MODELAGEM E APLICAÇÃO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA NEGOCIAÇÃO EM ALTA FREQUÊNCIA EM BOLSA DE VALORES

EVERTON JOSUÉ DA SILVA

MODELAGEM E APLICAÇÃO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA NEGOCIAÇÃO EM ALTA FREQUÊNCIA EM BOLSA DE VALORES

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

ORIENTADOR: ADRIANO CÉSAR MACHADO PEREIRA

Belo Horizonte

Maio de 2015

Silva, Everton Josué da.

S586e Modelagem e aplicação de técnicas de aprendizado de máquina para negociação em alta frequência em bolsa de valores / Everton Josué da Silva — Belo Horizonte, 2015.
xx, 46. : il. ; 29cm.

Dissertação (Mestrado) - Universidade Federal de Minas Gerais – Departamento de Ciência da Computação

Orientador: Adriano César Machado Pereira
Coorientador: Humberto César Brandão de Oliveira

1. Computação - Teses. 2. Modelagem de dados -Teses. 3. Aprendizado do computador – Teses 4. Bolsa de valores – Teses. I. Orientador. II Coorientador. III. Título.

519.6*82(043)

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

FOLHA DE APROVAÇÃO

Modelagem e aplicação de técnicas de aprendizado de máquina para negociação em alta frequência em bolsa de valores

EVERTON JOSUÉ DA SILVA

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

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

PROF. HUMBERTO CÉSAR BRANDÃO DE OLIVEIRA - Coorientador
Bacharelado em Ciência da Computação - UNIFAL

PROF. ARTHUR RODRIGO BOSCO DE MAGALHÃES
Departamento de Física e Matemática - CEFET

PROF. BRUNO PEREZ FERREIRA
Departamento de Ciências Administrativas - UFMG

Belo Horizonte, 29 de maio de 2015.

À Lucinda, Josué, Jéssica e Jéferson.

Agradecimentos

Agradeço primeiramente à Deus por me guiar, iluminar, me dar força e tranquilidade para alcançar meus objetivos.

Agradeço aos meus pais e irmãos por estarem ao meu lado em todos os momentos. Pelo incentivo e apoio em minhas escolhas e decisões.

Gostaria de agradecer ao meu orientador, Adriano Machado, pela sua constante ajuda e encorajamento ao longo do projeto. Por compartilhar sua sabedoria e conhecimento. Obrigado pela confiança depositada durante todo o desenvolvimento deste trabalho e por me ajudar a caminhar pelas novas vias de trabalho de pesquisa e publicações.

Quero expressar também meu profundo agradecimento aos meus amigos e sócios, Humberto e Douglas, pela sincera amizade e confiança. Obrigado por sempre poder contar com a ajuda de vocês nos momentos mais críticos. Por contribuírem para o meu crescimento profissional e por serem também um exemplo a ser seguido. A participação de vocês foi fundamental para a realização deste trabalho.

Agradeço aos meus amigos pelos momentos de distração e companheirismo ao longo desta caminhada.

Em especial, a minha minha afilhada Júlia e meu sobrinho Nícolas, anjos que Deus colocou na minha vida para alegrar meus finais de semana.

Resumo

Algoritmos de negociação (*algotrading*) têm desempenhado um papel importante no mercado de ações eletrônico. Entretanto, esses algoritmos, sem qualquer poder de previsão, não são seguros para realizar suas negociações. Neste contexto, a previsão do mercado de ações sempre foi um tema de pesquisa interessante entre os pesquisadores, principalmente devido ao potencial de ganho ao negociar ações e/ou para compreender as informações originadas dos dados do mercado de ações. Muitos algoritmos de aprendizado de máquina e modelos estatísticos têm sido propostos por pesquisadores para previsão do preço das ações, como também a previsão de movimento do preço das ações. Neste trabalho, desenvolvemos e implementamos um sistema de negociação, que inclui um gerador de sinal de tendência baseado em técnicas de aprendizado de máquina e uma estratégia direcional de negociação (*directional strategy*), a qual realiza suas operações ao identificar uma tendência de curto prazo. Depois de vários experimentos com diferentes técnicas de aprendizado de máquina, optamos por criar modelos utilizando redes neurais MultiLayer Perceptron (MLP) e um modelo *ensemble*, que combina duas MLPs, para prever a direção do preço das ações em um curto intervalo de tempo, mais especificamente, para prever uma tendência de alta. Esses modelos atuam como suporte ao algoritmo de negociação proposto neste trabalho. O algoritmo utiliza a saída do modelo para tomar decisões ao realizar suas negociações.

O principal objetivo deste trabalho foi modelar e usar técnicas de aprendizado de máquina para maximizar o retorno obtido pela estratégia de negociação. Utilizando um grande volume de dados (*tick data*), conduzimos o *back-testing* e simulação em um simulador realístico da Bolsa de Valores de São Paulo. Através dos resultados empíricos obtidos, mostramos que técnicas de aprendizado de máquina foram capazes de melhorar a eficácia nesse processo de tomada de decisão. Demonstramos que a precisão da previsão e resultados obtidos através da simulação realística são melhores com a abordagem *ensemble*.

Os resultados alcançados abriram novas oportunidades de pesquisa: 1) Aperfeiçoar os modelos de previsão para reduzir o número de falso-positivos. Essa redução impacta diretamente nos resultados financeiros obtidos em simulação, pois aumentará a taxa de acerto da estratégia de negociação; 2) Utilizar técnicas de aprendizado de máquina para dar suporte

a outros tipos de estratégias de negociação em alta frequência.

Palavras-chave: negociação em alta frequência, aprendizado de máquina, previsão do mercado de ações, algoritmos de negociação, processo de tomada de decisão.

Abstract

Algorithmic trading has performed an important rule in the electronic stock market. However, these algorithms, without any forecasting capability, are not safe to perform trading. In this context, the stock market prediction was always an interesting research topic for the researchers, mainly due to its capacity of making profit on stock trading and/or in order to understand the information originated from the stock market‘s data. Many machine learning algorithms and statistic models have been proposed by researchers for forecast of stock price and stock price movement. In this work, it was developed and implemented a trading system, that include a trend signal generator based on machine learning techniques and a directional trading strategy, which perform their operations by identifying a short term trend. After developing many experiments with different machine learning techniques, it was opted for developing models using neural networks called Multilayer Perceptron (MLP) and an ensemble model, which combine two MLPs, to predict uptrends. These models act as a support for the trading algorithm proposed by this work. The algorithm uses the model output for taking decisions on performing trading.

This work‘s main objective was to model and use machine learning techniques to maximize the trading strategy‘s return. Using a massive volume of tick data, it was conducted back-testing and simulation in a realistic simulator of the São Paulo‘s stock market. From the empirics results obtained, it was demonstrated that the machine learning techniques were capable of increasing the effectiveness of the decision making process. It was demonstrated that the prediction‘s precision and the results obtained from the realistic simulation are better with the ensemble approach.

The achieved results opened new research opportunities: 1) Improving the forecasting models to reduce the false-positive numbers. This reduction directly impacts on the financial results obtained in simulation, because it is going to increase the trading strategy hit rate; 2) Using machine learning techniques to support other high frequency trading strategies.

Keywords: high-frequency trading, machine learning; stock market prediction, algotrading, decision making process.

Lista de Figuras

Lista de Tabelas

Sumário

Capítulo 1

Introdução

A negociação online vem transformando o mundo dos investimentos a uma velocidade vertiginosa. Antes do advento dos computadores, as operações executadas pelos investidores eram realizadas por meios de controle estatísticos mais elementares, através de análises técnicas e fundamentalistas. Com o crescimento dos investimentos e negociações no mercado acionário, uma busca contínua por melhores ferramentas para prever com maior precisão o mercado de ações tornou-se necessária. Essas ferramentas têm como objetivo aumentar os lucros e diminuir os riscos dos investimentos. Atualmente, existem vários algoritmos de negociação que utilizam plataformas eletrônicas para inserir ou enviar ordens no mercado sem a intervenção humana [Lin, 2013].

Um tipo particular de algoritmo de negociação é a negociação em alta frequência (*High-Frequency Trading* - HFT). Esta categoria de negociação emergiu nas últimas décadas e recentemente tem ganhado atenção de pesquisadores acadêmicos. Estudos mostraram que HFT representam cerca de 50 à 75% de todo o volume negociado na NASDAQ OMX. O HFT é divido em quatro grandes classes: Arbitragem, Estratégia Direcional, Formadores de Mercado e estratégias de Detecção de Liquidez. O tipo mais comum de HFT é a execução de um processo chamado de formador de mercado (market making) [Hagströmer & Norden, 2013; Aldridge, 2013; Menkveld, 2013]. Um formador de mercado refere-se genericamente a estratégias de negociação que são responsáveis por definir preços de compra e venda, de acordo com suas políticas, para um ativo durante todo o período de negociação. Outras estratégias bastante utilizadas, são estratégias direcionais, que realizam suas operações através da identificação de tendências ou momentos de curto prazo. Esta classe de negociação em alta frequência incluem estratégias direcionadas a eventos e estratégias com base na previsão de movimentos de preços a curto prazo. No entanto, estratégias de HFT sem qualquer sinal de previsão não são seguras para realizar negociações.

Neste contexto, assim como outros processos de tomada de decisão, técnicas de apren-

dizado de máquina podem auxiliar os algoritmos de negociação em alta frequência. Uma fator importante na tomada de decisão em HFT é saber quando uma estratégia deve realizar suas operações no mercado de ações. Por exemplo, se o formador de mercado é configurado para executar uma operação rentável esperando um pequeno aumento no preço das ações, mas o preço vai para o lado oposto, então perdas podem ser acumuladas. Este é um problema muito desafiador porque envolve prever pequenas variações nos preços das ações em curtos períodos de tempo.

Neste trabalho, mostramos que técnicas de aprendizado de máquina podem melhorar a precisão das estratégias de HFT. Mais especificamente, redes neurais artificiais e um modelo híbrido entre as mesmas (*ensemble*), foram utilizadas para prever uma tendência de alta em um curto intervalo de tempo. Em outras palavras, nossa estratégia direcional de negociação em alta frequência realiza suas operações de compra e venda no mercado de ações com base na previsão feita pelos modelos de aprendizado de máquina.

1.1 Motivação

O processo de globalização resultou em um intenso intercâmbio entre países e com isso, cada vez mais o mercado acionário vem adquirindo uma crescente importância no cenário financeiro internacional. Neste contexto, os países em desenvolvimento procuram abrir sua economia para receber investimento externo. Além de ser um canal na captação de recursos que permitem o desenvolvimento de empresas, o mercado acionário também torna-se uma importante opção de investimento para pessoas e instituições.

Com o avanço da tecnologia, sugiram os algoritmos de negociação (*algo trading*). A automação de estratégias de negociação não é um processo recente. Algoritmos de negociação existem desde que os mercados se tornaram eletrônico em meados dos anos 80 [Aldridge, 2013]. Com o progresso tecnológico na última década, um novo campo de negociação foi caraterizado e bem definido: negociação em alta frequência (*High Frequency Trading* - HFT). Nesta categoria, algoritmos de negociação podem comprar e vender ações em menos de um milissegundo.

Esse segmento de algoritmos de negociação em alta frequência tem aumentado significativamente. Seguindo esta tendência, bolsas de valores e corretoras do mundo todo modernizaram seus sistemas e principalmente suas infraestruturas para suportar um crescimento sustentável da utilização desses algoritmos de negociação.

Nos EUA, os algoritmos de negociação em alta frequência são responsáveis por aproximadamente 65% de todo o volume negociado em bolsa. No Brasil, o HFT ainda está sendo difundido e vem cada vez mais ganhando força e relevância. Segundo os dados operacio-

nais da BM&FBOVESPA, o volume de negócios realizados via DMA (Acesso Direto ao Mercado) movimentou $R\$\ 104,5$ bilhões em fevereiro de 2012.

As estratégias de HFT tem desempenhado um papel importante no mercado de ações eletrônico. Entretanto, essas estratégias, sem qualquer poder de previsão, não são seguras para realizar negociações. Como em outros processos de tomada de decisão, técnicas de aprendizado de máquina podem ajudar os algoritmos de HFT. Uma importante decisão é saber quando o algoritmo de negociação precisa realizar suas operações. Este é um problema muito desafiador porque envolve prever pequenas tendências nos preços das ações em curtos períodos de tempo.

1.2 Objetivos

O principal objetivo deste trabalho é modelar, desenvolver e avaliar estratégias de negociação em alta frequência, mais especificamente estratégias direcionais de negociação. Como mencionado anteriormente, essas estratégias precisam de algum poder de previsão para ter sucesso em suas operações. Neste contexto, a ideia é explorar a capacidade de predição de diferentes técnicas abordadas para previsão de pequenas tendências de alta nos preços das ações. Em outras palavras, nosso algoritmo de negociação consulta os modelos gerados para identificar se um pequena tendência de alta ocorrerá. Caso seja identificada uma tendência, a estratégia de negociação inicia uma operação de compra e venda de ações. Com isso, espera-se melhorar a eficiência das negociações realizadas a fim de aumentar o lucro e reduzir o risco em suas operações.

