

Universidade de Brasília – UnB
Engenharia Elétrica

**Comparação da performance de algoritmos de Q-Learning e
Double Q-Learning no mercado financeiro**

Autor: Iago Campos Mameri

Orientador: Prof. Dr. Alexandre Ricardo Soares Romariz

Brasília, DF

2019



Iago Campos Mameri

Comparação da performance de algoritmos de Q-Learning e Double Q-Learning no mercado financeiro

Monografia submetida ao curso de graduação em Engenharia Elétrica da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Elétrica.

Universidade de Brasília – UnB

Orientador: Prof. Dr. Alexandre Ricardo Soares Romariz

Brasília, DF

2019

Resumo

Este trabalho tem como objetivo analisar a performance de agentes inteligentes treinados por meio de algoritmos de aprendizagem de máquina Q-Learning e Double Q-Learning em 17 ações da Bolsa de Valores de São Paulo com diferentes níveis de volatilidade e indicadores fundamentalistas. A performance dos agentes foi testada com ciclos de treinamento de alta, baixa e deslocamento lateral. Os agentes obtiveram uma performance melhor quando treinados em períodos de alta e superaram a valorização do índice Ibovespa em aproximadamente 1/3 dos casos. Foi possível perceber uma correlação entre o desempenho do algoritmo e a variação da volatilidade das ações analisadas, assim como a vantagem que algoritmos de Q-Learning possuem em ciclos de alta.

Palavras-chaves: Aprendizado por Reforço, Indicador Fundamentalista, Indicador Técnico, Ações, Q-Learning, Double Q-Learning, Aprendizagem de Máquinas

Abstract

This work aims at analyzing the performance of intelligent agents trained through Q-Learning and Double Q-Learning in 17 stocks of the Sao Paulo Stock Exchange with different levels of volatility and fundamentalist indicators. The agents' performance was evaluated when trained during bearish, bullish and neutral market cycles. The agents performed better when trained during bullish cycles and surpassed the Ibovespa index in approximately one-third of the cases. It was possible to identify a correlation between the performance of the algorithm and the variation of the volatility of the analyzed stocks, as well as the advantage that algorithms of Q-Learning have in bullish cycles.

Key-words: Reinforcement Learning, Fundamental Indicator, Technical Indicator, Stocks, Q-Learning, Double Q-Learning, Machine Learning

Lista de Ilustrações

Figura 1 - Valorização do Índice Ibovespa e Nasdaq de 1995 a 2018	18
Figura 2 - Matriz de classificação dos ativos	22
Figura 3 - Estrutura do gráfico de ações	36
Figura 4 - Componentes de Candlesticks	37
Figura 5 - Gráfico com tendências de alta, baixa e lateral	38
Figura 6 - Exemplo de padrões: W e Ombro-cabeço-ombro	39
Figura 7 - Exemplo de triângulo ascendente	40
Figura 8 - Exemplo de modelo linear	59
Figura 9 - Quantidade de empresas listadas na bolsa e disponibilidade de dados de 2003 a 2018	67
Figura 10 – Percentual de convergência, desvio padrão geral e desvio padrão para os casos dos algoritmos de Q-Learning e Double Q-Learning que convergiram, em diferentes épocas, operando PETR4, quando treinados nos ciclos de alta apresentados na Seção Análise e Coleta de Dados.....	71
Figura 11 – Performance média de 10 resultados dos algoritmos de Q-Learning e Double Q-Learning operando PETR4, quando submetidos a 1.000 iterações de treinamento com diferentes variáveis de estado	74
Figura 12 - Performance média de 10 resultados dos algoritmos de Q-Learning e Double Q-Learning operando PETR4, quando submetidos a 1.000 iterações de treinamento em diferentes ciclos da bolsa	75
Figura 13 - Performance média dos 10 retornos obtidos pelos algoritmos de Q-Learning e Double Q-Learning para as 17 ações, treinados nos 9 ciclos de alta apresentados na Seção Análise e Coleta de Dados	76
Figura 14 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a volatilidade máxima das ações	78
Figura 15 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a variação do indicador P/E das ações	79
Figura 16 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a variação do indicador P/VPA das ações.....	80
Figura 17 - Retorno, variação do indicador P/E e volatilidade máxima das ações	81

Figura 18 - Relação entre indicador P/E inicial, desvio padrão da volatilidade e retorno apresentado	82
Figura 19 - Comparação do retorno obtido pelos algoritmos de Q-Learning e Double Q-Learning, Ibovespa e Selic.....	83
Figura 20 - Comparação do retorno obtido pelos algoritmos de Q-Learning e Double Q-Learning e estratégia de Buy and Hold	83

Lista de Tabelas

Tabela 1 - Diferença entre principais tipos de captação para empresas	24
Tabela 2 - Pontos negativos e positivos acerca da entrada no mercado de capitais	25
Tabela 3 - Relação entre resultado, expectativa e movimento provável das ações	27
Tabela 4 – Dinâmica de emissões de ordem na boleta (1/4)	31
Tabela 5 - Dinâmica de emissões de ordem na boleta (2/4)	32
Tabela 6 - Dinâmica de emissões de ordem na boleta (3/4)	32
Tabela 7 - Dinâmica de emissões de ordem na boleta (4/4)	32
Tabela 8 – Boleta com ordens de compra e venda incompatíveis (1/2)	34
Tabela 9 - Boleta com ordens de compra e venda incompatíveis (2/2).....	34
Tabela 10 - Ciclos de treinamento utilizados.....	65
Tabela 11 - Critérios utilizados para escolha das ações envolvidas no estudo	68

Sumário

1. Introdução	17
1.1. Contexto.....	17
1.2. Definição do problema	20
1.3. Limitações do escopo.....	21
1.4. Divisão do texto.....	22
2. Desenvolvimento.....	24
2.1. Mercado de Capitais	24
O Mercado de Capitais.....	24
O Mercado de Ações.....	24
O Funcionamento do Mercado de Ações Secundário	26
Ações no Brasil	29
Um Exemplo da Dinâmica de Compra e Venda de Ações	31
Ordens de Compra e Venda de Ações	33
2.2. Análise Técnica.....	34
Fundamentos da Análise Técnica	34
Críticas à Análise Técnica.....	35
Estrutura dos Gráficos.....	36
Análise Gráfica.....	37
Padrões Gráficos	38
Indicadores Técnicos	41
2.3. Análise Fundamentalista	43
Introdução à Análise Fundamentalista	44
Críticas à Análise Fundamentalista.....	44
Metodologias Fundamentalistas: Introdução.....	45
Metodologias Fundamentalistas: Comparação de Preço e Valor.....	45

Metodologias Fundamentalistas: Comparação de Indicadores.....	50
3. Aprendizagem de Máquina	55
3.1. O Surgimento da Inteligência Artificial.....	55
3.2. Inteligência Artificial.....	56
3.3. Aprendizagem de Máquina	57
3.4. Algoritmos Fundamentais.....	58
3.5. Aplicações.....	60
3.6. Q-Learning	60
3.7. Double Q-Learning	62
4. Materiais e Métodos	64
4.1. Materiais.....	64
4.2. Métodos	64
Análise e Coleta de Dados.....	64
Desenvolvimento do Algoritmo.....	68
Avaliação da Performance do Algoritmo	72
5. Análises e Discussão	73
5.1. Comparação da performance em relação a quantidade e tipos de variáveis de estados utilizados	73
5.2. Comparação da performance para diferentes ciclos de treinamento ..	74
5.3. Comparação de performance entre Q-Learning e Double Q-Learning	76
5.4. Comparação da performance para diferentes níveis de volatilidade...	77
5.5. Comparação da performance para diferentes P/E e P/VPA.....	78
5.6. Matriz de classificação das ações.....	80
5.7. Benchmarks de performance	82
6. Conclusões e Próximos Passos.....	84
7. Referencial Bibliográfico.....	86

1. Introdução

1.1. Contexto

A capacidade dos seres humanos de utilizar o raciocínio lógico para imaginar o futuro e avaliar os possíveis cenários é um dos principais fatores que diferenciam a espécie humana dos outros seres vivos (HARARI, 2015). Essa habilidade tem sido um instrumento importante não apenas para a manutenção da espécie, mas também para o desenvolvimento dela, pois permite ao ser humano:

- Avaliar e ponderar as possíveis consequências de determinados eventos (GIANNETTI, 2012); e
- Preparar-se para possíveis mudanças e se adaptar a elas, conforme a necessidade e os objetivos dos indivíduos;

Essa capacidade de avaliação de possíveis cenários e a tentativa de antecipar eventos antes de acontecerem ocorre, principalmente, a partir de dois processos:

- Identificação dos fatores influenciadores e causadores;
- Determinação das relações de impacto entre os fatores e o resultado;

A partir desses procedimentos, o ser humano é capaz de executar simulações de possíveis eventos futuros. Com a evolução tecnológica, o desenvolvimento da estatística aplicada e otimização da performance dos computadores, o ser humano se tornou capaz de desenvolver modelos computacionais que possibilitam a análise do funcionamento e da performance de sistemas cada vez mais complexos (NANCE & SARGENT, 2002), que incluem, por exemplo, o mercado financeiro.

Apesar de se basear em uma lei básica, da oferta e da demanda, os mercados financeiros são sistemas complexos, dada a sua abrangência, a quantidade de fatores que influenciam sua movimentação e o envolvimento de diferentes agentes, que podem responder de diversas maneiras a um mesmo evento (WEAVER, 1948).

Apesar de complexo, o desenvolvimento de sistemas capazes de simular a movimentação dos mercados financeiros é alvo de interesse de diversas pessoas e empresas, dada a importância desses sistemas para a economia mundial e a possibilidade de grandes lucros em curtos espaços de tempo.

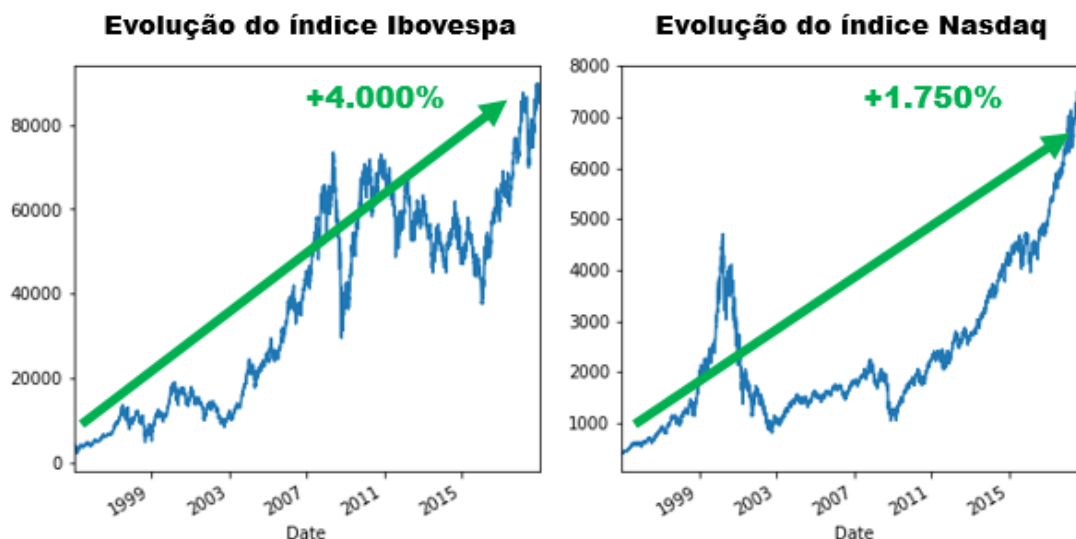


Figura 1 - Valorização do Índice Ibovespa e Nasdaq de 1995 a 2018

Como se pode perceber a partir da Figura 1, investidores são capazes de fazer fortunas se conseguirem de escolher adequadamente seus ativos, operar de maneira racional e ter paciência para colher os frutos dos investimentos no futuro (GRAHAM, 2003). Entretanto, grande parte dos investidores individuais, principalmente os que utilizam técnicas de trading, tende a perder dinheiro nos mercados financeiros.

O comportamento dos seres humanos no mercado financeiro já foi estudado por diversos especialistas e percebe-se que existem 3 fatores que contribuem significativamente para esse baixo desempenho dos investidores individuais. Os fatores são apresentados a seguir por ordem de importância.

O primeiro fator está relacionado à influência da emoção na tomada de decisões. Teoricamente, a escolha dos ativos que serão alvo de investimentos deveria ser uma decisão predominantemente racional. Entretanto, os processos de decisão são constantemente influenciados pelas emoções (DAMASIO, 1994). Isso afasta o ser humano de um modelo completamente racional de tomada de

decisão e coloca a escolha dos ativos sob influência das emoções sentidas pelo investidor.

Dessa maneira, o investidor está sujeito a tomar decisões que não fazem sentido do ponto de vista lógico, o que pode resultar na perda de oportunidades e, no pior caso, no prejuízo e na conseqüente perda de patrimônio (GRAHAM, 2003). Um exemplo dessa influência da emoção na tomada de decisão ocorre nos ciclos de supervalorização e desvalorização da bolsa. Durante os ciclos de supervalorização, a ganância e o otimismo dos investidores costuma resultar na valorização de projeções e promessas otimistas em detrimento de indicadores econômicos e da realidade local. Por outro lado, em movimentos de baixa do mercado, a situação se inverte de modo que o pessimismo dos investidores leva empresas a perderem até 25% do valor de mercado em um único dia, como o que aconteceu com a companhia Vale S.A. no dia da tragédia de Brumadinho.¹

O segundo fator está relacionado à importância dada às emoções negativas em relação às positivas e o impacto desse resultado na tomada de decisão. A assimetria dos sentimentos positivo e negativo foi alvo de estudo de diversos pesquisadores (ANDERSON, 1965; PEETERS & CZAPINSKI, 1990; SKOWRONSKI & CARLSTON, 1989). Durante essas pesquisas, foi verificado que impressões negativas são formadas de modo mais rápido e são mais resistentes do que impressões positivas. Dessa maneira, eventos que possuem um significado negativo predominante terão um impacto maior no indivíduo.

O principal impacto desse fator está na aversão ao risco dos investidores. Dessa maneira, investidores tendem a fazer aplicações mais seguras, com um rendimento menor e, algumas vezes, com rendimento real negativo, e evitar investimentos mais arriscados, que podem trazer rendimentos superiores. Por outro lado, outra conseqüência desse fator está na quantidade de investidores. Isso é visível principalmente no Brasil, tendo em vista que apenas 0,25% da população possui aplicações no mercado de ações, um número sem relevância comparado aos 70% da população norte americana.

¹ Conforme apresentado na notícia: <https://g1.globo.com/economia/noticia/2019/01/28/perda-de-r-70-bilhoes-em-1-dia-da-vale-e-a-maior-da-historia-do-mercado-brasileiro.ghtml>, Acesso em 22/02/2019.

O terceiro fator está relacionado à assimetria das informações. Eugene Fama, em meados da década de 60, propôs a Teoria dos Mercados Eficientes (EMH, do inglês *Efficient Market Hypothesis*). De acordo com essa teoria, o mercado precifica e reflete completamente as informações disponíveis (FAMA, 1965). Dessa maneira, caso alguns investidores tenham acesso às informações antes de outros (técnica ilegal, conhecida como *insider trading*), esses investidores teriam conhecimento dos movimentos do mercado antes dele ocorrer e poderiam tirar vantagem dessa situação.

No longo prazo, o efeito da assimetria de informações é minimizado, considerando que isso não impacte os fatores levados em consideração na tomada de decisão dos investimentos. Entretanto, a curto prazo, esse elemento pode ter impactos significantes no rendimento, a partir das variações diárias dos preços dos ativos, e na tomada de decisão dos investidores, considerando que o preço dos ativos pode ficar mais ou menos atrativo.

Dessa maneira, tanto a baixa educação dos investidores individuais, como o reduzido contato com as empresas de capital aberto contribuem para que o rendimento desses investidores, em termos gerais, seja reduzido em comparação com investidores profissionais, fundos de investimento e outras instituições.

1.2. Definição do problema

Percebe-se que a presença de sentimentos na tomada de decisão é um dos fatores que impactam negativamente o rendimento dos investidores no mercado financeiro.

Dessa maneira, o presente projeto busca desenvolver um algoritmo embasado na análise técnica utilizando inteligência artificial para operar diferentes ativos no mercado de ações brasileiro e, assim:

- Avaliar a existência de fatores que podem contribuir para melhorar a performance dos algoritmos;
- Avaliar qual algoritmo apresenta a melhor performance diante do cenário investigado; e

- Avaliar a classe de ativos mais atraente para aplicação desse tipo de algoritmo, de acordo com parâmetros de liquidez e indicadores fundamentalistas que caracterizam o ativo.

1.3. Limitações do escopo

O presente projeto considera as seguintes limitações de escopo: de tempo, de mercado, de ativos e das técnicas aplicadas no algoritmo.

O estudo tem como alvo a análise de 17 ativos do mercado brasileiro. Os algoritmos serão treinados em diferentes períodos entre 2003 e 2012 e serão avaliados entre os anos de 2012 e 2018, utilizando técnicas de análise técnica.

O mercado brasileiro foi escolhido com o objetivo de mostrar, aos não investidores brasileiros, que é possível obter ganhos monetários em diversas etapas do ciclo do mercado de ações. Dessa maneira, espera-se incentivar a população brasileira a incluir os estudos do mercado de ações entre suas atividades diárias e envolvê-lo no seu portfólio de investimentos.

Foram escolhidas 17 ações brasileiras para compor o alvo do estudo da performance do algoritmo. Dessa maneira, é possível combinar as características de cada uma e formar 4 grupos de análise, conforme a matriz exibida na Figura 2.

