

PORTFÓLIO DIVERSIFICADO COM MÚLTIPLAS RESTRIÇÕES DE
CARDINALIDADE

Leon Victor Medeiros de Lima

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciência da Computação, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ, como parte dos requisitos necessários à obtenção do grau de mestre.

Orientadores:
Pedro Henrique González Silva

PORTFÓLIO DIVERSIFICADO COM MÚLTIPLAS RESTRIÇÕES DE CARDINALIDADE

Dissertação de Mestrado em Ciência da Computação, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ.

Leon Victor Medeiros de Lima

Aprovada por:

Presidente, Prof. Pedro Henrique González , D.Sc. (orientador)

Glauco Fiorott Amorim, D.Sc.

Laura Silva de Assis, D.Sc.

Israel Mendonça dos Santos, D.Sc.

Rio de Janeiro,
14 de novembro de 2023

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

L932 Lima, Leon Victor Medeiros de
Portfolio diversificado com múltiplas restrições de cardinalidade
/ Leon Victor Medeiros de Lima. — 2023.
51f. : il. color. , enc.

Dissertação (Mestrado) Centro Federal de Educação
Tecnológica Celso Suckow da Fonseca, 2023.

Bibliografia : f. 49-51

Orientador: Pedro Henrique González Silva

1. Finanças pessoais. 2. Aprendizado de Máquina. 3. Redes neurais (Computação). 4. Investimentos. I. Silva, Pedro Henrique González (Orient.). II. Título.

CDD 332. 024

DEDICATÓRIA

As pessoas que não puderam me ver concluir este mestrado. Ao meu irmão Laslo Victor Medeiros de Lima (em memória), ao meu pai Luis Medeiros de Lima (em memória), as minhas amadas tias: Ivoneide Fidélis da Silva (em memória), tia Ninha, e Maria Rita Lima (em memória), tia Rita. Além dessas, a todas as pessoas vítimas da COVID-19.

AGRADECIMENTOS

A Deus.

À Juliana, minha esposa, companheira, meu amor e amiga, por ser a luz e o caminho quando não enxergo e não posso andar. Por ser a força das minhas fraquezas, compartilhar e me apoiar nos melhores e piores momentos.

À minha mãe e ao meu irmão (Júnior), por tudo pelo que passamos, e sem o apoio de vocês, isso não teria sido possível.

Ao meu pai (em memória), Luis Medeiros de Lima, meu espelho, que me ensinou os valores corretos a seguir.

Ao meu amado irmão (em memória), Laslo Victor Medeiros de Lima, que nos deixou imprevisivelmente, levando consigo uma parte de mim. A você, meu irmão, todo o meu amor.

Ao meu professor e orientador, Pedro Henrique González, pela paciência e por me ajudar nesta trajetória.

Aos amigos, Vinícius Siston, que me ajudou muito a estudar nos momentos mais difíceis, Gabriel Herzog, António Pata, Jorge Eugênio, entre outros.

RESUMO

PORTFÓLIO DIVERSIFICADO COM MÚLTIPLAS RESTRIÇÕES DE CARDINALIDADE

Leon Victor Medeiros de Lima

Orientadores:

Pedro Henrique González Silva

Resumo da Dissertação submetida ao Programa de Pós-graduação em Ciência da Computação do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ como parte dos requisitos necessários à obtenção do grau de mestre.

Os mercados financeiros e investimentos desempenham um papel crucial na economia. Além disso, os avanços na tecnologia e a disponibilidade de serviços abertos estão afetando a integração dos mercados financeiros e das instituições financeiras. Isso tem tornado os processos de tomada de decisão financeira ainda mais complexos, à medida que são aplicados métodos precisos para lidar com restrições específicas. O problema de seleção de portfólio (PSP) é um dos problemas mais importantes e relevantes no campo das finanças. O objetivo é escolher um conjunto de ativos que maximizem ou minimizem o retorno ou risco, respectivamente. Este trabalho apresenta uma nova formulação do problema de seleção de portfólio com um único objetivo e restrições de cardinalidade, visando minimizar o risco do portfólio considerando várias categorias de ativos. O objetivo é compor um portfólio multiativo com um número fixo de três categorias de ativos diversas (ações brasileiras, ações americanas e REITs) para minimizar o risco do portfólio, dado um retorno esperado. Além disso, este trabalho utiliza as Redes Neurais Recorrentes de arquitetura *Long Short-Term Memory* (LSTM) e os algoritmos de Regressão de Vetor de Suporte (SVR) para prever o retorno esperado dos ativos. Também apresenta modelos matemáticos para a realização do processo de composição da carteira, juntamente com experimentos computacionais que validam o proposto. Nossos resultados principais validaram o modelo proposto mostrando que é possível obter portfólios diversificados dado um nível de retorno, superando os índices de mercado.

Palavras-chave:

Otimização de Portfólio, Diversificação, Múltiplas Categorias, Aprendizado de Máquina, Redes Neurais Recorrentes - LSTM

Rio de Janeiro,

14 de novembro de 2023

ABSTRACT

PORTFÓLIO DIVERSIFICADO COM MÚLTIPLAS RESTRIÇÕES DE CARDINALIDADE

Leon Victor Medeiros de Lima

Advisors:

Pedro Henrique González Silva

Abstract of dissertation submitted to Programa de Pós-graduação em Ciência da Computação - Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ as partial fulfillment of the requirements for the degree of master.

The Financial markets and investments are crucial to the economy. Furthermore, advances in technology and open services are impacting the integration of financial markets and institutions. This has made financial decision-making processes even more complex, as precise methods are applied to deal with specific constraints. The portfolio selection problem (PSP) is one of the principal and most relevant problems in the finance area. The objective is to choose a set of assets that maximize or minimize return or risk, respectively. This work presents a new formulation of the portfolio selection problem to single-objective with cardinality constraints for minimizing the portfolio risk with multiple categories of assets. Its objective is to compose a multi-asset portfolio with a fixed number of three diverse asset categories (Brazilian stocks and American stocks and REITs) to minimize portfolio risk given an expected return. Additionally, this work uses Recurrent Neural Networks of the Long Short-Term Memory (LSTM) architecture and Support Vector Regressor (SVR) algorithms to predict the expected asset return. It also presents mathematical models for the portfolio composition process, along with computational experiments that validate the proposed approach. Our main results validated the proposed model, showing that it is possible to obtain diversified portfolios given a level of return, outperforming market benchmarks.

Key-words:

Portfolio Optimization, Diversification, Multiple Categories, Machine Learning, Recurrent Neural Networks - LSTM

Rio de Janeiro,

14 de novembro de 2023

Sumário

I	Introdução	12
II	Trabalhos Relacionados	15
III	Referencial Teórico e Formulação Matemática	18
III.1	Mercado de Ações	18
III.2	Retorno de Investimento e Risco da Carteira	19
III.3	Descrição e Fomulação do Modelo Matemático de Média-Variância (MV) de Markowitz	20
III.4	Aprendizado do Máquina	21
III.4.1	Rede Neural Recorrente LSTM	21
III.4.2	Máquinas de Vetores de Suporte	23
III.4.3	Métricas dos Modelos de Predição	24
IV	Métodos	26
IV.1	Predições de Retorno com Referências Móveis	26
IV.2	Modelo Matemático Proposto	29
IV.2.1	Portfólio Diversificado com Múltiplas Restrições de Carnalidade - Modelo de Programação Quadrática	30
IV.2.2	Portfólio Diversificado com Múltiplas Restrições de Carnalidade - Modelo de Programação Linear	31
V	Experimentos e Resultados Computationais	32
V.1	Dados e Configurações	32
V.2	Experimentos	33
V.3	Resultados Computationais	34
V.3.1	Volatilidade	39
V.3.2	Retorno	39
V.3.3	Melhores Carteiras	40
VI	Conclusões	48

Lista de Figuras

III.1 Célula LSTM. Fonte: Graves and Jaitly [2014]	23
III.2 Hierplano SVR. Fonte: Drucker et al. [1996]	25
IV.1 Modelo MLOpt Fonte: Autor	26
V.1 Comparações dos modelos quadrático e linear - modelo LSTM Fonte: Autor	35
V.2 Comparações dos modelos quadrático e linear - modelo SVR Fonte: Autor	39
V.3 Volatilidade das carteiras - modelo LSTM Fonte: Autor	42
V.4 Volatilidade das carteiras - modelo SVR Fonte: Autor	43
V.5 Retorno das carteiras - modelo LSTM Fonte: Autor	44
V.6 Retorno das carteiras - modelo SVR Fonte: Autor	45
V.7 Melhores carteiras em retorno e volatilidade - modelo LSTM Fonte: Autor	46
V.8 Melhores carteiras em retorno e volatilidade - modelo SVR Fonte: Autor	47

Lista de Tabelas

IV.1	Paramêtros testados - rede neural LSTM	27
IV.2	Paramêtros testados do SVR	28
IV.3	Média das métricas dos 300 modelos.	29
V.1	Guia dos experimentos.	34
V.2	Resultados dos experientos - modelo LSTM	36
V.3	Resultados dos experientos - modelo SVR	37
V.4	Retorno e volatilidade - modelo LSTM.	38
V.5	Retorno e volatilidade - modelo SVR.	38

Capítulo I Introdução

No Brasil a busca por melhores investimentos vem se tornando cada vez mais comum entre a população, em principal devido as recentes mudanças econômicas e políticas do país. Uma dessas mudanças foi a reforma da previdência ocorrida em 2019, a qual aumentou o tempo de trabalho, idade e diminuiu diversos outros benefícios, tornando assim a obtenção da aposentadoria pelo serviço social (INSS) um alvo mais difícil de ser alcançado. Outra modificação ocorrida por esta reforma é a dificuldade na obtenção do valor máximo do benefício. Tendo em vista essas e outras mudanças, a aplicação em investimentos, como no mercado financeiro, tornou se essencial para uma estabilidade financeira e bem estar futuros.

Os mercados financeiros se referem amplamente a qualquer mercado onde ocorre a negociação de títulos, incluindo o mercado de ações, mercado de títulos, mercado de câmbio e mercado de derivativos, entre outros. Eles desempenham um papel fundamental no funcionamento das economias capitalistas, oferecendo possibilidades de retorno ao longo do tempo. Talvez os mais comuns dos mercados financeiros sejam os mercados de ações (ou bolsas de valores). Esses são locais onde as empresas listam suas ações, que são compradas e vendidas por *traders* e investidores [Hayes, 2023].

Outro tipo de investimento que normalmente pode ser adquirido nas bolsas de valores são cotas de um fundo de investimento imobiliário (FIIs ou REITs do inglês, *Real State Investment Trust*). Investidores de REITs detêm cotas em um fundo que possui e administra um conjunto de propriedades imobiliárias ou hipotecas, enquanto investidores em ações adquirem ações na propriedade de uma empresa pública [Gresham, 2017].

Uma carteira de investimentos (ou portfólio de investimentos) é um conjunto de ativos negociáveis, como títulos, ações e valores mobiliários, mantidos por um investidor. Logo, a busca por uma composição de um carteira de investimentos é uma questão fundamental na gestão de ativos e está principalmente preocupada em encontrar a melhor combinação deles para atender às necessidades de um investidor, o qual expressa suas preferências e restrições em relação ao risco (propensão a perdas) e à taxa de retorno (tipicamente ganhos em investimentos) esperada da carteira [Pai and Michel, 2009].

Neste sentido, a diversificação é uma medida importante para lidar com o risco de uma carteira de investimentos. Alguns destes riscos estão inter-relacionados, como o crescimento econômico, variações dos ativos na carteira e questões políticas. Já outros riscos estão associados ao ativo, por

exemplo, litígios ou greves repentinas na empresa. Todos esses riscos afetam a flutuação dos ativos (ou seja, a liquidez).

Por isso, ter uma boa estratégia de investimento é vital para investidores e gestores financeiros nos mercados altamente competitivos de hoje. Uma estratégia tradicional de distribuir o capital entre diferentes títulos em vez de investir em um único título pode reduzir o risco do investimento. Por esse motivo, escolher uma combinação favorável de ativos com proporções corretas a serem mantidas em uma carteira é crucial para a diminuição do risco e aumento dos lucros [Kalayci et al., 2020].

Markowitz [1952] apresentou o estudo pioneiro em seu artigo, “Seleção de Portfólio”, em 1952, para resolver o problema de seleção de portfólio usando um modelo de média-variância. Graças a esse trabalho pioneiro na teoria financeira, Markowitz foi premiado com o Prêmio Nobel de Economia em 1991. A teoria está preocupada com dois objetivos: o retorno esperado e o risco de seus respectivos investimentos. O objetivo é minimizar o nível de risco para um determinado nível de retorno esperado, o que resulta na fronteira eficiente irrestrita, onde cada ponto representa o nível mínimo de risco a ser assumido para um retorno esperado. Isso deu origem ao chamado *Problema de Otimização de Portfólio* (POP), no qual busca-se a minimização do risco com base em um retorno mínimo esperado exigido pelo investidor [Doering et al., 2016].

Para Kolm et al. [2014] um princípio importante neste trabalho é a diversificação de portfólio, que se baseia na ideia de que o risco de um portfólio depende das correlações entre seus componentes, não apenas do risco médio de suas participações separadas.

Uma outra crítica ao modelo de Markowitz, segundo Pai and Michel [2009], é que ele assume um mercado perfeito, ignorando custos de transação e impostos, onde a venda a descoberto é proibida e os ativos são negociáveis em frações não negativas. Além disso, o modelo de Markowitz considera o retorno dos ativos como uma distribuição normal, o que raramente ocorre na prática [Chang et al., 2009].

Após o estudo de Markowitz, surgiram diversos outros modelos de seleção de carteiras que utilizam suas premissas básicas, e que também buscaram aprofundar se e trazer mais condições realistas em sua aplicação. Alguns exemplos são: *i) Semi-Variance* (S-V) [Markowitz, 1959], *ii) Variance with Skewness* (VwS) [Samuelson, 1970], *iii) Mean-Absolute Deviation* (MAD) [Konno and Yamazaki, 1991], *iv) Value-at-Risk* (VaR) [Jorion, 1997], *v) Minimax* (MM) [Young, 1998], entre outros [Chagas, 2023].

Dentre as características estudadas incluí-se, a restrição do número de ativos na carteira (cardinalidade) e restrições de pré-atribuição que refletem as preferências do investidor. No entanto, devido à sua complexidade, ao considerar instâncias realistas, esse problema geralmente se torna NP-difícil [Panadero et al., 2020].

Além disso, todos estes modelos de portfólios assumem que as séries de retornos de cada ativo seguem uma distribuição Normal, e geralmente adotam a média do retorno histórico dos ativos como o retorno esperado, o que produz uma inadequada precisão sobre os retornos futuros dos ativos [Ma et al., 2020].