1.3 Objetivos Específicos

- Pesquisar, implementar e analisar diversos indicadores relacionados a oscilação de preço das ações;

- Modelar, desenvolver e avaliar estratégias de negociação em alta frequência;

- Modelar, implementar e avaliar o uso de diferentes técnicas de aprendizado de máquina para decidir quando uma estratégia de negociação deve realizar suas negociações;

- Testar e avaliar as estratégias direcionais propostas em um simulador realístico a fim de obter resultados próximos de testes reais.

1.4 Contribuições Esperadas

- Mostrar que estratégias de HFT sem nenhum poder de previsão não suficientes para obter bons resultados.
- Mostrar que modelos de aprendizado de máquina criados podem melhorar a eficiências da operações realizadas pelas estratégias de HFT, reduzindo riscos e aumentando o lucro.
- Mostrar que o uso de um simulador realístico é um fator importante para testar e avaliar estratégias de HFT.

1.5 Organização do Trabalho

O restante do trabalho está organizado da seguinte maneira. O capítulo 2 apresenta alguns conceitos básicos e discute trabalhos relacionados. O capítulo 3 descreve a metologia aplicada neste trabalho, detalhando cada etapa a ser realizada para conclusão deste trabalho. O capítulo 4 descreve as configurações dos testes realizados e os resultado da simulação. O capítulo 5 apresenta as conclusões e trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Neste capítulo será apresentada uma visão geral dos principais conceitos fundamentais relacionados ao desenvolvimento deste trabalho. Na Seção 2.1 apresentamos uma introdução ao Mercado de Ações e a descrição dos termos mais comuns utilizados neste ambiente.

2.1 Introdução ao Mercado de Ações

O Mercado de Ações está diretamente relacionado com a economia de um país. Quanto mais desenvolvida é uma economia, mais ativo é seu mercado de capitais, o que se traduz em mais oportunidades para as pessoas, empresas e instituições aplicarem suas poupanças. Uma empresa abre seu capital para obter uma fonte de captação de recursos permanente. Isso acontece quando a empresa lança suas ações ao público, ou seja, emite ações e as negocia nas bolsas de valores. Uma ação é uma unidade de propriedade de uma empresa. Um maior número de ações conduzirá a uma maior participação acionária na empresa [Bovespa, 2008; Brokers, 2010].

O mercado de ações é uma instituição, um tipo de organização social com regras pública ou privada, onde pessoas negociam (compram e/ou vendem) ações, opções e contratos futuros, e commodities. Compradores e vendedores negociam entre si através de uma plataforma fornecida pela bolsas de valores, que são os mais importantes centros de negociação de ações. No Brasil, o mercado de ações é representado pela Bolsa de Valores de São Paulo (BOVESPA), a qual proporciona todas as condições e sistemas necessários para as negociações de compra e venda de ações de forma transparente [Bovespa, 2008; Brokers, 2010].

Os investidores compram ou vendem ações de acordo com suas análises. O preço das ações de uma empresa é determinado pela oferta e procura do ativo. Assim, se qualquer pessoa quer comprar, alguém tem que vender, e vice-versa. Nesse contexto, se existe um maior número de pessoas dispostas a comprar ações de uma empresa do que pessoas dispostas a

vender, o preço das ações irá subir. Da mesma forma, se um número existe um maior número de pessoas dispostas a vender do que comprar ações, o preço das ações irá cair. No Mercado de Ações, as pessoas fazem lucro comprando ações que elas acreditam que vai valer mais no futuro. Também podem obter lucro realizando um operação contrária, denominada venda a descoberto, que consiste na venda de um ativo financeiro ou derivativo que não se possui, esperando que seu preço caia para então comprá-lo de volta e lucrar na transação com a diferença [Mahato, 2014].

A negociação *online* vem transformando o mundo dos investimentos a uma velocidade vertiginosa. Com o crescimento dos investimentos e negociações no mercado acionário, uma busca contínua por melhores ferramentas para prever com maior precisão o mercado de ações tornou-se necessária. Essas ferramentas têm como objetivo aumentar os lucros e diminuir os riscos dos investimentos. Atualmente, existem vários algoritmos de negociação que utilizam plataformas eletrônicas para inserir ou enviar ordens no mercado sem a intervenção humana [Lin, 2013].

Na próxima seção apresentaremos os principais tipos de algoritmos de negociação, com destaque para os algoritmos de negociação em alta frequência. Apresentaremos também a importância das técnicas de aprendizado de máquina para a previsão do mercado de ações.

2.2 Algoritmos de Negociação em Alta Frequência e Previsão do Mercado de Ações

Duas das principais funções dos mercados financeiros são a de transmitir sinais de preços para a economia e a de permitir que agentes possam obter ganhos através de negociações no mercado. A inovação financeira é útil na medida em que aumenta a eficiência com que os mercados disponibilizam estas funções.

Negociação em alta frequência (*HFT*) é uma das recentes grandes inovações nos mercados financeiros. Estimou-se em 2010, pela consultoria *Tabb Group*, que HFT foram responsáveis por 56% do comércio de ações nos USA e 38% na Europa [Biais & Woolley, 2011].

Nos últimos anos, a negociação em alta frequência tem ganhado uma posição relevante nos mercados financeiros, habilitado e impulsionado por uma interação de medidas legislativas, pelo aumento da concorrência entre plataformas de execução e avanços significativos em tecnologia da informação. Em contraste com as estratégias de negociação tradicionais, operadores de alta frequência não têm como objetivo estabelecer e manter posições de longo prazo. Em vez disso, eles entram em posições de curto prazo e terminam o dia de negociação sem ficar posicionado para o próximo dia útil. Em outras palavras, os operadores de alta

frequência não deixam ativos (ações, mini-índices, contratos futuros, etc) acumulados para o próximo dia [Gomber et al., 2011].

HFT utiliza sofisticados programas de computador para analisar dados de mercado na busca de oportunidades de negociação. Estes programas mapeiam essas informações em estratégias de negociação e roteiam ordens de negociação ao mercado, tudo sem a intervenção direta de um humano. A velocidade desse processo é desconcertante. O latência entre a chegada de informações no computador e a execução de ordens é da ordem de milissegundos, muito mais rápido do que os seres humanos podem até mesmo registrar a informação inicial. Neste contexto, investidores de alta frequência competem por velocidade tanto quanto por computadores mais poderosos, conexões e programas. Entretanto, os operadores de alta frequência têm um custo adicional para usufruir destes privilégios e colocarem seus computadores o mais próximo possível do local de negociação [Biais & Woolley, 2011].

A evolução da tecnologia para apoiar esses algoritmos aconteceu em paralelo com o progresso das técnicas de inteligência computacional, durante os últimos 30 anos. Devido ao comportamento não-linear e volátil do mercado de ações, não é possível tomar decisões rentáveis e obter retornos de forma consistente sem prever corretamente o preço das ações. Neste contexto, precisamos de um modelo que prevê o mercado de ações com uma precisão satisfatória. Este problema pode resolvido através da aplicação de técnicas de inteligência artificial, como a mineração de dados ou aprendizagem de máquina [Mahato, 2014].

Como em outras áreas, um passo natural da comunidade científica foi investigar um grande número de técnicas para ajudar os sistemas de negociação na tomada de decisão ao realizar suas operações na bolsa de valores. Até o final da década de 90, um grande número de pesquisas foram conduzidas para tentar prever as séries temporais de preços, liquidez e volatilidade. A maioria delas discretizam a hora em intervalos diários, semanais, quinzenais ou mensais.

HFT é categoria de negociação que emergiu nas últimas décadas e recentemente tem ganhado atenção de pesquisadores acadêmicos. Apesar das divergências sobre a definição precisa do HFT, a maioria dos participantes do mercado concordam que as estratégias de HFT pertencem as quatro grandes classes seguintes:

1. Arbitragem (*Arbitrage*)

2. Estratégia direcional (*Directional strategy*)

3. Formadores de mercado (*Market making*)

4. Detecção de liquidez (*Liquidity detection*)

A estratégia de negociação utilizada neste trabalho trata-se de uma estratégia direcional. As estratégias direcionais identificam tendências ou momentos de curto prazo. Esta classe de negociação em alta frequência, inclui estratégias direcionadas a eventos e estratégias com base na previsão de movimentos de preços a curto prazo. Neste contexto, implementamos e avaliamos uma estratégia direcional de negociação juntamente com um modelo de previsão capaz de identificar tendências de curto prazo.

Capítulo 3

Trabalhos Relacionados

A automação de estratégias de negociação para o mercado financeiro não é um processo recente. Algoritmos de negociação (*Algo Trading*) existem desde que os mercados se tornaram eletrônicos, em meados dos anos 80 [Aldridge, 2013]. A Internet também desempenhou um papel importante na popularização de algoritmos de negociação. A evolução da tecnologia para apoiar os algoritmos de negociação aconteceu paralelamente com o avanço das técnicas de inteligência computacional, durante os últimos 30 anos. Como em outras áreas, uma etapa natural da comunidade científica foi investigar um grande número de técnicas para ajudar os sistemas a tomar decisões automáticas de investimentos nas bolsas de valores.

Com o progresso tecnológico na última década, um novo campo de negociação foi criado e bem definido: negociação em alta frequência (*High-Frequency Trading - HFT*). Nessa categoria, algoritmos de negociação podem comprar e vender ações em menos de um milissegundo. Isso é interessante por duas razões principais: *1)* Os algoritmos podem girar um grande volume de capital com pouco dinheiro em suas contas; *2)* Fechar posições em curtos intervalos de tempo, controlando facilmente e satisfatoriamente o risco. Isso acontece porque o valor em carteira é menos dependente das oscilações de preço das ações.

Todos os algoritmos de negociação em alta frequência são *day traders*, i.e, compram e vendem ações no mesmo dia. É uma modalidade de negociação utilizada em bolsas de valores, que tem como objetivo obter lucro com a oscilação de preço, ao longo do dia, de ativos financeiros. Mas *day traders* existem desde o início do mercado de ações em 1972 [Lukeman, 2000]. Os seres humanos têm especulado sobre preços desde o início da moeda, e essa tendência vai continuar, independentemente do estado do mercado. Além disso, antes do advento dos computadores, investidores do mercado de ações realizavam suas negociações principalmente na intuição. Com o crescimento dos investimentos e negociações no mercado acionário, uma busca contínua por melhores ferramentas para gerar sinais de negociação tornou-se necessária. Esses sinais geralmente são gerados por técnicas de inteligência

computacional, as quais são responsáveis por prever preços das ações, volatilidade, liquidez, tendência do mercado, entre outros.

Recentemente, vários modelos de previsão do mercado e técnicas de inteligência computacional têm sido desenvolvidas a fim de apoiar a decisão de investimentos em operações *intraday* no mercado de ações. Essas operações *intraday* consistem de estratégias de negociação de médio e longo prazo, assim como os algoritmo de negociação em alta frequência, como formadores de mercado. Para HFT, o campo de inteligência computacional tem sido investigado usando redes neurais artificiais (*Artificial Neural Networks - ANNs*) [Putra & Kosala, 2011] e sistemas *fuzzy* [Kablan & Ng, 2010; Kablan, 2009].

Regressão estatística, análise técnica, análise fundamentalista, análise de séries temporais, teoria do caos são algumas das técnicas que têm sido adotadas para prever tendências de mercado [Thirunavukarasu, 2009]. Contudo, estas técnicas não são capazes de gerar consistentemente uma previsão correta do mercado de ações, gerando muitas dúvidas entre os analistas sobre o uso de muitas dessas abordagens. Então, prever o movimento de preço no mercado de ações e tentar tomar decisões corretas de investimento é uma das principais necessidades e desafios neste contexto [Thalheimer & Ali, 1979].

O primeiro estudo significativo na aplicação de modelos de redes neurais para previsão do mercado financeiro foi publicado por *White* [White, 1988]. Ele apresentou alguns resultados usando técnicas de modelagem de aprendizado através de redes neurais para descobrir e decodificar regularidades não-lineares no movimento de preço de ações. O estudo foi realizado nas ações da *IBM* e com granularidade de retorno diário. Tendo que lidar com as principais características dos dados econômicos, a inferência estatística desempenhou um papel importante e alterações técnicas foram necessárias para o processo de aprendizagem. Depois desse estudo, vários esforços de pesquisa têm sido conduzidos para examinar a efetividade da previsão do mercado de ações utilizando redes neurais [de Oliveira et al., 2013; Niaki & Hoseinzade, 2013].

