



TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À ESTRATÉGIA DE SWING TRADE DO MERCADO FINANCEIRO

Pedro Henrique Barbosa Nori

Projeto de Graduação apresentado ao Curso de Engenharia Eletrônica e de Computação da Escola Politécnica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Engenheiro.

Orientador: Heraldo Luis Silveira de Almeida

Rio de Janeiro

Julho de 2021

TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À
ESTRATÉGIA DE SWING TRADE DO MERCADO
FINANCEIRO

Pedro Henrique Barbosa Nori

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO CURSO
DE ENGENHARIA ELETRÔNICA E DE COMPUTAÇÃO DA ESCOLA PO-
LITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE ENGENHEIRO ELETRÔNICO E DE COMPUTAÇÃO

Autor:

Pedro Henrique Barbosa Nori

Orientador:

Heraldo Luis Silveira de Almeida, D. Sc.

Examinador:

Prof xxxxx

Examinador:

Prof xxxx

Rio de Janeiro

Julho de 2021

Declaração de Autoria e de Direitos

Eu, *Pedro Henrique Barbosa Nori* CPF 134.129.077-82, autor da monografia *TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À ESTRATÉGIA DE SWING TRADE DO MERCADO FINANCEIRO*, subscrevo para os devidos fins, as seguintes informações:

1. O autor declara que o trabalho apresentado na disciplina de Projeto de Graduação da Escola Politécnica da UFRJ é de sua autoria, sendo original em forma e conteúdo.
2. Excetua-se do item 1. eventuais transcrições de texto, figuras, tabelas, conceitos e idéias, que identifiquem claramente a fonte original, explicitando as autorizações obtidas dos respectivos proprietários, quando necessárias.
3. O autor permite que a UFRJ, por um prazo indeterminado, efetue em qualquer mídia de divulgação, a publicação do trabalho acadêmico em sua totalidade, ou em parte. Essa autorização não envolve ônus de qualquer natureza à UFRJ, ou aos seus representantes.
4. O autor pode, excepcionalmente, encaminhar à Comissão de Projeto de Graduação, a não divulgação do material, por um prazo máximo de 01 (um) ano, improrrogável, a contar da data de defesa, desde que o pedido seja justificado, e solicitado antecipadamente, por escrito, à Congregação da Escola Politécnica.
5. O autor declara, ainda, ter a capacidade jurídica para a prática do presente ato, assim como ter conhecimento do teor da presente Declaração, estando ciente das sanções e punições legais, no que tange a cópia parcial, ou total, de obra intelectual, o que se configura como violação do direito autoral previsto no Código Penal Brasileiro no art.184 e art.299, bem como na Lei 9.610.
6. O autor é o único responsável pelo conteúdo apresentado nos trabalhos acadêmicos publicados, não cabendo à UFRJ, aos seus representantes, ou ao(s) orientador(es), qualquer responsabilização/ indenização nesse sentido.
7. Por ser verdade, firmo a presente declaração.

Pedro Henrique Barbosa Nori

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Escola Politécnica - Departamento de Eletrônica e de Computação

Centro de Tecnologia, bloco H, sala H-217, Cidade Universitária

Rio de Janeiro - RJ CEP 21949-900

Este exemplar é de propriedade da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmear ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es).

DEDICATÓRIA

À minha mãe engenheira mecânica que tanto amo no meu coração.

AGRADECIMENTO

Agradeço ao meu país, que esconde um povo tão sofrido e ao mesmo tempo tão amoroso. Este trabalho é apenas um pedacinho do pagamento da minha dívida.

RESUMO

Todos os dias, diversas negociações são realizadas nas bolsas de valores no mundo inteiro. Com os mais diversos objetivos, investidores buscam um aumento crescente de patrimônio de forma consistente. Paralelamente, inteligências artificiais vem substituindo cada vez mais atividades antes desempenhadas pelo homem.

Nesse sentido, este trabalho visa a aplicação de técnicas de aprendizado de máquina para aumento de performance de uma estratégia de swing trade no mercado de ações brasileiro (B3). Para isso, é realizada a reprodução aproximada da estratégia, seguida pela substituição dos critérios de decisão de entrada nas operações e preços alvos de venda por um modelo de aprendizado de máquina.

Palavras-Chave: Machine Learning, Análise Técnica, Swing Trade, Mercado Financeiro.

ABSTRACT

Insert your abstract here. Insert your abstract here. Insert your abstract here.
Insert your abstract here. Insert your abstract here.

Key-words: word, word, word.

