

Universidade de Brasília - UnB
Faculdade de tecnologia - FT
Engenharia de Redes de Comunicação

Redes Neurais Aplicadas à Predição De Ativos e Alocação De Capital

**Autores: Bruno Emerich Ferreira, Henrique Simas Coutinho
Barbosa**

Orientador: Geovany Araújo Borges

Brasília, DF

2021



Bruno Emerich Ferreira, Henrique Simas Coutinho Barbosa
Redes Neurais Aplicadas à Predição De Ativos e Alocação De Capital/ Bruno
Emerich Ferreira, Henrique Simas Coutinho Barbosa. – Brasília, DF, 2021-
71 p. : il. (algumas color.) ; 30 cm.

Orientador: Geovany Araújo Borges

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade de tecnologia - FT , 2021.

1. Markowitz. 2. IA. 3. finances. 4. neural networks. I. Geovany Araújo Borges.
II. Universidade de Brasília. III. Faculdade de Tecnologia - FT. IV. Redes Neurais
Aplicadas à Predição De Ativos e Alocação De Capital

CDU 02:141:005.6

Bruno Emerich Ferreira, Henrique Simas Coutinho Barbosa

Redes Neurais Aplicadas à Predição De Ativos e Alocação De Capital

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

Brasília, DF, 16 de Novembro de 2021:

Geovany Araújo Borges
Orientador

Alexandre Ricardo Soares Romariz
Examinador Interno

Daniel Guerreiro e Silva
Examinador Interno

Brasília, DF
2021

Agradecimentos

Gostaríamos de agradecer aos familiares, amigos e docentes presentes durante a graduação. Vocês nos inspiram a ser profissionais mais dedicados e amáveis todos os dias de nossas vidas.

Resumo

A predição dos direcionamentos do mercado é algo muito visado pelos investidores. Em uma sociedade cada vez mais digitalizada e globalizada, a conectividade e o acesso à informação moldam um mercado que se torna mais especulativo e imprevisível. O *Deep Learning* é uma técnica poderosa que vem sendo aprimorada nos últimos anos e surge com grande potencial de extrair informações de dados que se retroalimentam gerando novas entradas em sistemas com uma modelagem complexa capaz de superar a compreensão humana. O projeto explorará as oportunidades de investimento a partir da predição de preços de ativos na bolsa em conjunto com análise técnica de investimentos e estratégias de *trading*, guiadas por métricas de rentabilidade, margem de erro de predição e análise de risco. A implementação de uma rede neural é utilizada para extrair as informações necessárias para executar duas estratégias de *swing-trade*, a estratégia de *momentum* onde se identificará tendências de subida ou de queda a partir de valores de ativos preditos no futuro e comparados em uma determinada janela de tempo. A estratégia de *market making* busca executar ordens de compra e de venda quando os valores flutuarem dentro um intervalo de predição de preço. Finalmente, utiliza-se o critério de alocação de Markowitz em que é realizada a busca das fronteiras eficientes de investimento do capital para um dado conjunto de ativos financeiros. A composição dessa carteira de investimento é feita a partir de análise de rentabilidade e gerenciamento de risco no investimento. O acoplamento entre o alocador e o preditor não foi concluído e será considerado em trabalhos futuros. Os objetos de estudo têm a sua atuação avaliada separadamente dentro de seus próprios escopos de atuação, expondo resultados eficientes para introduzir todo o potencial de *Deep Learning* aplicado a área de análise financeira.

Palavras-chave: Aprendizado Profundo, Investimentos, Redes Neurais, Bolsa de Valores.

Abstract

Predicting market directions is something investors are really interested. In an increasingly digitalized and globalized society, connectivity and access to information shape a market that becomes even more speculative and unpredictable. Deep Learning is a powerful technique that has been improved in recent years and has great potential to extract information from data that feed back into systems with a complex modeling capable of surpassing human comprehension. The project seeks to explore investment opportunities from stock market prediction along with technical analysis of investments and trading strategies, guided by profitability metrics, prediction error margin and risk analysis. The implementation of a neural network is used to extract the information necessary to execute two well-known swing-trade strategies, the momentum strategy, which seeks to identify stock price tendencies predicted in the future and compare them with past values within a time window. And the market making strategy, where you seek to execute buy and sell orders when values fluctuate within a price prediction range. Finally, the Markowitz allocation criterion is applied, in which the search for efficient investment frontiers for a set of financial assets is performed. The investment portfolio is made from profitability analysis and investment risk management. Coupling between allocator and predictor has not been completed and will be considered future work. The objects of study have their performance evaluated separately within their own scope of action, exposing efficient results to introduce the full potential of Deep Learning applied to the area of financial analysis.

Keywords: Deep Learning, Investments, Neural Networks, Stock Market

Sumário

1	INTRODUÇÃO	1
1.1	Diversificação	1
1.2	Análise do ativo	1
1.3	Objetivo	2
1.4	Resultados Alcançados	2
2	RENDA VARIÁVEL — BOLSA DE VALORES	4
2.1	Ações	4
2.2	O que é a Bolsa de Valores?	5
2.2.1	Comissão de Valores Mobiliários	6
2.2.2	História da Bolsa Brasileira	7
2.2.2.1	Funcionamento da Bolsa de Valores	8
2.2.3	Índices de mercado	10
2.3	Análise técnica	11
2.3.1	Gráfico de <i>Candlesticks</i>	11
2.3.2	Tendências	12
2.3.3	Suporte e Resistência	13
2.3.4	Teoria de Dow Jones	15
2.3.5	Indicadores financeiros	17
2.4	Teoria de Markowitz	18
3	MACHINE LEARNING	21
3.1	Métodos de Aprendizagem em Machine Learning	22
3.1.1	Aprendizagem Supervisionada	22
3.1.2	Aprendizagem Não Supervisionada	23
3.1.3	Aprendizagem Semi-supervisionada	23
3.1.4	Aprendizagem por Reforço	24
3.2	Análise de Performance do Modelo de ML	24
3.3	Redes Neurais Artificiais	26
3.3.1	Camada de entrada	27
3.3.2	Camada oculta	28
3.3.3	Camada de saída	29
3.4	Deep Learning, RNN e LSTM	29
4	MACHINE LEARNING EM SISTEMAS DE TRADING	37
4.1	Estrutura do Sistema	37

4.1.1	Plataforma	37
4.1.2	Estatísticas financeiras	38
4.1.3	Dataset	44
4.1.4	Alocador	46
4.1.5	Rede Neural	48
4.2	Metodologia	49
5	TESTES E RESULTADOS	54
5.1	Estratégia <i>Momentum</i>	54
5.2	Estratégia <i>Market Making</i>	56
5.3	Alocação do capital	59
5.3.1	Setores distintos	59
5.3.2	Setores iguais	62
6	CONSIDERAÇÕES FINAIS	65
6.1	Escopo e limitações	65
6.2	Conclusão	65
6.3	Trabalhos futuros	66
	REFERÊNCIAS	68
	ANEXOS	71

Lista de ilustrações

Figura 2.1 – Organizacional - Comissão de Valores Mobiliários. Fonte: gov.br	7
Figura 2.2 – Mercado Primário — IPO. Fonte: (INVESTIMENTOS, 2021)	9
Figura 2.3 – Mercado secundário - Compra e venda de papéis entre investidores. Fonte: (INVESTIMENTOS, 2021)	10
Figura 2.4 – Índice IBOVESPA - Janela de 2016-atual, mensal. Fonte: dos autores .	11
Figura 2.5 – Leitura <i>candlesticks</i> . Fonte: dos autores	12
Figura 2.6 – Padrões de <i>candlesticks</i> — (AGROCLASI.COM, 2021)	13
Figura 2.7 – PETR4 — Formação das 3 possíveis tendências. Fonte: dos autores . .	13
Figura 2.8 – Princípio da Inversão. Fonte: dos autores	15
Figura 2.9 – Pilar teoria de <i>Dow Jones</i> - As três tendências de mercado. Fonte: dos autores	16
Figura 2.10 – Indicações financeiros transpostos em gráfico. Fonte: dos autores	18
Figura 2.11 – Gráfico da Teoria de Markowitz. Fonte: dos autores	20
Figura 3.1 – Programação clássica (A) versus paradigma de aprendizado supervisi- onado de máquina (B). Adaptado de (CHOI et al., 2020)	21
Figura 3.2 – Modelo de rede neural (ANN). Adaptado de (HENRIQUE, 2018)	27
Figura 3.3 – Ativação de um neurônio. Fonte: dos autores.	29
Figura 3.4 – Lista de funções de ativação. Fonte: Sebastian Raschka	30
Figura 3.5 – Estrutura básica de uma RNN. Fonte: (ZHU ZHILE YANG, 2019)	31
Figura 3.6 – Porta de esquecimento de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)	34
Figura 3.7 – Porta de entrada de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)	34
Figura 3.8 – Porta de saída de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)	34
Figura 3.9 – Estrutura de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)	36
Figura 4.1 – Dataset - Módulo Rede Neural. Fonte: dos autores	45
Figura 4.2 – Arquitetura - Módulo Rede Neural. Fonte: dos autores	49
Figura 4.3 – Modelo de treinamento da rede neural. Fonte: dos autores	50
Figura 5.1 – Top 5 — Ativos mais rentáveis na estratégia de <i>momentum</i> . Fonte: dos autores	55
Figura 5.2 – Top 5 - Ativos com maior valor de erro sobre investimento na estratégia de <i>momentum</i> . Fonte: dos autores	55
Figura 5.3 – Top 5 - Ativos com maior erro quadrático médio na estratégia de <i>mo- mentum</i> . Fonte: dos autores	56

Figura 5.4 – Top 5 — Ativos mais rentáveis na estratégia de <i>market making</i> . Fonte: dos autores	57
Figura 5.5 – Top 5 — Ativos com maior valor sobre o investimento em <i>market making</i> . Fonte: dos autores	58
Figura 5.6 – Top 5 — Ativos que mais negociam na estratégia de <i>market making</i> . Fonte: dos autores	58
Figura 5.7 – Carteira multisetores — Distribuição de papéis. Fonte: dos autores . .	59
Figura 5.8 – Carteira multisetores — Distribuição de carteiras. Fonte: dos autores .	60
Figura 5.9 – Carteira multisetores — Distribuição de carteiras. Fonte: dos autores .	61
Figura 5.10–Carteira setORIZADA — Distribuição de papéis. Fonte: dos autores	62
Figura 5.11–Carteira setORIZADA — Distribuição de papéis. Fonte: dos autores	63
Figura 5.12–Carteira setORIZADA — Distribuição de papéis Fonte: dos autores	64

Lista de abreviaturas e siglas

B3 = Brasil, Bolsa, Balcão

CVM = Comissão de Valores Mobiliários

SLSQP = *Sequential Least Squares Programming*

API = *Application Programming Interface*

ML = *Machine Learning*

SMA = *Simple Moving Average*

EMA = *Exponential Moving Average*

MACD = *Moving Average Convergence Divergence*

RSI = *Relative Strength Index*

OBV = *On Balance Volume*

IPO = *Initial Public Offering*

OHLC = *Open, High, Low, Close*

DL = *Deep Learning*

IA = Inteligência Artificial

ANN = *Artificial Neural Network*

RNN = *Recurrent Neural Network*

LSTM = *Long Short-Term Memory Network*

EQM = Erro Médio Quadrático

RQM = Retorno Médio Quadrático

1 Introdução

A bolsa de valores sempre foi um ambiente caótico e desconhecido pelos brasileiros, desta forma a proporção de indivíduos que possuem cadastro em corretoras para realização de operações financeira sempre foi apenas uma pequena e singela fatia da população, fato que tem se alterado neste últimos anos com a crescente entrada de novos investidores. O número de CPFs cadastrados na bolsa de valores apresenta um crescimento vertiginoso acumulando uma alta de 30% apenas no último ano, sendo o investidor recente com um perfil mais jovem e com aporte menores (B3, 2021a).

1.1 Diversificação

Apesar do maior número de investidores na bolsa de valores ainda se tem um perfil de investidor que aplica a maior parte (se não todo) do capital em renda fixa devido ao risco inerente ao mercado de renda variável, porém muitas vezes o capital tem esta destinação devido à falta de conhecimento ou mesmo formas de operacionalizar o investimento no mercado lucrativo de ações. Atualmente também se tem um perfil de investidor que se adverte de realizar o investimento de capitais em mais de uma categoria de ativo, logo o seu portfólio não possui nenhuma diversificação e assim as oportunidades de melhores investimentos é desperdiçada (B3, 2021a).

A diversificação no investimento de ativos é justamente um das doutrinas adotadas ao longo da Teoria de Markowitz, pois mesmo que um ativo individual tenha um baixo risco intrínseco ainda há uma operação aberta com uma taxa de risco maior que o “portfólio ótimo” de Markowitz, deste modo é sempre com a composição de um portfólio diversificado que teremos uma mitigação de riscos e uma maior possibilidade de retorno (BRUM, 2008).

1.2 Análise do ativo

Com o perfil de investimento conservador majoritário presente na bolsa de valores tem-se que a análise de ativos, como ações, não é executado e há a ocorrência de execuções de investimentos por indicações de analistas que não acontece alinhadamente com os interesses pessoais do investidor, logo, há a necessidade de que o investidor tenha ferramentas e maneiras de poder tomar decisões de maneira mais consciente e assertiva para atendimento de suas expectativas.

O estudo de um ativo financeiro pode ser realizado através de diversos prismas

e abordagens distintas, uma das opções consiste na análise técnica, focada em realizar o estudo matemático e estatístico da cotação de um ativo para a orientação da tomada de decisão correta em relação a uma operação financeira (BOAINAIN; PEREIRA, 2009). Devido à análise técnica de um ativo estar intimamente relacionada a cálculos e estatísticas, é possível utilizar-se do advento da tecnologia para fazer o uso de redes neurais que possam identificar e correlacionar diversos indicadores que visam oportunidades de investimentos lucrativos.

A rede neural é uma forma revolucionária de realizar a análise de sinais financeiros e apresenta resultados que por vezes pode ser superior à aplicação de um método clássico, visto que tais estruturas realizam a identificação de padrões e possuem memória referente aos termos históricos da cotação, o que acaba permitindo a realização de uma predição de valor mais consistente de um dado ativo (FREITAS, 2001).

1.3 Objetivo

O escopo do trabalho é fornecer um auxílio inicial para o novo investidor da bolsa de valores, permitindo que ele realize a seleção de ativos de sua preferência e, com isso, o sistema aqui composto realiza a distribuição do capital disponível e fornece predições de valores futuros do ativo, para assim realizar o balizamento de decisões a respeito do posicionamento no mercado.

A automação de alocações de capital do investidor buscará sempre realizar da melhor maneira possível a relação de retorno por risco que seja mais rentável para o investidor, utilizando a Teoria de Markowitz (BRUM, 2008).

A predição de futuras cotações de um ativo será realizada com o uso de uma rede neural, que irá realizar a identificação de padrões que correlacionam os indicadores financeiros com o histórico de cotações e desta maneira proverá as predições. Durante a execução das posições será abordada a estratégia de *market making*, logo as cotações fornecidas pela rede neural funcionarão como patamares de entrada e saída das transações. Ao final do documento é esperado que se tenha uma clara identificação dos benefícios de adoção da estratégia de *market making*, em detrimento da estratégia de *momentum*.

1.4 Resultados Alcançados

Foi possível identificar a influência das técnicas de *Deep Learning* como preditor e extrator de informações visto que os dados de saída da rede neural formam uma curva de predição que se aproxima dos valores reais e quando comparadas as curvas, consegue-se obter bons parâmetros de rentabilidade em ambas estratégias de negociação implementadas, visto que na estratégia de *market making* foi possível se obter retornos de 106%

(aliado ao fato da valorização do ativo). O problema, no entanto, está na limitação encontrada na rede neural LSTM, que não consegue aproximar-se o suficiente dos valores reais para se tornar um preditor mais consistente.

2 Renda Variável — Bolsa de Valores

Neste capítulo serão abordados fundamentos teóricos importantes para o entendimento e desenvolvimento deste projeto. O capítulo se dispõe a definir os conceitos advindos da área de finanças e investimentos.

2.1 Ações

Os ativos mais negociados em bolsa de valores são indiscutivelmente as ações das empresas que possuem sociedade anônima e estão registradas junto aos órgãos competentes (na seção 2.2.1 é realizado um maior detalhamento) para poderem realizar tal comércio, porém, para entender esse mercado é necessário antes realizar uma definição formal do que é um “ação”. De maneira simplificada, porém sem menor completude, as ações são representações de uma fração do capital social de uma empresa, logo ao realizar a compra de uma ação o investidor está diretamente se tornando proprietário de uma pequena parcela daquela empresa e assim tendo os riscos inerentes do negócio com a proporcional parcela dos lucros e prejuízos da companhia em questão.

O mercado de ações movimenta bilhões de reais diariamente e desta maneira atrai investidores pelos benefícios em ser um sócio minoritário de uma companhia, pois as ações são um meio de se obter rendimentos mais vantajosos quando se compara com a renda fixa (por exemplo, poupança), logo as mesmas estão envoltas no mercado de renda variável, onde não se tem garantia de lucros (ou prejuízos) porém há possibilidades de maiores rendimentos. As ações podem ser classificadas em diversas categorias, pois cada “fatia” vendida de um empresa carrega consigo alguma denominação.

- Ações ordinárias

As ações ordinárias são aquelas que dão direito a voto nas assembleias gerais e participação nas decisões das empresas, porém o peso dos votos é ponderado pelo número de ações em posse do acionista, logo os acionistas majoritários são aquele que possuem o chamado “poder de decisão”, devido aos mesmos deterem uma maior quantidade e conseqüentemente proporção da empresa.

- Ações preferenciais

As ações preferenciais são aquelas que concedem ao acionista uma preferência para o recebimento de dividendos, porém ao custo de não se possuir direito de voto nas assembleias gerais. Os proprietários das ações preferenciais também recebem antecipadamente compensações, no caso de falência.

- *Units*

As ações *units* são uma composição mista de ações ordinárias e ações preferenciais, logo quando o investidor realiza a compra de uma *unit* está adquirindo uma mistura de tipos de ação e com isso pode se beneficiar das particularidades individuais de cada categoria, como o direito a voto nas assembleias gerais e a preferência no recebimento de dividendos.

- *Small caps*

Essas ações são de empresas com menor capitalização, conseqüentemente com uma menor liquidez, assim os valores desse tipo de ações geralmente é mais baixo e atrativo devido ao seu potencial de ganhos, porém ao custo de um risco mais elevado que as ações das grandes empresas, pois pode ocorrer do investidor realizar a aquisição da mesma e não conseguir negociar em bolsa devido à falta de liquidez. Os investidores dessa categoria de ativo geralmente estão mais aptos a correrem maiores riscos.

- *Medium caps*

Assim como as *small caps*, as ações *medium caps* se diferenciam por serem de empresas de porte médio, logo trazem consigo de maneira análoga a *small cap* uma maior possibilidade de ganhos ao custo de um maior risco. As ações dessa categoria são consideradas por muitos como o “*sweet spot*” dentre custo e a relação retorno/risco.

- *Blue chips*

Essas são as ações que possuem maior volume de negociação nos pregões da bolsa de valores, pois são pertencentes às maiores empresas listadas na bolsa e conseqüentemente apresentam uma alta taxa de liquidez, porém tais papeis apresentam um valor mais elevado quando comparados com empresas de menor porte. As *blue chips* são ações que geralmente se encaixam em qualquer perfil de investidor de renda variável, dado a sua facilidade de negociação e também sua viável taxa de risco.

2.2 O que é a Bolsa de Valores?

A bolsa de valores é uma entidade que proporciona um ambiente de mercado para que compradores e vendedores possam negociar de forma transparente, segura e eficiente ativos financeiros, como ações, derivativos de ações, títulos de renda fixa, títulos públicos federais, derivativos financeiros, moedas à vista e *commodities*. O ambiente em que são realizadas as negociações financeiras é totalmente digitalizado, atualmente, e deve garantir por lei a equidade, transparência, eficiência e segurança das operações de compra e venda dos ativos negociáveis em bolsa, assim compete à Comissão de Valores Mobiliários (CVM) realizar a regulamentação da bolsa brasileira.

2.2.1 Comissão de Valores Mobiliários

A CVM é uma entidade pública e autárquica vinculada ao Ministério da Economia, assim, é totalmente independente na esfera jurídica e possui capital próprio, sendo auto administrada. A CVM foi criada em 1976, regida pela Lei Nº 6.385, sendo seus objetivos: disciplinar, fiscalizar e desenvolver o mercado de valores mobiliários. As atribuições exercidas pela CVM são todas previstas em lei ([MOBILIÁRIOS, 2013](#)).

A primeira e fundamental atribuição da CVM é garantir a segurança e acesso do investidor ao mercado de ativos financeiros, desta maneira o órgão tem a competência de assegurar os direitos do investidor e realizar a instrução do mesmo quanto aos temas de educação financeira e acesso a informações dos ativos negociados pela bolsa de valores. A competência da CVM se estende também a cuidar da integridade do mercado de capitais, portanto busca garantir segurança, isonomia e transparência do ambiente de investimento, desta maneira a organização realiza o credenciamento e fiscalização ativa de todas as corretoras licenciadas e validadas para poder atuar no mercado de capitais, além de atuar para diminuição de quaisquer burocracias no tocante à realização das atividades financeiras ([YAZBEK, 2006](#)).

A entidade possui um organograma bem definido, conforme pode ser observado na Figura 2.1, assim o colegiado da instituição é composto por um presidente e quatro diretores, nomeados pelo Presidente da República e posteriormente aprovados pelo Senado Federal, sendo que tais cadeiras podem somente ser ocupadas por indivíduos com ilibada reputação e reconhecida competência em matéria de mercado de capitais. A estrutura executiva da CVM é completada pelas Superintendências Regionais de São Paulo e Brasília.

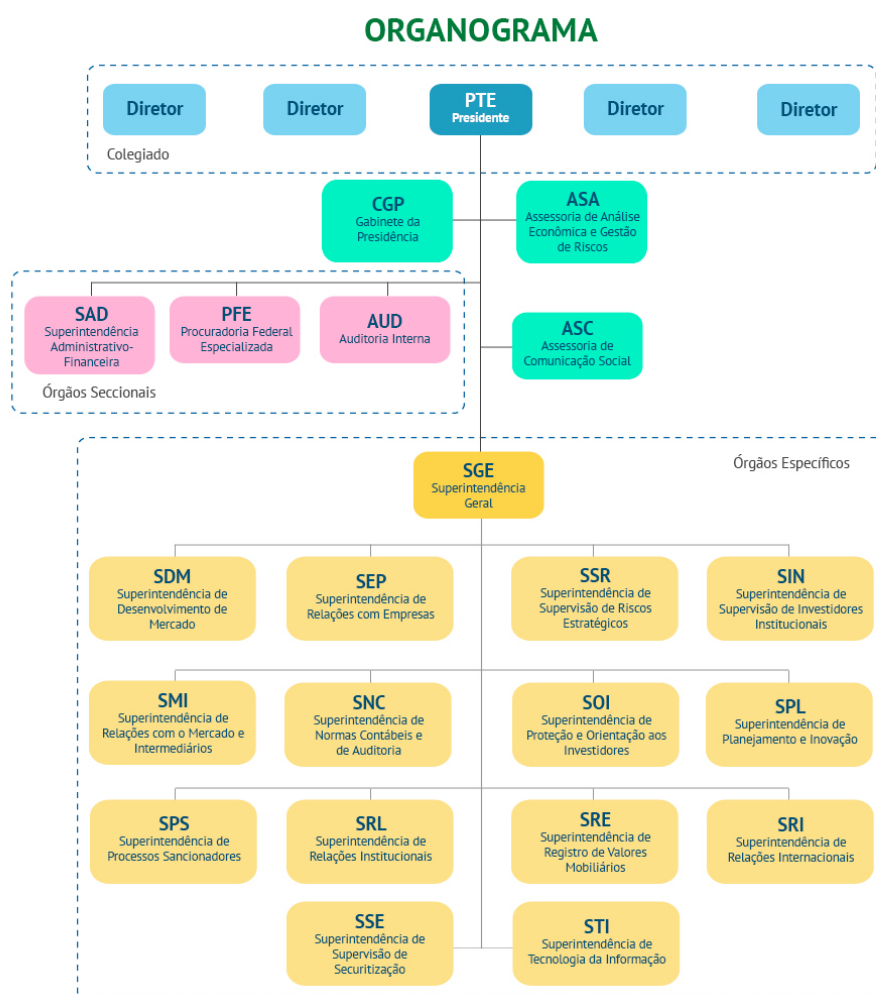


Figura 2.1 – Organizacional - Comissão de Valores Mobiliários. Fonte: gov.br

2.2.2 História da Bolsa Brasileira

O atual ponto de maturidade da bolsa brasileira só se tornou possível após inúmeros marcos históricos antecedentes que estruturam o molde atual em que se encontra a Brasil, Bolsa, Balcão (B3).