No trabalho [Kutsurelis, 1998], os autores investigaram o uso de redes neurais para prever a tendência futura do índice do mercado de ações. Um novo modelo para prever o preço dos ativos do Mercado de Ações de Singapura (*Singapore Stock Exchange - SES*) usando os ativos da *Singapore Airlines (SIA)* é apresentado em [Suan & Chye, 1998]. O modelo foi construído para prever o preço de fechamento da *SIA* dentro de uma semana, com base no conhecimento recente e histórico do preço máximo, mínimo e de fechamento e também o volume negociado. A camada de saída da rede neural tem apenas uma saída que armazena o valor do preço de fechamento fornecida por uma semana a partir do dia atual. A rede neural foi treinada usando aprendizado supervisionado através da regra delta generalizada. Eles verificaram que o preço de fechamento fornecido é muito perto do real. Dos 50 casos de teste, 47 (94%) apresentam erros absolutos menor que 5% e 35 previsões (70%)

mostram erros absolutos abaixo de 1%. O trabalho de [Adebiyi Ayodele et al., 2012] relata que a aplicação de redes neurais para prever as variáveis do mercado financeiro e o uso de análise técnica e fundamentalista é um fator predominante para a previsão do mercado de ações. O artigo apresenta um modelo híbrido que combina o uso de variáveis de análise técnica e fundamentalista como indicadores do mercado de ações para prever os preços futuros. Os resultados obtidos mostraram uma melhoria significativa em comparação ao uso apenas de variáveis da análise técnica. Em outro contexto, a abordagem de previsão também foi satisfatória como um guia para os comerciantes e investidores na tomada de decisões qualitativas. Em [McCluskey, 1993], os autores usam algoritmos para treinar modelos para prever o índice *Standard&Poors 500* e então compara os resultados usando uma abordagem de programação genética.

Chen & Pennock [2010] apresenta uma revisão da literatura sobre mecanismos de previsão, incluindo previsão de mercados e sistemas de previsão em pares (*peer prediction systems*). Eles se concentram no processo de design, com destaque para os objetivos e as propriedades que são importantes na concepção de bons mecanismos de previsão.

No contexto de *HFT*, a abordagem em [Putra & Kosala, 2011] usa um modelo de RNA que tem como entrada indicadores técnicos populares para prever sinais de negociação, que podem ser úteis para realizar negociações diárias. A base de dados utilizada para construir o modelo são *ticks* de alta frequência de ativos *intraday* de alguns setores da indústria da Bolsa de Valores da Indonésia. O estudo compara o desempenho do modelo com a estratégia simples de *buy-and-hold* e o perfil de lucro máximo. Os resultados experimentais mostraram que o modelo proposto tem um desempenho melhor do que a estratégia simples. Assim, o autor concluiu que a rede neural artificial é útil para gerar sinais de negociação para negociações *intra-day* a partir do conjunto de dados de alta frequência. Em [Kablan & Ng, 2010], uma nova abordagem para execução de ordens em *HFT* é proposta usando uma nova forma de análise dinâmica que faz uso de mecanismos lógica *fuzzy*. O algoritmo proposto também considera a volatilidade *intra-day* do mercado para minimizar os custos de negociação. O estudo relata que encontrar as melhores taxas de execução de ordens é um problema intrigante e para corretores de negociação de grandes ordens, o efeito do tamanho da ordem, a tendência do mercado e volatilidade são cruciais para a programação da ordem.

Em relação a algoritmos de formação de mercado, [Li et al., 2014] desenvolveram e implementaram um *framework* com duas camadas, que inclui um gerador de sinal de negociação baseado em uma abordagem de aprendizado supervisionado e uma estratégia de formação de mercado orientada e eventos. O gerador proposto utiliza informação da microestrutura do livro de ofertas e notícias do mercado para prever a tendência do mercado. Na segunda camada, a estratégia de formação de mercado realiza sua negociações com base nos sinais gerados pela primeira camada e evita-se de perdas de lucro conduzidas por tendências

do mercado. Os testes e simulação foram conduzidos em um simulador industrial, usando dados de *tick* [1] dos preços da bolsa de valores de Tokyo (*Tokyo Stock Exchange - TSE*) e da bolsa de valores de Shanghai (*Shanghai Stock Exchange - TSE*). Através de resultados empíricos, eles encontraram que: *1)* estratégias com sinais têm uma melhor performance que as estratégias sem nenhum sinal em termos de média diária de lucro e perda (Profit and Loss - PnL) e *sharpe ratio* (SR). Esta última é uma medida de relação de retorno e risco; *2)* Previsões corretas podem ajudar estratégias de formação de mercado a reajustar suas cotações juntamente com tendências do mercado, evitando que as estratégias ativem um procedimento de perda.

Todos esses trabalhos motivam nossa investigação e fornecem informações que mostram a relevância do nosso principal objetivo nessa pesquisa. Entretanto, existem algumas diferenças entre nosso trabalho e os artigos citados anteriormente:

1) A principal diferença, é que muitos dos trabalhos relacionados apenas criam modelos de previsão do mercado de ações, mas não têm uma estratégia ou algoritmo de negociação bem definido que utiliza as previsões dos modelos para realizar suas operações no mercado de ações.

2) É importante destacar que nossa estratégia de negociação foi projetada para o mercado de ações real. Utilizamos um simulador realístico da BOVESPA, ao invés de simulação com série de preços gerada sinteticamente. A simulação realística é um etapa essencial para testar e validar estratégias de negociação, principalmente quando se trata de estratégias de negociação em alta frequência.

3) Nosso trabalho também apresenta uma metologia completa, com etapas bem definidas e que pode ser aplicada para desenvolver, testar e validar qualquer estratégia de negociação baseada nas previsões realizadas por técnicas de aprendizado de máquina.

[1] Movimento ascendente ou descendente mínimo no preço de um título. Fonte: http://www.investopedia.com/terms/t/tick.asp

Capítulo 4

Metodologia

Este capítulo descreve a metologia aplicada neste trabalho, a qual é ilustrada pela Figura 4.1. Cada etapa será explicada nas próximas seções.

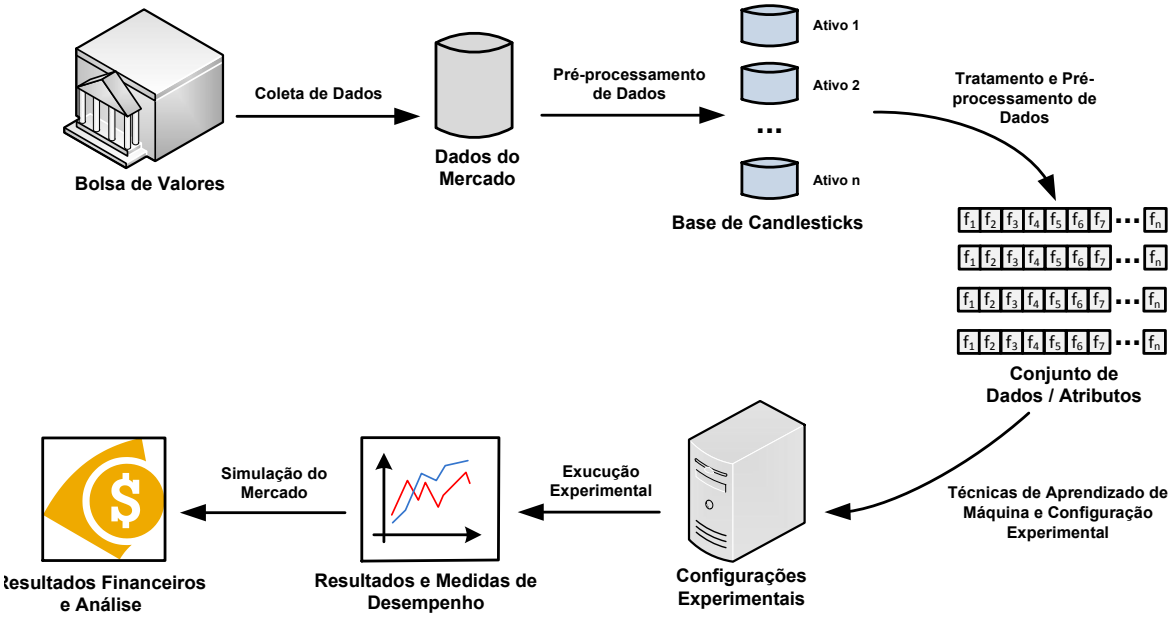

Figura 4.1: Metodologia

4.1 Coleta de Dados

O conjunto de dados utilizado neste trabalho consiste de dados reais da Bolsa de Valores de São Paulo (BM&FBOVESPA), que é a oitava maior bolsa do mundo. Esse conjunto de dados foi obtido de um fornecedor oficial (financial vendor), que provê dados para empresas

financeiras, *traders* [1] e investidores. Estes dados contém todas informações de cada evento ocorrido na BM&FBOVESPA, como por exemplo, se uma ordem de compra/venda foi enviada ao mercado, se uma ordem foi cancelada ou modificada, se uma negociação ocorreu, entre outras. Como mostrado na Figura 4.1, os dados coletados constituem nossa base de dados de mercado (Dados de Mercado).

4.2 Pré-processamento de Dados

Pré-processamento de dados envolve transformar dados brutos em um formato compreensível, preparando-os para posterior processamento. Neste contexto, os dados de mercado obtidos na etapa anterior foram pré-processados para gerar *candlesticks*. Um *candlestick* representa a variação dos preços de um determinado ativo em uma unidade de tempo (por exemplo, 15 minutos), como mostrado na Figura 4.2, onde:

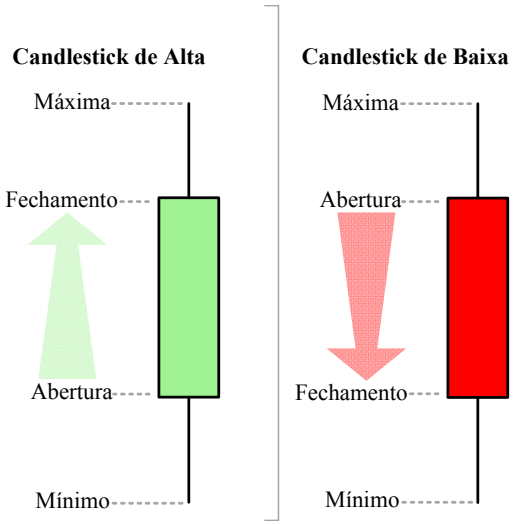

Figura 4.2: Representação de um *candlestick*.

- Abertura - o preço da primeira negociação no período;
- Fechamento - o preço da última negociação no período;
- Máximo - o preço máximo de todas as negociação no período;
- Mínimo - o preço mínimo de todas as negociação no período;

[1] Em finanças, um *trader* é uma pessoa ou empresa que compra e vende instrumentos financeiros como ações, mercadorias (*commodities*) e derivativos.

Nós usamos um método diferente para gerar os *candlesticks*, aplicando o conceito de janela deslizante (*rolling window*). Assim, os *candlesticks* não sintetizam um intervalo de tempo mutuamente exclusivos. Considere que a granularidade de um *candlestick* é, por exemplo, de 5 unidades de tempo. Em representações comuns, cada *candlestick* resume esse período de uma maneira estática e única, i.e, começa no tempo t_0 e compreende informações até o tempo t_4. Outro *candlestick* inicia em t_5 e finaliza em t_9, e assim por diante. Desse modo, eles não têm informações em comum com os anteriores, sendo mutuamente exclusivos. Usando o conceito de janela deslizante, aplicado também com granularidade de 5 unidades, o primeiro *candlestick* começa no tempo t_0 e termina no tempo t_4. Entretanto, o segundo inicia no tempo t_1 e finaliza no tempo t_5, o terceiro inicia no tempo t_2 e finaliza no tempo t_6, como mostrado na Figura 4.3.

Aplicando o pré-processamento, construímos uma base de dados de *candlesticks* para cada ativo (Figura 4.1 - Base de Dados de *Candlestick*). Os *candlesticks* foram processados e tratados para extração de atributos de dados, como mostrado na Seção seguinte.

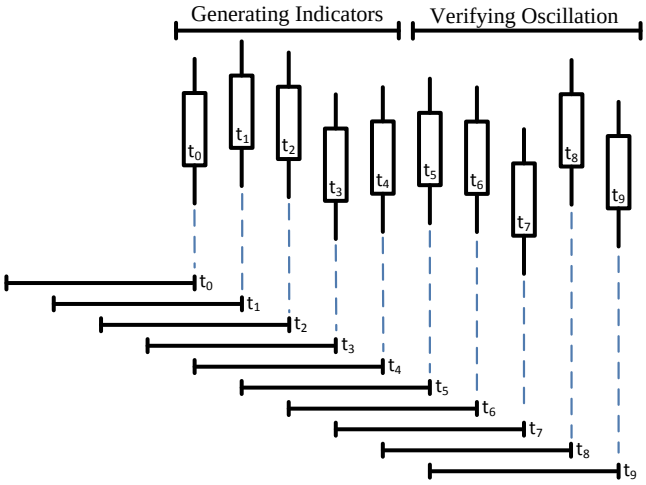

Figura 4.3: Construção dos modelos de previsão

4.3 Processamento e Tratamento de Dados

A fim de decidir os momentos para realizar um operação no mercado de ações, utilizamos técnicas de aprendizado de máquina para realizar a previsão de tendência de alta de um determinado ativo. Em outras palavras, nosso algoritmo de negociação consulta os modelos de previsão para saber se haverá um tendência de alta nos próximos t minutos. Caso houver, uma ordem limite de compra e venda de uma ação são enviadas simultaneamente ao mercado nos seguintes preços: x reais e $x + y$ reais, respectivamente, sendo y reais, o lucro que se

deseja obter. A partir desse momento, é esperado que o preço da ação oscile e ambas as ordens sejam executadas.