SIGLAS

AF - Análise Fundamentalista

API - *Application Programming Interface*

ANS - Aprendizado Não Supervisionado

AS - Aprendizado Supervisionado

AT - Análise Técnica

B3 - Bolsa, Brasil, Balção

CSV - *Comma-separated values*

DT - *Decision Tree*

FS - *Feature Selection*

iBovespa - Índice Bovespa

JSON - *JavaScript Object Notation*

k-NN - *K Nearest Neighbors*

LR - *Logistic Regression*

ML - *Machine Learning*

RF - *Random Forest*

SVM - *Support Vector Machines*

UFRJ - Universidade Federal do Rio de Janeiro

Sumário

1	Introdução	1
1.1	Tema	1
1.2	Delimitação	1
1.3	Justificativa	2
1.4	Objetivos	3
1.5	Metodologia	3
1.6	Descrição	4
2	Fundamentação Teórica	5
2.1	Mercado de Capitais, Bolsa de Valores e Ações	5
2.1.1	Hipótese do Mercado Eficiente	6
2.1.2	Índice de Bolsa de Valores	8
2.1.3	Mercado Fracionário	8
2.2	Tipos de Análises	9
2.2.1	Análise Fundamentalista	9
2.2.2	Análise Técnica	9
2.3	Aprendizado de Máquina	10
2.3.1	Aprendizado Supervisionado	10
2.3.2	Problema de Regressão	11
2.3.3	Problema de Classificação	11
2.4	Trabalhos Relacionados	11
2.4.1	Modelos Baseadas em Indicadores Técnicos	11
2.4.2	Modelos Baseados em Processos Estocásticos	12
2.4.3	Modelos Baseados em Inteligência Artificial	13

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Tema

O tema deste trabalho se resume no aperfeiçoamento de uma estratégia de swing trade na bolsa de valores através de métodos de aprendizado de máquina.

Nesse contexto, o problema a ser abordado é a identificação do momento apropriado para compra de um determinado ativo, como também os preços alvos determinantes para venda, tendo em vista uma variação positiva de seu preço.

1.2 Delimitação

Este trabalho se limita aos ativos negociados na Bolsa de Valores de São Paulo, a B3, cujos dados diários são de domínio público e foram adquiridos através da plataforma Yahoo Finance pela API open-source yfinance, disponível em Python. Não são levadas em consideração informações sobre proventos (dividendos e juros sobre capital próprio) devido à inconsistência dos mesmos na API supracitada e à dificuldade técnica para automatização da busca de tais dados.

A duração das operações tem em vista um horizonte mínimo de um dia, sendo portanto operações de swing trade. Não são realizadas vendas a descoberto, portanto só há lucro em variações positivas dos ativos. Apenas uma operação por ativo pode existir em um determinado instante de tempo para uma estratégia. Em outras

palavras, só é possível comprar mais ações de uma empresa após a venda completa das ações da mesma, caso existam.

A incidência de impostos devidos (e.g., imposto de renda) está fora do escopo, assim como a utilização de critérios baseados em análise fundamentalista, por causa da dificuldade de obtenção dessas informações de maneira automatizada e estruturada.

1.3 Justificativa

O crescimento do número de investidores na bolsa de valores brasileira [1] demonstra um maior interesse da população na busca por um complemento da renda familiar ou até na substituição da fonte de renda principal.

No cenário global, o aumento do uso de robôs de trading (ou algoritmos) tem se mostrando expressivo [2], sejam por pessoas físicas ou fundos de investimento, de forma total ou parcial em suas estratégias. Por outro lado, tal crescimento não vem sendo igualmente representado no Brasil devido às peculiaridades do mercado de capitais nacional, como a alta volatilidade e a alta sensibilidade a notícias [3].

Paralelamente, estudos relacionados a aprendizado de máquina vem trazendo resultados práticos no dia-a-dia das pessoas, desde o clássico exemplo de reconhecimento de mensagens de spam em um caixa de email à identificação do perfil de consumo de clientes em uma loja. Da mesma forma, instituições financeiras e bancos centrais também estão, com cautela, incorporando aplicações de aprendizado de máquina em tarefas internas [4].

Apesar das dificuldades inerentes ao cenário atual do mercado de capitais brasileiro, não se pode ignorar o potencial que os algoritmos podem trazer. Desta forma, o presente trabalho visa a união de técnicas de aprendizado de máquina a estratégias de trading de forma a trazer uma melhor performance, colaborando assim para uma maior variedade de opções de investimentos à população brasileira.

1.4 Objetivos

O objetivo geral deste trabalho é implementar um software capaz de simular uma estratégia de swing trade e gerar uma nova estratégia baseada na anterior utilizando aprendizado de máquina a fim de melhorar sua performance. Especificamente, o software deve: (1) Criar um ambiente automatizado que permita buscar, atualizar e armazenar dados diários da bolsa brasileira de forma simples e conforme necessidade do usuário da aplicação; (2) Simular a estratégia de swing trade do trader André Moraes da forma mais fidedigna que a janela de dados diária permita; (3) Criar e simular um novo algoritmo baseado na estratégia anterior utilizando aprendizado de máquina; (4) Criar e simular uma estratégia de baseline, referente à estratégia de aprendizado de máquina; (5) Analisar os modelos gerados.