A história da bolsa de valores teve início com o comércio de mercadorias entre o Brasil e o mercado europeu, juntamente ao fato de que a família real portuguesa se encontrava no Brasil e, com isso, a importância do mercado exterior tomou forma e assim se instaurou a primeira bolsa de valores no Brasil, sediada no Rio de Janeiro. Nesta primordial bolsa a negociação de fretes navais para importação e exportação de bens era a principal atividade a ser executada, as quais eram executadas através de pregões ao ar livre no que ficou conhecido como Praça do Comércio, assim a primeira tentativa de realizar uma organização do mercado. Com o passar dos anos e a recente importância da cidade de São Paulo, foi instaurado na cidade a Bolsa Livre em 1890, porém a mesma teve uma vida útil curta encerrando suas atividades no ano seguinte devido à política do

encilhamento que causou uma grande crise econômica no país. Somente no ano de 1895 é que começaram a ser realizadas as atividades de negociação da Bolsa de Fundos Públicos de São Paulo, que 40 anos depois foi nomeada como “Bolsa Oficial de Valores de São Paulo”. Com o surgimento das sociedades corretoras em 1967, o nome mudou para Bolsa de Valores de São Paulo (BOVESPA), trazendo assim a familiaridade com a estrutura existente atualmente (SILVA, 2016) (UNUCSEH, 2020).

A BOVESPA foi ganhando cada vez mais importância nacional e assim unificando as nove bolsas de valores estaduais operantes que existiam na época, deste modo começaria a negociação centralizada das ações. Ainda no processo de unificação e centralização surgiu assim a BM&F (Bolsa de Mercadorias e Futuros) em 1986, onde foram negociadas as principais *commodities* (como o café, algodão e gado) e derivativos, ficando assim o mercado de ações centralizado na BOVESPA, que no que lhe concerne era composta por uma sociedade de empresas, corretoras de valores e bancos (UNUCSEH, 2020).

A partir do processo de intensificação da globalização, em meados dos anos 2000, gerou-se a intensa necessidade de se realizar a centralização das operações de ativos financeiros e, com isso, houve a unificação de todas as bolsas operantes além da finalização do pregão de viva-voz, que teve o seu fim no ano de 2005, tornando a bolsa totalmente eletrônica. No ano de 2008 tem-se finalmente a fundação da BM&FBovespa dada pela união da BOVESPA com a Bolsa de Mercadorias e Futuros, tornando assim a instituição conhecida ainda atualmente e possui papéis sendo negociados em bolsa até os dias atuais através do *ticker* B3SA3 (SILVA, 2016).

Em 2017 tivemos então a última etapa da história para chegarmos até o ponto de maturação atual, pois neste ano seria aprovado pela CVM e pelo CADE (Conselho Administrativo de Defesa Econômica) a fusão entre a BM&FBovespa e a CETIP, criando assim a atual B3 (Brasil, Bolsa, Balcão) que se situou em 5^o lugar no *ranking* das bolsas com maior volume de negociação. Atualmente na B3 temos a negociação de ações, ativos de Renda fixa públicos, privados e cotas de fundos, derivativos, *commodities* e moedas.

2.2.2.1 Funcionamento da Bolsa de Valores

Atualmente temos centenas de ativos sendo negociados no ambiente estabelecido pela B3, desta forma o investidor tem acesso a uma infinidade de opções para “operar” no mercado, porém é necessário se entender como tais ativos são disponibilizados neste ambiente de negociação.

Inicialmente é importante entender haver uma série de motivações que levam à decisão de uma empresa abrir o seu capital no mercado de ativos financeiros, dos quais podemos destacar o levantamento de capital (por vezes mais benéfico que a tomada de empréstimos devido a juros e prazos, uma vez que as ações não possuem prazo de validade e os juros são pagos como dividendo aos seus investidores), liquidez (permitindo assim



Figura 2.2 – Mercado Primário — IPO. Fonte: (INVESTIMENTOS, 2021)

aos acionistas realizarem a transformação de suas ações em dinheiro) e visibilidade comercial. Para que uma empresa possa abrir seu capital é necessário a mesma ser constituída como sociedade anônima, além de cumprir uma série de requisitos para se registrar como companhia aberta junto à CVM, com isso temos a primeira etapa do processo chamada IPO (sigla do inglês para *Initial Public Offering*) em que é executado o mercado primário, onde entram em cena as corretoras para realizar a organização das vendas destas frações da empresa para os investidores interessados, fazendo assim com que a companhia receba o dinheiro referente às ações. A Figura 2.2 mostra o esquemático do processo do mercado primário.

Dado que a empresa possua ações sendo negociadas em bolsa de valores, novas ações podem ofertadas através de propostas públicas conhecidas como *follow-on*, assim praticado novamente o mercado primário (INVESTIMENTOS, 2021).

Por sua vez, o mercado secundário ocorre com as ações da empresa que estão em posse dos investidores, portanto neste momento as frações da empresa são negociadas entre dois investidores sendo um deles o vendedor e o outro o comprador. No mercado secundário também temos a presença das corretoras atuando com entidades organizadoras e executoras das ordens de compra e venda pelos investidores, onde as mesmas são responsáveis por disponibilizar no sistema eletrônico da B3 o registro da ordem, mesmo quando executada internamente na corretora (por exemplo, comprador e vendedor de uma dada operação são constituintes clientes de uma mesma corretora). A Figura 2.3 mostra o esquemático do processo do mercado secundário. (INVESTIMENTOS, 2021)



Figura 2.3 – Mercado secundário - Compra e venda de papéis entre investidores. Fonte: (INVESTIMENTOS, 2021)

2.2.3 Índices de mercado

No mercado financeiro é muito comum o uso de índices e estatísticas para se poder realizar uma estimativa do rumo das cotações dos ativos, por conseguinte tais indicadores podem ser entendidos como o “termômetro” do mercado. Na bolsa brasileira existem diversos índices que buscam dar um panorama do ambiente econômico, sendo tais índices classificados em categorias como amplos, de sustentabilidade, de governança, setoriais e segmentos.

Os índices amplos são estatísticas que tentam mensurar e medir a variação das cotações de papéis de diversas empresas que podem ou não pertencer a setores completamente diferentes, portanto o índice reflete o comportamento médio de tais cotações e assim permite uma visão do panorama, o método de seleção das empresas varia em cada um dos índices devido ao propósito individual e único. O índice BOVESPA é o mais conhecido desta categoria (mostrado pela Figura 2.4), e como característica do mesmo tem-se que o balanceamento é quadrimestral e composto conforme as regras presentes em (B3, 2021b), assim o mesmo é composto por um número variável de ações que juntas representam em média 80% do volume negociado em bolsa. Outro exemplo de índice desta categoria é o Índice Brasil 100 (IBrX 100) que possui rebalanceamento também quadrimestral sendo composto pelas 100 ações mais negociadas na bolsa.

Os índices de sustentabilidade são ferramentas comparativas que avaliam as empresas listadas em bolsa no tocante ao tema de sustentabilidade empresarial, contempla a emissão de carbono, eficiência econômica e outros fatores, sendo o ICO2 um exemplo de índice desta categoria. Índices setoriais são ferramentas que buscam avaliar o comporta-

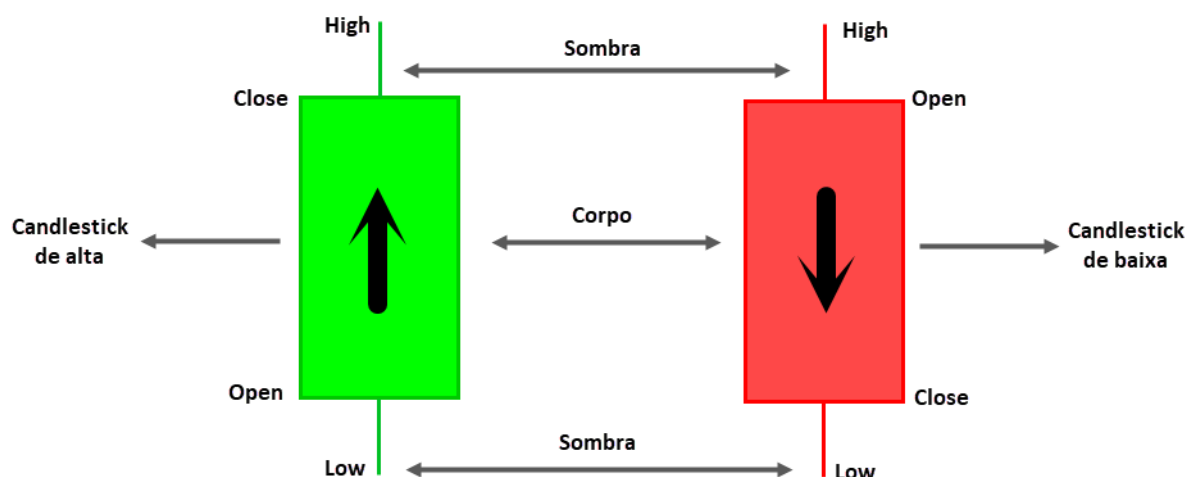


Figura 2.5 – Leitura *candlesticks*. Fonte: dos autores

realizar a leitura de um gráfico de *candlesticks* é necessário primeiro entender que cada *candlestick* irá representar o passeio executado pelo ativo naquele determinado período de tempo (expressão comumente traduzida para *timeframe* no mercado financeiro), assim a leitura individual da vela deve se iniciar pela identificação se a mesma é de alta ou de baixa, logo deve-se verificar a notação de cores utilizadas (no caso da Figura 2.5 verde e vermelho simbolizam respectivamente alta e baixa), pois assim será possível determinar o topo de uma vela representa a abertura ou fechamento. Cada *candlestick* representa quatro valores dados pelo acrônimo OHLC (*Open*, *High*, *Low*, *Close*) sendo assim numa vela de alta o ativo inicia a cotação no ponto *Open*, alcança a máxima *High* e mínima *Low* dada pelas sombras (em qualquer ordem) e finaliza o período pelo ponto *Close* (LACERDA, 2021).

A utilização de gráficos de *candlesticks* permite identificar a formação de alguns padrões que podem ser gatilhos de operações de compra ou venda, desta maneira proporciona ao investidor oportunidades de operações com lucro. Dentro da ramificação da análise gráfica existe uma área de estudo destinada a realizar a identificação de padrões e seu respectivo significado no contexto de oscilação do ativo, conforme pode ser visto em (LU; SHIU, 2012). A Figura 2.6 mostra alguns exemplos de padrões formados por *candlesticks*.

2.3.2 Tendências

Durante o passeio normal da cotação de um ativo inevitavelmente teremos a formação de uma das três possíveis tendências: alta, baixa ou lateral. A tendência de alta pode ser caracterizada pela sucessão de topos e fundos superiores, ao passo que a tendência de

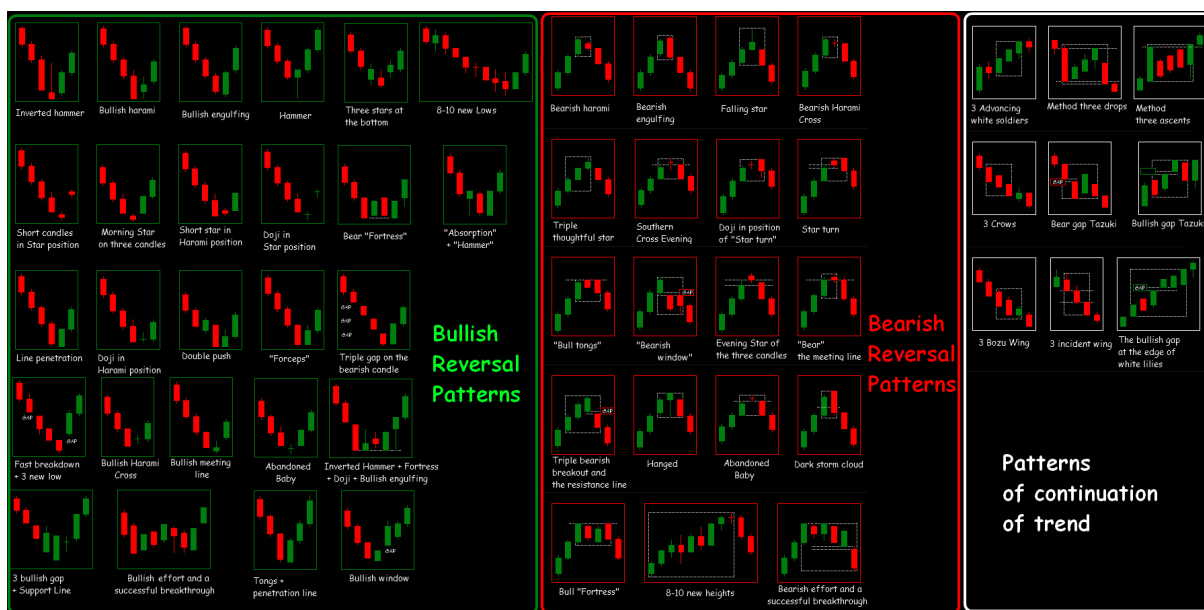


Figura 2.6 – Padrões de *candlesticks* — (AGROCLASI.COM, 2021)



Figura 2.7 – PETR4 — Formação das 3 possíveis tendências. Fonte: dos autores

baixa é dado pelo oposto, sucessão de topos e fundos inferiores. A tendência lateral é a falta de direção (superior ou inferior) de consecutivos topos e fundos, conseqüentemente o ativo oscila em torno de uma linha horizontal e não apresenta viés direcional em sua cotação. A Figura 2.7 ilustra claramente na ação preferencial da Petrobras©(PETR4) a formação das três possíveis tendências.

2.3.3 Suporte e Resistência

Os patamares de suporte e resistência são linhas características na evolução de preço de um ativo financeiro, pois representam limiares que a cotação enfrenta dificuldades de exceder, pois, é um momento que o movimento comprador ou mesmo vendedor perde força e o movimento conseqüentemente não tem o “fôlego” necessário para realizar o

rompimento do patamar. As linhas de suporte e resistência podem ser encontradas com a união de fundos e topos consecutivos, respectivamente, assim as linhas de tais níveis podem se encontrar e formar figuras importantes como a bandeira (CORREIA, 2008).

A definição empregada por (ELDER, 2004) define o suporte como um nível de preço que a força compradora é suficientemente grande para reverter o movimento de declínio do preço do ativo, assim a mesma lógica aplicada para o nível de resistência porém neste momento é a força vendedora que possui enfoque para reverter a tendência de alta.

Outro ponto importante no tema é o Princípio da Inversão (também conhecido como princípio da mudança de polaridade), é um fenômeno que ocorre no momento em que há uma troca de comportamento de um determinado patamar de preço, assim nesse momento podemos ter uma resistência virando um suporte e vice-versa (CORREIA, 2008).

Um exemplo do Princípio da Inversão pode ser observado na Figura 2.8 em que existe uma tendência de alta impedida de continuar a guinada devido a forte zona de resistência em que se realizou anteriormente o teste (por mais de uma vez, conforme pode ser visto pelas setas laranjas) e não houve rompimento, conseqüentemente a confiança dos investidores está voltada a realizar a abertura de posições de venda no próximo toque (data de 22/10/2019), pois especulam que será um novo *bouncing* (efeito que remete ao ato da cotação “quicar” em dado patamar) da cotação e o valor começaria a cair novamente, porém, o movimento de venda não tem força e há rompimento da resistência devido à força compradora, neste momento ocorre o Princípio da Inversão e assim o preço de R\$ 27,92 passa a se comportar como um suporte. Os investidores com posições de venda abertas permanecem na operação com prejuízo, pois esperam encerrar a operação empatados. Logo quando ocorrer a correção para o patamar de R\$ 27,92 haverá presente uma força compradora no ativo para liquidez das posições, fato que empurrará a cotação do ativo para uma tendência de alta que ocorre de fato no dia 03/12/2019 sendo testada mais a frente novamente conforme mostrado pela seta roxa.



Figura 2.8 – Princípio da Inversão. Fonte: dos autores

2.3.4 Teoria de Dow Jones

Charles Dow (1851-1902) foi um jornalista americano que escrevia colunas para editoriais do *Wall Street Journal* a respeito do comportamento de mercado financeiro, sendo de sua autoria o primeiro índice do mercado de ações denominado *Dow Jones Index Average*. Anos após a morte de Dow, William Peter Hamilton realizou a organização das colunas publicas por Charles Dow para realizar a estruturação da teoria de Dow, que em 1932 ganhou uma obra de Robert Rhea chamada “*Dow Theory*”. A teoria de Dow é baseada em seis pilares que visam fundamentar e validar os movimentos executados pelo mercado, conforme mostrado em (BOTELHO, 2004) (NORONHA, 1995).

O primeiro pilar da tese é que os preços descontam tudo, de forma que toda e qualquer informação relevante para o ativo está embutida no preço, sejam essas informações provenientes de demonstrações financeiras, relações políticas, caráter especulativo, cunho jornalístico, etc. Neste ponto da teoria temos a definição dos “atos divinos” que são fenômenos/eventos que não tem previsibilidade como calamidades públicas, catástrofes naturais, atentados terroristas e pandemia (BORGES, 2010), que provocaram uma oscilação repentina do ativo que será absorvida pelo mercado e assim refletida na precificação de mercado (GUARNIERI, 2006).

O segundo fundamento da teoria é que o mercado possui três tendências e as realiza simultaneamente. A tendência primária, que dita a direção do movimento geral possui longa duração (escopos anuais) ao passo que a tendência secundária é de médio prazo (escopo de mensais) e pode ou não estar no mesmo sentido do movimento principal, ora realizando movimentos de correções como os *pullgaps*, e por último temos a tendência terciária que é de curto prazo (escopo de dias a semanas) em que ocorre às flutuações da cotação e sua direção pode ou não corresponder as outras tendências. A Figura 2.9 evidencia a existência de tais movimentos.

O terceiro princípio é referente às tendências primárias terem três fases, sendo



Figura 2.9 – Pilar teoria de *Dow Jones*- As três tendências de mercado. Fonte: dos autores

assim em movimentos de guinada *bullish*, temos as fases: acumulação, alta sensível e euforia. A fase inicial de acumulação ocorre quando os investidores bem informados se posicionam de forma comprada, seguida da fase de alta sensível, quando os seguidores de mercado começam a executar ordens de compra e se posicionar no mercado, finalmente a fase de euforia é definida pelo momento em que existe uma alta especulação do mercado e os bem informados começam a realizar operações ao passo que investidores mais leigos começam a se posicionar em ordens de compra, como resultado do movimento em geral temos uma sequência de reversão para um movimento descendente.

O movimento de declínio *bearish* da tendência primária é composta pelas fases: distribuição, pânico e a fase de baixa lenta. No primeiro instante temos a distribuição definida pelo momento de entrada dos investidores mais municiados de informação, seguida imediatamente pela fase do pânico, a qual é composta pelo movimento de declínio causado pelos investidores do “efeito manada” que executam ordens de venda e finaliza na fase de baixa lenta que é a representação que a queda perdeu força e momento ideal para a realização de lucros dos investidores bem informados (GUARNIERI, 2006).

O quarto fundamento da teoria é que as reversões precisam ser confirmadas, desta maneira uma reversão apenas será confirmada quando dois ou mais índices apontarem para tal, assim podendo ser confirmado pelo movimento paralelo de duas ações do mesmo setor, por exemplo.

O quinto pilar faz alusão de que o volume deve confirmar a tendência em vigor, deste modo em um movimento *bullish* o volume das negociações deve aumentar quando a cotação estiver se valorizando e diminuir quando a cotação estiver em declínio (GUARNIERI, 2006).

O último fundamento da teoria diz que a tendência perdura até a confirmação contrária, sendo assim uma tendência de alta é apenas rompida no momento em que houver uma falha na sucessão de topos e fundos superiores, neste momento é decretada a

quebra de uma tendência seja por lateralização ou reversão.

2.3.5 Indicadores financeiros

Com o avanço da computação e ferramentas de análise gráfica a partir da década de 80, todo um conjunto de ferramentas matemáticas e estatísticas foi criado para auxiliar a análise técnica a possuir mais assertividade e ferramental de decisão ao tomar posições financeiras, sendo assim dando oportunidades de realização de operações com menor risco e conseqüente margem de lucro induzida maior (GUARNIERI, 2006).

Os indicadores financeiros nada mais são que transposições gráficas de fórmulas matemáticas em que os dados de entrada de tais fórmulas são compostos dos valores históricos da cotação e volume do ativo, desta forma tais ferramentas proporcionam um senso mais crítico que buscam validar consensos de forma empírica. Os indicadores possuem as funções de realizar o filtro de flutuações de preço, facilitar a interpretação do gráfico e auxiliar no momento de abertura de posição. Diante dos objetivos que buscam ser atendidos pelos indicadores financeiros, é importante que o uso dos mesmos seja realizado de forma calibrada, pois à medida que é reduzida a amostragem, haverá maior geração de falsos sinais, ao mesmo passo que se houver uma sobreamostragem podemos perder oportunidades lucrativas, por esse prisma é visto que a calibração dos indicadores é primordial.

Há uma infinidade de indicadores disponíveis para serem utilizados na análise gráfica, em que cada um possui uma série de parâmetros configuráveis, logo o número de possibilidades é enorme e por este motivo é interessante que o investidor realize o uso e possua o conhecimento acerca de um conjunto seletivo de estatísticas para que possam ser aperfeiçoadas as estratégias de operação em bolsa. Pode-se dividir em grandes grupos os indicadores de momento, volume e tendência, porém é possível que uma estatística apresente características de mais de um grupo, logo durante a escolha de um conjunto é preferível tomar elementos que compartilham categorias para que mutualmente possa-se confirmar, ao invés de termos exclusão de indicação devido às características diferentes escolhidas em cada estatística (CORREIA, 2008).

Dentre os indicadores disponíveis temos que os mais comuns são: *Simple Moving Average (SMA)*, *Exponential Moving Average (EMA)*, *Standard Deviation (SDT)*, *Moving Average Convergence/Divergence (MACD)*, *Bollinger Bands*, *Relative Strength Index (RSI)*, *On Balance Volume (OBV)*, *Stochastic (STO)*, *Ichimoku Clouds*. Os indicadores elencados anteriormente foram implementados e terão seu detalhamento na seção 4.1.3. A Figura 2.10 mostra a aplicação da *SMA*, *MACD*, *RSI* e *Bollinger Bands* sobre o gráfico da Petrobras©.

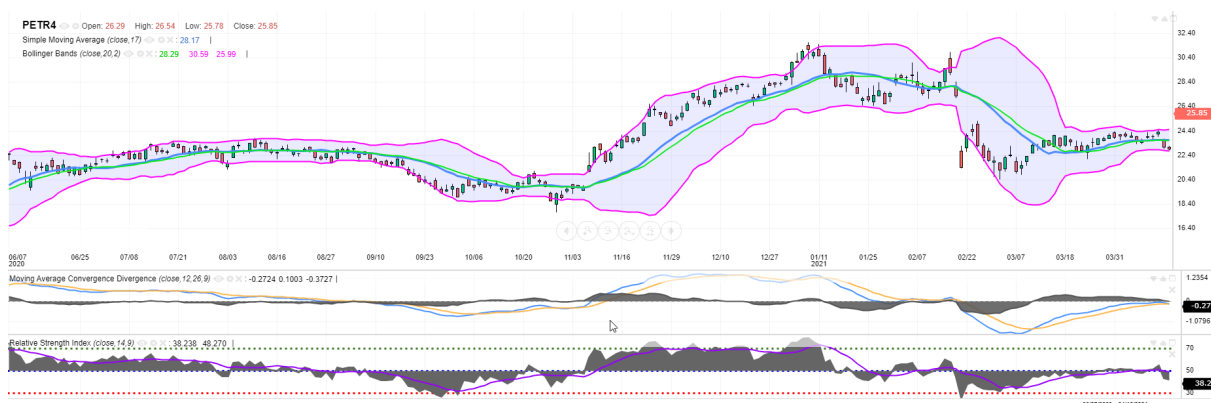


Figura 2.10 – Indicações financeiros transpostos em gráfico. Fonte: dos autores

2.4 Teoria de Markowitz

A escolha dos ativos e seus respectivos pesos dentro da carteira é um pilar fundamental e de extrema importância para a rentabilidade das operações financeiras, desta maneira a seleção dos ativos pode ser feita de maneira subjetiva a cada investidor, pois é do interesse individual o investimento em determinados setores e marcas. A alocação, pode ser realizada de formas arbitrárias uma vez que é definido o conjunto de ativos constituintes na carteira do investidor, porém para um dado conjunto de ativos sempre haverá somente uma alocação possível que apresenta a melhor relação entre risco e retorno, logo é neste ponto que a brilhante Teoria de Markowitz atua em busca do portfólio ótimo (ZANINI; FIGUEIREDO, 2021).