Entretanto, nas últimas décadas ocorreu o surgimento de vários métodos de aprendizagem de máquina para predições de séries temporais, tais como as redes neurais artificiais, as máquinas de vetor de suporte, entre outros [Freitas et al., 2009].

A proposta desta pesquisa é desenvolver uma formulação que permita a construção de uma carteira ótima com base em um número fixo de três categorias de ativos distintas, utilizando modelos de aprendizado de máquina para predição do retorno futuro esperado, visando minimizar o risco da carteira, dado um retorno mínimo esperado.

O objetivo desta pesquisa é contribuir para a literatura por meio da formulação do Problema de Seleção Portfólio (PSP) com a inclusão de categorias distintas de ativos, bem como utilizando a aplicação de modelos de aprendizado de máquina para predição de retorno de preços futuros dos ativos. Para isso, são utilizados dados históricos dos índices financeiros do Brasil (BR) e dos Estados Unidos da América (EUA) no período de 2014 a 2021. Os resultados do método proposto obtiveram volatilidades menores e retornos maiores em comparação com os índices da Bovespa (BR), Nasdaq (EUA) e FTSE (EUA). Essa abordagem busca fornecer informações valiosas sobre a construção de carteiras de investimentos ótimas, levando em consideração múltiplas categorias de ativos e utilizando dados reais dos mercados financeiros do Brasil e dos Estados Unidos.

Após esta introdução, o Capítulo II aborda a revisão da literatura nas áreas correlacionadas a este trabalho, destacando os desenvolvimentos recentes na aplicação de modelos de otimização em conjunto de técnicas de predição para o retorno de ativos. O Capítulo III, descreve os referenciais teóricos de risco e retorno, e a formulação matemática modelo de Markowitz, além de incluir uma breve definição e apresentar os modelos utilizados nesta pesquisa. O Capítulo IV mostra os métodos e a formulação matemática do método utilizado nas avaliações experimentais realizadas neste trabalho. O Capítulo V discute os dados, experimentos e resultados empregados na avaliação do desempenho do modelo de seleção de carteira. Por fim, o Capítulo VI aborda as conclusões e perspectivas de trabalhos futuros.

Capítulo II Trabalhos Relacionados

O surgimento de novas tecnologias, os avanços computacionais, a utilização e disponibilidade massiva de dados, vem contribuindo e potencializando a utilização de técnicas de inteligência artificial e computacional na exploração de novas abordagens para melhores estimativas do retorno esperado e risco na construção de portfólios otimizados. Deste modo, diversos estudos e abordagens tem sido aplicados na tentativa de aproximar e/ou encontrar novas abordagens baseadas no modelo inicial da média-variância. A seguir, são discutidos alguns trabalhos relacionados de modelos de otimização utilizando técnicas de predição de retorno de ativos.

Os modelos de otimização de portfólio baseados em previsão são uma classe de modelos que usam previsões de retornos futuros de ativos para otimizar a composição de um portfólio de investimento. Esses modelos são usados por investidores para maximizar o retorno esperado do portfólio e minimizar o risco.

Um algoritmo para remoção de *outliers* em previsões de matrizes de covariâncias foi proposto por Hellström [2000]. O algoritmo produziu uma redução significativa no erro de previsão das matrizes de covariâncias de 24 ativos do mercado de ações sueco e 29 ativos do índice americano Dow Jones. Em seguida, o autor otimizou carteiras com base nestas previsões e verificou que a remoção dos *outliers* não afetou significativamente o cálculo das carteiras ótimas.

O estudo de de Freitas et al. [2006], propõe um novo modelo de otimização de portfólio que usa uma rede neural autorregressiva para prever os retornos esperados das ações. O modelo proposto usa os erros de previsão da rede neural para calcular o risco do portfólio. Os resultados experimentais do estudo mostraram que o modelo proposto superou o modelo de variância média (do inglês, *Mean-Variance* - MV), um modelo simples que usa a média dos retornos passados das ações para prever os retornos futuros, gerando um retorno melhor para o mesmo risco.

Além desse estudo, Freitas et al. [2009] formula um modelo de otimização de portfólio para a seleção de carteiras com o objetivo de minimização do risco, o qual é medido como a variância da combinação linear do retorno dos seus ativos. Em seu trabalho, Freitas et al. [2009] utiliza um modelo de rede neural autorregressiva móvel de referência (do inglês, *Autoregressive Moving Reference Neural Network* - AR-MRNN) para prever os retornos futuros desses ativos. Os resultados experimentais do estudo mostraram que o modelo de otimização de portfólio proposto superou o modelo de variância média clássico.

Os estudos de Freitas et al. [2006] e Freitas et al. [2009] fornecem evidências de que os modelos de otimização de portfólio baseados em previsão podem ser uma ferramenta eficaz para os investidores que desejam maximizar o retorno de seus investimentos e minimizar o risco.

Hao et al. [2013] propuseram um modelo de seleção de portfólio baseado em previsão. O modelo usa SVR para prever os retornos futuros das ações e a variância dos erros de previsão como risco para otimizar o portfólio. Os autores compararam seu modelo com outro presente na literatura e mostraram que seu modelo teve um desempenho melhor. Eles também mencionam que uma melhor previsão do retorno futuro das ações resultou em um melhor desempenho do modelo.

Krauss et al. [2017] implementou e comparou a performance de alguns modelos, dentre eles: *Gradient-Boost Tree*, redes neurais de multilaminadas (do inglês, *Multilayer Perceptron* - MLP), *Random Forest*, entre outros para arbitragem estatística. Foram construídos portfólios com os resultados preditivos desses modelos, investindo nas k ações com as maiores previsões de retorno e vendendo a descoberto as k ações com as menores previsões de retorno. Os experimentos mostraram que um portfólio baseado em um modelo *ensemble* que combina MLP, *gradient-boost tree* e *random forest* gerou um retorno médio diário de 0,45% sem considerar as taxas de transação.

O estudo de Chong et al. [2017] analisou três métodos de extração de recursos não supervisionados para redes neurais profundas: análise de componentes principais, *autoencoder* e máquina de Boltzmann restrita. Os resultados sugerem que as redes neurais profundas podem extrair informações adicionais dos resíduos do modelo autorregressivo e melhorar o desempenho da previsão. Além disso, a estimativa de covariância também é notavelmente melhorada quando a rede preditiva é aplicada à análise de estrutura de mercado baseada em covariância. O estudo conclui que as redes neurais profundas são uma ferramenta promissora para análise e previsão do mercado de ações.

Fischer and Krauss [2018] descobriram que a rede neural LSTM é um modelo eficaz para prever o retorno futuro das ações. Eles construíram um portfólio com base nos resultados preditivos da rede neural LSTM e descobriram que esse portfólio superou o mercado geral de 1992 a 2009. No entanto, o portfólio se deteriorou em 2010 com os retornos da LSTM oscilando em torno de zero.

Yang et al. [2019] desenvolveram um modelo de previsão de ações baseado em máquina de aprendizagem extrema. Eles usaram o retorno preditivo para construir um modelo de otimização de portfólio que também considerava outros índices técnicos. Eles usaram algoritmos de evolução diferencial para resolver o problema de otimização de portfólio. Os resultados do estudo, que usaram dados do mercado de ações da China, mostraram que o modelo proposto teve um desempenho melhor do que os métodos tradicionais, o que sugere que a previsão de ações pode ser uma ferramenta eficaz para a seleção de ações.

Lee and Yoo [2020] foram os primeiros a comparar três tipos de redes neurais recorrentes para prever o desempenho das ações: simples, *Gated Recurrent Unit* (GRU) e LSTM. Os resultados

mostraram que a rede neural LSTM é a mais precisa. Eles também propuseram uma nova estratégia de investimento baseada nos resultados da rede neural LSTM, que também foi bem-sucedida.

Ma et al. [2020] propõem um novo modelo de otimização de portfólio baseado em previsões que usa três redes neurais profundas (do inglês, *Deep Neural Networks* - DNN): Perceptron Multicamadas Profundo (do inglês, *Deep Multilayer Perceptron* - DMLP), LSTM e Redes Neurais Convolucionais (do inglês, *Convolution Neural Networks* - CNN). As DNNs são usadas para prever os retornos futuros das ações, e os erros de previsão das DNNs são usados para medir o risco das ações. O modelo de otimização de portfólio proposto é construído integrando os retornos preditivos e o semi-desvio absoluto dos erros preditivos. Os resultados experimentais mostram que o modelo de otimização de portfólio proposto tem o melhor desempenho entre os modelos comparados sob diferentes retornos de portfólio desejados. O trabalho apresentado em Ma et al. [2020] conclui que as DNNs podem ser usadas para construir modelos de otimização de portfólio baseados em previsões que têm um desempenho superior aos modelos tradicionais.

Pode-se observar que os trabalhos apresentados na literatura utilizam diferentes métodos de seleção de portfólio, alguns destes usam modelos de predição atuando na predição de retornos e na seleção de portfólios, enquanto outros trabalhos combinam modelos de predição com os de otimização. Além disso, a análise realizada por essas pesquisas são principalmente direcionadas para a avaliação e comparação dos modelos de predição de retorno utilizados. Embora muitos destes modelos sejam aplicados para minimização do risco, eles não fazem uso de categorias distintas de ativos como um fator determinante para diminuição do risco. Um outro ponto importante e diferente das literaturas, neste trabalho é avaliado a qualidade das soluções encontradas pelos métodos propostos, bem como a comparação de tais métodos.

Capítulo III Referencial Teórico e Formulação Matemática

Este capítulo apresenta os conceitos de mercado de ações, retorno e risco de um portfólio de investimento e a definição matemática do modelo média-variância de Markowitz para o problema de seleção de portfólio (PSP). Por fim, descreve brevemente sobre os modelos de rede neurais LSTM e aprendizado de máquina SVR e suas metrificações utilizadas neste trabalho.

III.1 Mercado de Ações

As empresas desempenham um papel fundamental na organização socioeconômica de uma sociedade e têm várias funções e responsabilidades, como por exemplo: geradoras de emprego, produção de bens e serviços [Sá, 1987].

Uma empresa é uma organização particular, pública ou mista (pública e privada) com busca à obtenção de lucros, e é constituída, de modo mais comum, na forma de sociedades limitadas e sociedades anônimas. As sociedades limitadas são constituídas por meio de contrato social, contrato entre sócios, a qual estabelece as participações deles na empresa. As sociedades anônimas, o capital da empresa é dividido em ações e a responsabilidade dos sócios (detentores das ações) é proporcional à quantidade de suas cotas.

As ações de uma empresa são valores mobiliários emitidos pelas companhias, que representam as menores parcelas nas quais se divide o capital da companhia [Sá, 1987]. As ações podem ser negociadas no mercado primário, quando uma companhia emite uma quantidade de ações, ou no mercado secundário, as ações são negociadas nas Bolsas de Valores ou mercado de Balcão.

Outro tipo de valor mobiliário e também negociado nas Bolsas de Valores é o REIT, ou *Real Estate Investment Trust*. O REIT é uma forma de investimento que permite aos investidores participar no mercado imobiliário sem ter que comprar propriedades diretamente. Os REITs são empresas que possuem, operam ou financiam imóveis ou outros ativos relacionados ao setor imobiliário. Eles oferecem aos investidores uma maneira de investir em imóveis comerciais ou residenciais, bem como em ativos como hipotecas lastreadas em imóveis.

A principal bolsa de valores brasileira é a Bolsa de Valores de São Paulo - B3 - Brasil, Bolsa, Balcão. Nos Estado Unidos, as principais bolsas de valores são NASDAQ (*National Association of Securities Dealers Automated Quotations*) e a NYSE (*New York Stock Exchange*).

III.2 Retorno de Investimento e Risco da Carteira

Um investimento é obtido a partir de um período de tempo denominado *horizonte de investimento*. Nele um capital inicial é investido, C_i , e após um determinado tempo é recuperado, C_f . Este capital final que é relativo ao investido inicialmente é uma medida de desempenho do investimento realizado, pois pode ser comparado seu valor final com o de início, no entanto torna-se sem utilidade para investimentos em que o capital inicial aportado é diferente. Deste modo, é necessário que uma forma independente do capital inicial seja utilizada para avaliar a variação do capital ao longo do horizonte de investimento, essa medida é denominada *retorno do investimento*, ou apenas *retorno*. Para uma ativo (ações, por exemplo), o *retorno da ação* pode ser definido conforme a Equação III.1, onde P_t e P_{t-L} são os preços da ação nos tempos t e $t - L$:

$$r_t = \frac{P_t - P_{t-L}}{P_{t-L}} \quad (\text{III.1})$$

Uma carteira (ou portfólio) é uma conjunto de ativos, em que cada ativo possui seu respectivo peso. Cada participação X_i , $i = 1, \dots, M$, com $0 \leq X_i \leq 1$, e $\sum X_i = 1$, representa a fração do valor da carteira investido na ação i . O retorno realizado da carteira no tempo t , r_t , é dado pela combinação das participações e dos retornos individuais dos ativos, em que M é o número total de ativos na carteira, X_i é a participação da ação i na carteira e r_{pt} o retorno realizado do ativo i no tempo t [de Freitas et al., 2006], e está definido na Equação III.2:

$$r_{pt} = \sum_{i=1}^M X_i r_{it} \quad (\text{III.2})$$

A proposição original de Markowitz trata o retorno esperado de uma ação como a média aritmética da série de retornos desta ação, onde \hat{r} é o retorno esperado da ação no tempo $t = N + 1$, onde N é o tamanho da série histórica e r_t o retorno realizado no tempo t , definida na Equação III.3 como:

$$\hat{r} = \frac{1}{n} \sum_{i=1}^N r_t \quad (\text{III.3})$$

Da mesma forma, o retorno esperado da carteira é definido como a combinação linear das participações e dos retornos esperados dos seus ativos [de Freitas et al., 2006].

