamento descreve a prioridade das ordens entre si de acordo com tais dimensões. Uma ordem de maior prioridade é mais provável de ser executada que uma ordem de menor prioridade.

Quando as ordens são **criadas**, as regras que determinam suas prioridades são as seguintes:

1. Para ordens de compra, aquelas de maior preço possuem prioridade sobre aquelas de menor preço.

Para ordens de venda, aquelas de menor preço possuem prioridade sobre aquelas de maior preço;

2. Para ordens de mesmo preço, as mais antigas possuem prioridade sobre as mais recentes.

Quando ordens são **modificadas**, há mudança na prioridade das mesmas. Em geral, ordens modificadas são tratadas como novas ordens, cancelando-se suas versões anteriores. A prioridade dessas ordens modificadas segue as regras já descritas.

Ordens **canceladas** saem de posicionamento. Mesmo que se tratem de ordens parcialmente executadas, elas podem ser canceladas, mas suas negociações anteriores não são revertidas.

Ordens são **executadas**, completamente ou parcialmente, quando existe concordância entre os preços de compra e venda entre elas. Caso uma delas possua um tamanho maior que a outra, essa primeira será apenas parcialmente executada.

2.1.3 Definição de preço inicial

Diversas bolsas de valores ao redor do mundo não estão abertas o tempo todo. Isto é, elas impõem períodos bem definidos durante o dia para a negociação dos papéis que gerenciam. Esse é o caso da bolsa de valores brasileira, *Brasil, Bolsa, Balcão*, ou simplesmente, B3 [2019a].

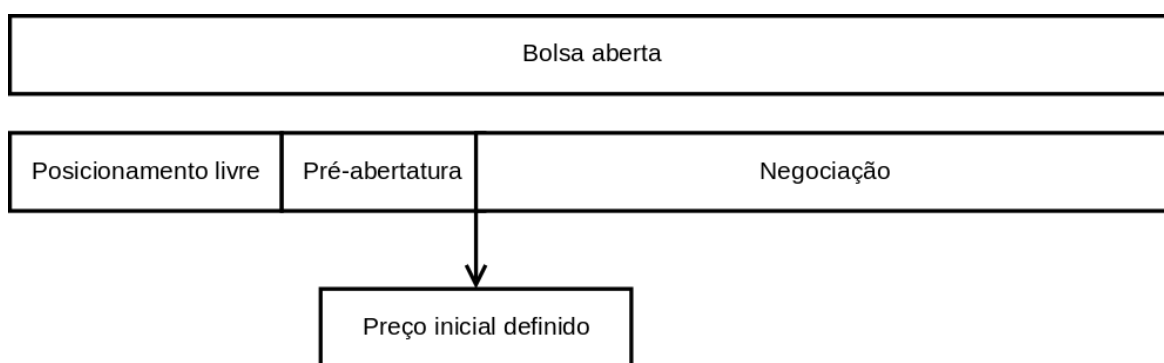


Figura 2.1. Período da bolsa aberta e suas divisões

De forma simplificada, pode-se dividir o período em que a bolsa de valores está aberta em três fases, *posicionamento livre*, *pré-abertura* e *negociação*, assim como mostrado na Figura 2.1. Essas fases tem as seguintes características:

1. **Posicionamento livre:** Durante esse período, as ordens podem ser enviadas e canceladas de forma livre, mas nenhuma delas é executada;

2. **Pré-abertura:** Já nesse período, permite-se apenas o posicionamento de ordens, e as mesmas já não podem mais ser canceladas. Elas ainda não serão executadas;
3. **Negociação:** As ordens são posicionadas livremente e são executadas de acordo com oferta e demanda.

Existem outras fases durante o período de negociação que podem variar de acordo com as necessidades de cada bolsa de valor. No caso da *B3*, existem fases de *pós-negociação* B3 [2019d], que acontecem ao fim do período da bolsa aberta, mas não tratadas nesse trabalho.

Como exibido na Figura 2.1, a definição do preço inicial acontece na abertura do mercado, entre as fases de pré-abertura e negociação. O mecanismo utilizado para esse fim é baseado nas características do *preço de equilíbrio* Varian [1992]. O preço escolhido como preço inicial de determinado papel deve ser aquele que atende aos seguintes requisitos, ordenados do mais significativo aos menos significativo:

- Executa a maior quantidade possível de ordens posicionadas;
- Deixa de executar o mínimo de ordens possíveis, em caso de empate;
- É o mais próximo do preço final do período de negociação anterior, no caso de empate.

Ao término da fase de pré-abertura, um algoritmo é executado sobre as ordens posicionadas e então define-se o preço inicial do papel em questão. Todas as ordens executadas são executadas nesse preço.

2.2 Livro de Ofertas

Com as regras de negociação definidas, é possível implementar o sistema capaz de executá-las. Esse sistema é o *livro de ofertas eletrônico*. Por meio dele, centraliza-se o recebimento de todas ordens de determinado papel e garante-se a execução justa dessas ordens. A Figura 2.2 exemplifica um livro de ofertas com algumas ordens posicionadas.

2.2.1 Estrutura

O livro de ofertas organiza as *ordens limite* de acordo com suas propriedades. Essas propriedades derivam três conceitos do livro de ofertas: *lado*, *níveis de preço* e *fila de ordens*. As *ordens a mercado* são executadas imediatamente e portanto não ficam

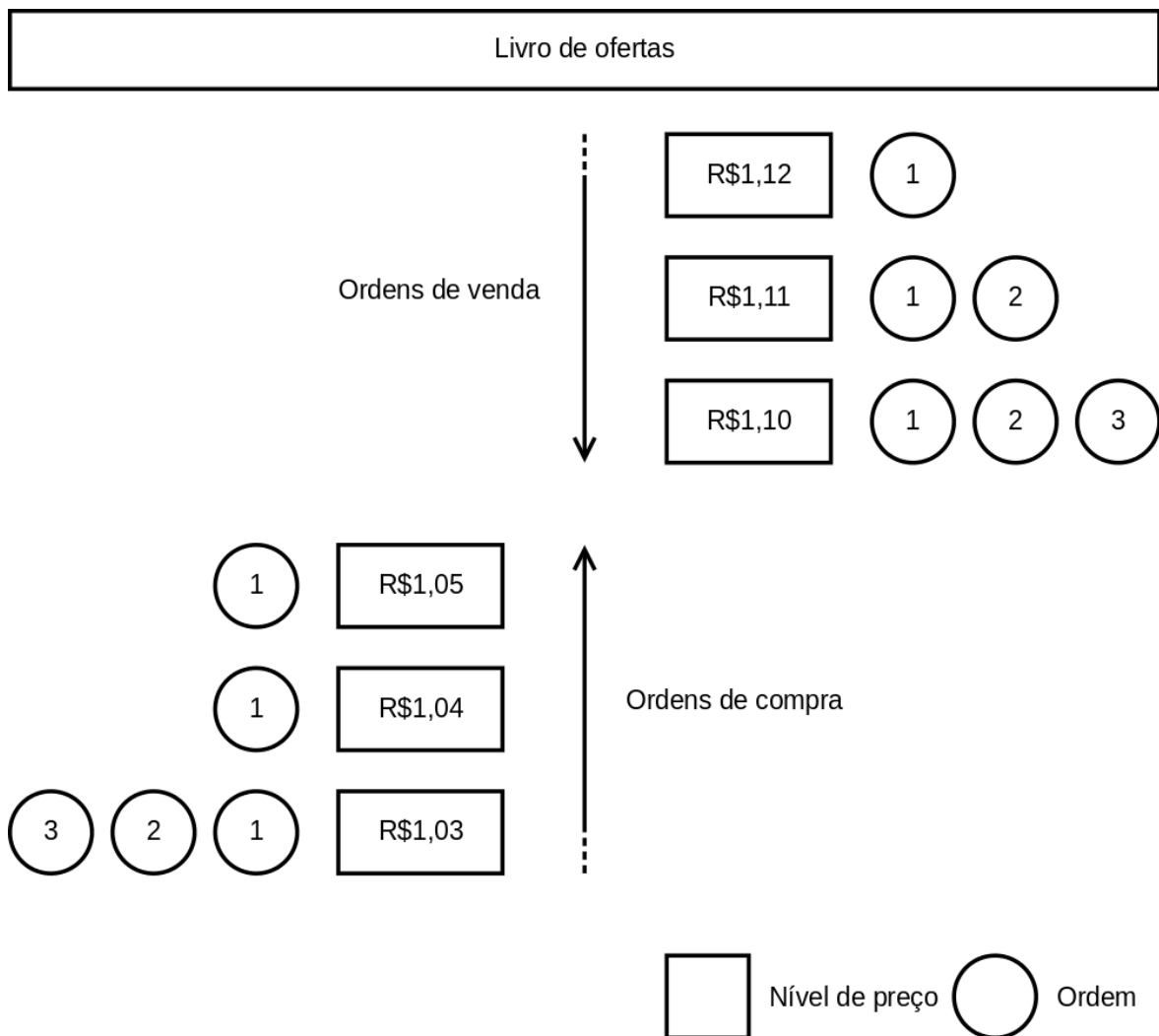


Figura 2.2. Representação de um livro de ofertas.

registradas no livro de ofertas, mas alteram seu estado, pois compõem negócios com ordens limite presentes nele.

Cada um dos conceitos possui o objetivo de agrupar as *ordens limite* de alguma forma. Esse agrupamento se dá de forma hierárquica seguindo a seguinte ordem:

1. **Lado:** O livro de ofertas divide as ordens em dois grupos de acordo com seus sentidos: o lado dos *compradores* e o lado do *vendedores*.

O lado dos compradores acumula as ordens de compra, enquanto que o lado dos vendedores acumula as ordens de venda;

2. **Nível de preço:** Todas as ordens de mesmo valor são agrupadas em um nível de preço. Os níveis de preço são ordenados dependendo do lado do livro de ofertas considerado.

Do lado dos compradores, os níveis de preço são ordenados decendentemente, de forma a priorizar ordens de compra de maior valor. Do lado dos vendedores, a ordenação é ascendente, dando-se mais prioridade às ordens de menor preço;

3. **Fila de ordens:** Considerando-se um nível de preço, as ordens são ordenadas entre si ascendente, considerando-se seu momento de registro. Dessa forma, ordens mais antigas possuem prioridade maior a ordens mais recentes.

A partir desses agrupamentos, surge o conceito de *topo do livro de ofertas*. O topo do livro de ofertas é composto pelos níveis de preço de maior prioridade de cada um dos lados do livro de ofertas. Considerando-se a Figura 2.2, o topo é composto pelos níveis $R\$1,10$ e $R\$1,05$ dos lados de vendedores e compradores respectivamente.

2.2.2 Formação de preço

O preço de determinado papel é definido como sendo o preço do último *negócio* executado naquele papel, ou seja, o negócio mais recente. Um negócio é formado por duas ordens, uma de compra e uma de venda. Em geral, dentre essas ordens, uma delas é considerada *agressora* e a outra *receptora*:

- **Ordem agressora:** É a ordem responsável pela geração do negócio. Quando se trata de uma ordem de compra, essa possui preço igual ou superior às ordens de venda do topo do livro de ofertas. Quando se trata de uma ordem de venda, essa possui preço igual ou inferior às ordens de compra do topo do livro de ofertas. Diz-se que a ordem ‘agride’ o lado de sentido oposto do livro de ofertas. Em geral, se trata de uma *ordem a mercado*;
- **Ordem receptora:** É a ordem limite de maior prioridade do livro de ofertas que casa com a ordem agressora. É de venda quando a ordem agressora é de compra, e é de compra quando a ordem agressora é de venda.

O preço do negócio é definido pelo preço da ordem receptora. Isso ocorre sem prejuízos para a ordem agressora. Isto é, uma ordem agressora de compra, com preço superior às ordens de venda do topo do livro de ofertas, não perde por pagar menos pelo papel. O mesmo ocorre para a operação no sentido inverso.

Capítulo 3

Trabalhos relacionados

Existem diversos trabalhos recentes relacionados a microestrutura de mercado, aprendizado de máquina e estratégias de investimento. Isso é consequência da disponibilização de grandes massas de dados, possibilitando simulações e reproduções realísticas dos eventos do mercado financeiro. A seguir, cada um dos principais trabalhos correlatos são brevemente descritos e, ao final, sistematizados numa tabela comparativa com este projeto de mestrado.

1. *Deep Learning can Replicate Adaptive Traders in a Limit-Order-Book Financial Market* le Calvez & Cliff [2018]: Nesse artigo os autores utilizam técnicas de *deep learning* para aprender o comportamento de investidores reais durante o horário de negociação da bolsa. O modelo construído por eles se adapta a mudanças do próprio investidor no tempo. Eles utilizam, como entrada, os dados financeiros e as decisões de compra e venda realizadas pelos investidores reais. Os resultados são uma prova de conceito de que uma *rede neuronal profunda* poderia aprender estratégias de investidores reais a partir da observação de suas decisões.
2. *Cooperative Multi-Agent Reinforcement Learning Framework for Scalping Trading*: Jo et al. [2019]: Os autores constroem um ferramental para treinar agentes a partir de *aprendizado por reforço*, buscando-se imitar as ações de investidores praticantes de *scalping* Cheng [2007]. Os dados utilizados para treinar os agentes consistem de agregações dos dados do livro de ofertas, das negociações realizadas e de dados de investidores reais.
3. *Investigating Limit Order Book Characteristics for Short Term Price Prediction: a Machine Learning Approach* Qureshi [2018]: Nesse artigo, procura-se construir preditores de preço de curto prazo para diferentes papéis. Utiliza-se, como en-

trada, os valores e quantidade de ordens de vários níveis do livro de ofertas. O primeiro resultado é que a acurácia obtida pelos modelos é sensível ao desbalançamento do livro de ofertas nos dados utilizados. Além disso, outro resultado importante é que quanto maior a quantidade utilizada de níveis do livro de ofertas, maior a acurácia dos modelos.