A teoria da seleção ótima de portfólios com base em um conjunto de instrumentos financeiros foi desenvolvida por Harry Markowitz em 1952, sendo seu principal objetivo realizar a melhor alocação possível do capital disponível em busca da otimização da relação retorno por risco. Um fator importante da teoria é que não há qualquer diferenciação ou mesmo impedimento de generalização, sendo assim contempla o universo de todos os bens financeiros que podem ser negociados em bolsa, visto que a teoria é totalmente isonômica quanto ao tipo do ativo, pois para a mesma precisa-se saber apenas o risco e retorno do

ativo desejado (BRUM, 2008)(ZANINI; FIGUEIREDO, 2021).

$$P_n = \sum_{n=1}^n \frac{C_n}{C_{n-1}}, \quad (2.1)$$

$$\sigma_n = \sqrt{\frac{\sum_{i=0}^n (C_n - \bar{C})^2}{n-1}}, \quad (2.2)$$

$$P = \sum_{n=1}^n X_n P_n, \quad (2.3)$$

$$\sigma = \sqrt{\sum_{i=1}^i \sum_{j=1}^j \rho_{ij} \sigma_i \sigma_j X_i X_j}, \quad (2.4)$$

$$\sum_{n=1}^n X_n = 1, \quad (2.5)$$

em que C_n , P_n , σ_n , P , σ e X_n são respectivamente a cotação n-ésima, retorno do n-ésimo ativo, risco do n-ésimo ativo, retorno do portfólio, risco do portfólio e Quantidade investida no n-ésimo ativo.

Inicialmente é necessário entender o que é compreendido como risco e retorno, no contexto da teoria de Markowitz. O retorno P_n de um n-ésimo ativo é correspondente à média das razões entre períodos consecutivos, conforme pode ser visto na equação (2.1), ao passo que o risco de um ativo n-ésimo é dado pelo desvio padrão σ_n da cotação do período de análise, como é mostrado na equação (2.2), porém para o cálculo do risco associado ao portfólio é necessário considerar o coeficiente de correlação ρ_{ij} entre dois ativos i e j quaisquer, desta maneira o risco total σ de um dado portfólio pode ser definido como mostrado na equação (2.4) (SEIDEL, 2018).

Com a definição dos conceitos relacionados a teoria de Markowitz, podemos imaginar uma situação em que um investidor tenha uma quantia de capital X e o queira alocar em diferentes ativos, em que para cada ativo $1, 2, \dots, n$ (no caso X_n representa o valor alocado no n-ésimo ativo) já temos a definição do seu retorno P_n e risco σ_n , logo com a aplicação de minimizações e da Teoria de Markowitz é possível encontrar um portfólio ótimo que apresente a maior taxa de retorno possível para um determinado patamar de risco, ou por outro ponto de vista, se encontrar um portfólio que apresente o menor risco possível para uma dada taxa de retorno. A carteira eficiente deve, de forma logica, utilizar todo o capital disponível, sendo assim a condição mostrada na equação (2.5) tem que ser respeitada (BRUM, 2008).

A execução do Teoria de Markowitz resulta em uma fronteira de alocação eficiente em que temos o máximo retorno para um dado nível de risco assumido, sendo assim, o gráfico mostrado na Figura 2.11 mostra a disposição de algumas carteiras teóricas, das quais as vermelhas se caracterizam por serem carteiras não eficientes por não apresentarem o maior retorno possível para o seu inerente nível de risco. Já as carteiras verdes

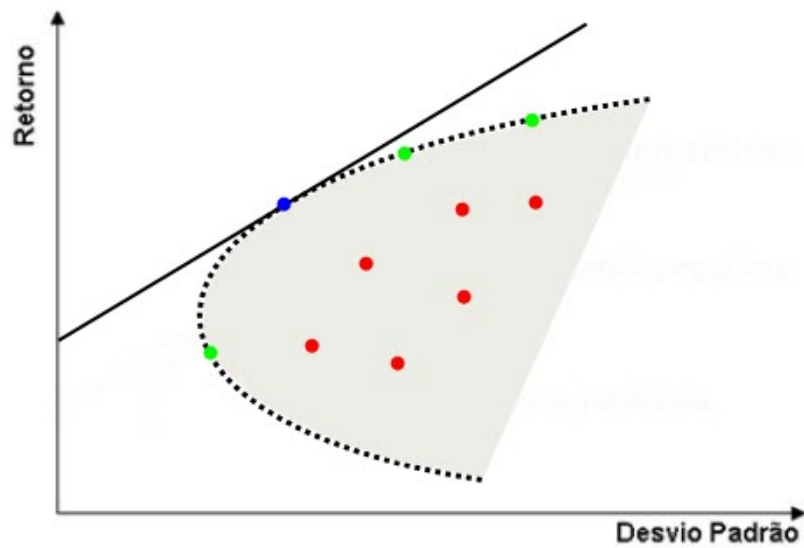


Figura 2.11 – Gráfico da Teoria de Markowitz. Fonte: dos autores

são carteiras eficientes, pois utilizam o capital da forma mais rentável possível, porém a carteira azul é comumente chamada de carteira de mercado, pois apresenta a melhor razão retorno por risco possível.

Vale a informação de que a teoria possui limitações inerentes à forma pela qual foram moldados seus objetivos primários, pois não temos a consideração de impostos e custos relacionados ao capital investido e a alocação é feita de forma determinística sem considerar nenhuma racionalidade pelo investidor, além de que o modelo contém a simplificação de que o mercado funciona de forma ideal em que todas as ordens de negociação serão executadas independentemente de qualquer interferência externa (BRUM, 2008).

3 Machine Learning

Machine Learning (ML) ou Aprendizado de Máquina é um campo que se concentra no aspecto de aprendizagem, dentro do escopo de inteligência artificial, com foco em desenvolvimento de algoritmos que melhor representam um conjunto de dados (CHOI et al., 2020). Em contraste com a programação clássica (Fig. 3.1A) na qual um algoritmo pode ser explicitamente codificado usando características conhecidas. ML utiliza subconjuntos de dados para gerar um algoritmo que pode usar combinações novas ou diferentes de características e pesos que podem ser derivados do primeiro paradigma (Fig. 3.1B) (GARETH et al., 2013) (T TIBSHIRANI R, 2009). ML contém muitos modelos e métodos, incluindo aprendizado profundo (*Deep Learning*) e redes neurais artificiais (*ANN - Artificial Neural Networks*).

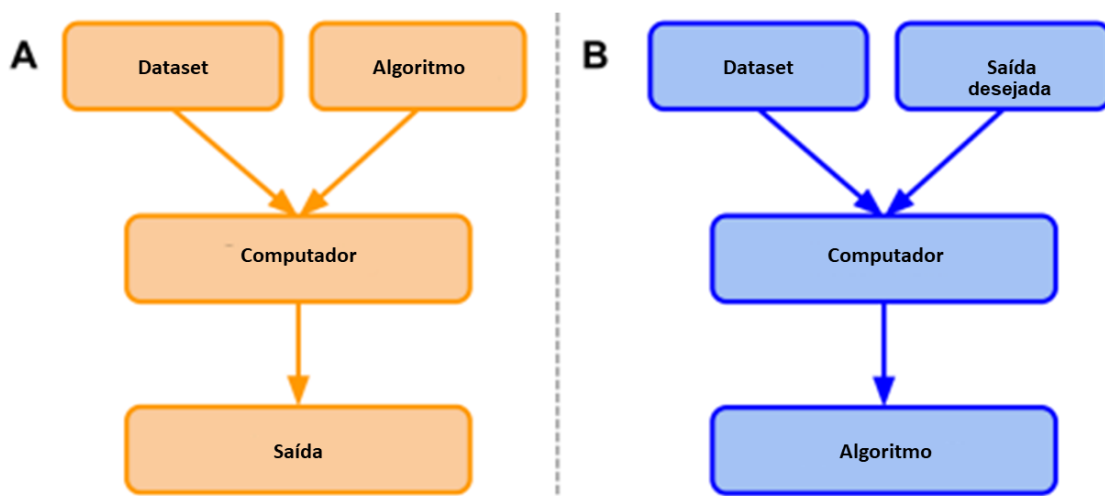


Figura 3.1 – Programação clássica (A) versus paradigma de aprendizado supervisionado de máquina (B). Adaptado de (CHOI et al., 2020)

3.1 Métodos de Aprendizagem em Machine Learning

Existem quatro métodos de aprendizagem comumente utilizados, cada um útil para resolver tarefas diferentes: aprendizagem supervisionada, não supervisionada, semi-supervisionada, e por reforço (GARETH et al., 2013)(T TIBSHIRANI R, 2009)(NVIDIA, 2018).

3.1.1 Aprendizagem Supervisionada

Suponha que uma imobiliária gostaria de prever o preço de uma casa com base em características específicas da casa. Para começar, a empresa iria primeiro reunir um conjunto de dados (*dataset*) que contém diversas instâncias(GARETH et al., 2013) (T TIBSHIRANI R, 2009) (MOLNAR, 2019). Cada instância representa uma observação singular do dataset, no caso do exemplo, de uma casa e características associadas. As características (*features*) são os registros de propriedade de uma casa que podem ser utilizados para prever preços (como por exemplo metragem quadrada total, número de andares, a presença de um pátio, etc) (GARETH et al., 2013) (T TIBSHIRANI R, 2009)(MOLNAR, 2019). O alvo é a característica a ser prevista, neste caso, o preço da habitação. *Datasets* são geralmente divididos em conjuntos de treinamento, validação e teste (os modelos sempre terão um desempenho ótimo com os dados em que são treinados)

A aprendizagem supervisionada usa padrões identificados no conjunto de dados de treinamento a fim de mapear suas características para o alvo, de forma a possibilitar que um algoritmo possa fazer previsões em *datasets* futuros. É chamada de supervisionada pois o modelo infere um algoritmo de pares de característica-alvo e é informado, pelo alvo, se ele previu corretamente (CHOI et al., 2020). Ou seja, as características, x são mapeadas para o alvo Y , que aprende a função de mapeamento f , de modo que as predições podem ser aproximadas usando o algoritmo $Y = f(x)$. O desempenho do algoritmo é avaliado no *dataset* de teste, que contém dados que o algoritmo nunca viu antes (CHOI et al., 2020). As etapas básicas do aprendizado de máquina supervisionado são:

1. Adquirir um conjunto de dados e dividi-lo em conjuntos de dados de treinamento, validação e teste;
2. Usar os conjuntos de dados de treinamento e validação para construir um modelo da relação entre os recursos e o alvo;
3. Avaliar o modelo por meio do conjunto de dados de teste para determinar a qualidade da predição. Em cada iteração, o desempenho do algoritmo nos dados de treinamento é comparado com o desempenho no conjunto de validação. Dessa forma, o algoritmo é ajustado pelos dados do conjunto de validação. Na medida em que o conjunto de

validação pode ser diferente do conjunto de teste, o desempenho do algoritmo pode ou não generalizar, conceito que será discutido mais adiante na seção 3.2.

As tarefas de aprendizagem supervisionada mais comuns são regressão e classificação (CHOI et al., 2020). A regressão envolve previsão de dados numéricos, como resultados de testes, valores laboratoriais, ou preços de um determinado item (CHOI et al., 2020). A classificação, por outro lado, envolve prever a qual categoria pertence um exemplo. Um exemplo clássico no escopo de ciência de dados consiste na leitura de imagens de cães e gatos e classificação nas categorias: 0 - Este exemplo é um cão; 1 - Este exemplo é um gato (CHOI et al., 2020).

3.1.2 Aprendizagem Não Supervisionada

Em contraste com o aprendizado supervisionado, o aprendizado não supervisionado visa detectar padrões em um *dataset* e categorizar instâncias individuais presentes neste. (GARETH et al., 2013) (T TIBSHIRANI R, 2009) (NVIDIA, 2018) Esses algoritmos são não supervisionados pois os padrões, que podem ou não existir em um *dataset*, não são informados pelo alvo e são deixados para serem determinados pelo algoritmo. Algumas das tarefas de aprendizagem não supervisionada mais comuns são agrupamento, associação e detecção de anomalias (GARETH et al., 2013) (T TIBSHIRANI R, 2009) (NVIDIA, 2018).

O agrupamento, como o nome sugere, agrupa as instâncias do *dataset* em *clusters* separados com base em combinações específicas de suas características. (GARETH et al., 2013) (T TIBSHIRANI R, 2009) (NVIDIA, 2018). No caso do exemplo da imobiliária, após uma investigação mais aprofundada, pode-se descobrir que os *clusters* representam os três arquitetos separados responsáveis por projetar as casas presentes no *dataset*, característica que não estava presente no *dataset* de treinamento.

3.1.3 Aprendizagem Semi-supervisionada

A aprendizagem semi-supervisionada pode ser considerada o “meio-termo” entre a aprendizagem supervisionada e não supervisionada e é particularmente útil para *datasets* que contêm dados rotulados e não rotulados (ou seja, todas as características estão presentes, mas nem todas possuem alvos associados) (NVIDIA, 2018). Esta situação normalmente acontece quando rotular imagens torna-se demorado e custo proibitivo. O aprendizado semi-supervisionado é frequentemente usado para imagens médicas, em que um médico pode rotular um pequeno subconjunto de imagens e usá-las para treinar um modelo. Este modelo é então usado para classificar o resto das imagens não rotuladas no *dataset*. O *dataset* rotulado resultante é então usado para treinar um modelo de trabalho que deve, em teoria, superar os modelos não supervisionados.

3.1.4 Aprendizagem por Reforço

Finalmente, a aprendizagem por reforço é a técnica de treinar um algoritmo para uma tarefa específica onde nenhuma resposta é correta, mas um resultado geral é desejado (TIBSHIRANI R, 2009) (NVIDIA, 2018). É indiscutivelmente a tentativa mais próxima de modelar a experiência de aprendizagem humana porque também aprende por tentativa e erro ao invés de dados apenas (TIBSHIRANI R, 2009) (NVIDIA, 2018). Imagine que alguém pudesse treinar um algoritmo para jogar o videogame *Super Mario Bros*, em que o objetivo do jogo é mover o personagem Mario do lado esquerdo da tela para o lado direito de modo a alcançar o mastro da bandeira no final de cada nível, evitando perigos como inimigos e poços. Não há sequência correta de entradas do controlador; há sequências que levam à vitória e outras que não. Na aprendizagem por reforço, um algoritmo poderia “jogar” por conta própria. Ele tentaria simular diversas entradas de controle diferentes e quando finalmente movesse Mario para frente (sem receber danos), o algoritmo seria “recompensado” (ou seja, o comportamento seria reforçado). Por meio desse processo, o algoritmo começa a aprender qual comportamento é desejado (por exemplo, avançar é melhor do que retroceder, pular sobre os inimigos é melhor do que correr para eles). Eventualmente, o algoritmo aprende como se mover do início ao final da fase do jogo.

3.2 Análise de Performance do Modelo de ML

Para maximizar a chance de generalização para o desempenho do algoritmo em dados fora da amostra, o *dataset* de treinamento é geralmente dividido em um conjunto de treinamento ligeiramente menor e um outro conjunto de validação separado (GARETH et al., 2013)(TIBSHIRANI R, 2009).

As métricas usadas para avaliação de um modelo dependem do próprio modelo e se ele está em fase de treinamento ou teste. O *dataset* de validação tem o objetivo de imitar o *dataset* de teste e ajudar os cientistas de dados a ajustar o algoritmo, identificando quando um modelo pode generalizar bem e funcionar em uma nova população. Devido ao *dataset* de validação ser uma pequena amostra da população verdadeira (maior), ele pode não representar com precisão a própria população devido a um viés de amostragem desconhecido. Portanto, o desempenho e a generalização do modelo não devem ser avaliados por meio do desempenho do conjunto de validação. É concebível que um cientista de dados possa criar um *dataset* de validação com um viés desconhecido e use-o para ajustar um determinado modelo. Embora o modelo possa ter um bom desempenho no *dataset* de validação, provavelmente não teria um bom desempenho em um conjunto de teste muito maior (ou seja, não seria um modelo generalizável).

O desempenho do modelo geralmente é monitorado por meio de alguma forma de precisão nos *datasets* de treinamento e validação durante esta etapa. Desde que a precisão

do modelo no conjunto de treinamento (X%) e no conjunto de validação (Y%) estejam aumentando e convergindo após cada iteração de treinamento, o modelo é considerado como aprendendo. Se ambos convergem, mas a um patamar baixo (por exemplo, X converge em Y em 50%), o modelo não está aprendendo e pode estar com os dados desajustados, ou seja, pode não ter aprendido o suficiente sobre a relação entre recursos e alvos de uma maneira aplicável à outra população. Por fim, se o desempenho do *dataset* de treinamento aumentar muito mais que o desempenho o conjunto de validação (por exemplo, o modelo tem uma precisão de 99% nos dados em que foi treinado, mas apenas 80% nos dados de validação), conclui-se que ele se ajusta muito bem ao *dataset* de treinamento mas se mostra ineficaz para prever novos resultados, este fenômeno é conhecido como sobreajuste (*overfitting*). O sobreajuste implica que o modelo aprendeu recursos específicos para a população do conjunto de dados de treinamento em detrimento da generalização para outra população. Embora o conjunto de dados de validação não seja usado especificamente para treinar o algoritmo, ele é usado para ajustar o algoritmo de forma iterativa. Portanto, o conjunto de dados de validação não é necessariamente um indicador confiável de desempenho do modelo em dados não observados (GARETH et al., 2013)(T TIBSHIRANI R, 2009).

Após a conclusão da fase de treinamento e validação, espera-se obter um modelo treinado altamente generalizável; no entanto, isso deve ser confirmado por meio de um *dataset* de teste separado, que contém dados não observados pelo modelo.

No caso de aprendizagem supervisionada (metodologia de aprendizagem utilizada para este projeto), o desempenho de um modelo treinado pode ser avaliado de diversas maneiras, a mais comum com base na precisão do preditor (classificação) ou métricas de erro e resíduos (regressão)(GARETH et al., 2013)(T TIBSHIRANI R, 2009). Conforme mencionado anteriormente, o *dataset* de teste contém instâncias do *dataset* original que não foram observadas pelo algoritmo durante a fase de treinamento. Se a capacidade de predição de um modelo é forte no conjunto de treinamento, mas ineficiente no *dataset* de teste, então o modelo é muito específico para os padrões dos dados de treinamento e é considerado sobreajustado para este conjunto de dados(GARETH et al., 2013)(T TIBSHIRANI R, 2009). O modelo memorizou padrões em vez de aprender um modelo generalizável. Um modelo subajustado, por outro lado, é aquele que tem um desempenho insatisfatório em dados de treinamento e teste e não aprendeu nem memorizou dados deste conjunto, além de ainda não ser generalizável (GARETH et al., 2013)(T TIBSHIRANI R, 2009). Um modelo idealmente ajustado é aquele que tem um forte desempenho em ambos os conjuntos de dados, sugerindo que é generalizável (ou seja, terá um bom desempenho

em *datasets* semelhantes) (GARETH et al., 2013)(T TIBSHIRANI R, 2009).

$$\text{EQM} = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n}, \quad (3.1)$$

$$\text{RQM} = \sum_{i=1}^n \frac{(y_i - \hat{y}_{i-H_t})^2}{n}, \quad (3.2)$$

$$\sigma = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n}. \quad (3.3)$$

Modelos que buscam prever preços de ativos verificam seu desempenho calculando erros de predição. O erro quadrático médio (EQM) (3.1) ou *Mean Squared Error* é um indicador de performance que mede o quão próximo um valor predito está do valor real correspondente, elevando os resultados ao quadrado e dividindo pelo número total de instâncias do *dataset*. O retorno quadrático médio (RQM) (3.3) é uma métrica que quantifica o retorno financeiro a ser obtido H_t dias a frente. O desvio padrão (σ) no contexto de renda variável permite observar a volatilidade do ativo, ou seja o quanto ele está disperso em relação a média. Estas são apenas algumas dentre diversas outras medidas de desempenho para modelos de regressão que estão fora do escopo deste trabalho.

Sendo que nas equações (3.1) e (3.3) y_i corresponde ao valor real da i -ésima característica predita, \hat{y}_i é o valor predito da i -ésima característica, n o total de instâncias (dias) do *dataset*, H_t a janela de tempo em que se deseja observar o retorno e \bar{y} o valor médio da característica.

3.3 Redes Neurais Artificiais

As redes neurais artificiais (*Artificial Neural Network* - ANNs) ou redes neurais simuladas (*Simulated Neural Networks* - SNNs) são um campo da IA concebido com o intuito de mimetizar o sistema neurológico humano. Esta técnica de ML permite que programas de computador reconheçam padrões e resolvam problemas comuns nas áreas de IA, e é considerada como o coração dos algoritmos de *Deep Learning*. ANNs são compostas por nós interconectados representando neurônios do cérebro humano, que recebem sinais, os transformam e transmitem informação para outros neurônios, análogo a axônios e dendritos.

Cada nó possui um peso e limite associado de maneira que se a saída de qualquer nó individual estiver acima do valor limite especificado, esse nó será ativado, enviando dados para a próxima camada da rede. Caso contrário, o neurônio é retirado da coleta de dados de treinamento, portanto, nunca é ativado durante o treinamento se tornando um neurônio considerado “morto”. Os neurônios mortos então não conseguirão atualizar seus pesos, pois as derivadas para os pesos de cada neurônio morto serão muito pequenas ou nulas e os erros não se propagarão, afetando assim outros neurônios na rede. As redes

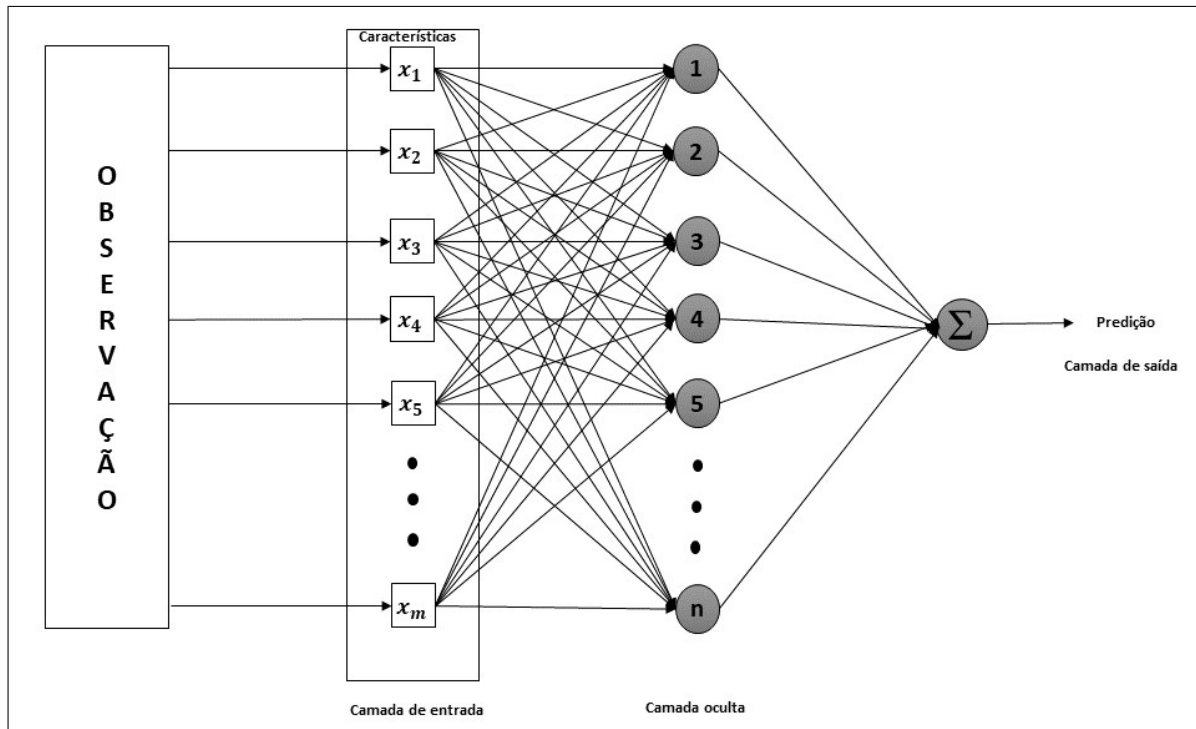


Figura 3.2 – Modelo de rede neural (ANN). Adaptado de (HENRIQUE, 2018)

neurais são orientadas a dados de treinamento para aprender e melhorar sua acurácia ao longo do tempo. Uma vez que esses algoritmos de aprendizagem são ajustados, eles são ferramentas poderosas, permitindo classificar e agrupar dados em alta velocidade.

Uma ANN é modelada como uma função matemática que recebe diversas entradas, as variáveis de predição $x_1, x_2, x_3, \dots, x_m$, e as utiliza para gerar uma saída y , que representa a classificação da observação ou instância x , como no trabalho de (LABOISSIERE; FERNANDES; LAGE, 2015). Uma rede desses modelos de neurônios recebe o nome de ANN, composta de uma camada de entrada, uma ou mais camadas chamadas ocultas e a camada de saída, com a classificação ou valor desejado. Cada camada é composta por um determinado número de modelos de neurônios, como ilustrado pela Figura 3.2 (HENRIQUE, 2018).