O risco da carteira proposto por Markowitz é a variância da combinação linear das participações dos ativos na carteira e das suas séries de retornos, definida em termos da covariância entre cada par de ativo. O risco total da carteira é composto pela contribuição de cada risco individual de cada ativo na carteira e pela contribuição do risco de cada par de ativo combinado. O efeito da covariância dos pares de ativos no risco da carteira é fundamental para a obtenção de carteiras com

risco inferior à soma ponderada dos riscos individuais das ações. Essa é a principal contribuição do modelo de média-variância de Markowitz [Freitas et al., 2009]. A composição do risco total da carteira é mostrada na Eq. III.4, em que V é o risco total da carteira, σ^2 é a variância individual de cada ativo i , e γ_{ij} a covariância dos pares de ativos i e j .

$$V = \sum_{i=1}^M \sigma_i^2 X_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{ij} \quad (\text{III.4})$$

III.3 Descrição e Formulação do Modelo Matemático de Média-Variância (MV) de Markowitz

O problema clássico de otimização de portfólio, o modelo Média-Variância (MV) de Markowitz, é um problema de otimização de objetivo único com uma das seguintes formulações do modelo: o investidor minimiza a exposição ao risco sujeito a um retorno esperado mínimo atingível ou o investidor maximiza o retorno esperado para um determinado nível de risco [Doering et al., 2016]. Ambos os casos usam a covariância como medida de risco. Além disso, o problema pode ser formulado como um problema de otimização multiobjetivo para maximizar o retorno esperado e minimizar o risco de um portfólio. O modelo Markowitz de objetivo único é descrito em Di Tollo and Roli [2008] em sua forma canônica do seguinte maneira:

$$\min \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \quad (\text{III.5})$$

$$\text{subject to: } \sum_{i=1}^n r_i x_i \geq R_p \quad (\text{III.6})$$

$$\sum_{i=1}^n x_i = 1 \quad (\text{III.7})$$

$$x_i \geq 0 \quad i = 1, \dots, n, \quad (\text{III.8})$$

onde n é o número de ativos disponíveis, r_i é o retorno esperado mínimo do ativo i , R_p representa o retorno esperado mínimo da carteira, σ_{ij} é a covariância entre os ativos i e j , e x_i é a variável de decisão que representa a proporção de dinheiro investido no ativo i . A Equação (III.5) minimiza o risco da carteira (medido pela covariância) enquanto o retorno esperado é dado pela Equação (III.6), a qual é a restrição do retorno desejado que assegura o resultado do portfólio. A Equação (III.7) define a restrição orçamentária e garante que a soma dos pesos dos ativos seja igual a um. A Equação (III.8) afirma a não negatividade para todos os pesos. Esta formulação (Equações (III.5) - (III.8)) é um problema simples de programação quadrática que pode ser resolvido por métodos numéricos clássicos.

III.4 Aprendizado do Máquina

As técnicas de aprendizado de máquina (do inglês, *Machine Learning* - ML) utilizam um princípio de inferência chamado indução, no qual obtém-se conclusões genéricas a partir de um conjunto particular de exemplos [Lorena and De Carvalho, 2007].

No aprendizado supervisionado, um professor externo fornece o conhecimento do ambiente por meio de conjuntos de exemplos, onde cada exemplo é composto por uma entrada e uma saída desejada [Haykin, 2010].

No aprendizado não-supervisionado não há a presença de um professor, ou seja, não existem exemplos rotulados. Em vez disso o algoritmo busca agrupar as entradas de dados submetidas de acordo com uma medida de qualidade [Lorena and De Carvalho, 2007].

O *Deep learning* (aprendizado profundo) é um subcampo de *machine learning* (ML) (aprendizado de máquina) são um paradigma de computação que se inspira no modelo computacional do cérebro humano. Portanto, é um método de aprendizado de máquina que se utiliza de uma rede neural artificial para aprender por meio de exemplos.

As redes neurais artificiais são construídas a partir de modelos matemáticos de componentes do cérebro humano, e conseguem reproduzir, em determinados níveis, diversas propriedades deste. Entre as propriedades mais importantes nas redes neurais, destacam-se: aprendizado, não-linearidade, plasticidade e generalização [Freitas et al., 2009].

As redes neurais são construídas a partir da interconexão de neurônios artificiais organizados em camadas que recebem como entrada as saídas das camadas anteriores, denominadas *sinapse*, e projetam suas saídas na camada posterior. Em uma rede neural de N camadas, as camadas iniciais e finais são denominadas camadas de entrada e saída, enquanto as camadas intermediárias são chamadas de camadas ocultas. O nome desta organização é denominada de *topologia* ou *arquitetura* [Haykin, 2010].

Para que uma rede neural possa ser utilizada em uma tarefa específica, como na predição de séries temporais por exemplo, ela necessita adquirir previamente o conhecimento necessário. O treinamento supervisionado. Desta forma, o erro entre a saída desejada e a saída produzida pela rede neural pode ser calculado para cada par de treinamento. O algoritmo *back-propagation*, apresentado por Rumelhart et al. [1986], propaga o erro da saída da rede no sentido inverso das camadas, calculando em cada camada os ajustes dos pesos para minimizar o seu erro.

III.4.1 Rede Neural Recorrente LSTM

Redes Neurais Recorrentes (do inglês, *Recurrent Neural Networks* - RNNs), têm demonstrado grande sucesso na resolução de tarefas que envolvem sequências de dados, como fala, processamento

de linguagem natural, previsão de séries temporais e muito mais [Fischer and Krauss, 2018]. Entre as variantes das RNNs, as de arquitetura *Long Short-Term Memory* (LSTM) se destacam devido à sua habilidade notável em capturar dependências de longo prazo.

Apresentada em Schmidhuber et al. [1997], a arquitetura de uma unidade LSTM é construída em torno de uma célula de memória, que é uma parte fundamental do seu funcionamento. Essa célula possui portas que controlam o fluxo de informações ao longo do tempo, permitindo assim uma melhor previsão dos estados futuros. As portas regulam o fluxo de informações de forma eficaz. Para cada passo de tempo, um neurônio LSTM recebe uma entrada x_t , e dentro da célula, cada porta (ou *gate*) utiliza unidades de ativação para determinar se devem ser ativadas ou não. Isso influencia a mudança de estado e a incorporação de novas informações à célula.

Dada uma sequência de entradas $x = (x_1, \dots, x_T)$, uma RNN padrão computa a sequência do vetor oculto, $h = (h_1, \dots, h_T)$, e a sequência do vetor de saída, $y = (y_1, \dots, y_T)$, interagindo pelas equações de $T = 1$ até T :

$$h_t = H(W_{ih}x_t + W_{hh}h_{t-1} + b_h) \quad (\text{III.9})$$

$$y_t = W_{ho}h_t + b_o \quad (\text{III.10})$$

em que o termo W denota o peso das matrizes (por exemplo: W_{ih} é a matriz de pesos da entrada oculta), o termo b denota o vetor viés (por exemplo: b_h é o vetor de viés oculto) e H é a função de ativação da camada oculto, a qual é implementada pela composição das Equações (III.11) - (III.15):

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (\text{III.11})$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (\text{III.12})$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (\text{III.13})$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (\text{III.14})$$

$$h_t = o_t \tanh(c_t) \quad (\text{III.15})$$

onde σ , é a função sigmóide logística, e i , f , o e c são respectivamente os vetores de ativação do portão de entrada, do portão de esquecimento, do portão de saída e da célula, todos do mesmo tamanho que o vetor oculto h . Os subscritos da matriz de pesos têm um significado óbvio, por exemplo W_{hi} é a matriz de portão de entrada oculta, W_{xo} é a matriz de portão de saída de entrada etc. As matrizes de pesos da célula para os vetores de portão (por exemplo, W_{ci}) são diagonais, portanto, o elemento m em cada vetor de portão só recebe entrada do elemento m do vetor de

célula. Os termos de viés (que são adicionados a i , f , c e o) foram omitidos para maior clareza [Graves and Jaitly, 2014].

O parâmetro c_t representa o estado da célula no momento t , retendo informações importantes das etapas anteriores. A porta de esquecimento f_t decide quais informações devem ser descartadas da célula, enquanto a porta de entrada i_t controla a atualização do estado da célula com base nas entradas. Por fim, o gate de saída o_t decide qual saída deve ser produzida com base na entrada atual e na memória da célula [Graves and Jaitly, 2014]. A Figura (III.1) mostra a configuração de uma célula LSTM.

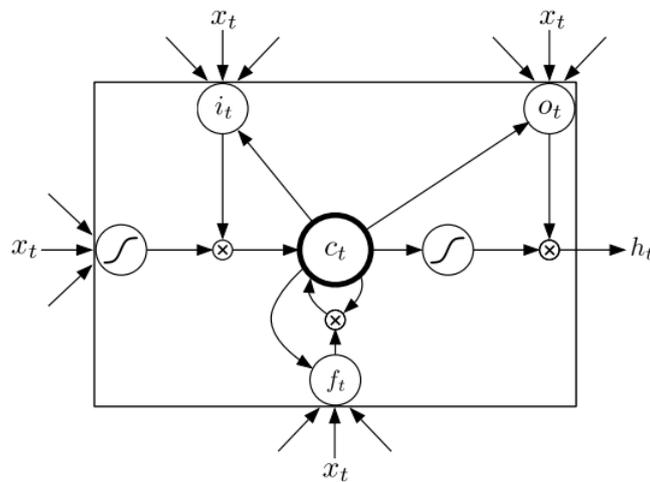


Figura III.1: Célula LSTM.
Fonte: Graves and Jaitly [2014]

III.4.2 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (SVMs) (do inglês, *Support Vector Machines*) constituem uma técnica de aprendizado que foi primeiramente desenvolvida por Vapnik [1995] na Teoria de Aprendizado Estatístico. Seus resultados são comparáveis e em muitas vezes superiores aos obtidos por algoritmos como das redes neurais para aplicações de regressão e classificação, especialmente em séries temporais [Yang et al., 2002].

As SVMs são máquinas de aprendizado que implementam o princípio indutivo de minimização do risco estrutural para obter boas generalizações em um limitado número de padrões de aprendizagem. Uma SVM constrói um hiperplano ou um conjunto de hiperplanos em um espaço dimensional grande ou infinito [Yang et al., 2002].

Algumas das principais características das SVMs que tornam seu uso atrativo são: (i) boa capacidade de generalização, (ii) robustez em grandes dimensões e (iii) convexidade da função objetivo. Entre as características citadas, o destaque das SVMs está em sua capacidade de generalização.

A técnica do SVR foi introduzida por Drucker et al. [1996] para aplicação do SVM em problemas

de regressão. O SVR é um método de regressão não linear baseado na função *Kernel*, o qual tenta localizar um hiperplano de regressão com o menor risco em um espaço com características de alta dimensão [Ma et al., 2020].

Eles (os SVRs) são configurados pelo princípio da minimização do risco estruturado ao invés da minimização do risco empírico. Portanto, o risco está apenas relacionado com o número de amostras de entrada e os seus parâmetros de estrutura determinados automaticamente a partir da amostra [Hao et al., 2013]. Assim, o SVR deve usar uma função de custo para medir o risco empírico, a fim de minimizar o erro de regressão.

Neste sentido, o método de SVR é baseada na análise de regressão que tem como objetivo gerar o número máximo de vetores de suporte com valores pequenos de erros a fim de separar os dados da melhor forma possível [Drucker et al., 1996]. Uma de suas características distintivas é a capacidade de ser insensível a erros dentro de uma determinada margem de tolerância (ϵ). Isso significa que o algoritmo permite que os pontos de dados se afastem um pouco da linha de regressão (hiperplano) sem afetar significativamente a qualidade do modelo. Caso o valor observado esteja acima das previsões, ξ_i é a diferença positiva entre o valor observado e a margem de tolerância. Caso contrário, ξ_i^* é a diferença negativa. A Figura (III.2) exhibe o hiperplano do SVR.

A Equação (III.16) define a função de base radial como a função de *Kernel* utilizada neste trabalho. Onde γ é o parâmetro da função de base radial, v_i e v_j é a distância Euclidiana de dois pontos. O SVR ainda contém o hiperparâmetro de regularização C , o qual controla o equilíbrio entre a busca por um modelo com menor erro de treinamento e a busca por um modelo com menor margem de erro. Maiores detalhes sobre SVM estão disponíveis em Drucker et al. [1996] e Lorena and De Carvalho [2007].

$$K_{(v_i, v_j)} = \exp(-\gamma \|v_i - v_j\|^2) \quad (\text{III.16})$$

III.4.3 Métricas dos Modelos de Predição

A metrificação do desempenho dos preditores é efetuada considerando os valores dos dados reais dos preços (valores das séries temporais) e o valor predito pelo modelo e é avaliado no estágio de validação.

O Erro Médio Absoluto (do inglês, *Mean Absolute Error* - MAE), é a média aritmética da soma dos erros absolutos, ou seja, mede a diferença absoluta entre a soma do valor real e o valor predito, e é determinado como:

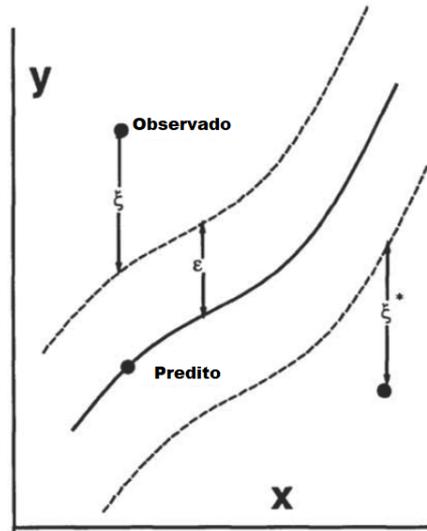


Figura III.2: Hierplano SVR.
Fonte: Drucker et al. [1996]

$$\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (\text{III.17})$$

onde n é o tamanho das séries históricas, e y_t e \hat{y}_t são o preço real (realizado) e o preço predito, respectivamente, no tempo t . O erro MAE é uma medida comum de erro de previsão na análise de séries temporais.

O Erro Quadrático Médio (MSE) (do inglês, *Mean Squared Error*) ou Desvio Quadrático Médio (MSD) (do inglês, *Mean Squared Deviation*) mede a diferença quadrática média entre os valores estimados (preditos) e o valor real. Como é derivado do quadrado da distância euclidiana, o MSE é sempre um valor positivo que diminui à medida que o erro se aproxima de zero. O MSE é estabelecido como:

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (\text{III.18})$$

O erro MAPE (do inglês *Mean Absolute Percentage Error*, ou Erro Percentual Absoluto Médio, é uma métrica de erro que mede a diferença percentual média entre os valores estimados e os valores observados. É uma medida de desempenho muito utilizada para avaliar a precisão de um modelo de previsão devido à sua interpretação muito intuitiva em termos de erro relativo, e é caracterizado como:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (\text{III.19})$$

Capítulo IV Métodos

Neste capítulo é apresentado o modelo de otimização de carteiras com múltiplas restrições de cardinalidade proposto neste trabalho. Este modelo de otimização atua em conjunto com um modelo de aprendizado de máquina que é utilizado para efetuar as predições dos retornos futuros dos ativos.

A combinação de algoritmos de aprendizado de máquina e modelo de otimização pode levar a soluções mais eficazes e eficientes na construção de portfólios diversificados que minimizam os riscos para um determinado nível de retorno esperado, pois os modelos de aprendizado de máquina podem capturar padrões complexos nos dados (na predição do desempenho de ativos), enquanto os modelos de otimização podem tomar decisões precisas com base nessas informações.