4. *Predicting and Forecasting the Price of Constituents and Index of Cryptocurrency Using Machine Learning* Chowdhury et al. [2019]: Nesse trabalho, utiliza-se diversos modelos para fazer regressão e previsão das séries de preço de diversas criptomoedas e de seu índice associado. O foco do trabalho é voltado à aplicação de técnicas já conhecidas a criptomoedas e avaliar o desempenho delas nesse tipo de papel.
5. *Algorithmic trading in a microstructural limit order book model* Abergel et al. [2017]: Esse trabalho trás um ferramental teórico de modelagem de microestrutura de mercado para estudar políticas ótimas para *formadores de mercado*. Esse ferramental considera ordens a limite, a mercado e cancelamentos, mantendo-se o estado do livro de ofertas ao longo do tempo. *Processos de decisão de Markov e programação dinâmica* são utilizados para encontrar essas estratégias ótimas. Por fim, as estratégias são comparadas com estratégias ingênuas de *formadores de mercado*.
6. *Market Microstructure Knowledge Needed for Controlling an Intra-Day Trading Process* Lehalle [2013]: Esse trabalho aborda o conhecimento necessário para tratar o problema de liquidar ordens de grande volume, obedecendo-se diversas regulações vigentes entre as bolsas de valores ao redor do mundo. Avalia-se o problema do ponto de vista temporal, isto é, escalonamento temporal do envio de ordens, e do ponto de vista espacial, isto é, roteamento de ordens por meio de diversos agentes intermediários, ocultando-se a origem das mesmas.
7. *Market Liquidity and Convexity of Order-Book* Qiao [2012]: Nesse artigo, constrói-se um indicador denominado *convexity* para estimar a liquidez potencial provida por um livro de ofertas. A partir dos dados reais de alta frequência do índice *SSE50* da bolsa de valores de Xangai, várias propriedades estatísticas desse indicador são avaliadas. Além disso, avalia-se a relação do indicador com outros indicadores do livro de ofertas.
8. *The Impact of Heterogeneous Trading Rules on the Limit Order Book and Order Flows* Chiarella et al. [2009]: Nesse artigo, um modelo orientado a ordens, onde

investidores definem ordens de compra e venda, com regras previamente definidas, é construído. Os investidores são definidos como agentes seguindo estratégias distintas de investimento. Essas estratégias são definidas entre fundamentalistas, grafistas e operadores de ruído. A partir dessas definições, um estudo sobre o impacto de cada um dos grupos de estratégias na formação de preço é realizado. Por fim, obtém-se evidências de que grandes variações de preço estão relacionadas a grandes lacunas no livro de ofertas.

9. *How markets slowly digest changes in supply and demand* Bouchaud et al. [2009]: Esse artigo estuda o efeito de ordens de grande volume sobre o preço de papéis no longo prazo. Ele revisita diversos aspectos fundamentais e teóricos relacionados a microestrutura de mercado.
10. *Asymmetric statistics of order books: The role of discreteness and evidence for strategic order placement* Zaccaria et al. [2010]: Nesse artigo, estudam-se estatísticas de *spread* Demsetz [1968] em livro de ofertas reais. Avalia-se que a assimetria presente nesses livros de oferta são consequências dos valores discretos pares ou ímpares de tal *spread*. Um modelo teórico é construído buscando-se reproduzir tais estatísticas de livros de ofertas reais. Utilizam-se dados de alta-frequência da bolsa de Nova York para validação experimental do modelo.
11. *Optimal trade execution and price manipulation in order books with time-varying liquidity* Fruth et al. [2014]: Nesse trabalho, estudam-se estratégias para liquidação de portfólios Markowitz [1952]. Por meio delas, existem possibilidades de manipulação de preço dependendo do estado do *spread* do livro de ofertas. O trabalho se foca em mostrar a possibilidade de tal manipulação de forma teórica.
12. *Effective Trade Execution* Cesari et al. [2012]: Nesse artigo, avalia-se o impacto do uso de agentes autônomos de investimento no mercado financeiros. Analisam-se estatísticas relacionadas a tipos de ordens enviadas e de negociações. Estudos dos custos de negociação são avaliados e abordam-se as estratégias de negociação popularmente utilizadas.
13. *Optimizing Market Making using Multi-Agent Reinforcement Learning* Patel [2018]: Aplica-se *aprendizado por reforço* ao problema dos *formadores de mercado* buscando-se execução bem sucedida de ordens limite. O modelo construído é dividido em dois: um decide a direção das operações e o outro decide o posicionamento no livro de ofertas. Utiliza-se bases de dados de alta-frequência de criptomoedas para validação experimental do modelo.

14. *Market Making via Reinforcement Learning* Spooner et al. [2018]: Já nesse trabalho, desenvolve-se um simulador de alta fidelidade de livros de oferta para treinar agentes via *aprendizado por reforço* para resolver o problema dos *formadores de mercado* Radcliffe [1997].
15. *Order-book modelling and market making strategies* Lu & Abergel [2018]: Esse trabalho é dividido em duas partes. Na primeira, avaliam-se diversas características de livros de ofertas reais que contestam a utilização pura de *livros de ofertas de Markov*. Na segunda parte, essas características são utilizadas para estender o modelo teórico do livro de ofertas. Mostra-se que essa extensão melhora o desempenho de estratégias para *formadores de mercado* em contextos realísticos. A validação é feita por meio de simulações e por meio de reproduções passadas, a partir de dados de alta-frequência.
16. *Behind the Price: On the Role of Agent's Reflexivity in Financial Market Microstructure* Barucca & Lillo [2017]: Avalia-se a relação entre a *hipótese de mercado eficiente* e a dinâmica de formação de preço no mercado financeiro. Apresentam-se os conceitos de *impacto de mercado* e *fluxo de ordens*, avaliando-se suas propriedades empíricas e seus usos para estratégias de investimento. O trabalho confirma a importância do uso de dados reais para validar hipóteses qualitativas sobre o comportamento dos investidores.
17. *Machine Learning Techniques for Price Change Forecast Using the Limit Order Book Data* Han et al. [2015]: Nesse trabalho, aplicam-se diferentes técnicas de *aprendizado máquina* para prever variações de preço de papéis. Utilizam-se diversos indicadores derivados do livro de ofertas como entrada de dados para os modelos construídos. Avaliam-se os resultados dos modelos considerando-se métricas padrões de avaliação, como *acurácia* e *recall*.
18. *Enhancing trading strategies with order book signals* Cartea et al. [2018]: Dados de alta frequência da bolsa de valores *Nasdaq* são utilizados para construir um indicador de assimetria de volume do livro de ofertas. Mostra-se que tal indicador é um bom preditor para a direção da próxima ordem a mercado. A partir disso, constrói-se um modelo utilizando *cadeias de Markov*, buscando-se prever variações de preço de curto prazo
19. *Cryptocurrency portfolio management with deep reinforcement learning* Jiang & Liang [2017]: Treina-se uma *rede neuronal profunda* utilizando-se técnicas de *aprendizado por reforço* com o objetivo de gerar um portfólio de criptomoedas

que maximize o retorno acumulado do período de treino. Esse portfólio é então avaliado em um período de tempo posterior e comparada com outras técnicas de otimização de portfólio.

20. *Machine learning for market microstructure and high frequency trading* Kearns & Nevmyvaka [2013]: Esse é um trabalho de múltiplas frentes. Avaliam-se três aplicações de aprendizado de máquina a problemas clássicos de mercado financeiro. Primeiro, estuda-se a aplicação de *aprendizado por reforço* para otimizar negociações de grandes volumes de ações. Em seguida, busca-se prever variações de preço de curto prazo a partir do estado do livro de ofertas. Por fim, buscam-se estratégias de escalonamento de ordens por meio de diversos agentes intermediários.
21. *Price jump prediction in limit order book* Zheng et al. [2012]: Utiliza-se dados de alta-frequência das ações presentes no índice francês *CAC40* para extrair indicadores do livro de ofertas e prever variações de preço de curto prazo dessa ações.
22. *Modelling asset prices for algorithmic and high-frequency trading* Cartea & Jaimungal [2013]: Esse trabalho apresenta o impacto de negociações de alta frequência no mercado financeiro americano. Compara-se o fluxo de eventos entre 2001 e 2008. Além disso, treina-se um *modelo escondido de Markov* com o objetivo de posicionar ordens no livro de ofertas e ganhar sobre variações de *spread*.
23. *A high-frequency trade execution model for supervised learning* Dixon [2018]: Nesse artigo, introduz-se um modelo de negociação de alta frequência para avaliar o impacto econômico de agentes treinados a partir de *aprendizado de máquina supervisionado*. Avaliam-se o desempenho de tais agentes por meio de *matrizes de confusão* e por meio de *matrizes de informação de negociação*, que consideram lucros e prejuízos das operações dos agentes. Utilizam-se os dados históricos do índice *S&P500* para validação experimental do modelo de negociação.
24. *Buy low, sell high: A high frequency trading perspective* Cartea et al. [2014]: Nesse trabalho, utiliza-se uma estratégia de investimento baseada em alta velocidade de processamento de dados e baixa latência para posicionamento de ordens no livro de ofertas buscando-se lucro sobre o *spread*. Além disso, modela-se o impacto de notícias e ordens a mercado sobre variações de preço de curto prazo.
25. *Limit order placement optimization with Deep Reinforcement Learning: Learning from patterns in cryptocurrency market data* Juchli [2018]: Nesse trabalho,

utiliza-se *aprendizado por reforço profundo* para tentar resolver o problema de posicionamento de ordens no livro de ofertas, buscando-se sua execução em um horizonte de cem segundos. Dados históricos de alta frequência de criptomoedas são utilizadas para validação experimental.

26. *Optimal high-frequency trading with limit and market orders* Guilbaud & Pham [2013]: Esse trabalho propõe um ferramental para estudar políticas de negociação de *formadores de mercado*. Utilizam-se tanto ordens limite quanto ordens a mercado. Além disso, modela-se a prioridade de execução de tais ordens de forma a obter melhores políticas. O sistema é modelado com o uso de *cadeias de Markov* e resolvido utilizando-se técnicas de *programação dinâmica*. Conclui-se que utilizar a prioridade das ordens na modelagem do problema tem impacto financeiro positivo nas políticas resultantes.
27. *The price impact of order book events* Cont et al. [2014]: Estuda-se o impacto de diferentes tipos de ordens sobre a formação de preço de cinquenta ações americanas. É demonstrado empiricamente que, em pequenos intervalos de tempo, as variações de preço são controladas principalmente pelo desbalanceamento entre oferta e demanda das melhores ordens de compra e venda dos papéis.
28. *Multiple kernel learning on the limit order book* Fletcher et al. [2010]: Um conjunto de indicadores simples, derivados a partir do livro de ofertas do papel da moeda *EURUSD*, é utilizado para construir um conjunto de *kernels*. Esses *kernels* são utilizados para treinar modelos *SVM* para prever variações de preço futuras. Os resultados revelam o potencial preditivo dos *kernels* e sua importância para as tarefas de predição.
29. *More statistical properties of order books and price impact* Potters & Bouchaud [2003]: Nesse estudo, avaliam-se novas estatísticas a partir de dados do livro de ofertas de ações inglesas e francesas. Primeiro, foca-se nas distribuições de preço de ordens limite que chegam ao livro de ofertas. Em seguida, avalia-se o formato médio do livro de ofertas, considerando-se o volume de ordens em cada nível de preço. Por fim, estuda-se o tempo de vida de ordens considerando-se a distância de preço entre elas e as ordens de melhor preço.
30. *The flash crash: High-frequency trading in an electronic market* Kirilenko et al. [2017]: Nesse artigo, avalia-se quantitativamente os eventos que lavaram ao *Flash Crash* de 2010, uma queda abrupta de preço do mini-contrato futuro americano do *S&P500*. Avalia-se que a pressão de venda do dia fez com que agentes de alta

frequência provocassem grande volatilidade no dia. O evento levantou diversas questões que viriam a se tornar regulamentações futuras sobre a atividade de negociação de alta frequência.

31. *Market microstructure: A survey* Madhavan [2000]: Esse trabalho apresenta as bases teóricas e empíricas relacionadas a microestrutura de mercado. Ele expõe tópicos como formação de preços, estrutura e projeto de mercados e transparência no que se refere à informação acessível aos agentes presentes no mercado.
32. *Design and evaluation of automatic agents for stock market intraday trading* Jabbur et al. [2014]: Nesse trabalho o autor propõe o uso do conceito das *Ondas de Elliot* Poser [2003] em conjunto a agentes autônomos operando no mercado *intraday*. O modelo usa como entrada informações de preço dos papéis estudados em diferentes granularidades de tempo, gerando regras por meio de aprendizado de máquina para o processo de tomada de decisão.
33. *A neural network based approach to support the market making strategies in high-frequency trading* Silva et al. [2014]: Nesse trabalho o autor propõe o uso de *redes neuronais*, buscando identificar e prever oscilações positivas de preço em curtos períodos de tempo.
34. *Forecasting stock market movement direction with support vector machine* Huang et al. [2005]: Esse trabalho faz comparações entre as máquinas de vetores de suporte, *SVMs*, e outros métodos de aprendizado de máquina na previsão de tendências de preço.

Com os trabalhos descritos, pode-se classificá-los quanto aos seguintes aspectos:

- **Estratégia de investimento (EDI)**: Se o trabalho propõe e avalia financeiramente uma estratégia de investimento;
- **Estatísticas do livro de ofertas (ELO)**: Se o trabalho deriva ou estuda estatísticas de eventos relacionados ao livro de ofertas e aspectos de microestrutura de mercado;
- **Dados de alta frequência (DAF)**: Se o trabalho utiliza dados de alta frequência para simulação do livro de ofertas e validação experimental;
- **Indicadores (IND)**: Se o trabalho deriva indicadores a partir do livro de ofertas e os utiliza como entrada para preditores ou modelos de regressão;

- **Aprendizado de máquina** (AMQ): Se o trabalho aplica técnicas e algoritmos de aprendizado de máquina para seus objetivos;
- **Teórico** (TEO): Se o trabalho aborda o assunto de microestrutura de mercado de forma teórica, apresentando conceitos da área.

A Tabela 3.1 apresenta as classificações de cada um dos trabalhos apresentados de acordo com os aspectos citados. Ao final da tabela, considera-se o posicionamento deste trabalho de mestrado em relação a eles.

A partir da Tabela 3.1, mostra-se que esse trabalho de mestrado é multidisciplinar. Descrevem-se os fundamentos necessários sobre microestrutura de mercado (*TEO*), derivam-se indicadores a partir do livro de ofertas (*IND*), aplicam-se técnicas de aprendizado de máquina (*AMQ*), constrói-se uma estratégia de investimento (*EDI*) e, por fim, validam-se os resultados utilizando-se dados reais de alta frequência (*DAF*). Trata-se de um trabalho que aborda o problema de investir utilizando-se agentes autônomos de investimento de ponta a ponta, abordando-se diversos dos aspectos apresentados.