1.5 Metodologia

O trabalho teve início na criação de um ambiente propício à simulação de estratégias, bem como sua configuração e manutenção. Consequentemente, a fim de: otimizar o tráfego de dados pela internet; minimizar o processamento necessário para a geração de dados derivados (pré-processamento); e armazenar os resultados das estratégias de forma organizada, foi utilizado um banco de dados PostgreSQL. Dentre as atividades realizadas durante o pré-processamento dos dados, anteriores à simulação, é possível citar a geração de candles semanais a partir de candles diários, a identificação de picos, os momentos de tendência de alta do mercado e as médias móveis exponenciais dos preços de fechamento.

Em seguida, a etapa de simulação começa na leitura de um arquivo JSON contendo todos parâmetros necessários para a execução das estratégias. Nesta etapa, o programa itera dia após dia para cada estratégia configurada verificando os momentos e os valores de compra e de venda para cada ativo que compõe as carteiras. Ao final, registram-se no banco todas as operações executadas, independente da obtenção de lucro, junto com as informações estatísticas necessárias para a avaliação da performance. Aqui são criadas e executadas: a estratégia base, que é uma adaptação do André Moraes; a estratégia aprimorada, que utiliza aprendizado de máquina; e

a estratégia de baseline, respectivamente.

Por fim, com o objetivo de facilitar a análise dos resultados gerados, criou-se um dashboard responsável por centralizar todas as informações pertinentes a uma execução de estratégia em uma única página web.

Observa-se que além do uso de estruturas do banco de dados PostgreSQL, como triggers e functions, o código foi construído em Python devido à ampla variedade de bibliotecas, especialmente de Data Science, e ao suporte da comunidade, apesar da desvantagem de desempenho por ser uma linguagem interpretada. Bastante foco foi dado à escalabilidade e à manutenção do código, que contou com as bibliotecas e as APIs yfinance, pandas, numpy, scikit-learn, multiprocessing, matplotlib e dash. Também utilizou-se containers Docker para simplificar a execução.

1.6 Descrição

No capítulo 2 é desenvolvida a fundamentação teórica acerca do mercado de capitais, da bolsa de valores e das ações, com enfoque no contexto brasileiro.

Capítulo 2

Fundamentação Teórica

Neste capítulo, são introduzidos alguns conceitos chave para o entendimento do projeto. Nas próximas seções, são feitas contextualizações sobre o Mercado de Capitais, Bolsa de Valores, Ações e Aprendizado de Máquina.

2.1 Mercado de Capitais, Bolsa de Valores e Ações

O Mercado de Capitais, também conhecido como Mercado de Valores Mobiliários, é um dos segmentos do sistema financeiro responsável por fazer o intermédio entre agentes superávitaros, que tem capital de investimento, e agentes deficitários, que buscam capital para rentabilizá-lo, através da compra e venda valores mobiliários (i.e., ativos financeiros) [5]. Consequentemente, gera-se uma maior liquidez destes ativos e também uma melhora no fluxo de capitais entre os agentes econômicos, sejam eles os governos por meio dos bancos centrais, os bancos privados, as instituições financeiras ou até mesmo as pessoas físicas.

No Brasil, o Mercado de Capitais é regulado e fiscalizado pela CVM (Comissão de Valores Mobiliários), uma autarquia federal vinculada ao Ministério da Fazenda e criada em 1976 através da Lei nº 6.385 [6].

A Bolsa de Valores é uma plataforma onde se negociam os valores mobiliários do Mercado de Capitais, dentre eles ações (i.e., fatias, pedaços) de sociedades anônimas (ou companhias). No Brasil, a única Bolsa de Valores oficial existente é a B3 (Brasil, Bolsa, Balcão) [7], que administra os sistemas de negociação, compensação,

liquidação, depósito e registro para todas as principais classes de ativos.

O processo de abertura de capital de uma empresa é uma iniciativa que possui vantagens estratégicas [8] como: o aumento da confiança na perspectiva do mercado, seja para o consumidor final ou para parceiros comerciais; a solução de problemas decorrentes de processos sucessórios; e também a captação de capital de investimento, a fim de contribuir para o crescimento ou para a consolidação da companhia. Esse processo acontece através de uma oferta pública [9], ou IPO (Initial Public Offering), onde as ações que compõem o capital social [10] de uma companhia são vendidas pela primeira vez ao público geral. Uma vez encerrado o IPO, estas mesmas ações passam para o mercado secundário [11], onde investidores as negociam entre si. Em retorno ao capital adquirido pela companhia, surgem algumas responsabilidades, dentre elas a publicação de demonstrações financeiras [12], auditadas pela própria CVM [13].

Para o acionista de uma sociedade anônima, existem duas formas de se obter lucro: através de proventos (dividendos e juros sobre capital próprio) [14]; ou através de operações de compra e de venda de ações, mediante oscilações de seu valor de mercado. Conforme a expectativa corretamente induz, o lucro é comumente aferido durante a venda de um determinado papel (i.e., ação) posteriormente à sua aquisição a um preço de compra inferior. No entanto, também é possível trabalhar com posições vendidas (short selling) [15], onde um investidor aluga ações de outro investidor por meio de um contrato. Em seguida as vende para posteriormente recomprá-las a um preço inferior, devolvendo-as assim ao respectivo dono. Neste caso, o lucro é obtido quando expectativa de queda de um ativo se mostra verdadeira.