O *workflow* exibido na Figura IV.1 apresenta a combinação dessas técnicas utilizadas neste trabalho, na qual as predições de retornos da saída do modelo de aprendizado de máquina alimentam o modelo de otimização, o qual realiza a seleção de cada subcarteira, assim gerando o portfólio de investimento.

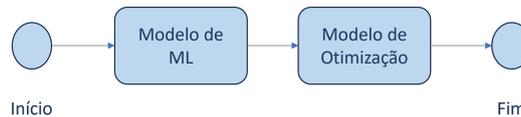


Figura IV.1: Modelo MLOpt
Fonte: Autor

Neste capítulo descrevemos o treinamento dos modelos de predição LSTM e SVR utilizados e o modelo matemático proposto para este trabalho formando assim nossa solução MLOpt.

IV.1 Predições de Retorno com Referências Móveis

O treinamento dos modelos de predição LSTM e SVR para os retornos dos ativos foram realizados com base nas configurações e aplicação dos modelos apresentando no estudo de Ma et al. [2020].

A rede neural LSTM, como a maioria das redes neurais profundas, contém diversos hiperparâmetros. Deste modo e conforme adotado no trabalho de Ma et al. [2020], para a busca da melhor combinação de hiperparâmetros foi utilizado a técnica *Grid Search*, a qual consiste em uma

busca exaustiva de combinação. Para redução do *overfitting* (quando um modelo não consegue ser generalista, ou seja, obter boas previsões em novos dados) foi utilizado a técnica *early stopping*, a qual basicamente interrompe o treinamento caso não haja melhora nas métricas de avaliação para os dados de validação. A Tabela IV.1 exibe os parâmetros avaliados para a construção do modelo.

Tabela IV.1: Parâmetros testados - rede neural LSTM

Parâmetros	Valor
Nós ocultos	{5, 10, 20, 50}
Camadas ocultas	{1}
Taxa de aprendizagem	{0,0001; 0,001; 0,001; 0,1}
Patience	{0, 5, 10}
Batch size	{32, 50, 100, 200}
Taxa de dropout	{0; 0,1; 0,2; ... ;0,5}
Taxa de dropout recorrente	{0; 0,1; 0,2; ... ;0,5}
Função de ativação	{relu; tanh}
Função de perda	{Erro quadrático médio (MSE)}

Após uma análise exaustiva, definiu-se as configurações dos parâmetros escolhidos que obtiveram os menores erros entre as combinações testadas, $MSE = 0,025$, $MAE = 0,1258$ e $MAPE = 9,754\%$, a função de ativação, que são aplicadas em diferentes partes das unidades LSTM para controlar o fluxo de informações e atualização da memória, como sendo a tangente hiperbólica (*Tanh*).

A quantidade de nós, que indica quantas unidades LSTM ou células de memória a rede terá, foi configurada para 50 unidades, enquanto o número de camadas utilizadas na rede LSTM é configurado para 1.

O parâmetro de *dropout* e *dropout* recorrente foram configurados para 0. O *dropout* é usado para controlar o *overfitting*, desativando aleatoriamente uma fração das unidades LSTM durante o treinamento. Enquanto o parâmetro *dropout* recorrente, é uma técnica de regularização que envolve desativar aleatoriamente um conjunto de unidades (neurônios) durante cada passo de tempo da sequência durante o treinamento.

O parâmetro *patience*, no *early stopping*, foi configurado para 0. Ele é o número de épocas adicionais que você permite que o treinamento continue após a melhora da métrica de validação ter cessado.

A taxa de aprendizagem em uma LSTM é um parâmetro crítico que controla o tamanho dos passos durante o treinamento. Neste modelo ela foi ajustada para 0,001.

O otimizador escolhido foi o *Adam*. O otimizador define o algoritmo de otimização usado para ajustar os pesos da rede durante o treinamento.

A quantidade de épocas foi definida igual 50. A quantidade de épocas representa o número de vezes que todo o conjunto de treinamento é passado pela rede durante o treinamento.

O *batch size*, especifica quantos exemplos de treinamento são processados em cada passo de atualização dos pesos da rede. O número de *batch* foi ajustado para 32.

A função de perda (também conhecida como função de custo ou função objetivo) em uma rede neural LSTM é uma métrica que quantifica o quão bem o modelo está performando em relação à tarefa que está sendo realizada. Neste modelo ela é configurada como o erro quadrático médio (MSE), pois ele penaliza erros maiores de forma mais acentuada do que erros menores devido à sua natureza quadrática e trabalha na mesma e gera resultados na mesma unidade das variáveis dependentes.

O modelo SVR é uma técnica clássica de aprendizado de máquina, a qual é utilizada amplamente para realizar previsões de preço devido a sua capacidade de lidar com relações complexas e não lineares entre as variáveis de entrada e as de saídas. Para este modelo é utilizado a função de base radial como a função de *kernel* do SVR. As funções de base radial são frequentemente usadas em redes neurais artificiais para aproximar funções e realizar classificação [Ma et al., 2020]. Os principais parâmetros do SVR são os parâmetros de regularização, C , e o parâmetro da função de base radial, γ . Após diversas tentativas e também utilizar a técnica de *Grid Search*, obtivemos as melhores métricas de erros com $MSE = 0.006$, $MAE = 0.065$ e $MAPE = 4,899\%$ nas seguintes configurações: $C = 2^4$ e $\gamma = 2^{-5}$. A Tabela IV.2 exhibe os parâmetros testados na construção do modelo SVR.

Tabela IV.2: Paramêtros testados do SVR

Parâmetros	valor
C	$2^1, 2^2, 2^3, 2^4$
γ	$2^{-1}, 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}$

Nós construímos um modelo de LSTM e um modelo de SVR para cada ativo dos 150 ativos utilizados nos experimentos e que foram selecionados aleatoriamente do conjunto de dados coletados, maiores informações sobre a seleção desses ativos são relatadas no Capítulo V. Nós utilizamos os valores semanais para cada uma das 150 ações utilizando a série histórica de preços diários dos fechamentos observados às quarta-feiras no período dos anos de 2014 até 2017. Assim obtivemos 150 séries temporais, cada uma com 208 observações de preços semanais. A escolha da amostra no meio da semana ocorreu para evitar os efeitos do início e do fim da semana nos preços das ações [Elton et al., 2009]. Por se tratar de séries temporais de ativos de mercados distintos, os dados faltantes foram preenchidos com os valores passados mais próximos.

O treinamento e predição do preço da ação no futuro com cada modelo utilizou uma janela temporal deslizante com 10 semanas dos 208 preços semanais, período de 2014 até 2017. Estas janelas foram deslocadas até completar o ciclo total de 208 semanas disponíveis, ou seja, por 198 semanas (208 - 10). Portanto, no total obtemos 199 previsões (198 das semanas mais 01 do início

janela). A janela deslizante de 10 semanas foi obtida através da realização de testes de combinação, a qual variou de 04 até 20 semanas.

O conjunto de dados relatados anteriormente foram separados em 85% para treino e 15% para validação dos modelos. O conjunto de treinamento é utilizado para o ajuste dos modelos enquanto o de validação para avaliar o seu desempenho. Os dados foram normalizados no intervalo de $[0, 1]$. A Tabela IV.3 exibe a média das métricas avaliadas (que foram apresentadas na seção III.4.3) para os 300 modelos treinados (150 ativos para cada modelo). Após a validação dos parâmetros e da janela para cada um dos modelos (SVR e LSTM), foi realizado o treinamento com os dados completos para obtenção dos 300 modelos finais de cada ativo.

Tabela IV.3: Média das métricas dos 300 modelos.

	MAE	MSE
LSTM	0,100221	0,090331
SVR	0,106190	0,096143

IV.2 Modelo Matemático Proposto

O modelo proposto neste trabalho apresenta algumas diferenças em relação aos modelos existentes na literatura. Em primeiro lugar, o objetivo deste modelo é avaliar a composição de um portfólio de investimentos visando a redução do risco, ou seja, o foco principal não é avaliar o retorno do portfólio, mas sim a minimização da sua volatilidade. Em segundo lugar, a composição de um portfólio diversificado é realizada considerando uma quantidade mínima de ativos de diferentes categorias no portfólio, ou seja, o portfólio é obrigatoriamente composto por ativos de classes distintas. Por fim, embora não seja o objetivo principal, é garantido um retorno mínimo esperado para a composição das carteiras.

Neste modelo, busca-se encontrar um portfólio onde a soma das variâncias entre os ativos que compõem a carteira seja a menor dentre todas as combinações possíveis no conjunto. Seja S o conjunto das possíveis carteiras que respeitam a quantidade de ativos a serem escolhidos de cada um dos tipos pré-definidos de soluções possíveis, em que s e \bar{s} representam estes portfólios selecionados, $\hat{V}_{i,j}$ é o risco total da carteira, que é igual à variância da combinação linear de cada ação na carteira entre os ativos i e j . Sua formulação matemática é descrita na Equação IV.1.

$$P = \left\{ s \in S \mid \sum_{i,j \in S} \hat{V}_{i,j} \leq \sum_{i,j \in \bar{S}} \hat{V}_{i,j}, \forall \bar{s} \neq s, \bar{s} \in S \right\} \quad (\text{IV.1})$$

IV.2.1 Portfólio Diversificado com Múltiplas Restrições de Cardinalidade - Modelo de Programação Quadrática

O modelo proposto em IV.1, pode ser representado como um modelo de Programação Quadrática com restrição de cardinalidade. O objetivo é obter um portfólio com risco mínimo por meio da seleção de subportfólios. Cada subportfólio compõe categorias de ativos distintas, portanto, um portfólio multiativo diversificado. Portanto, temos que selecionar ativos para cada subportfólio que, em conjunto, devem minimizar o risco do portfólio inteiro. A seguir é definido sua forma quadrática:

$$\bigcap_{k \in K} V(k) = \emptyset, \bigcup_{k \in K} V(k) = V$$

$$\min \sum_{i \in V} \sigma_i^2 y_i + \sum_{i \in V} \sum_{j \in V | i < j} 2y_i y_j \gamma_{ij} \quad (\text{IV.2})$$

$$\text{s.a.} \quad \sum_{i \in V(k)} y_i = C_k \quad \forall k \in K, \quad (\text{IV.3})$$

$$\frac{1}{n} \sum_{i \in V(k)} \hat{r}_i y_i \geq T \quad \forall k \in K, \quad (\text{IV.4})$$

$$y_i \in \{0, 1\} \quad \forall i \in V \quad (\text{IV.5})$$

O conjunto K representa cada conjunto de categorias distintas de ativos, e V é a união desses conjuntos distintos de ativos, ou seja, todos os ativos disponíveis. A variância do ativo i é expressada pela variável σ_i^2 . A variável γ_{ij} representa a covariância entre os ativos i e j . Por sua vez, C_k indica o número de ativos de cada uma das diferentes categorias. A variável \hat{r}_i representa o retorno predito do ativo, enquanto n é a quantidade de ativos totais na carteira escolhida C e T é a taxa de referência do retorno mínimo esperado.

A função objetivo (IV.2) minimiza o risco da carteira. A família de Restrições (IV.3) limita a quantidade máxima de ativos que devem ser inseridos de cada uma das diferentes categorias, garantindo que o somatório de cada categoria dos ativos escolhidos (y_i) seja limitado pelo tamanho de cada “subcarteira” da categoria de ativos (C_k). Já as Restrições (IV.4) e (IV.5) definem o retorno da carteira desejado e variável binária y_i , que possui valor igual a um para indicar se o ativo i foi selecionado para o portfólio e 0 caso contrário.

IV.2.2 Portfólio Diversificado com Múltiplas Restrições de Carnalidade - Modelo de Programação Linear

O modelo proposto, apresentado anteriormente nas Equações (IV.2) a (IV.5), pode ser reformulado em um modelo de Programação Linear Inteira através de uma linearização simples. Para isso, é necessário substituir o produto dos vetores de decisão, y_i e y_j por $z_{i,j}$, e adicionar algumas restrições para limitar seus limites superior e inferior. Sua formulação é descrita a seguir:

$$\bigcap_{k \in K} V(k) = \emptyset, \quad \bigcup_{k \in K} V(k) = V$$

$$\min \quad \sum_{i \in V} \sigma_i^2 y_i + \sum_{i \in V} \sum_{j \in V | i < j} 2z_{ij} \gamma_{ij} \quad (\text{IV.6})$$

$$\text{s.a.} \quad \sum_{i \in V(k)} y_i = C_k \quad \forall k \in K, \quad (\text{IV.7})$$

$$\frac{1}{n} \sum_{i \in V(k)} \hat{r}_i y_i \geq T \quad \forall k \in K, \quad (\text{IV.8})$$

$$z_{ij} \leq y_i \quad \forall i, j \in V | i \neq j, \quad (\text{IV.9})$$

$$z_{ij} \leq y_j \quad \forall i, j \in V | i \neq j, \quad (\text{IV.10})$$

$$z_{ij} \geq y_i + y_j - 1 \quad \forall i, j \in V | i \neq j, \quad (\text{IV.11})$$

$$y_i \in \{0, 1\} \quad \forall i \in V, \quad (\text{IV.12})$$

$$z_{ij} \in \{0, 1\} \quad \forall i, j \in V \quad (\text{IV.13})$$

O conjunto K representa cada conjunto de categorias distintas, e V é o conjunto de todos os ativos disponíveis. A função objetivo (IV.6) minimiza o risco da carteira. A variância é representada por σ_i^2 , a covariância entre os ativos i e j pela variável γ_{ij} . z_{ij} é uma variável binária com valor igual a 1 ou 0, para indicar se os ativos i e j foram selecionados ou não para a carteira, respectivamente.

A restrição (IV.7) garante que a soma da quantidade de ativos selecionados y_i seja igual ao tamanho de sua respectiva “subcarteira” C_k . Portanto, C_k define o tamanho da k -ésima “subcarteira” e, conseqüentemente, o tamanho total do portfólio. A restrição (IV.8) define o retorno da carteira desejado para n ativos totais na carteira escolhida.

As variáveis z_{ij} são limitadas pelas variáveis y_i e y_j introduzidas no modelo em (IV.9) a (IV.11). Se os ativos i e j forem escolhidos para a carteira, essas restrições garantem que seu valor seja um ($z_{ij} = 1$). Caso apenas um dos ativos seja escolhido, o valor será zero ($z_{ij} = 0$). Sendo assim, as restrições (IV.9) - (IV.11) relacionam o valor de z_{ij} às variáveis binárias y_i e y_j .

Finalmente, as Restrições (IV.12) e (IV.13) definem que z_{ij} e y_i são variáveis binárias.