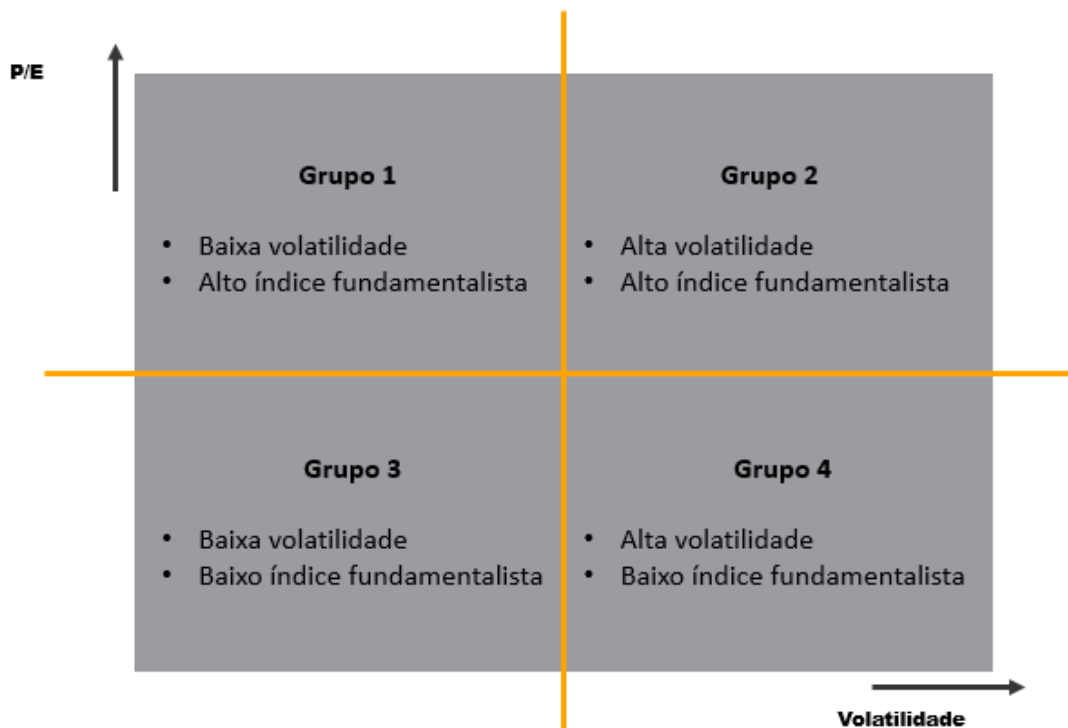


Figura 2 - Matriz de classificação dos ativos

Por fim, o algoritmo utilizará conceitos da aprendizagem de máquina aplicada a análise técnica de ações, de modo a excluir o fator sentimental presente nas decisões tomadas pelo ser humano, que pode comprometer o rendimento do portfólio.

1.4. Divisão do texto

O texto é dividido, sucintamente, em 6 capítulos, cada um com suas respectivas subdivisões:

- Introdução: busca-se apresentar as motivações para o tema escolhido, bem como o objetivo geral do estudo e as restrições de escopo;
- Desenvolvimento: busca-se apresentar os conceitos teóricos dos componentes utilizados no estudo. Dessa maneira, serão abordados:
 - Mercado de Capitais;
 - Diferenças entre Análise Técnica e Fundamentalista;

- Aprendizagem de Máquina: busca-se apresentar os conceitos mais importantes relacionados à aprendizagem de máquina, com foco em algoritmos de Q-Learning e Double Q-Learning;
- Materiais e Métodos: busca-se apresentar como foram levantados os artigos de referência do presente estudo, bem como a seleção dos ativos analisados e a estruturação do algoritmo de investimento;
- Análises e Discussão: procura-se apresentar as análises feitas e as consequências dos resultados obtidos; e
- Conclusão e Próximos passos: procura-se apresentar uma visão geral dos resultados, os principais complicadores da implementação das análises e os próximos passos.

2. Desenvolvimento

2.1. Mercado de Capitais

O Mercado de Capitais

O Mercado de Capitais compreende um conjunto de instituições e instrumentos que negociam títulos de valores mobiliários, nos quais estão incluídas ações de empresas de capital aberto, dívidas de ações, derivativos, opções, entre outros instrumentos. Esse tipo de mercado permite que empresas tenham condições de captar recursos de maneira mais direta com os investidores e realizem investimentos de médio e longo prazo.

Essa captação de investimentos pode ocorrer de duas formas, a partir da emissão de dívidas e de ações. Essas duas maneiras de captação apresentam algumas características diferentes, ilustradas na Tabela 1.

Tipo de captação	Ação	Dívida
Forma de pagamento	Valorização da ação ou dividendos	Juros sobre o principal
Relacionamento com a empresa	Acionista	Financiador
Tempo de relacionamento	Indefinido	Até o vencimento da dívida
Liquidez	Maior	Menor

Tabela 1 - Diferença entre principais tipos de captação para empresas

Como o mercado de ações é o foco principal do presente projeto, ele será alvo de maior aprofundamento daqui em diante.

O Mercado de Ações

Designa-se por mercado o local no qual agentes econômicos procedem à troca de bens por uma unidade monetária ou por outros bens (ROSTOW, 1959). No mercado acionário, os bens trocados são as ações, que, de acordo com a Comissão de Valores Mobiliários – CVM, são valores mobiliários que correspondem à menor parcela do capital social de uma companhia ou sociedades anônimas.

O surgimento do Mercado de Ações ocorreu em meados do século XVII, em Amsterdam, quando os diretores da Companhia Holandesa das Índias Ocidentais – VOC (do holandês, *Vereenigde Oost-Indische Compagnie*) resolveram tornar as ações da empresa disponíveis ao público (PETRAM, 2014).

O mercado de ações é constituído por duas partes: o mercado primário e o mercado secundário. O mercado primário consiste na emissão das ações pela empresa emissora e a consequente venda dessas ações para os investidores. Essa emissão de ações pode ser de dois tipos:

- Oferta Pública Inicial (IPO, do inglês *Initial Public Offer*), que consiste na abertura de capital da empresa, ou seja, quando ela emite suas primeiras ações no mercado; ou
- Novas emissões, que consiste na venda de mais ações de empresas que já são de capital aberto.

A emissão de ações no mercado primário é o mecanismo pelo qual os controladores vendem parte da empresa para o público em geral e podem, portanto, utilizar o capital levantado para realizar investimentos, expandir e manter as operações. A entrada no mercado de capitais por meio da venda de ações possui, como toda decisão, pontos positivos e negativos, conforme apresentado na Tabela 2.

Pontos positivos	Pontos negativos
Aumento da visibilidade na mídia	Necessidade de transparência das informações
Aumento da liquidez de ações	Possibilidade de aquisições hostis
Maior credibilidade no mercado	Pressão por entrega de resultados aos acionistas
Possibilidade de ganho de valor da empresa	Valor da empresa sujeito à especulação do mercado

Tabela 2 - Pontos negativos e positivos acerca da entrada no mercado de capitais

O mercado secundário, por outro lado, compreende a venda e a compra das ações já disponíveis do mercado por investidores. A venda e a compra de uma ação são, em geral, motivadas por dois fatores:

- Valor da empresa; e

- Preço da empresa.

Os conceitos e as metodologias utilizadas para avaliar o valor e o preço de uma empresa serão discutidos em maior detalhe na Seção [Análise Fundamentalista](#). Mas, de maneira geral, investidores buscam avaliar a tendência das empresas se tornarem mais ou menos valiosas no futuro a partir de diversos fatores e decidir, a partir do preço que ela se encontra no momento da decisão, se vale a pena adquirir uma parte da empresa (no caso da tendência de valorização) ou vender uma parte da empresa (no caso da tendência de desvalorização).

O Funcionamento do Mercado de Ações Secundário

O funcionamento do Mercado de Ações Secundário envolve, de maneira geral, empresas, investidores e corretoras. As empresas, geralmente, não participam das movimentações do mercado secundário, a não ser em dois momentos:

- Recompra de ações: empresas decidem por recomprar totalmente ou parcialmente suas ações disponíveis no mercado;
- e
- Diluição de ações: empresas decidem por dividir as ações atualmente disponíveis no mercado com o objetivo de terem maior liquidez.

Exceto durante esses dois momentos, as empresas participam do mercado secundário indiretamente por meio da apresentação de resultados e sofrem as consequências do mercado secundário diretamente na perda ou ganho de valor.

Conforme dito anteriormente, a compra ou venda de ações é, em grande parte, baseada nas expectativas futuras dos investidores em relação às empresas. De acordo com a Comissão de Valores Mobiliários – CVM – as empresas de capital aberto têm a obrigação de divulgar trimestralmente os resultados obtidos no último trimestre de exercício. Dessa maneira, a divulgação dos resultados das empresas em relação às expectativas dos investidores e as prováveis consequências são apresentadas na Tabela 3.

Relação resultado - expectativa	Movimento geral das ações
Resultado abaixo do esperado	Queda de preços
Resultado em linha com o esperado	Estabilidade de preços
Resultado acima do esperado	Alta de preços

Tabela 3 - Relação entre resultado, expectativa e movimento provável das ações

Desse modo, dependendo do resultado apresentado pelas empresas ou pela mudança na expectativa desses resultados por parte dos investidores devido a algum fator, informação ou evento, as empresas impactam o mercado e a consequente variação no seu valor de mercado reflete esse impacto.

A figura do investidor é fundamental para o a existência e o funcionamento dos mercados de capitais. Considerando-se os diferentes níveis de competência financeira e influência no mercado, os investidores podem ser classificados em dois grupos: investidores institucionais e individuais.

Os investidores individuais são pessoas físicas ou jurídicas que participam diretamente no mercado, seja comprando ou vendendo ações, por si próprios, assumindo sozinhos o risco (PINHEIRO, 2001). Normalmente, possui menos know-how e menor volume de patrimônio. Por isso, investidores individuais podem ingressar em fundos ou clubes de investimento, para adquirir mais conhecimento ou para possibilitar um gerenciamento de patrimônio mais profissional.

Investidores individuais possuem um nível de fragilidade relevante no mercado que se deve a 3 principais fatores: baixa disponibilidade, carência de conhecimento e distância dos grandes agentes do mercado.

A baixa disponibilidade dos investidores individuais está relacionada ao objetivo pelo qual eles entraram no mercado financeiro. Geralmente, investidores individuais buscam uma renda adicional no mercado de ações e possuem uma outra fonte principal de renda. Com isso, é fácil perceber que o foco da maioria dos investidores individuais é manter sua fonte de renda principal e investir o excedente no mercado financeiro para, se possível, gerar uma renda adicional. Entretanto, ao dedicar uma grande parte do tempo diário à fonte de renda principal, os investidores individuais se colocam em uma posição fragilizada em relação aos investidores institucionais, pois dispõem de

um tempo menor para acompanhar, estudar e operar o mercado, e isso impacta, principalmente, o nível de conhecimento.

Considerando a correlação positiva entre o tempo de estudo e o desempenho acadêmico (JEZ & WASSMER, 2013) e a menor quantidade de tempo dedicado aos estudos do mercado, a maioria dos investidores individuais carecem de conhecimento sobre o mercado.

Por fim, a distância entre os investidores individuais e os grandes agentes do Mercado de Ações também contribui para a intensificação desse déficit educacional e para o acesso a informação. Considerando que investidores institucionais estão em constante contato com empresas de capital aberto e com gestores de grandes fundos de investimento, os investidores individuais são marginalizados pelo sistema financeiro e isso intensifica ainda mais a situação de fragilidade dessas pessoas.

Os investidores institucionais, por outro lado, são pessoas jurídicas que, por determinações governamentais, são obrigadas a investir parte de seu capital no mercado de ações, constituindo uma carteira de investimento. Esses investidores são, geralmente, dotados de maior conhecimento, experiência e patrimônio, e, por isso, possuem maior influência no Mercado de Ações.

Por agregarem profissionais com formações acadêmicas e experiências profissionais orientadas para finanças, os investidores institucionais se equipam com um alto nível de conhecimento acerca do mercado acionário, o que contribui para uma tomada de decisão mais bem fundamentada e resultados mais positivos quando comparados aos dos investidores individuais.

Além disso, a disponibilidade dos investidores institucionais não é um ponto de atenção, dado que a função dos profissionais atuantes nessas instituições é voltada para o estudo, análise e acompanhamento do mercado.

Por último, os investidores não apenas estão próximos dos agentes tomadores de decisão, mas também, podem constituir esses próprios agentes. Isso é resultado do tamanho do patrimônio e da habilidade de análise e decisão desse tipo de investidor.

Além do contato próximo com o alto nível das empresas de capital aberto, governos e gestores de fundos de pensão, fundos de investimento, fundos de hedge, entre outros, os investidores institucionais conseguem, a depender do posicionamento do mercado, influenciar o movimento de preços (LAKONISHOK, SCHLEIFER & VISHNY, 1992). Isso pode ocorrer por dois caminhos:

- Participação ativa em empresas: o grande patrimônio gerido por investidores institucionais pode, em alguns casos, ser utilizado para adquirir uma participação relevante em empresas de capital aberto e, com isso, garantir uma certa relevância na tomada de decisão da empresa. Dessa maneira, por poder influenciar o modo de operação de grandes empresas, investidores institucionais são capazes de tomar decisões que podem alterar a performance das companhias investidas e, conseqüentemente, o preço das ações (BRAV, JIANG & KIM, 2013);
- Realocação de capital: o grande patrimônio gerido por investidores institucionais é capaz de influenciar a precificação de ativos, ao passo que o volume transacionado possui uma correlação positiva com a variação de preços (KARPOFF, 1987). Dessa maneira, realocações de capital relevantes decorrentes de uma possível mudança na estratégia de fundos podem influenciar o movimento do mercado.

Além do papel desempenhado pelos agentes compradores e vendedores, o funcionamento do Mercado de Ações depende também da existência das empresas corretoras. O papel dessas empresas é conectar os vendedores e os compradores no Mercado de Ações. Dessa maneira, elas criam e aumentam a liquidez no mercado e executam as ordens de compra e venda emitidas pelos clientes, cobrando destes uma taxa de corretagem.

Ações no Brasil

Conforme mencionado anteriormente, as ações são valores mobiliários emitidos por sociedades anônimas que representam a menor parcela do seu capital social. De maneira geral, as ações são representadas no mercado na

forma de *tickers*. Os *tickers* são os identificadores das ações e são compostos por duas partes:

- A primeira parte do *ticker* é composta por 4 caracteres alfanuméricos e fazem referência à empresa emissora da ação;
- A segunda parte do *ticker* é composta por números, que indicam o tipo de papel que a ação representa.

Para entender melhor a diferença entre os tipos de papéis que as ações podem representar, é necessário introduzir o conceito de dividendos.

Grande parte das empresas, e principalmente as empresas de capital aberto, têm como principal objetivo a geração de lucro, para dar retorno aos acionistas. Esse lucro gerado pelas empresas é contabilizado trimestralmente e é resultado da diferença entre a quantidade de receita gerada pela empresa durante o período e os gastos que a empresa teve durante o período.

O lucro, entretanto, pode ser utilizado pelas empresas de duas formas. Ele pode tanto ser reinvestido na empresa para expandir operações, pagar dívidas, entre outras aplicações, como ser distribuído aos acionistas na forma de dividendos.

A forma de utilização do lucro por parte das empresas é definida de acordo com a estratégia da empresa. No Brasil, existem tanto empresas brasileiras com políticas de distribuição de dividendos regulares, como empresas com distribuição de dividendos esporádicas, que dependem do resultado da empresa.

As empresas com distribuição de dividendos regulares possuem dois tipos de ações emitidas: ações preferenciais e ações ordinárias, que se diferenciam principalmente pela preferência no recebimento dos dividendos e no poder de voto.

- As ações preferenciais (tipo PN) no mercado brasileiro são representadas geralmente pela numeração 4 (como, CMIG4, PETR4, VALE4, entre outras). Esse tipo de ação tem prioridade no recebimento dos dividendos das empresas;

- As ações ordinárias (tipo ON) no mercado brasileiro são representadas geralmente pela numeração 3 (como, CMIG3, PETR3, VALE3, entre outras). Esse tipo de ação não possui prioridade no recebimento de dividendos, mas, por outro lado, possui poder de voto.

Um Exemplo da Dinâmica de Compra e Venda de Ações

A dinâmica de negociação da compra e venda de ações no mercado segue a lei da oferta e da demanda. Essa lei determina que o preço de um determinado bem ou produto é resultado da relação entre a oferta daquele produto (por parte dos vendedores) e a demanda por aquele produto no mercado (SMITH, 1996).

A seguir, exemplifica-se a dinâmica de compra e venda de um lote (100 unidades) da ação PETR4 em uma corretora do mercado de ações brasileiro.

Supõe-se um cenário composto por 4 investidores (C1, C2, V1 e V2). Os investidores C1 e C2 estão interessados na compra de ações da PETR4 e os investidores V1 e V2 estão interessados na venda de ações da PETR4. Abaixo, o racional de cada investidor do cenário e o status das boletas que abrigam as ordens de compra, venda e status dos investidores são apresentados.

- O investidor C1, por algum motivo, acredita que o preço das ações de PETR4 está abaixo do que a empresa realmente vale e ele deseja se aproveitar da situação do mercado para lucrar com o movimento potencial de alta dessa ação. Como ele acha que essa ação pode alcançar o valor de R\$30,00 em um determinado período, ele está disposto a pagar até R\$27,50 por ação e emite uma ordem de compra de um lote (100 unidades) de PETR4 a R\$27,50, como ilustra a Tabela 4.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	PETR4	Compra	100	27,50	Pendente

Tabela 4 – Dinâmica de emissões de ordem na boleta (1/4)

- O investidor V1 adquiriu ações da PETR4 há um ano atrás e acredita que existem outras oportunidades melhores no mercado

atualmente para serem exploradas. Esse investidor, portanto, deseja vender suas ações da PETR4 e utilizar o dinheiro para realizar outras aplicações. Entretanto, ele acredita que as ações da empresa valham R\$28,00 e emite uma ordem de venda de um lote de PETR4 a R\$28,00, como mostra a Tabela 5.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	PETR4	Compra	100	27,50	Pendente
V1	PETR4	Venda	100	28,00	Pendente

Tabela 5 - Dinâmica de emissões de ordem na boleta (2/4)

- O investidor C2, atento aos preços do mercado, percebe que existe uma disparidade entre as ordens de compra e venda de PETR4. Ele acredita que o valor das ações da empresa devem subir, mas está disposto a pagar o que o investidor V1 está pedindo por suas ações. O investidor C2, portanto, emite uma ordem de compra para um lote de ações de PETR4 a R\$28,00, conforme apresentado na Tabela 6.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	PETR4	Compra	100	27,50	Pendente
V1	PETR4	Venda	100	28,00	Executada
C2	PETR4	Compra	100	28,00	Executada

Tabela 6 - Dinâmica de emissões de ordem na boleta (3/4)

- O investidor V2, por outro lado, acha que as ações da PETR4 valem, no mínimo, R\$30,00 atualmente e não tem pressa para vendê-las. Mesmo vendo que existem compradores no mercado que estão dispostos a pagar até R\$27,50 pelas ações, o investidor V2 decide emitir uma ordem de venda de um lote de ações de PETR4 a R\$30,00, como ilustra a Tabela 7.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	PETR4	Compra	100	27,50	Pendente
V1	PETR4	Venda	100	28,00	Executada
C2	PETR4	Compra	100	28,00	Executada
V2	PETR4	Venda	100	30,00	Pendente

Tabela 7 - Dinâmica de emissões de ordem na boleta (4/4)

Ordens de Compra e Venda de Ações

No tópico anterior, foi exemplificada uma situação onde existiam 4 investidores agentes no mercado que negociavam ações da PETR4. Apesar da simplicidade da dinâmica que foi apresentada para retratar algumas situações normais de operação do mercado, é interessante explorar outros conceitos relacionados às ordens de compra e venda de ações.