	EDI	ELO	DAF	IND	AMQ	TEO
31 Madhavan [2000]						X
29 Potters & Bouchaud [2003]		X	X			X
34 Huang et al. [2005]				X	X	
8 Chiarella et al. [2009]		X	X			X
9 Bouchaud et al. [2009]			X			X
10 Zaccaria et al. [2010]		X	X			X
28 Fletcher et al. [2010]	X		X	X	X	
7 Qiao [2012]		X	X	X		
12 Cesari et al. [2012]	X	X	X			
21 Zheng et al. [2012]			X	X	X	
6 Lehalle [2013]						X
26 Guilbaud & Pham [2013]	X				X	X
20 Kearns & Nevmyvaka [2013]	X		X		X	X
22 Cartea & Jaimungal [2013]	X	X	X		X	
11 Fruth et al. [2014]	X					X
32 Jabbur et al. [2014]	X			X	X	
33 Silva et al. [2014]	X			X	X	
27 Cont et al. [2014]		X	X			
24 Cartea et al. [2014]	X	X	X			
17 Han et al. [2015]			X	X	X	
30 Kirilenko et al. [2017]			X			X
16 Barucca & Lillo [2017]		X	X	X		X
19 Jiang & Liang [2017]	X				X	
5 Abergel et al. [2017]	X				X	X
1 le Calvez & Cliff [2018]			X		X	
2 Jo et al. [2019]			X		X	
3 Qureshi [2018]		X	X	X	X	
13 Patel [2018]	X		X		X	
14 Spooner et al. [2018]	X		X		X	
15 Lu & Abergel [2018]		X	X			X
18 Cartea et al. [2018]			X	X	X	
23 Dixon [2018]	X				X	X
25 Juchli [2018]	X		X		X	
4 Chowdhury et al. [2019]	X				X	
Este trabalho	X		X	X	X	X

Tabela 3.1. Relação dos trabalhos citados com os aspectos apresentados. Na última linha, apresenta-se a relação desse trabalho com esses aspectos.

Capítulo 4

Metodologia

Esse projeto foi desenvolvido em quatro seções: *simulador*, *rotulação de séries de preço*, *indicadores* e *construção e validação da estratégia* de investimento. A Figura 4.1 trás um diagrama apresentando tais seções e seus respectivos detalhes.

Na seção de *simulador*, aborda-se os detalhes específicos relacionados à simulação do período de negociação e uma notação é apresentada de forma a abstrair os detalhes técnicos da mesma. Na seção de *rotulação de séries de preço*, discute-se como a série de preços é rotulada. Na seção de *indicadores*, denota-se os indicadores construídos para esse projeto. Por fim, na seção de *construção e validação da estratégia*, apresenta-se a construção da estratégia de investimento e seu processo de avaliação experimental.

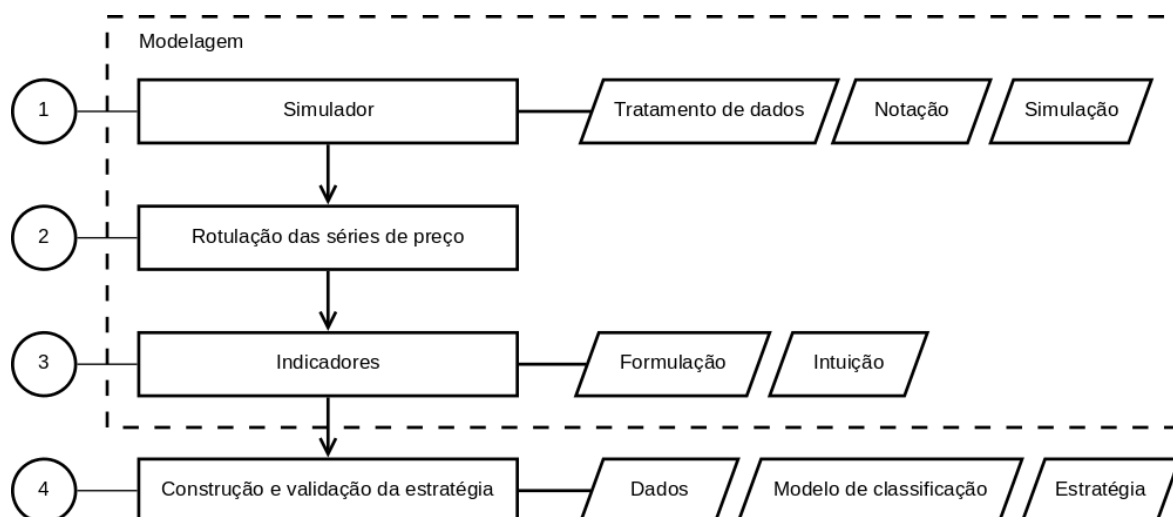


Figura 4.1. Metodologia e seus detalhes.

4.1 Simulador

Objetivando-se obter resultados verossímeis, um modelo fiel à forma como as negociações da bolsa funcionam é necessário. Dessa forma, construiu-se um simulador capaz de reproduzir os eventos reais do período de negociação de determinado papel, sendo possível reproduzir qualquer dia de operação. Nas próximas seções descreve-se o formato dos dados, seu tratamento, o funcionamento do simulador e, por fim, uma notação para abstrair tal funcionamento.

4.1.1 Tratamento de dados

Utiliza-se de insumo para o simulador os dados reais da bolsa de valores *B3*. Esses dados estão disponíveis no repositório de dados B3 [2019c] da mesma.

Os dados presentes nesse repositório são arquivos de registros de todas as *ordens* e *negociações* de cada dia da bolsa. Cada dia é dividido em três arquivos: um arquivo das ordens de compra, um arquivo das ordens de venda e um arquivo das negociações realizadas. Esses arquivos são textuais e com posicionamento bem definido.

Os arquivos das ordens de compra e venda possuem as seguintes propriedades:

- **Data da sessão:** Dia de operação ao qual o arquivo se refere;
- **Símbolo do instrumento:** Símbolo da papel operado, como, por exemplo, *PETR4* para ações preferenciais da empresa *Petrobras*;
- **Sentido:** Sentido da ordem, compra ou venda;
- **Número de sequência:** Identificador único daquela oferta;
- **Identificador de geração:** Identificador secundário da ordem, utilizado para versionar a mesma. Cada versão de uma ordem representa um estado em que a mesma esteve em determinado momento no tempo;
- **Evento gerador:** Evento que gerou a ordem. Os eventos possíveis são: *nova ordem*, *modificação*, *cancelamento pelo operador*, *negociação*, *reentrada*, *novo preço de parada*, *rejeição*, *remoção pelo sistema*, *preço de parada atingido* e *expiração*;
- **Momento de registro:** Horário no qual a ordem foi registrada no sistema. É o campo utilizado para definir a prioridade da ordem dentro de um mesmo nível de preço;
- **Preço:** Valor pelo qual a ordem deve ser negociada;

- **Indicador de prioridade:** É o segundo parâmetro para prioridade dentro de um mesmo nível de preço, após o momento de registro;
- **Quantidade ofertada:** É o tamanho da ordem enviada, contendo a quantidade máxima de papéis a ser negociada;
- **Quantidade negociada:** Representa a quantidade de papéis, daquela ordem, que foi negociada;
- **Data de entrada da ordem:** É a data em que a ordem foi registrada, podendo ser anterior à data de sessão;
- **Estado da ordem:** Representa o estado da versão mais atual da ordem. O estado inicial de uma ordem é sempre *nova ordem*. A partir dele, a ordem pode ir para os estados intermediários, isto é, os estados *parcialmente negociada* e *modificada*. Por fim, uma ordem chega a seus estados terminais, os estados *totalmente negociada*, *rejeitada*, *cancelada* e *expirada*;
- **Condição da ordem:** Representa se a ordem é *neutra*, *agressora* ou *receptora*. Ordens neutras são registradas sem ser executadas de imediato. Já ordens agressoras e receptoras, são aquelas que fazem parte de uma negociação;
- **Corretora:** O código numérico referente à corretora da bolsa de valores à qual a ordem pertence B3 [2019e].

Os arquivos de negociações possuem diversas propriedades compartilhadas com os arquivos de ordens, como *data da sessão* e *símbolo do instrumento*. Como um negócio é formado entre duas ordens, uma de compra e uma de venda, os arquivos de negociações possuem as propriedades de *data de entrada*, *número de sequência*, *identificador de geração*, *condição* e *corretora* de ambas. Além desses, as seguintes propriedades estão definidas no mesmo:

- **Número do negócio:** Identificador único do negócio;
- **Preço do negócio:** É o valor pelo qual o negócio foi executado;
- **Quantidade:** Quantidade de papéis negociada;
- **Momento da negociação:** Horário no qual o negócio foi executado;
- **Indicador de anulação:** Indica que determinados negócios já executados foram revertidos e suas respectivas ordens canceladas. Quando isso ocorre, a bolsa informa a seus investidores do fato e é feito um novo leilão;

- **Indicador de negociação cruzada:** Uma negociação cruzada ocorre quando uma corretora realiza troca de papéis entre seus próprios investidores e então notifica à bolsa de valores. Esse indicador define se uma negociação cruzada dentro de uma corretora foi intencional ou não. Negociações cruzadas são limitadas às regras da bolsa de valores.

Com as propriedades de cada arquivo bem definidas, pode-se definir o processo de tratamento de dados. Esse processo é definido em diversas etapas assim como exibido na Figura 4.2.

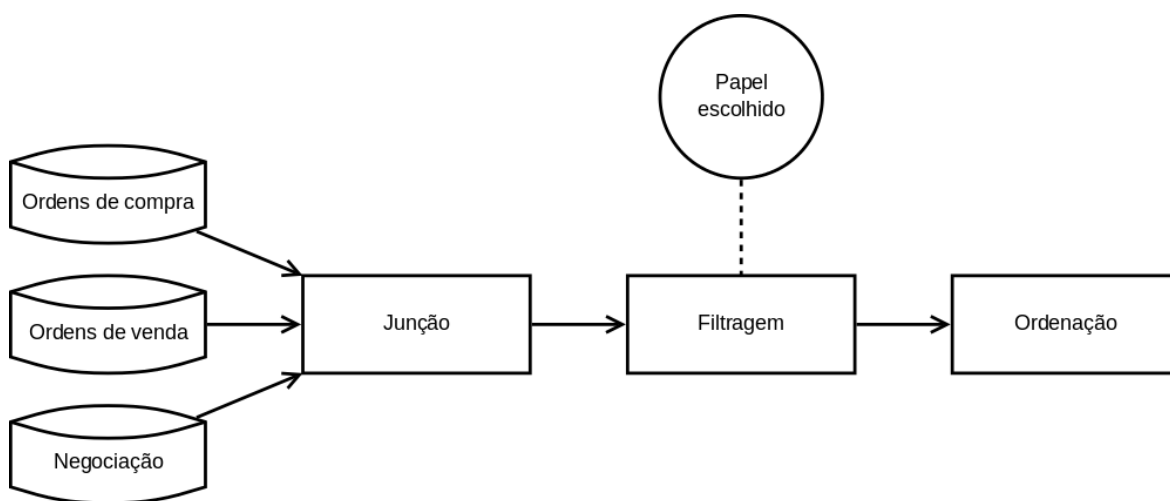


Figura 4.2. Etapas do tratamento de dados.

A primeira etapa do tratamento de dados é a junção dos três arquivos descritos. Somente com os dados de todos eles é possível reproduzir aquele dia de negociação. Por se tratar de arquivos de texto bem definidos, o processo de junção é trivial.

A etapa seguinte é a filtragem. Nessa etapa escolhe-se o papel alvo para extração. Para papéis do tipo de ações, basta escolher o papel e extrair os eventos listados com o símbolo do mesmo. Já para contratos futuros, cujo símbolo muda de acordo com o vencimento dos mesmos, deve-se considerar o dia da sessão de forma a selecionar os eventos do contrato com vencimento mais próximo ou o de preferência.

Com os registros extraídos, esses devem ser ordenados de acordo com diferentes propriedades. As propriedades utilizadas para esse ordenação são os seguintes, em ordem de preferência:

1. **Momento de registro:** Ordenam-se as ordens e negociações de acordo com seus momentos de registro e execução;

2. **Identificador de geração:** Para eventos que ocorreram em um mesmo momento, utiliza-se seu identificador de geração, já que esse é um campo estritamente crescente durante o dia de operação. Já para negócios, considera-se o identificador de geração da ordem mais recente;
3. **Tipo de evento:** Por fim, ordens possuem prioridade sobre negociações no processo de ordenação.

A partir dessa ordenação, algumas consequências surgem. A primeira delas é que eventos que ocorrem em um mesmo momento são atômicos, isto é, não podem ser divididos e portanto devem ser tratados de tal forma durante a simulação. A segunda é que a ordenação apresentada garante que as ordens presentes em uma negociação são exibidas antes de suas respectivas negociações.

Por fim, com todos os eventos extraídos e ordenados, grava-se um arquivo binário dos mesmos. Esse arquivo é utilizado para execução no simulador.

4.1.2 Simulação

Com os dados tratados, utiliza-se o simulador construído para reproduzir fielmente os eventos de determinado papel na bolsa de valores de forma determinística. Os detalhes técnicos de sua implementação não são explicitados por não ser o objetivo. O funcionamento do mesmo pode ser dividido em duas entidades: *emissor de eventos* e *livro de ofertas*. A Figura 4.3 representa o funcionamento desse simulador.

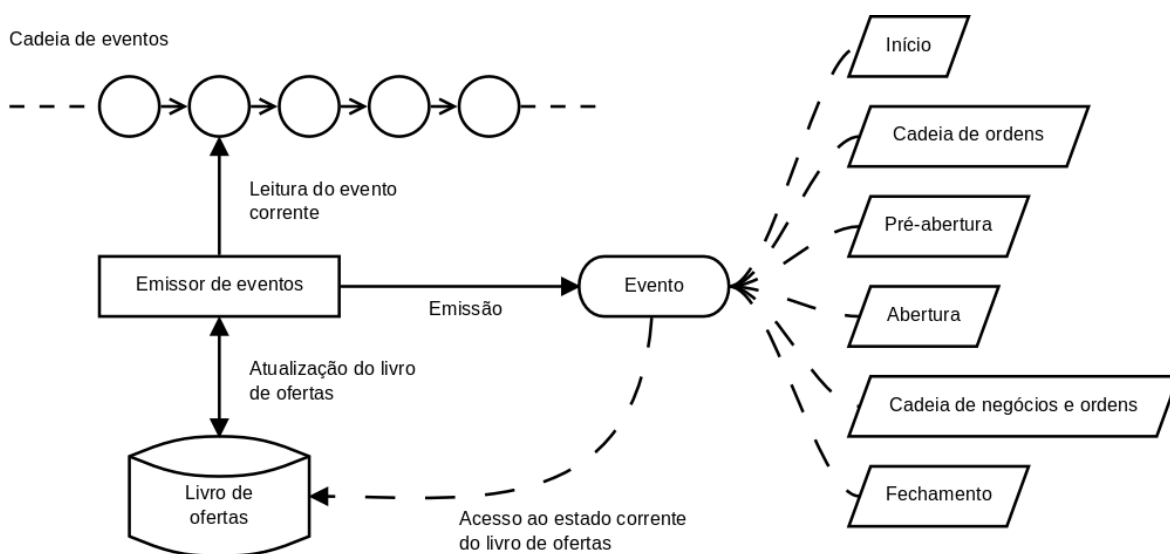


Figura 4.3. Representação do funcionamento do simulador.