Capítulo V Experimentos e Resultados Computacionais

Neste capítulo, apresentamos os dados e configurações utilizados neste trabalho, bem como os experimentos e resultados dos métodos definidos no Capítulo III, utilizando os modelos MLOpt para a seleção de portfólios diversificados com múltiplas restrições de cardinalidade na carteira.

V.1 Dados e Configurações

O método proposto foi programado em Python (versão 3.8.8). Todos os testes foram executados utilizando o *IBM ILOG CPLEX Optimization Studio* versão 20.1.0 em um computador com as seguintes configurações: processador Intel Core i5 de 2,30 GHz, memória RAM de 8 GB e sistema operacional Windows 10 de 64 bits. Os modelos de aprendizado de máquina: rede neural LSTM e SVR, foram treinados utilizando o ambiente da *Google Colab Repository* na sua versão gratuita que oferece opções de ambientes de computação acelerados que incluem GPU e TPU.

Para avaliar o desempenho dos modelos propostos, utilizamos conjuntos de dados financeiros históricos de três tipos de ativos: *i*) ações brasileiras, *ii*) ações americanas e *iii*) REITs americanos. Os dados foram coletados considerando as cotações diárias fechadas obtidas através do “*Yahoo Finance*” por meio da biblioteca *yfinance* em Python.

Os conjuntos de dados foram coletados dos mercados financeiros NASDAQ (Estados Unidos da América) e B3 (Brasil), para o período de 2014 a 2021. Os dados dos anos de 2014 a 2017 foram utilizados para realizar os experimentos, ou seja, para a seleção da carteira, enquanto os dados dos anos de 2018 a 2021 foram utilizados para avaliar o desempenho das carteiras selecionadas. Devido a pandemia da COVID-19, os dados dos ativos coletados das Bolsas de Valores apresentam um comportamento distinto em suas séries históricas a partir do ano de 2020.

Além dos dados individuais de cada ativo, também coletamos os índices correspondentes: Nasdaq Composite (EUA), FTSE Nareit All REITs Index (EUA) e Bovespa (BR), nos anos de 2018 a 2021. Esses índices foram utilizados para fins de comparação com as soluções geradas pelos modelos propostos

Para esses dados públicos, são adotados alguns métodos do processo de limpeza de dados para evitar ruído, dados incompletos e/ou inconsistentes, por exemplo: eliminação de séries curtas. Consideramos apenas ações e REITs listados no período de análise e removemos dados com valores

ausentes. Assim, também são eliminados os ativos como os de oferta pública inicial (IPO) durante esse período. Várias abordagens existem para resolver efetivamente o problema de valores ausentes, por exemplo, preencher com o valor mais apropriado ou remover os dados. Neste caso, eles foram removidos.

V.2 Experimentos

A quantidade de ativos disponíveis coletados através da biblioteca *yfinance* é maior do que a quantidade disponibilizada para a execução dos experimentos. Portanto, foi realizada uma seleção aleatória proporcional para cada categoria de ativo, a fim de escolher os ativos que fariam parte dos experimentos. Tal seleção se fez necessária porque, considerando os modelos propostos, a disponibilidade de uma grande quantidade de ativos impossibilitaria encontrar soluções ótimas, dificultando a avaliação de sua qualidade pelos modelos propostos.

As previsões dos retornos dos ativos são gerados utilizando cada um dos 300 modelo criados. As previsões foram realizadas para obter valores das 205 semanas no período que compreende os anos de 2018 até 2021, assim um período total de 4 anos. Assim, foram gerados um total de 61.500 previsões de preços semanais (205 semanas x 300 ativos). As previsões foram realizadas utilizando uma janela deslizante de 10 dias, conforme também efetuado no treinamento. Após a obtenção dos valores preditos, o retorno de cada ativo no período é calculado. Por fim, esses retornos são utilizados como restrição na seleção das carteiras.

Considerando a restrição do retorno da carteira, utilizamos a taxa CDI (Certificado de Depósito Interbancário) que representa as taxas de juros de empréstimos realizados entre bancos no período analisado (2018 até 2021). Normalmente ela possui um valor próximo ao da taxa básica de juros, taxa Selic, que é o principal instrumento de política monetária utilizado pelo Banco Central (BC) para controlar a inflação. Neste intervalo de tempo seu valor consolidado foi de 20,96%.

Cada instância (ou experimento) contém uma amostra sem reposição correspondente à proporção dos ativos totais disponíveis extraídos com tamanhos de 15 a 150 ativos. Em cada instância, a seleção da carteira é diferente em termos do número escolhido, de 3 a 30 ativos, dependendo do número de ativos disponíveis para a instância. Todas as instâncias são realizadas para obter o custo, tempo de execução (segundos), *gap* (%), e a carteira selecionada (PS) com três horas como limite de tempo. Quando o limite de tempo é excedido, a instância é interrompida e seus dados são coletados.

A aplicação dos testes envolve as duas configurações de modelos MLOpt: uma que utiliza o modelo de rede neural LSTM em conjunto com os modelos de otimização, e outra que emprega o modelo de aprendizado de máquina SVR em conjunto com os modelos de otimização.

A Tabela V.1 apresenta um resumo dos testes realizados. A coluna 1 indica a divisão das

instâncias dos testes, enquanto as colunas 2 a 4 mostram a quantidade de ativos disponíveis em cada categoria. A coluna 5 exibe a disponibilidade total de ativos, e a coluna 06 mostra a quantidade de ativos a serem selecionados para cada instância.

Ainda na Tabela V.1 observamos os diferentes testes realizados. Os Testes foram divididos em faixas numeradas de 01 a 38. Cada teste foi caracterizado pela quantidade de ativos disponíveis em cada categoria (ações dos EUA, REITs dos EUA e ações do Brasil). A coluna “Total disponíveis” indica a quantidade total de ativos disponíveis considerando todas as categorias, enquanto a coluna “Ativos para selecionar” indica a quantidade de ativos a serem selecionados em cada instância de teste (categoria). Por exemplo, no Teste 01, foram selecionados 3 ativos, sendo 1 ativo de cada categoria. Já no Teste 02, foram selecionados 6 ativos, sendo 2 de cada categoria. Os demais experimentos seguem a mesma lógica e, assim, a quantidade total de ativos a serem selecionados na carteira é a soma das subcarteiras de cada categoria.

Tabela V.1: Guia dos experimentos.

Testes	Ações (EUA)	REITs (EUA)	Ações (BR)	Total disponíveis	Ativos para selecionar
01 - 03	05	05	05	15	03, 06, 12
04 - 08	10	10	10	30	03, 06, 12, 18, 24
09 - 14	20	20	20	60	03, 06, 12, 18, 24, 30
15 - 20	25	25	25	75	03, 06, 12, 18, 24, 30
21 - 26	30	30	30	90	03, 06, 12, 18, 24, 30
27 - 32	40	40	40	120	03, 06, 12, 18, 24, 30
33 - 38	50	50	50	150	03, 06, 12, 18, 24, 30

V.3 Resultados Computacionais

Nas Tabelas V.2 e V.3 são mostrados os resultados dos modelos quadrático e linear em relação as predições realizadas com os modelos LSTM e SVR, respectivamente. A coluna 1 indica as instâncias executadas, totalizando 38. As colunas 2 a 4 mostram o custo, o *gap* (diferença em relação à solução ótima) e o tempo de execução em segundos para as soluções encontradas pelo modelo quadrático para cada instância. As colunas 5 a 7 exibem os mesmos valores para o modelo linear. A coluna 8 indica se a carteira escolhida pelos modelos quadrático e linear é a mesma.

Na Tabela V.2, são mostrados os resultados utilizando o modelo LSTM. Pode-se observar que o modelo quadrático encontrou soluções ótimas e provou suas otimalidades em todas instâncias executadas. O modelo linear não foi capaz de encontrar e provar a otimalidade em apenas aproximadamente 24% das instâncias. Em relação ao tempo de execução, o modelo quadrático concluiu sua execução sem atingir o limite de tempo de 3 horas (ou 1.800 segundos) para todas as instâncias, e atingiu o seu maior tempo no experimento 38. Por outro lado, o modelo linear atingiu o limite de tempo em aproximadamente 21% das instâncias (8 instâncias), mas conseguiu obter os mesmos portfólios ótimos que o modelo quadrático.

Na Tabela V.3, vemos os resultados dos métodos matemáticos utilizando o modelo SVR. Assim como nos resultados mostrados anteriormente para o LSTM, o modelo quadrático encontrou soluções ótimas em todas as instâncias com exceção dos experimentos 03 e 08, pois eles não encontraram carteiras que atendessem todas as restrições. Já o modelo linear não foi capaz de encontrar soluções tão boas quanto o quadrático, sendo que aproximadamente 26% das soluções encontradas alcançaram o tempo limite. Observando a igualdade dos portfólios, vemos que nas instâncias 37 e 38 o modelo linear encontrou piores soluções, ou seja, os portfólios encontrados não são iguais aos ótimos (selecionados pelo modelo quadrático).

A Figura V.1 exibe a comparação dos resultados dos modelos quadráticos e linear dos experimentos executados com o modelo LSTM. O modelo quadrático atingiu o valor ótimo ($Gap = 0\%$) em todas as execuções enquanto o modelo linear não garantiu a otimalidade para 8 instâncias. Olhando os custos, ambos os modelos escolheram carteiras próximas com o valor médio do custo da função objetivo em torno de 0,005. Entretanto, quando comparamos a qualidade das soluções pelo Gap e tempo de execução, vemos que o modelo quadrático obteve melhores resultados e conseguiu encontrar as soluções ótimas em mais experimentos e em um menor tempo do que com o modelo linear.

A Figura V.2 mostra a comparação dos resultados dos modelo quadrático e linear dos experimentos executados com o modelo SVR. O modelo quadrático garantiu a otimalidade em todas as execuções, enquanto que o linear atingiu o tempo limite em 07 experimentos. Nestes experimentos, para 02 instâncias o método não conseguiu encontrar soluções que atendessem as restrições, são eles: experimento 03 e 08.

As Tabelas V.4 e V.5 exibem os resultados calculados para o retorno e volatilidade de cada experimento das soluções ótimas encontradas. Conforme mostrados nas Equações III.3 e III.4, o retorno dos portfólios é calculado como a soma das variações percentuais dos preços (rendimento), enquanto a volatilidade é medida com a variância. Os resultados contidos nestas tabelas serão explorados a seguir nas seções V.3.1 e V.3.2.

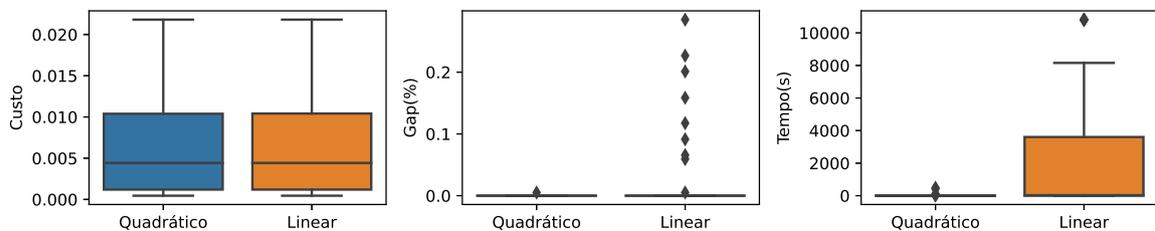


Figura V.1: Comparações dos modelos quadrático e linear - modelo LSTM

Fonte: Autor

Tabela V.2: Resultados dos experientos - modelo LSTM

Exp.	Quadrático			Linear			Portfólio Iguais
	Custo	Gap(%)	tempo(s)	Custo	Gap (%)	tempo(s)	
1	0,00	0	0,02	0,00	0	0,05	X
2	0,00	0	0,02	0,00	0	0,05	X
3	0,01	0	0,02	0,01	0	0,03	X
4	0,00	0	0,03	0,00	0	0,22	X
5	0,00	0	0,05	0,00	0	0,53	X
6	0,01	0	0,03	0,01	0	0,34	X
7	0,01	0	0,02	0,01	0	0,09	X
8	0,02	0	0,03	0,02	0	0,16	X
9	0,00	0	0,09	0,00	0	0,09	X
10	0,00	0	0,17	0,00	0	0,17	X
11	0,00	0	0,47	0,00	0	15,56	X
12	0,01	0	0,33	0,01	0	109,70	X
13	0,01	0	0,34	0,01	0	37,89	X
14	0,02	0	0,24	0,02	0	18,36	X
15	0,00	0	0,19	0,00	0	3,66	X
16	0,00	0	0,53	0,00	0	2,09	X
17	0,00	0	0,67	0,00	0	107,25	X
18	0,01	0	1,64	0,01	0	3437,55	X
19	0,01	0	0,45	0,01	0	3660,22	X
20	0,02	0	0,52	0,02	0,52	1466,00	X
21	0,00	0	0,44	0,00	0	7,03	X
22	0,00	0	0,77	0,00	0	3,75	X
23	0,00	0	4,19	0,00	0	614,42	X
24	0,01	0	4,66	0,01	0	8156,86	X
25	0,01	0	1,61	0,01	9,13	10805,60	X
26	0,02	0	0,81	0,02	5,93	10806,20	X
27	0,00	0	1,14	0,00	0	24,16	X
28	0,00	0	1,70	0,00	0	7,58	X
29	0,00	0	6,36	0,00	0	418,00	X
30	0,01	0	8,25	0,01	6,57	10810,40	X
31	0,01	0	30,00	0,01	15,85	10805,90	X
32	0,02	0	6,00	0,02	20,12	10804,50	X
33	0,00	0	2,45	0,00	0	43,14	X
34	0,00	0	4,08	0,00	0	19,33	X
35	0,00	0	10,63	0,00	0	905,38	X
36	0,01	0	122,58	0,01	11,74	10803,50	X
37	0,01	0	417,38	0,01	22,67	10807,90	X
38	0,01	0	488,61	0,01	28,50	10815,20	X

Tabela V.3: Resultados dos experimentos - modelo SVR

Exp.	Quadrático			Linear			Portfólio Iguais
	Custo	Gap(%)	tempo(s)	Custo	Gap (%)	tempo(s)	
1	0,00	0	0,25	0,00	0	0,02	X
2	0,00	0	0,22	0,00	0	0,03	X
3	-	-	-	-	-	-	
4	0,00	0	0,05	0,00	4,70	0,01	X
5	0,00	0	0,03	0,00	3,10	0,52	X
6	0,01	0	0,03	0,00	3,10	0,27	X
7	0,01	0	0,03	0,00	3,10	0,13	X
8	-	-	-	-	-	-	
9	0,00	0	0,09	0,00	0	1,75	X
10	0,00	0	0,11	0,00	0	1,63	X
11	0,01	0	0,56	0,01	0	7,92	X
12	0,01	0	0,28	0,01	0	9,36	X
13	0,02	0	0,16	0,02	0	6,03	X
14	0,03	0	0,09	0,03	0	1,75	X
15	0,00	0	0,16	0,00	0	2,95	X
16	0,00	0	0,09	0,00	0	2,69	X
17	0,00	0	0,70	0,00	0	10,44	X
18	0,01	0	0,56	0,01	0	216,86	X
19	0,02	0	0,56	0,02	0	460,30	X
20	0,02	0	0,25	0,02	0	39,58	X
21	0,00	0	0,23	0,00	0	5,67	X
22	0,00	0	0,20	0,00	0	7,91	X
23	0,00	0	2,67	0,00	0	34,70	X
24	0,01	0	1,61	0,01	0	866,31	X
25	0,02	0	0,88	0,02	0	7674,26	X
26	0,02	0	1,39	0,02	0	5512,73	X
27	0,00	0	0,61	0,00	0	20,95	X
28	0,00	0	0,47	0,00	0	11,92	X
29	0,00	0	4,89	0,00	0	66,05	X
30	0,01	0	3,09	0,01	4,62	10809,00	X
31	0,01	0	41,03	0,01	10,91	10806,20	X
32	0,02	0	9,52	0,02	15,38	10809,20	
33	0,00	0	1,36	0,00	0	43,23	X
34	0,00	0	1,70	0,00	0	102,22	X
35	0,00	0	10,92	0,00	0	299,14	X
36	0,01	0	28,17	0,01	8,75	10800,30	X
37	0,01	0	258,56	0,01	16,55	10812,40	
38	0,02	0	462,81	0,02	20,20	10817,80	

Tabela V.4: Retorno e volatilidade - modelo LSTM.