Para realizar as previsões, modelamos e construímos um conjunto de padrões que são usados como entradas para os modelos de previsão. A Figura 4.3 ilustra a criação dos modelos, que é dividido em duas etapas, geração de atributos e verificação de tendência, respectivamente. Os atributos geradas na primeira etapa são compostas de indicadores de análise técnica. Na segunda etapa é gerado o atributo binário de saída, 0 ou 1, que determina se a variação (positiva) de preço desejada irá ocorrer ou não. Em outras palavras, o valor 1 indica que o formador de mercado deve ser inicializado e 0, caso contrário. Para verificar se o preço da ação terá a variação de preço esperada, verificamos o *candlestick* que contém informações futuras. Por exemplo, considerando uma janela deslizante de 5 unidades de tempo, para verificar se uma operação deve ser realizada no tempo t_5, verificamos se a variação desejada ocorre no tempo t_9, pois o *candlestick* nesse instante contém informações de preço no intervalo de tempo entre t_5 a t_9. Essa verificação é executada como segue:

$$var = maximo(t_j) - abertura(t_i) \tag{4.1}$$

onde t_i e t_j são os instantes de tempo que uma operação é iniciado e finalizada, respectivamente, sendo t_j igual a $t_i + rolling_window$. $Abertura(t_i)$ é o preço inicial do *candlestick* no tempo t_i e $Maximo(t_j)$ é o maior preço obtido no intervalo de t_i até t_j. Finalmente, var contém a variação entre o preço máximo e o preço de abertura nesse intervalo.

A próxima Seção apresenta os indicadores de análise técnica.

4.3.1 Indicadores de Análise Técnica

A metodologia proposta neste trabalho considera a caracterização e análise de vários indicadores de análise técnica. Um indicador pode ser definido como uma série de pontos de dados, os quais são derivados a partir de informações de preço de uma ação aplicada a uma fórmula matemática. O conjunto de dados dos atributos de preços de cada ação é uma combinação dos preços de abertura, fechamento, máximo e mínimo ao longo de um período de tempo Thawornwong et al. [2001].

Analistas de mercado geralmente usam um ou mais indicadores para suas análises Thawornwong et al. [2001]. Esses indicadores são usualmente escolhidos avaliando a acurácia do modelo. Frequentemente, muitos indicadores são omitidos e um bom modelo pode nunca ser construído para um particular ativo, i.e, quanto mais informação tivermos, melhor será o resultado do modelo. Se os dados de entrada não são relevantes para a saída desejada,

provavelmente o modelo não irá aprender bem as associações entre os dados de entrada e saída. Assim, o primeiro passo é usar um conjunto de indicadores comumente usados em análise técnica tradicionais.

Atributos adequados têm de ser escolhidos para prever a direção do preço em um curto prazo. Este é um passo extremamente importante, pois atributos mal escolhidos podem demonstrar nenhuma previsibilidade ou simplesmente ignorar o bom senso. A fim de especificar os atributos utilizados neste trabalho, realizamos previamente uma caracterização geral de vários indicadores de análise técnica, com o objetivo de inferir qual deles tem potencial para compor nosso conjunto de atributos. Primeiramente selecionamos os indicadores mais utilizados como atributos para algoritmos de aprendizado de máquina. Em seguida, escolhemos os quais têm relação com a saída de nosso modelo, ou seja, aqueles que estão relacionados à velocidade e mudança do movimento de preço e indicam tendência ou direção do preço.

A seguir, descrevemos brevemente cada indicador que compõe nosso conjunto de atributos [2].

- *Relative Strength Index (RSI)*

  RSI é um indicador que mede a velocidade e mudança do movimento de preço. É calculado usando o valor da variação dos preços de alta e queda ao longo de um período de tempo especificado. A fórmula para computar o RSI é como segue:

$$RSI = 100 - \left(\frac{100}{1 + RS}\right) \tag{4.2}$$

$$RS = \frac{AG}{AL} \tag{4.3}$$

  onde RS é a média dos períodos de alta (AG - *Average Gain*) divido pela média dos períodos de baixa (AL - *Average Loss*). Um período de alta ou baixa é caracterizado quando o preço de um ativo sofre uma variação positiva.

- *Simple Moving Average (SMA)*

[2] Para mais informações sobre indicadores de análise técnica, ver http://www.stockcharts.com, http://www.investopedia.com and [Thawornwong et al., 2001]

Médias móveis simples suavizam os preços para formar um indicador de tendência. Eles não preveem a direção de preço, mas definem a direção. São úteis para a eliminação de ruído nos dados brutos, produzindo uma breve descrição das tendências.

$$SMA = \frac{\sum_{k=1}^{n} ClosingPrice_k}{n} \tag{4.4}$$

- *Exponential Moving Average (EMA)*

  A média móvel exponencial é uma extensão da média móvel simples, utilizando a suavização da mesma para diminuir a quantidade de sinais de compra ou venda. A média móvel exponencial é uma média ponderada de observações passadas e pode ser calculada através da seguinte fórmula:

$$EMA_x = EM(x-1) + K * ClosingPrice(x) - SMA(x-1) \tag{4.5}$$

  onde: EMA_x representa a média móvel exponencial no período x, $EM(x-1)$ representa a média móvel exponencial no período $x-1$, N é o número de períodos para os quais se quer o cálculo e uma contante $K = 2/(N+1)$.

  Média móvel exponencial reduz o atraso adicionando mais pesos ao preços recentes, sendo mais sensível aos valores mais recentes. Assim, na exponencial os dados mais novos possuem maior importância.

- *Moving Average Convergence/Divergence (MACD)*

  $MACD$ é um exemplo específico de oscilador no preço e é usado nos preços de fechamento de uma ação para detectar tendência, mostrando a relação entre duas médias móveis. Basicamente consiste de dois elementos: a linha $MACD$ e a linha de sinal. A linha $MACD$ é formada pela diferença entre duas médias móveis exponenciais (EMA), uma de curto e outra de longo prazo, geralmente sendo computadas com 12 e 26 períodos, respectivamente.

$$MACD = EMA[12] - EMA[26] \tag{4.6}$$

- *Average Directional Movement Index (ADX)*

O indicador ADX determina a força de uma tendência, podendo assumir valores entre 0 e 100. Valores baixos (< 20) indicam uma tendência fraca, enquanto valores altos (> 40) uma tendência forte. Entretanto, quando combinado com outros dois indicadores, *Plus Directional Indicator* ($+DI$) e *Minus Directional Indicator* ($-DI$), define a direção da tendência.

- *Aroon Indicator*

  É um indicador usado para identificar tendências de um ativo e a probabilidade de que essas tendências irão se reverter. Ela é composta por duas linhas: uma chamada de $Aroonup$, que mede a força da tendência de alta, e a outra denominada $Aroondown$, que mede a tendência de baixa. Essas linhas são definidas como segue:

$$Aroon_{Up} = \frac{N - MAX}{N} * 100 \tag{4.7}$$

$$Aroon_{Down} = \frac{N - MIN}{N} * 100 \tag{4.8}$$

  onde N é o número de períodos, MAX e MIN é o número de períodos desde o preço máximo e mínimo dos N períodos, respectivamente. Assim, $Aroon$ é calculado através desses dois indicadores:

$$Aroon_{indicator} = Aroon_{UP} - Aroon_{Down} \tag{4.9}$$

  Quanto mais forte é a tendência, ou seja, quanto mais alta encontra-se a linha ADX, maior a confiabilidade dos sinais de compra e venda.

- *Bollinger Bands*

  Bollinger Bands são bandas de volatilidade colocadas acima e abaixo de um média móvel. Esse indicador possui uma forte relação com a volatilidade. Assim, quando maior a volatilidade um ativo, maior seu desvio padrão. As bandas são constituídas por um conjunto de três curvas calculadas em relação aos preços. Elas são traçadas a partir de uma determinada distância de uma média móvel. Um exemplo do cálculo das bandas de *bollinger* pode ser visto a seguir:

$$UpperBand = SMA[20] + 2 * SD[20] \tag{4.10}$$

$$MiddleBand = SMA[20] \tag{4.11}$$

$$LowerBand = SMA[20] - 2 * SD[20] \tag{4.12}$$

onde SMA é a média móvel simples de 20 períodos, e SD é o desvio padrão.

- *Commodity Channel Index (CCI)*

  A CCI, do inglês *Commodity Channel Index*, é um indicador desenvolvido para identificar movimentos cíclicos. Ele assumi que o preço das ações sem movem em ciclos, com altas e baixas aparecendo em períodos de intervalos constantes. É um indicador versátil que pode ser usado para identificar uma nova tendência ou aviso de condições extremas. Em geral, CCI mede o nível de preço atual em relação a um nível médio de preços ao longo de um determinado período de tempo. CCI é relativamente alto quando os preços estão muito acima de sua média e é relativamente baixo quando os preços estão muito abaixo de sua média. É calculado como:

$$CCI = \frac{TP - SMA(TP)}{0.015 * SD(TP)} \tag{4.13}$$

$$TP = \frac{high + low + close}{3} \tag{4.14}$$

  onde SMA é a média móvel simples de 20 períodos, e SD é o desvio padrão. *High*, *low* e *close* referem-se ao preço máximo, mínimo e preço de fechamento que o ativo assumiu em um determinado período.

- *Chande Momentum Oscillator (CMO)*

  O CMO é um indicador de momento que foi desenvolvido por Tushar Chande e introduzido em seu livro, *The New Trader* de 1994. O CMO é projetado para medir o que Chande denomina de 'momento puro' e dados de retorno como uma linha que oscila entre $+100$ e -100. A seguir é apresentada a fórmula para calcular o CMO:

$$CMO = 100 * \frac{Up - Down}{Up + Down} \tag{4.15}$$

onde Up é a soma dos períodos de alta no período em análise e $Down$ é a soma dos períodos de baixa no período em análise.

- *Rate of Change (ROC)*

  O indicador ROC mostra a dinâmica de um ativo como uma porcentagem. É calculado subtraindo-se o preço de um número de períodos atrás, do preço atual, dividindo-se pelo preço de um número de períodos atrás, e depois multiplicar por 100 para obter uma porcentagem. É um indicador técnico simples que mostra a diferença percentual entre o preço atual e o preço de fechamento de N períodos anteriores. ROC é classificado como um indicador de dinâmica de preços ou um indicador de velocidade porque mede a taxa de mudança ou a força de impulso de mudança.

$$ROC = \frac{CP - CPA}{CPA} * 100 \tag{4.16}$$

onde CP é o preço de fechamento e CPA é o preço de fechamento de N períodos atrás.

Note que um indicador pode gerar mais de um atributo. Por exemplo, o indicador *Aroon* é construído por duas linhas, denominadas *aroon up* e *aroon down*. Uma mede a força de uma tendência de alta, e outra a força de uma tendência de baixa, respectivamente. Assim, cada padrão é composto de 26 atributos de entrada seguidos de um atributo de saída binário (0 ou 1). É importante deixar claro que os atributos passam por um processo de normalização. Esse processo transforma os atributos de entradas em valores dentro do intervalo de 0 a 1.

Com o processamento e tratamento da base de dados de *candlesticks*, geramos um conjunto de dados (veja Figura 4.1 - Conjunto de Dados/*Features*). Esse conjunto de dados é divido em conjunto de treinamento e teste para avaliar as técnicas de aprendizado de máquina utilizadas. A próxima Seção descreve como esse processo é realizado.

4.4 Técnicas de Aprendizado de Máquina e Configuração Experimental

Os experimentos foram realizados seguindo o conceito semelhante à janela deslizante utilizado no processamento de *candlesticks*. Separamos os dados, criados na etapa anterior (4.3), em conjunto de treinamento e teste, definidos da seguinte maneira: usamos um período que contém n dias para treinar os modelos, e testamos sua capacidade de previsão no próximo dia $(n+1)$. Um dia i compreende os padrões gerados usando os *candlesticks* exclusivamente desse dia.

A Figura 4.4 mostra como o conjunto de treinamento e teste são criados, utilizando uma janela deslizante de 4 dias. Seguindo este exemplo, a primeira configuração do experimento usaria os dias D_0, D_1, D_2 e D_3 para treinar um modelo, e o dia D_4 como conjunto de teste. O tamanho da janela N é fixado e a janela se move para a direita um dia de cada vez, enquanto um dia do lado esquerdo é descartado. Para cada deslocamento da janela, um novo modelo é gerado para a previsão do dia seguinte. Em nossos experimentos, usamos diferentes tamanhos de janelas deslizantes para os ensaios de *backtesting*. Assim, diferentes modelos foram gerados com essas configurações e nós avaliamos a efetividade de cada um usando algumas medidas de desempenho. A próxima Seção apresenta uma breve descrição das técnicas de aprendizado de máquina utilizadas neste trabalho.