O *livro de ofertas* é responsável por controlar e organizar todas as ordens recebidas por ele. Por meio dele é possível medir diversas características de interesse em relação ao papel escolhido. Ele é construído e atualizado à medida que os registros são lidos. Vale lembrar que ordens em estados terminais não ficam posicionadas no livro de ofertas.

O *emissor de eventos* é núcleo de funcionamento do simulador. Ele fica responsável pela leitura dos eventos contida no arquivo produzido pelo processo anterior. A partir dessa leitura, ele acumula eventos gerados em um mesmo momento, produzindo cadeias de eventos atômicas. Em segundo plano, ele atualiza o livro de ofertas à medida que recebe ordens de compra e venda. Assim que uma cadeia de eventos é acumulada, um evento é emitido notificando a cadeia de eventos fechada.

Os eventos produzidos pelo *emissor de eventos*, considerando-se as divisões do período de negociação, são os seguintes:

1. **Início:** A simulação está prestes a iniciar, mas nenhum evento foi processado ainda;
2. **Cadeia de ordens:** Esse evento só é emitido enquanto o mercado não está aberto. A lista de ordens processadas é disponibilizada juntamente à versão mais atual do livro de ofertas. Nesse momento, o livro de ofertas permite o cruzamento de ordens sem a produção de negócios;
3. **Pré-abertura:** A fase pré-abertura chegou ao fim. Nesse momento as primeiras negociações vão acontecer. A definição do preço inicial se dá pelo algoritmo de preço de equilíbrio;
4. **Abertura:** A primeira cadeia de negócios e ordens é emitida. A lista de ordens e negócios processado é disponibilizada juntamente à versão mais atual do livro de ofertas. Depois desse momento, não mais se permite que ordens se cruzem no livro de ofertas sem a ocorrência de negócios;
5. **Cadeia de negócios e ordens:** A lista de ordens e negócios processado é disponibilizada juntamente à versão mais atual do livro de ofertas;
6. **Fechamento:** Reporta o fim das operações. Disponibiliza a última cadeia de eventos. O livro de ofertas nesse momento é esvaziado.

Com base nesses eventos, todos os estudos desse projeto se tornam viáveis. Esse arcabouço reproduz de forma determinística os eventos decorridos em determinado papel da bolsa de valores, contando-se que os arquivos de registro estejam disponíveis.

4.1.3 Notação

Com o simulador descrito, pode-se explicar como os indicadores foram construídos. Porém, de forma a tornar a descrição mais simples e direta, apresenta-se uma notação matemática representando o simulador. A partir dessa notação, torna-se trivial a explicação dos indicadores construídos.

Como já foi descrito, todos os eventos são agrupados em cadeias de mesmo momento. Pode-se dizer que cada uma dessas cadeias possui um índice, que representa a ordem em que aconteceu. Cada um desses índices representa um momento específico no tempo. Com base nisso, define-se:

- i : É o índice de uma cadeia de eventos. Está atrelado a um momento específico no tempo. $i + 1$ é o índice seguinte e, portanto, refere-se a um momento no tempo imediatamente posterior a i ;
- B_i : Representa a i -ésima cadeia de eventos processada pelo simulador. Pode-se tratar tanto das cadeias anterior à abertura de mercado quanto às posteriores. Considerando-se uma cadeia B_i fixada, pode-se definir os seguintes símbolos:
 - e_{ji} : É o j -ésimo evento presente na cadeia de eventos B_i . Pode-se definir os limites de j no intervalo $[0, n]$, de forma que e_{0i} e e_{ni} representam o primeiro e último eventos da cadeia, respectivamente;
 - o_{ji} : Considerando-se apenas as ordens de compra e venda dessa cadeia de eventos, pode-se definir o_{ji} como a j -ésima ordem presente na cadeia de eventos. Define-se j no intervalo $[0, m]$. Como as ordens podem ser de compra ou venda, definem-se:
 - * o_{ji}^+ : A j -ésima ordem de **compra** da cadeia de eventos, com j no intervalo $[0, y]$;
 - * o_{ji}^- : A j -ésima ordem de **venda** da cadeia de eventos, com j no intervalo $[0, w]$.
 - t_{ji} : Representa a j -ésima negociação presente na cadeia de eventos, com j definido no intervalo $[0, h]$. Assim como foi feito com as ordens, pode-se definir a seguinte variação para as negociações:
 - * t_{ji}^+ : A j -ésima negociação de **compra** da cadeia de eventos. Nesse caso, diz-se que a negociação é de compra no caso de a ordem de compra ser a **ordem agressora**. Nesse caso, j está definido no intervalo $[0, k]$;

- * t_{ji}^- : A j -ésima negociação de **venda** da cadeia de eventos. Nesse caso, diz-se que a negociação é de venda no caso de a ordem de venda ser a **ordem agressora**. Nesse caso, j está definido no intervalo $[0, z]$.
- A quantidade total de eventos é igual à quantidade de ordens somada à quantidade de negociações. Desta forma, $n = m + h = (y + w) + (k + z)$.
- p_i : O preço do papel escolhido no instante referenciado pelo índice i . Deve-se lembrar que esse valor se refere ao preço da última negociação executada até o índice selecionado. Dessa forma, esse símbolo é indefinido enquanto o mercado está fechado;
- O_i : O estado do livro de ofertas após o processamento da cadeia de eventos B_i . Pode-se dizer que é uma matriz multi-dimensional que endereça as ordens de acordo com as regras já explicitadas. Os seguintes símbolos detalham melhor esse estado do livro de ofertas:
 - O_i^+ : O lado dos compradores do livro de ofertas;
 - O_i^- : O lado dos vendedores do livro de ofertas;
 - O_{vi}^+ : A lista de prioridade cujas ordens possuem índice de valor igual a v . O índice v representa a posição entre os níveis de preço definidos no lado dos compradores. Dessa forma, O_{0i}^+ representa o nível de preço de maior valor no lado dos compradores do livro de ofertas;
 - O_{vi}^- : A lista de prioridade cujas ordens possuem índice de valor igual a v . O índice v representa a posição entre os níveis de preço definidos no lado dos vendedores. Dessa forma, O_{0i}^- representa o nível de preço de menor valor no lado dos vendedores do livro de ofertas;
 - P_{vi}^+ : Representa o valor definido pelo nível de preço de índice v do lado dos compradores;
 - P_{vi}^- : Representa o valor definido pelo nível de preço de índice v do lado dos vendedores.

Além dos símbolos descritos, algumas funções básicas devem ser definidas:

- $x(e)$: Extrai o preço do evento de entrada e . Se tal evento e for uma ordem o , $x(o)$ é o preço de compra ou venda de tal ordem. No caso de e ser uma negociação t , $x(t)$ é o preço da negociação em questão;

- $s(e)$: Extrai o volume de um evento de entrada e . Caso e seja uma ordem o , $s(o)$ é quantidade máxima de papéis a ser negociada por tal ordem. Pode-se definir também $\hat{s}(o)$, representando a quantidade restante de ordens a serem executadas por tal ordem, caso esta esteja parcialmente executada. Caso e seja uma negociação t , $s(t)$ é a quantidade de papéis negociados;
- $u(i)$: Extrai o momento real indexado por i .

Os símbolos e funções definidos representam e abstraem o comportamento do simulador. Com eles é possível descrever completamente quaisquer características desejadas.

4.2 Rotulação das séries de preço

Uma série de preços de determinado papel pode ser classificada de forma simples como *crecente*, *decrecente* ou *neutra*. Ela é *crecente* quando seu valor varia positivamente e *decrecente* quando seu valor varia negativamente. Além disso, pode-se dizer que existem momentos em que o mercado está ‘lateralizado’ e, portanto, nesse caso, a série é *neutra*.

Em geral, cada estratégia de investimento utiliza uma rotulação própria e, então, busca-se estabelecer regras que tentem aproximar tal rotulação. Dessa forma, uma estratégia de rotulação se faz necessária. Com uma série rotulada, sabe-se qual ação deve ser tomada por um agente autônomo. Isto é, se o rótulo da série é *crecente*, deve-se comprar o papel em busca de lucro. Se o rótulo da série é *decrecente* deve-se vender o papel. Se é *neutra*, não se deve estar posicionado.

Uma estratégia ótima de rotulação deve possuir os seguintes requisitos:

- **Opera no momento ideal:** Garante-se que as operações ocorrem no momento de máximo lucro, isto é, a decisão de compra ou venda do papel não será mais lucrativa sendo tomada mais cedo ou mais tarde;
- **Lucro mínimo:** Para cada par de compra e venda de determinado papel, o lucro não deve ser menor que um valor previamente determinado;

Considerando-se os requisitos apresentados, construiu-se um algoritmo de rotulação capaz de produzir as saídas esperadas. Essa algoritmo alcança tal objetivo acessando momentos futuros da série de preços, simulando-se um *operador oráculo*. A partir desses acessos, consegue-se definir mínimos e máximos locais na série de preços

que obedeçam a tais requisitos. Dessa forma, define-se que, entre mínimos e máximos, respectivamente, possui-se séries *crescentes* de preço. Já entre máximos e mínimos, respectivamente, possui-se séries *decrecentes* de preço. Inclui-se apenas uma regra adicional a tal algoritmo de forma que o mesmo trate dos casos da série em estado *neutro*.

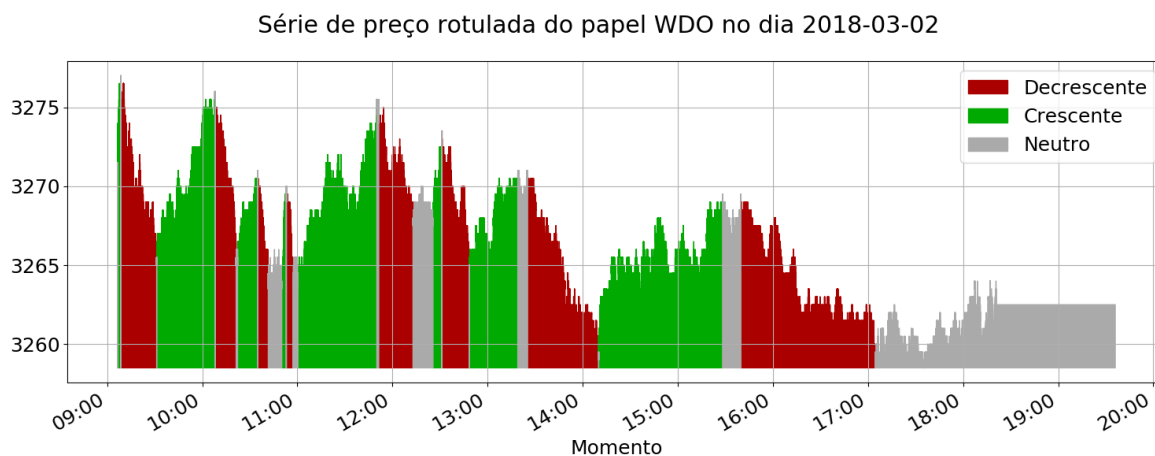


Figura 4.4. Série de preço rotulada do mini-dólar futuro no dia 02 de Março de 2018.

A Figura 4.4 apresenta um exemplo de saída do algoritmo. As regiões em vermelho são séries de preço *decrecentes*. As regiões em verde são séries de preço *crescentes*. Por fim, as regiões em cinza são séries de preço *neutras*. Os melhores momentos de operação nessas séries rotuladas são nos momentos de transição, isto é, quando a série muda de um rótulo para outro.

4.3 Indicadores

Nesta seção são apresentados os indicadores utilizados nesse projeto. A notação apresentada anteriormente é fortemente utilizada nessa seção evitando-se a necessidade de explicitar pseudo-códigos de cada indicador. Para cada um deles, apresenta-se sua fórmula, uma intuição e um exemplo gráfico do mesmo. Os exemplos apresentados são do mini-dólar futuro no dia 02 de Março de 2018.

4.3.1 *Spread e Mid-price*

O *spread* corresponde à distância entre os preços das melhores ordens de venda e compra presentes no livro de ofertas. Ele só pode ser definido após a abertura do mercado,

considerando-se que, antes disso, é possível que haja cruzamento de ordens no livro de ofertas. O *spread* pode ser calculado da seguinte forma:

$$spr_i = spr(O_i) = P_{0i}^- - P_{0i}^+ \quad (4.1)$$

Empiricamente, o *spread* carrega informação sobre a volatilidade de preço em curtos períodos de tempo. Quanto maior for o *spread*, mais volátil é o preço naquele momento. Por outro lado, quanto menor for o *spread*, mais estável é o preço no período. Além disso, valores altos do indicador mostram que compradores e vendedores não estão de acordo com o preço do papel.

O *mid-price* é definido como a média entre os preços das melhores ordens de venda e compra presentes no livro de ofertas. Ele representa uma medida central do preço real para as transações imediatamente procedentes. Ele é definido como:

$$mid_i = mid(O_i) = \frac{P_{0i}^- + P_{0i}^+}{2} \quad (4.2)$$

Ao se usar ambos os indicadores juntos, é estimar o preço da próxima transação. Considerando-se que apenas uma unidade do papel em questão é transacionada, o preço de tal transação pode ser definido da seguinte forma:

$$\begin{aligned} p_{i+1} &= mid_i \pm \frac{spr_i}{2} \\ p_{i+1} &= mid_i + \frac{spr_i}{2}, \text{ em uma negociação de compra} \\ p_{i+1} &= mid_i - \frac{spr_i}{2}, \text{ em uma negociação de venda} \end{aligned}$$

A Figura 4.5 apresenta uma visualização de ambos os indicadores. O espaço entre as regiões vermelha e verde compõem o *spread*, enquanto que os pontos azuis representam o *mid-price*.

No exemplo apresentado, o *spread* é quase sempre constante, assim como demonstrado pelo histograma lateral. No histograma fica visível que o valor do *spread* fica, na maior parte do tempo, com o valor mínimo de variação de preço do papel escolhido. Esse é o comportamento esperado de um papel muito negociado quanto o *mini-dólar futuro*. Vale lembrar que esse papel tem variação mínima de preço igual a 0.5 pontos.