3.3.1 Camada de entrada

A camada de entrada é constituída de neurônios artificiais que desempenham um papel diferente. Os neurônios desta camada não recebem informações das camadas anteriores visto que formam a primeira camada da rede. A camada de entrada envia os dados para as camadas subsequentes, nas quais os neurônios têm entradas ponderadas por pesos (*weights*) na camada oculta. Também define as condições de entrada que ditam o treinamento da rede (KARSOLIYA, 2012)(ASTHANA; BAHRDWAJ, 2017) (GUPTA; K.RAZA, 2020).

3.3.2 Camada oculta

A camada oculta é uma camada intermediária entre a entrada e a saída da rede. Pode ser representada como conjuntos de neurônios que transferem dados e padrões extraídos do treinamento para outras camadas da rede por meio de um processo conhecido como ativação dos neurônios (KARSOLIYA, 2012) (XUA; L.CHEN, 2008). O número de camadas ocultas na rede varia de acordo com a modelagem implementada para resolução do problema.

A função de ativação φ utiliza as entradas x_1, \dots, x_m ponderadas pelos pesos w_{1j}, \dots, w_{mj} na camada oculta número j e as combina em um modelo de regressão linear, no qual

$$h_j = \sum_{i=1}^m w_{ij} * x_i \quad (3.4)$$

Em que h_j corresponde a regressão linear dos neurônios da j -ésima camada.

Pense em cada nó individual como seu próprio modelo de regressão linear, composto de dados de entrada, pesos, um viés (ou limite - *bias*) e uma saída, dada por

$$o_j = \varphi(h_j + b_j) \quad (3.5)$$

No qual o_j corresponde a ativação não-linear dos neurônios da j -ésima camada. A Figura 3.3 ilustra o processo completo de ativação de uma camada.

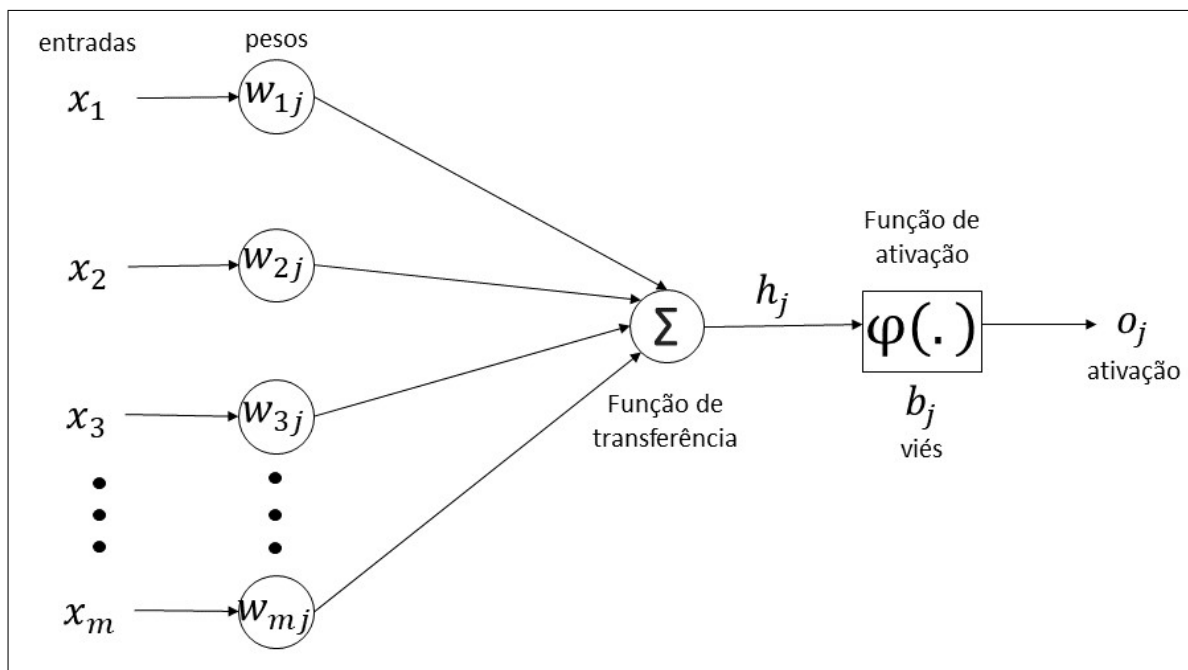


Figura 3.3 – Ativação de um neurônio. Fonte: dos autores.

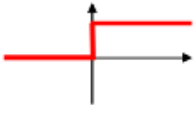
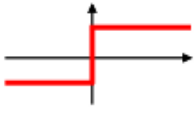
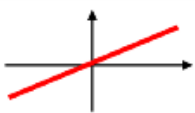

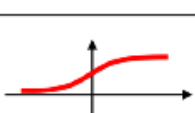

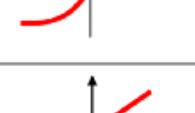

3.3.3 Camada de saída

A camada de saída recebe os padrões processados pela camada oculta e os utiliza para classificar ou prever os valores e configurar tomadas de decisão. Representa a saída correspondente a entrada fornecida à rede. Os neurônios nas camadas de saída têm relação direta com a complexidade do problema ou com o trabalho dado à rede (KARSOLIYA, 2012). Os neurônios da camada de saída são projetados de forma diferente das outras camadas, seu foco consiste em agilizar e melhorar os resultados finais do processo iterativo.

3.4 Deep Learning, RNN e LSTM

Deep Learning ou Aprendizagem Profunda é uma subárea de ML que utiliza estrutura de grafos responsável por modelagem matemática de alto nível utilizando várias camadas de processamento não lineares. A aplicação de arquiteturas DL permite a construção de sistemas cada vez mais inteligentes e autônomos, eliminando parte da intervenção humana necessária. Os avanços tecnológicos consistem em: visão computacional; reconhecimento de fala; processamento de linguagem natural; e reconhecimento de áudio.

No contexto de Redes Neurais aplicadas a DL, as Redes Neurais Profundas, como são conhecidas, utilizam diversas camadas ocultas que possuem função de ativação não linear para extrair recursos dos dados da rede, formando uma complexa camada de processamento, chamada Camada Densa. A Figura 3.4 lista algumas das funções de ativação mais utilizadas em algoritmos ML.

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = \max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^z)$	Multi-layer Neural Networks	

Copyright © Sebastian Raschka 2016
(<http://sebastianraschka.com>)

Figura 3.4 – Lista de funções de ativação. Fonte: [Sebastian Raschka](#)

Recurrent Neural Network - RNN ou Redes Neurais Recorrentes é um tipo de rede neural artificial especializada em aprendizagem de dados sequenciais que retroalimentam a rede a partir de uma estrutura cíclica, o aprendizado anterior se reflete no aprendizado atual. Resolve as limitações do aprendizado de dados contínuo, repetitivo e sequencial existente estabelecendo uma relação de tempo entre os períodos de aprendizagem. A Figura 3.5 ilustra o algoritmo RNN.

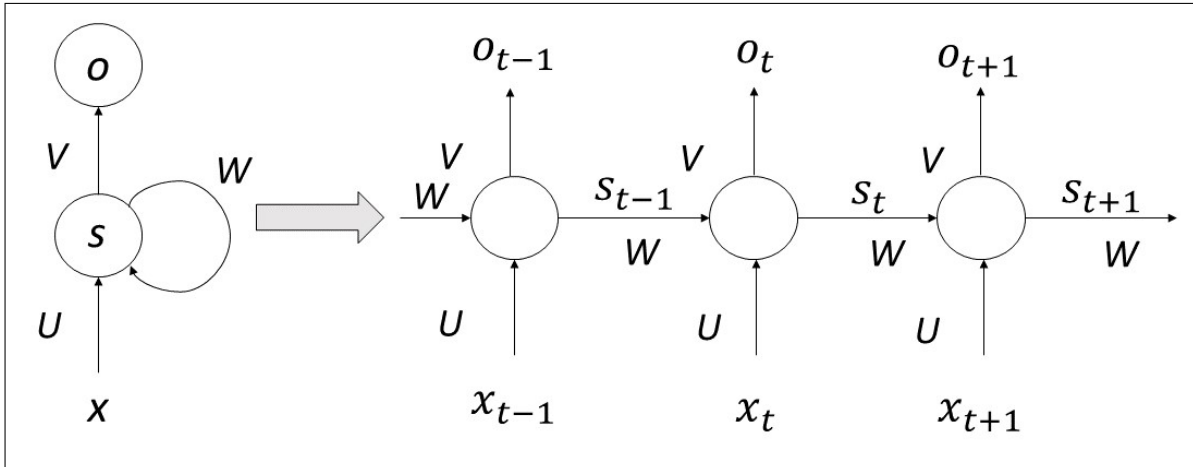


Figura 3.5 – Estrutura básica de uma RNN. Fonte: (ZHU ZHILE YANG, 2019)

Para o tempo t :

$$s_t = \varphi(Ux_t + Ws_{t-1} + b_1) \quad (3.6)$$

$$o_t = \varphi(Vs_t + b_2) \quad (3.7)$$

$$\hat{y}_i = \varphi(o_t) \quad (3.8)$$

Em que x_t , s_t e o_t , denotam unidade de entrada, unidade oculta e de saída em t , respectivamente. Os pesos das conexões de rede são denotados por V , W , U . Além disso, b e \hat{y} representam o viés e o valor de saída previsto, e φ denota a função de ativação.

A desvantagem das RNN dá-se pelos valores dos pesos das camadas iniciais, que diminuem gradualmente à medida que são aprendidos pela rede, devido ao grande número de camadas ocultas e a retropropagação que ocorre em cada ciclo de aprendizagem, este fator acaba degradando a capacidade de aprendizagem em aplicações com conjuntos de dados com camadas densas, problema conhecido como *Gradient Vanishing*. Uma arquitetura que busca resolver isto é a rede *Long Short-Term Memory Network* (LSTM). Ela adiciona outra camada de recurso, chamada estado da célula, para decidir se deve manter pesos associados a memórias de longo prazo, o que contorna o problema do *Gradient Vanishing*.

Long Short Term Memory Network (LSTM)

A estrutura LSTM mantém a estrutura lógica da rede recorrente. A principal diferença entre LSTM e RNN é que o método LSTM adiciona um “processador” ao algoritmo para determinar se a informação de entrada é útil ou não. O processador é denominado “célula” e inclui todos os recursos dos módulos LSTM. Conforme mostrado na Figura 3.9, três portas são projetadas na célula denominadas porta de entrada (i_t), porta de esquecimento (f_t) e porta de saída (o_t), respectivamente, para manter e atualizar informações

valiosas dos dados antes do tempo t . O método de treinamento do modelo para LSTM é o amplamente adotado *back-propagation through time (BPTT)* (WERBOS, 1990).

Estado Oculto (*Hidden State*)

O estado oculto é responsável por memorizar os padrões da rede LSTM, ou seja a modelagem das características que mais se aproxima do resultado desejado pela rede em cada intervalo de tempo. Supondo um trabalho de Processamento de Linguagem Natural (onde busca-se reconhecer e prever corretamente padrões em linguagens que utilizem símbolos como as letras de um alfabeto). Ao processar a frase “o céu é azul, portanto o elefante bebê está chorando”, por exemplo. Se o desejo for que a rede LSTM seja capaz de classificar o sentimento de uma palavra no contexto da frase, o estado oculto em $t = 3$ seria uma versão codificada do caractere “é”, que então seria processado posteriormente (por um mecanismo fora da rede LSTM) de modo a se obter o sentimento previsto. Se o objetivo for que a rede LSTM seja capaz de prever a próxima palavra com base na série atual de palavras, o estado oculto em $t = 3$ seria uma versão codificada da previsão para a próxima palavra (idealmente, “azul”), que seria processado novamente fora da rede LSTM a fim de obter a palavra prevista. Os padrões extraídos do estado oculto diferem conforme o projeto da rede LSTM. É tão flexível porque, em última análise, a forma que se projeta a LSTM dita a maneira de treino e o tipo de dado utilizado; os pesos serão ajustados e orientados da melhor forma a se aproximar da resposta esperada. O estado oculto não é igual à saída ou previsão da rede, é apenas uma codificação dos padrões extraídos no intervalo de tempo mais recente que pode ser processado junto as outras entradas de forma a se obter novos dados significativos para o funcionamento da rede nos próximos instantes de tempo. É denotado por h_t (INFOLKS, 2019).

Estado ou Memória da Célula (*Cell State*)

O estado da célula, preocupa-se com todos os dados processados até o momento (ou seja as características de entrada no instante de tempo atual e comportamento da rede). Se você estiver processando agora a palavra "elefante", o estado da célula contém informações de todas as palavras desde o início da frase. É possível notar ao observar a Figura 3.9 que cada vez que os dados de um determinado intervalo de tempo passam por uma célula LSTM, uma cópia destes dados é filtrada por uma porta de esquecimento e outra cópia pela porta de entrada; o resultado de ambas as portas é incorporado ao estado da célula a partir do processamento do intervalo de tempo anterior e é passado adiante para ser modificado pelo próximo intervalo de tempo. Os pesos na porta de esquecimento e na porta de entrada são responsáveis por extrair recursos de tais informações, pois estes determinam quais intervalos de tempo são importantes (pesos de esquecimento altos), quais não são (pesos de esquecimento baixos) e como codificar informações do intervalo de tempo atual para o estado da célula (pesos de entrada). Como resultado, nem todos os intervalos de tempo são incorporados igualmente ao estado da célula, alguns são mais

significativos do que outros. Isso é o que dá aos LSTMs sua capacidade característica de decidir dinamicamente até onde voltar na história ao trabalhar com dados de séries temporais. De forma resumida, o estado da célula é basicamente a memória global ou agregada da rede LSTM em todos os instantes de tempo. É denotado por C_t (INFOLKS, 2019).

Porta de Esquecimento (*Forget Gate*)

Existem algumas informações do estado anterior da célula LSTM que não são necessárias para a unidade presente em uma LSTM. A porta de esquecimento é responsável por remover essas informações do estado da célula. As informações que não são mais necessárias para a LSTM entender ou as informações de menor importância são removidas por meio da multiplicação por um filtro σ . Utilizado para otimizar o desempenho da rede LSTM. Em outras palavras, pode-se dizer que ele determina quanto do estado anterior deve ser passado para o próximo estado. A porta tem duas entradas x_t e h_{t-1} . h_{t-1} representa a saída da célula oculta em $t - 1$ e x_t é a entrada no intervalo de tempo t . As entradas fornecidas são multiplicadas pelas matrizes de peso e um viés é adicionado. A porta de esquecimento (f_t) lê as informações em h_{t-1} e x_t , em seguida, produz um valor entre 0 e 1 para a célula de estado C_{t-1} , 1 denota “completamente reservado” e 0 representa “completamente descartado” (INFOLKS, 2019). A Figura 3.6 ilustra a porta de esquecimento de uma LSTM.

Porta de Entrada (*Input Gate*)

A porta de entrada i_t decide a quantidade de informação nova a ser adicionada à célula de estado. A combinação de x_t e h_{t-1} passa por uma função *sigmoid* do filtro σ que decide quais informações serão atualizadas. Em seguida, a combinação de x_t e h_{t-1} passa por uma camada *tanh*, a qual gera um vetor contendo todos os valores possíveis que podem ser adicionados a célula, denotado por \tilde{C}_t . Essa porta permite garantir que apenas informações importantes e não redundantes sejam adicionadas ao estado da célula. A memória da célula atual C_t é gerada pelo acúmulo dos parâmetros da célula anterior e da porta de entrada. (INFOLKS, 2019) A Figura 3.7 ilustra a porta de entrada de uma LSTM.

Porta de Saída (*Output Gate*)

Um vetor é criado após a aplicação da função *tanh* ao estado da célula C_t . Este vetor é combinado com a saída de um filtro σ que utiliza os valores de h_{t-1} e x_t como entrada, de modo que possa regular os valores que precisam ser produzidos a partir do vetor criado acima. Este filtro novamente emprega uma função *sigmoid*. Em seguida, ambos são multiplicados para formar a saída de estado da célula h_t (INFOLKS, 2019). A Figura 3.8 ilustra a porta de saída de uma LSTM.

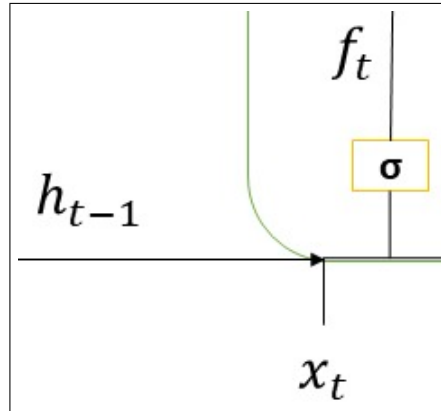


Figura 3.6 – Porta de esquecimento de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)

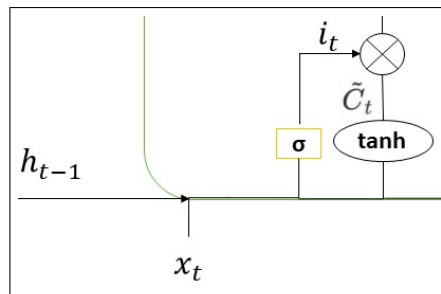


Figura 3.7 – Porta de entrada de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)

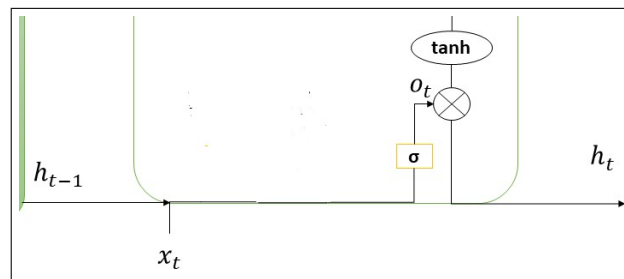


Figura 3.8 – Porta de saída de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)

O esquema de atualização dos estados e parâmetros da célula é mostrado a seguir:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \quad (3.9)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (3.10)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \quad (3.11)$$

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1}, \quad (3.12)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad (3.13)$$

$$h_t = o_t * \tanh(C_t), \quad (3.14)$$

Em que i_t , f_t e o_t correspondem às portas de entrada, esquecimento e saída, respectivamente. h_{t-1} é a saída no instante $t - 1$, e x_t é a entrada no momento atual, e C_{t-1} é a memória da célula anterior. W e b representam, respectivamente, a matriz de pesos e o viés, σ e \tanh são funções de ativação enunciadas pelas equações (3.15) e (3.16).

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \quad (3.15)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (3.16)$$

A Figura 3.9 Ilustra a arquitetura completa de uma célula LSTM.

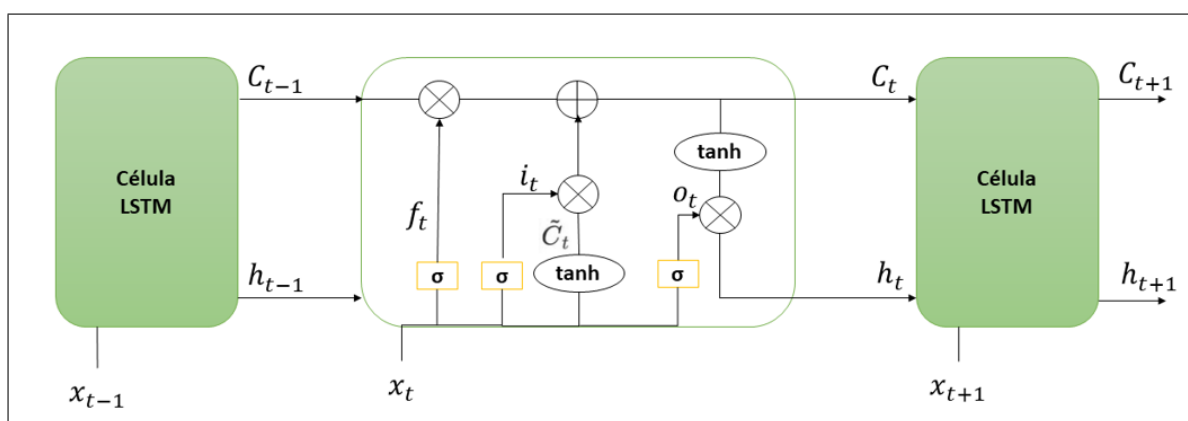


Figura 3.9 – Estrutura de uma rede LSTM. Adaptado de: (ZHU ZHILE YANG, 2019)

4 Machine Learning em Sistemas de Trading

Neste capítulo é apresentada a concepção de um sistema de *trading* utilizando técnicas de *Machine Learning*, acoplado a um alocador de papéis que controla a quantidade de títulos a serem investidos em um determinado ativo. Constam, também, o detalhamento de como foram feitos os módulos deste projeto, suas funcionalidades, a metodologia em que foram desenvolvidos e ferramentas utilizadas para viabilizar os modelos implementados.

4.1 Estrutura do Sistema

O sistema proposto é dividido em cinco objetos de estudo que podem ser acoplados, são eles:

- Plataforma
- Estatísticas financeiras
- *Dataset*
- Alocador de Papéis
- Rede Neural

4.1.1 Plataforma

A utilização de uma plataforma financeira é indispensável para a realização do trabalho proposto por este documento, dessa maneira precisamos da plataforma para que possa ser realizado a coleta de cotações e o envio de ordens (compra e venda, por exemplo) para o mercado. Neste projeto foi utilizada a plataforma *MetaTrader 5* que por sua vez é uma plataforma gratuita e largamente utilizada no mercado financeiro, pois apresenta um excelente conjunto de ferramentas, como a API para automatização e criação de *scripts* para realizar operações no *live market*.

Foi realizado o desenvolvimento de uma simples interface em *python* para interagir com a plataforma *MetaTrader* através do pacote *MetaTrader5* disponível no link [MetaTrader5](#), em que o objetivo do módulo é realizar a extração de cotações e listagem dos ativos disponíveis.

Inicialmente na implementação do módulo é realizada a definição de algumas constantes para representação dos possíveis *timeframes* disponíveis no *MetaTrader*, na sequência são definidas as funções básicas para interação da plataforma como *login* e *shutdown*,

por último a classe também contém métodos para realizar a coleta de cotações referentes aos ativos negociados em amplo mercado.

A função *currency_time_window* é responsável por realizar a coleta da cotação de um ativo específico em uma granularidade e janela de tempo especificada pelo usuário. Para invocar a função temos o uso do parâmetro *timeframe* (por sua vez pertencendo a representação de *timeframes* específicos), esse parâmetro será responsável por definir a granularidade de extração das cotações desta maneira especificando se as cotações serão a cada minuto, hora, dia, semana, mês e etc. Outro parâmetro informado para a função é a janela de tempo que a cotação deve ser obtida, sendo assim informada pelos parâmetros *start_date* e *end_date*. De maneira similar temos a implementação da função *multiple_currency_time_window* que apresenta a mesma funcionalidade que a anterior porém a mesma coleta múltiplos ativos ao mesmo tempo, sendo estes informados em seu vetor de *input*.

Durante a implementação da interface com a plataforma *MetaTrader 5* foi possível notar que existe uma limitação de extração de dados para janelas de tempo muito longas, logo a função aqui implementada tem o intuito de tornar transparente para a aplicação a extração de dados financeiros da plataforma, pois nesta função temos a lógica que realiza a concatenação e ordenação das cotações para qualquer janela de tempo informada.

4.1.2 Estatísticas financeiras

O módulo de estatísticas financeiras foi criado para realizar a implementação da teoria vista na subseção 2.3.5, sendo assim temos aqui o cálculo de alguns dos indicadores mais utilizados no mercado financeiro, na análise gráfica. O uso de estatísticas financeiras para a realização do projeto permite que a rede neural tenha uma visão mais analítica do dado bruto das cotações que é extraído pelo módulo 4.1.1, sendo assim possibilita que a mesma possa durante o seu treinamento realizar a criação de padrões de análise que envolvem diversos indicadores financeiros e seus respectivos valores (por exemplo, média móvel aritmética de 8 períodos está acima do valor da cotação atual). As funções implementadas neste módulo apenas realizam o cálculo matemático formulado por cada um dos indicadores, sendo assim suas respectivas implementações podem ser vistas como a aplicação da fórmula matemática na cotação histórica dos ativos.

A implantação deste módulo foi feito em *python* com o uso das bibliotecas *numpy* e *pandas*, sendo assim tal implementação ocorre de forma funcional, em que cada função realiza o cálculo de um indicador financeiro que será detalhado abaixo. Todas os indicadores financeiros e suas respectivas funções trabalham dados estruturados em *pandas.Series*.