2.1.1 Hipótese do Mercado Eficiente

A Hipótese do Mercado Eficiente, definida por FAMA [16], afirma que idealmente o preço de um ativo reflete toda a informação disponível sobre seu valor intrínseco. Em outras palavras, quanto menor o efeito de fatores que contribuam para uma inércia no fluxo de capital de investidores e na transmissão de informações, mais o mercado tende a ser eficiente. São estudados os três níveis de hipóteses:

- HME fraca: O preços atuais refletem o todo o histórico de informações dispo-

nibilizados publicamente.

- HME semi-forte: Engloba a HME fraca, acrescentando a existência de uma mudança instantânea que os preços sofrem ao surgirem novas informações.
- HME forte: Engloba a HME semi-forte, porém entende que a mudança instantânea dos preços acompanha toda e qualquer informação existente sobre o ativo. Assim, absolutamente nenhum investidor conseguiria obter lucro superior à média do mercado, pois não há como acessar nenhuma informação privilegiada, uma vez que ela já estaria refletido no preço corrente do ativo.

O autor menciona que o HME forte não é estritamente válida na realidade, o que é uma afirmação coerente quando se verifica a existência de casos em que o vazamento de informações confidenciais trouxe aos acusados uma lucratividade significativa [17].

A HME fraca foi verificada devido à consistência da correlação dos preços dia após dia de determinadas ações, mesmo que esta fosse baixa.

A hipótese semi-forte também foi sustentada por alguns fatores, dentre eles a verificação de que os futuros pagamentos de dividendos das companhias se refletem, em média, no preços das ações [18].

Em resumo, o estudo das Hipóteses de Mercado Eficiente traz informações relevantes quanto se avalia a teoria por trás da possibilidade de aplicação de estratégias de trading no mercado financeiro. No entanto, é importante ressaltar que outros autores questionam ao menos parcialmente os estudos realizados por FAMA, sejam por resultados inconclusivos ou por anomalias detectadas no comportamento do mercado. Por exemplo, SHOSTAK [19] critica abertamente a premissa de que todos os investidores teriam a mesma expectativa sobre os retornos da empresa. O ganhador do prêmio Nobel em ciências econômicas Paul Samuelson, que afirma que o a HME funciona muito melhor para ações individuais do que para o mercado como um todo [20]. Já o investidor Jack Schwager afirma que a HME está correta pelos motivos errados [21], pois é muito difícil bater a média do mercado de forma consistente ao mesmo tempo que investidores possuem habilidades diferentes, portanto a informação não é interpretada e aplicada por todos da mesma forma.

2.1.2 Índice de Bolsa de Valores

Índices de Bolsas de Valores [22] são métricas criadas para avaliar a saúde de um determinado grupo de ações negociadas na bolsa. Cada índice possui uma regra própria de criação que define quais ações são englobadas e com quais pesos, como por exemplo:

- S&P 500: Um dos mais conhecidos no mercado. É a média ponderada pelo capital social das 500 maiores companhias do mercado americano.
- Dow Jones Industrial Average: É a média ponderada pelo preço da ação das 30 maiores blue-chips industriais e financeiras do mercado americano (i.e., companhias bem conhecidas, bem estabelecidas e com grande capital social).
- Ibovespa: Principal indicador de desempenho do mercado brasileiro. Possui alguns critérios específicos, mas basicamente é composto pelas ações com maior volume de negociação na B3 [23].

Índices não são negociáveis pois não passam de métricas de mercado. Para isso existem fundos de investimentos chamados ETFs (Exchange-Traded Funds) [24], especializados em seguir um determinado índice.

No Brasil, um investidor que deseja que uma parte de seu capital acompanhe um rendimento equivalente ao iBovespa deverá investir no ETF, cujo código de negociação é BOVA11.

2.1.3 Mercado Fracionário

Ações são negociadas em múltiplos de um lote, que representa uma quantidade mínima de papéis a transacionar. Nesse contexto, o Mercado Fracionário [25] surge com o objetivo de facilitar negociações de volumes menores que o lote mínimo permitido. Na prática, ações fracionárias são agrupadas até formarem um lote para então serem negociadas. Normalmente o Mercado Fracionário possui menor liquidez e maior volatilidade, mas sempre acompanha o preço do ativo negociado no mercado aberto.

Ações fracionárias podem ser criadas devido: a um desdobramento de ações que não gera resultado par (e.g., 3 para 2); ou a fusões e aquisições de empresas que combinam suas ações em uma razão predeterminada.