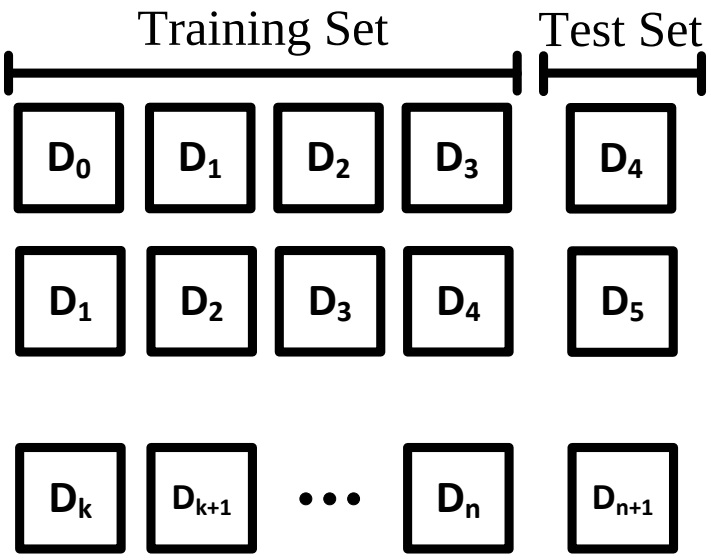

Figura 4.4: Conjunto de treinamento e teste

4.4.1 Técnicas de Aprendizado de Máquina

A fim de verificar se o preço da ação apresenta a variação desejada em um intervalo de tempo, utilizamos duas técnicas de aprendizado de máquina.

- Multilayer Perceptron - MLP: consiste de várias camadas de elementos simples (ou dois estados) de processamento sigmoidal, ou neurônios que interagem usando conexões ponderadas [Fahlman & Hinton, 1987]. Depois de uma camada de entrada, há normalmente qualquer número de camadas intermediárias, ou escondidas, seguida por uma camada de saída [Pal & Mitra, 1992; Rosenblatt, 1961]. MLP utiliza uma técnica de aprendizado supervisionado chamado *backpropagation* para treinamento da rede, é uma modificação da perceptron linear padrão e pode distinguir os dados que não são linearmente separáveis. Tem a vantagem de ser um método simples, em termos de complexidade computacional. A Figura 4.5 mostra a arquitetura de uma rede neural MLP com uma camada de entrada, 2 camadas escondidas, e uma camada de saída. Neste trabalho, as entradas (X_m) são os diferentes indicadores de análise técnica. A saída y_p pode assumir os valores 1 ou 0, que define a saída do modelo. O valor 1 indica que o preço da ação terá uma variação positiva maior que R$ x, e 0 o caso contrário.

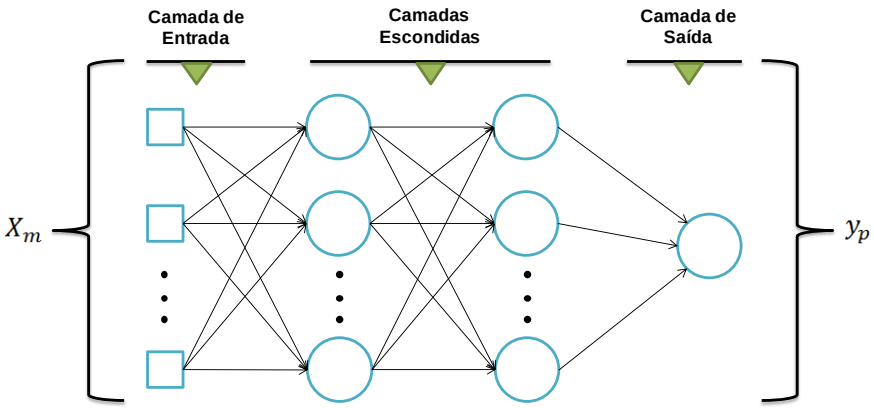

Figura 4.5: Arquitetura de uma rede neural *multilayer perceptron* (MLP) com duas camadas escondidas

- *Ensemble*: um classificador *ensemble* é um conjunto de classificadores cujas decisões individuais são combinadas de alguma forma (normalmente por votação ponderada ou não ponderada) para classificar novos exemplos. Métodos *ensemble* usam essa combinação para melhorar o desempenho preditivo que poderia ser obtido de qualquer um dos classificadores individuais [Dietterich, 2000]. Muitos métodos para construir *ensembles* têm sido desenvolvidos, os quais podem ser aplicados a diferentes algoritmos

de aprendizagem. Neste trabalho, nós construímos um *ensemble* combinando as decisões de duas MLP. A saída do *ensemble* será 1, se e somente se a saída de ambos modelos pertence a classe 1. Em todos os outros casos, a saída do *ensemble* será 0.

4.4.2 Execução Experimental

Este passo consiste em executar todas as configurações experimentais geradas na fase anterior. O resultado de cada configuração experimental foi avaliado utilizando-se diferentes medidas de desempenho.

4.4.2.1 Medidas de Desempenho para Classificação

Diferentes medidas de desempenho estão disponíveis para avaliar a efetividade de um classificador. Em um problema de classificação, um classificador rotula exemplos como positivo ou negativo. A decisão tomada pelo classificador pode ser representada em uma estrutura conhecida como matriz de confusão ou tabela de contingência [Davis & Goadrich, 2006]. A matriz de confusão é frequentemente utilizada para organizar e exibir informações utilizadas para avaliar o desempenho de um algoritmo, geralmente um algoritmo de aprendizagem supervisionada. Cada coluna da matriz representa as instâncias de uma classe prevista, enquanto cada linha representa os casos em uma classe real [Stehman, 1997]. Em uma classificação binária, a matriz de confusão tem duas linhas e duas colunas, como mostrado na Tabela 4.1.

Tabela 4.1: Matriz de Confusão

		Valor Previsto		
		p	**n**	**Total**
Valor Real	**p′**	Verdadeiro Positivo	Falso Negativo	P′
	n′	Falso Positivo	Verdadeiro Negativo	N′
	Total	P	N	

Verdadeiros positivos (VP) são exemplos rotulados corretamente como positivos, falsos positivos (FP) referem-se a exemplos negativos incorretamente rotulados como positivo,

verdadeiros negativos (VN) correspondem a negativos rotulados corretamente como negativos e falsos negativos (FN) referem-se a exemplos positivos incorretamente rotulados como negativo [Davis & Goadrich, 2006].

Neste trabalho nós usamos diferentes medidas de desempenho para valiar os classificadores, isto é, sua capacidade de tomar as decisões corretas de classificação. As principais medidas de desempenho utilizadas neste trabalho foram: verdadeiro positivo (VP), falso negativo (FN), falso positivo (FV), verdadeiro negativo (VN) e precisão. Mais informações sobre medidas de desempenho podem ser encontradas em [Sokolova & Lapalme, 2009]. A seguir será apresentada a importância de cada uma delas no contexto deste trabalho.

VP são os *triggers* (execução de um formador de mercado) que foram corretamente classificados, isto é, as ordens de compra e venda foram executadas. FP é o número de *triggers* que foram classificados incorretamente. FN representa os *triggers* que existem mas não foram identificados. Neste caso, perdemos a oportunidade de iniciar um formador de mercado. Por último, VN é quando não exite um *trigger* e o classificador identificou isso corretamente. Entretanto, nós não avaliamos essa métrica porque estamos interessados em cenários que proveem algum lucro. Neste contexto, das várias medidas de performance listadas, damos maior importância a precisão, porque ela estabelece a relação entre VP e FP. A precisão (P) de uma classe x é a razão entre o número de exemplos classificados corretamente (VP) e o total de exemplos previstos para a classe x (VP + FP):

$$P = \frac{VP}{VP + FP} \tag{4.17}$$

Os resultados dos experimentos são apresentados na Capítulo 5. Para cada dia de teste, uma matriz de confusão é gerada. Além das informações presentes na matriz de confusão também é calculada a precisão de cada classificador. A fim de analisar e discutir os resultados, tabelas e gráficos CDF (*cumulative distribution function*) são sintetizados com os resultados obtidos. O principal objetivo dessa análise e discussão é identificar o melhor classificador, assim como a melhor configuração: tamanho da janela deslizante de treinamento e a janela de tempo da previsão. Assim, usamos um simulador HFT realístico para avaliar o resultado financeiro obtido das previsões de cada técnica de aprendizado de máquina utilizada neste trabalho.

A próxima Seção descreve o simulador realístico utilizado neste trabalho para testar e avaliar a estratégia de formação de mercado proposta.

4.5 Simulação do Mercado

Para estudar, entender e avaliar a qualidade do uso da estratégias de negociação em alta frequência na BOVESPA, e analisar as hipóteses propostas neste trabalho, o mercado de ações, a corretora, a estrutura de comunicação, e outros componentes chave foram emulados usando uma técnica chamada Simulação Evento-Discreto (SED). Esse simulador nos permite executar testes sem precisar comprar e vender ações no mercado real.

O mecanismo geral do simulador utilizado neste trabalho, Figura 4.6, é o mesmo utilizado em [Oliveira, 2012]. O simulador possui uma estrutura (uma fila de prioridade) que armazena seus eventos de forma ordenada em função da hora de acontecimento. Assim, o evento que possui a maior prioridade para sair da fila e ser processado, é aquele que possui o menor instante de chegada.

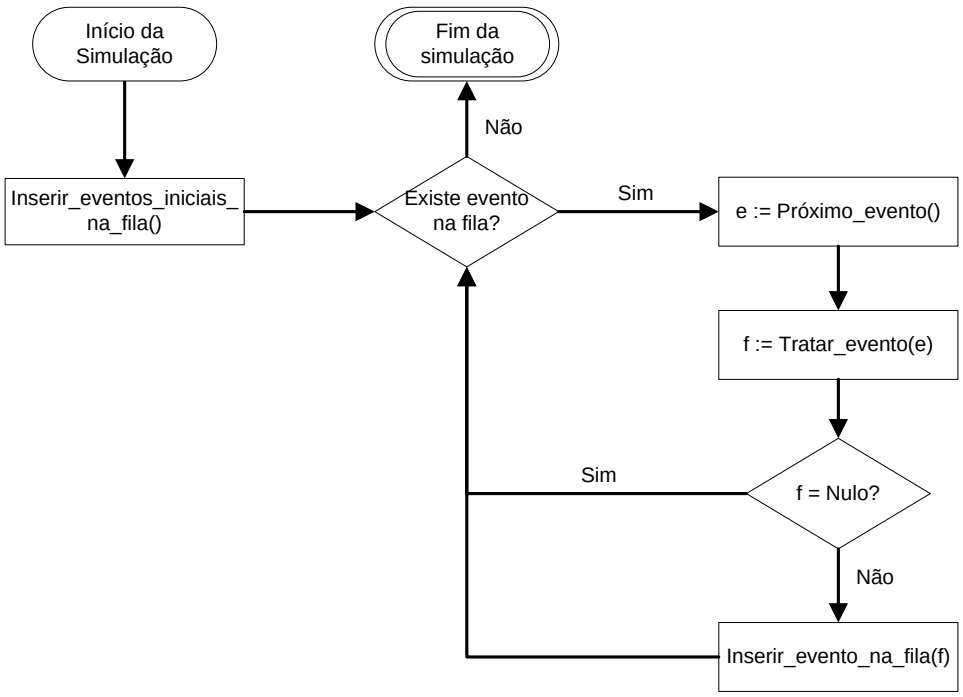

Figura 4.6: Fluxograma simplificado da Simulação Evento-Discreto [Oliveira, 2012]

Ao longo do processo de simulação, novos eventos podem ser inseridos na fila, que mantém seus elementos ordenados. Sempre o próximo evento a ser tratado pelo simulador, é o próximo evento na linha do tempo. Este mecanismo permite considerar o avanço do tempo em uma velocidade não constante, mas nunca permitindo o processamento de um evento antes de um evento antecessor, impossibilitando a geração de inconsistências no estado do objeto da simulação [Oliveira, 2012].

A função Tratar Evento pode alterar o estado do objeto da simulação. Eventos tratados pelo simulador podem gerar, dinamicamente, novos eventos para o ambiente, por exemplo:

- Uma estratégia de negociação recebe uma informação de que uma oferta (compra ou venda) de um cliente qualquer foi enviada a bolsa. Com base nessa informação, a estratégia pode tomar uma decisão que gera outros eventos, como, por exemplo, enviar uma oferta de compra e uma venda ao mercado.

- Uma estratégia de negociação pode receber uma notificação de que sua oferta enviada ao mercado foi rejeitada. Assim, a estratégia pode desistir de efetuar aquela oferta ou realizar uma nova tentativa de oferta ao mercado.

O simulador considera todos eventos recebidos pela BM&FBOVESPA em um dia de negociação. Ou seja, a instância da simulação contém todos os eventos que aconteceram no mercado real de ações. Inicialmente, todos os eventos da instância (eventos estáticos), são inseridos na fila de prioridades do SED. Os eventos dinâmicos surgirão em função da reação do algoritmo de negociação em análise. O algoritmo é informado, com atraso temporal, sobre todos os eventos que o mercado distribuiu. Sempre que a estratégia visualizar o estado do objeto do mercado, este estado pode ser a representação do passado do mercado de ações em função do atraso na comunicação. Esse mecanismo de atraso é importante, pois no mundo real isso sempre acontece. Assim, um algoritmo de negociação em alta frequência possui as mesmas informações que seriam obtidas no mercado real.