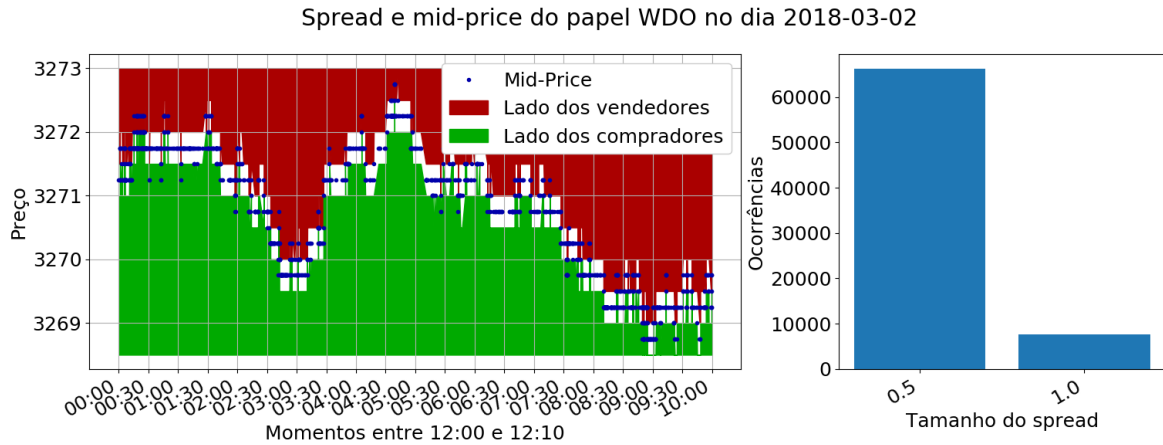


Figura 4.5. *Spread e do mid-price.*

4.3.2 Compras, vendas e volume

Pode-se contar a quantidade de papéis negociados em uma cadeia de eventos. As contagens possíveis são *compras*, *vendas* e *volume*, como definido a seguir:

$$pur_i = pur(B_i) = \sum_{j=0}^k s(t_{ji}^+) \quad (4.3)$$

$$sal_i = sal(B_i) = \sum_{j=0}^z s(t_{ji}^-) \quad (4.4)$$

$$vol_i = vol(B_i) = pur(B_i) + sal(B_i) = \sum_{j=0}^h s(t_{ji}) \quad (4.5)$$

Essas contagens são úteis para definir outros indicadores de forma simplificada. A Figura 4.6 mostra a quantidade de compras e vendas durante o período de negociação. A partir dos histogramas, é possível perceber a simetria da quantidade de negócios de compra e venda desse dia. Essa característica tende a se manter para papéis muito negociados.

4.3.3 Posição e desbalanceamento

O indicador de *posição* é definido como a soma cumulativa dos indicadores de *compras*, *vendas* e *volume*. Esse indicador, por construção, é monotonicamente crescente. Suas três definições são:

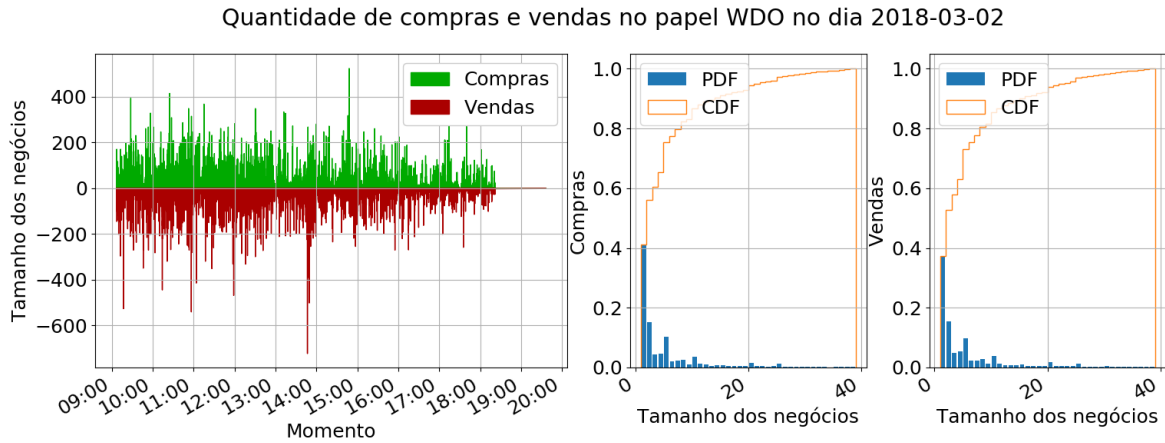


Figura 4.6. Quantidade de compras e vendas durante o período de negociação. Apresentam-se nos painéis laterais as distribuições de probabilidade dos volumes desses indicadores.

$$ppos_i = pos(pur, i) = \sum_{j=0}^i pur_j = ppos_{i-1} + pur_i \quad (4.6)$$

$$spos_i = pos(sal, i) = \sum_{j=0}^i sal_j = spos_{i-1} + sal_i \quad (4.7)$$

$$vpos_i = pos(vol, i) = \sum_{j=0}^i vol_j = vpos_{i-1} + vol_i \quad (4.8)$$

O indicador de *desbalanceamento* é definido como a diferença entre os indicadores *posição de compra* e *posição de venda*. Quando esse indicador possui valor positivo, ele mostra que houve preferência por compras sobre vendas. Quando esse indicador possui valor negativo, ele mostra que houve preferência por vendas sobre compras. Ele é definido como:

$$imb_i = ppos_i - spos_i \quad (4.9)$$

A partir da Figura 4.7 é possível ver que ambas posições de compra e venda crescem uma próxima da outra. Ao se analisar o indicador de desbalanceamento, a semelhança de sua direção com a direção do preço do papel é notável. Essa característica dá a intuição de que a direção de oscilação do preço é controlada pela direção agregada das negociações no curto prazo. Além disso, ao fim do dia, percebe-se que a

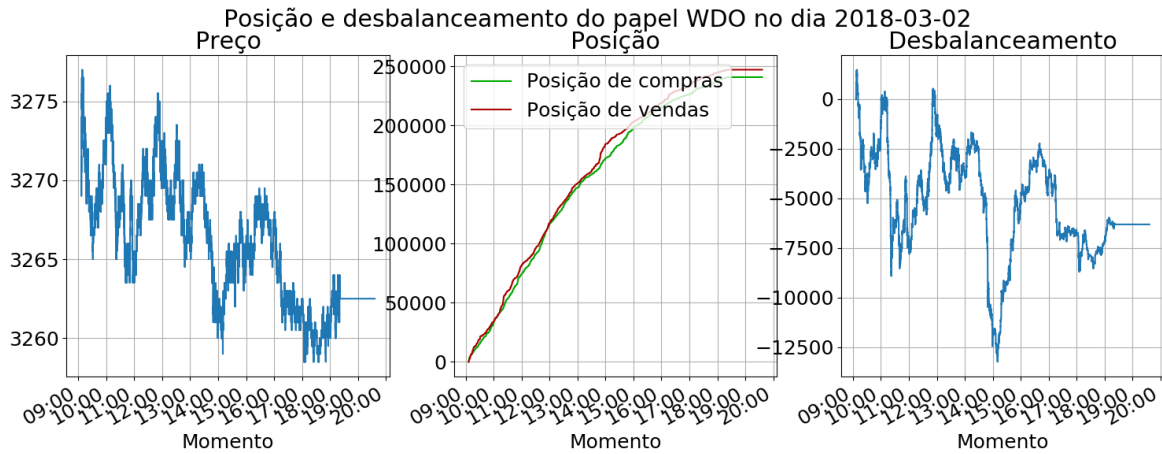


Figura 4.7. Indicadores de posição e desbalanceamento.

posição de venda supera a *posição de compra*, mostrando que houveram mais vendas a mercado que compras a mercado nesse dia. Essa informação concorda com o fato de que a variação de preço do papel no dia foi negativa. Não se trata de uma regra geral, dado que a magnitude da variação é controlada pelo livro de ofertas. Portanto, é possível que o desbalanceamento seja negativo e a variação de preço do papel seja positiva. De toda forma, o indicador busca a direção da movimentação do preço no curto prazo.

4.3.4 Velocidade

O indicador de *velocidade* é definido como a diferença entre as posições de compra e venda sobre janelas de tempo. O tamanho dessas janelas é parametrizável pelo parâmetro q , que considera a quantidade de cadeias de evento a se considerar. A velocidade é definida como:

$$pspe_i^q = spe(ppos, i, q) = \frac{ppos_i - ppos_{i-q}}{u(i) - u(i-q)} \quad (4.10)$$

$$sspe_i^q = spe(spos, i, q) = \frac{spos_i - spos_{i-q}}{u(i) - u(i-q)} \quad (4.11)$$

$$vspe_i^q = spe(vpos, i, q) = \frac{vpos_i - vpos_{i-q}}{u(i) - u(i-q)} \quad (4.12)$$

Para valores relativamente grandes de q , o indicador de velocidade aproxima a **velocidade média** do crescimento do indicador de posição. Para valores relativamente pequenos de q , o indicador aproxima a **velocidade instantânea** do crescimento do

indicador de posição.

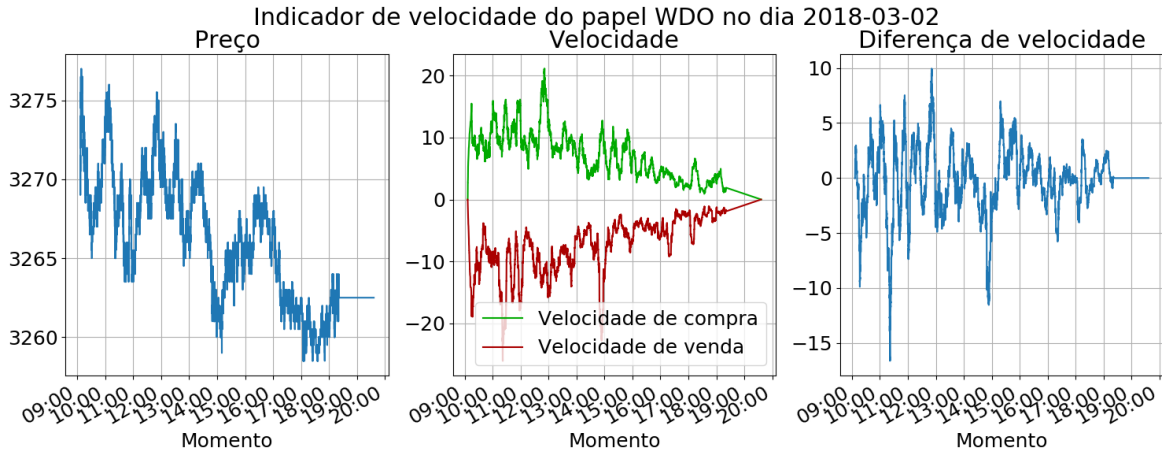


Figura 4.8. Representação do indicador de velocidade de compra e de venda. O parâmetro q foi definido de forma a obter uma janela de tempo de 5 minutos para esse exemplo.

O indicador de *velocidade* é útil na captura do quão rápido as negociações estão ocorrendo. Dessa forma, quando há diferença significativa entre velocidades de compra e venda, ganha-se a intuição da direção futura do preço do papel. A Figura 4.8 apresenta os indicadores e a respectiva diferença entre eles.

4.3.5 Densidade e profundidade

O indicador de *densidade* é definido como a quantidade de papéis posicionados em um nível de preço v . Ele trás a informação da quantidade de papéis que devem ser comprados ou vendidos de forma a ‘ultrapassar’ aquele nível de preço v .

$$den_{vi}^+ = \sum \{\hat{s}(o) | o \in O_{vi}^+\} \quad (4.13)$$

$$den_{vi}^- = \sum \{\hat{s}(o) | o \in O_{vi}^-\} \quad (4.14)$$

O indicador de *profundidade* soma a quantidade de papéis que devem ser comprados ou vendidos de forma que o preço daquele papel mude para o preço alvo. Isso significa que, buscando-se que o preço do papel varie de p_i para P_{vi}^+ , com v sendo o índice de tal preço, a quantidade necessária de papéis a serem vendidos é:

$$dep_{vi}^+ = 1 + \sum_{j=0}^{v-1} den_{ji}^+ \quad (4.15)$$

O somatório vai apenas até $v - 1$ ao invés de v . O motivo disso é que apenas mais um papel precisa ser vendido quando todos os níveis de preço, exceto v , estão vazios. Considerando-se o alvo P_{vi}^- , a quantidade de papéis a ser comprada é:

$$dep_{vi}^- = 1 + \sum_{j=0}^{v-1} den_{ji}^- \quad (4.16)$$

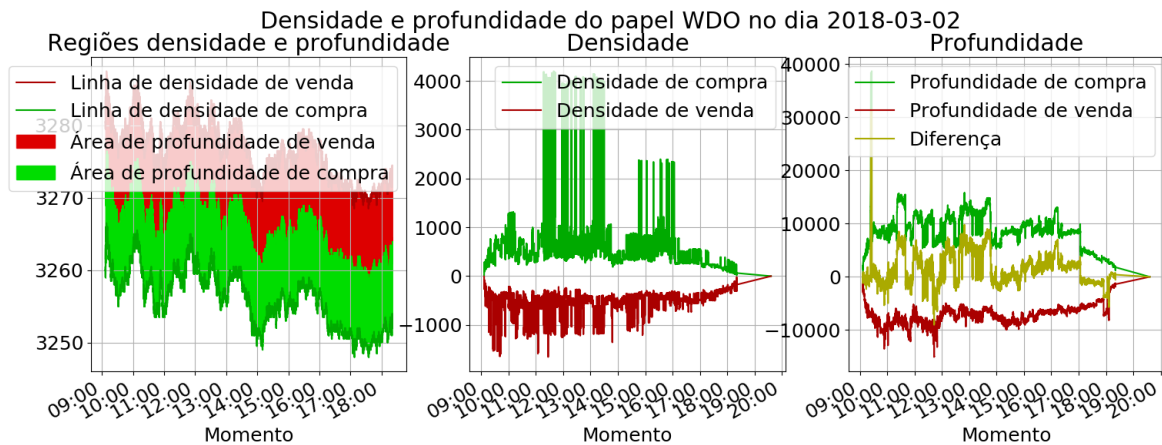


Figura 4.9. Os indicadores de densidade e profundidade. O valor v escolhido é definido de forma que a variação seja de 10 pontos acima e abaixo dos preço do papel.