Grandes investidores e fundos de investimentos não possuem problemas quanto ao capital mínimo necessário para a compra de um lote de ações, visto que negociam em quantidades muito maiores. O problema surge quando um investidor com pouco aporte financeiro deseja entrar no mercado e não consegue encontrar ativos cujo lote mínimo esteja dentro de seu orçamento.

No Brasil, o lote mínimo é de 100 ações e o Mercado Fracionário permite a compra de no mínimo 1 ação.

2.2 Tipos de Análises

2.2.1 Análise Fundamentalista

A Análise Fundamentalista (AF) é muito utilizada para identificar tendências de flutuação no preço de ações tendo em vista um horizonte de longo prazo [26]. Ela se baseia em fatores econômicos relacionados à companhia, como: o quadro de diretores e dirigentes maiores; o fluxo de caixa; a saúde e a situação financeira; o contexto político do país; os concorrentes de mercado; as circunstâncias climáticas; os desastres climáticos, naturais ou não, dentre outros fatores.

Devido à natureza desorganizada e desestruturada dos dados que representam os fatores mencionados, torna-se muito difícil implementar uma automação.

2.2.2 Análise Técnica

A Análise Técnica (AT) busca identificar tendências de curto prazo na série temporal de preços de ações através da identificação de padrões e da criação de informações derivadas (indicadores técnicos) [27, 28]. Em teoria, o preço das ações é consequência de todos os acontecimentos relacionados direta ou indiretamente a uma companhia, assim como as análises feitas por investidores grandes.

Diferentemente da AF, a automação desta análise é muito mais fácil pois os dados normalmente são organizados e estruturados. No entanto, como são obtidos a posteriori, a dificuldade desta análise se dá na separação entre o que é ruído e o que é de fato tendência de mercado, além da criação de informações derivadas que se mostram relativamente úteis.

2.3 Aprendizado de Máquina

Aprendizado de Máquina (Machine Learning) é um campo de estudo dentro de Inteligência Artificial [29] que engloba estatística e ciência da computação. O objetivo é extrair conhecimento a partir de um conjunto de dados [30]. A terminologia foi criada por um pesquisador da IBM chamado SAMUEL em 1959 [31] para um estudo de caso do jogo de damas [32].

Em geral, algoritmos de ML buscam realizar tarefas extremamente complexas computacionalmente sem serem explicitamente programadas caso a caso. Alguns exemplos de aplicações que deixam evidente os benefícios deste método são: visão computacional, recomendação de produtos, identificação de rosto em imagens, identificação transações financeiras fraudulentas, suporte a diagnósticos médicos, dentre diversos outros.

Algoritmos de ML podem ser baseados em Aprendizado Supervisionado, Aprendizado Não Supervisionado ou até mesmo um modelo híbrido. Este trabalho utiliza apenas AS para a criação de modelos.

2.3.1 Aprendizado Supervisionado

Uma das metodologias mais comuns de ML, seu objetivo é a predição de um resultado a partir de um conjunto de dados de entrada, com a condição de que o modelo tem acesso a vários exemplos de entrada e saída de dados para uma melhor performance [30]. Existem dois tipos de problemas associados ao AS, os problemas de Regressão e os problemas de Classificação.

2.3.2 Problema de Regressão

2.3.3 Problema de Classificação

2.4 Trabalhos Relacionados

Tendo em vista o conflito de interesses existente por trás de trabalhos de cujo tema está relacionado à previsibilidade do mercado financeiro, pode-se questionar se as estratégias mais promissoras de fato são encontradas em domínio público. Isso ocorre pois a democratização de uma estratégia lucrativa poderia implicar na redução das lucratividades individuais, especialmente se for utilizada em escala.

Segundo KIM [33], somente a partir dos anos 80 que as corretoras começaram a utilizar protocolos de comunicação eletrônica para substituir a corretagem por voz. Essa inovação permitiu o desenvolvimento do Algorithmic Trading, que é a automatização da tomada de decisões de estratégias por um computador capaz de enviar ordens de compra e venda diretamente ao mercado.

Para efeito de simplificação, as estratégias de AT aplicadas ao mercado financeiro serão agrupadas em três metodologias centrais: modelos baseados em indicadores técnicos; modelos baseados em processos estocásticos; e modelos baseados em inteligência artificial.

2.4.1 Modelos Baseadas em Indicadores Técnicos

Este tipo de abordagem utiliza informações derivadas da série temporal de preços para criar uma combinação de indicadores que possuam algum poder de previsibilidade da tendência de mercado. Quando comparada aos outros tipos, é a metodologia mais simples e democrática, uma vez que pessoas com pouco ou nenhum conhecimento sobre estatística e inteligência artificial podem operar em estratégias próprias.

Dentre os indicadores mais famosos e portanto utilizados, podemos citar: o volume financeiro; a identificação de tendências de alta, de baixa e de consolidação de acordo com a Teoria de Dow [34]; as linhas de suporte e resistência do mercado; as médias

móveis; as bandas de Bollinger [35]; e o MACD (Moving Average Convergence-Divergence) [36].