Índice/Inst	Retorno (%)	Volat. no Período (%)	Índice/Inst	Retorno (%)	Volat. Período (%)
Nasdaq	123,28	23,80	Exp. 19	147,72	22,71
FTSE	43,49	24,17	Exp. 20	126,20	22,99
Bovespa	34,58	28,49	Exp. 21	22,10	35,48
Exp. 1	27,53	30,98	Exp. 22	20,34	24,20
Exp. 2	76,62	25,06	Exp. 23	218,83	21,26
Exp. 3	88,38	24,92	Exp. 24	174,85	21,92
Exp. 4	84,02	21,58	Exp. 25	152,45	21,88
Exp. 5	95,95	25,74	Exp. 26	123,54	22,92
Exp. 6	88,76	24,57	Exp. 27	42,13	30,48
Exp. 7	86,89	22,73	Exp. 28	30,49	23,11
Exp. 8	97,93	23,92	Exp. 29	206,97	18,24
Exp. 9	22,10	35,48	Exp. 30	167,75	20,58
Exp. 10	20,34	24,20	Exp. 31	136,85	20,51
Exp. 11	83,40	22,57	Exp. 32	120,87	20,59
Exp. 12	78,32	22,25	Exp. 33	42,13	30,48
Exp. 13	68,62	22,29	Exp. 34	34,86	22,30
Exp. 14	57,96	23,80	Exp. 35	29,15	19,42
Exp. 15	22,10	35,48	Exp. 36	154,08	18,72
Exp. 16	20,34	24,20	Exp. 37	136,40	20,29
Exp. 17	213,97	21,04	Exp. 38	117,17	20,37
Exp. 18	179,52	21,18			

Tabela V.5: Retorno e volatilidade - modelo SVR.

Índice/Inst	Retorno (%)	Volat. no Período (%)	Índice/Inst	Retorno (%)	Volat. Período (%)
Nasdaq	123,28	23,80	Exp. 19	53,44	23,94
FTSE	43,49	24,17	Exp. 20	87,21	22,93
Bovespa	34,58	28,49	Exp. 21	45,75	32,77
Exp. 1	24,92	32,73	Exp. 22	37,75	22,81
Exp. 2	90,92	25,74	Exp. 23	214,10	23,10
Exp. 3	-	-	Exp. 24	65,04	22,57
Exp. 4	31,72	29,06	Exp. 25	56,83	22,49
Exp. 5	100,15	23,40	Exp. 26	52,78	21,94
Exp. 6	81,84	24,35	Exp. 27	45,75	32,77
Exp. 7	81,43	22,69	Exp. 28	70,15	23,35
Exp. 8	-	-	Exp. 29	201,49	20,99
Exp. 9	67,36	27,64	Exp. 30	48,86	22,13
Exp. 10	100,15	23,40	Exp. 31	134,23	20,39
Exp. 11	55,83	24,48	Exp. 32	114,55	20,32
Exp. 12	68,35	22,20	Exp. 33	45,75	32,77
Exp. 13	103,25	22,70	Exp. 34	79,64	25,40
Exp. 14	133,00	22,44	Exp. 35	67,37	21,28
Exp. 15	45,75	32,77	Exp. 36	45,91	22,02
Exp. 16	87,58	24,75	Exp. 37	151,93	20,90
Exp. 17	208,40	22,65	Exp. 38	120,55	20,46
Exp. 18	60,46	23,62			

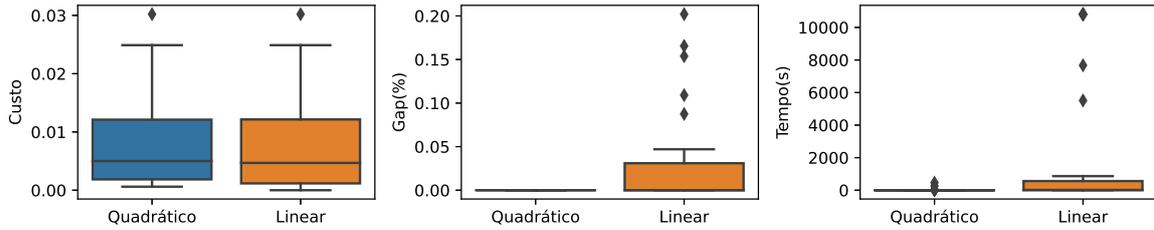


Figura V.2: Comparações dos modelos quadrático e linear - modelo SVR

Fonte: Autor

V.3.1 Volatilidade

A Figura V.3 exibe o resultado da volatilidade das carteiras ótimas selecionadas pelo modelo quadrático em relação aos índices de *benchmarking* (Nasdaq, FTSE e Bovespa) utilizando o LSTM para predição dos retornos. Em torno de 60% das carteiras selecionadas resultaram em uma volatilidade menor comparado aos índices de *benchmarking*.

A respeito das carteiras em que as volatilidades ficaram maiores que as dos índices, para todas elas sua composição é de apenas 03 ativos. Portanto, uma maior volatilidade é esperada devido a ter uma quantidade pequena de ativos para sua diversificação. É importante esclarecer que o valor das volatilidades que são iguais nestes experimentos são devido aos mesmos ativos terem sido selecionados na composição da carteira. Neste sentido, se desconsiderarmos estas carteiras observamos que os experimentos tiveram volatilidade próxima ou menor que as dos índices, caracterizando assim a composição de carteiras diversificadas através do método proposto.

A Figura V.4 exibe o resultado da volatilidade das carteiras selecionadas em relação aos índices de *benchmarking* utilizando o modelo SVR para predição dos retornos. O valor das volatilidades que são iguais nestes experimentos é devido terem sido selecionados os mesmos ativos na composição da carteira.

Nas carteiras selecionadas, 63% tiveram menor volatilidade que todos os índices de *benchmarking*, enquanto que apenas 15% delas resultaram em uma volatilidade maior comparado aos índices de *benchmarking* (carteiras 01, 04, 15, 21, 27 e 33). Logo, desconsiderando essas carteiras, observamos que os experimentos tiveram volatilidade próxima ou menor aos índices de *benchmarking*, caracterizando assim a composição de carteiras diversificadas através do método proposto com baixas volatilidades.

V.3.2 Retorno

Os resultados da Figura V.5 mostram os retornos dos portfólios selecionados utilizando o modelo de LSTM na predição de retornos futuros para os portfólios ótimos encontrados no modelo quadrático. Observamos que a grande parte dos experimentos superaram os índices de *benchmarking*.

king (obtiveram retorno acima dos índices), principalmente comparados aos índices da Bovespa e FTSE. Olhando mais atentamente, apenas 18% das instâncias (experimentos 01, 09, 10, 15, 16, 28 e 35) não superaram nenhum índice. Estes experimentos, que não superaram os índices, são os que possuem uma menor quantidade de ativos na carteira, 03 e 06 ativos.

Outro resultado importante é que 34% dos experimentos tiveram retornos acima dos índices Bovespa e FTSE (instâncias 02 a 09, 11 a 14, 32 e 38), e 31% superaram todos os índices e, inclusive, em muito o índice Nasdaq (instâncias 17 a 20, 23 a 25, 29 a 31, 36 e 37).

Nas carteiras selecionadas que superaram todos os índices (31%), algumas foram de grande destaque adquirindo retornos superior a 200% (experimentos 17, 23 e 29). Nestes portfólios a carteira selecionada é composta por um total de 12 ativos.

A Figura V.6 mostra os retornos dos portfólios selecionados utilizando o modelo de SVR na predição de retornos futuros para os portfólios ótimos encontrados no modelo quadrático. Observamos que poucos experimentos superaram todos os índices de *benchmarking* (obtiveram retorno acima dos índices), principalmente comparados ao índice Nasdaq. Olhando mais atentamente, apenas 5% das instâncias (experimentos 01, 04) não superaram nenhum índice. Estes experimentos, que não superaram os índices, são os que possuem uma menor quantidade de ativos disponíveis para escolha na composição da carteira (15 ativos).

Outro resultado importante é que 86% das carteiras selecionadas superaram os índices Bovespa e FTSE, ou seja, apenas os portfólios 01 e 04, relatados anteriormente, e a carteira 22 (experimento 22) não superaram esses índices. Em relação ao índice Nasdaq, 15% dos experimentos o conseguiram superar.

V.3.3 Melhores Carteiras

A Figura V.7 mostra as carteiras que utilizaram o algoritmo LSTM e obtiveram tanto o maior retorno quanto a menor volatilidade em comparação a todos os índices de *benchmarking*. A respeito disso, temos 34% dos experimentos com uma composição de carteira melhor tanto em retorno quanto volatilidade do que os índices. Destas carteiras, vemos que em sua maioria (experimentos 17 a 19, 23 a 25 e 29 a 31, 36 e 37) obtiveram retornos muito maiores que dos índices. O mesmo verificamos para as volatilidades, as carteiras 29 e 36 obtiveram as menores volatilidades entre todas as melhores carteiras selecionadas.

Por fim, temos a carteira 29 como o melhor resultado entre todos, ou seja, apresenta a menor volatilidade (em torno de 18%) e também o maior retorno (valorização em torno de 206%) entre todas selecionadas como melhores carteiras.

A Figura V.8 exibe as carteiras que tiveram retorno maior e volatilidade menor que os índices de *benchmarking* utilizando o modelo SVR nas predições de retorno. Um resultado de 06 carteiras (ou

experimentos) obtiveram melhores resultados tanto em retorno quanto em volatilidade comparando aos índices, o que corresponde a 15% do total de carteiras geradas nas instâncias. Acerca dos retornos, vemos que os portfólios 17, 23 e 29 atingiram um resultado em torno de 200%, muito maiores que os retornos observados nos índices.

Por último, não temos uma carteira que contenha a combinação de menor volatilidade e maior retorno entre todos os portfólios escolhidos, mas conseguimos ter ótimos resultados tanto em retornos bem superiores e quanto em menores volatilidades em comparação aos índices, por exemplo: carteiras 17, 29 e 37.