As ordens de compra e venda são comandadas pelos investidores e enviadas ao mercado pelas corretoras, responsáveis pela comunicação entre as diferentes partes envolvidas. As ordens são caracterizadas por 5 fatores:

- Tipo: compra ou venda;
- Ação: *ticker* da ação que está sendo negociada.
- Quantidade: quantidade de ações envolvida na execução da ordem;
- Preço: valor de compra ou de venda emitido na ordem; e
- Status: mostra se a ordem foi executada, cancelada ou se está pendente.

A quantidade de ações é um fator que deve ser melhor explicado. As ações são geralmente negociadas em lotes ou frações de lotes. Um lote comum de ações contém 100 unidades daquela ação. Portanto, o valor total da ordem é o valor da ação multiplicado por 100. No caso de investidores com restrição de capital, é possível negociar, apenas para alguns tipos de ações, frações de lotes. Frações de lotes de uma determinada ação são indicadas por um “F” no final do *ticker*. Nesse caso, por exemplo, o investidor não negociaria ações PETR4, e sim PETR4F.

Outro fator que merece atenção é o status das ordens. Uma ordem é executada quando existe uma ordem compatível com ela e que permite a execução dela. As ordens podem ser executadas completamente ou parcialmente, a depender das quantidades envolvidas. Por exemplo, supõe-se que existem duas investidoras C1 e V1, que negociam as ações VALE3. Na Tabela 8, são apresentadas as ordens emitidas por cada uma delas.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	VALE3	Compra	100	50,00	Pendente
V1	VALE3	Venda	200	50,00	Pendente

Tabela 8 – Boleta com ordens de compra e venda incompatíveis (1/2)

Percebe-se que, apesar de as ordens serem compatíveis no quesito preço e ação, elas não são compatíveis no quesito quantidade. Nesse caso, a ordem com menor quantidade de ações será executada, e a ordem com maior quantidade de ações será segmentada em duas: uma executada e a outra, pendente. Na Tabela 9, é apresentado o resultado da operação.

ID	Ação	Tipo	Quantidade	Preço	Status
C1	VALE3	Compra	100	50,00	Executada
V1	VALE3	Venda	100	50,00	Executada
V1	VALE3	Venda	100	50,00	Pendente

Tabela 9 - Boleta com ordens de compra e venda incompatíveis (2/2)

Por último, ainda em termos do status das ordens, normalmente, as ordens de compra e venda possuem validade de 1 dia. Ou seja, caso as ordens não sejam executadas no dia de emissão, elas são canceladas. Além disso, as ordens podem ser canceladas manualmente pelo próprio investidor no caso de ainda não terem sido executadas.

2.2. Análise Técnica

As duas principais escolas de pensamento que direcionam os investimentos no mercado financeiro são a análise técnica e a análise fundamentalista. Nesta seção, busca-se apresentar as bases da análise técnica e os indicadores mais utilizados por essa abordagem.

Fundamentos da Análise Técnica

A análise técnica consiste no estudo do movimento do mercado, principalmente por meio de gráficos, com o objetivo de prever os movimentos futuros dos preços dos ativos (MURPHY, 1999). No contexto de ações de empresas, o movimento do mercado é formado por dois componentes: preço e volume.

A análise técnica baseia-se em 3 premissas que, juntas, compõem seu racional de investimento:

1. O movimento de mercado desconta tudo;
2. Os preços se movimentam conforme as tendências; e
3. A história se repete.

A primeira premissa se baseia no fato de que, de acordo com a análise técnica, tudo que pode afetar o preço de uma ação – política, fundamentos, ambiente, humor, entre outros fatores – é refletido no movimento do mercado. Dessa maneira, para a análise técnica, se tudo que afeta o mercado é refletido no preço, o estudo dos preços é suficiente para prever os movimentos futuros.

A segunda premissa embasa a abordagem de seguir tendências gráficas. O propósito de estudar o movimento dos preços no mercado é identificar tendências em um estágio inicial de desenvolvimento e investir de acordo com essa tendência.

A terceira e última premissa da análise técnica se baseia na repetição da história. A análise técnica busca identificar padrões de alta e de baixa que, no fundo, refletem a psicologia humana em relação ao mercado (PRING, 1991). Dessa maneira, assume-se que a psicologia humana tende a permanecer a mesma e, por isso, entende-se que para prever o futuro, é necessário compreender o passado.

Críticas à Análise Técnica

A principal crítica à análise técnica deriva da Hipótese de Caminhadas Aleatórias (do inglês, *Random Walk Hypothesis*). Essa hipótese defende que a mudança de preços do mercado é um evento independente e que, por isso, são movimentos aleatórios e imprevisíveis (COOTNER, 1964).

A Hipótese de Caminhadas Aleatórias é consistente com a Hipótese de Mercados Eficientes (do inglês, *Efficient Market Hypothesis*), que defende que os preços refletem todas as informações disponíveis no mercado e que apenas uma estratégia baseada em informações privilegiadas pode gerar retornos acima do mercado (FAMA, 1965).

Dessa maneira, de acordo com essas hipóteses, não é possível obter retornos acima do mercado a partir do estudo de tendências gráficas do passado, pois os movimentos do mercado são imprevisíveis.

A análise técnica, entretanto, defende que, se o movimento dos mercados fosse realmente aleatório, nenhuma técnica de predição iria funcionar e que o entendimento geral do movimento passado é um importante fator na previsão de movimentos futuros do mercado (MURPHY, 1999).

Estrutura dos Gráficos

A Figura 3 exibe um gráfico diário da ação preferencial da Petrobrás no dia 8 de abril de 2019, às 12:08. Esse gráfico servirá como base para a análise dos componentes presentes em um gráfico de ações.



Figura 3 - Estrutura do gráfico de ações

No gráfico, o título revela a empresa que emitiu a ação, o ticker e as informações de preço diárias da ação analisada. O eixo vertical apresenta o intervalo de preços da ação conforme o período selecionado e o eixo horizontal é composto de duas partes: as escalas de divisão do eixo e as informações de volume negociado.

A escala de divisão do eixo varia conforme o escopo temporal selecionado para a ação. Além da segmentação temporal, o eixo horizontal também contém informações a respeito do volume negociado. As informações do volume são apresentadas de dois modos: tamanho e cor da barra. A partir do tamanho da barra, é possível verificar a quantidade de ações negociadas durante

o intervalo e, a partir da cor da barra, é possível verificar se as negociações resultaram em uma queda ou aumento do preço da ação.

A unidade básica dos gráficos utilizados na análise técnica se são as varas de vela (do inglês, *Candlestick*). Esse gráfico é utilizado pelos analistas técnicos, pois possui uma quantidade de informações apresentadas maior que outros tipos de gráfico, o que enriquece a análise. A Figura 4 apresenta dois *Candlesticks* e as informações apresentadas.

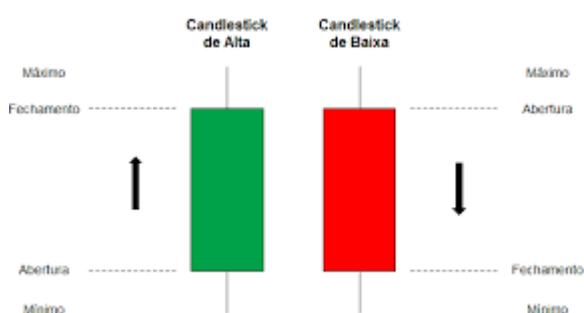


Figura 4 - Componentes de *Candlesticks*

Conforme percebe-se pela figura acima, cada *Candlestick* apresenta 5 informações: preço máximo, mínimo, de abertura e de fechamento e tipo de variação (aumento ou decréscimo de preço).

Análise Gráfica

Como mencionado anteriormente, o principal propósito da análise técnica é identificar as tendências em seu estágio inicial de desenvolvimento e operar os ativos conforme essa tendência (MURPHY, 1999).

Para a análise técnica, de maneira geral, uma tendência é, simplesmente, a direção dos picos e vales que compõem o movimento do mercado (EDWARDS & MAGEE, 2001). Dessa maneira, é possível identificar três tipos de tendência, que são apresentadas na Figura 5.



Figura 5 - Gráfico com tendências de alta, baixa e lateral

Conforme pode-se notar, tendências dependem da identificação de padrões dos picos e vales que compõem o gráfico de preços. Na linguagem técnica, os picos e vales recebem denominações específicas: resistência e suporte, respectivamente.

Suportes são níveis de preço em que o interesse de compra supera o interesse de venda e, por isso, dificultam a queda dos preços a partir daquele ponto (OSLER, 2000). Resistências agem da maneira inversa, de modo que consistem em níveis de preço em que o interesse de venda supera o de compra e, por isso, dificultam o aumento dos preços.

A partir da identificação de sucessivos suportes e resistências em um gráfico, é possível conectar os pontos para formar os canais de tendências, que indicam a faixa de preço esperada do ativo considerando a relação das forças compradoras e vendedoras mapeadas (MURPHY, 1999).

O rompimento das linhas de suporte e resistência dos canais pode indicar uma tendência mais forte de venda (no caso do rompimento da linha de suporte) ou de compra (no caso do rompimento da linha de resistência).

Padrões Gráficos

Padrões gráficos são figuras ou formações que aparecem nos gráficos de preços que podem ser classificadas em diferentes categorias e possuem

potencial preditivo (BULKOWSKI, 2000). Esses padrões gráficos, podem ser de dois tipos: reversão e continuação.

Como sugerido na própria denominação dos padrões gráficos, padrões de reversão indicam uma inversão na tendência e padrões de continuação indicam uma manutenção da tendência. Como pré-requisito para a identificação desses padrões, é necessário que haja uma tendência identificada previamente.

Existem dois grandes padrões de inversão, que são exibidos na Figura 6: cabeça e ombros e vales duplos.



Figura 6 - Exemplo de padrões: W e Ombro-cabeça-ombro

Os padrões serão explicados para tendências de alta, mas é importante salientar que esses padrões também podem ser identificado em tendências de baixa e, no caso, as orientações e os termos são invertidos.

O padrão cabeça e ombros é, provavelmente, o padrão mais conhecido e mais confiável de todos os padrões de reversão (MURPHY, 1999). Esse padrão é formado por pontos de resistência cada vez mais altos que, após atingirem um ponto máximo, começam a mostrar uma tendência de queda. Nesse formato, a maior resistência é denominada de “cabeça”, e a resistência que sucede e precede a “cabeça” são denominadas de “ombros”. Nesse padrão, também é possível identificar a “linha de pescoço”, que é formada pela linha de suporte abaixo da cabeça. Quando o padrão é revertido, essa linha geralmente representa uma linha de resistência que dificulta a retomada do movimento de alta.

O padrão de topos duplos consiste em dois pontos de resistência consecutivos que são sucedidos de um movimento de baixa de preços. Em movimentos de alta, esse padrão geralmente é conhecido como “M” e, em movimentos de baixa, como “W”.

Padrões de continuação geralmente indicam que os movimentos laterais do mercado não são nada além de uma pausa da tendência anterior e que próximo movimento tenderá a seguir essa tendência. Outra diferença em comparação com os padrões de reversão é a duração, pois padrões de continuação tendem a durar menos tempo e representam padrões de curto-médio prazo.

Existem dois grandes tipos de padrões de continuação: triângulos e canais. Triângulos consistem em duas linhas de tendência de suporte e resistência não paralelas enquanto os canais tendem a apresentar linhas paralelas.

De acordo com a inclinação das linhas, os triângulos podem ser ascendentes, descendentes ou simétricos. A Figura 7 apresenta um triângulo ascendente com rompimento na resistência.



Figura 7 - Exemplo de triângulo ascendente

De maneira semelhante, os canais podem apresentar tendências de alta ou de baixa de acordo com sua inclinação.

Indicadores Técnicos

Além dos padrões gráficos identificados a partir das movimentações de preços, a análise técnica também oferece indicadores técnicos para direcionar as ações de investidores e traders.

Entre os três principais indicadores técnicos, pode-se citar: médias móveis, osciladores e volume.

Médias móveis consistem em médias de preço de um ativo em um determinado período. As médias móveis podem ser de dois tipos: Médias Móveis Simples e Médias Móveis Exponenciais.

Uma Média Móvel Simples (SMA_n) consiste na média de preços do ativo nos últimos n dias (ADRIAN, 2015). Dessa maneira, ela é calculada conforme a fórmula:

$$SMA_n = \frac{1}{n} \sum_{n=1}^N P_n \quad (2.1)$$

Onde:

- P_n é o preço do ativo no dia do cálculo da média;
- “ n ” é a quantidade de termos (dias) envolvidos no cálculo da média; e
- SMA_n é a média calculada a partir dos preços dos últimos “ n ” dias.

Uma Média Móvel Exponencial (EMA_n) consiste na média ponderada de preços do ativo nos últimos n dias, dando mais relevância para os preços mais atuais e amortizando variações abruptas que tenham ocorrido anteriormente (MURPHY, 1999). O cálculo da EMA_n é:

$$EMA_n = (P_n - EMA_{n-1}) \frac{2}{n+1} + SMA_{n-1} \quad (2.2)$$

As médias móveis, além da própria importância como indicador técnico, são os componentes centrais das Bandas de Bollinger, que são resultado da adição de uma linha de tendência com uma distância equivalente a dois desvios padrão acima e abaixo da média móvel. A distância de dois desvios padrões garante que as Bandas de Bollinger incluam 95% da variação dos preços do

ativo e, ao tocar as bandas, entende-se que existe um movimento extremo de compra e/ou venda.

Além das médias móveis, outro indicador importante são os osciladores. Existem três osciladores importantes para a análise técnica: Momentum, Taxa de Mudança e Índice de Força Relativa.

Os osciladores são utilizados quando as ações apresentam movimentações laterais e dão indícios das forças de compra e de venda – ou momentum (MURPHY, 1999). O momentum mede a velocidade da mudança de preços e indica a força das variações. Dessa maneira, em termos percentuais, o Momentum é medido conforme a fórmula abaixo.

$$M_n = \frac{P - P_n}{P} \quad (2.3)$$

Onde:

- M_n é o momentum de “n” dias da ação;
- P é o preço atual da ação; e
- P_n é o preço de fechamento da ação a “n” dias atrás.

Dessa maneira, quando plotado em um gráfico, o momentum indica se a variação de preços está ficando mais forte ou mais fraca ao longo do tempo e, dessa maneira, representa a força de uma tendência (MURPHY, 1999).

Um segundo oscilador importante é a Taxa de Mudança ou ROC (do inglês, *Rate of Change*). Esse oscilador mede a variação percentual entre o preço atual e o preço em um determinado dia. O cálculo da Taxa de Mudança é apresentado de diversas maneiras na literatura, mas, de maneira, geral, utiliza-se o formato abaixo (EDWARDS & MAGEE, 2001).

$$ROC_n = \frac{P}{P_n} - 1 \quad (2.4)$$

Onde:

- ROC_n é a Taxa de Mudança de “n” dias da ação;
- P é o preço atual da ação; e

- P_n é o preço de fechamento da ação a “n” dias atrás.

O terceiro oscilador que é vastamente utilizado na análise técnica é o Índice de Força Relativa ou RSI (do inglês, *Relative Strength Index*). Esse indicador foi desenvolvido com o intuito de suavizar distorções de preço e oferecer uma faixa fixa de valores (WILDER, 1978). Dessa maneira, o RSI indica um valor entre a faixa de 0 a 100 e é calculado conforme a fórmula abaixo.

$$RSI_n = 100 - \frac{100}{1 + RS_n} \quad (2.5)$$

$$RS_n = \frac{U_n}{D_n} \quad (2.6)$$

Onde:

- RSI_n é o Índice de Força Relativa de “n” dias;
- RS_n é a força relativa de “n” dias;
- U_n é a média de valores (contados na moeda de análise) ganhos durante os dias de alta nos últimos “n” dias; e
- D_n é a média de valores (contado na moeda de análise) perdidos durante os dias de baixa nos últimos “n” dias.

É importante salientar que, para todos os indicadores, o termo “dias” está sendo empregado considerando-se uma análise de um gráfico diário de preços. No caso de análises que utilizam gráficos semanais, mensais ou anuais, esse termo deve ser devidamente substituído para o equivalente.

2.3. Análise Fundamentalista

Nesta seção, serão apresentados os principais fatores relacionados a outra grande escola de investimento: a análise fundamentalista. O tópico se divide em três partes:

- Introdução à Análise Fundamentalista, onde se apresenta o conceito da análise fundamentalista e as principais diferenças entre esta e a análise técnica;
- Críticas à Análise Fundamentalista, onde são apresentadas as principais críticas dessa escola de pensamento; e

- Indicadores Fundamentalistas, onde são apresentados os principais indicadores dessa escola de pensamento e seu método de cálculo.

Introdução à Análise Fundamentalista

De maneira geral, enquanto a análise técnica se baseia no estudo dos gráficos, a análise fundamentalista se concentra no estudo das empresas (FONTANILLS & GENTILE, 2001). Esse estudo utiliza, como fonte principal, o estudo de demonstrativos financeiros para determinar o valor intrínseco da empresa, considerando seu potencial de ganhos futuros (FAERBER, 2007).

Dessa maneira, pode-se dizer que a análise fundamentalista possui 3 objetivos principais:

- Determinar o valor intrínseco da empresa a partir do estudo histórico dos demonstrativos financeiros da companhia e da projeção de resultados futuros;
- Comparar o valor intrínseco da empresa com o valor atribuído a ela pelo mercado; e
- Comparar as diferentes precificações do mercado para diferentes ações de um mesmo setor.

Dessa maneira, é possível identificar empresas que estão supervalorizadas ou subvalorizadas e basear a estratégia de investimento nessa diferença.

Críticas à Análise Fundamentalista

As escolas de investimento não constituem uma ciência exata e o nível de acerto das decisões de cada uma depende tanto da capacidade de processamento das informações disponíveis (LEUDERS, 2017), como do horizonte de tempo disponível (GRAHAM & ZWEIG, 2003).

A principal crítica à análise fundamentalista é a Hipótese de Mercados Eficientes – EMH (do inglês, *Efficient Market Hypothesis*), que defende que investidores individuais não são capazes de obter retornos acima do mercado a partir do estudo de informações disponíveis (FAMA, 1970). A Hipótese de

Mercados Eficientes argumenta que a análise de informações e níveis históricos, base das análises fundamentalistas e técnicas, respectivamente, não leva a retornos acima do mercado.