1. Média Móvel Simples

A média móvel simples (*Simple Moving Average*, em inglês) é um dos indicadores financeiros mais utilizados em análises gráficas. Esse indicador é útil para mostrar a tendência de valor de um determinado ativo sem a interferência de pontuais picos e ruídos que ocorrem no oscilar normal de um ativo. O cálculo do indicador é realizado sob uma janela de tempo desejada em que são tomadas as últimas n -ésimas cotações do ativo. A implementação do cálculo,

$$SMA_n = \sum_{i=0}^n \frac{C_n}{n}, \quad (4.1)$$

em que SMA_n é a n -ésima média Móvel simples e, C_n a cotação do n -ésimo período pode ser encontrada na função `simple_moving_average` do módulo, sendo assim seu único parâmetro é o período de análise.

2. Média Móvel Exponencial

A média móvel exponencial é outro indicador que é bastante útil por evidenciar a tendência de preços de um ativo em uma determinada direção, porém diferentemente da média móvel simples a EMA (sigla do inglês para *Exponential Moving Average*) possui uma maior ponderação para elementos mais recentes dos preços dos ativos, logo mostra pode confirmar de forma mais responsiva a guinada ou declínio no valor de um ativo e assim confirmar uma tendência de forma mais prematura. A função `exponential_moving_average` realiza a implementação do cálculo matemático de maneira análoga e simplificada (pois o fator de decaimento poderia ser um parâmetro) e possui apenas o parâmetro do período de análise. Diferentemente da SMA temos que o cálculo da EMA ocorre de forma recursiva enunciado por,

$$EMA_n = C_n \left(\frac{2}{n+1} \right) + EMA_{n-1} \left(1 - \frac{2}{n+1} \right), \quad (4.2)$$

$$EMA_0 = C_0, \quad (4.3)$$

em que EMA_n é a Média Móvel Exponencial n -ésima e C_n a cotação do n -ésimo período.

3. Desvio Padrão

O desvio padrão é uma das medidas estatísticas mais utilizadas no estudo de sinais financeiros pois permite ao analista realizar uma estimativa acurada de qual é o passeio médio executado pela cotação de um determinado ativo, sendo assim o mesmo tem o intuito de mensurar o que pode ser considerado como uma volatilidade

de preço normal ou *outliner* para um ativo em um determinado horizonte de tempo, logo o mesmo não apresenta qualquer viés de guinada ou declínio do valor do ativo e sim apenas uma indicação de que o ativo está ou não acima de sua volatilidade regular. Esse indicador pode ser implementado de maneira simples através da função *standard_deviation* em que é realizado cálculo estatístico,

$$SD_n = \sqrt{\frac{\sum_{i=0}^n (C_n - \bar{C})^2}{n - 1}}, \quad (4.4)$$

tal que SD_n corresponde ao desvio padrão n-ésimo, C_n a cotação do n-ésimo período e \bar{C} a média da cotação do período. O único parâmetro do método é o período de análise.

4. MACD

De maneira análoga às médias móveis apresentadas o MACD também é um indicador que busca mostrar tendências nos valores do ativo. O mesmo é composto por três EMA de períodos diferentes, sendo elas chamadas de *fast*, *slow* e *MACD*. Esse indicador é útil para se ter uma análise da futura tendência de um ativo, pois é possível notar a convergência ou divergência entre as curvas *fast* e *slow*, o que define, assim, uma tendência de guinada ou declínio do valor do ativo. A linha *MACD* é obtida pela EMA da diferença entre a linha *fast* e *slow*, conforme mostrado na equação (4.7). As implementação do indicador se dá pela função *MACD* que é o resultado da codificação das equações (4.5), (4.6) e (4.7), em que o resultado do valor de retorno da função são as três curvas que compõem o indicador (fato que flexibiliza o uso do retorno da função). A implementação do indicador permite a especificação dos períodos de cada uma das curvas. Enuncia-se então as equações,

$$FAST_f = C_n \left(\frac{2}{f + 1} \right) + EMA_{f-1} \left(1 - \frac{2}{f + 1} \right), \quad (4.5)$$

$$SLOW_s = C_n \left(\frac{2}{s + 1} \right) + EMA_{s-1} \left(1 - \frac{2}{s + 1} \right), \quad (4.6)$$

$$MADC_n = EMA_m (FAST_f - SLOW_s), \quad (4.7)$$

em que $FAST_f$, $SLOW_s$ e C_n correspondem a EMA rápida de período f-ésimo, EMA lenta de período s-ésimo e a cotação do n-ésimo período respectivamente.

5. Bollinger Bands

Esse é um indicador também muito utilizado por analistas do mercado financeiro, sendo composto por três curvas que representam, respectivamente, a SMA do ativo,

banda superior e banda inferior. Esse é um indicador voltado a ilustrar a volatilidade de um ativo em uma dada janela de tempo, desta maneira as bandas proporcionam um intervalo de valores (e conseqüentemente a volatilidade induzida) para o ativo, sendo assim valores fora das bandas são considerados *outliers* e bons pontos de entrada para operações financeiras. O indicador possui a sua banda superior computada como a SMA acrescida de um número k de desvios padrões e a banda inferior como a SMA decrescida de k desvios padrões (geralmente $k = 2$). Para a implementação do método, pode ser transcritos em código as equações (4.9), (4.10) e (4.8). A implementação realizada para a estatística financeira possui o período e o número k de desvios padrões como parâmetro. Têm-se as equações,

$$BBM_n = \sum_{i=0}^n \frac{C_n}{n} = SMA_n, \quad (4.8)$$

$$BBU_n = \sum_{i=0}^n \frac{C_n}{n} + k \sqrt{\frac{\sum_{i=0}^n (C_n - \bar{C})^2}{n-1}} = SMA_n + k \cdot SD_n, \quad (4.9)$$

$$BBD_n = \sum_{i=0}^n \frac{C_n}{n} - k \sqrt{\frac{\sum_{i=0}^n (C_n - \bar{C})^2}{n-1}} = SMA_n - k \cdot SD_n, \quad (4.10)$$

$$(4.11)$$

nas quais BBM , BBU e BBD são, respectivamente: banda de Bollinger média de período f -ésimo, banda de Bollinger superior de período f -ésimo e a banda de Bollinger inferior de período f -ésimo. K e C_n correspondem ao número de desvios padrao para ser acrescido e removido e a cotação do n -ésimo período.

6. Relative Strength Index

O RSI (sigla do inglês para *Relative Strength Index*) é um indicador financeiro, que compõe a categoria dos osciladores, muito utilizado para se avaliar as condições do mercado, como a alteração do preço de um ativo ao longo de uma janela de tempo específica e as posições tomadas pelo mercado. Este indicador apresenta valores que oscilam de 0 a 100 e, assim, pode se considerar que um determinado ativo se encontra sobrecomprado caso o RSI seja maior que 70 e sobrevendido quando o mesmo se possui um valor abaixo de 20. Para realizar a implementação do RSI é necessário em um primeiro momento calcular a razão entre o EMA dos ganhos e perdas para os períodos de uma dada janela de tempo, logo após é utilizada a fórmula (4.15) para realizar o cálculo do indicador. A codificação realizada permite que se especifique o período dentro da janela de tempo. Para o cálculo dos ganhos e perdas são utilizadas as formulas mostradas nas equações (4.12) e (4.13).

$$G = \begin{cases} \frac{C_n}{C_{n-1}} & \text{para } \frac{C_n}{C_{n-1}} \geq 0 \\ 0 & \text{para } \frac{C_n}{C_{n-1}} < 0 \end{cases}, \quad (4.12)$$

$$L = \begin{cases} \frac{C_n}{C_{n-1}} & \text{para } \frac{C_n}{C_{n-1}} < 0 \\ 0 & \text{para } \frac{C_n}{C_{n-1}} \geq 0 \end{cases}, \quad (4.13)$$

$$RS_n = \frac{G_n \left(\frac{2}{n+1} \right) + EMA_{n-1} \left(1 - \frac{2}{n+1} \right)}{L_n \left(\frac{2}{n+1} \right) + EMA_{n-1} \left(1 - \frac{2}{n+1} \right)} = \frac{EMA_n(G)}{EMA_n(L)}, \quad (4.14)$$

$$RSI_n = 100 - \frac{100}{1 + RS_n}, \quad (4.15)$$

$$(4.16)$$

tal que G , L , G_n , L_n correspondem respectivamente ao vetor com os ganhos da janela temporal, o vetor com as perdas da janela temporal, ganho n -ésimo e a perda n -ésima. C_n , C_{n-1} e RS_n condizem a cotação do n -ésimo período, cotação anterior ao n -ésimo período e ao vetor de força relativa, respectivamente.

7. On Balance Volume

Esse é um indicador diferente dos anteriores mas igualmente muito utilizado para análise técnica, pois esta ferramenta tem o intuito de evidenciar o rumo que o volume de um ativo está tomando. Sendo assim, o mesmo é acumulativo e independe de janela temporal. O OBV possui a característica de relacionar o preço de um ativo com o seu volume, logo é possível de se identificar quando temos uma tendência prestes a perder força e conseqüentemente inverter o movimento principal.

A implementação do OBV consiste basicamente de um acumulador, é realizada a diferença entre o fechamento de um período n com o período $n - 1$ em que, caso o resultado seja positivo, é acrescido no acumulador o volume do período n e, caso contrário, é retirado o volume do período do acumulador. Para ilustrar de forma mais clara temos a fórmula (4.17) que mostra o cálculo realizado. As fórmulas estão enunciadas de maneira que,

$$OBV_n = \begin{cases} OBV_{n-1} + V_n & \text{para } C_n - C_{n-1} > 0 \\ OBV_{n-1} & \text{para } C_n - C_{n-1} = 0 \\ OBV_{n-1} - V_n & \text{para } C_n - C_{n-1} < 0 \end{cases}, \quad (4.17)$$

sendo que C_n , C_{n-1} e V_n são, respectivamente, a cotação do n -ésimo período, cotação anterior ao n -ésimo período e o volume do n -ésimo período.

8. Oscilador Estocástico

O oscilador estocástico é um indicador bastante utilizado para identificar momentos em que o ativo passa por uma reversão de tendência. O indicador apresenta valores que oscilam de 0 a 100 e, assim como o RSI, pode indicar se um ativo se encontra sobrecomprado ou sobrevendido baseado em limiares de análise (usualmente mais de 80 para sobrecomprado e abaixo de 20 para sobrevendido). O estocástico é composto pela linha k e pela linha sinal denotada por d .

Para realizar o cálculo do indicador primeiramente é necessário se localizar o valor máximo e mínimo do ativo em uma janela temporal, tal operação pode ser feita usando a codificação das equações (4.18) e (4.19), logo após é utilizado a fórmula (4.20) para realizar o cálculo final dos termos que compõem a linha k do indicador. A linha d do indicador pode ser encontrada realizado uma SMA dos termos presentes na linha k , conforme mostrado na equação (4.21). Enunciamos as equações (4.18), (4.19), (4.20), (4.21) como

$$C_{max} = \text{Max}(C_{n-j}, C_{n-j+1}, \dots, C_{n-1}, C_n), \quad (4.18)$$

$$C_{min} = \text{Min}(C_{n-j}, C_{n-j+1}, \dots, C_{n-1}, C_n), \quad (4.19)$$

$$\text{STOK}_n = \left(\frac{C_n - C_{min}}{C_{max} - C_{min}} \right), \quad (4.20)$$

$$\text{STOD}_n = \sum_{i=0}^n \frac{\text{STOK}_n}{n} = \text{SMA}_n(\text{STOK}_n), \quad (4.21)$$

em que C_n , C_{n-j} , C_{max} e C_{min} correspondem a cotação do n -ésimo período, cotação de j períodos anteriores ao n -ésimo período, cotação máxima dentro da janela temporal e a cotação mínima dentro da janela temporal respectivamente.

9. Ichimoku

Esse é o indicador mais complexo implementado no projeto descrito neste documento. É um indicador que busca identificar qual a tendência de nivelamento dos preços em um mercado futuro assim como também realiza a identificação de suportes e resistências ao longo da janela temporal de observação da oscilação de um ativo. O indicador é composto por 5 linhas que estão distribuídas no passado, presente e futuro do valor atual do ativo. A linha do passado é denominada *Chinkou Span*, as linhas do presente são chamadas de *Kijun-sen* e *Tenkan-sen* enquanto que as linhas que realizam o reflexo do ativo no futuro são chamadas de *Senkou Span A* e *Senkou Span B*.

As linhas que estão no “presente” são basicamente compostas por SMA de períodos distintos (em que a linha *Kijun-sen* é a que possui um horizonte mais longo), sendo

assim caso o preço do ativo venha a realizar o cruzamento com a linha *Kijun-sen*, pode representar o início de uma inversão de mercado, além de tais linhas funcionarem como uma espécie de suporte e resistência para o ativo. Já as linhas do “futuro” visam realizar a formação de nuvens de *Ichimoku* que podem ser de guinada ou declínio do preço do ativo, logo as mesmas funcionam como uma forma de indicar a força do movimento de acordo com a sua largura.

Para realizar a implementação do indicador *Ichimoku* inicialmente começamos com a linha *Chinkou Span* que é basicamente o reflexo do preço do ativo no passado, logo se trata da repetição da cotação atrasada de j períodos. Continuando na codificação do indicador podemos realizar a implementação das linhas *Kijun-sen* e *Tenkan-sen* com o uso de SMAs. Ambas as linhas são implementadas da mesma maneira, porém apenas diferindo nos períodos adotados em seus cálculos. De maneira análoga as linhas *Senkou-span-a* e *Senkou Span B* são implementadas usando SMA porém deslocadas a frente para realizar a formação das nuvens detalhadas no parágrafo anterior.

$$Chinkou-span_n = C_{n-j}, \quad (4.22)$$

$$Kijun-sen_n = \sum_{i=0}^j \frac{C_{n-j}}{j} = SMA_j, \quad (4.23)$$

$$Tenkan-sen_n = \sum_{i=0}^n \frac{C_{n-k}}{n} = SMA_k, \quad (4.24)$$

$$Senkou-span-a_{n-j} = \sum_{i=0}^j \frac{C_{n-j}}{j} = SMA_j, \quad (4.25)$$

$$Senkou-span-b_{n-j} = \sum_{i=0}^k \frac{C_{n-k}}{k} = SMA_k, \quad (4.26)$$

sendo que C_n , C_{n-j} , j e k correspondem a cotação do n -ésimo período, cotação de j períodos anteriores ao n -ésimo período, período longo de observação e período curto de observação.

4.1.3 Dataset

A construção de um *dataset* se faz necessário para alcance do objetivo a ser trabalhado. O uso de *dataset* será feito de forma extensiva durante toda a implementação do projeto a partir da extração de dados da plataforma *MetaTrader 5*. O *dataset* será responsável por realizar a formatação dos dados conforme molde da Figura 4.1 e será modificado com a agregação das estatísticas financeiras mencionadas na seção para realizar o treino da rede neural que é mostrada em mais detalhes na seção 4.1.5. Esse módulo foi criado de

forma a receber qualquer tipo de dados formatado em forma de um *dataframe* parametrizado. O escopo deste projeto será focado no uso dos dados já formatados e provenientes da plataforma *MetaTrader 5*.

	date	open	high	low	close	vol
0	18191.0	16.94	17.14	16.77	17.04	345200.0
1	18192.0	17.02	17.12	16.87	16.97	376100.0
2	18193.0	16.96	16.98	16.78	16.92	332400.0
3	18194.0	16.88	17.03	16.74	16.94	225500.0
4	18197.0	17.05	17.44	17.04	17.13	568600.0
..
449	18856.0	14.98	14.98	14.28	14.59	1486400.0
450	18857.0	14.61	15.24	14.54	14.85	1463200.0
451	18858.0	14.64	15.06	14.55	15.03	740700.0
452	18859.0	14.97	15.20	14.87	15.10	596300.0
453	18862.0	15.24	15.51	15.12	15.34	786600.0

[454 rows x 6 columns]

Figura 4.1 – Dataset - Módulo Rede Neural. Fonte: dos autores

A implementação deste módulo é feita através de uma classe *python* e com o uso das bibliotecas *pandas* e *numpy*. A criação de uma instância desta classe precisa apenas do conjunto de dados iniciais formatado em um *dataframe*.

A classe contém um conjunto de funções úteis que visam realizar a formatação dos dados, importação e exportação dos dados para arquivos, assim temos a função *append_column* e *remove_column* que são responsáveis por adicionar e remover, respectivamente, colunas do *dataframe* que compõe a única propriedade pública da classe, tais funções trabalham para que a estrutura seja sempre mantida, logo a adição de colunas ocorre apenas ao final da estrutura e a remoção de uma coluna intermediária realiza a reorganização das demais colunas.

As funções de importação e exportação de dados trabalham com arquivos em formatação *csv* e *xlsx* (formato muito utilizado no *Microsoft Excel*). Tem-se o conjunto de funções *save_to_excel*, *save_to_csv*, *read_from_excel* e *read_from_csv*. Outro ponto que vale a pena enfatizar é que as funções descritas anteriormente foram implementadas com a filosofia de corrigir quaisquer erros, para assim poder realizar a transferência do *dataset* de forma segura.

Neste método temos implementadas funções voltadas para a formatação dos dados de maneira adequada para realizar o treinamento da rede neural descrita em 4.1.5. Temos as funções *window_samples* e *window_rolling* que realizam a criação de uma janela deslizando sobre o *dataset* original, desta maneira é provida para a rede neural uma série de

“fatias” da base de dados em ordem cronológica. A diferença de ambas as funções é o fato de que a *window_rolling* considera uma nova “fatia” de dados para cada índice enquanto que a função *window_samples* apenas considera termos espaçados de k elementos (em que k é o tamanho da janela informada para a função).

$$BC(x) = \begin{cases} -1 & \text{para } x \in [-\infty, 0) \\ 0 & \text{para } x = 0 \\ 1 & \text{para } x \in (0, +\infty] \end{cases} \quad (4.27)$$

As últimas funções do escopo desta classe são voltadas para a realização de comparações entre as colunas de dados e uma coluna referência, é provido o método *create_binary_comparation* que realiza uma comparação de magnitude entre os elementos de colunas do *dataset* e os discrimina conforma mostrado na equação (4.27), sendo que x é dado pela diferença entre a coluna referência e a coluna comparada no momento. Uma característica desta função é que a mesma concatena os seus resultados no *dataset* original, logo cria mais colunas com o uso da função *append_column*. A função *create_percentage_comparation*, possui um comportamento semelhante com a ultima função detalhada, porém neste momento é utilizada uma comparação percentual entre a coluna referência e as demais, desta maneira o resultado é igualmente concatenado com o restante do *dataset* original.

As funções de comparação presentes neste módulo são de suma importância para a criação de colunas de dados adicionais no *dataset*, o uso dessas colunas adicionais no treinamento de redes neurais permite que tenhamos uma convergência da mesma de forma mais concisa e com uma menor quantidade de dados, visto que as colunas adicionais são formadas por comparativos e assim permitem que a rede neural tenha em sua *input* dados já processados que podem revelar a correlação das colunas presentes no *dataset*.

4.1.4 Alocador

A alocação do capital que será utilizado nas operações financeiras é uma dos módulos principais do projeto, sendo assim o mesmo é o responsável por gerir a quantidade a ser investida em cada ativo financeiro. O sistema de alocação de capital busca realizar a distribuição de capital da forma mais rentável possível, pois o objetivo do mesmo é definir o melhor portfólio possível para um fixo patamar de risco que o investidor considera aceitável ou mesmo definir o melhor portfólio possível em uma relação de retorno por risco.

A abordagem executada durante a realização do projeto foi utilizada a Teoria de Markowitz em que é realizada a busca das fronteiras eficientes de investimento do capital para um dado conjunto de ativos financeiros, em que é levado em consideração o risco

e retornos envolvidos na alocação do capital no conjunto dos ativos e não de maneira isolada e individual.

A implementação da Teoria de Markowitz foi realizada com a utilização das bibliotecas *pandas*, *numpy* e *scipy* e com a utilização do módulo de indicadores financeiros mostrados na seção 4.1.3, desta maneira a implementação consiste em uma classe *python* que permite assim a instanciação de objetos com passagens de parâmetros.

A alocação adotada no projeto é realizada de forma dinâmica para uma janela de tempo selecionada sendo este um parâmetro para realizar a instância da classe. A alocação baseada em janelas temporais é uma das características da implementação da Teoria de Markowitz, pois a classe pode realizar em janelas a alocação do capital visando reajustar o portfólio de acordo com os retornos e riscos envolvidos no histórico recente de um ativo.

Inicialmente temos a elaboração do portfólio com base na série histórica de um conjunto de ativos informado, desta maneira é executado o método *portfolio_definition* que computa o retorno e risco estimado histórico de cada ativo na janela de tempo informada no momento de criação do objeto. O retorno e risco são definidos como

$$RET_{x_i}(X) = \log\left(\frac{X_i}{X_{i-1}}\right) \quad \text{e} \quad RISK_{x_i}(X) = STD(X)$$

sendo que X_i é a série histórica do i -ésimo ativo formatada conforme mostrado na subseção 4.1.3.

Com a definição do portfólio é possível se utilizar da função *optimal_portfolio* para que seja encontrado a distribuição de carteira que apresenta o melhor *sharpe* e assim traga a melhor relação retorno/risco. O *sharpe* é definido pela razão entre a média da diferença do retorno do portfólio R_p e o retorno livre de risco R_f e o desvio padrão do portfólio σ_p .

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \quad (4.28)$$

A função *optimal_portfolio* executa de forma numérica a minimização do risco para uma dada taxa de retorno e com isso é obtido o denominado *market portfolio*. O método numérico escolhido foi o *Sequential Least Squares Programming (SLSQP)* (KRAFT, 1988). Para ter coesão do processo de minimização é utilizado como vetor inicial o vetor mostrado em (4.29) e a restrição mostrada em (4.30), em que n é o número de ativos informados no momento de instância da classe e W_i é o peso percentual de patrimonio

investido no i -ésimo ativo.

$$\left[\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right] \quad (4.29)$$

$$\sum_{i=1}^n W_i = 1 \quad (4.30)$$

$$[K_1, K_2, \dots, K_n] \quad (4.31)$$

$$RET_{x_i}(X) = RISK_{x_i}(X) \quad (4.32)$$

A definição do portfólio permite também a implementação da função *optimal_curve* que realizará o cálculo da fronteira eficiente para a alocação do capital. Inicialmente essa função realiza o cálculo de retorno e riscos para diversas carteiras aleatórias que satisfazem a restrição (4.30), logo para cada carteira K_i teremos que o retorno da mesma é dada pela fórmula (2.1), sendo assim é definido um vetor (4.31) com inúmeras carteiras possíveis, que contém da carteira com menor retorno K_1 até a que possui o maior retorno K_n .

Com o vetor definido a função é incumbida de realizar a minimização do risco para cada elemento do vetor (4.31), utilizando o método numérico *SLSQP* com o vetor inicial mostrado em (4.29) e as restrições mostradas em (4.30) e (4.32), sendo essa segunda condição necessária para podermos encontrar qual o risco $RISK_{x_i}$ que gerou o retorno dado pela carteira K_i , logo será dessa maneira possível realizar a plotagem de uma curva que mostra o menor risco para um dado nível de retorno.

4.1.5 Rede Neural

O módulo consiste em utilizar valores *Open, High, Low, Close* (OHLC) e indicadores financeiros utilizados em análise técnica com o volume de cotações históricas baseadas em títulos negociados na Bolsa de Valores Oficial do Brasil (B3). Visa-se encontrar oportunidades de entrada no mercado a partir da predição de valores de máximo e mínimo. Os valores OHLC e indicadores financeiros do *dataset* parametrizado alimentam uma rede neural LSTM que por sua vez gera as predições, que serão utilizadas para avaliar resultados em duas estratégias de *trading* distintas. A Figura 4.2 ilustra o sistema descrito.