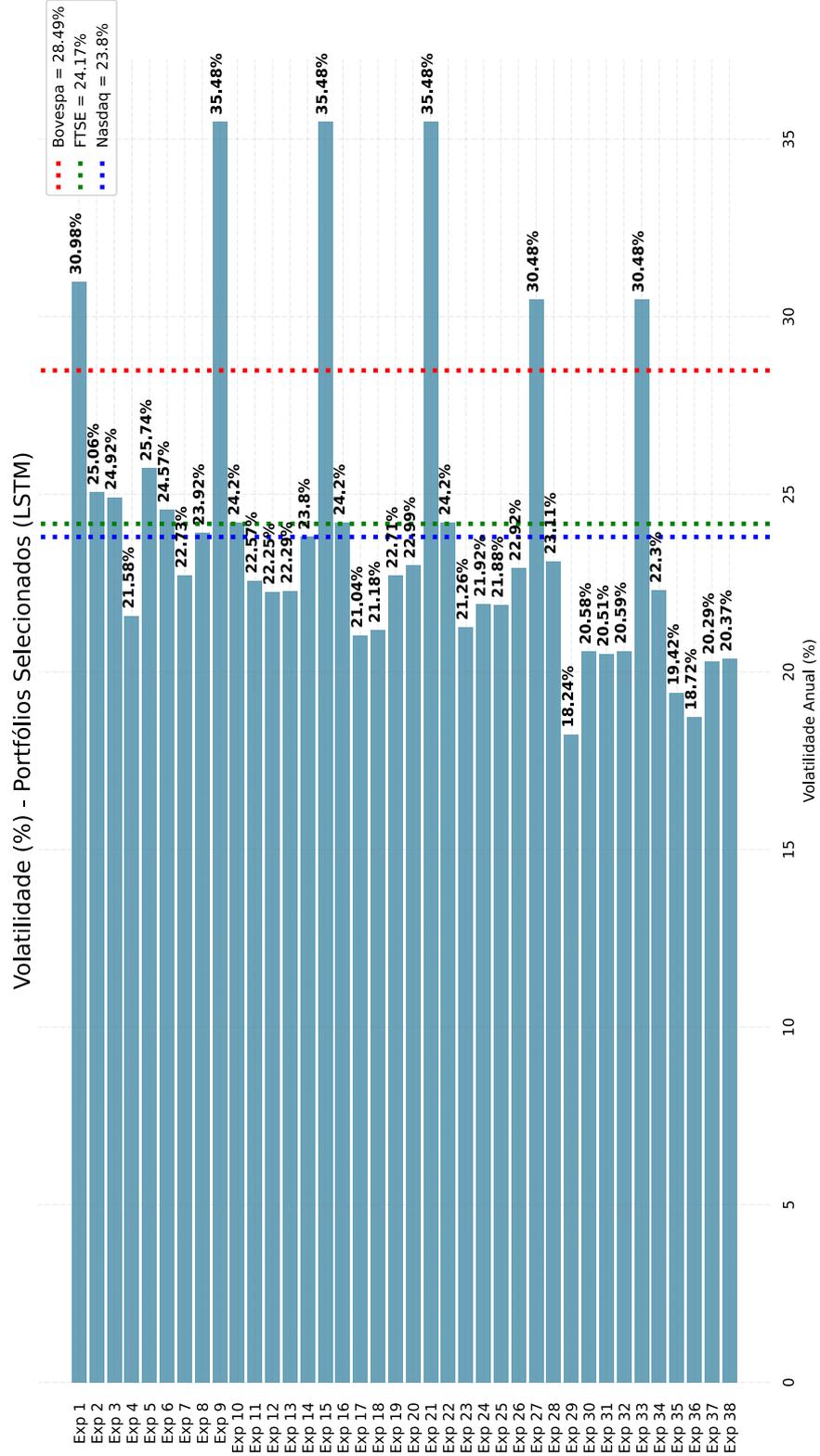


Figura V.3: Volatilidade das carteiras - modelo LSTM
 Fonte: Autor

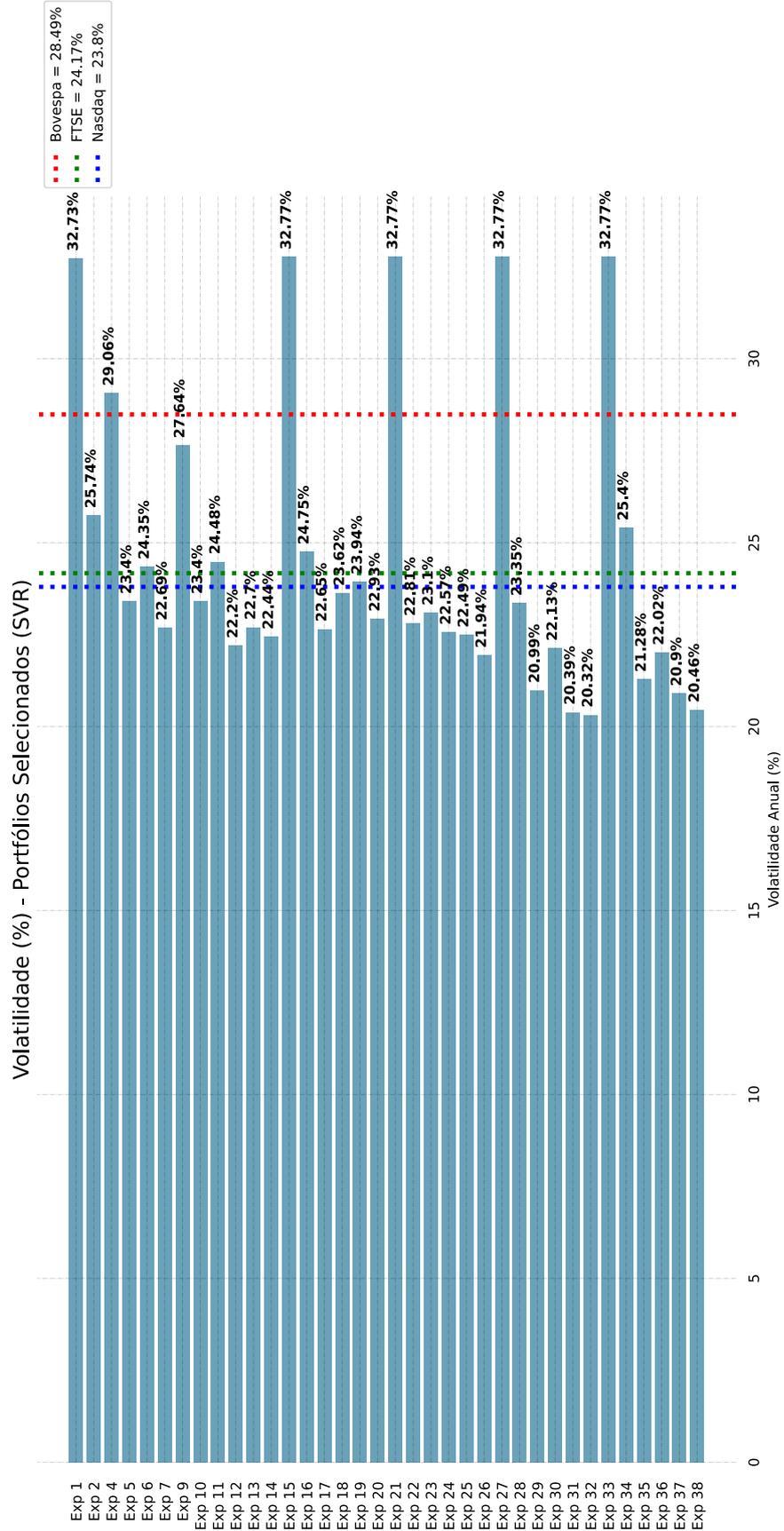


Figura V.4: Volatilidade das carteiras - modelo SVR

Fonte: Autor

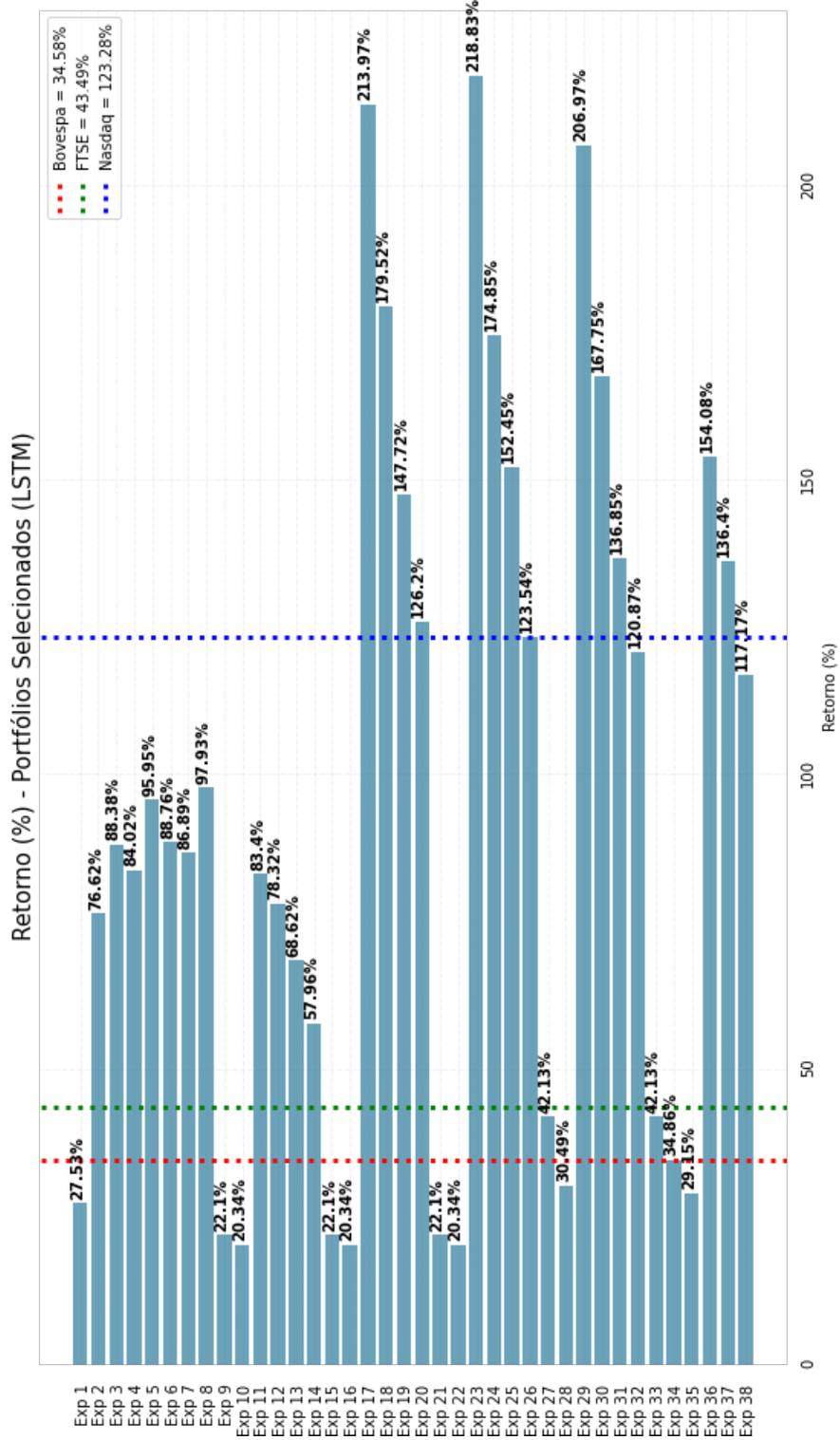


Figura V.5: Retorno das carteiras - modelo LSTM
 Fonte: Autor

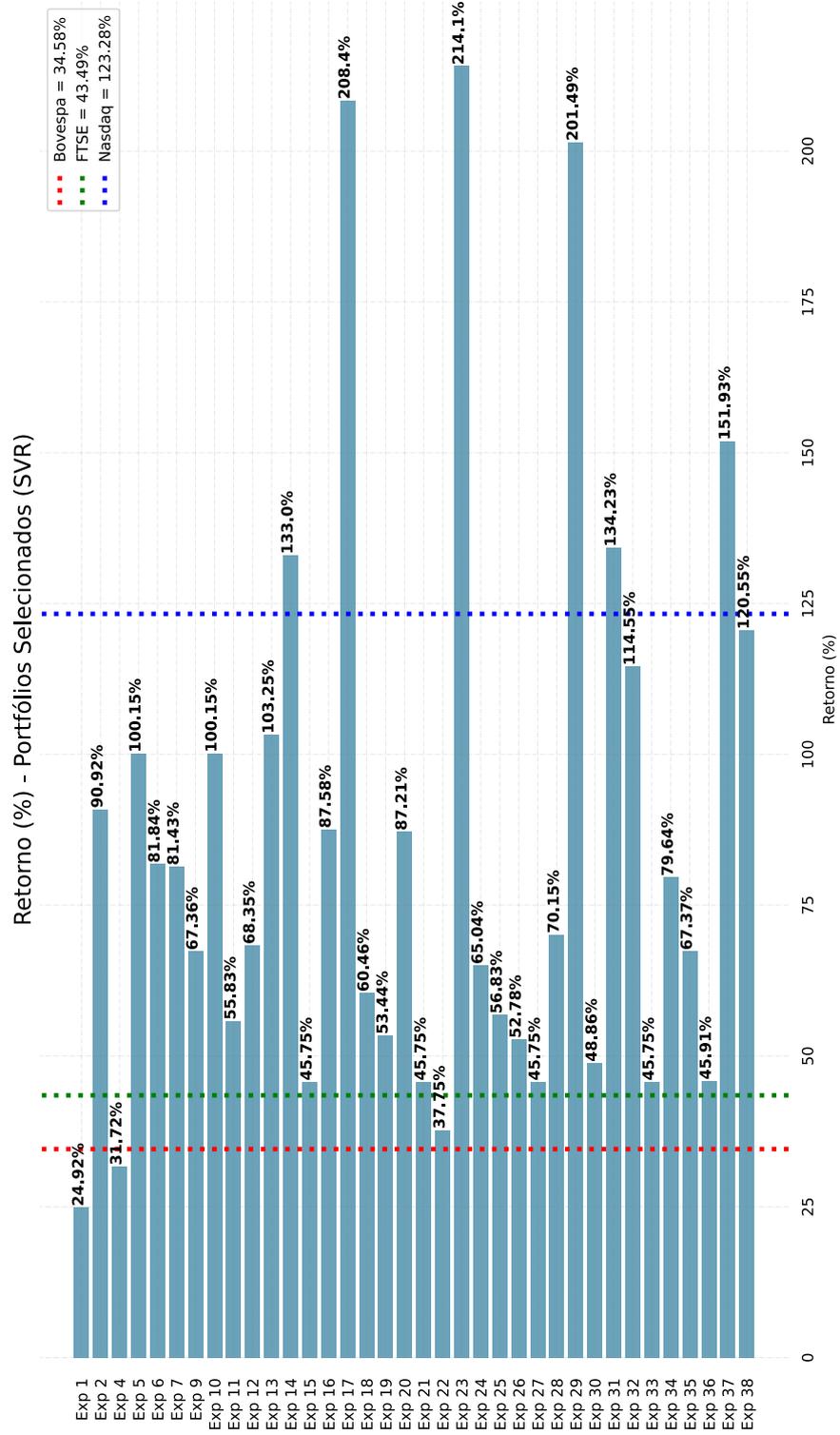


Figura V.6: Retorno das carteiras - modelo SVR
 Fonte: Autor

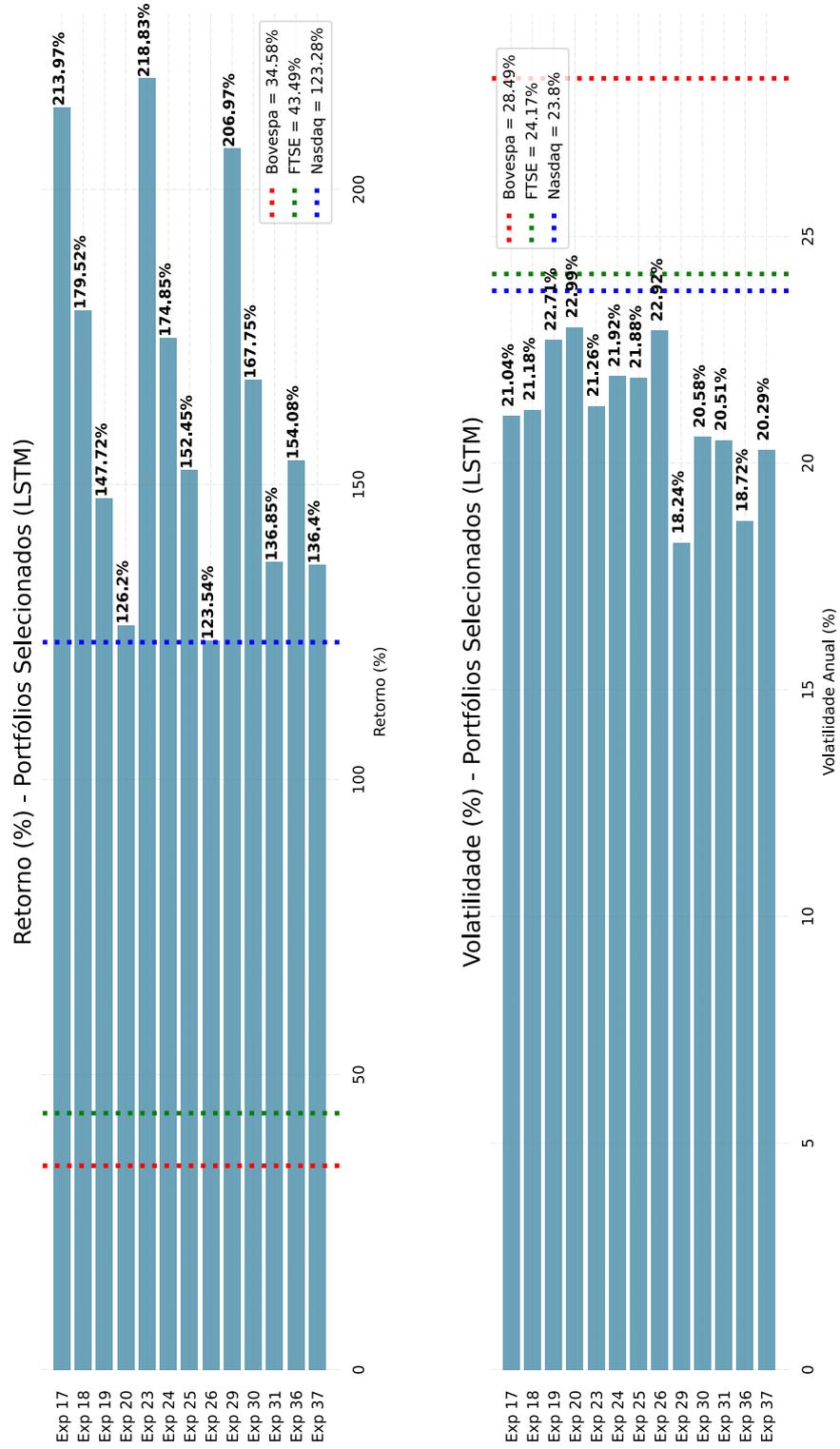


Figura V.7: Melhores carteiras em retorno e volatilidade - modelo LSTM
 Fonte: Autor