Entretanto, como alguns eventos podem possuir um efeito prolongado, alguns autores defendem que o mercado, em algumas situações, pode não estar em equilíbrio (SHOSTAK, 1997) e que, portanto, o estudo do histórico pode oferecer benefícios aos investidores. Nessa linha, é possível notar que modelos simples, baseados apenas em índices tradicionais de análise fundamentalista, são capazes de diferenciar portfólios vitoriosos de perdedores (HAUGEN, 2007) e que, no longo prazo, o mercado tende a equilibrar as diferenças fundamentalistas dos ativos (GRAHAM & ZWEIG, 2003).

Metodologias Fundamentalistas: Introdução

Como mencionado anteriormente, a análise fundamentalista busca determinar se uma ação está subvalorizada ou supervalorizada com base no estudo dos resultados históricos. A análise fundamentalista se baseia em duas metodologias principais:

- Comparação do preço e do valor intrínseco de uma empresa;
- Comparação de Indicadores Fundamentalistas.

A seguir, são apresentadas as diferentes metodologias e as etapas para cálculo de cada uma.

Metodologias Fundamentalistas: Comparação de Preço e Valor

Para entender a primeira metodologia fundamentalista, é necessário, primeiramente, entender a diferença entre o preço o valor de uma empresa.

O preço de uma empresa é oferecido pelo mercado, a partir das relações entre oferta e procura das ações de uma empresa (LUEDERS, 2017). O cálculo do preço de uma empresa é dado pela fórmula a seguir.

$$P = \sum_{n=1}^N P_n N_n \quad (2.7)$$

Onde:

- P é o preço da empresa;

- P_n é o preço de um ticker da empresa (por exemplo, a ação ordinária de uma empresa);
- N_n é a quantidade de ações existentes com aquele ticker; e
- “n” é a quantidade de diferentes ações existentes no mercado relacionada àquela empresa.

Diferentemente do preço, o valor de uma empresa é uma estimativa obtida a partir da análise dos resultados históricos dela e das suas perspectivas futuras de lucratividade, que podem ser justificados a partir de fatos e informações históricas e disponíveis sobre a empresa, setor e mercado (GRAHAM & DODD, 1951).

Mesmo que dois indivíduos tenham as mesmas informações como base para o cálculo do valor de uma empresa, é possível, e provável, que eles atribuam diferentes valores para ela. Isso ocorre pois, para determinar o valor de uma empresa, é necessário definir premissas e cenários que embasarão as perspectivas de crescimento futuro da empresa em estudo e essas premissas dependem da visão pessoal, baseada em fatos, a respeito do futuro (DAMODARAN, 2010).

Apesar das estimativas poderem apresentar divergências consideráveis nos valores, as metodologias de determinação do valor de empresas costumam ser semelhantes. Entre as metodologias utilizadas para avaliar empresas, pode-se destacar duas principais:

- Modelo de Desconto de Dividendos – DDM (do inglês, *Dividend Discount Model*); e
- Desconto de Fluxos de Caixa – DCF (do inglês, *Discounted Cash Flow*).

O Modelo de Desconto de Dividendos considera que o valor de uma ação é determinado pelo valor dos dividendos futuros relacionados àquela ação, descontados a valor presente (FARRELL, 1985). Dessa maneira, considerando a perpetuidade do pagamento de dividendos por uma determinada ação, o valor dela pode ser calculado pela fórmula:

$$V = \frac{D}{k} \quad (2.8)$$

Onde:

- V é o valor da ação;
- D são os dividendos por ação;
- k é a taxa de desconto.

Caso assumamos que os dividendos pagos por ação aumentam a uma determinada taxa constante (g), a fórmula deve ser ajustada:

$$V = \frac{D}{k-g} \quad (2.9)$$

Dessa maneira, a partir do valor atribuído a cada ação a partir da projeção dos dividendos por ação, é possível obter o valor da empresa multiplicando-se o valor por ação pela quantidade de ações.

A metodologia de Fluxo de Caixa Descontado é uma das principais metodologias utilizadas atualmente (FERNÁNDEZ, 2007) e busca determinar o valor de uma empresa a partir da projeção dos fluxos de caixa gerados por ela no futuro que são descontados a valor presente por uma taxa proporcional ao risco da empresa (KOLLER, 2015). A fórmula abaixo é o ponto de partida para calcular o valor de uma empresa utilizando o DCF.

$$EV = \frac{CF_1}{1+k} + \frac{CF_2}{(1+k)^2} + \frac{CF_3}{(1+k)^3} + \dots + \frac{CF_n+VR}{(1+k)^n} \quad (2.10)$$

$$VR = \frac{CF_n(1+g)}{k-g} \quad (2.11)$$

Onde:

- EV é o valor presente da empresa;
- CF_n é o fluxo de caixa gerado no período “n”;
- k é a taxa de desconto utilizada;
- VR é o valor residual da empresa; e
- g é a taxa de crescimento perpétuo assumida para a empresa a partir do ano “n”.

Muitos autores utilizam o Fluxo de Caixa Livre (FCF) como Fluxo de Caixa e consideram o custo de capital ponderado (WACC) como taxa de desconto.

Fluxo de Caixa Livre (FCF) consiste no fluxo de caixa operacional, ou seja, o fluxo de caixa gerado pelas operações de uma empresa (ROSS, WESTERFIELD & JORDAN, 2000). As fórmulas necessárias para o cálculo do FCF são:

$$FCF = EBIT(1 - t) + D\&A - \Delta Capex - \Delta WC \quad (2.12)$$

$$EBIT = Receitas - COGS - SG\&A - D\&A \quad (2.13)$$

$$WC = AC - PC \quad (2.14)$$

Onde:

- FCF: Fluxo de Caixa Livre;
- EBIT: Lucro Operacional (do inglês, *Earnings Before Interest and Taxes*);
- D&A: Depreciação e Amortização;
- Capex: Investimentos necessários para manter a operação (do inglês, *Capital Expenditures*);
- WC: Capital de Giro (do inglês, *Working Capital*);
- COGS: Custos do Produto Vendido (do inglês, *Costs of Goods Sold*);
- SG&A: Despesas gerais (do inglês, *Sales, General & Administration*);
- AC: Ativos correntes;
- PC: Passivos correntes;
- t: taxa de impostos; e
- Δ representa a variação anual de um determinado valor.

Por outro lado, o custo de capital ponderado – WACC (do inglês, *Weighted Average Cost of Capital*) consiste na taxa utilizada para descontar os fluxos de caixa futuros a valor presente (FERNÁNDEZ, 2011) e é calculado a partir da fórmula:

$$WACC = K_e \frac{E}{D+E} + K_d \frac{D}{D+E} (1 - t) \quad (2.15)$$

$$K_e = R_f + (R_m - R_f)\beta \quad (2.16)$$

Onde:

- WACC: custo de capital ponderado;
- K_e : custo de capital para o acionista;
- K_d : custo de capital para o credor;
- E : Capital próprio da empresa (*Equity*);
- D : Dívidas da empresa (*Debt*);
- t : taxa de imposto;
- R_f : taxa de retorno livre de risco;
- R_m : taxa de retorno esperada do mercado; e
- β : correlação entre a variação da ação e do índice de mercado.

É interessante salientar que o custo de capital ponderado geralmente é diferente para duas empresas distintas, pois depende da estrutura de capital adotada pela empresa, correlação de suas ações ou de ações de empresas semelhantes com o índice de mercado, tipo de dívida contraída e período de contração da dívida, entre outros fatores.

Após descontar os fluxos de caixa projetados pelo custo de capital adequado, é possível chegar ao valor total da empresa – EV (do inglês, *Enterprise Value*). O próximo passo é chegar no valor esperado por ação, que pode ser feito a partir da fórmula:

$$V = EV - (ND + PS + MI) \quad (2.17)$$

$$SV = \frac{V}{NoS} \quad (2.18)$$

Onde:

- V: Valor (patrimonial) da empresa;
- ND: Dívida líquida (do inglês, *Net Debt*);
- PS: Ações preferenciais (do inglês, *Preffered Stock*);
- MI: Interesse Minoritário (do inglês, *Minority Interest*);
- SV: Valor por ação (do inglês, *Stock Value*); e
- NoS: Número de ações (do inglês, *Number of Stocks*).

Percebe-se que tanto o Modelo de Desconto de Dividendos, como o método de Fluxo de Caixa Descontado permitem calcular o valor apropriado das ações de uma determinada empresa. Esse valor calculado pode ser comparado com o valor atual das ações da empresa em análise para determinar o grau de atratividade do investimento (percentual de diferença entre o valor real e o valor calculado) e o tipo de operação (compra ou venda).

Metodologias Fundamentalistas: Comparação de Indicadores

Conforme visto anteriormente, a metodologia de comparação entre o preço e valor de uma empresa pode demandar um esforço relativamente grande por parte dos investidores (ROSENBAUM, 2013). Uma metodologia alternativa é a comparação de indicadores, que exige um esforço menor e é menos susceptível a distorções de percepções por parte dos investidores, dado que não leva em consideração a projeção de resultados (LUEDERS, 2017).

De maneira geral, a comparação de indicadores se baseia em duas etapas:

- Cálculo de indicadores que relacionam o preço e elementos financeiros de performance de uma empresa; e
- Comparação do valor do indicador com o histórico da empresa e com seus pares no momento da análise.

Nesse tópico, são apresentados os 4 principais indicadores fundamentalistas utilizados por referências na área de investimentos, o significado de cada um e as metodologias de cálculo.

O indicador fundamentalista mais utilizado para definir a atratividade de um investimento é a relação entre o preço da empresa e seu lucro líquido (GRAHAM & DODD, 1951). O cálculo desse indicador é executado conforme a fórmula:

$$P/E = \frac{PPS}{EPS} \quad (2.19)$$

Onde:

- P/E: Indicador Preço-Lucro (do inglês, *Price to Earnings*);
- PPS: Valor por Ação (do inglês, *Price per Share*); e
- EPS: Lucro por Ação (do inglês, *Earnings per Share*).

O indicador Preço-Lucro (P/E) representa o preço pago pelo acionista para cada unidade de lucro da empresa. Dessa maneira, quanto maior o valor do indicador, mais caro o acionista está pagando para cada unidade de lucro da empresa.

A identificação de empresas com um P/E baixo (quando comparado à média do setor ou histórico da empresa) geralmente revela um potencial de valorização futuro, típico de empresas de valor (GRAHAM & DODD, 1951).

A eficácia desse indicador no longo prazo é percebida por investidores e teóricos relevantes na indústria de investimento, ao passo que uma carteira composta por ações com indicadores P/E em níveis baixos teria superado uma carteira composta por ações com indicadores P/E em níveis altos em mais de 7% ao ano nos últimos 80 anos (DAMODARAN, 2010).

O segundo indicador mais relevante para os acionistas é o Preço-Valor Patrimonial por Ação (P/VPA). Esse indicador é obtido a partir da divisão entre o preço por ação e o Valor Patrimonial por ação de uma determinada empresa, conforme exibido nas fórmulas:

$$P/VPA = \frac{PPS}{EQPS} \quad (2.20)$$

$$EQPS = \frac{A-P}{NoS} \quad (2.21)$$

Onde:

- P/VPA: Indicador Preço-Valor Patrimonial por Ação;
- PPS: Valor por Ação (do inglês, *Price per Share*);
- EQPS: Valor Patrimonial por Ação (do inglês, *Equity per Share*);
- A: Ativo total da empresa;
- P: Passivo total da empresa; e
- NoS: Número de ações (do inglês, *Number of Stocks*).

O indicador Preço-Valor Patrimonial por Ação representa quanto os investidores estão dispostos a pagar pelo Valor Patrimonial de uma empresa. Esse indicador, quando menor que 1.0, pode indicar empresas precificadas com desconto, pois, no caso de situações de fechamento de capital, recompra de ações ou venda da empresa, é provável que haja uma valorização significativa no preço das ações, de modo que estas tendem a, no mínimo, igualar o Valor Patrimonial da empresa (GRAHAM & ZWEIG, 2003).

Benjamin Graham, uma grande referência em teorias de investimento e mentor de Warren Buffet, um dos maiores investidores do século, defendia que, para uma empresa ser atrativa, um dos requisitos para a compra das ações da empresa era que seu P/VPA x P/E fosse menor que 22,5 (GRAHAM & ZWEIG, 2003).

Entretanto, é importante salientar que essa condição é relativa ao setor da empresa e ao contexto macroeconômico (FABOZZI, FOCARDI & KOLM, 2010), dado que em situações de mercado em alta, esses múltiplos podem apresentar valores bastante superiores aos apresentados por Graham.

Outro indicador importante que direciona a análise de empresas no mercado de ações é a relação Dividendo-Preço – DY (do inglês, *Dividend Yield*). Esse indicador é bastante relevante no contexto econômico brasileiro devido ao sistema tributário nacional, que, por meio do artigo 10 da Lei no. 9.249/1995,

concede isenção de imposto de renda incidente sobre a distribuição de lucros e dividendos aos sócios e acionistas de pessoas jurídicas (TELLES, 2016).

Conforme o próprio nome sugere, o cálculo desse indicador é obtido a partir da divisão do valor distribuído aos investidores na forma de dividendos pelo preço da ação.

$$DY = \frac{DPS}{PPS} \quad (2.22)$$

Onde:

- DY: Indicador Preço-Dividendo;
- DPS: Dividendo por Ação (do inglês, *Dividend per Share*);
- PPS: Preço por Ação (do inglês, *Price per Share*).

É importante salientar a necessidade de se considerar o histórico de dividendos da empresa antes de executar o cálculo do DY.

Ao se considerar o histórico de distribuição de dividendos por empresas, o DY, além de representar quantos por cento do preço da ação uma determinada distribui para seus acionistas na forma de dividendos, pode mostrar, de forma indireta, em quanto tempo o investidor pode obter um retorno de 100% o valor das ações na forma de dividendos.

Dessa maneira, se uma determinada empresa possui uma ação que hoje vale R\$ 15,00 e distribui anualmente dividendos de R\$ 0,50, é possível notar que ela possui um DY igual a 3,3% e é possível obter um retorno de 100% sobre o valor da ação em dividendos em 30 anos.

Por fim, o quarto indicador fundamentalista mais utilizado por investidores é a relação Preço-EBITDA. Esse indicador ganhou relevância desde o final do século XX, considerando a supervalorização das ações de empresas de tecnologia que dariam início à Bolha da Internet (MORRIS & ALAM, 2008). Nessa época, boa parte das empresas mais atraentes do mercado apresentavam prejuízos financeiros e, como forma de encontrar outro indicador nos demonstrativos financeiros relacionado à atividade operacional, definiu-se o

Preço-EBITDA. O cálculo do indicador Preço-EBITDA ocorre a partir das fórmulas:

$$P/EBITDA = \frac{PPS}{EBITDAPS} \quad (2.23)$$

$$EBITDA = Receita - COGS - SG\&A \quad (2.24)$$

Onde:

- P/EBITDA: Indicador Preço-EBITDA;
- PPS: Preço por Ação (do inglês, *Price per Share*);
- EBITDAPS: EBITDA por Ação (do inglês, *EBITDA per Share*);
- COGS: Custos do Produto Vendido (do inglês, *Costs of Goods Sold*); e
- SG&A: Despesas Gerais (do inglês, *Selling, General & Administration*).

O indicador Preço-EBITDA, como visto, é uma alternativa a ser utilizada quando as empresas analisadas não possuem resultado financeiro positivo e, conseqüentemente, apresentam um DY e um P/E com significância reduzida (LUEDERS, 2017). Nesse sentido, o EBITDA oferece uma boa opção para o resultado operacional da empresa, sinalizando, com as devidas ressalvas, uma proximidade com o fluxo de caixa operacional.

Conforme os próprios elementos utilizados para o cálculo do EBITDA, é possível desconsiderar altos valores de depreciação e amortização, pagamento de juros e impostos, que contribuem para um possível resultado financeiro negativo e, conseqüentemente, impactam a visibilidade da empresa.

3. Aprendizagem de Máquina

Neste capítulo, são apresentados os principais fatores relacionados ao surgimento, desenvolvimento e estado atual da inteligência artificial e aprendizagem de máquinas.

3.1. O Surgimento da Inteligência Artificial

Apesar da grande relevância do assunto atualmente, os conceitos que embasam o ramo de Inteligência Artificial já são imaginados e descritos pelo ser humano há bastante tempo (BUCHANAN, 2006). Pode-se verificar que na literatura de ficção científica, como no Mágico de Oz (BAUM, 1939) e Frankenstein (SHELLEY, 1818), a presença de personagens artificiais inteligentes pode ser vista desde o século XIX.

Além das contribuições provenientes da fertilidade da imaginação humana, o desenvolvimento da Inteligência Artificial foi influenciado por outros dois fatores: desenvolvimento de teorias de probabilidade e econômicas, e capacidade tecnológica (RUSSEL & NORVIG, 2010).

Entre as principais teorias probabilísticas e econômicas que contribuíram para a formação das bases da Inteligência Artificial, pode-se citar três:

- Teorema de Bayes;
- Teoria da Decisão; e
- Teoria dos Jogos.

O Teorema de Bayes contribuiu para a Inteligência Artificial por meio da sua aplicação em modelos estatísticos. Bayes propôs regras que atualizam probabilidades com base em novas evidências e, com isso, formou a base para uma linha de abordagens modernas para o raciocínio incerto de sistemas artificiais.

A Teoria da Decisão e a Teoria dos Jogos (VON NEUMANN & MORGENTHAU, 1944) contribuem para a formação de uma estrutura formal e completa para a tomada de decisão em situações de incerteza e propõem que agentes racionais devem agir de forma causal.

Por fim, o principal fator que permitiu o avanço nos estudos sobre o impacto da Inteligência Artificial no contexto mundial atual foi a evolução tecnológica (RUSSEL & NORVIG, 2010). A evolução tecnológica compreende dois aspectos: o primeiro aspecto é o aumento da geração dos dados, o segundo é o aumento da capacidade de processamento das máquinas.

Em relação ao primeiro aspecto, o aumento da conectividade dos seres humanos e explosão da geração de dados foram fundamentais para o desenvolvimento de aplicações de Inteligência Artificial, ao passo que permitiram que as máquinas obtivessem treinamentos sobre grandes bases de dados e, conseqüentemente, se preparassem para uma quantidade maior de situações.

Em relação ao segundo aspecto, o aumento da capacidade de processamento das máquinas contribuiu para que computadores fossem capazes de operar grandes volumes de dados e realizar operações cada vez mais complexas em um curto período.

3.2. Inteligência Artificial

Conceitualmente, a Inteligência Artificial já foi definida por diversos autores. Abaixo, uma definição mais recente e outra mais antiga são apresentadas.

- Inteligência Artificial é a automação de atividades que associamos ao pensamento humano, como tomadas de decisão, resolução de problemas, aprendizado, entre outros (BELLMAN, 1978);
- Inteligência Artificial é o estudo do design de agentes inteligentes (POOLE et. al., 1998).