A Figura 4.9 apresenta os indicadores de densidade e profundidade. Quando a *profundidade de compra* é superior à *profundidade de venda*, tem-se a intuição de que é mais provável que o preço varia positivamente. Isto significa que menos negócios são necessária para variar o preço positivamente, por meio de compras a mercado, do que negativamente, por meio de vendas a mercado. Além disso, no painel central existem picos de *densidade de compra*, mostrando-se que existem níveis de preço com um grande número de ordens posicionadas.

4.4 Construção e validação da estratégia

De forma a avaliar o potencial desses indicadores no cenário financeiro, constrói-se uma estratégia de investimento implementada por meio de um agente autônomo simples, utilizando-se o simulador descrito, os respectivos indicadores e a estratégia de rotulação. A estratégia de investimento utiliza um modelo gerado por aprendizado de máquina e regras de execução para operar. Dessa forma, trata-se de uma estratégia híbrida de investimento. A Figura 4.10 apresenta o processo geral de avaliação do agente autônomo de investimento.

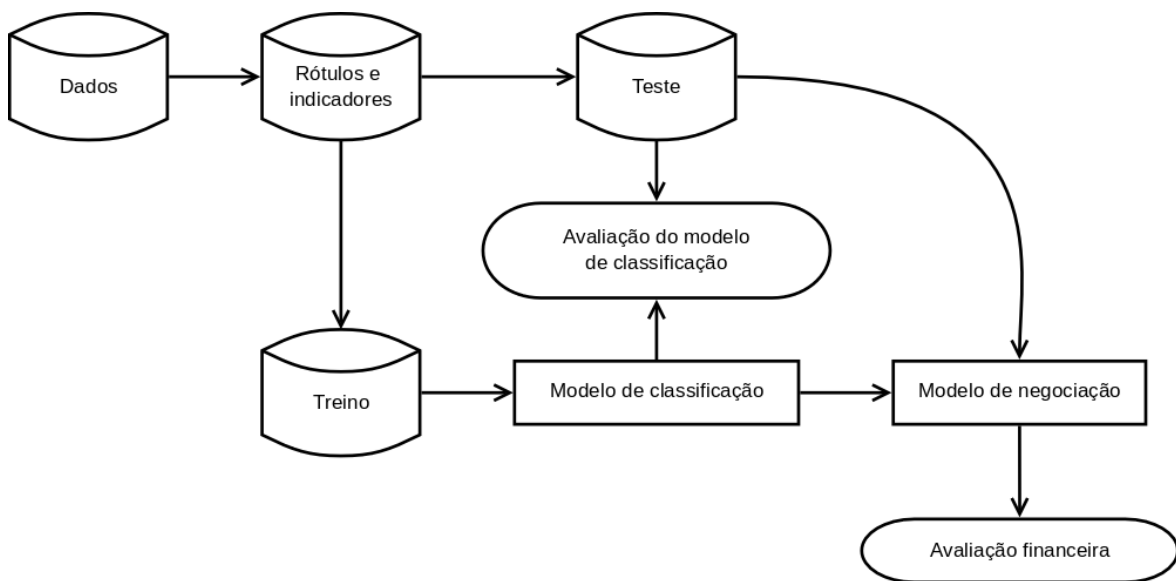


Figura 4.10. Processo de avaliação da estratégia de investimento.

4.4.1 Dados

A primeira etapa é processar os eventos do papel alvo. Esse processamento gera como resultado um conjunto de dados contendo os indicadores e os rótulos das cadeias de evento.

A partir do conjunto de dados produzido, deve-se separá-lo em dois subconjuntos. Um subconjunto de *treino* e um subconjunto de *teste*:

- **Subconjunto de treino:** Possui os dados que são utilizados para construir um modelo de classificação para a série de preços;
- **Subconjunto de teste:** Possui os dados que são utilizados para avaliar o modelo de classificação construído e então fazer a avaliação financeira do modelo de negociação.

Vale lembrar que a separação desses dois subconjuntos não pode ser arbitrária. Isto é, por se tratarem de séries temporais, deve-se tomar o cuidado de não levar informações dos dados de teste para os dados de treino por meio de normalizações ou outros ajustes. Além disso, é ideal que o conjunto de treino esteja em um momento temporal anterior ao de teste, buscando avaliar o poder preditivo dos indicadores e a generalização do modelo para séries futuras.

4.4.2 Modelo de classificação

A partir do dados de treino, constrói-se um classificador de séries. Esse classificador tem como saída o rótulo mais provável da série naquele momento, a partir dos indicadores dados como entrada.

Foram considerados três algoritmos de aprendizado de máquina para construção do modelo de classificação: *Logistic Regression* Cox [1958], *Support Vector Machines* Boser et al. [1992] e *Random Forest* Barandiaran [1998]. A *regressão logística* foi considerada por se trata de um método simples e de rápido treino. O *SVM* foi considerado por gerar modelos robustos, porém, dada a natureza do problema, não se conseguiu utilizá-lo de forma satisfatória. O principal obstáculo para utilizá-lo reside em sua complexidade de tempo quando tem-se um grande volume de dados de treino. Por fim, escolheu-se utilizar o *Random Forest* por ser eficiente na etapa de treino, mesmo com grandes volumes de dados, e ser robusto contra *overfitting*.

A avaliação do modelo de classificação gerado é feita a partir de *matrizes de confusão* e a métrica de *acurácia*:

- **Matriz de confusão** Powers [2011]: Constrói-se uma matriz comparando-se os rótulos estimados pelo modelo de classificação com os rótulos reais dos dados de teste. A partir dessa matriz é possível ver a performance geral do modelo construído;
- **Acurácia**: Corresponde à proporção correta de rótulos produzida pelo modelo de classificação. Pode ser derivada diretamente a partir da matriz de confusão. Utiliza-se essa métrica para avaliação do modelo de classificação pois ela está diretamente relacionada à estratégia de investimento construída, que opera nas duas direções possíveis, compra e venda.

Outros detalhes específicos à área de aprendizado de máquina não são discutidos por estarem fora do escopo da dissertação.

4.4.3 Estratégia

Com o modelo de classificação construído, o mesmo é utilizado como entrada para estratégia de investimento simples, implementada como um agente autônomo de negociação. Tal modelo possui três estados: **não posicionado**, **comprado** e **vendido**.

- **Não posicionado:** O agente não possui nenhum papel em sua posse. Esse estado é o estado inicial e final do agente;
- **Comprado:** O agente possui um papel em sua posse;
- **Vendido:** O agente deve um papel à bolsa de valores. Ele atinge esse estado por meio de uma *venda a descoberto*. É uma aposta contra o mercado, esperando que o preço do papel varie negativamente.

Considerando-se a saída do modelo de classificação e o estado corrente do agente autônomo, modelam-se as seguintes transições de estado:

- **(Não posicionado; Série crescente):** Um papel é comprado e o estado é definido como *comprado*;
- **(Não posicionado; Série decrescente):** Um papel é vendido e o estado é definido como *vendido*;
- **(Comprado; Série decrescente):** Dois papéis são vendidos e o estado é definido como *vendido*;
- **(Vendido; Série crescente):** Dois papéis são comprados e o estado é definido como *comprado*;
- **(Qualquer estado; Série neutra), (Comprado; Série crescente) e (Vendido; Série decrescente):** Mantém a posição e o estado.

Ao fim do período de negociação, o agente vai para o estado *não posicionado* vendendo-se seus papéis restantes ou comprando-se seus papéis devidos. Isso significa que o agente não fica posicionado entre dias de negociação diferentes.

A avaliação do modelo de negociação é feita em por meio de duas métricas, *taxa de acerto* e *ganho financeiro*. A taxa de acerto mede a quantidade de pares de compra e venda que acertaram a direção da variação de preço. O ganho financeiro avalia o ganho real do agente, podendo-se considerar custo sobre as transações. A diferença entre ambos é que por mais que um par tenha acertado a direção da variação de preço, isso não significa que o par de operações foi lucrativo.

Capítulo 5

Experimentos: Resultados e análise

Nesse capítulo é realizada a validação experimental do trabalho. Essa validação baseia-se nos conceitos apresentados nas seções 4.1, 4.2 e 4.3. Em seguida, aplica-se a estratégia de investimento descrita na seção 4.4. Além disso, são apresentados cenários que consideram custos de operação e latência nas simulações.

5.1 Dados

O papel escolhido para avaliação experimental é o mini-contrato futuro de dólar (*WDO*) da *B3*. Esse papel foi escolhido dada sua grande liquidez e seu potencial de retorno. Esse contrato varia de meio em meio ponto. Cada ponto nesse papel corresponde a dez reais e, portanto, sua variação mínima é de cinco reais.

O período de dados escolhido para os experimentos é o ano de 2018. Os dados foram repartidos da seguinte forma:

- **Janeiro e Fevereiro:** Base de dados para treino;
- **Março a Dezembro:** Base de dados para teste e avaliação financeira.

Para a base de dados de treino, utilizou-se uma técnica de amostragem que dá mais preferência para amostras mais próximas às zonas de transição de classe da série de preço. Ao fazer isso, diminuiu-se a acurácia geral do modelo de classificação, mas, ao mesmo tempo, as regras de rotulação ficaram baseadas em amostras de maior relevância, que vem a refletir diretamente nos resultados financeiros.

5.2 Modelo de classificação

Utilizando-se a base de dados de treino, construiu-se um modelo de classificação a partir do algoritmo *Random Forest*. A Tabela 5.1 apresenta a matriz de confusão gerada pela aplicação do modelo aos dados de teste entre os meses de Março e Abril. Apresentam-se duas versões da avaliação, uma aplicando-se o modelo à todas as amostras e, depois, às amostras mais relevantes, resultantes do processo de amostragem descrito:

	Todas as amostras		Amostras mais relevantes	
	Série crescente	Série decrescente	Série crescente	Série decrescente
SCE	631.204 (19,71%)	830.223 (25,92%)	9.034 (35,36%)	3.674 (14,38%)
SDE	1.052.244 (32,85%)	689.448 (21,52%)	3.468 (13,58%)	9.371 (36,68%)

Tabela 5.1. Matriz de confusão gerada a partir da base de testes nos meses de Março e Abril, considerando-se todas as amostras e, depois, apenas as amostras mais relevantes. SCE: Série crescente estimada. SDE: Série decrescente estimada

Vale lembrar que para esses experimento optou-se por utilizar a rotulação binária das séries de preço, ao invés da rotulação ternária, ou seja, utilizou-se os rótulos *crescente* e *decrescente* para as séries de preço, não utilizando-se os rótulos *neutros*. A Tabela 5.1 mostra que a acurácia geral do modelo gerado é de **41,23%**, enquanto que a acurácia para as amostras mais relevantes é de **72,04%**. A acurácia para as amostras mais relevantes é refletida diretamente nos resultados financeiros apresentados na seção seguinte.

5.3 Modelo de negociação

A avaliação financeira do agente autônomo construído é feita em diferentes cenários. Esses cenários são obtidos a partir da variação de dois parâmetros fundamentais:

- **Custo por ordem:** Valor nominal cobrado pelo registro de ordens. Considera-se nesse valor todos os custos relacionados a impostos e corretagem;
- **Latência:** Intervalo de tempo entre a ação do agente autônomo e confirmação da mesma.

A princípio, avalia-se o desempenho do agente em um cenário teórico, considerando-se custo zero e latência zero. Em seguida, avalia-se o agente considerando-se custo não nulo e latência zero. O terceiro cenário considera latência não nula e custo zero. Por fim, apresenta-se um estudo geral do desempenho do mesmo considerando os dois parâmetros não nulos.

5.3.1 Custo zero, latência nula

Esse cenário é completamente teórico, pois não leva em consideração nenhum dos parâmetros citados. Nesse cenários, estamos interessados no potencial bruto do agente autônomo.

As Figuras 5.1, 5.2 e 5.3 apresentam o capital acumulado em pontos em pontos do agente autônomo nos três primeiros dias de mercado de Março de 2018.

Resultado financeiro do papel WDO no dia 2018-03-01 - 14788 ordens corretas / 18322 ordens totais
Capital acumulado

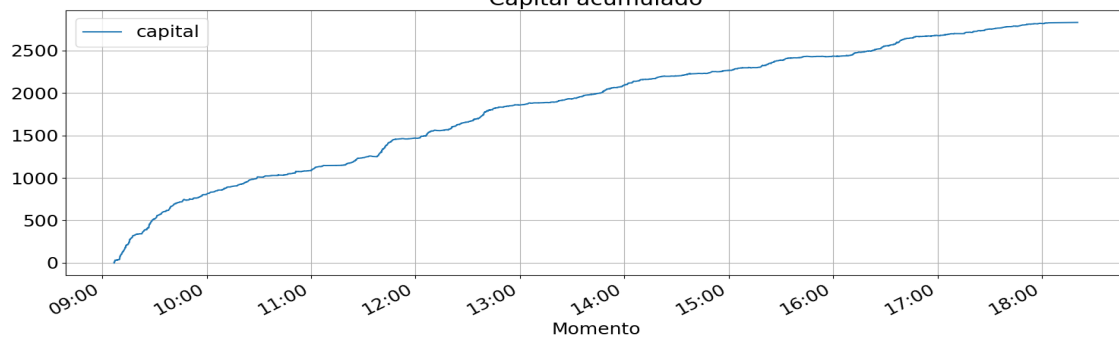


Figura 5.1. Capital acumulado ao longo do dia 01 de Março de 2018.

Resultado financeiro do papel WDO no dia 2018-03-02 - 12286 ordens corretas / 14780 ordens totais
Capital acumulado

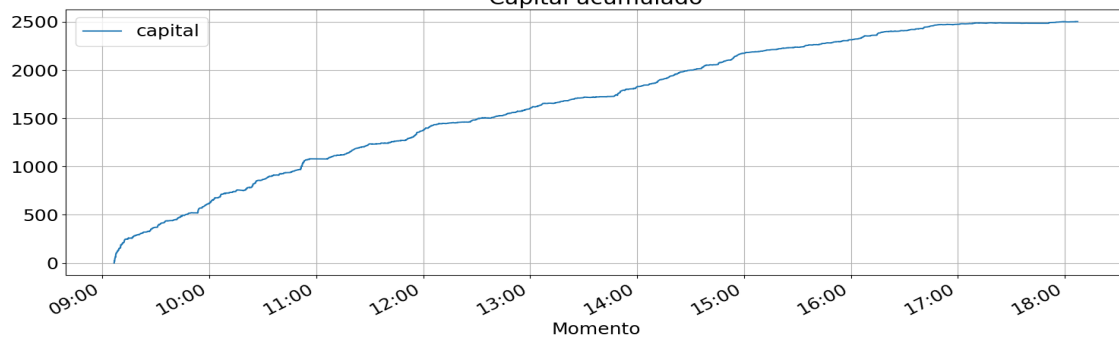


Figura 5.2. Capital acumulado ao longo do dia 02 de Março de 2018.