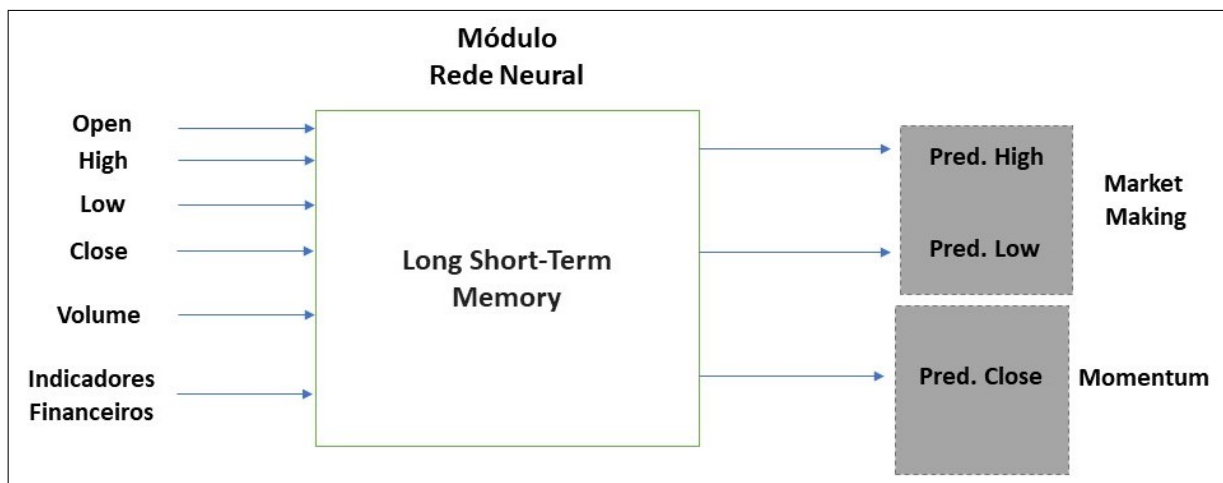


Figura 4.2 – Arquitetura - Módulo Rede Neural. Fonte: dos autores

4.2 Metodologia

Carregando as bases de dados

As bases utilizadas consistem em séries históricas de um determinado período dispostas na escala D1 (composição de um único *candle* por dia) que estão disponíveis publicamente na plataforma *MetaTrader 5*. Os parâmetros são extraídos como um *dataset* com as colunas: *Open*, *High*, *Low*, *Close* e *Volume* de cada data correspondente. Os valores das datas são lidos como valores inteiros a fim de se realizar operações aritméticas com as datas.

Modelo de Treinamento

A estratégia de treinamento definida observa os valores de entrada dentre uma janela de tempo H_T dias anteriores ao dia referencial D e gera como saída uma previsão no horizonte H_P do dia D . A janela de treinamento e o horizonte de previsão deslocam-se a partir de D por toda a série histórica em $[D + H_P]$. Os dados foram normalizados para fins de refinamento de modo a evitar que o modelo seja enviesado por algum parâmetro de entrada.

- Na estratégia de *trading* por *momentum* a previsão gerada corresponde ao preço de fechamento do ativo.
- O *trading* por *market making*, por sua vez, são geradas previsões para o preço de alta e de baixa do ativo.

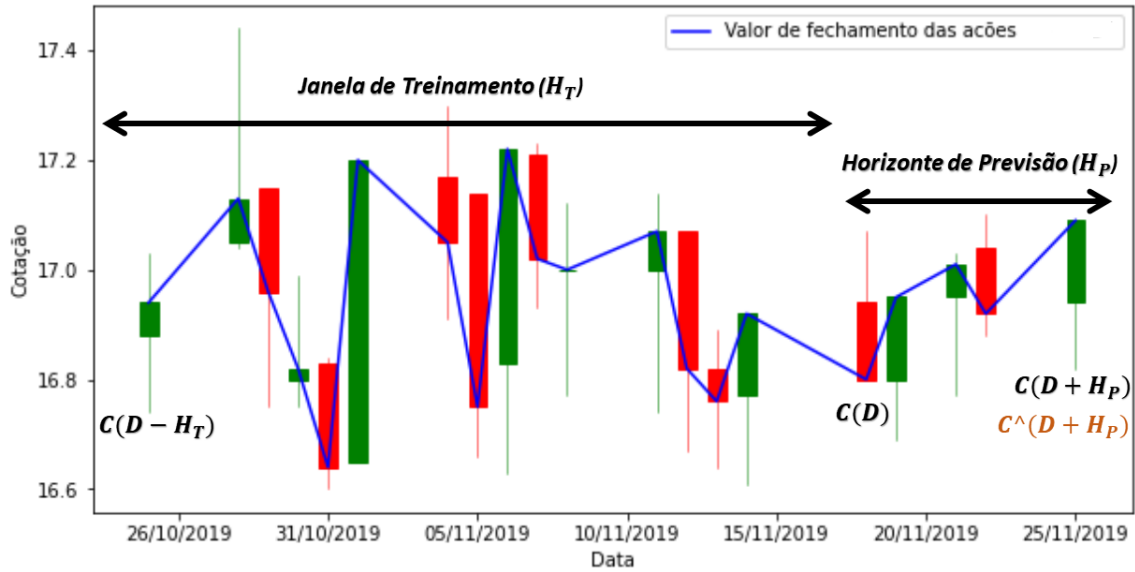


Figura 4.3 – Modelo de treinamento da rede neural. Fonte: dos autores

Rede Neural

O preditor foi projetado como um módulo *Tensorflow* composto de três células LSTM e uma camada densa que processam as entradas no formato de vetor de dimensão correspondente ao número de *features*. O otimizador Adam (KINGMA, 2015) foi utilizado para assegurar maior acurácia as previsões e a função de perda utilizada para medir performance foi o erro quadrático médio que será mostrado mais a frente.

Análise de Performance

Foram adotadas duas métricas principais para avaliar a qualidade da previsão em ambas as estratégias de *trading*, são elas: *erro quadrático médio (EQM)* e *retorno quadrático médio (RQM)*. Em que

$$EQM = \frac{1}{n} \sum_{D=1}^n (V_{Real}(D) - V_{Predito}(D))^2, \quad (4.33)$$

$$RQM = \frac{1}{n} \sum_{D=1}^n (V_{Real}(D - H_P) - V_{Predito}(D))^2, \quad (4.34)$$

Em que D, V correspondem ao dia e valor associado ao parâmetro que foi predito.

A relação entre retorno e erro é dada por:

$$\text{Razão} = \frac{EQM}{RQM}. \quad (4.35)$$

A razão estabelece uma relação entre risco e retorno por erro de predição, quanto maior a Razão significa que mais afastado a predição está do valor real, e quanto menor, maior é o retorno sobre o investimento.

A rentabilidade acumulada é enunciada pela razão entre o valor final e inicial da carteira subtraído por -1 (variação em relação ao total)

$$R_{cum} = \frac{V_{final}}{V_{inicial} - 1}. \quad (4.36)$$

Lógica de Negociação Momentum

Na lógica de negociação por *momentum*, foi definida uma rentabilidade mínima (r_{min}) para determinar o posicionamento do algoritmo no mercado, de modo que,

$$F_{Predito}(D) > F_{Real}(D - H_P) * (1 + r_{min}). \quad (4.37)$$

O que significa que se o preço predito estiver acima da rentabilidade mínima estabelecida pelo algoritmo adota-se uma posição de *compra* (*buy*) do ativo, pois haverá margem para lucrar a rentabilidade mínima no dia D .

A posição de *venda* (*sell*) ocorre quando

$$F_{Predito}(D) < F_{Real}(D - H_P) * (1 - r_{min})$$

que implica que o preço predito prevê queda de r_{min} ou maior em relação ao dia $D - H_P$

$$F_{Predito}(D) < F_{Real}(D - H_P) * (1 - r_{min}) \quad (4.38)$$

Lógica de Negociação Market Making

Na lógica de negociação por *market making*, o erro e retorno quadrático médio são dados, para cada cotação, por:

$$EQM_{HIGH} = \frac{1}{n} \sum_{D=1}^n (H_{Real}(D) - H_{Predito}(D))^2, \quad (4.39)$$

$$EQM_{LOW} = \frac{1}{n} \sum_{D=1}^n (L_{Real}(D) - L_{Predito}(D))^2, \quad (4.40)$$

No qual H representa o preço de alta do dia D e L representa o preço de mínima cotação.

O retorno quadrático médio para *market making* é calculado por

$$RQM_{HIGH} = \frac{1}{n} \sum_{D=1}^n (H_{Real}(D - H_P) - H_{Predito}(D))^2, \quad (4.41)$$

$$RQM_{LOW} = \frac{1}{n} \sum_{D=1}^n (L_{Real}(D - H_P) - L_{Predito}(D))^2. \quad (4.42)$$

Em que H_P é o horizonte de previsão.

A operação de ativos nessa estratégia de negociação utiliza ordens de compra e de venda que são automaticamente executadas assim que o preço do ativo “entra” em um determinado intervalo, seja ele de compra ou de venda. Diferentemente da estratégia de *momentum* onde se realiza uma negociação a partir da observação da tendência do ativo, a qual é obtida por meio da análise de onde o preço estará H_P dias à frente em relação ao valor da cotação atual.

A operação inicia no estado “não posicionado” onde a mesma coloca uma ordem de compra, e espera ela ser executada.

Uma ordem de compra (*buy*) é executada quando:

$$L_{Predito}(D) \geq L_{Real}(D). \quad (4.43)$$

Em seguida, coloca-se uma ordem de venda a fim de resgatar o lucro do valor investido

Aplica-se uma ordem de venda (*sell*) se:

$$H_{Predito}(D) \leq H_{Real}(D). \quad (4.44)$$

Após a ordem de venda, a operação retorna para o estado inicial “não posicionado” e um novo ciclo de operação se inicia.

Nos cálculos de *market making* também se considera o cálculo do “*sharpe*” que é colocado dessa forma devido a esta variável ser uma aproximação do índice de *sharpe* real calculado pela equação (4.28) considerando o retorno livre de riscos R_f como nulo dado que a aproximação é.

No momento que a ordem de compra ou venda é definida pela rede neural, temos que o papel do alocador de capital é simplesmente ditar qual seria a quantidade de capital do montante do investidor a ser investido na operação, desta maneira temos a divisão das responsabilidades do alocador e rede neural em que o primeiro define a quantidade de investimento e a segunda parte define o momento a ser investido.

Para que se possa exemplificar o processo fim a fim podemos tomar a situação em que um investidor gostaria de realizar uma operação financeira e possui X de capital. Inicialmente o investidor realiza a definição de qual seria o ativo de interesse (ou um conjunto de ativos $C_0, C_1, C_2, \dots, C_n$ sendo passado os mesmos para que a rede neural possa realizar o treinamento e começar a identificar as possíveis oportunidades. Quando a rede neural enxerga uma oportunidade (dado por uma predição de valor futuro maior que o valor atual) no ativo C_n a mesma dispara uma ordem de compra do ativo sendo que o valor x_n de compra será definido pela teoria de Markowitz, logo uma ordem para o ativo C_n com valor X_n . Após a execução da ordem de compra temos que a rede neural irá fazer a predição da cotação do próximo período e tomando com base a estratégia de *market making* obrigatoriamente a próxima ordem será de venda, logo que o valor do ativo alcance um determinado valor a venda será executada.

5 Testes e Resultados

Este capítulo apresenta os testes realizados e resultados obtidos pelo preditor ou seja pelo módulo rede neural com valores base OHLC, pautado nas estratégias de *trading* de *momentum* e de *market making*. O alocador de capital e o *dataset* com indicadores financeiros funcionam momentaneamente como módulos *standalone* o qual não trabalham junto com o preditor e serão desenvolvidos em trabalhos futuros.

Para a coleta dos resultados, foram analisadas 50 cotações de ativos da B3, seguindo as duas linhas de investimento já citadas. Os resultados foram coletados considerando 50% dos dados para treino. Foram utilizadas as métricas dispostas na seção 4.2 a fim de se aferir a qualidade da resposta dos ativos durante a simulação.

Os parâmetros de treino foram definidos da seguinte forma $H_t = 15$, $H_P = 5$, $epochs = 50$, $batch\ size = 32$.

5.1 Estratégia *Momentum*

Na estratégia de *momentum* destacam-se os seguintes pontos dentre outros resultados:

- O AALR3 obteve uma rentabilidade de 79% o que mostra que a técnica de *momentum* também é efetiva para padrões de *Machine Learning*.
- Quase 50% dos ativos sofreram prejuízo — rentabilidade negativa o que indica alto grau de risco para aplicação desta técnica em conjunto com a rede neural utilizada.
- 14 ativos realizaram o maior número de trades — 36, o que indica uma movimentação intensa nesse mercado.
- a razão erro/retorno é um bom indicativo de *momentum* visto que quanto mais próxima de 1 menor é o risco e maior a oportunidade de uma boa rentabilidade, para valores absolutos maiores de erro médio e retorno quadrático médio teremos maior rentabilidade.

Os resultados de *momentum* dos outros ativos estão disponíveis no Anexo II, deste documento. As linhas verdes em Anexo correspondem aos 10 ativos mais rentáveis. As linhas em vermelho representam ativos com rentabilidade acumulada negativa. As Figuras 5.1 ,5.2 e 5.3 permite visualizar o top 5 em alguns resultados importantes.

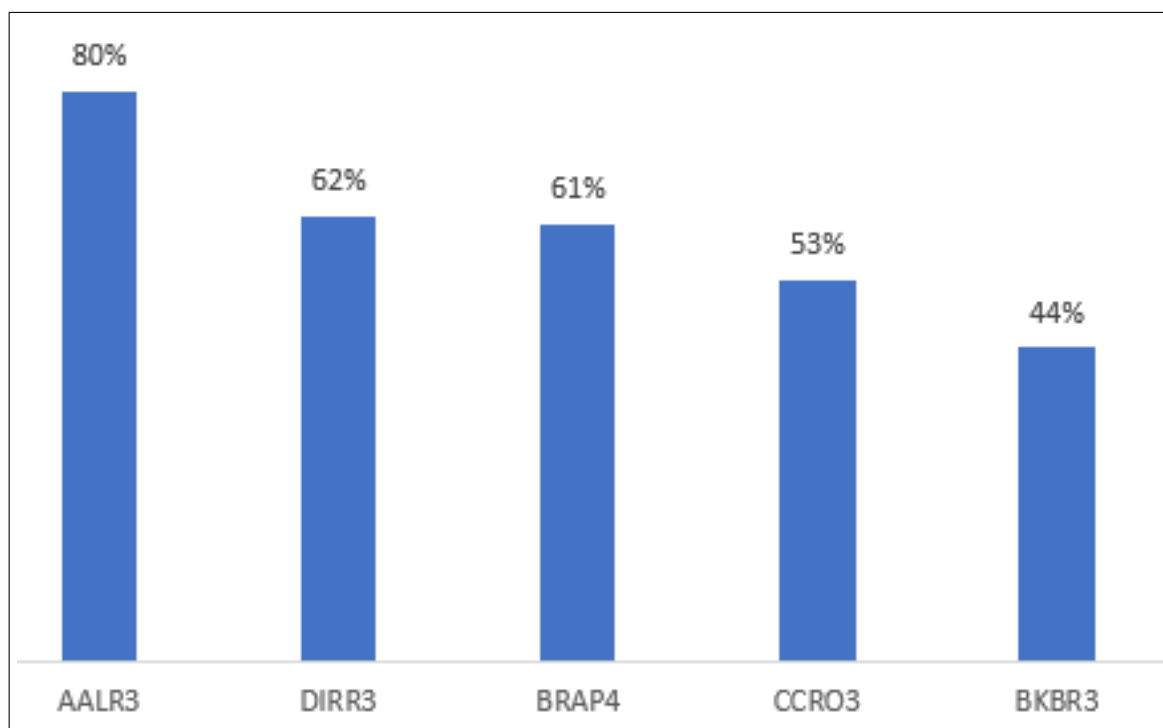


Figura 5.1 – Top 5 — Ativos mais rentáveis na estratégia de *momentum*. Fonte: dos autores

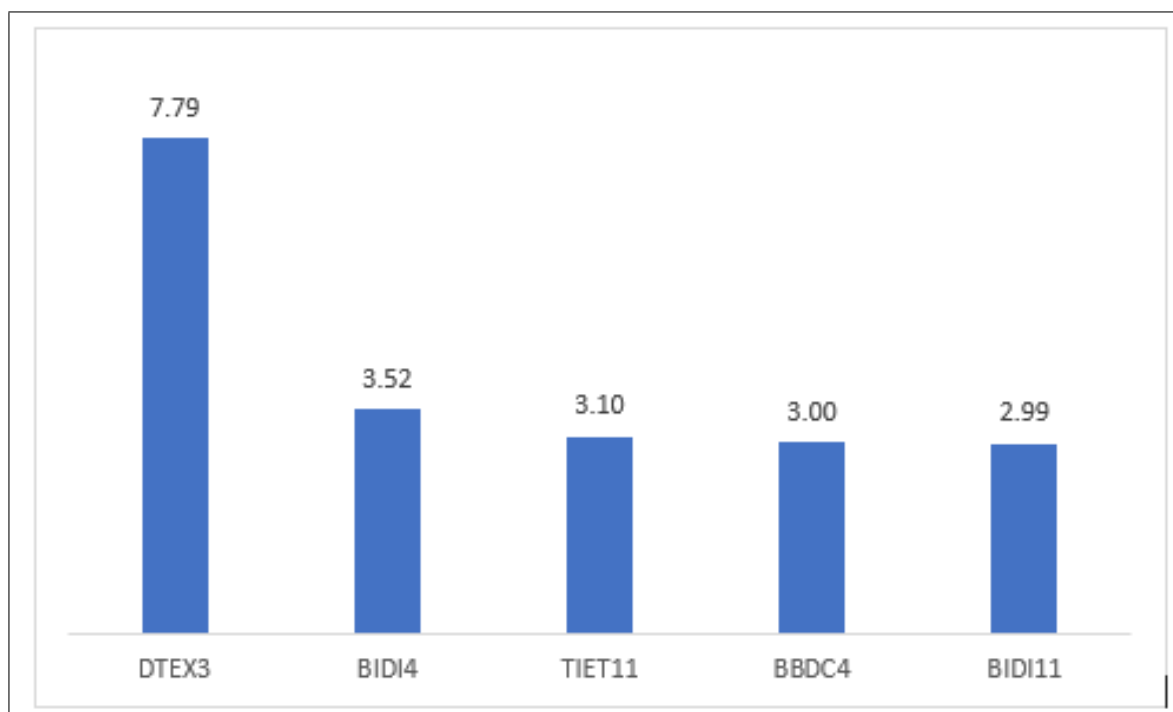


Figura 5.2 – Top 5 - Ativos com maior valor de erro sobre investimento na estratégia de *momentum*. Fonte: dos autores

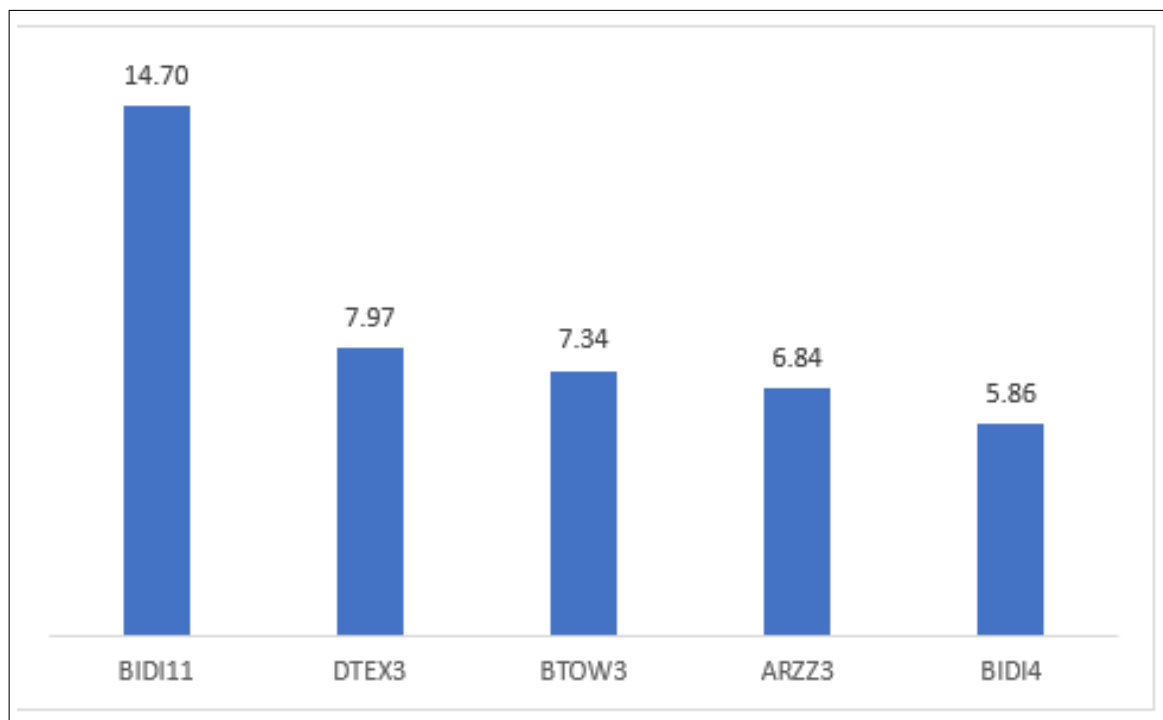


Figura 5.3 – Top 5 - Ativos com maior erro quadrático médio na estratégia de *momentum*.
Fonte: dos autores

5.2 Estratégia *Market Making*

A estratégia de *market making* gerou boa rentabilidade em boa parte dos ativos, todos realizaram pelo menos um *trade* durante a simulação. Dos resultados observados, destacam-se:

- Cerca de 42% obtiveram uma rentabilidade acima de 30% e apenas 10% das ações acumulou rentabilidade negativa, o que demonstra que o risco de investimento é pequeno tamanha a margem lucro e baixa perda de investimento.
- A CIEL3 possui o menor “*sharpe*”, o que indica que este ativo é o mais arriscado de se investir.
- A BRPR3 possui o menor risco de investimento em operações de alta e baixa na bolsa - 1.11 e 1.06 respectivamente.
- a BRAP4 vendeu o maior número de papéis — 20, obtendo uma rentabilidade acumulada de 70%. O BRKM5 foi o ativo que teve maior rendimento entre todos, embora tenha vendido metade dos papéis de BRAP4 - apenas 10. Isso se deve ao risco de investimento do BRAP4 ser maior, o que é evidenciado pelo menor “*sharpe*” e pelo maior erro quadrático médio em valores de alta onde se busca lucrar com a venda.

Os resultados de *market making* dos outros ativos estão disponíveis no Anexo I. As linhas verdes em Anexo correspondem aos 10 ativos mais rentáveis desta estratégia. As linhas em vermelho representam ativos com rentabilidade acumulada negativa.

Os líderes em alguns quesitos estão presentes nas Figuras 5.4 ,5.5 e 5.6, que permitem visualizar o top 5 em alguns resultados importantes.