Qualquer estratégia, em avaliação, ao processar cada evento ocorrido no mercado, pode, como descrito na Figura 4.6, enviar eventos ao mercado simulado.

O simulador é baseado em dados *intra-day*, capaz de receber e executar ordens e também calcular a prioridade de uma ordem no livro de ofertas que não tenha sido executada ou cancelada. O sistema aceita ordens limite, ordens a mercado e cancelamentos.

Esse processo de simulação realística é de suma importância para testar e avaliar nosso algoritmo de negociação, além de validar as previsões realizadas pelos modelos de aprendizado de máquina. Neste contexto, os melhores modelos de previsão foram usados para identificar os momentos que a estratégia deve realizar suas operações no simulador, e, como resultado, vamos apresentar uma análise do retorno financeiro durante o nosso período de teste.

Capítulo 5

Resultados Experimentais

Este capítulo apresenta os resultados obtidos com a aplicação da metodologia previamente apresentada. Os dados usados nos experimentos consistem de um conjunto de dados reais da Bolsa de Valores de São Paulo (BM&FBOVESPA). Esses dados contém todas informações de 51 dias negociação, de Janeiro a Abril de 2015, incluindo ordens, negociações, alterações no livro de ofertas e todos os outros dados entregues pelo protocolo *FIX*[1]. É importante esclarecer que o número de dias é considerado satisfatório para a análise experimental, pois o volume de dados é grande, considerando que o problema é sobre HFT (*High-Frequency Trading*), podendo gerar milhares de padrões com poucos dias no conjunto de dados [Aldridge, 2013].

Os ativos das empresas analisadas neste trabalho são: Itaú Unibanco Holding S.A. (ITUB4), Petróleo Brasileiro S.A. - Petrobras (PETR4) e Vale S.A. (VALE5). Estas empresas fazem parte de um grupo seleto ativos da BM&FBOVESPA, chamados Índice Bovespa (iBov ou IBovespa)[2]. O Índice Bovespa é composto por uma carteira teórica com as ações que representam 80% do volume negociado nos últimos 12 meses e que foram negociados em pelo menos 80% dos dias de negociação. Ele é revisado trimestralmente a fim de manter a sua representatividade do volume negociado e, em média, os componentes do Ibovespa representam 70% de todas os valores de ações negociadas. As empresas que participam do Índice Bovespa têm maior liquidez e representatividade da Bolsa de Valores Brasileira. Assim, os ativos que escolhemos para as nossos experimentos são considerados líquidos e muito importante no mercado brasileiro.

Em HFT, é comum realizar várias operações em um dia, tentando obter um pequeno ganho em cada operação. Neste contexto, nossos experimentos destinam-se a prever se um

[1]FIX Protocol: www.fixtradingcommunity.org

[2]Ibovespa index:
http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx? Indice=Ibovespa&Idioma=en-us

determinado ativo terá uma variação de preço positiva igual ou superior a um valor predefinido: R$ 0,02 (três centavos). Estamos interessados nessa variação porque podemos ganhar R$ 0,02 por cada ação negociada, inserindo uma ordem de compra e venda no preços R$ x e R$ x + 0,02, respectivamente. Assim, comprando e vendendo um lote padrão de ações (100 ações), temos um ganho de R$ 2,00 para cada tendência de alta identificada com sucesso. Repetindo esse processo, centenas ou milhares de vezes durante uma sessão de negociação, nossa estratégia de negociação pode obter ganhos consideráveis.

É importante deixar claro que o valor de R$ 0,02 foi obtido por meio da análise de custos de operação na Bolsa de Valores Brasileira. Esta variação também depende do valor da ação negociada, podendo ser maior ou menor que R$ 0,02. Testamos outras variações, por exemplo, R$ 0,03 e 0,04. Entretanto, optamos por usar R$ 0,02 para ter mais oportunidades e deixar o conjunto de dados mais balanceado em relação à distribuição das classes 0 e 1.

Conduzimos vários ensaios de *back-testing* com diferentes tamanhos de janela deslizante de treinamento e janela de tempo da previsão. A janela de tempo de previsão é o mesmo tempo usado para fechar uma negociação HFT, ou seja, quando uma oportunidade é identificada e duas ordens (compra e venda) são enviadas ao mercado, o algoritmo de negociação espera o tempo equivalente a janela de tempo da previsão para fechar posição.

Para o tamanho da janela deslizante de treinamento, usamos 1, 5, 8, 10, 14 e 20 dias. A janela de tempo de previsão recebe os seguintes valores: 5, 8 e 10 minutos. Neste contexto, encontramos que com 14 dias de treinamento, os modelos gerados têm um melhor desempenho no próximo dia de teste.

Através de gráficos CFD (*cumulative distribution function*) e tabelas, apresentaremos os melhores resultados encontrados. As tabelas mostram os valores das medidas de desempenho para os diferentes ensaios de *back-testing* executados. A CDF nos fornece uma maneria de descrever como as probabilidades são associadas aos valores ou intervalos de valores de uma variável aleatória de valor real X. Para cada número real x, a CDF é dada por:

$$F(x) = P(X \leq x), x \in \Re \tag{5.1}$$

A função F é igual à probabilidade de que a variável aleatória X assuma um valor inferior ou igual a determinado x.

Em nossos experimentos, definimos duas variáveis aleatórias para criar as CDFs: precisão e valor financeiro. Assim, os gráficos e tabelas nos permitem comparar os diferentes modelos de previsão gerados e também os resultados financeiros obtidos pela nossa estratégia de negociação. A seguir, apresentaremos uma análise e discussão das tabelas e CFDs.

As Tabelas 5.1, 5.2 e 5.3 apresentam a média dos resultados das medidas de desempe-

nho para três ativos, ITUB4, PETR4 e VALE5, respectivamente. A primeira coluna é a janela de tempo da previsão (TW). Para cada TW, temos os resultados de desempenho para cada modelo gerado utilizando diferentes técnicas de aprendizado de máquina: duas MLPs (MLP e MLP_1) e um *Ensemble*. Com esses resultados podemos concluir que: 1) A precisão dos modelos de aprendizado de máquina são sempre significativamente maior do que um modelo aleatório (que segue a distribuição de classe 1). Por exemplo, se nossa base de dados é composta por 1000 exemplos, sendo 400 da classe 1 e 600 da classe 0, a distribuição da classe 1 representa 40% de nossa base de dados; 2) Os melhores resultados (destacados em negrito), foram alcançados pelo *Ensemble*, que às vezes são semelhantes aos resultados obtidos pela MLP_1; 3) É importante esclarecer a diferença entre a precisão obtida pelos modelos no *back-testing* e taxa de acerto obtida em simulação. A precisão do modelo (Equação 4.17) estabelece a relação entre o número de gatilhos (oportunidades) identificados corretamente (VP) e o número de gatilhos identificados incorretamente (FP). A soma de VP e FP refere-se a todas oportunidades de negociação identificadas pelo modelo, ou seja, todas as vezes que a estratégia negociação realizou um operação no mercado, inserindo uma ordem de compra e outra de venda. Entretanto, em simulação realística, alguns gatilhos classificados pelo modelo como falso-positivos podem ter sucesso em simulação. Isso acontece por causa de como nosso modelo de previsão foi definido.

Através da Figura 5.1, podemos ver como um exemplo classificado falso-positivo pode gerar um gatilho de sucesso em simulação, ou seja, ambas as ordens de compra e venda serão executadas. Suponha que compramos um lote de 100 ações de ITUB4 a $R\$\ 34,46$, que a janela de previsão é de 5 minutos e que nosso modelo identificou uma oportunidade nesse momento, ou seja, previu que o preço da ação atingirá $R\$\ 34,49$, tendo uma variação positiva maior que $R\$0,02$. Neste contexto, uma ordem de venda é enviada ao mercado no preço de $R\$\ 34,48$. Esse gatilho é classificado como falso-positivo, pois o preço da ação atinge $R\$\ 34,49$ nos próximos 5 minutos. Entretanto, nossa ordem de venda pode ser executada, mas depende de sua prioridade no livro de ofertas. Como podemos ver na Figura 5.1, depois de 5 minutos, nossa ordem de venda é a próxima ser consumida. Neste caso, ela é consumida e o preço da ação volta a cair, não atingindo a variação prevista. Embora esse gatilho seja classificado como falso-positivo, em simulação realística, ele é executado com sucesso. Modelamos nossa função de previsão desta maneira para garantir que quando o modelo identificar uma oportunidade verdadeira (o preço da ação terá uma variação positiva maior que $R\$\ 0,02$), as ordens de compra e venda inseridas sempre serão executadas.

Seguindo com a análise e discussão dos resultados, as Figuras 5.2 a 5.10 apresentam os gráficos CDF para analisarmos o comportamento dos modelos em relação a distribuição da classe 1, que também pode ser entendida como a precisão de um modelo aleatório. Os

Tabela 5.1: Medidas de Desempenho para ITUB4

Tamanho de TW	Técnica	Distribuição	VP	FP	Precisão do Modelo	Taxa de Acerto
TW5	MLP	0.45	51,20	56,43	0,50	0,61
	MLP_1		52,97	58,00	0,50	0,62
	ENSEMBLE		29,54	30,83	**0,51**	**0,62**
TW8	MLP	0,55	110,89	83,03	0,58	0,67
	MLP_1		112,26	90,34	0,58	0,68
	ENSEMBLE		77,60	58,94	**0,60**	**0,70**
TW10	MLP	0,59	157,94	109,03	0,60	0,70
	MLP_1		154,83	104,14	0,61	0,71
	ENSEMBLE		130,49	88,63	**0,61**	**0,71**

Tabela 5.2: Medidas de Desempenho para PETR4

Tamanho de TW	Técnica	Distribuição	VP	FP	Precisão do Modelo	Taxa de Acerto
TW5	MLP	0,21	16,68	31,24	0,34	0,38
	MLP_1		15,14	31,00	0,33	0,43
	ENSEMBLE		8,53	13,75	**0,40**	**0,44**
TW8	MLP	0,29	28,03	46,59	0,38	0,44
	MLP_1		29,46	49,78	0,38	0,43
	ENSEMBLE		17,60	27,86	**0,40**	**0,45**
TW10	MLP	0,33	35,74	56,80	0,37	0,42
	MLP_1		41,89	66,23	0,35	0,42
	ENSEMBLE		25,00	35,66	**0,37**	**0,42**

exemplos pertencentes a classe 1 referem-se aos padrões que indicam oportunidades (gatilhos) de negociação. Além da CDF do modelo aleatório (linha amarela), o gráfico apresenta

Tabela 5.3: Medidas de Desempenho para VALE5

Tamanho de TW	Técnica	Distribuição	VP	FP	Precisão do Modelo	Taxa de Acerto
TW5	MLP	0,31	27,54	50,91	0,37	0,43
	MLP_1		27,91	57,17	0,36	0,42
	ENSEMBLE		14,95	25,11	**0,40**	**0,46**
TW8	MLP	0,40	50,34	68,89	0,43	0,50
	MLP_1		50,83	63,66	0,44	0,52
	ENSEMBLE		17,60	27,86	**0,48**	**0,52**
TW10	MLP	0,45	64,91	74,09	0,47	0,54
	MLP_1		70,77	75,29	0,47	0,53
	ENSEMBLE		46,46	45,71	**0,49**	**0,57**

também as CDFs da precisão dos modelos *ensemble* sobre o conjunto de teste (linha azul) e a precisão em simulação realística (linha verde). Esta análise está sintetizada para os diferentes tamanhos de janela de tempo da previsão e para cada ativo avaliado. Os resultados compreendem todo o período usado para o conjunto de teste.

Analisando os gráficos podemos ver que o comportamento do modelo *ensemble* é sem-

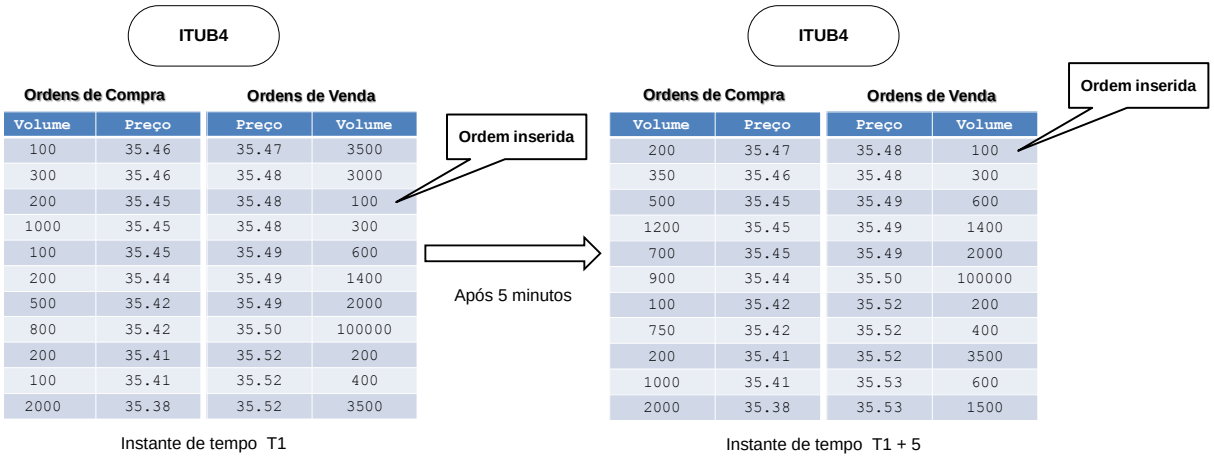

Figura 5.1: Estados do livro de ofertas para exemplificar que um gatilho falso-positivo pode ser executado com sucesso em simulação realística.

pre melhor que o modelo aleatório e, que a taxa de acerto em simulação (linha verde) apresenta uma CDF melhor se comparado com as outras. Como explicado anteriormente, isso acontece porque na simulação realística, alguns gatilhos classificados pelo modelo como falso-positivos podem ter sucesso em simulação.