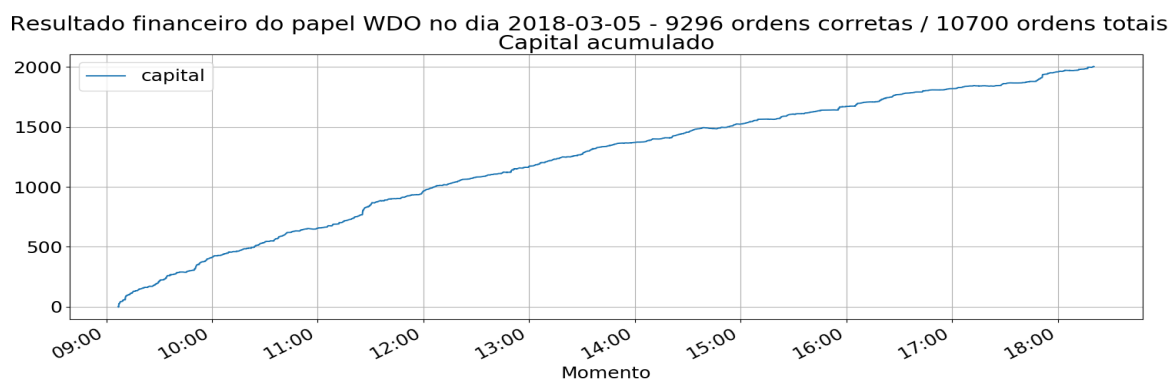


Figura 5.3. Capital acumulado ao longo do dia 05 de Março de 2018.

As figuras mostram um acúmulo positivo de capital do agente nesses dias. Uma ideia do potencial do mesmo é evidente. De forma a fazer uma análise mais profunda do mesmo, apresenta-se a Figura 5.4, que mostra o capital acumulado em pontos do agente para cada dia entre os meses de Março e Dezembro de 2018.

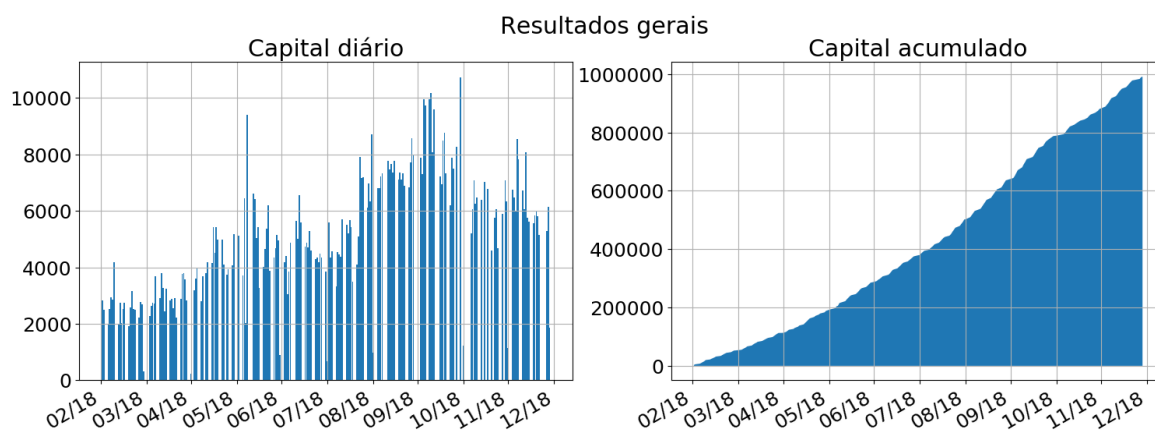


Figura 5.4. Capital diário e acumulado entre os meses de Março e Dezembro de 2018.

Por fim, a Tabela 5.2 apresenta os resultados sumarizados desse cenário.

Total de ordens	5.742.990
Média de ordens por dia	29.451
Total de ordens enviadas corretas	(82.3%) 4.766.519
Capital acumulado (em pontos)	991.336
Capital diário médio (em pontos)	5.083
Capital financeiro acumulado	R\$ 9.913.360,00
Capital financeiro diário médio	R\$ 50.837,74

Tabela 5.2. Resultados do cenário de custo zero e latência nula.

O resultado financeiro apresentado pela Tabela 5.2 é bastante positivo. Porém, como já foi mencionado, ele ocorre em um cenário completamente teórico e, portanto, não é factível na realidade.

5.3.2 Custo positivo, latência nula

Esse cenário é muito similar ao anterior, pois não altera nenhum comportamento referente à simulação do agente. Dessa forma, pode-se apenas aplicar as taxas de operação ao resultado final.

Total de ordens	5.742.990
Média de ordens por dia	29.451
Total de ordens corretas	(82.3%) 4.766.519
Capital acumulado (em pontos)	991.336
Capital diário médio (em pontos)	5.083
Capital financeiro acumulado	R\$ 4.170.370,00
Capital financeiro diário médio	R\$ 21.386,51

Tabela 5.3. Resultados do cenário de custo positivo e latência nula.

A Tabela 5.3 apresenta os mesmos resultados da Tabela 5.2, com exceção dos resultados financeiros finais. Considerou-se que o custo por ordem nesse cenário foi de R\$1,00 por ordem. É visível que, mesmo após a aplicação do custo de envio de ordens, o resultado financeiro continua bastante positivo. Porém, tal resultado é mais de duas vezes menor que o anterior. Desta forma, deve-se avaliar cenários mais realísticos, considerando-se latências para tais operações e, então, avaliar a factibilidade do agente no mundo real.

5.3.3 Custo zero, latência positiva

Nesse cenário, o agente é reavaliado considerando-se o tempo de rede necessário para o transporte das ordens enviadas por ele. Ignora-se os custos de envio de ordens nesse cenário de forma a melhor identificar o efeito da latência sobre o agente.

As Figuras 5.5, 5.6 e 5.7 apresentam o capital acumulado em pontos do agente autônomo nos três primeiros dias de mercado de Março de 2018 considerando diferentes latências. Como é apresentado nessas figuras, quanto maior a latência considerada, menor o capital final do agente.

A diferença do cenário teórico de zero latência para esse é significativa. Isso mostra a necessidade de considerar tais parâmetros de forma a melhor aproximar a

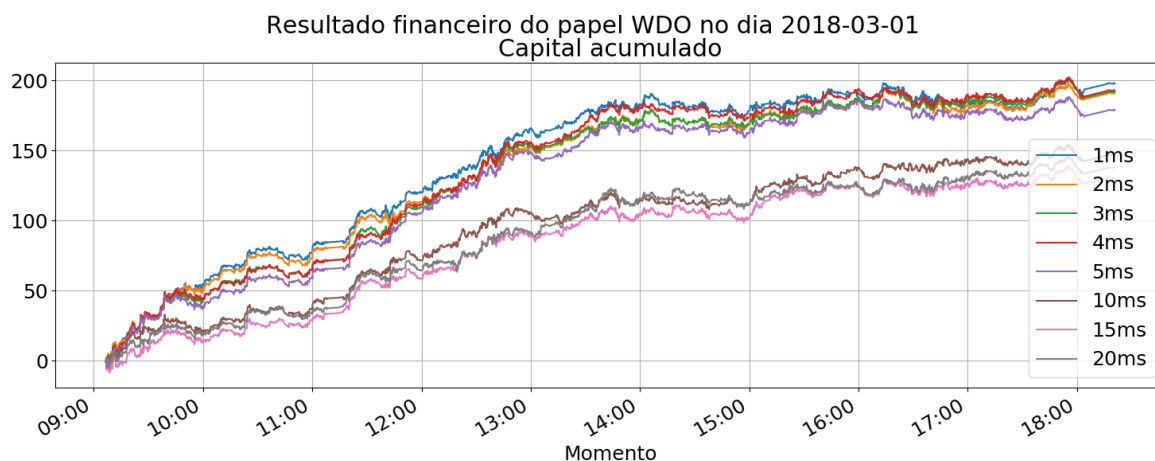


Figura 5.5. Capital acumulado ao longo do dia 01 de Março de 2018 para diferentes latências.

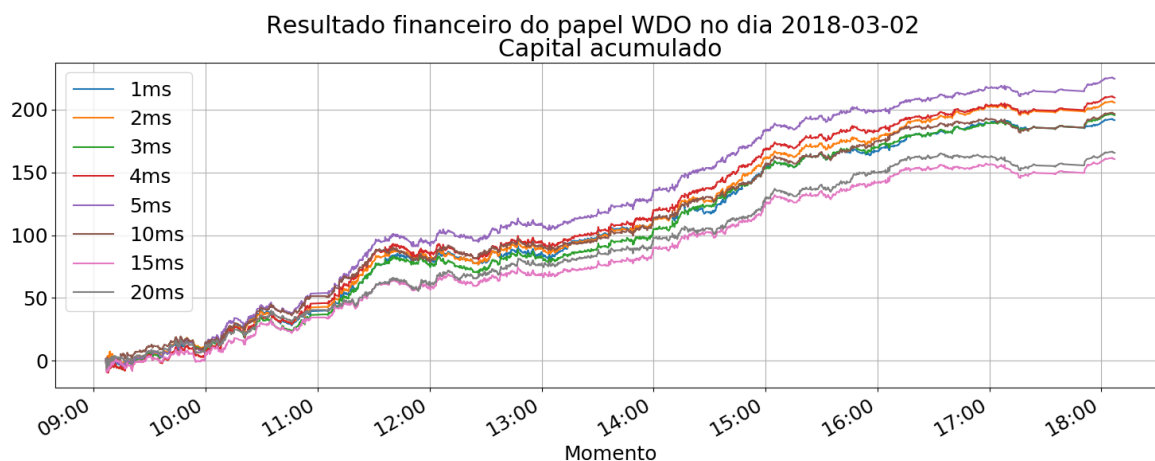


Figura 5.6. Capital acumulado ao longo do dia 02 de Março de 2018 para diferentes latências.

realidade. A Figura 5.8 apresenta o capital acumulado em pontos pelo agente ao longo dos meses de Março a Dezembro de 2018.

Por fim, a Tabela 5.4 apresenta os resultados sumarizados desse cenário. Os resultados apresentados demonstram a sensibilidade do agente autônomo à variações de latência. Em geral, quanto maior a latência, menor a performance geral do agente. Por fim, deve-se avaliar sua performance variando-se os parâmetros de custo e latência ao mesmo tempo.

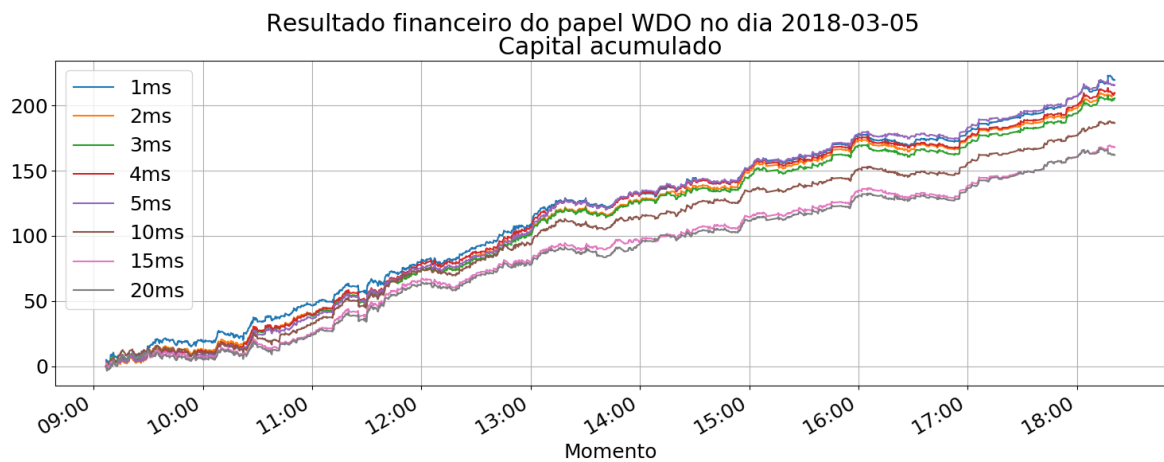


Figura 5.7. Capital acumulado ao longo do dia 05 de Março de 2018 para diferentes latências.

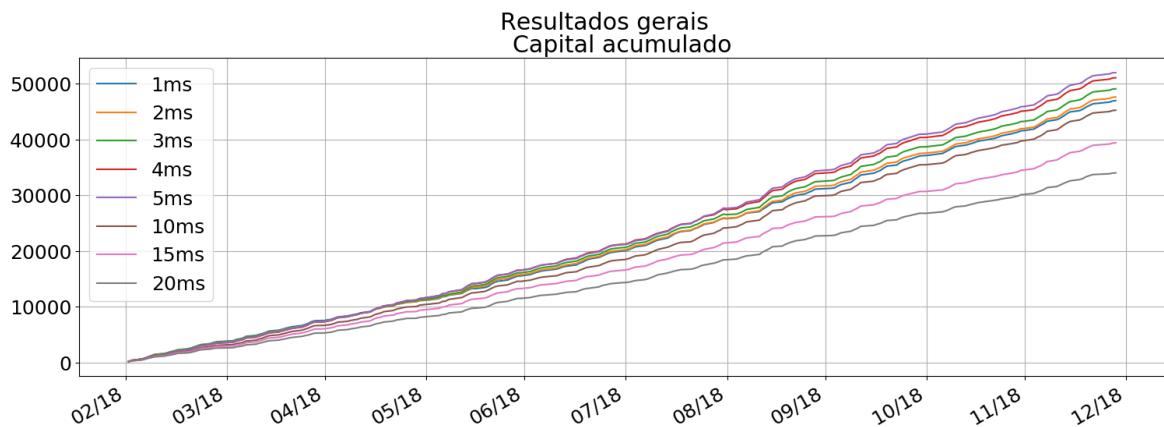


Figura 5.8. Capital acumulado entre os meses de Março e Dezembro de 2018 para diferentes latências.

5.3.4 Avaliação geral

De forma a avaliar a aplicabilidade do agente autônomo no mercado real, avaliam-se diferentes valores para os parâmetros já citados. A partir dessa avaliação é possível obter as condições necessárias para uma aplicação lucrativa do agente, se elas existem.

Os valores escolhidos para o parâmetro de latência estão definidos no intervalo de 1 a 20 milissegundos. Já os valores escolhidos para o custo por ordem estão definidos no intervalo de R\$0,00 a R\$1,00 de forma arbitrária. A partir disso, a Tabela 5.5 é construída, onde cada célula mostra o *capital financeiro diário médio*, sendo lucro quando positivo e prejuízo quando negativo.