Citar MORAES [37].

2.4.2 Modelos Baseados em Processos Estocásticos

De acordo com GODFREY [38], a hipótese de que a flutuação de preços no mercado de ações poderia ser explicada por uma Random Walk¹ foi feita por BACHELIER [39]. A partir da década de 60, muitos trabalhos acadêmicos foram realizados nessa linha na tentativa de entender o comportamento e a previsibilidade do mercado [16, 40, 41], assim como estratégias [42]. Nota-se que até hoje utiliza-se Random Walks para testar a hipótese de eficiência de mercados [43].

Outra abordagem utilizada são os Modelos Ocultos de Markov (do inglês Hidden Markov Model, ou HMM) [44]. Uma Cadeia de Markov é um processo estocástico que modela um sistema por meio de uma sequência finita de estados. A mudança ou a permanência em cada estado é determinada por probabilidades que dependem somente do estado atual. Em uma Cadeia de Markov, pressupõe-se que seus estados sejam observáveis, o que para algumas aplicações, pode não ser verdade. Nesse sentido surge o modelo HMM, que busca aprender sobre um processo não observável (oculto) a partir de um processo observável.

Em sua pesquisa, JADHAV et al [45] utiliza um modelo HMM para previsão do preço de fechamento do dia seguinte para ações FAANG². A partir da série histórica de preços OHLC³, seu modelo atinge uma eficiência de 97%-99%, calculado a partir do erro percentual absoluto médio⁴.

¹Processo aleatório definido pela equação $y_t = y_{t-1} + X$, onde X é uma variável aleatória e y é a variável resultante.

²Facebook, Amazon, Apple, Netflix, Google

³Open, High, Low, Close. Em português: Abertura, Máximo, Mínimo, Fechamento

⁴Mean Absolute Percentage Error (MAPE): $\frac{1}{N} \sum_{i=1}^N \frac{|Predicted(i) - Actual(i)|}{Actual(i)}$

Uma outra aplicação de modelos HMM é dada por DE ANGELIS et al [46], que criou uma metodologia a partir de índices da bolsa americana capaz de identificar períodos estáveis e instáveis (i.e. crises econômicas), assim como as probabilidades de transição entre um estado e o outro.

Por fim, pode-ser mencionar o uso de modelos ARCH⁵ (Autoregressive Conditional Heteroskedasticity). A ideia central está na modelagem de uma variância (volatilidade) condicional, ou seja, que muda de acordo o instante da série [47]. Essa característica se faz muito útil em séries que possuem períodos de alta volatilidade se alternando com períodos de baixa volatilidade. Para um modelo genérico ARCH(q), seja ϵ_t o erro (resíduo) no instante t e α_0 um ruído branco, pode-se descrever a variância condicional de acordo com a Equação 2.1.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad (2.1)$$

O modelo ARCH foi proposto por ENGLE em 1982 para estimar a variância da inflação do Reino Unido [48]. A partir daí, várias derivações surgiram, como por exemplo: GARCH ⁶ por BOLLERSLEV [49] em 1986, EGARCH ⁷ por NELSON em 1991, NGARCH ⁸ por HIGGINS e BERA[50] em 1992, TGARCH ⁹ por ZAKOIAN e RABEMANANJARA [51] em 1993, dentre outros. Alguns dos modelos da família ARCH podem ser encontrados nos trabalhos de FRANSES e DIJK [52], de MARCUCCI [53] e de ALBERG et al [54].

2.4.3 Modelos Baseados em Inteligência Artificial

Existem registros de estudos sobre inteligência artificial aplicados ao mercado financeiro por volta da década de 70 [55], porém ainda em um estágio embrionário devido às dificuldades de acesso à informação e de processamento computacional

⁵Em português: Heteroscedasticidade Condicional Auto-regressiva

⁶Generalised ARCH

⁷Exponential Generalised ARCH

⁸Non-linear Generalised ARCH

⁹Threshold Generalised ARCH

na época. Conforme ambos os fatores foram evoluindo, a dificuldade de se testar algoritmos cada vez mais complexos diminuiu.

Combinações de GARCH com ANN [56]