De maneira geral, pode-se dizer que a Inteligência Artificial é toda e qualquer técnica que permite que máquinas imitem o comportamento humano. Entretanto, algumas técnicas específicas possuem uma denominação própria, como a Aprendizagem de Máquina – ML (do inglês, *Machine Learning*) e a Aprendizagem Profunda – DP (do inglês, *Deep Learning*).

Conceitualmente, a Aprendizagem de Máquina é um ramo da Inteligência Artificial que permite que máquinas aprendam a resolver problemas a partir de algoritmos baseados em modelos estatísticos aplicados a uma grande

base de dados (BURKOV, 2019). Dessa maneira, a Aprendizagem de Máquinas são técnicas de Inteligência Artificial que permitem que computadores aprendam sem serem explicitamente programados para isso.

Já a Aprendizagem Profunda consiste em uma classe de técnicas de ML que explora diversos níveis de processamento de informações não lineares para extração, transformação, análise de padrões e classificação de maneira supervisionada ou não supervisionada (DENG & YU, 2013).

3.3. Aprendizagem de Máquina

Conforme visto anteriormente, a Aprendizagem de Máquina engloba técnicas de Inteligência Artificial que permitem que computadores aprendam sem serem explicitamente programados para isso. O processo de aprendizagem é verificado quando há uma melhoria de performance em uma determinada atividade conforme a máquina (ou agente) aumenta a quantidade de observações feitas (RUSSEL & NORVIG, 2013).

O processo de aprendizagem de máquina depende da relação entre as entradas fornecidas ao agente, as saídas fornecidas pelo agente e o feedback oferecido ao agente após o processo. Nesse sentido é necessário definir os conceitos de entrada, saída, feedback e agente:

- Entrada: consiste nos dados (números, imagens, sons, textos, entre outros) oferecidos ao agente;
- Saída: consiste na saída fornecida pelo agente a partir dos dados oferecidos;
- Feedback: consiste no retorno dado ao agente a partir da saída fornecida por ele;
- Agente: sistema lógico capaz de perceber um ambiente a partir de um sensor e agir nesse ambiente a partir de um atuador.

De acordo com a existência e o tipo de feedback dado para o agente que executa as atividades, o processo de aprendizagem pode ser de três tipos:

- Supervisionada;

- Não supervisionada; ou
- Por Reforço.

Na aprendizagem não supervisionada, o agente aprende os padrões das entradas mesmo sem receber feedbacks (RUSSEL & NORVIG, 2013). O tipo mais comum de aprendizagem não supervisionada é o agrupamento (ou *clustering*) de dados a partir das informações recebidas.

Na aprendizagem por reforço, o agente aprende a partir de uma série de punições ou recompensas, que estimulam saídas certas e desestimulam as saídas erradas. Nesse sentido, o objetivo da aprendizagem por reforço é que o agente entenda um comportamento e maximize sua recompensa (BURKOV, 2019).

Na aprendizagem supervisionada, o agente observa pares de entrada e saída oferecidos e busca elaborar uma função que relaciona a saída com as entradas (BURKOV, 2019).

3.4. Algoritmos Fundamentais

Dentre os diversos algoritmos que se enquadram em técnicas de aprendizagem de máquinas, pode-se destacar cinco de maior relevância encontrados na literatura:

- Regressão linear;
- Regressão logística (BISHOP, 2006);
- Árvore de decisão (BREIMAN et al., 1984; QUINLAN, 1979; QUINLAN, 1993);
- Máquinas de Vetores de Suporte (KECMAN, 2005); e
- k-vizinhos (YANG, 1999).

Das técnicas destacadas, a regressão linear será destacada nesta seção, considerando sua alta relevância para a formulação do algoritmo de Q-Learning e Double Q-Learning utilizado para este projeto.

Algoritmos de regressão linear buscam aprender modelos que consistem na combinação linear dos fatores contidos nas entradas (BURKOV, 2019).

Existem três componentes importantes que fazem parte do algoritmo de regressão linear: entradas, modelo de regressão linear e a função de perda.

As entradas, para a regressão linear, devem ser oferecidas na forma $\{(x_i; y_i)\}_{i=1}^N$, onde N é o tamanho da amostra oferecida, x_i é um vetor de entradas de dimensão D e y_i é o valor desejado.

Nesse caso, o objetivo do algoritmo é construir um modelo de regressão linear $f_{w,b}(x)$, que consiste em uma combinação linear dos fatores contidos na amostra x , que minimize a função de perda. O modelo de regressão linear e a função de perda são:

$$f_{x,b}(x) = wx + b \quad (3.1)$$

$$f_l = \frac{1}{N} \sum_{i=1}^N (f_{w,b}(x_i) - y_i)^2 \quad (3.2)$$

Onde:

- w é um vetor de parâmetros de dimensão D ;
- b é um número real; e
- f_l é a função de perda.

Dessa maneira, busca-se fazer com que o modelo possa oferecer uma linha que seja capaz de projetar um rótulo y_{novo} a partir de uma nova entrada não rotulada x_{novo} . Um exemplo de modelo linear para uma determinada amostra é apresentado na Figura 8.

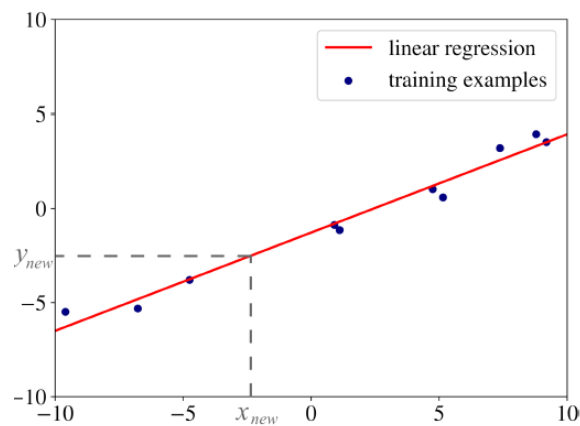


Figura 8 - Exemplo de modelo linear

3.5. Aplicações

Conforme o interesse no ramo de Inteligência Artificial aumentou, pesquisadores e empresas perceberam que as técnicas propostas podiam ser aplicadas em diversas áreas e com diversos propósitos. Abaixo, são apresentados quatro setores onde a inteligência artificial é aplicada atualmente:

- Comunicação (TAGLIAMONTE & DENIS, 2008; ARON, 2011; HILL, FORD & FARRERAS, 2015);
- E-commerce (LEE, LIU & LU, 2002; BALABANOVIC & SHOHAM, 1997; JOACHIMS, FREITAG & MITCHELL, 1997; MOONEY & ROY, 2000; DUHIGG, 2012);
- Saúde (JIANG et al., 2017; SOMASHEKHAR et al., 2017; ESTEVA et al., 2017; BOUTON et al., 2016; FARINA et al., 2017; DILSIZIAN & SIEGEL, 2014); e
- Finanças.

No campo das finanças, destaque do presente projeto, é possível verificar a utilização de ferramentas de Aprendizagem de Máquinas em diversos campos, conforme apresentado abaixo:

- Desenvolvimento de sistemas de categorização e filtragem para detecção de fraudes (QUAH & SRIGANESH, 2008);
- Desenvolvimento de sistemas inteligentes para gerenciamento de riscos (AZIZ & DOWLING, 2018); e
- Desenvolvimento de algoritmos de investimento (CHABOUD et al., 2009; ALVES, CAARLS & LIMA, 2018; BERTOLUZZO & CORAZZA, 2012).

3.6. Q-Learning

A utilização de algoritmos de Q-Learning em trading foi alvo de estudo de diversos autores (DANTAS, 2018; CORAZZA & BERTOLUZZO, 2014; DU, ZHAI & LV, 2016) e, neste projeto, pretende-se dar continuidade a esse tema. O Q-Learning (WATKINS, 1989) consiste em um algoritmo de aprendizagem de máquina por reforço utilizado para solucionar, de maneira otimizada, os

Processos de Decisão de Markov – MDP (do inglês, *Markov Decision Processes*). O MDP consiste em 4 partes: um espaço de estados S, uma função A que oferece as ações possíveis para cada estado, uma função de transição T e uma recompensa R (ROSS, 1983).

A principal propriedade do MDP é que o estado atual do sistema é a única informação necessária para que o agente decida a ação que deve ser tomada, ou seja, o conhecimento do estado atual torna dispensável o conhecimento sobre o histórico do sistema (WATKINS, 1989).

Considerando um ambiente como um espaço controlado de Markov, onde o agente atua como controlador, o algoritmo de Q-Learning passa por uma sequência de episódios distintos que contribuem para a formação da experiência do agente (WATKINS & DAYAN, 1992). Na n ésima experiência, o agente:

- Observa o estado atual - x_n ;
- Seleciona e executa a ação definida - a_n ;
- Observa o próximo estado ao qual a ação o levou - y_n ;
- Recebe a recompensa - r_n ; e
- Ajusta o valor de Q com base no fator de aprendizado α_n .

O termo Q representa a função ação-valor do algoritmo e consiste na recompensa descontada esperada para executar a ação a no estado x e seguir a política a partir de então (WATKINS & DAYAN, 1992). Matematicamente, o Q-Learning é expressado pela equação abaixo:

$$Q_t(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t(s_t, a_t) \left[r_t + \gamma \max_a Q_t(s_{t+1}, a^*) - Q_t(s_t, a_t) \right] \quad (3.3)$$

Onde:

- s_t : estado do agente;
- a_t : ação do agente;
- a^* : ação que maximiza o valor da função $Q_t(s_{t+1}, a)$;
- Q_t : valor-ação atual;

- α_t : taxa de aprendizagem;
- r_t : recompensa; e
- γ : fator de desconto.

O Q-Learning é considerado um algoritmo *off-policy*. Isso decorre do fato do valor do $Q_t(s_t, a_t)$ ser atualizado com base no valor $Q_t(s_{t+1}, a^*)$, onde a^* é a ação com o maior valor de Q no estado s_{t+1} . Ou seja, independente da política utilizada para escolher a ação, Q é atualizado com base em uma política gananciosa.

Em relação ao fator de desconto γ , este tem o objetivo de oferecer uma taxa de desconto para trazer as recompensas da sequência de ações futuras a valor presente (PITIS, 2019). Dessa maneira, o tempo que levaria para obter a recompensa se torna um fator que impacta a decisão.

Em relação à superestimação dos valores de ações, apesar da popularidade do Q-Learning em algoritmos de aprendizagem por reforço, o fator $\max_a Q_t(s_{t+1}, a^*)$, que determina o valor do próximo estado, pode superestimar o valor de algumas ações. Dessa maneira, a performance do algoritmo pode ser prejudicada devido a um otimismo na tomada da decisão (HASSELT, 2010). Nesse sentido, uma possível abordagem seria a utilização do Double Q-Learning.

3.7. Double Q-Learning

Ao contrário do Q-Learning, que utiliza apenas uma função Q para estimar o valor do próximo estado, o Double Q-Learning utiliza duas funções para estimar esse valor e, dessa maneira, minimiza o fator otimista na tomada de decisão (HASSELT, 2010).

O Double Q-Learning armazena duas funções Q_a e Q_b , onde cada uma delas é atualizada com o valor da outra para o próximo estado. A partir dessas duas funções, é possível que ambas aprendam a partir de um conjunto de experiências diferentes, construam modelos diferentes e reduzam o otimismo na tomada de decisão pelo algoritmo.

Introduzindo-se as mudanças propostas no Double Q-Learning, a função de atualização dos valores das ações é dada conforme abaixo.

$$Q_a(s_t, a_t) = Q_a(s_t, a_t) + \alpha_t(s_t, a_t)[r_t + \gamma Q_b(s_{t+1}, a^*) - Q_a(s_t, a_t)] \quad (3.4)$$

$$Q_b(s_t, a_t) = Q_b(s_t, a_t) + \alpha_t(s_t, a_t)[r_t + \gamma Q_a(s_{t+1}, b^*) - Q_b(s_t, a_t)] \quad (3.5)$$

Pode-se observar a presença do termo $Q_b(s_{t+1}, a^*)$ como substituto do termo $\max_a Q_t(s_{t+1}, a_t)$, onde a^* é a ação que máxima o valor no estado s_{t+1} de acordo com a função Q_a .

4. Materiais e Métodos

Neste capítulo, são apresentados os materiais utilizados para a execução do presente estudo, bem como os métodos utilizados para realização das análises e o desenvolvimento do algoritmo de aprendizagem de máquinas.

4.1. Materiais

Para o desenvolvimento do presente estudo, foram utilizados 3 tipos de materiais:

- Materiais de revisão bibliográfica: artigos científicos, estudos e livros;
- Bases de dados: séries históricas das ações do mercado financeiro brasileiro e dados sobre os resultados das empresas;
- Algoritmos: estruturação, funcionamento e componentes de algoritmos de trading que utilizam técnicas de aprendizagem de máquinas.

4.2. Métodos

Os métodos compreendem o conjunto de procedimentos utilizados na concretização do estudo (CÁCERES, GANDARA & PUGLICI, 2011). Pode-se segmentar os métodos utilizados no desenvolvimento do presente projeto em 3 partes:

- Análise e coleta de dados;
- Desenvolvimento do algoritmo; e
- Definição dos parâmetros de avaliação da performance.

Análise e Coleta de Dados

O escopo do estudo envolveu a análise de dados de empresas de capital aberto na B3 (Bolsa de Valores do Brasil). Nesse sentido, foram realizados os seguintes procedimentos:

- Definição da extensão temporal analisada;
- Definição dos critérios analisados;

- Definição da quantidade de empresas analisadas; e
- Definição das empresas analisadas.

Na primeira etapa, buscou-se definir um período de análise que englobasse ciclos de alta e ciclos de baixa da bolsa brasileira. Dessa maneira, definiu-se que o período entre os anos de 2003 e 2018 (16 anos) seria uma boa amostra para execução das análises, por três motivos:

- Utiliza 71% do período de existência do índice da bolsa de valores utilizado atualmente; e
- Engloba pelo menos dois grandes ciclos de alta e de baixa da bolsa brasileira; e
- Engloba ciclos de alta e de baixa com velocidades diferentes.

Ao todo, foram definidos 9 miniciclos de treinamento (com duração de 5 a 8 meses), que são apresentados na Tabela 10.

ID	Início	Fim	Duração (meses)	Variação	Tipo
1	01/07/2003	01/12/2003	5,00	64%	Alta
2	01/06/2004	01/02/2005	8,00	34%	Alta
3	01/07/2005	01/02/2006	7,00	49%	Alta
4	01/09/2006	01/03/2007	6,00	26%	Alta
5	02/04/2007	01/10/2007	6,00	33%	Alta
6	02/06/2008	03/11/2008	5,00	-50%	Baixa
7	01/04/2009	01/12/2009	8,00	46%	Alta
8	01/04/2011	01/09/2011	5,00	-24%	Baixa
9	01/04/2010	01/10/2010	6,00	0%	Lateral

Tabela 10 - Ciclos de treinamento utilizados

Após a definição do escopo temporal, foram definidos os critérios utilizados nas análises. O estudo tem o objetivo de avaliar empresas em duas dimensões (volatilidade e atratividade fundamentalista) utilizando-se um algoritmo de aprendizagem de máquinas que se baseia em critérios de análise técnica para tomada de decisão.

Nesse sentido, é necessário incluir no estudo três tipos de critérios:

- Volume e volatilidade;
- Indicadores fundamentalistas; e
- Indicadores técnicos.

O volume é necessário para indicar a liquidez das ações em estudo, de modo que quanto maior o volume de uma ação, maior a sua liquidez. Além do volume, o indicador de volatilidade (variabilidade do preço) foi utilizado para avaliar se ações com uma variabilidade maior podem oferecer oportunidades maiores de retorno a partir da utilização do algoritmo.

Os indicadores fundamentalistas, teoricamente, devem indicar a atratividade de uma determinada ação, de modo que se uma ação está subvalorizada ela é atrativa para o investidor e vice-versa. Nesse sentido, foram escolhidos os dois indicadores fundamentalistas mais utilizados por investidores: P/E e P/VPA.

Os indicadores técnicos, no presente estudo, têm a função de guiar a decisão de investimento do algoritmo. Nesse sentido, foram escolhidos os indicadores técnicos mais utilizados por investidores: médias móveis simples e exponenciais.

Dessa maneira, foram executados quatro procedimentos para determinação da quantidade de empresas analisadas.

1. Mapear as empresas que foram negociadas na bolsa brasileira de 2003 a 2018;
2. Definição dos indicadores fundamentalistas, volatilidade e volume para as ações e distribuição no agrupamento;
3. Avaliação da disponibilidade dos indicadores e dados para as empresas; e
4. Definição da quantidade de empresas analisadas.

O mapeamento das empresas negociadas na bolsa levou em consideração a data de abertura de capital de cada uma e, a partir da quantidade

de empresas presentes na bolsa de valores brasileira a em 2003 e a quantidade de empresas presentes na bolsa de valores em 2018, definiu-se a Figura 9.

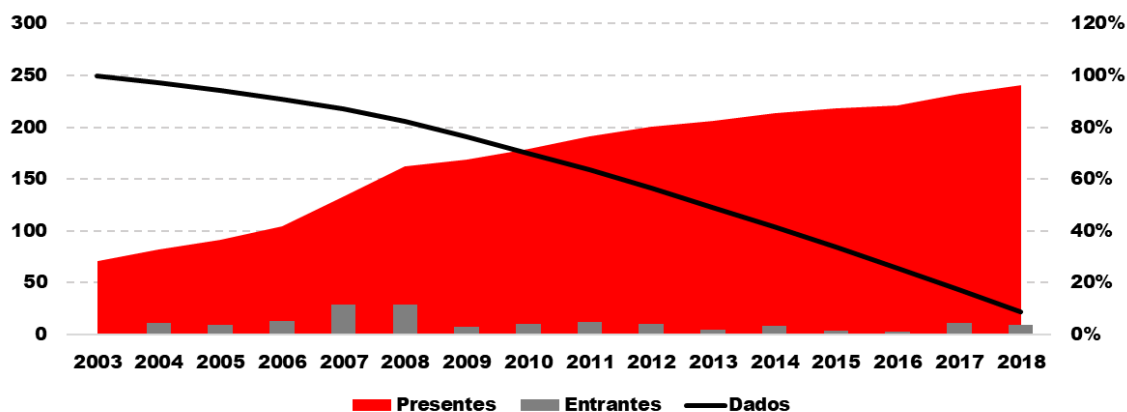


Figura 9 - Quantidade de empresas listadas na bolsa e disponibilidade de dados de 2003 a 2018

Percebe-se que, apesar de a quantidade de empresas presentes na bolsa em 2018 superar a quantidade presente em 2003 em mais de 200%, a quantidade de dados disponíveis em 2018 é consideravelmente menor que a de 2003 em termos percentuais.