A Tabela 5.5 mostra que o capital financeiro diário médio cai à medida que se aumenta o custo por ordem e a latência. Esse é um efeito esperado, dadas as avaliações

L	TO	MOD	TOC	CA	CDM	CFA	CFDM
1ms	3.348.606	17.172	2.765.285 (82.6%)	46.955	240	R\$ 469.550,00	R\$ 2.407,95
2ms	3.328.772	17.070	2.735.091 (82.2%)	47.591	244	R\$ 475.910,00	R\$ 2.440,56
3ms	3.315.948	17.004	2.710.793 (81.8%)	49.056	251	R\$ 490.560,00	R\$ 2.515,69
4ms	3.308.476	16.966	2.690.955 (81.3%)	51.029	261	R\$ 510.295,00	R\$ 2.616,90
5ms	3.300.780	16.927	2.670.998 (80.9%)	51.965	266	R\$ 519.655,00	R\$ 2.664,90
10ms	3.251.576	16.674	2.563.713 (78.8%)	45.226	231	R\$ 452.260,00	R\$ 2.319,28
15ms	3.217.860	16.501	2.470.362 (76.8%)	39.378	201	R\$ 393.780,00	R\$ 2.019,38
20ms	3.189.152	16.354	2.382.150 (74.7%)	33.989	174	R\$ 339.895,00	R\$ 1.743,05

Tabela 5.4. Resultados sumarizados: Latência (L), Total de Ordens (TO), Média de Ordens Diária (MOD), Total de Ordens Corretas (TOC), Capital Acumulado (CA), Capital Diário Médio (CDM), Capital Financeiro Acumulado (CFA) e Capital Financeiro Diário Médio (CFDM).

L/C	0,00	0,01	0,02	0,05	0,10	0,20	1,00
0ms	50.837,74	50.543,23	50.248,72	49.365,18	47.892,62	44.947,50	21.386,51
1ms	2.407,95	2.236,23	2.064,50	1.549,33	690,71	-1.026,52	-14.764,39
2ms	2.440,56	2.269,86	2.099,15	1.587,03	733,50	-973,56	-14.630,06
3ms	2.515,69	2.345,64	2.175,60	1.665,45	815,21	-885,28	-14.489,17
4ms	2.616,90	2.447,23	2.277,57	1.768,57	920,24	-776,41	-14.349,65
5ms	2.664,90	2.495,63	2.326,36	1.818,54	972,19	-720,52	-14.262,18
10ms	2.319,28	2.152,53	1.985,79	1.485,54	651,81	-1.015,67	-14.355,47
15ms	2.019,38	1.854,37	1.689,35	1.194,29	369,20	-1.280,98	-14.482,46
20ms	1.743,05	1.579,51	1.415,96	925,32	107,59	-1.527,87	-14.611,57

Tabela 5.5. Avaliação geral do agente para diferentes valores de custo e latência: Latência (L) e Custo (C). Cada célula apresenta o *CFDM* para a latência e custo definido por ela.

anteriores. É possível ver que, para latências não nulas, quando o custo por ordem é superior a R\$0,20, o agente só gera prejuízo. Por outro lado, considerando-se um custo por ordem de até R\$0,10, pode-se ter valores mais realísticos de latência, garantindo-se o lucro do agente. É possível ver que, para uma latência de 20 milissegundos e um custo de R\$0,10, o agente gera lucro, apesar de se tratar de um valor relativamente pequeno em relação às simulações teóricas.

Por fim, o agente autônomo construído é passível de aplicação no mundo real, desde que se mantenha as condições apresentadas. Tais condições são alcançáveis apenas em cenários de *high frequency trading (HFT)*, dependendo-se de grandes investimento em infraestrutura e negociação de custo por ordem com as corretoras a serem utilizadas.

Capítulo 6

Conclusão

Esse projeto foi desenvolvido buscando-se atingir três objetivos, construir um simulador verossímil de negociação na bolsa de valores, derivar indicadores financeiros e princípios de rotulação de séries de preço e, por fim, construir um agente autônomo de investimento utilizando tais indicadores. A conclusão desses objetivos produzem, por si só, algumas contribuições, descritas a seguir.

Em primeiro lugar, o objetivo de se construir um simulador, capaz de reproduzir fielmente o funcionamento de dias de negociações passados, obedecendo-se todas as regras explicitadas anteriormente, foi concluído com êxito. Com tal simulador, abre-se a possibilidade de estudar diversos aspectos de microestrutura de mercado e até simular cenários mais realísticos para outras estratégias de investimento. A partir do mesmo, também é possível gerar novas bases de dados para estudos na área.

Em segundo lugar, ao se derivar novos indicadores financeiros, abre-se a porta para melhor entender a dinâmica de movimentação de preço de diferentes papéis. Tais indicadores, junto à sua formulação, podem ser estendidos e ampliados para múltiplos objetivos. Além disso, ao se definir princípios básicos para rotular séries de preço, formaliza-se tal processo independentemente das características específicas do modelo de negociação formulado.

Por fim, a construção do agente autônomo a partir dos indicadores citados, valida experimentalmente suas capacidades preditivas para o problema em questão. A avaliação experimental gera o conjunto de cenários em que a aplicação de tal agente é viável. Além disso, o ferramental construído para essa avaliação pode ser utilizado para outros agentes e estratégias.

O projeto foi concluído com êxito dentro do escopo definido a princípio. Abordou-se diversos aspectos inerentes à mercados financeiros, focando-se em operações de alta frequência e detalhes de microestrutura de mercado. O agente construído tem poten-

cial de ser aplicado no mercado real, desde que se obedeça os requisitos explicitados. Melhoras nesse agente podem ampliar seus parâmetros de viabilidade.

Esse trabalho concluí que o uso de indicadores baseados em microestrutura de mercado têm grande potencial em melhorar estratégias de investimento e agentes autônomos. Além disso, o estudo mais detalhado de tal estrutura pode ampliar os conhecimentos na área financeira, isto é, tal estrutura possui mais informação que apenas agregados e preço e volume no tempo e, logo, representam uma nova dimensão para estudos em agente autônomos de investimento.

Como trabalhos futuros, alguns aspectos podem ser considerados de forma a melhorar os resultados apresentados nesse trabalho:

- **Rotulação das séries de preço:** Uma rotulação mais detalhada das séries de preço podem melhorar o desempenho desse agente. Tal rotulação pode tentar aproximar o modelo de classificação do modelo de negociação, de forma que a otimização do modelo de classificação signifique resultado financeiro positivo do agente;
- **Controle probabilístico dos gatilhos:** Além da classe emitida pelo modelo de classificação, pode-se usar sua probabilidade de estar correta para tomar decisões mais precisas durante o período de negociação;
- **Outros modelos de classificação:** Pode-se aplicar outras técnicas de aprendizado de máquina de forma a gerar modelos de classificação mais robustos para o agente autônomo.

Referências Bibliográficas

- Abergel, F.; Huré, C. & Pham, H. (2017). Algorithmic trading in a microstructural limit order book model. *arXiv preprint arXiv:1705.01446*.
- B3 (2019a). B3. http://www.b3.com.br/pt_br/. Acessado: 2019-05-01.
- B3 (2019b). Cvm aprova fusão entre bm&fbovespa e cetip. <https://br.investing.com/news/stock-market-news/cvm-aprova-fusao-entre-bmfbovespa-e-cetip-234503>. Acessado: 2019-05-01.
- B3 (2019c). Dados puros de negociação da b3. <ftp://ftp.bmf.com.br/MarketData/>. Acessado: 2019-05-10.
- B3 (2019d). Horários de negociação no mercado de ações. http://www.b3.com.br/pt_br/solucoes/plataformas/puma-trading-system/para-participantes-e-traders/horario-de-negociacao/acoes/. Acessado: 2019-05-10.
- B3 (2019e). Lista de corretoras da b3. http://www.b3.com.br/pt_br/produtos-e-servicos/participantes/busca-de-participantes/busca-de-corretoras/. Acessado: 2019-05-10.
- Barandiaran, I. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8).
- Barucca, P. & Lillo, F. (2017). Behind the price: On the role of agent's reflexivity in financial market microstructure. Em *Methods and Finance*, pp. 51--61. Springer.
- Baumol, W. J. (2007). Economic theory. *Measurement and ordinal utility, The New Encyclopædia Britannica*, 17:719.
- Bhattacharya, U. & Daouk, H. (2002). The world price of insider trading. *The Journal of Finance*, 57(1):75--108.

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.
- Boser, B. E.; Guyon, I. M. & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. Em *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144--152. ACM.
- Bouchaud, J.-P.; Farmer, J. D. & Lillo, F. (2009). How markets slowly digest changes in supply and demand. Em *Handbook of financial markets: dynamics and evolution*, pp. 57--160. Elsevier.
- Cartea, Á.; Donnelly, R. & Jaimungal, S. (2018). Enhancing trading strategies with order book signals. *Applied Mathematical Finance*, 25(1):1--35.
- Cartea, A. & Jaimungal, S. (2013). Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, 20(6):512--547.
- Cartea, A.; Jaimungal, S. & Ricci, J. (2014). Buy low, sell high: A high frequency trading perspective. *SIAM Journal on Financial Mathematics*, 5(1):415--444.
- Cesari, R.; Marzo, M. & Zagaglia, P. (2012). Effective trade execution.
- Cheng, G. (2007). *7 Winning Strategies For Trading Forex: Real and actionable techniques for profiting from the currency markets*. Harriman House Limited.
- Chiarella, C.; Iori, G. & Perelló, J. (2009). The impact of heterogeneous trading rules on the limit order book and order flows. *Journal of Economic Dynamics and Control*, 33(3):525--537.
- Chowdhury, R.; Rahman, M. A.; Rahman, M. S. & Mahdy, M. (2019). Predicting and forecasting the price of constituents and index of cryptocurrency using machine learning. *arXiv preprint arXiv:1905.08444*.
- Cont, R.; Kukanov, A. & Stoikov, S. (2014). The price impact of order book events. *Journal of financial econometrics*, 12(1):47--88.
- Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 215--242.
- CVM (2019). Comissão de valores mobiliários. <http://www.cvm.gov.br/>. Acessado em: 2019-05-10.
- Demsetz, H. (1968). The cost of transacting. *The quarterly journal of economics*, 82(1):33--53.

- Dixon, M. (2018). A high-frequency trade execution model for supervised learning. *High Frequency*, 1(1):32--52.
- Fletcher, T.; Hussain, Z. & Shawe-Taylor, J. (2010). Multiple kernel learning on the limit order book. Em *Proceedings of the First Workshop on Applications of Pattern Analysis*, pp. 167--174.
- Fruth, A.; Schöneborn, T. & Urusov, M. (2014). Optimal trade execution and price manipulation in order books with time-varying liquidity. *Mathematical Finance*, 24(4):651--695.
- Guilbaud, F. & Pham, H. (2013). Optimal high-frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79--94.
- Han, J.; Hong, J.; Sutardja, N. & Wong, S. F. (2015). Machine learning techniques for price change forecast using the limit order book data. Relatório técnico, Working Paper.
- Harris, L. (2003). *Trading and exchanges: Market microstructure for practitioners*. OUP USA.
- Huang, W.; Nakamori, Y. & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10):2513--2522.
- Jabbur, E.; Silva, E.; Castilho, D.; Pereira, A. & Brandao, H. (2014). Design and evaluation of automatic agents for stock market intraday trading. Em *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 03*, pp. 396--403. IEEE Computer Society.
- Jiang, Z. & Liang, J. (2017). Cryptocurrency portfolio management with deep reinforcement learning. Em *2017 Intelligent Systems Conference (IntelliSys)*, pp. 905--913. IEEE.
- Jo, U.; Jo, T.; Kim, W.; Yoon, I.; Lee, D. & Lee, S. (2019). Cooperative multi-agent reinforcement learning framework for scalping trading. *arXiv preprint arXiv:1904.00441*.
- Juchli, M. (2018). Limit order placement optimization with deep reinforcement learning: Learning from patterns in cryptocurrency market data.

- Kearns, M. & Nevmyvaka, Y. (2013). Machine learning for market microstructure and high frequency trading. *High Frequency Trading: New Realities for Traders, Markets, and Regulators*.
- Kirilenko, A.; Kyle, A. S.; Samadi, M. & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3):967--998.
- le Calvez, A. & Cliff, D. (2018). Deep learning can replicate adaptive traders in a limit-order-book financial market. Em *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1876--1883. IEEE.
- Lehalle, C.-A. (2013). Market microstructure knowledge needed for controlling an intra-day trading process. *arXiv preprint arXiv:1302.4592*.
- Lu, X. & Abergel, F. (2018). Order-book modelling and market making strategies. *arXiv preprint arXiv:1806.05101*.
- Madhavan, A. (2000). Market microstructure: A survey. *Journal of financial markets*, 3(3):205--258.
- Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1):77--91.
- O'hara, M. (1995). *Market microstructure theory*, volume 108. Blackwell Publishers Cambridge, MA.
- Patel, Y. (2018). Optimizing market making using multi-agent reinforcement learning. *arXiv preprint arXiv:1812.10252*.
- Poser, S. W. (2003). *Applying Elliot Wave theory profitably*, volume 169. John Wiley & Sons.
- Potters, M. & Bouchaud, J.-P. (2003). More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):133--140.
- Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation.
- Qiao, K. (2012). Market liquidity and convexity of order book (evidence from china). *arXiv preprint arXiv:1211.2078*.
- Qureshi, F. I. (2018). Investigating limit order book characteristics for short term price prediction: a machine learning approach. *arXiv preprint arXiv:1901.10534*.

- Radcliffe, R. C. (1997). *Investment: concepts, analysis, strategy*. Addison-Wesley Boston.
- Russell, S. J. & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.
- Silva, E.; Castilho, D.; Pereira, A. & Brandao, H. (2014). A neural network based approach to support the market making strategies in high-frequency trading. Em *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 845--852. IEEE.
- Spooner, T.; Fearnley, J.; Savani, R. & Koukorinis, A. (2018). Market making via reinforcement learning. Em *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 434--442. International Foundation for Autonomous Agents and Multiagent Systems.
- Varian, H. R. (1992). *Microeconomic analysis*, ww norton&company. Inc, New York, New York.
- Zaccaria, A.; Cristelli, M.; Alfi, V.; Ciulla, F. & Pietronero, L. (2010). Asymmetric statistics of order books: The role of discreteness and evidence for strategic order placement. *Physical Review E*, 81(6):066101.
- Zheng, B.; Moulines, E. & Abergel, F. (2012). Price jump prediction in limit order book. *arXiv preprint arXiv:1204.1381*.