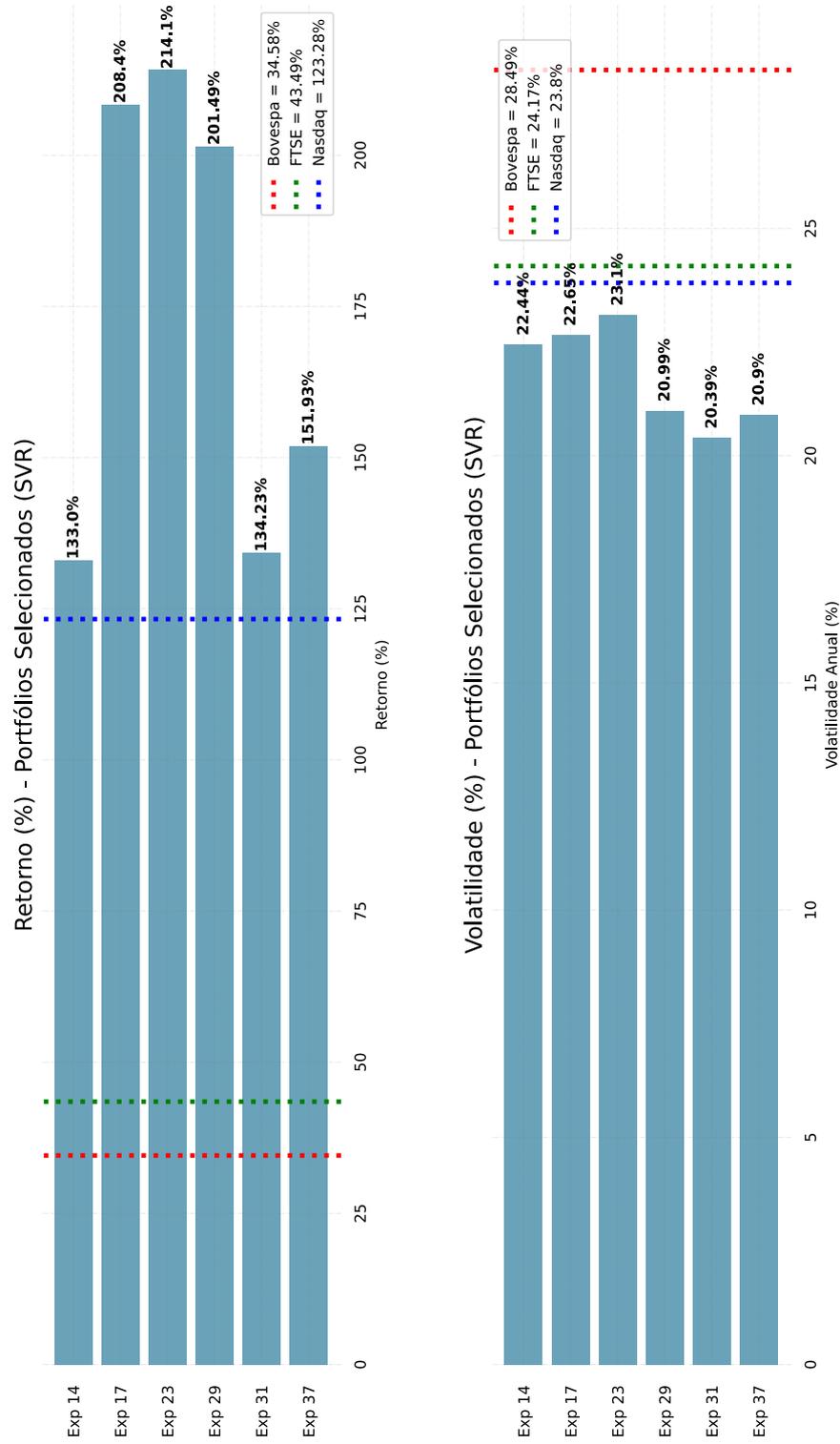


Figura V.8: Melhores carteiras em retorno e volatilidade - modelo SVR
Fonte: Autor

Capítulo VI Conclusões

A seleção de carteiras é um problema fundamental na área de finanças e com grande aprofundamento no mercado e na literatura. Desde sua implementação inicial por Markowitz, surgiram diversas técnicas que podem contribuir para a seleção de portfólios, e outras evoluíram, incluindo técnicas de aprendizado de máquina e redes neurais. Essas abordagens podem, em muitos casos, representar escolhas superiores em comparação com a simples utilização do retorno médio das ações como retorno esperado, que é comumente aplicada pelos modelos clássicos de seleção de carteira. Isto abre oportunidades na aplicação destas técnicas juntamente com a utilização de modelos de otimização para propor novas resoluções na composição portfólios.

Nesse contexto, foi formulado a hipótese de construir uma carteira diversificada e equilibrada, composta por diversos ativos de diferentes mercados, e empregar técnicas de previsão para alcançar o retorno mínimo desejado. Com base nessa premissa, foi desenvolvido o modelo MLOpt para a “*seleção de portfólio diversificado com múltiplas restrições de cardinalidade*”, que incorpora uma rede neural LSTM e um modelo de aprendizado de máquina SVR para prever os retornos futuros, em conjunto com um modelo de otimização. Em seguida, comparou se o risco, medido pela variância, e o retorno das carteiras selecionadas com os índices Nasdaq, Bovespa e FTSE.

Foi efetuado um grande número de experimentos com dados reais e históricos do mercado de REITs, e ações americanas e ações brasileiras, por meio do qual simulamos a seleção de carteiras utilizando o retorno calculado de cada um dos modelos. Os resultados obtidos validaram o modelo proposto de diversificação da carteira de investimentos com múltiplos ativos.

O método apresentado neste trabalho obteve resultados preliminares promissores, entretanto, a fim que torne se mais próximo de uma aplicação prática, algumas sugestões de trabalho futuro são:

1. Utilizar diferentes métricas de avaliação do risco do portfólio aderente aos utilizados no mercado como risco *sharpe* e a fronteira eficiente;
2. Incluir restrição de volatilidade mínima para limitar o risco exposto na formulação da carteira de investimento;
3. Efetuar a escolha dos ativos disponíveis mais cuidadosamente, utilizando critérios como os dos índices de *benchmarking*, observando o volume dos ativos negociados, o balanço saudável

- das empresas;
4. Considerar portfólios e pesos diferentes na formulação da carteira, tendo em vista um perfil de investidor com uma abordagem mais agressiva, moderado ou neutra.
 5. Utilizar de modelos de aprendizado de máquina distintos para cada categoria de ativos.
 6. Investigação e aprimoramento dos métodos de predição dos retornos utilizados. Nesse sentido, partir da aplicação de predição dos retornos futuros como diversas abordagens na literatura.
 7. Aplicação e exploração de novas técnicas e modelos de predição de retorno futuro. Explorando algoritmos distintos como ARIMA, Redes Neurais Auto Regressivas, etc.
 8. Aplicar algoritmos metaheurísticos para avaliar a qualidade das soluções e explorar um conjunto maior de ativos disponíveis para solução.

Referências Bibliográficas

- Alex Graves and Navdeep Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In *International conference on machine learning*, pages 1764–1772. PMLR, 2014. , 23
- Harris Drucker, Christopher J Burges, Linda Kaufman, Alex Smola, and Vladimir Vapnik. Support vector regression machines. *Advances in neural information processing systems*, 9, 1996. , 23, 24, 25
- Adam Hayes. Financial markets: Role in the economy, importance, types, and examples, Jan 2023. URL <https://www.investopedia.com/terms/f/financial-market.asp>. Accessed: 2023-02-09. 12
- Tom Gresham. Reits vs. stocks, Nov 2017. URL <https://finance.zacks.com/reits-vs-stocks-3920.html>. Accessed: 2023-02-09. 12
- GA Vijayalakshmi Pai and Thierry Michel. Evolutionary optimization of constrained k -means clustered assets for diversification in small portfolios. *IEEE Transactions on Evolutionary Computation*, 13(5):1030–1053, 2009. 12, 13
- Can B Kalayci, Olcay Polat, and Mehmet A Akbay. An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization. *Swarm and Evolutionary Computation*, 54: 100662, 2020. 13
- Harry Markowitz. Portfolio selection in the journal of finance vol. 7. 1952. 13
- Jana Doering, Angel A Juan, Renatas Kizys, Angels Fito, and Laura Calvet. Solving realistic portfolio optimization problems via metaheuristics: A survey and an example. In *International Conference on Modeling and Simulation in Engineering, Economics and Management*, pages 22–30. Springer, 2016. 13, 20
- Petter N. Kolm, Reha Tütüncü, and Frank J. Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356–371, 2014. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2013.10.060>. URL <https://www.sciencedirect.com/science/article/pii/S0377221713008898>. 60 years following Harry Markowitz’s contribution to portfolio theory and operations research. 13

- Tun-Jen Chang, Sang-Chin Yang, and Kuang-Jung Chang. Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Systems with applications*, 36(7):10529–10537, 2009. 13
- Harry M Markowitz. Portfolio selection: Efficient diversification of investments. *Cowles Foundation Monograph*, 16, 1959. 13
- Paul A Samuelson. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *The Review of Economic Studies*, 37(4):537–542, 1970. 13
- Hiroshi Konno and Hiroaki Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management science*, 37(5):519–531, 1991. 13
- Philippe Jorion. Value at risk: the new benchmark for controlling market risk. *(No Title)*, 1997. 13
- Martin R Young. A minimax portfolio selection rule with linear programming solution. *Management science*, 44(5):673–683, 1998. 13
- Eduardo de Carvalho Chagas. Modelo de otimização de portfólios restrito: um modelo para custos para b3 e um estudo do impacto das restrições de cardinalidade e fronteira. 2023. 13
- Javier Panadero, Jana Doering, Renatas Kizys, Angel A Juan, and Angels Fito. A variable neighborhood search simheuristic for project portfolio selection under uncertainty. *Journal of Heuristics*, 26(3):353–375, 2020. 13
- Yilin Ma, Ruizhu Han, and Weizhong Wang. Prediction-based portfolio optimization models using deep neural networks. *Ieee Access*, 8:115393–115405, 2020. 14, 17, 24, 26, 28
- Fabio D Freitas, Alberto F De Souza, and Ailson R De Almeida. Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, 72(10-12):2155–2170, 2009. 14, 15, 16, 20, 21
- Thomas Hellström. *Outlier removal for prediction of covariance matrices with an application to portfolio optimization*. Univ., 2000. 15
- Fábio Daros de Freitas, Alberto Ferreira De Souza, and Ailson Rosetti de Almeida. A prediction-based portfolio optimization model. In *Proc. 5st Int. Symp. Robot. Automat.*, pages 520–525, 2006. 15, 16, 19
- Cuiyan Hao, Jiaqian Wang, Wei Xu, and Yuan Xiao. Prediction-based portfolio selection model using support vector machines. In *2013 Sixth International Conference on Business Intelligence and Financial Engineering*, pages 567–571. IEEE, 2013. 16, 24

- Christopher Krauss, Xuan Anh Do, and Nicolas Huck. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the s&p 500. *European Journal of Operational Research*, 259(2):689–702, 2017. 16
- Eunsuk Chong, Chulwoo Han, and Frank C Park. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83:187–205, 2017. 16
- Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, 270(2):654–669, 2018. 16, 22
- Fengmei Yang, Zhiwen Chen, Jingjing Li, and Ling Tang. A novel hybrid stock selection method with stock prediction. *Applied Soft Computing*, 80:820–831, 2019. 16
- Sang Il Lee and Seong Joon Yoo. Threshold-based portfolio: the role of the threshold and its applications. *The Journal of Supercomputing*, 76(10):8040–8057, 2020. 16
- GT Sá. Mercado de ações e bolsas de valores. *Rio de Janeiro: Editora Técnica*, pages 39–48, 1987. 18
- Giacomo Di Tollo and Andrea Roli. Metaheuristics for the portfolio selection problem. *International Journal of Operations Research*, 5(1):13–35, 2008. 20
- Ana Carolina Lorena and André CPLF De Carvalho. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, 14(2):43–67, 2007. 21, 24
- Simon Haykin. Neural networks: a comprehensive foundation. 1999. *Mc Millan, New Jersey*, pages 1–24, 2010. 21
- David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986. 21
- Jürgen Schmidhuber, Sepp Hochreiter, et al. Long short-term memory. *Neural Comput*, 9(8):1735–1780, 1997. 22
- Vladimir Vapnik. *The nature of statistical learning theory*. Springer, 1995. 23
- Haiqin Yang, Laiwan Chan, and Irwin King. Support vector machine regression for volatile stock market prediction. In *International conference on intelligent data engineering and automated learning*, pages 391–396. Springer, 2002. 23
- Edwin J Elton, Martin J Gruber, Stephen J Brown, and William N Goetzmann. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009. 28