Outra observação está relacionado a janela de tempo da previsão (TW). Quanto maior o valor de TW, mais os resultados entre *ensemble* e o modelo aleatório se aproximam. Como exemplo, para o ativo ITUB4 (TW8), 80% da distribuição da precisão para o modelo *ensemble* está acima de 50% (ver Figura 5.2), enquanto na simulação realística é aproximadamente 98%. Ainda analisando ITUB4, mas com TW igual a 10, o modelo ensemble tem aproximadamente 91% da distribuição com precisão acima de 50%. O tamanho de TW também aumenta o risco nas operações realizadas pela estratégia de negociação, pois a estratégia mantém por mais tempo suas ordens no livro de ofertas. Isso pode levar a maiores perdas caso a previsão estiver errada.

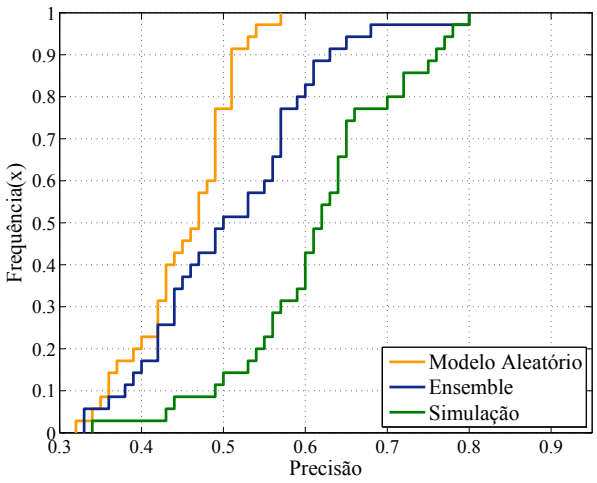

Figura 5.2: CDF com janela de tempo de previsão igual a 5 (TW5) para ITUB4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

A fim de avaliar os resultados financeiros reais obtidos pelo nosso algoritmo de negociação, realizamos uma simulação realística usando dados reais do sistema de negociação da Bolsa de Valores de São Paulo. Como esperado, pela análise dos resultados anteriores, os melhores resultados foram obtidos pelo modelo *ensemble* proposto, que combina duas redes neurais distintas. As Figuras 5.11, 5.12 e 5.13 mostram as CDFs dos resultados financeiros para os ativos ITUB4, PETR4 e VALE5. Cada gráfico apresenta três CDFs para diferentes janelas de tempo de previsão, 5,8 e 10 minutos. Note que aplicamos um *zoom-in* nas curvas, onde o eixo x representa o retorno financeiro (em $R\$$) das operações de compra e venda realizadas pela estratégia de negociação. O eixo y representa a frequência acumulada (valores

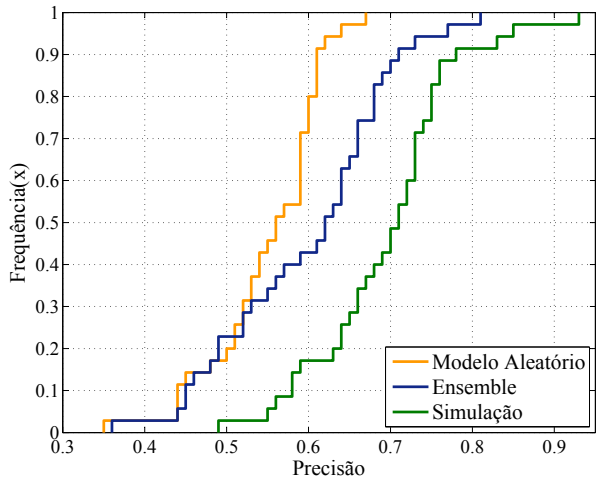

Figura 5.3: CDF com janela de tempo de previsão igual a 8 (TW8) para ITUB4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

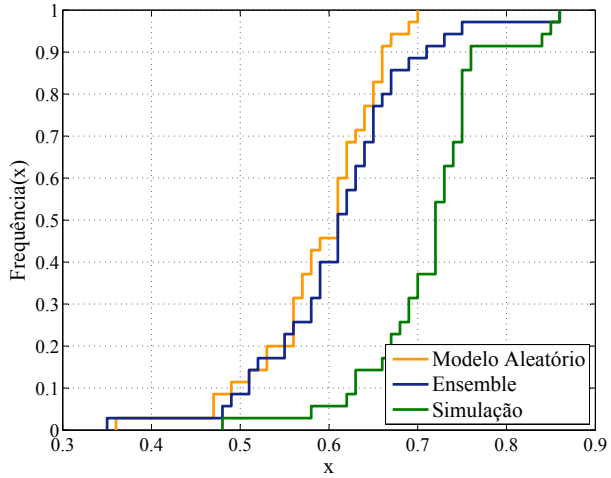

Figura 5.4: CDF com janela de tempo de previsão igual a 10 (TW10) para ITUB4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

entre 0 e 1) que os valores de x assume. É importante destacar que todos os gatilhos executados com sucesso sempre têm um retorno financeiro de $R\$\ 2,00$. Cada operação consiste em comprar e vender um lote de 100 ações e como nossa estratégia objetiva ganhar $R\$\ 0,02$ por ação, isso totaliza um retorno financeiro de $R\$\ 2,00$ por cada operação executada com sucesso.

Como podemos observar, as CDFs geradas nos mostram que, para todos os símbolos, os melhores resultados são alcançados com janelas de tempo maiores (8 e 10 minutos), ex-

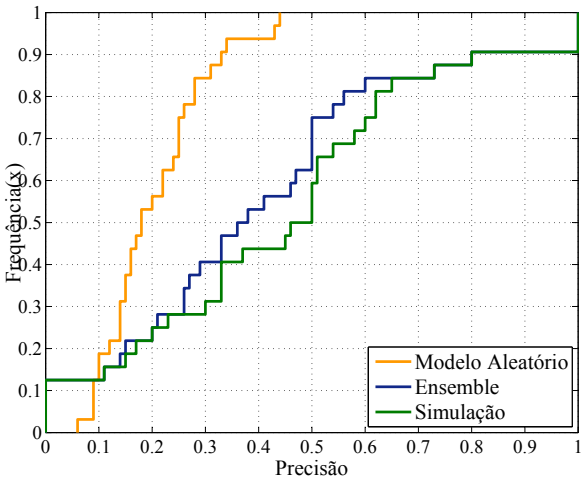

Figura 5.5: CDF com janela de tempo de previsão igual a 5 (TW5) para PETR4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

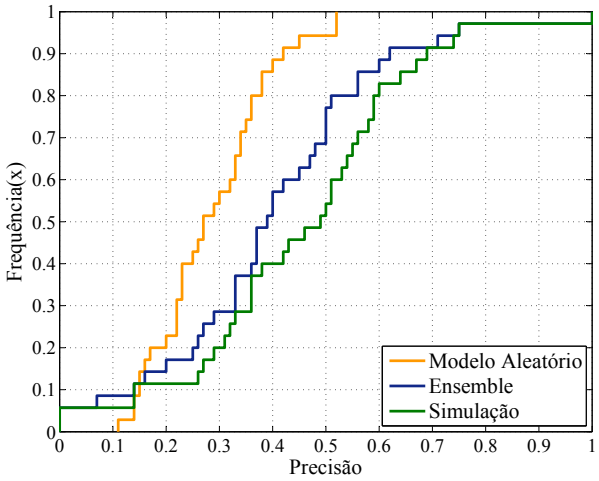

Figura 5.6: CDF com janela de tempo de previsão igual a 8 (TW8) para PETR4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

ceto para a PETR4. No entanto, nestes casos, a quantidade de perdas financeiras também são altas, uma vez que o risco da operação é alto devido a ocorrência de falso-positivos, ou seja, gatilhos que não deveriam ser executados e mesmo assim o modelo indicou como uma oportunidade para negociar. Para ITUB4, Figura 5.11), temos que, das gatilhos identificados, 64% (TW5) a 71% (TW10) foram finalizados com sucesso, isto é, suas ordens de compra e venda foram executadas nos preços desejados. Para a PETR4, Figura 5.12), 45% a 51% e para VALE5, 45% a 60%.

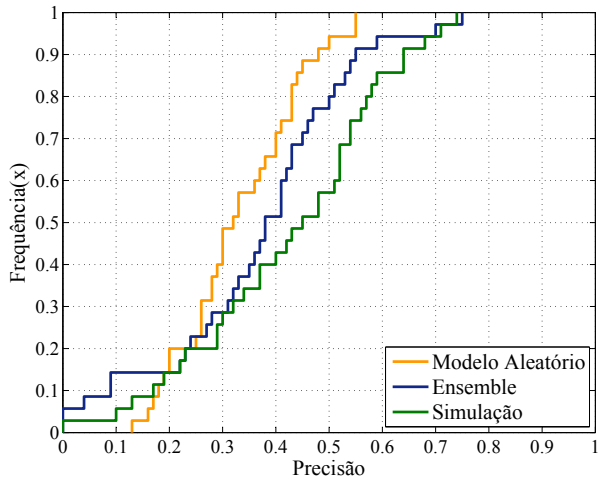

Figura 5.7: CDF com janela de tempo de previsão igual a 10 (TW10) para PETR4; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

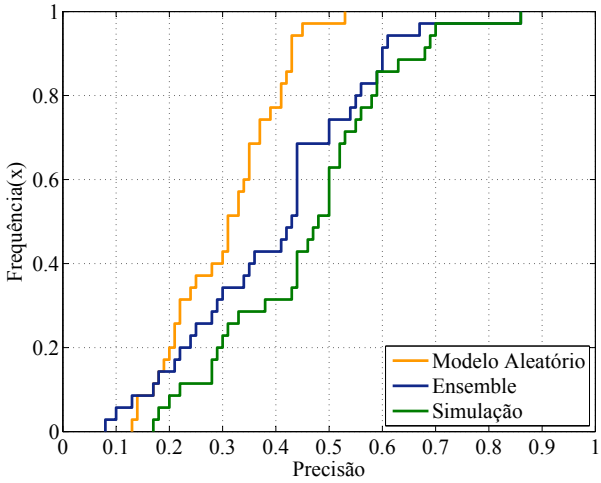

Figura 5.8: CDF com janela de tempo de previsão igual a 5 (TW10) para VALE5; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

Estes resultados são promissores, uma vez que o nosso modelo de aprendizagem de máquina já atingiu uma boa quantidade de sinais positivos. Como trabalho futuro, precisamos minimizar o número de gatilhos falso-positivos (FP) e, consequentemente aumentar o lucro e reduzir o risco. Assim, nossa estratégia de de negociação poderá fornecer resultados viáveis e rentáveis para negociação de alta frequência (HFT).

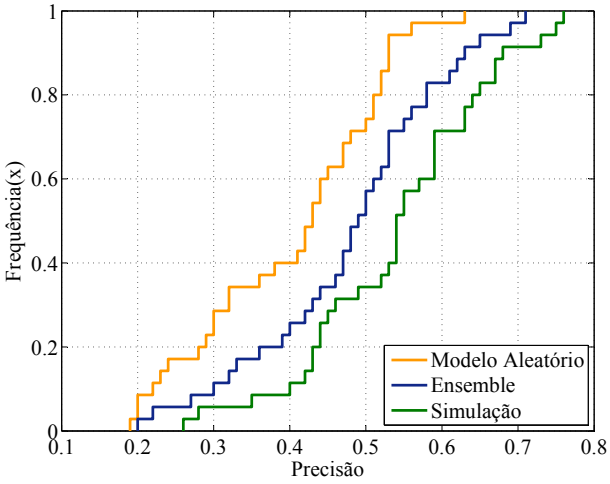

Figura 5.9: CDF com janela de tempo de previsão igual a 8 (TW8) para VALE5; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

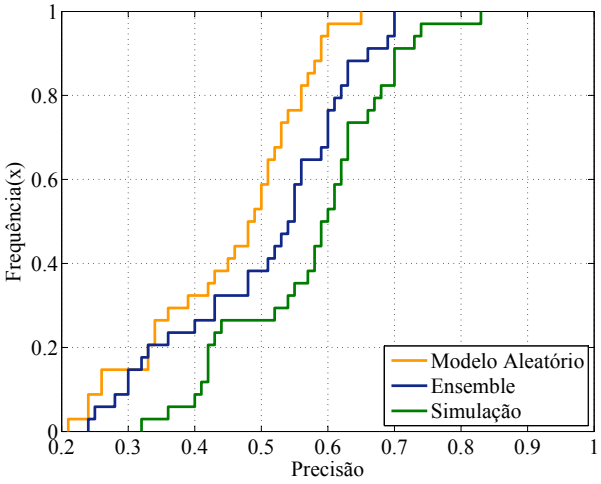

Figura 5.10: CDF com janela de tempo de previsão igual a 10 (TW10) para VALE5; frequência distribuição aleatória (linha laranja), precisão do modelo ensemble (linha azul) e taxa de acerto em simulação (linha verde).