Referências Bibliográficas

- [1] B3, “B3 atinge 5 milhões de contas de investidores em renda variável em janeiro”, https://www.b3.com.br/pt_br/noticias/5-milhoes-de-contas-de-investidores.htm, (Acesso em 21 de Março de 2022).
- [2] INFOMONEY, “Robôs de investimentos já controlam mais de US\$ 200 bilhões ao redor do mundo”, <https://www.infomoney.com.br/onde-investir/robos-de-investimentos-ja-controlam-mais-de-us-200-bilhoes-ao-redor-do-mundo>, (Acesso em 22 de Março de 2022).
- [3] INFOMONEY, “No Brasil, robôs de investimento não conseguem bater melhores fundos”, <https://www.infomoney.com.br/onde-investir/no-brasil-robos-de-investimento-nao-conseguem-bater-melhores-fundos>, (Acesso em 22 de Março de 2022).
- [4] FERNÁNDEZ, A., “Artificial intelligence in financial services”, *Banco de Espana Article*, v. 3, pp. 19, 2019.
- [5] CVM, “Entendendo o Mercado de Valores Mobiliários”, https://www.investidor.gov.br/menu/primeiros_passos/entendendo_mercado_valores.html, (Acesso em 24 de Março de 2022).
- [6] BRASIL, “Lei nº 6.385, de 7 de dezembro de 1976. Dispõe sobre o mercado de valores mobiliários e cria a Comissão de Valores Mobiliários.”, http://www.planalto.gov.br/ccivil_03/leis/l6385.htm.
- [7] B3, “Uma das principais empresas de infraestrutura de mercado financeiro do mundo”, https://www.b3.com.br/pt_br/b3/institucional/quem-somos/, (Acesso em 24 de Março de 2022).

- [8] B3, “Ações”, https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/acoes.htm, (Acesso em 24 de Março de 2022).
- [9] BRASIL, “Lei nº 6.404, de 15 de dezembro de 1976. Dispõe sobre as Sociedades por Ações.”, Capítulo XC, Seção VII, http://www.planalto.gov.br/ccivil_03/leis/l6404consol.htm, (Acesso em 24 de Março de 2022).
- [10] CVM, “Lei 6.404/76: Exposição de Motivos”, Capítulo II, Seção I, <https://www.gov.br/cvm/pt-br/acesso-a-informacao-cvm/institucional/sobre-a-cvm/>, (Acesso em 24 de Março de 2022).
- [11] INVESTIMENTOS, X., “Mercado secundário: entenda as diferenças com o mercado primário”, <https://conteudos.xpi.com.br/aprenda-a-investir/relatorios/mercado-secundario/>, (Acesso em 24 de Março de 2022).
- [12] BRASIL, “Lei nº 6.404, de 15 de dezembro de 1976. Dispõe sobre as Sociedades por Ações.”, Capítulo XV, Seção II, Art. 176, http://www.planalto.gov.br/ccivil_03/leis/l6404consol.htm, (Acesso em 24 de Março de 2022).
- [13] BRASIL, “Lei nº 6.404, de 15 de dezembro de 1976. Dispõe sobre as Sociedades por Ações.”, Capítulo XXI, Seção IV, Art. 275, § 4º, http://www.planalto.gov.br/ccivil_03/leis/l6404consol.htm, (Acesso em 24 de Março de 2022).
- [14] INFOMONEY, “Proventos: O que são, como funcionam e como ganhar dinheiro com eles?”, <https://www.infomoney.com.br/guias/proventos/>, (Acesso em 24 de Março de 2022).
- [15] B3, “Posições vendidas no mercado de ações”, https://www.b3.com.br/pt_br/noticias/short-selling.htm, (Acesso em 24 de Março de 2022).
- [16] FAMA, E. F., “Efficient capital markets: A review of theory and empirical work”, *The Journal of Finance*, v. 25, n. 2, pp. 383–417, 1970.

- [17] INVESTOPEDIA, “Four Scandalous Insider Trading Incidents”, <https://www.investopedia.com/articles/stocks/09/insider-trading.asp#:~:text=Four> (Acesso em 25 de Março de 2022).
- [18] FAMA, E. F., FISHER, L., JENSEN, M., *et al.*, “The adjustment of stock prices to new information”, *International economic review*, v. 10, n. 1, 1969.
- [19] SHOSTAK, F., “In defense of fundamental analysis: A critique of the efficient market hypothesis”, *The Review of Austrian Economics*, v. 10, n. 2, pp. 27–45, 1997.
- [20] JUNG, J., SHILLER, R. J., “Samuelson’s dictum and the stock market”, *Economic Inquiry*, v. 43, n. 2, pp. 221–228, 2005.
- [21] SCHWAGER, J. D., *Market Sense and Nonsense: How the Markets Really Work (and how They Don’t)*. John Wiley & Sons, 2012.
- [22] FORBES, “Investing Basics: What Is A Market Index?”, <https://www.forbes.com/advisor/investing/stock-market-index/>, (Acesso em 28 de Março de 2022).
- [23] B3, “Ibovespa B3”, https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm, (Acesso em 28 de Março de 2022).
- [24] B3, “ETF de Renda Variável”, https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/etf-de-renda-variavel.htm, (Acesso em 28 de Março de 2022).
- [25] INVESTOPEDIA, “Fractional Share”, <https://www.investopedia.com/terms/f/fractionalshare.a> (Acesso em 28 de Março de 2022).
- [26] BULKOWSKI, T. N., *Fundamental Analysis and Position Trading: Evolution of a Trader*, v. 605. John Wiley & Sons, 2012.
- [27] MURPHY, J. J., *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin, 1999.