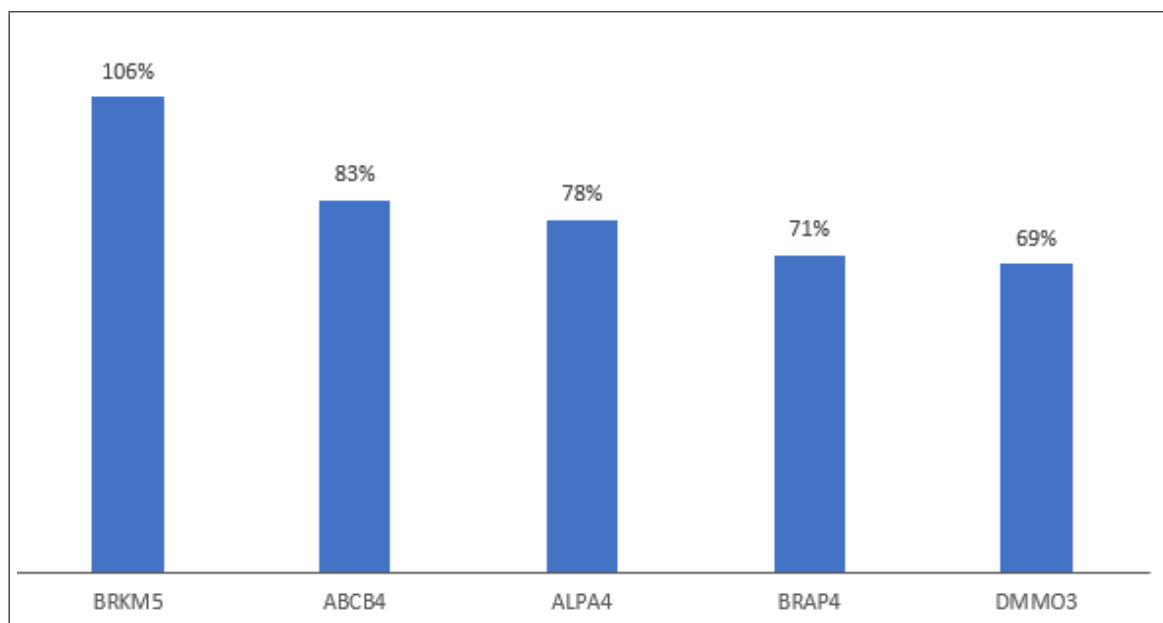


Figura 5.4 – Top 5 — Ativos mais rentáveis na estratégia de *market making*. Fonte: dos autores

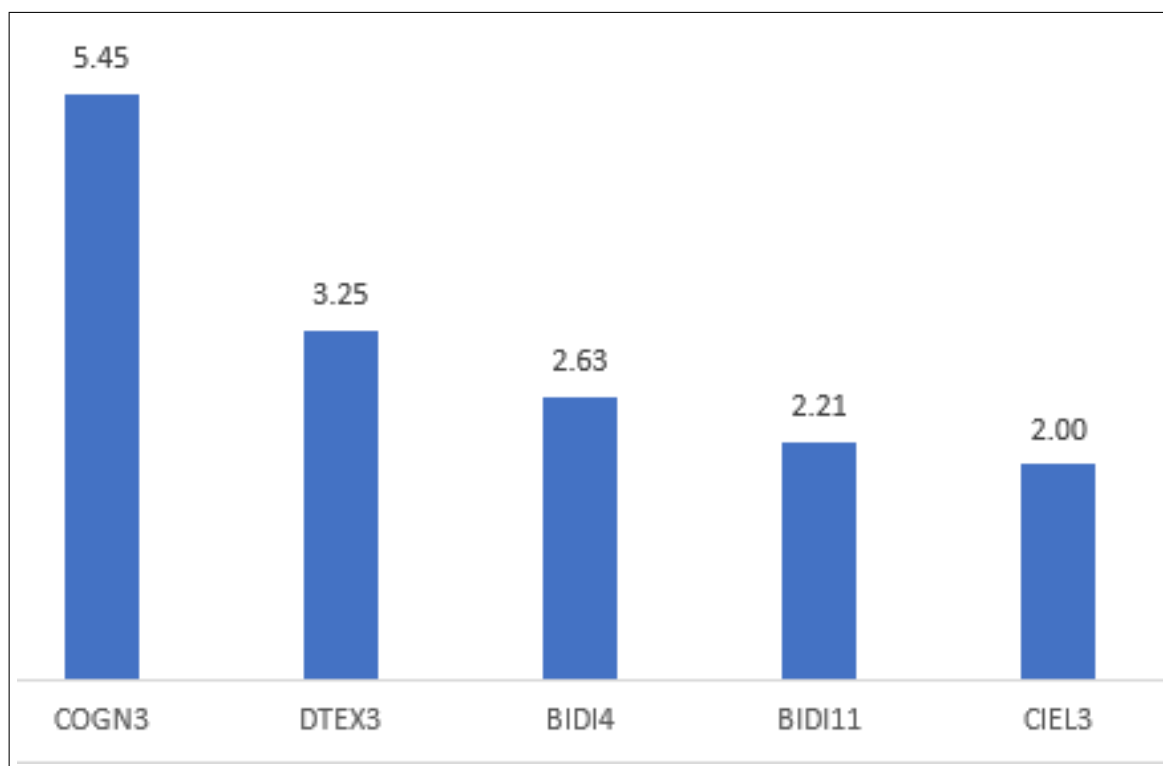


Figura 5.5 – Top 5 — Ativos com maior valor sobre o investimento em *market making*.
 Fonte: dos autores

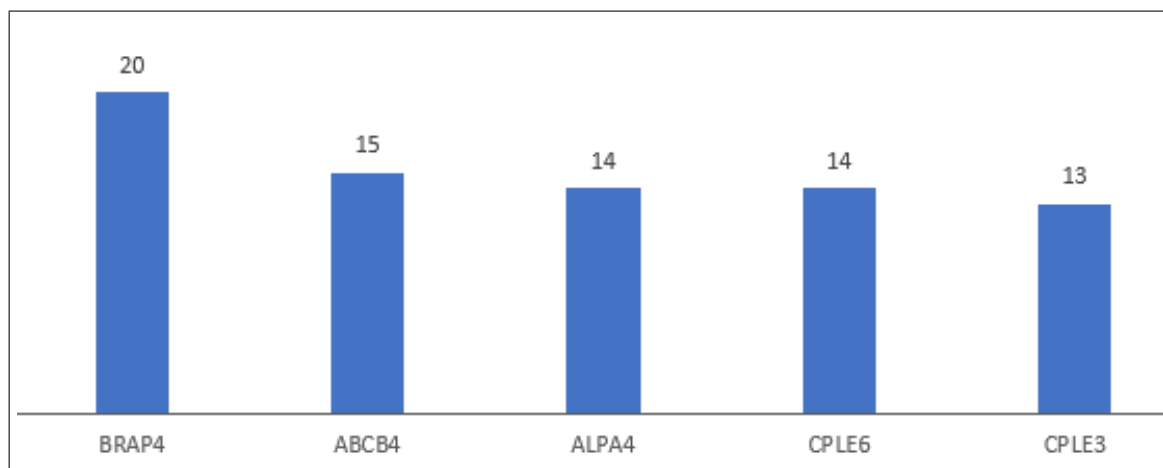


Figura 5.6 – Top 5 — Ativos que mais negociam na estratégia de *market making*. Fonte:
 dos autores

5.3 Alocação do capital

A alocação do capital foi realizada através da Teoria de Markowitz, sendo desta maneira realizada de forma matemática em busca do melhor portfólio possível, para a execução e elaboração dos resultados foram utilizados dois diferentes conjuntos de ações provenientes de empresas negociadas na B3.

Para a realização do primeiro portfólio foi utilizado um conjunto de ações de empresas que são de setores distintos, pois assim se minimiza a correlação entre os ativos. Em uma segunda análise foi realizado o oposto, em que vários ativos do mesmo setor econômico foram utilizados para se poder ter uma maior correlação e assim avaliar o desempenho da carteira ideal.

5.3.1 Setores distintos

Com a execução da teoria de Markowitz sobre o conjunto de ações que são provenientes de empresas que atuam sobre setores distintos, é possível visualizar na Figura 5.7 que temos uma boa distribuição dos papéis no plano de Markowitz, pois os mesmos não apresentam uma correlação direta um com o outro, pois são empresas que atuam em setores distintos da economia. A baixa correlação entre as ações pode ser visto quando comparamos lado a lado o histórico de preços das mesmas, visto que uma queda em uma ação não é refletida diretamente na outra.

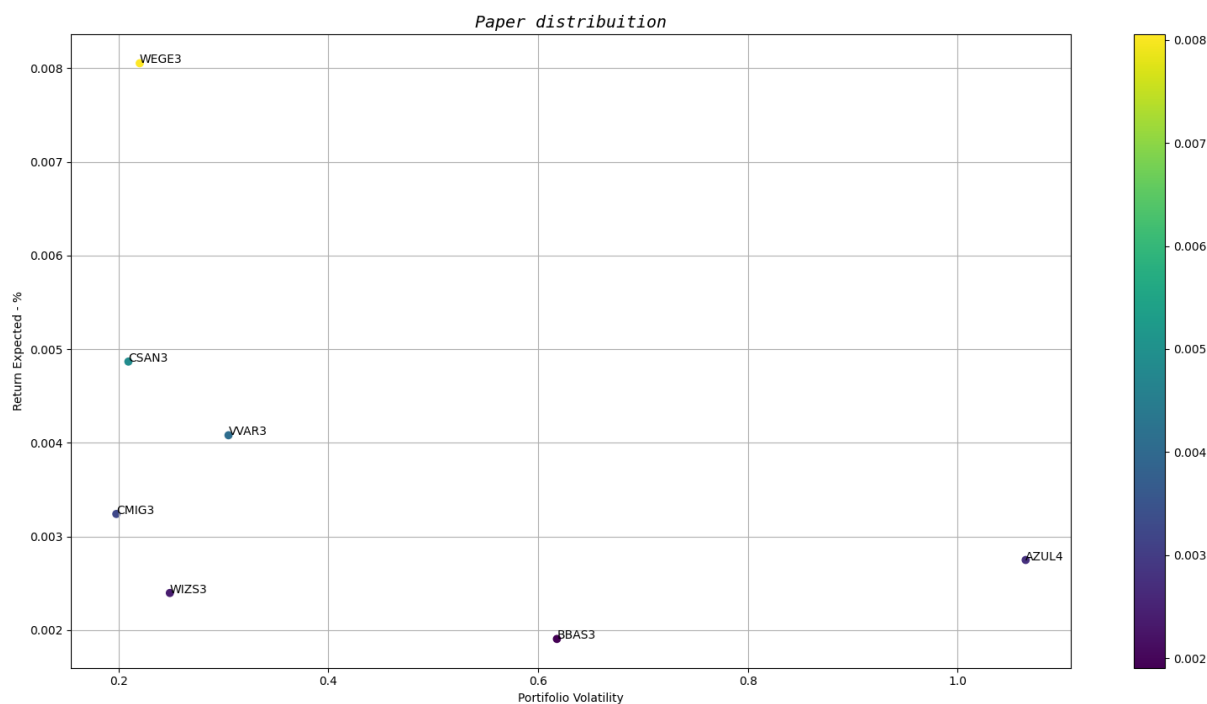


Figura 5.7 – Carteira multisetores — Distribuição de papéis. Fonte: dos autores

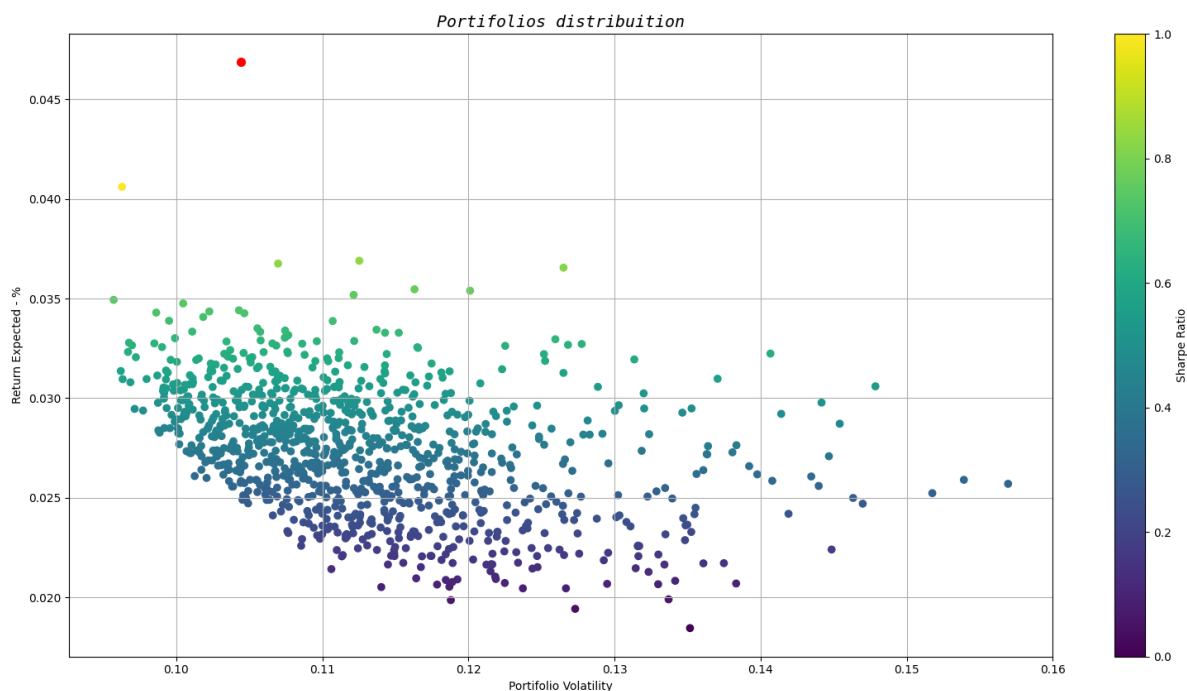


Figura 5.8 – Carteira multisetores — Distribuição de carteiras. Fonte: dos autores

A execução da alocação do capital dentre as ações que compõem o conjunto é mostrada na Figura 5.8. Nesta figura são mostradas carteiras aleatórias que alocam o capital nas respectivas ações, sendo assim possível notar que todas as carteiras montadas saíram com lucros por mais que não fossem as carteiras ótimas e se encontram bem distribuídas horizontalmente, conseqüentemente apresentam vários níveis de risco para o mesmo retorno.

Outra observação que pode ser avaliada é que as carteiras aleatórias apresentam uma rentabilidade por risco (medida também conhecida como *sharpe*) muito mais baixa que a carteira ideal, mostrada em vermelho, pois utilizam os recursos do investidor de uma forma não otimizada e assim se incubem a riscos maiores. Ainda analisando as carteiras aleatórias é possível notar que temos apenas uma que apresenta um nível de risco equiparável ao seu lucro, sendo a mesma plotada em amarelo. O ponto vermelho que simboliza a carteira ótima também mostra que quando estamos com o melhor *sharpe* possível não necessariamente estamos com o menor risco induzido, sendo assim um fator já esperado quando o assunto é renda variável, pois é necessário se assumir riscos para poder encontrar melhores retornos.

Com o uso da Figura 5.9 é possível notar que nenhuma das carteiras aleatórias pode ser considerada ótima, pois não estão sob a curva de Markowitz, sendo este um fator esperado, pois o uso descontrolado e sem planos do capital nunca irá convergir para o melhor resultado que uma carteira pode oferecer, desta maneira a única carteira que se

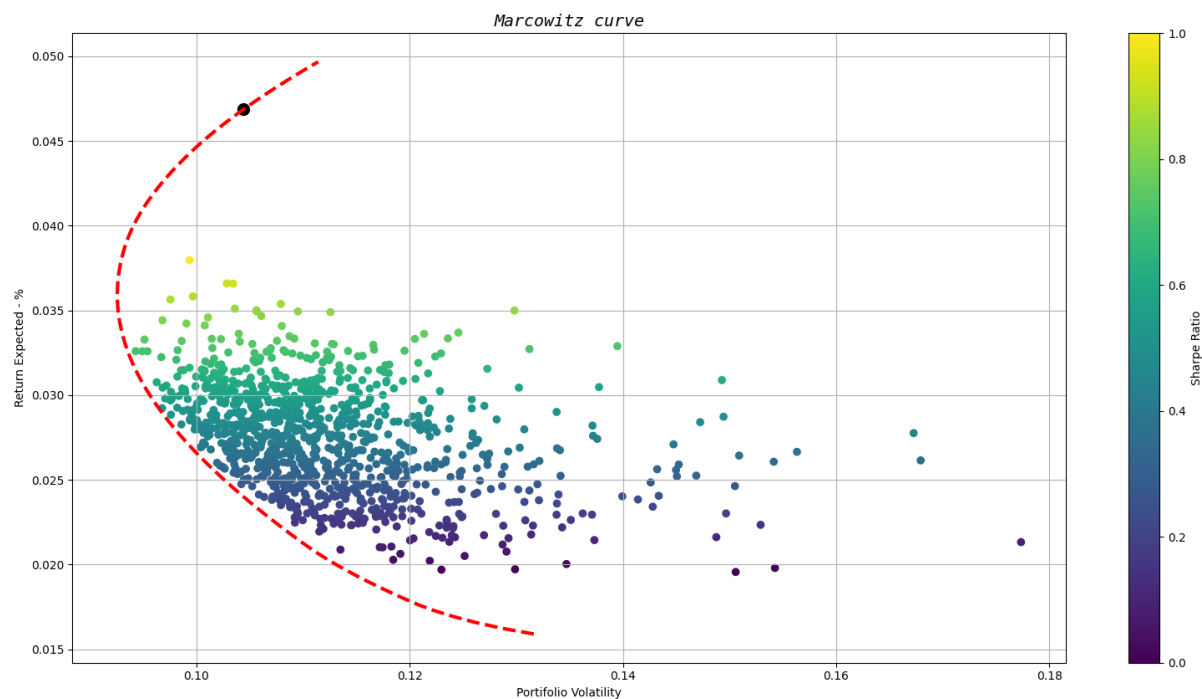


Figura 5.9 – Carteira multisetores — Distribuição de carteiras. Fonte: dos autores

encontra sob a curva é a carteira obtida aplicando a teoria de Markowitz.

Outro fator que pode ser observado é que dentre as carteiras aleatórias, temos algumas que estão próximas à fronteira de Markowitz, sendo assim elas não necessariamente apresentam o melhor retorno.

5.3.2 Setores iguais

Nesta etapa dos resultados foi utilizado ações de empresas que atuam no setor bancário, sendo assim empresas que possuem ações que apresentam uma alta correlação visto que o desempenho do setor como um todo irá refletir individualmente nas ações das mesmas. Quando comparamos a Figura 5.7 com a Figura 5.10 é possível ver, pelas escalas do gráfico, que o segundo conjunto de ações está mais concentrado, pois os mesmos apresentaram desempenhos semelhantes ao longo da janela de cotações obtidas, visto a atuação no mesmo setor econômico.

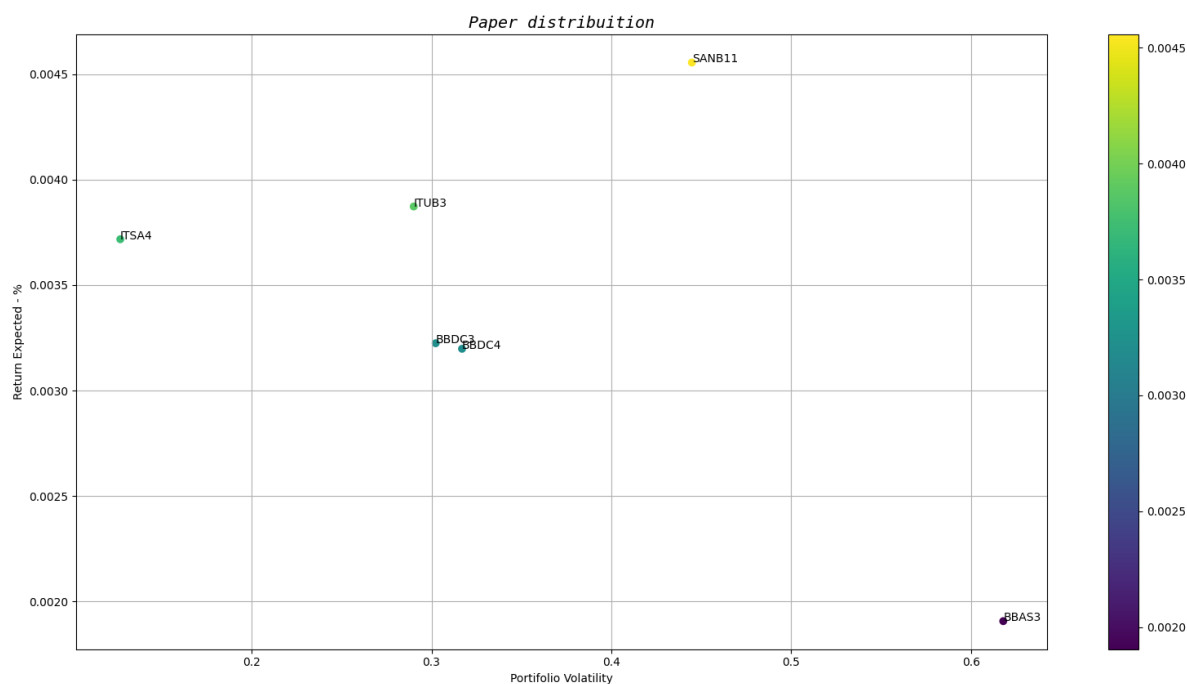


Figura 5.10 – Carteira setORIZADA — Distribuição de papéis. Fonte: dos autores

A distribuição das carteiras aleatórias e a carteira ótima no que lhe concerne possui um padrão único mostrado na Figura 5.11 sendo totalmente diferente daquele mostrado na Figura 5.8, pois desta vez o padrão mostrado é diagonalizado. A disposição diagonal das carteiras é totalmente justificável, pois o risco aumenta para menores rendimentos, visto que a correlação dos ativos causa que o resultado individual possua uma maior ponderação no resultado do portfólio como um todo, sendo assim as carteiras com uma coloração roxa apresentam um investimento proporcionalmente maior na ação BBSA3, dentre as existentes no conjunto é a que apresenta o maior risco.

Ainda se analisando a Figura 5.11 é possível se notar que o maior número de carteiras é disposto em uma colocação media esverdeada, pois dentre o conjunto de ações escolhidas há a existência de uma alta correlação entre os ativos BBDC3, BBDC4 e ITUB4 e as mesmas possuem resultados individuais médios que correspondem a escala

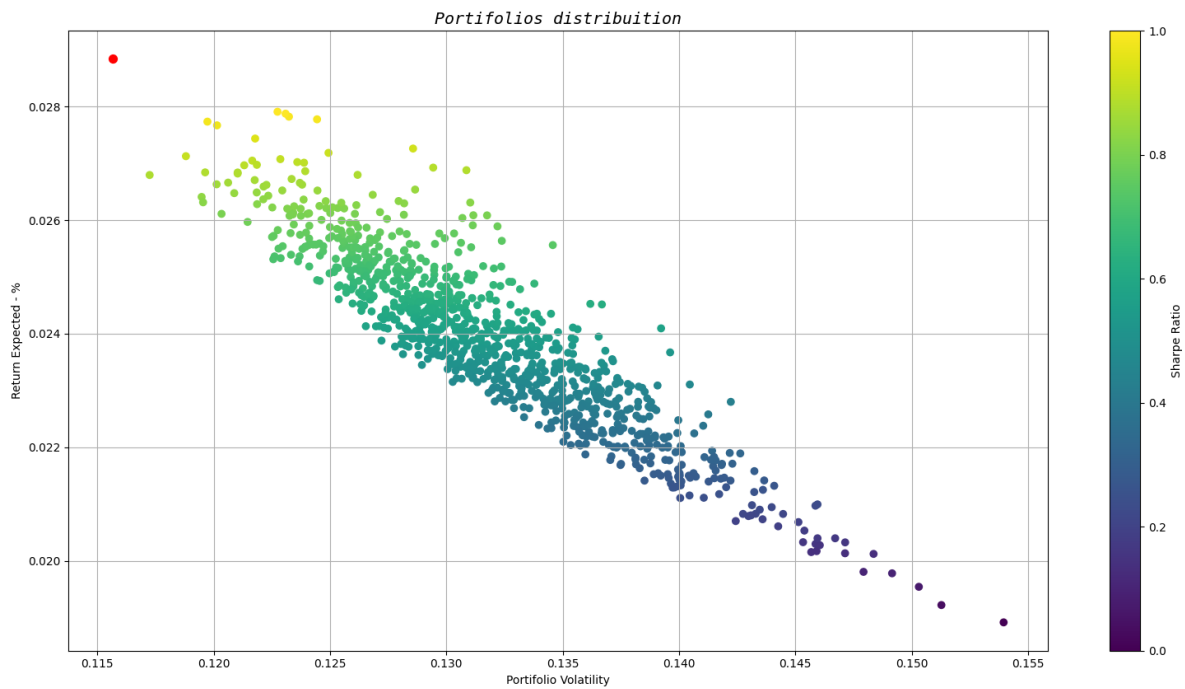


Figura 5.11 – Carteira setORIZADA — Distribuição de papéis. Fonte: dos autores

esverdeada do gráfico, logo o portfólio em que tais ativos possuam uma maior composição, inevitavelmente terá um resultado que será bastante semelhante ao desempenho individual das mesmas. Já a carteira ótima que é plotada em vermelho possui uma maior parte do capital investido na ação ISA4 e pouco(ou mesmo nenhum) capital investido em BBSA3, podendo ser inferido, pois, como mencionado anteriormente o retorno do portfólio neste cenário é bem mais relacionado ao desempenho individuais das ações.

A Figura 5.12 mostra a fronteira de Markowitz e neste caso é possível ver que como o perfil da curva segue as carteiras aleatórias e novamente apenas a carteira ótima se encontra sob a fronteira, fato também mostrado na Figura 5.9. A carteira ótima neste momento se encontra bem mais próximo ao ponto de inflexão da fronteira de Markowitz, pois assim como mencionado no parágrafo anterior esta carteira é majoritariamente composta pela ação ITSA4 e tal ação apresenta uma performance bem similar a tal carteira. Comparando a fronteira de Markowitz dos conjuntos de ações provenientes de empresas de setores distintos com empresas do mesmo setor é possível se notar que no último caso o melhor a curva apresenta um perfil mais “fechado” devido à posição mais acumulada dos ativos no plano retorno por risco.