Dessa maneira, a primeira seleção manteve as empresas que abriram seu capital até o ano de 2003, pois este ano possui a maior disponibilidade de dados para treinamento, que é um fator essencial no treinamento do algoritmo.

A segunda seleção avaliou a disponibilidade de dados em duas dimensões:

- Disponibilidade de indicadores fundamentalistas de 2003-2018; e
- Disponibilidade de preços de 2003-2018;

Considerando-se uma disponibilidade de dados superior a 98% em relação ao total de dias utilizados no estudo, para possibilitar um treinamento adequado do algoritmo, a população do estudo foi reduzida de 71 empresas para 31.

O último dado considerado na definição da população do estudo foi a liquidez das ações. A liquidez foi avaliada considerando-se o volume de ações negociadas anualmente. Para esta seleção, o volume mínimo negociado anualmente foi definido como 1 milhão de ações por dia.

Na Tabela 11, são apresentadas as etapas, critérios e empresas que atendiam aos critérios definidos em cada etapa.

Etapa	Critério	Quantidade de empresas que atendem ao critério
Data de abertura de capital	≤ 2003	71
Disponibilidade de P/E e P/VPA	$\geq 98\%$	31
Volume negociado	> 1 milhão por dia	17

Tabela 11 - Critérios utilizados para escolha das ações envolvidas no estudo

A partir destes critérios, foram selecionadas 17 empresas para compor a população e a amostra do presente estudo.

Desenvolvimento do Algoritmo

O algoritmo utilizado no projeto foi implementado na linguagem Python, utilizando os valores diários para avaliar cenários e tomar decisões. Além do agente e do ambiente, existem 3 elementos que compõem um sistema de aprendizagem de máquina por reforço: política, recompensa e função de valor (SUTTON & BARTO, 2015). Considerando que a política de um agente relaciona o estado do agente e a ação realizada por ele naquele estado (SUTTON & BARTO, 2015), o desenvolvimento do algoritmo de operação com aprendizagem de máquinas foi estruturado em seis etapas:

- Definição das variáveis de estado;
- Definição das ações realizadas pelo agente;
- Definição das recompensas;
- Definição da quantidade de iterações de treinamento;
- Definição da função de aproximação; e
- Escolha dos métodos de aprendizagem de máquinas comparados.

No presente estudo, o estado do agente inteligente é determinado pela combinação de dois tipos de inputs:

- Retorno diário; e
- Indicadores técnicos.

Considerando-se que o algoritmo de aprendizagem de máquinas utilizado aproxima a função a partir de uma regressão linear, a utilização de uma ampla gama de estados pode não contribuir para uma melhora no desempenho. Dessa maneira, buscou-se utilizar uma quantidade adequada que tivesse impacto no desempenho do algoritmo.

Para seleção dos estados utilizados, foram realizados testes de performance no algoritmo para cada estado isoladamente e, a partir disso, foram mapeados os que contribuíam para uma maior performance e aceleravam o tempo de aprendizado.

Após a definição dos estados, foram avaliadas as possíveis ações que o agente poderia tomar. O escopo de ações possíveis foi definido com base nos princípios básicos de negociação de ações no mercado financeiro. Nesse sentido, é possível considerar a existência de três possíveis ações que podem ser tomadas pelo algoritmo:

- Compra;
- Venda; e
- Repouso.

Quando o agente determina a compra de uma ação, ele se coloca em um estado comprado e a valorização da ação resulta em um retorno para o agente. No caso da venda de uma ação, ocorre o oposto, a desvalorização de uma ação resulta em um retorno negativo para o agente.

O repouso pode ocorrer em estados de compra, venda ou nenhum dos dois. No caso de o repouso ocorrer em nenhum estado, o agente se coloca como mero observador das ações, sem ter prejuízo nem lucro com a variação dos preços.

Após a definição das ações que o agente pode tomar no ambiente, é importante definir as recompensas. Essa recompensa tem, como objetivo, incentivar a adoção de ações que valorizem o portfólio e desestimular ações que causem desvalorizações. Nesse sentido, três caminhos são possíveis:

- Oferecer uma recompensa universal para o acerto e uma punição de mesma intensidade para o erro;
- Oferecer uma recompensa universal para o acerto e uma punição mais intensa para o erro, ou vice-versa;
- Oferecer recompensas e punições proporcionais à variação do portfólio.

O oferecimento de uma recompensa universal não é interessante para o presente projeto, dado que isso não estimularia o agente a atuar de maneira a desenvolver modelos que maximizassem a performance, pois no caso de um aumento de portfólio de 1% ou 10% ele seria recompensado da mesma maneira.

Por outro lado, a utilização de recompensas e punições fixas, com uma punição mais intensa que uma recompensa pode levar o agente a determinar estratégias mais seguras e que levem a um retorno menor, mas reduzam o risco de perdas.

Por fim, a utilização de recompensas e punições proporcionais à variação do portfólio é interessante, pois oferece recompensas e punições diretamente relacionadas ao objetivo do algoritmo. Desse modo, buscou-se avaliar a performance do algoritmo nesse último caso, conforme a fórmula:

$$R_n = \frac{P_n - P_{n-1}}{P_{n-1}} \quad (4.1)$$

Onde:

- R_n é o valor da recompensa no dia “n”;
- P_n é o valor do portfólio do agente no dia “n”; e
- P_{n-1} é o valor do portfólio do agente no dia “n-1”.

Outro fator importante que foi considerado na realização do presente projeto foi a quantidade de iterações de treinamento, que consistem na

quantidade de vezes que o algoritmo é submetido ao mesmo período de treinamento (MOHRI, ROSTAMIZADEH & TALWALKAR, 2012). Para definição da quantidade de iterações de treinamento, foram mapeados os desvios padrões, em termos percentuais, para uma determinada ação, de acordo com o número de épocas de treinamento. Os resultados são apresentados na Figura 10.

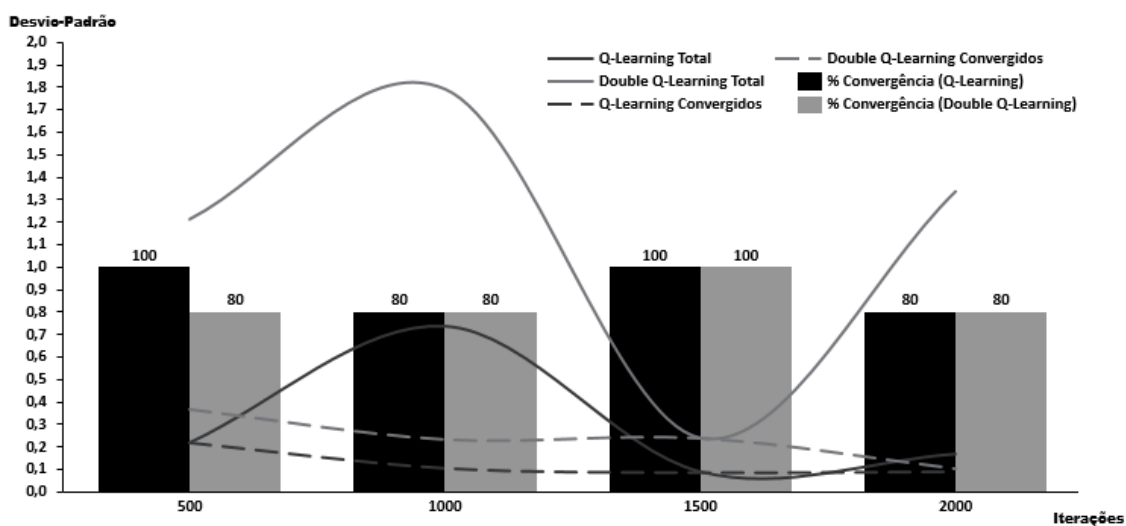


Figura 10 – Percentual de convergência, desvio padrão geral e desvio padrão para os casos dos algoritmos de Q-Learning e Double Q-Learning que convergiram, em diferentes épocas, operando PETR4, quando treinados nos ciclos de alta apresentados na Seção [Análise e Coleta de Dados](#)

A Figura 10 apresenta três tipos de informações: o desvio padrão dos algoritmos considerando-se todos os 10 ciclos de teste, o desvio padrão dos algoritmos considerando-se apenas os testes que convergiram e o percentual de convergência dos algoritmos, para diferentes quantidades de iterações.

A convergência do algoritmo foi medida a partir do armazenamento das decisões tomadas pelo algoritmo durante o período de treinamento. Se o algoritmo, a partir de um determinado número de iterações, repetisse as mesmas decisões em um determinado estado, considerou-se que ele convergiu.

Conforme apresentado na Figura 10, é possível mapear as aproximações de erro que cada quantidade de iterações pode trazer ao projeto. Considerando-se uma relação entre o tempo disponível para execução do projeto e o erro apresentado de acordo com as iterações, optou-se por utilizar 1.000 iterações para o treinamento do algoritmo nas análises realizadas.

Por fim, optou-se por aproximar a função Q utilizando o algoritmo de LMS (do inglês, *Least Mean Square*), que minimiza o erro quadrático médio entre a função criada e os exemplos utilizados a partir da atualização dos pesos adotados no filtro (HAYKIN & WIDROW, 2003).

Avaliação da Performance do Algoritmo

Por se tratar de um algoritmo cujo objetivo é a valorização da carteira, definiu-se como métrica de performance o retorno obtido pelas operações no período de teste, desconsiderando-se o custo de corretagem.

Dessa maneira, para avaliar o desempenho do agente, foram definidos três benchmarks no mesmo período:

- Rentabilidade da taxa básica de juros anual Selic;
- Retorno do Ibovespa no período; e
- Retorno da ação em análise obtida no período de acordo com uma estratégia simples de *Buy and Hold*.

5. Análises e Discussão

A partir dos métodos apresentados no capítulo 4, desenvolveram-se três análises, que comparam a performance dos algoritmos de Q-Learning e Double Q-Learning considerando-se quatro diferenças:

- Quantidade e tipo de estados definidos;
- Tipos de ciclos de treinamento;
- Nível de volatilidade; e
- Indicadores fundamentalistas.

5.1. Comparação da performance em relação a quantidade e tipos de variáveis de estados utilizados

A comparação da performance dos algoritmos de Q-Learning e Double Q-Learning em relação a quantidade e tipo de estados utilizados foi realizada para a ação PETR4.

Essa ação foi submetida a dois tipos de simulação:

- Simulação conjunta: considera 5 variáveis de estado; e
- Simulação isolada: considera apenas 2 variáveis de estado.

As simulações conjuntas foram realizadas considerando-se as seguintes variáveis de estado: retorno diário, bandas de Bollinger, convergência de médias móveis, oscilador estocástico e volume. As simulações isoladas, por sua vez, foram realizadas considerando-se a combinação do retorno diário e cada variável de estado isoladamente.

A performance medida para cada cenário, considerando-se 1.000 iterações de treinamento, é apresentada na Figura 11.

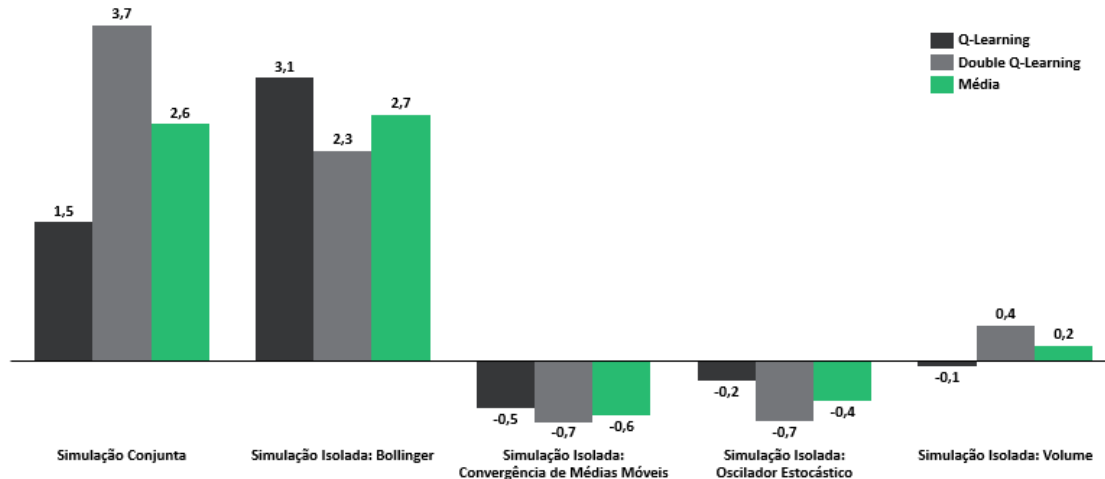


Figura 11 – Performance média de 10 resultados dos algoritmos de Q-Learning e Double Q-Learning operando PETR4, quando submetidos a 1.000 iterações de treinamento com diferentes variáveis de estado

Percebe-se que, apesar de a simulação conjunta apresentar uma variedade maior de estados e, conseqüentemente, maior informação para a tomada de decisão do agente, a simulação isolada com as Bandas de Bollinger foi a que apresentou a maior média de performance para os algoritmos de Q-Learning e Double Q-Learning.

É interessante notar que o algoritmo de Double Q-Learning superou a performance do algoritmo de Q-Learning em mais de 100% no caso da simulação conjunta, o que indica que esse tipo de algoritmo pode ser mais adequado nos casos onde o agente possui uma maior quantidade de informações disponíveis.

Além disso, percebe-se pelos resultados obtidos que os outros indicadores técnicos utilizados na simulação isolada obtiveram retornos negativos ou bastante inferiores aos obtidos na simulação com as Bandas de Bollinger.

5.2. Comparação da performance para diferentes ciclos de treinamento

A comparação da performance para diferentes ciclos de treinamento também foi executada para o ativo PETR4. Nesse caso, foram comparados os retornos obtidos pelo algoritmo quando submetido aos ciclos curtos (de 5-8

meses) de alta, baixa e deslocamento lateral indicados no item [Análise e Coleta de Dados](#).

A performance do algoritmo para cada tipo de ciclo curto, considerando-se 1.000 épocas de treinamento, é apresentada na Figura 12.

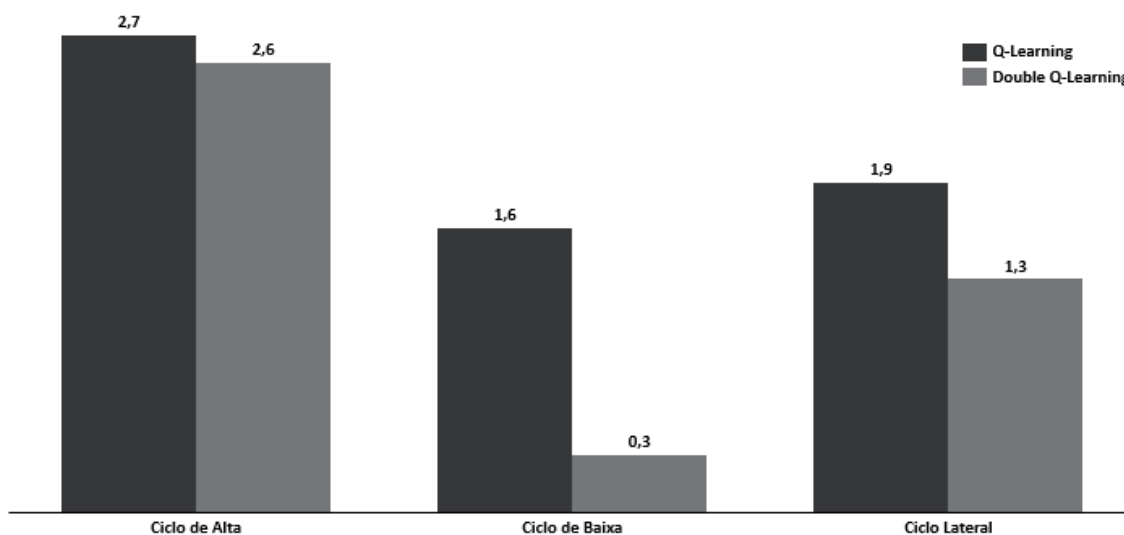


Figura 12 - Performance média de 10 resultados dos algoritmos de Q-Learning e Double Q-Learning operando PETR4, quando submetidos a 1.000 iterações de treinamento em diferentes ciclos da bolsa

A partir da figura acima, é possível tirar duas conclusões:

- Treinamentos em ciclos de alta geralmente possuem um retorno médio maior que treinamentos em ciclos de baixa e ciclos laterais; e
- O algoritmo de Q-Learning apresenta um retorno médio maior que o algoritmo de Double Q-Learning em todos os cenários.

É possível levantar algumas hipóteses a respeito das conclusões acima. A primeira hipótese é que, como o período de aplicação do algoritmo envolve um longo ciclo de alta no final (de 2016 a 2018), o otimismo característico do algoritmo de Q-Learning (HASSELT, 2010) pode contribuir para uma melhor performance em comparação com o algoritmo de Double Q-Learning.

Além disso, percebe-se que, como treinamentos em ciclos de alta submetem o algoritmo a um aprendizado mais intensificado em situações de aumento de valor das ações, pode haver uma predisposição do algoritmo a

executar ações de compra, o que favorece a performance no caso em que ele é submetido a um ciclo de aplicação de alta. Dessa maneira, percebe-se que a variedade das ações em que o algoritmo é submetido no ciclo de treinamento contribui para uma performance melhor, mas que a quantidade de eventos repetidos aos quais o agente é exposto também possui um efeito determinante na performance.

5.3. Comparação de performance entre Q-Learning e Double Q-Learning

Considerando os resultados obtidos nas Figuras 11 e 12, optou-se por avaliar a performance dos algoritmos de Q-Learning e Double Q-Learning utilizando-se duas variáveis de estado - retorno diário e Bandas de Bollinger – e treinados nos ciclos de alta, destacados na Seção [Análise e Coleta de Dados](#). A Figura 13 apresenta a média dos 10 resultados obtidos após a implementação dos algoritmos de Q-Learning e Double Q-Learning nas 17 ações escolhidas.