- [28] EDWARDS, R. D., MAGEE, J., BASSETTI, W. C., *Technical analysis of stock trends*. CRC press, 2018.
- [29] IBM, “Artificial Intelligence (AI)”, <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>, (Acesso em 4 de Abril de 2022).
- [30] MÜLLER, A. C., GUIDO, S., *Introduction to machine learning with Python: a guide for data scientists*. ”O’Reilly Media, Inc.”, 2016.
- [31] IBM, “Machine Learning”, <https://www.ibm.com/cloud/learn/machine-learning#:~:text=IBM>(Acesso em 4 de Abril de 2022).
- [32] ARTHUR, S., OTHERS, “Some studies in machine learning using the game of checkers”, *IBM Journal of research and development*, v. 3, n. 3, pp. 210–229, 1959.
- [33] KIM, K., *Electronic and algorithmic trading technology: the complete guide*. Academic Press, 2010.
- [34] KIRKPATRICK II, F. C. D., JULIE, R., “Dow theory”, *CMT Curriculum Level I 2022: An Introduction to Technical Analysis*, , 2022.
- [35] BOLLINGER, J., *Bollinger on Bollinger bands*. McGraw Hill Professional, 2002.
- [36] APPEL, G., DOBSON, E., *Understanding MACD*. Traders Press, 2007.
- [37] MORAES, A., *Se Afastando da Manada: Estratégias para vencer no Mercado de Ações*. Infomoney, 2016.
- [38] GODFREY, M. D., GRANGER, C. W., MORGENSTERN, O., “THE RANDOM-WALK HYPOTHESIS OF STOCK MARKET BEHAVIOR a”, *Kyklos*, v. 17, n. 1, pp. 1–30, 1964.
- [39] BACHELIER, L., “Théorie de la spéculation”. In: *Annales scientifiques de l’École normale supérieure*, v. 17, pp. 21–86, 1900.
- [40] SOLNIK, B. H., “Note on the validity of the random walk for European stock prices”, *The journal of Finance*, v. 28, n. 5, pp. 1151–1159, 1973.

- [41] COOPER, J. C., “World stock markets: Some random walk tests”, *Applied Economics*, v. 14, n. 5, pp. 515–531, 1982.
- [42] MALKIEL, B. G., *A random walk down Wall Street: the time-tested strategy for successful investing*. WW Norton & Company, 2019.
- [43] SAID, A., HARPER, A., “The efficiency of the Russian stock market: A revisit of the random walk hypothesis”, *Academy of Accounting and Financial Studies Journal*, v. 19, n. 1, pp. 42–48, 2015.
- [44] RABINER, L. R., “A tutorial on hidden Markov models and selected applications in speech recognition”, *Proceedings of the IEEE*, v. 77, n. 2, pp. 257–286, 1989.
- [45] JADHAV, A., KALE, J., RANE, C., *et al.*, “Forecasting FAANG Stocks using Hidden Markov Model”. In: *2021 6th International Conference for Convergence in Technology (I2CT)*, pp. 1–4, IEEE, 2021.
- [46] DE ANGELIS, L., PAAS, L. J., “A dynamic analysis of stock markets using a hidden Markov model”, *Journal of Applied Statistics*, v. 40, n. 8, pp. 1682–1700, 2013.
- [47] ENDERS, W., *Applied econometric time series*. John Wiley & Sons, 2008.
- [48] ENGLE, R. F., “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation”, *Econometrica: Journal of the econometric society*, pp. 987–1007, 1982.
- [49] BOLLERSLEV, T., “Generalized autoregressive conditional heteroskedasticity”, *Journal of econometrics*, v. 31, n. 3, pp. 307–327, 1986.
- [50] HIGGINS, M. L., BERA, A. K., “A class of nonlinear ARCH models”, *International Economic Review*, pp. 137–158, 1992.
- [51] RABEMANANJARA, R., ZAKOIAN, J.-M., “Threshold ARCH models and asymmetries in volatility”, *Journal of applied econometrics*, v. 8, n. 1, pp. 31–49, 1993.

- [52] FRANCES, P. H., VAN DIJK, D., “Forecasting stock market volatility using (non-linear) Garch models”, *Journal of forecasting*, v. 15, n. 3, pp. 229–235, 1996.
- [53] MARCUCCI, J., “Forecasting stock market volatility with regime-switching GARCH models”, *Studies in Nonlinear Dynamics & Econometrics*, v. 9, n. 4, 2005.
- [54] ALBERG, D., SHALIT, H., YOSEF, R., “Estimating stock market volatility using asymmetric GARCH models”, *Applied Financial Economics*, v. 18, n. 15, pp. 1201–1208, 2008.
- [55] FELSEN, J., “Artificial intelligence techniques applied to reduction of uncertainty in decision analysis through learning”, *Journal of the Operational Research Society*, v. 26, n. 3, pp. 581–598, 1975.
- [56] BILDIRICI, M., ERSIN, Ö. Ö., “Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange”, *Expert Systems with Applications*, v. 36, n. 4, pp. 7355–7362, 2009.