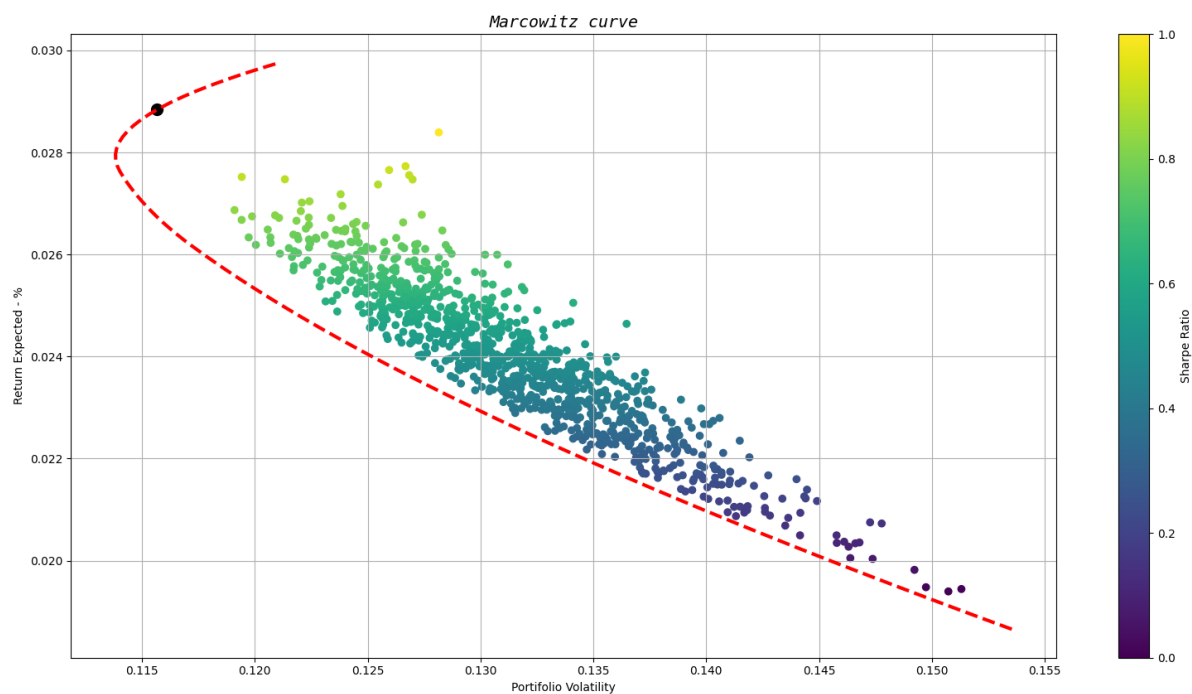


Figura 5.12 – Carteira setORIZADA — Distribuição de papéis Fonte: dos autores

6 Considerações Finais

O capítulo apresenta a compreensão dos resultados, impacto do desenvolvimento, dificuldades e limitações e trabalhos futuros.

6.1 Escopo e limitações

A complexidade dos diversos módulos de redes neurais e estratégias de *trading* permitem expandir muito além do que foi desenvolvido neste documento, visto as inúmeras modelagens matemáticas aplicáveis a área de *Deep Learning*. O escopo inicial do projeto utilizaria como entrada um *dataset* gerado pelos indicadores de análise técnica e OHLC em uma rede neural preditora que trabalharia junto com um classificador permitindo o desenvolvimento de estratégias de investimentos e técnicas de análise mais complexas. O potencial do *dataset* também não foi explorado, tendo apenas sido desenvolvido em sua estrutura estando funcional e aplicável para gerar diversos resultados que não foram testados a fundo devido a maior complexidade de compreensão e análise dos arranjos indicador-OHLC.

A estratégia de *momentum* se mostrou mais aquecida, com maior volume de negociações e entradas no mercado, porém também com maior risco de investimento. Visto que esta busca prever um determinado valor com exatidão, ela acaba sendo menos eficiente que a técnica de *market making* que embora possua uma movimentação menor no mercado, espera a hora certa e busca encontrar as melhores oportunidades, mitigando riscos de investimento, ou seja, gerando maior rentabilidade com menor risco.

6.2 Conclusão

O trabalho executado ao decorrer deste documento permitiu concluir que a análise gráfica é uma área de bastante estudo e que também permite boas análises de oportunidades de investimento, sendo essa abordagem de análise uma porta de entrada para a inserção de redes neurais em interpretações de sinais financeiros. A rede neural desenvolvida permitiu entender que não existe uma maneira determinística de modelagem que venha a conceder a melhor análise de todas as ações listadas na bolsa, pois temos que cada ativo tem o seu passeio normal de deslocamento e conseqüentemente suas respectivas características, logo um único modelo de rede neural não consegue realizar a generalização de forma a atender margens aceitáveis de rentabilidade.

Um outro prisma que foi possível obter com a realização do trabalho é que o uso

de redes neurais traz benefícios ao setor financeiro, pois com esse tipo de abordagem é possível identificar padrões recorrentes no histórico de um ativo que o paradigma de programação clássica não seria capaz de realizar, pois a rede neural possui um comportamento "orgânico" que permite interpretar os dados de forma a entender a conexão entre eles. Um outro ponto que é possível de se consumir é de que a rede modelada em LSTM realiza previsões que tendem ao valor real do ativo, sendo assim a mesma apresenta um comportamento semelhante a uma média móvel ou mesmo um RSI em que na esmagadora maioria dos casos se orienta na direção que o ativo realmente toma.

Embora os resultados colhidos tiveram uma boa rentabilidade, a rede neural LSTM nunca consegue interceptar a curva real, pois está sempre atrasada em relação a esta o que dificulta a mitigação de riscos, juntamente com certo grau de aleatoriedade e enviesamento do modelo, que carece de solução consistente para ser aplicada no mercado de ações, necessitando de uma outra modelagem não linear ou técnicas combinadas que possibilitem refinar e aproximar mais os modelos preditivos LSTM dos valores reais.

Um outro fundamento que é possível ser entendido com o escopo do projeto é que a abordagem utilizada na teoria de Markowitz é extremamente eficiente em se encontrar a melhor distribuição de patrimônio para um dado conjunto de ativos (renda variável, renda fixa, fundos de investimento e etc), logo se nota que o investidor pode utilizar a teoria para encontrar o maior lucro possível para uma determinada taxa de risco que o mesmo tolera.

A teoria de Markowitz não possui o objetivo de realizar a seleção de ativos, mas apenas realizar a distribuição do capital, porem a mesma também permite identificar a correlação entre ativos distintos para que assim o investidor possa realizar uma análise e realize a tomada de decisão do conjunto de ativos que o mesmo tomará em sua carteira, pois caso tenha muitos ativos correlacionados irá ter períodos em que há uma valorização de todos os ativos, ao mesmo tempo que tem o risco de todos desvalorizarem, logo a teoria enfatiza que a diversificação dos ativos é, em média, a melhor alternativa.

6.3 Trabalhos futuros

A realização do escopo proposto para esse trabalho mostra que existem próximas etapas que melhorariam a análise de sinais financeiros com o uso de redes neurais, pois é possível realizar um refinamento do modelo que poderia realizar a previsão com maior acurácia utilizando-se uma rede neural mesclada, apresentando camadas convolucionais e LSTM por exemplo, para realizar a previsão. Um refino possível seria realizar metodologias distintas de treinamentos para um mesmo modelo de RN em busca da melhor generalização.

Além do avanço do modelo da rede neural proposto, a utilização de outras estraté-

gias para lidar com os *triggers* é uma maneira viável de se aumentar a rentabilidade, sendo assim realizar o teste de desempenho de uma mesma rede neural submetida a diferentes estratégias de *trading*. Durante o projeto tivemos apenas a abordagem de realizar operações opostas a cada turno (compra depois vende), porém uma estratégia a ser testada e a de realização de múltiplas compras e vendas pois muitas vezes um ativo pode entrar em condições ainda mais favoráveis do que foi identificado na mesma operação anterior.

Um outro ponto de melhoria no projeto seria realizar uma alocação dinâmica de capital que fosse integrada ao sistema de decisão de operação (rede neural + módulo de estratégia) pois podendo até mesmo ser implementado com o uso de uma segunda rede neural que realiza a identificação de risco da operação e assim dimensiona a quantidade de capital a ser aplicada naquela operação, em que tal rede neural seja capaz de realizar a identificação de oportunidades futuras e assim ter perdas controladas.

Os pares indicador-OHLC podem gerar diversas informações com uma análise mais complexa e detalhada destes *datasets*, que podem ser desenvolvidos futuramente pelos autores ou pela comunidade científica.

Referências

- AGROCLASI.COM. *Candle patterns*. 2021. <<https://br.pinterest.com/pin/739012620101858657/>>. Citado 2 vezes nas páginas viii e 13.
- ASTHANA, A. P. S.; BAHRDWAJ, A. *Analysis of Multiple Hidden layers vs Accuracy in Performance using Back Propagation Neural network*", *Indian Journal of Science and Techonology*. 2017. Citado na página 27.
- B3. *Uma análise da evolução dos investidores na B3*. 2021. <https://www.b3.com.br/data/files/35/17/D6/0D/B37EC710BD0885C7AC094EA8/Book_PF_Set2021.pdf>. Citado na página 1.
- B3. *Índice Bovespa*. 2021. <https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm>. Citado na página 10.
- BOAINAIN, P. G.; PEREIRA, P. L. V. "ombro-cabeça-ombro": Testando a lucratividade do padrão gráfico de análise técnica no mercado de ações brasileiro. *Revista Brasileira de Finanças*, Sociedade Brasileira de Finanças, v. 7, n. 3, p. 265–303, 2009. Citado na página 2.
- BORGES, R. Eficiência e o mercado de renda variável brasileiro. 2010. Citado na página 15.
- BOTELHO, F. d. A. Análise técnica & estratégia operacional. *Curso de Análise Técnica, Florianópolis*, 2004. Citado na página 15.
- BRUM, F. M. Aplicação da teoria de markowitz e índice de sharpe em um clube de investimento. 2008. Citado 4 vezes nas páginas 1, 2, 19 e 20.
- CHOI, R. Y. et al. Introduction to machine learning, neural networks, and deep learning. *Translational Vision Science & Technology*, The Association for Research in Vision and Ophthalmology, v. 9, n. 2, p. 14–14, 2020. Citado 4 vezes nas páginas viii, 21, 22 e 23.
- CORREIA, J. S. *Operando na bolsa de valores utilizando análise técnica*. [S.l.]: Novatec Editora, 2008. Citado 2 vezes nas páginas 14 e 17.
- ELDER, A. *Como se transformar em um operador e investidor de sucesso*. [S.l.: s.n.], 2004. Citado na página 14.
- FREITAS, S. O. de. Utilização de um modelo baseado em redes neurais para a precificação de opções. Universidade Federal de Minas Gerais, 2001. Citado 2 vezes nas páginas 2 e 11.
- GARETH, J. et al. *An introduction to statistical learning: With applications in r*. New York: Springer, 2013. Citado 6 vezes nas páginas 21, 22, 23, 24, 25 e 26.
- GUARNIERI, O. C. Um estudo empírico da eficiência da análise técnica como instrumento na predição do comportamento dos preços das ações: o caso embraer. Brasil, 2006. Citado 3 vezes nas páginas 15, 16 e 17.

- GUPTA, T.; K.RAZA. *Optimizing Deep Neural network Architecture: A Tabu search Based Approach*. 2020. "Department of Computer Science, Jami Milia Islamia, New Delhi, Accounting and Finance, 5(3), 284–308.",110025". Citado na página 27.
- HENRIQUE, B. M. Predição da direção dos preços de ativos do mercado financeiro usando aprendizagem de máquina. 2018. Citado 2 vezes nas páginas viii e 27.
- INFOLKS, L. P. *"Recurrent Neural Network and Long Term Dependencies"*. 2019. <https://medium.com/tech-break/recurrent-neural-network-and-long-term-dependencies-e21773defd92>, Acessado em: 21/11/2021. Citado 2 vezes nas páginas 32 e 33.
- INVESTIMENTOS, T. *O que é a Bolsa de Valores? Conheça mitos, verdades e como funciona!* 2021. <<https://blog.toroinvestimentos.com.br/bolsa-de-valores-o-que-e-como-funciona>>. Citado 3 vezes nas páginas viii, 9 e 10.
- KARSOLIYA, S. *Approximating Number of Hidden layer Neurons in Multiple hidden Layer BPNN Architecture*. 2012. "International Journal of Engineering Trends and Technology Volume 3, issue 6. Citado 3 vezes nas páginas 27, 28 e 29.
- KINGMA, J. B. D. P. Adam: A method for stochastic optimization. *Cornell University*, 2015. Citado na página 50.
- KRAFT, D. "a software package for sequential quadratic programming". *DFVLR-FB 88-28*, 1988. Citado na página 47.
- LABOISSIERE, L. A.; FERNANDES, R. A.; LAGE, G. G. Maximum and minimum stock price forecasting of brazilian power distribution companies based on artificial neural networks. *Applied Soft Computing*, Elsevier, v. 35, p. 66–74, 2015. Citado na página 27.
- LACERDA, Á. L. Análise técnica e visualização de dados do mercado de ações utilizando python. Pontifícia Universidade Católica de Goiás, 2021. Citado na página 12.
- LU, T.-H.; SHIU, Y.-M. Tests for two-day candlestick patterns in the emerging equity market of taiwan. *Emerging markets finance and trade*, Taylor & Francis, v. 48, n. sup1, p. 41–57, 2012. Citado na página 12.
- MOBILIÁRIOS, C. C. de V. Comissão de valores mobiliários. *Instrução Normativa*, n. 153, 2013. Citado na página 6.
- MOLNAR, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Morrisville, North Carolina. 2019. <<https://christophm.github.io/interpretable-ml-book/>, Morrisville, North Carolina: LuluPress, Inc.> Citado na página 22.
- NORONHA, M. Análise técnica: teorias, ferramentas, estratégias. *São Paulo: Editec*, p. 30, 1995. Citado na página 15.
- NVIDIA. *NVIDIA Blog: Supervised vs Unsupervised Learning. The Official NVIDIA Blog*. 2018. <<https://blogs.nvidia.com/blog/2018/08/02/supervised-unsupervised-learning/>. Published August 2, 2018. Accessed October 28, 2021>. Citado 3 vezes nas páginas 22, 23 e 24.

SEIDEL, J. P. B. Revisitando o modelo de markowitz para a otimização de portfólios do btg pactual digital. Niterói, 2018. Citado na página 19.

SILVA, R. B. S. da. Abordagem histórica da bolsa de valores. *Maiêutica-Estudos Contemporâneos em Gestão Organizacional*, v. 4, n. 1, 2016. Citado na página 8.

T TIBSHIRANI R, F. J. H. The elements of statistical learning: Data mining, inference, and prediction. applications in r. New York: Springer, 2009. Citado 6 vezes nas páginas 21, 22, 23, 24, 25 e 26.

UNUCSEH, J. d. I. C. da. Desmistificando a bolsa de valores: História e atualidade. 2020. Citado na página 8.

WERBOS, P. "Backpropagation through time: What it does and how to do it. *Proc. IEEE* 1990,78 1550-1560.". 1990. [Http://dx.doi.org/10.1109/5.58337](http://dx.doi.org/10.1109/5.58337)). Citado na página 32.

XUA, S.; L.CHEN. "A novel approach for determining the optimal number of hidden later neurons for FNNs and its application in data mining",. 2008. "5th(ICITA)". Citado na página 28.

YAZBEK, O. A regulamentação das bolsas de valores e das bolsas de mercadorias e futuros e as novas atribuições da comissão de valores mobiliários. *Revista de Direito Bancário e do Mercado de Capitais*, v. 34, p. 687–709, 2006. Citado na página 6.

ZANINI, F. A. M.; FIGUEIREDO, A. C. As teorias de carteira de markowitz e de sharpe: uma aplicação no mercado brasileiro de ações entre julho/95 e junho/2000. *RAM. Revista de Administração Mackenzie*, SciELO Brasil, v. 6, p. 38–65, 2021. Citado 2 vezes nas páginas 18 e 19.

ZHU ZHILE YANG, M. M. Y. G. Y. Z. Y. C. Y. W. S. F. J. Electric vehicle charging load forecasting: A comparative study of deep learning approaches. *MDPI - Multidisciplinary Digital Publishing Institute*, 2019. Citado 4 vezes nas páginas viii, 31, 34 e 36.

Anexos

ANEXO I – ESTRATÉGIA DE MARKET MAKING

Ativo	erro médio quadrático high	erro médio quadrático low	retorno médio quadrático high	retorno médio quadrático low	rentabilidade acumulada	número de trades	coeficiente de variação	métrica_high	métrica_low
BRKM5	3,9434	4,0025	2,6419	2,4513	105,70%	10	2,6302	1,4926	1,6328
ABCB4	0,6437	0,6552	0,5734	0,5453	82,71%	15	0,9361	1,1226	1,2016
ALPA4	3,9335	3,8299	2,3365	2,2987	78,34%	14	0,9448	1,6835	1,6661
BRAP4	4,7073	3,9926	3,2755	3,1361	70,53%	20	0,9191	1,4371	1,2731
DMMO3	0,2679	0,2032	0,1479	0,0703	68,53%	5	0,6437	1,8118	2,8901
BKBR3	0,8096	0,7955	0,5947	0,5901	57,53%	9	1,3701	1,3614	1,3481
CRFB3	0,9810	0,9111	0,8289	0,7796	57,41%	10	0,9183	1,1834	1,1687
CCRO3	0,6662	0,6637	0,5642	0,5565	54,31%	8	1,8684	1,1809	1,1928
DIRR3	1,1019	0,9677	0,7053	0,7025	52,66%	8	0,8162	1,5624	1,3775
ABEV3	0,8743	0,9237	0,7058	0,6831	50,31%	11	0,8893	1,2387	1,3522
HGTX3	3,3177	3,3678	1,7684	1,5951	43,65%	10	2,2144	1,8761	2,1113
RLOG3	1,3208	1,6020	0,8684	0,8024	42,34%	5	1,2920	1,5210	1,9964
ALUP11	1,5654	1,6560	0,6507	0,6798	38,44%	12	1,3741	2,4057	2,4359
ARZZ3	6,7148	6,3139	3,1146	3,0349	37,48%	9	0,8085	2,1559	2,0804
BPAC11	1,9751	2,1221	1,2573	1,2606	36,94%	7	0,9427	1,5709	1,6835
BBDC4	2,7995	3,0548	0,9734	0,9347	36,35%	4	1,9959	2,8760	3,2683
BRPR3	0,3657	0,3973	0,3446	0,3571	35,35%	7	1,0128	1,0611	1,1127
BRSR6	0,6401	0,6265	0,4594	0,4345	34,47%	9	1,1656	1,3934	1,4420
CESP6	1,3554	1,2331	0,7632	0,7838	34,07%	8	1,1822	1,7760	1,5732
ANIM3	1,4847	1,6195	0,7199	0,6927	33,52%	8	0,5381	2,0623	2,3380
AALR3	1,0426	0,8505	0,8305	0,6998	31,89%	9	0,4898	1,2553	1,2153
BRML3	0,7293	0,6553	0,5088	0,5165	23,11%	7	0,4895	1,4333	1,2686
BBDC3	1,8043	1,8838	0,8181	0,7987	22,92%	6	1,0778	2,2055	2,3587
CVCB3	3,2255	3,2252	1,5661	1,4924	20,99%	5	1,7826	2,0596	2,1610
CEAB3	0,9011	1,0550	0,7038	0,6698	20,37%	4	1,8147	1,2803	1,5751
CYRE3	2,2214	2,2480	1,2860	1,3248	19,95%	10	0,3206	1,7273	1,6969
BPAN4	4,2155	3,4768	1,3025	1,2328	19,32%	6	0,6703	3,2364	2,8201
ALSO3	1,6952	1,8351	1,4134	1,4218	18,47%	6	0,4419	1,1993	1,2907
CPLE6	0,2998	0,3201	0,2434	0,2354	17,86%	14	0,3310	1,2319	1,3600
CPLE3	0,3012	0,3178	0,2314	0,2248	17,31%	13	0,2616	1,3017	1,4138
CSMG3	0,6482	0,9351	0,5708	0,5766	16,44%	5	1,2754	1,1357	1,6217
AZUL4	7,8756	7,8483	2,6832	2,5505	14,45%	2	3,2466	2,9351	3,0772
TIET11	0,8098	0,9444	0,5363	0,5326	14,26%	7	0,6990	1,5101	1,7733
CPFE3	1,2654	1,2991	1,0636	1,0553	12,02%	8	0,3150	1,1898	1,2310
BBSE3	1,5396	1,5034	0,8027	0,8360	11,84%	6	0,4507	1,9181	1,7982
DTEX3	4,0022	4,4075	1,0871	0,9628	11,82%	1	inf	3,6816	4,5776
BRFS3	2,6665	2,7564	1,4321	1,2513	11,80%	3	0,5773	1,8620	2,2027
CSAN3	3,6771	3,7888	1,0894	1,0817	9,34%	5	0,3913	3,3754	3,5025
BMGB4	0,3273	0,3138	0,2495	0,2515	9,25%	6	0,4336	1,3119	1,2478
CMIG3	1,1759	1,2246	0,6059	0,6026	6,66%	7	0,2123	1,9409	2,0323
CIEL3	0,4675	0,3706	0,2153	0,2006	6,06%	2	5,4502	2,1718	1,8476
CMIG4	1,2216	1,4936	0,4729	0,4877	5,64%	4	0,2901	2,5832	3,0627
BIDI4	8,6557	7,5119	1,7012	1,5703	3,42%	1	inf	5,0881	4,7836
BIDI11	21,3091	19,9959	5,0355	4,6343	2,07%	1	inf	4,2318	4,3148
BBAS3	2,1195	2,1932	1,4623	1,4274	1,61%	7	0,0697	1,4495	1,5365
BTOW3	8,2304	7,8240	5,1607	4,9360	-1,06%	8	0,0357	1,5948	1,5851
CAML33	0,7793	0,7721	0,4702	0,4444	-3,24%	9	-0,0285	1,6574	1,7374
ECOR3	1,5029	1,3857	0,5687	0,5572	-6,85%	3	-0,8666	2,6425	2,4869
CNTO3	4,1117	3,3743	1,6475	1,6305	-13,44%	3	-2,7043	2,4957	2,0695

ANEXO II – ESTRATÉGIA DE MOMENTUM					
Ativo	erro médio quadrático	retorno médio quadrático	rentabilidade acumulada	número de trades	métrica
AALR3	0,8869	0,8107	79,86%	34	1,0941
DIRR3	0,8638	0,7178	62,27%	35	1,2034
BRAP4	4,2393	3,3294	61,14%	33	1,2733
CCRO3	0,6098	0,5673	53,37%	35	1,0750
BKBR3	0,6527	0,6045	43,92%	34	1,0797
CMIG3	0,8375	0,6169	43,61%	34	1,3575
CMIG4	0,6287	0,4869	41,36%	33	1,2911
BMGB4	0,3409	0,2549	33,23%	33	1,3373
ANIM3	1,0127	0,7047	29,22%	35	1,4370
CESP6	1,0713	0,7892	25,43%	34	1,3574
ABCB4	0,6930	0,5671	24,31%	34	1,2221
CNTO3	3,5331	1,6914	24,22%	27	2,0888
CIEL3	0,3970	0,2119	21,43%	36	1,8741
BBAS3	1,9191	1,4999	21,09%	35	1,2795
BTOW3	7,3425	5,1866	20,26%	34	1,4157
CRFB3	0,9913	0,8194	15,86%	34	1,2099
BRPR3	0,4292	0,3536	15,85%	33	1,2136
DMMO3	0,1336	0,0858	15,51%	36	1,5580
BBSE3	1,1858	0,8376	13,61%	33	1,4157
BRML3	0,5914	0,5203	12,58%	35	1,1368
CVCB3	1,9052	1,5196	8,99%	35	1,2537
BBDC3	1,5853	0,8247	7,06%	36	1,9222
CPLE3	0,3000	0,2331	6,87%	33	1,2871
BRFS3	2,3466	1,3372	3,32%	35	1,7549
ALUP11	0,7709	0,6811	1,92%	33	1,1318
ALPA4	2,8003	2,3227	1,30%	34	1,2056
TIET11	1,6119	0,5201	0,90%	28	3,0991
CYRE3	1,7783	1,3517	0,69%	34	1,3156
CSMG3	0,7148	0,5785	-0,95%	33	1,2356
CEAB3	1,0086	0,6804	-1,29%	36	1,4824
RLOG3	1,2315	0,8760	-2,62%	26	1,4058
BRSR6	0,4951	0,4505	-3,13%	32	1,0992
ALSO3	1,6386	1,4375	-4,82%	34	1,1399
CAML33	0,7635	0,4460	-5,72%	34	1,7119
CPFE3	1,2243	1,0847	-5,80%	33	1,1287
DTEX3	7,9659	1,0230	-10,63%	36	7,7869
CSAN3	1,5445	1,0688	-11,08%	36	1,4451
BPAC11	2,1146	1,2534	-12,51%	35	1,6871
BPAN4	3,0261	1,2969	-13,07%	36	2,3333
ECOR3	0,7427	0,5674	-19,71%	34	1,3090
BRKM5	3,9182	2,5906	-22,70%	35	1,5125
BBDC4	2,9058	0,9677	-23,33%	36	3,0029
CPLE6	0,4595	0,2449	-26,78%	36	1,8760
ABEV3	1,2028	0,7212	-27,48%	35	1,6678
AZUL4	4,4186	2,6148	-27,70%	35	1,6899
ARZZ3	6,8365	3,1420	-33,49%	36	2,1759
HGTX3	3,1034	1,7016	-53,83%	36	1,8238
BIDI4	5,8570	1,6646	-57,34%	36	3,5185
BIDI11	14,6954	4,9191	-59,72%	36	2,9874