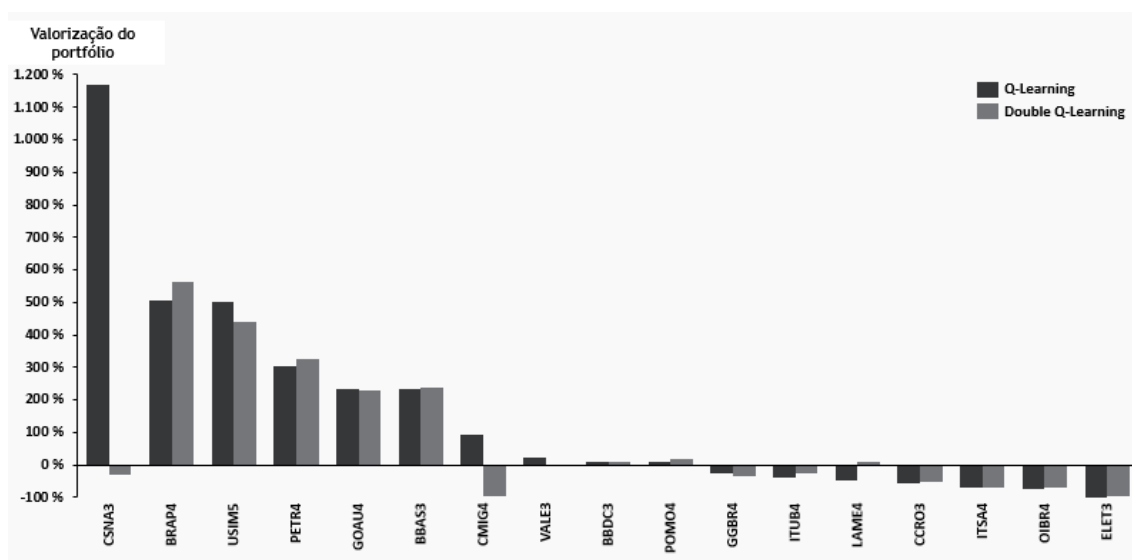


Figura 13 - Performance média dos 10 retornos obtidos pelos algoritmos de Q-Learning e Double Q-Learning para as 17 ações, treinados nos 9 ciclos de alta apresentados na Seção [Análise e Coleta de Dados](#)

A partir dos resultados acima, é possível verificar 4 fatores:

- Excluindo-se as ações CSNA3 e CMIG4, cujo desempenho do algoritmo de Q-Learning é consideravelmente maior que o do Double Q-Learning, a performance dos algoritmos é semelhante;

- Nos casos em que ambos os algoritmos apresentam retorno superior a zero, o retorno apresentado pelo algoritmo de Double Q-Learning tende a superar o retorno apresentado pelo algoritmo de Q-Learning em aproximadamente 70% das vezes;
- Nos casos em que ambos os algoritmos apresentam retorno inferior a zero, o prejuízo apresentado pelo algoritmo de Double Q-Learning é inferior ao prejuízo apresentado pelo algoritmo de Q-Learning em aproximadamente 70% das vezes;
- Considerando todas as ações avaliadas, o algoritmo de Q-Learning apresenta um retorno médio de 1,58, enquanto o algoritmo de Double Q-Learning apresenta um retorno médio de 0,80. Desconsiderando-se os casos onde há maior disparidade (CMIG4 e CSNA3), o retorno médio apresentado pelo algoritmo de Double Q-Learning é de 0,87 contra 0,94 do algoritmo de Q-Learning.

5.4. Comparação da performance para diferentes níveis de volatilidade

Conforme mencionado anteriormente, a volatilidade de um ativo representa uma estatística de dispersão do preço de um determinado ativo. A volatilidade está atrelada, na maioria dos casos, a um maior risco (ENGLE & NG, 1993). Dessa maneira, para avaliar o comportamento dos agentes em diferentes cenários de risco e, conseqüentemente, oportunidades de retorno, avaliou-se a performance obtida para as 17 ações conforme a volatilidade anual máxima de cada uma durante o período de aplicação. A relação entre a volatilidade anual máxima no período analisado e a performance média do algoritmo é apresentada na Figura 14.

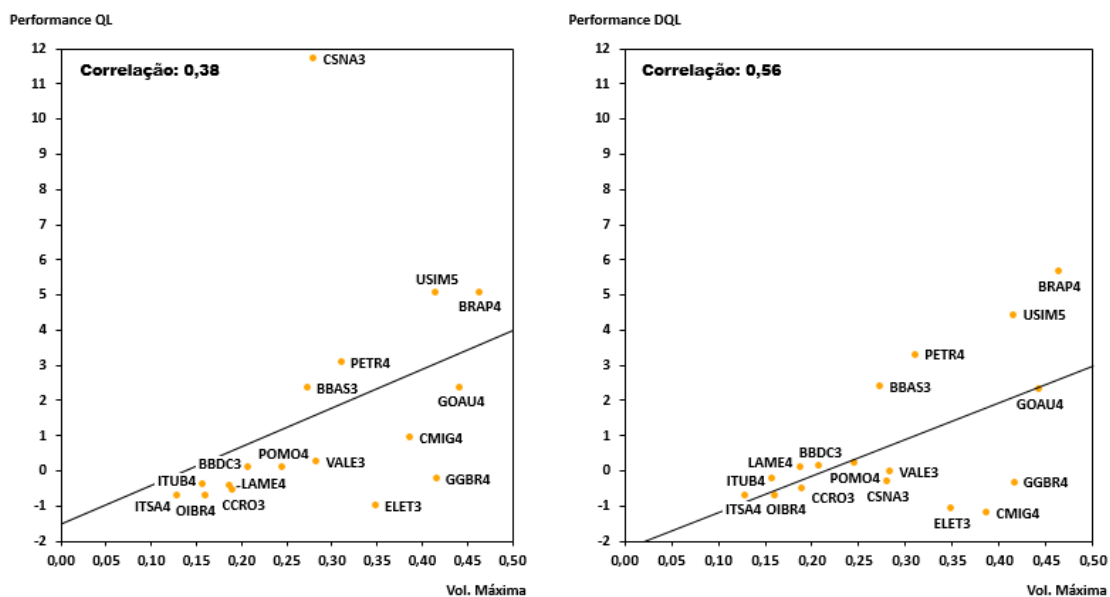


Figura 14 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a volatilidade máxima das ações

A partir da figura, percebe-se que existe, de fato, uma relação positiva entre a volatilidade apresentada pelas ações e o retorno obtido pelo algoritmo. A correlação apresentada entre o retorno apresentado pelo algoritmo de Q-Learning foi de 0,38, enquanto o algoritmo de Double Q-Learning apresentou uma correlação de 0,58 em relação à volatilidade anual máxima apresentada pelas ações analisadas.

5.5. Comparação da performance para diferentes P/E e P/VPA

Além da volatilidade, outro parâmetro de segmentação das ações foi a atratividade fundamentalista, representada pelos indicadores P/E e P/VPA. No caso, o valor dos indicadores varia conforme o preço da ação e os resultados obtidos pela companhia em cada trimestre. Dessa maneira, é interessante observar a variação apresentada por esses indicadores entre o início do período de aplicação e o final, bem como o valor do indicador P/E no início do período.

A Figura 15 mostra a performance obtida pelos dois algoritmos para cada ação conforme a variação do indicador P/E entre o período de aplicação do algoritmo.

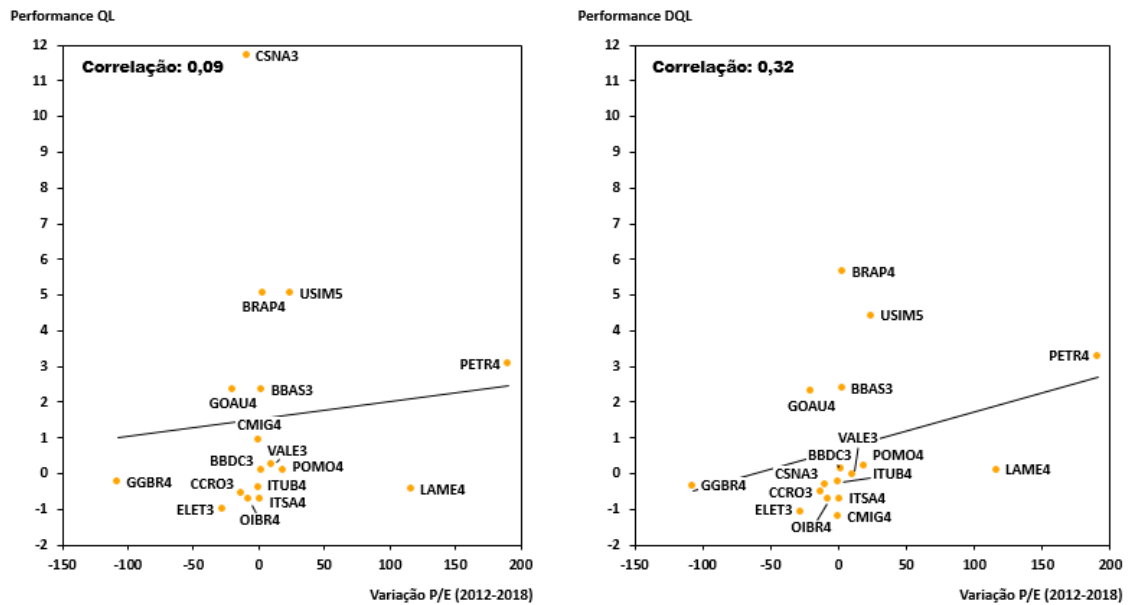


Figura 15 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a variação do indicador P/E das ações

A partir da figura acima, é possível verificar que existe uma correlação mais intensa entre o retorno atingido pelo algoritmo de Double Q-Learning e a variação do indicador P/E das ações, do que a apresentada pelo algoritmo de Q-Learning, que pode ser explicada, principalmente, pela presença do retorno obtido na ação CSNA3. Excluindo-se o retorno obtido nessa ação, a correlação apresentada pelo algoritmo de Q-Learning passa de 0,09 para 0,26.

Dessa maneira, entende-se que, como um aumento do indicador P/E está correlacionado a um aumento do retorno obtido pelo algoritmo quando estes operam uma determinada ação, a aquisição de ativos com indicadores P/E baixos pode levar a retornos mais atrativos para o investidor.

Outro indicador que foi analisado no retorno apresentado pelos algoritmos foi o indicador P/VPA. A Figura 16 mostra a performance obtida pelo algoritmo para cada ação, conforme a variação do indicador P/VPA no período analisado.

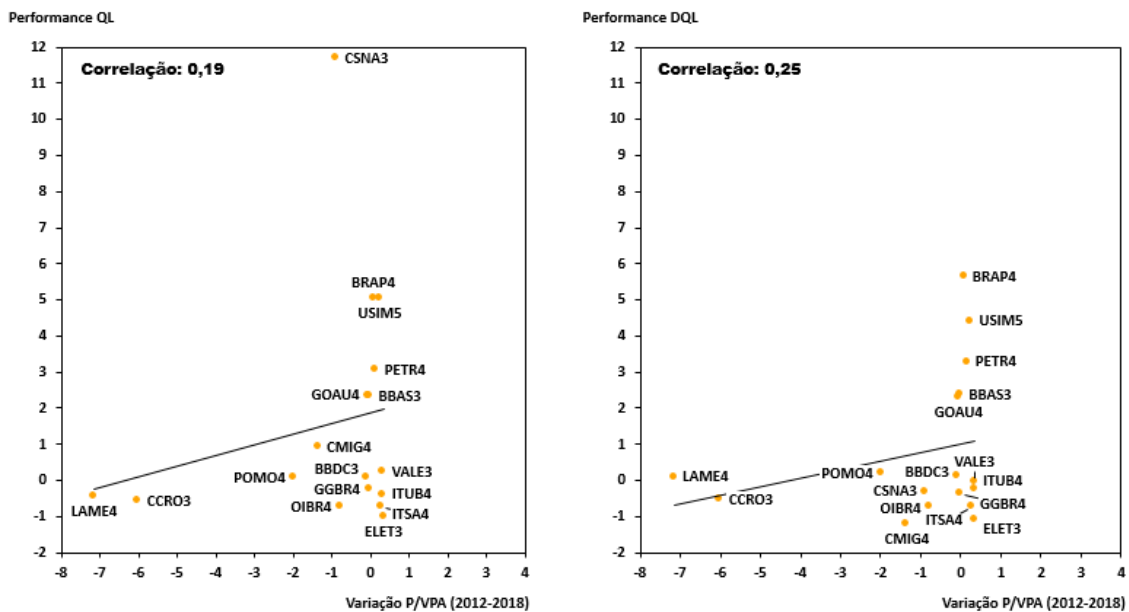


Figura 16 - Relação entre a performance dos algoritmos de Q-Learning e Double Q-Learning e a variação do indicador P/VPA das ações

A partir da figura acima, é possível verificar dois fatores:

- A maior parte das ações analisadas apresentou uma baixa variação no indicador P/VPA no período de aplicação do algoritmo;
- É possível perceber uma correlação entre o retorno apresentado e a variação do indicador P/VPA, mas essa relação é menor que a apresentada pelo indicador P/E.

5.6. Matriz de classificação das ações

Conforme proposição feita por (GRAHAM & DODD, 1951) as ações podem se dividir em dois grupos: ações de crescimento e ações de valor. Nesse sentido, a principal diferença entre esses grupos é a expectativa de crescimento futuro projetado por analistas, que, na maioria das vezes, se reflete em um alto pagamento pelos lucros da empresa no caso de ações de crescimento.

Dessa maneira, tomando como referência que investimentos em ações de valor com preços atrativos tendem a gerar retornos maiores no médio-longo prazo do que investimentos em ações de crescimento (SANTOS & MONTEZANO, 2011), as 19 ações investidas pelo algoritmo foram segmentadas

conforme sua volatilidade e sua atratividade fundamentalista, medida através do indicador P/E.

A inserção da volatilidade como parâmetro de segmentação busca avaliar a hipótese de que ações com maior volatilidade tendem a apresentar maiores variações de preços e, conseqüentemente, maiores chances de retorno acima da média.

A Figura 17 apresenta a relação entre a variação do indicador P/E, o retorno médio obtido pelos algoritmos e a volatilidade para as 17 ações analisadas de 01/01/2012 a 31/12/2018.

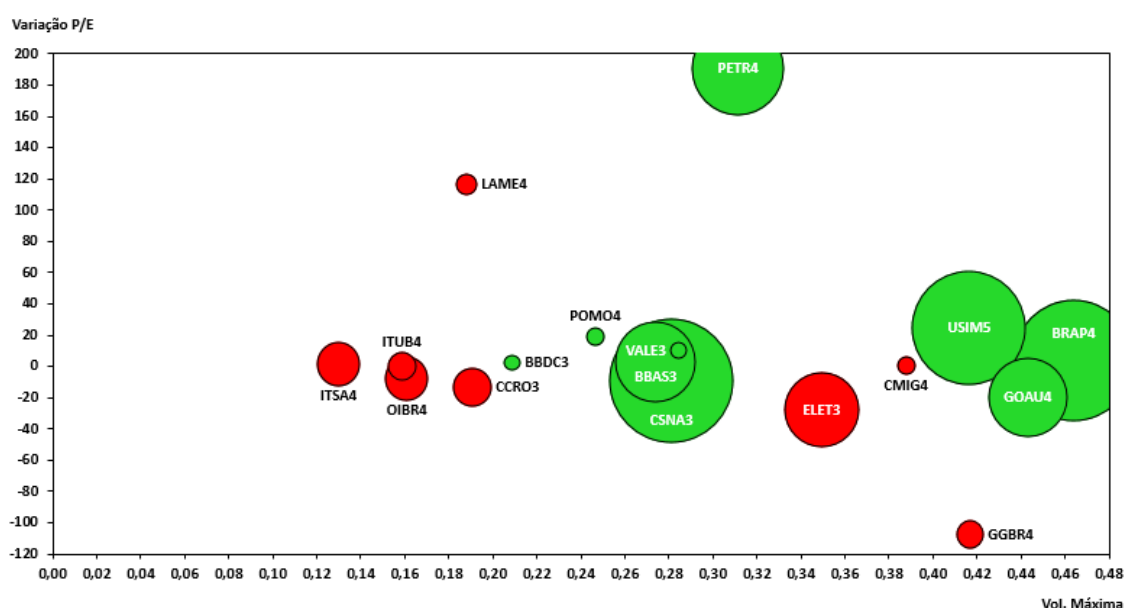


Figura 17 - Retorno, variação do indicador P/E e volatilidade máxima das ações

A partir do gráfico, é possível verificar que apesar da variação do indicador P/E influenciar os resultados do algoritmo, o algoritmo possui uma performance melhor com ações de alta volatilidade (acima de 20%) e com uma variação de P/E entre -30 e 30.

Entretanto, uma dúvida que surge está relacionada ao potencial de seleção de ações com essas características. Dessa maneira, percebe-se que é possível identificar, com um certo grau de precisão, ações mais atraentes para a aplicação dos algoritmos a partir da análise do P/E inicial da ação e do desvio-padrão médio da volatilidade histórica apresentada.

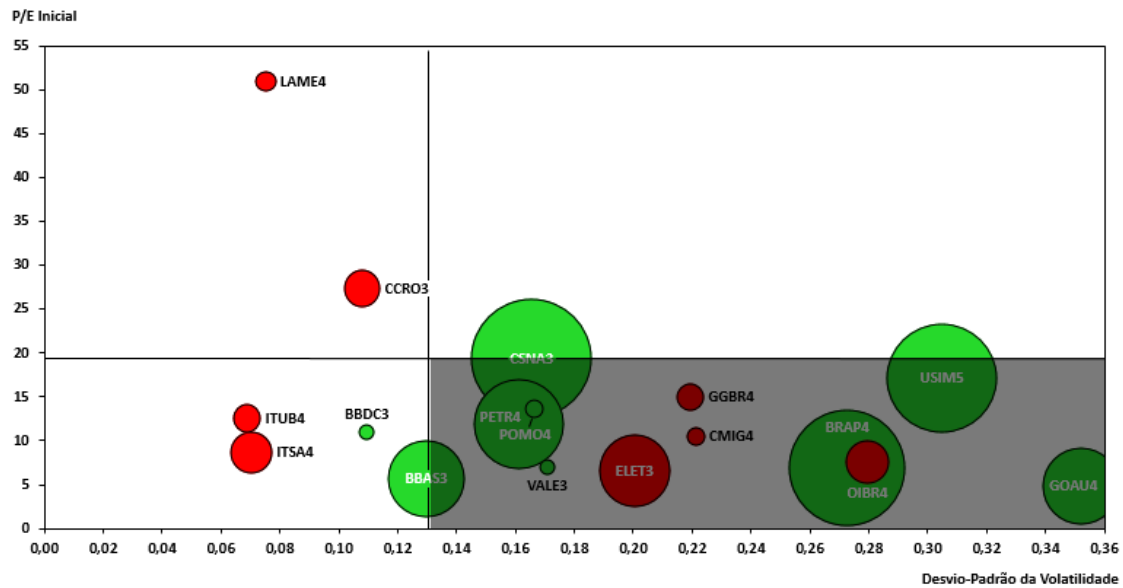


Figura 18 - Relação entre indicador P/E inicial, desvio padrão da volatilidade e retorno apresentado

Percebe-se que o algoritmo apresenta uma chance de retorno maior para ações localizadas no quadrante indicado na Figura 18. Dessa maneira, é interessante avaliar, antes da aplicação do algoritmo, os riscos aos quais o investidor estaria disposto a correr e se estes são contemplados pelas características apresentadas no gráfico.