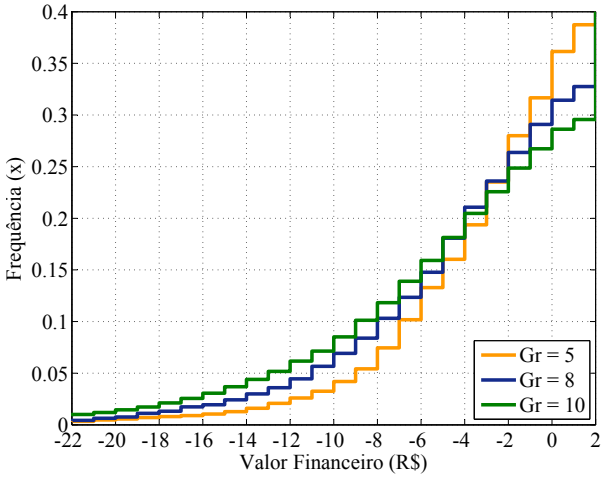

Figura 5.11: CDF do resultado financeiro para ITUB4; GR é o valor da janela de tempo da previsão. O eixo x representa o retorno financeiro de cada operação (oferta de compra e venda) realizado pela estratégia de negociação.

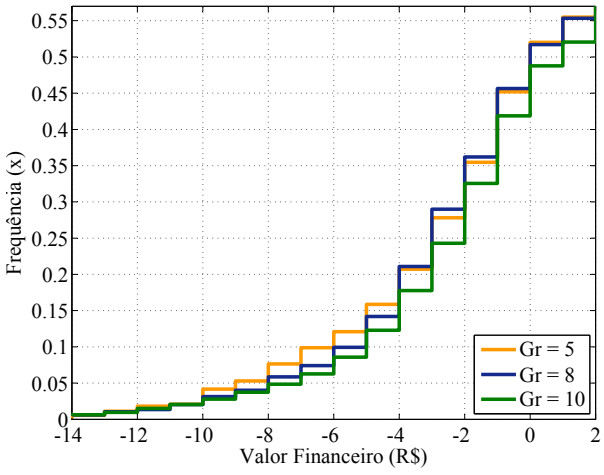

Figura 5.12: CDF do resultado financeiro para PETR4; GR é o valor da janela de tempo da previsão. O eixo x representa o retorno financeiro de cada operação (oferta de compra e venda) realizado pela estratégia de negociação.

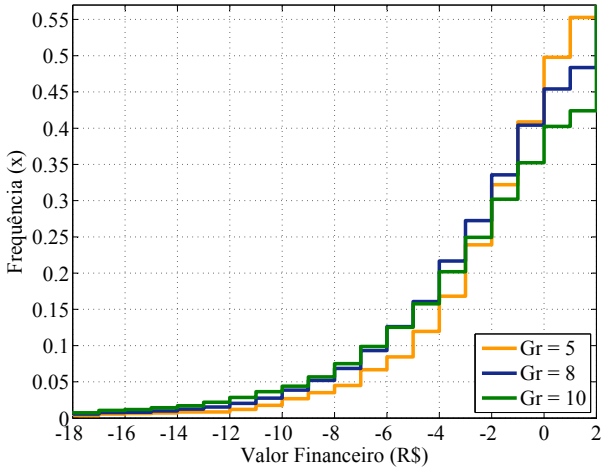

Figura 5.13: CDF do resultado financeiro para VALE5; GR é o valor da janela de tempo da previsão. O eixo x representa o retorno financeiro de cada operação (oferta de compra e venda) realizado pela estratégia de negociação.

Capítulo 6

Conclusão

O principal objetivo deste trabalho foi modelar e usar técnicas de aprendizado de máquina para maximizar o retorno obtido por uma estratégia direcional de negociação. Utilizando um grande volume de dados (*tick data*), conduzimos o *back-testing* e simulação em um simulador realístico da Bolsa de Valores de São Paulo. Através dos resultados empíricos obtidos, mostramos que técnicas de aprendizado de máquina foram capazes de melhorar a eficácia nesse processo de tomada de decisão. Demonstramos que a precisão e resultados obtidos através da simulação realística são melhores com a abordagem *ensemble*.

Como contribuição principal, um classificador *ensemble* foi construído para decidir se um momento específico é ou não propício para realizar uma operação de compra e venda de ações no mercado. Mostramos que o modelo *ensemble* proposto, que combina duas redes neurais artificiais, foi capaz de identificar com boa precisão essas oportunidades de negociação. Sua precisão foi sempre maior que a precisão de um modelo aleatório para todos os experimentos realizados. Além disso, os resultados do modelo ensemble também foram melhores que os resultados das redes neurais isoladas.

A metodologia proposta é genérica e pode ser aplicada para diferentes conjuntos de séries temporais financeiras, com a flexibilidade de escolher diferentes indicadores para compor o conjunto de atributos de entrada para o modelo, além de avaliar diferentes técnicas de aprendizado de máquina.

Este trabalho contextualiza uma nova contribuição no campo de algoritmo de negociação (*algotrading*), onde algoritmos de negociação em alta frequência têm especial importância.

Os resultados e conclusões alcançados abriram novas oportunidades de pesquisa. Assim prevemos as seguintes oportunidades para o trabalho futuro:

- Aperfeiçoar os modelos de previsão para reduzir o número de falso-positivos. Essa

redução impacta diretamente nos resultados financeiros obtidos em simulação, pois aumentará a taxa de acerto da estratégia de negociação;

- Utilizar técnicas de aprendizado de máquina para dar suporte a outros tipos de estratégias de negociação em alta frequência;
- Avaliar se tendência de baixa também pode ser identificada utilizando a mesma metodologia, apenas alterando a fórmula para calcular a saída do modelo. Assim, podemos utilizar duas estratégias de negociação em conjunto, uma realizando operações ao identificar tendência de alta e outra ao identificar um tendência de baixa. Isto poderia diminuir o risco de investimento e prover uma maior liquidez ao mercado.

Referências Bibliográficas

(2013). Editorial board. *Journal of Financial Markets*, 16(4):IFC –. ISSN 1386-4181. High-Frequency Trading.

Adebiyi Ayodele, A.; Ayo Charles, K. & Otokiti Sunday, O. (2012). Stock price prediction using neural network with hybridized market indicators.

Aldridge, I. (2013). *High-frequency trading: a practical guide to algorithmic strategies and trading systems*. John Wiley & Sons.

Benevenuto, F.; Magno, G.; Rodrigues, T. & Almeida, V. (2010). Detecting spammers on twitter. Em *Collaboration, electronic messaging, anti-abuse and spam conference (CEAS)*, volume 6, p. 12.

Biais, B. & Woolley, P. (2011). High frequency trading. *Manuscript, Toulouse University, IDEI*.

Bovespa (2008). Mercado de Capitais. `http://www.bmaiscompet.com.br/arquivos/MercadodeCapitaisBovespa.pdf`. [Acesso em: 02 de Maio de 2015].

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5--32.

Brokers (2010). O que é a Bolsa de valores e Mercado de Ações? `http://www.melhoresbrokers.com.br/defini%C3%A7%C3%A3o/o-que-%C3%A9-a-bolsa-de-valores-e-mercado-de-a%C3%A7%C3%B5es-explica%C3%A7%C3%A3o-e-defini%C3%A7%C3%A3o-de-termos.html`. [Acesso em: 02 de Maio de 2015].

Chakraborty, T. & Kearns, M. (2011). Market making and mean reversion. Em *Proceedings of the 12th ACM conference on Electronic commerce*, pp. 307--314. ACM.

Chen, Y. & Pennock, D. M. (2010). Designing markets for prediction. *AI Magazine*, 31(4):42–52.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273--297.

Costa, H.; Benevenuto, F. & Merschmann, L. H. (2013). Detecting tip spam in location-based social networks. Em *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pp. 724--729. ACM.

Davis, J. & Goadrich, M. (2006). The relationship between precision-recall and roc curves. Em *Proceedings of the 23rd international conference on Machine learning*, pp. 233--240. ACM.

de Oliveira, F. A.; Nobre, C. N. & Zárate, L. E. (2013). Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index - case study of petr4, petrobras, brazil. *Expert Systems with Applications*, 40(18):7596 – 7606. ISSN 0957-4174.

Dietterich, T. G. (2000). Ensemble methods in machine learning. Em *Multiple classifier systems*, pp. 1--15. Springer.

Fahlman, S. E. & Hinton, G. E. (1987). Connectionist architectures for artificial intelligence. *Computer;(United States)*, 20(1):100--109.

Gomber, P.; Arndt, B.; Lutat, M. & Uhle, T. (2011). High-frequency trading. *Available at SSRN 1858626*.

Hagströmer, B. & Norden, L. (2013). The diversity of high-frequency traders. *Journal of Financial Markets*, 16(4):741--770.

Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P. & Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10--18.

Kablan, A. (2009). Adaptive neuro fuzzy inference systems for high frequency financial trading and forecasting. Em *Advanced Engineering Computing and Applications in Sciences, 2009. ADVCOMP'09. Third International Conference on*, pp. 105--110. IEEE.

Kablan, A. & Ng, W. (2010). High frequency trading using fuzzy momentum analysis. Em *Proceedings of the World Congress on Engineering*, volume 1.

Kutsurelis, J. E. (1998). Forecasting financial markets using neural networks: An analysis of methods and accuracy. Relatório técnico, DTIC Document.

Li, X.; Deng, X.; Zhu, S.; Wang, F. & Xie, H. (2014). An intelligent market making strategy in algorithmic trading. *Frontiers of Computer Science*, 8(4):596--608.

Lin, T. C. (2013). The new investor. *UCLA L. Rev.*, 60:678--778.

Lukeman, J. (2000). *The Market Maker's Edge*. McGraw-Hill New York.

Mahato, P. K. (2014). Prediction of stock price movement using various ensemble models. Dissertação de mestrado, Department of Computer Engineering and Information Technology, College of Engineering, Pune.

McCluskey, P. C. (1993). Feedforward and recurrent neural networks and genetic programs for stock market and time series forecasting. *Master's thesis, Brown University*.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712 – 740. ISSN 1386-4181. High-Frequency Trading.

Niaki, S. & Hoseinzade, S. (2013). Forecasting s&p 500 index using artificial neural networks and design of experiments. *Journal of Industrial Engineering International*, 9(1). ISSN 1735-5702.

Oliveira, H. C. B. d. (2012). *Algoritmo Online para o Problema Dinâmico de Roteamento de Veívulos*. Tese de doutorado, Departamento de Ciência da Computação - UFMG, Belo Horizonte.

Pal, S. K. & Mitra, S. (1992). Multilayer perceptron, fuzzy sets, and classification. *Neural Networks, IEEE Transactions on*, 3(5):683--697.

Putra, E. F. & Kosala, R. (2011). Application of artificial neural networks to predict intraday trading signals. *Recent Researches in E-Activities*, pp. 174--179.

Rosenblatt, F. (1961). Principles of neurodynamics. perceptrons and the theory of brain mechanisms. Relatório técnico, DTIC Document.

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.

Silva, E.; Castilho, D.; Pereira, A. & Brandao, H. (2015). A binary ensemble classifier for high-frequency trading. Em *Neural Networks (IJCNN), 2015 International Joint Conference on*. IEEE.

Sokolova, M. & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427--437.

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote sensing of Environment*, 62(1):77--89.

Suan, T. S. & Chye, K. H. (1998). Neural network applications in accounting and business. Relatório técnico.

Thalheimer, R. & Ali, M. M. (1979). Time series analysis and portfolio selection: An application to mutual savings banks. *Southern Economic Journal*, pp. 821--837.

Thawornwong, S.; Dagli, C. H. & Enke, D. L. (2001). Using neural networks and technical analysis indicators for predicting stock trends. *Intelligent Engineering Systems through Artificial Neural Networks*.

Thirunavukarasu, P. (2009). Estimation of return on investment in share market through ann. *Global Journal of Finance and Management*, 1(2):113--122.

White, H. (1988). Economic prediction using neural networks: The case of ibm daily stock returns. Em *Neural Networks, 1988., IEEE International Conference on*, pp. 451--458. IEEE.