5.7. Benchmarks de performance

A partir dos resultados obtidos para as diferentes ações em que as simulações foram implementadas, é possível comparar a performance do algoritmo com os benchmarks mencionados no item [Avaliação da Performance do Algoritmo](#). A comparação da performance obtida com a taxa Selic e a variação do Ibovespa é apresentada na Figura 19.

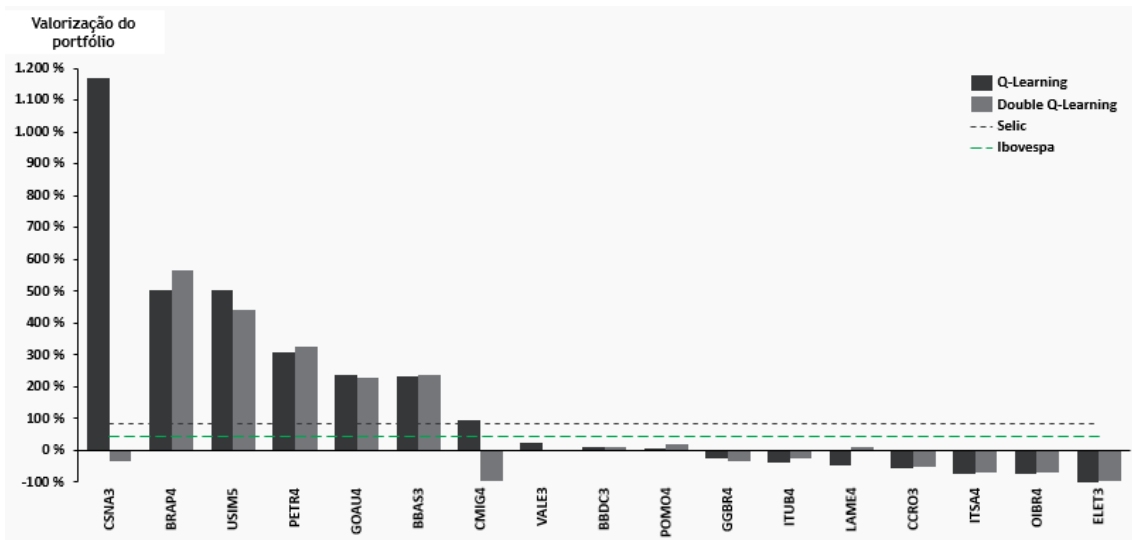


Figura 19 - Comparação do retorno obtido pelos algoritmos de Q-Learning e Double Q-Learning, Ibovespa e Selic

A partir da figura, é possível perceber que em apenas 5 dos 17 casos em estudo o rendimento oferecido pelos algoritmos de Q-Learning e Double Q-Learning superou o retorno obtido pela Selic, que ofereceu um retorno acumulado de 93% aos investidores entre 2012 e 2018.

Outro benchmark adotado foi o retorno apresentado pela ação no período, seguindo uma estratégia de *Buy and Hold*. A comparação da performance do algoritmo nesse caso é apresentada na Figura 20.

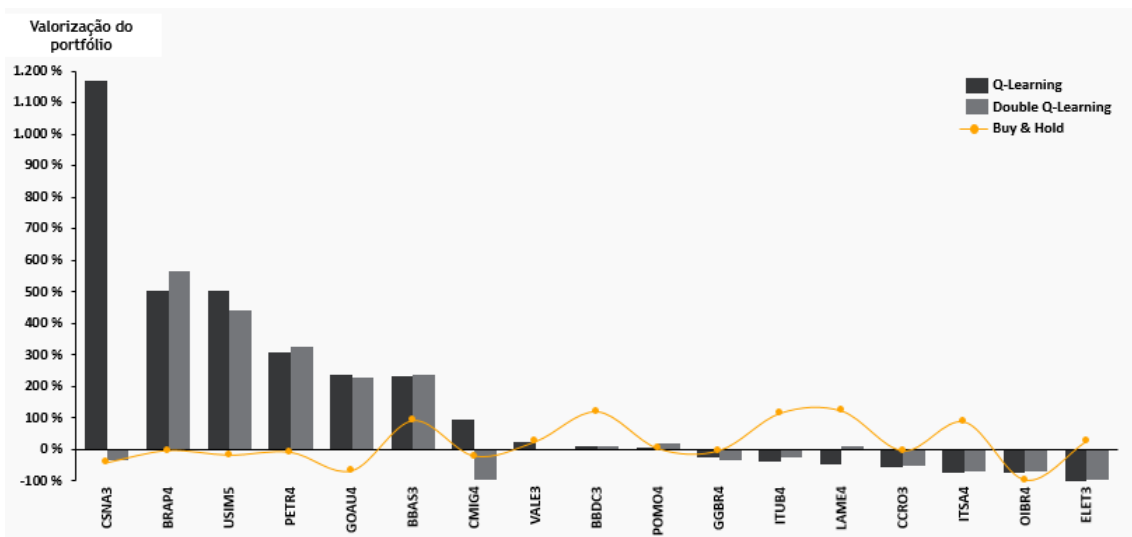


Figura 20 - Comparação do retorno obtido pelos algoritmos de Q-Learning e Double Q-Learning e estratégia de Buy and Hold

A partir da figura, é possível perceber que em 7 dos 17 casos em estudo ambos os algoritmos de Q-Learning e Double Q-Learning apresentaram retornos maiores que os obtidos em estratégias de *Buy and Hold*.

6. Conclusões e Próximos Passos

A partir das análises da performance média entre os algoritmos de Q-Learning e Double Q-Learning nos 17 ativos da bolsa brasileira selecionado, é possível verificar alguns fatores.

Primeiro, percebe-se que ambos algoritmos tendem a apresentar uma performance com retornos expressivos para ativos com maiores níveis de volatilidade.

Segundo, observa-se que ambos apresentam uma performance melhor quando são submetidos a períodos de treinamento onde predominam ciclos de alta. Dessa maneira, quando avaliados sobre um mesmo cenário de aplicação que apresenta um ciclo de baixa prolongado (perda de 40% do valor do Ibovespa em 4 anos) e um ciclo de alta intensa (ganho de 112% do valor do Ibovespa em 2 anos), a performance dos agentes é consideravelmente melhor quando eles são treinados em ciclos de alta do que quando são treinados em ciclos de baixa ou laterais.

Terceiro, percebe-se que os algoritmos de Q-Learning quando aplicados ao ambiente de investimento em ações apresentam uma performance relativamente melhor que os algoritmos de Double Q-Learning em cenários de alta do mercado, o que pode ser um efeito positivo do otimismo relacionado a esse tipo de aprendizado de máquina (HASSELT, 2010).

Por fim, apesar de apresentarem resultados expressivos em 7 das 17 ações analisadas, percebe-se que existem pontos que ainda devem ser investigados a respeito da performance dos algoritmos para avaliar o motivo das perdas causadas em alguns casos.

Como próximos passos, sugere-se a avaliação dos seguintes fatores:

- Utilização de indicadores fundamentalistas como variáveis de estado para os algoritmos;

- Avaliação de performance para diferentes taxas de aprendizagem;
- Avaliação da performance considerando metodologias alternativas ao *backtesting*; e
- Comparação de performance com algoritmos de redes neurais.

7. Referencial Bibliográfico

ADRIAN, Z. The sensitivity of moving average trading rules performance with respect to methodological assumptions. *Procedia Economics and Finance*, Vol. 32, 2015.

ALVES, A., CAARLS, W., LIMA, P. Weightless neural network for high frequency trading. 2018 International Joint Conference on Neural Networks, 2018.

ANDERSON, N. H. Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology*, no. 2, 1965.

ARON, J. Software tricks people into thinking it is human. *New Scientist*, 2011.

AZIZ, S., DOWLING, M. AI and Machine Learning for Risk Management. In *Disrupting Finance: FinTech and Strategy in the 21st Century*, Palgrave, 2019.

BALABANOVIC, M., SOHAM, Y. Content-based collaborative recommendation. *Communications of the ACM*, 1997.

BAUM, L. *The wizard of oz*. McGraw Hill, 1939.

BELLMAN, A. *An introduction to artificial intelligence: can computers think?* Boyd & Fraser, 1978.

BERTOLUZZO, F., CORAZZA, M. Testing different reinforcement learning configurations for financial trading: introduction and applications. *Procedia Economics and Finance*, no. 3, 2012.

BISHOP, C. *Pattern recognition and machine learning*. Cambridge: Springer, 2006.

BOUTON, C. et al. Restoring cortical control of functional movement in a human with quadriplegia. *Nature*, Vol. 533, 2016.

BRAV, A., JIANG, W., KIM, H. The real effects of hedge fund activism: productivity, asset allocation, and labor outcomes. SSRN, 2015.

BREIMAN, L. et al. *Classification and regression trees*. Wadsworth Books, 1984.

BUCHANAN, B. A (very) brief story of artificial intelligence. *AI Magazine*, Vol. 26, no. 4, 2006.

BULKOWSKI, T. *Encyclopedia of chart patterns*. John Wiley Sons, 2000.

BURKOV, A. *The hundred page machine learning book*. 2019.

CÁCERES, A., GÂNDARA, J., PUGLISI, M. Redação científica e a qualidade dos artigos: em busca de maior impacto. *Jornal da Sociedade Brasileira de Fonoaudiologia*, Vol. 23, no. 4, 2011.

CHABOUD, A. et al. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 2014.

COOTNER, P. *The random character of stock market prices*. Cambridge: MIT Press, 1964.

CORAZZA, M.; BERTOLUZZO, F. Q-learning-based financial trading systems with applications. University Ca' Foscari of Venice, Dept. of Economics Working Paper Series, 2014.

DAMASIO, A. *Descartes' error: emotion, reason and the human brain*. Nova York: G.P. Putnam, 1994.

DAMODARAN, A. *Equity risk premium (EPR): determinants, estimation and implications*. Nova York, 2010.

DANTAS, S., SILVA, D. Equity trading at the Brazilian stock market using a Q-learning based system. 2018 7th Brazilian Conference on Intelligent Systems (BRACIS), 2018.

DENG, L., YU, D. *Deep learning: methods and applications*. *Foundations and Trends in Signal Processing*, Vol. 7, no. 3, 2013.

DILSIZIAN, S., SIEGEL, E. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current Cardiology Report*, 2016.

DU, X.; ZHAI, J.; LV, K. Algorithm trading using q-learning and recurrent reinforcement learning. *Journal Positions*, v. 1, 2016.

DUHIGG, C. O poder do hábito: porque fazemos o que fazemos na vida e nos negócios. Rio de Janeiro: Objetiva, 2012.

EDWARDS, R., MAGEE, J. Technical analysis of stock trends. St. Lucie Press, 2001.

ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature, vol. 542, 2017.

FABOZZI, F., FOCARDI, S., KOLM, P. Quantitative equity investing. New Jersey: John Wiley & Sons, 2010.

FAERBER, E. All about stocks. McGraw Hill, 2007.

FAMA, E. Efficient capital markets: a review of theory and empirical work. The Journal of Finance, Vol. 25, 1970.

FAMA, E. The behavior of stock markets. The Journal of Business, Vol. 38, no. 1, 1965.

FARINA, D. et al. Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation. Nature Biomedical Engineering, Vol. 1, 2017.

FARRELL, J. The dividend discount model: a primer. Financial Analysis Journal, Vol. 41, no. 6, 1985.

FERNÁNDEZ, P. Company valuation methods: the most common errors in valuation. University of Navarra, 2007.

FONTANNILS, G., GENTILE, T. The stock market course. John Wiley & Sons, 2001.

GIANNETTI, E. O valor do amanhã. São Paulo: Editora Schwarcz, 2012.

GRAHAM, B. The intelligent investor. Nova York: HarperCollins, 2003.

GRAHAM, B., DODD, D. Security Analysis. Nova York: McGraw Hill, 1951.

GRAHAM, B., ZWEIG, J. The intelligent investor (revised edition). Nova York: Harper Business Essentials, 2003.

HAYKIN, S., WIDROW, B. Least-mean-square adaptive filters. John Wiley & Sons, v. 31, 2003.

HARARI, Y. N. Sapiens: Uma breve história da humanidade. Porto Alegre: L&PM Editores, 2015.

HASSETL, H. Double Q-Learning. Advances in Neural Information Processing, Vol. 23, 2010.

HAUGEN, R. OS segredos da bolsa. São Paulo: Pearson Prentice Hall, 2000.

HILL, J., FORD, W., FARRERAS, I. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in Human Behavior, Vol. 49, 2015.

JEZ, S., WASSMER, R. The impact of learning time on academic achievement. 2015.

JIANG, F. et al. Artificial intelligence in healthcare: past, present and future. Stroke and Vascular Neurology, Vol. 2, 2017.

JOACHIMS, T., FREITAG, D., MITCHELL, T. Webwatcher: a tour guide for the world wide web. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, 1997.

KARPOFF, J. The relation between price changes and trading volume: a survey. Journal of Financial and Quantitative Analysis, Vol. 22, no. 1, 1987.

KECMAN, V. Support vector machines: an introduction. Studies in fuzziness and soft computing, vol. 177, 2005.

KOLLER, T. Valuation: measuring and managing the value of companies. John Wiley & Sons, 2015.

LAKONISHOK, J., SHLEIFER, A., VISHNY, R. The impact of institutional trading on stock prices. Journal of Financial Economics, Vol. 32, 1992.

LEE, W., LIU, C., LU, C. Intelligent agent-based systems for personalized recommendations in internet commerce. Expert Systems with Applications, Vol. 22, 2002.

LEUDERS, A. Investindo em small caps. Rio de Janeiro: Alta Books, 2017.

LICHTENBERG, F. The impact of biomedical innovation on longevity and health. *Nordic Journal of Health Economics*, Vol. 5, no. 1, 2017.

MOHRI, M., ROSTAMIZADEH, A., TALWALKAR, A. *Foundations of machine learning*. Cambridge: The MIT Press, 2012.

MOONEY, R., ROY, L. Content-based book recommending using learning for text categorization. In *Proceedings of the SIGIR-99 Workshop on Recommender Systems: Algorithms and Evaluation*, 1999.

MORRIS, J., ALAM, P. *Analysis of the dot-com bubble of the 1990s*. 2008.

MURPHY, J. *Technical analysis of the financial markets: a comprehensive guide to trading methods and applications*. Nova York, New York Institute of Finance, 1999.

NANCE, R., SARGENT, R. Perspectives on the evolution of simulation. *Electrical Engineering and Computer Science*, no. 100, 2002.

NEWTON, I. *Principia*. Londres, 1696.

NYONG, M. Capital Market Development and Long run Economic Growth: Theory, Evidence and Analysis, *First Bank Review*, 1997.

OSLER, C. Support for resistance: technical analysis and intraday exchange rates. *Economic Policy Review*, 2000.

PEETERS, G., CZAPINSKI, J. Positive-negative asymmetry in evaluations: the distinction between affective and informational negativity effects. *European Review of Social Psychology*, no. 1, 1990.

PETRAM, L. *The world's first stock exchange*. Nova York: Columbia University Press, 2014.

PINHEIRO, J. *Mercado de Capitais*. São Paulo: Atlas, 2001.

PITIS, S. Rethinking the discount factor in reinforcement learning: a decision theoretic approach. *Association for the Advancement of Artificial Intelligence*, 2019.

POOLE, A., MACKWORTH, A., GOEBEL, R. *Computational intelligence: a logical approach*. Nova York: Oxford University Press, 1998.

PRING, M. Technical analysis explained: the successful investor's guide to spotting investment trends and turning points. McGraw Hill, 1991.

QUAH, J., SRIGANESH, M. Real-time credit card fraud detection using computational intelligence. Expert Systems with Applications, Vol. 35, 2008.

QUINLAN, J. C4.5: programs for machine learning. Sao Francisco, Morgan Kaufmann Publishers, 1993.

QUINLAN, J. Discovering rules by induction from large collections of examples. Expert systems in the micro electronic age, 1979.

ROSENBAUM, J. Investment banking: valuation, leveraged buyouts, and mergers & acquisitions. John Wiley & Sons, 2013.

ROSS, S. Introduction to stochastic dynamic programming. New York: Academic Press, 1983.

ROSS, S., WESTERFIELD, R., JORDAN, B. Princípios de administração financeira. São Paulo: Atlas, 2000.

ROSTOW, W. The stages of economic growth. The Economic History Review, Vol. 12, no. 1, 1959.

RUSSEL, S., NORWIG, P. Artificial intelligence: a modern approach. Harlow: Pearson Education, 2010.

SANTOS, L., MONTEZANO, R. Value and growth stocks in Brazil: risks and returns for one- and two-dimensional portfolios under different economic conditions. Revista Contabilidade e Finanças, Vol. 22, no. 56, 2011.

SHELLEY, M. Frankenstein. Londres, 1818.

SHOSTAK, F. In defense of fundamental analysis: a critique of the efficient market hypothesis. Review of Austrian Economics, Vol. 10, no. 2, 1997.

SIMON, H. Rational decision making in business organizations. Nobel Prize in Economics, 1978.

SKOWRONSKI, J., CARLSTON, D. Negativity and extremity biases in impression formation: a review of explanations. Psychological Bulletin, no. 105, 1989.

- SMITH, A. A riqueza das nações: investigação sobre sua natureza e suas causas. São Paulo: Nova Cultural, 1996.
- SOMASHEKHAR, S. et al. Abstract S6-07: double blinded validation study to access performance of IBM artificial intelligence platform, Watson for oncology in comparison with multiple multidisciplinary tumor board - First study of 638 breast cancer cases. Cancer research, Vol. 77, 2017.
- SUTTON, R., BARTO, A. Reinforcement learning: an introduction. Cambridge: The MIT Press, 2015.
- TAGLIAMONTE, S., DENIS, D. Linguistic ruin? Lol! Instant messaging and teen language. American Speech, Vol. 83, no. 1, 2008.
- TELLES, M. Tributação e distribuição de lucros e dividendos no Brasil. Brasília, 2016.
- VON NEUMANN, J., MORGENSTEIN, O. Theory of games and economic behavior. Princeton: Princeton University Press, 1944.
- WATKINS, C. Learning from delayed rewards. Londres: King's College, 1989.
- WATKINS, C., DAYAN, P. Technical note: Q-Learning. Machine Learning, Vol. 8, 1992.
- WEAVER, W. Science and complexity. American Scientist, no. 36, 1948.
- WILDER, J. New concepts in technical trading systems. Greensboro: Trend Research, 1978.
- YANG, Y. An evaluation of statistical approaches to text categorization. 1999.