

CENTRO UNIVERSITÁRIO FEI

ENZO H SATUCHENGO BENVENUTI

GABRIEL TANAKA DE CARVALHO

LEONARDO MODESTO GUERRA

**EXPLORANDO RELAÇÕES E DESEMPENHOS: UM SISTEMA PARA ANÁLISE DE
ATIVOS LISTADOS NA B3**

SÃO BERNARDO DO CAMPO

2024

ENZO H SATUCHENGO BENVENUTI
GABRIEL TANAKA DE CARVALHO
LEONARDO MODESTO GUERRA

**EXPLORANDO RELAÇÕES E DESEMPENHOS: UM SISTEMA PARA ANÁLISE DE
ATIVOS LISTADOS NA B3**

Trabalho de Conclusão de Curso, apresentado
ao Centro Universitário FEI, como parte dos
requisitos necessários para obtenção do título de
Bacharel em Ciência da Computação.
Orientador: Prof. Ricardo de Carvalho Destro.

SÃO BERNARDO DO CAMPO

2024

AGRADECIMENTOS

Gostaríamos de agradecer profundamente às famílias que nos acompanharam e apoiaram, com seu amor e suporte constantes. Aos colegas e amigos que prestaram auxílio e trouxeram apelos reconfortantes nessa jornada desafiadora.

À Deus que até aqui nos acompanhou e ao professor e orientador Ricardo de Carvalho Destro pela orientação, dedicação e escuta para garantir o sucesso na entrega do trabalho final. Assim também, aos demais professores que fizeram parte desses anos estudantis com aprendizados e conhecimentos que levaremos para vida.

Ao Centro Universitário FEI pela estrutura, oportunidades, experiências e desenvolvimento intelectual, nossa gratidão eterna.

Por último, mas não menos importante, à todos que participaram indiretamente do planejamento, desenvolvimento e processo nessa etapa acadêmica.

Todos são essenciais para essa finalização e entrega.

*"O tempo é o recurso mais precioso que temos, e a sabedoria está em saber como usá-lo.
Investir em aprendizado e paciência, no longo prazo, é a chave para o sucesso."
(BUFFETT, 1997, tradução nossa)*

RESUMO

O conhecimento sobre investimentos ainda é limitado para grande parte da população, dificultando a poupança e o acesso a decisões financeiras eficazes. A volatilidade do mercado financeiro reforça a necessidade de metodologias avançadas para maximização de retornos e redução de riscos. Este trabalho se baseia na Teoria Moderna de Portfólio (TMP) de Markowitz e nos modelos de Fama e French, destacando a importância da diversificação e da avaliação de riscos, conforme argumentado por Bodie, Kane e Marcus. Utilizando técnicas como clustering, Random Forest e Simulação de Monte Carlo, foram desenvolvidas ferramentas analíticas para identificar padrões entre ativos da B3 e otimizar carteiras. Essas metodologias integram avanços em ciência de dados e aprendizado de máquina, promovendo análises robustas e preditivas. A principal entrega do projeto é um software de otimização de portfólios, que combina previsões de retornos históricos com estratégias quantitativas, oferecendo suporte acessível e eficaz à tomada de decisão no mercado de investimentos. Este trabalho destaca como a integração de conceitos matemáticos e estatísticos pode transformar a prática de investimentos, promovendo inovação e acessibilidade no mercado financeiro brasileiro.

Palavras-chave: Investimento; Otimização; Riscos; Dados; Previsões.

ABSTRACT

Knowledge about investments is still limited for a large part of the population, making it difficult to save and access effective financial decisions. Financial market volatility reinforces the need for advanced methodologies to maximize returns and reduce risks. This paper is based on Markowitz's Modern Portfolio Theory (MPT) and the Fama and French models, highlighting the importance of diversification and risk assessment, as argued by Bodie, Kane and Marcus. Using techniques such as clustering, Random Forest and Monte Carlo Simulation, analytical tools were developed to identify patterns among B3 assets and optimize portfolios. These methodologies integrate advances in data science and machine learning, promoting robust and predictive analyses. The main deliverable of the project is a portfolio optimization software, which combines historical return forecasts with quantitative strategies, offering accessible and effective support for decision-making in the investment market. This work highlights how the integration of mathematical and statistical concepts can transform investment practices, promoting innovation and accessibility in the Brazilian financial market.

Keywords: Investments; Optimization; Risks; Data; Predictions.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo Análise de Correlação	14
Figura 2 – Exemplo funcionamento Random Forest	18
Figura 3 – Limpeza inicial dos dados	24
Figura 4 – Função para cálculos dos retornos diários	24
Figura 5 – Cálculos dos retorno esperado e matriz de covariância	25
Figura 6 – Cálculo do VaR e CVaR	26
Figura 7 – Cálculo do Drawdown Máximo	26
Figura 8 – Tratamento inicial dos dados	28
Figura 9 – Remover valores ausentes dos dados processados	28
Figura 10 – 10 maiores correlações dos ativos processados	30
Figura 11 – Carteira otimizada	31
Figura 12 – Retorno Acumulado da Carteira Otimizada e Ativos Individuais	32
Figura 13 – Drawdown ao longo do tempo para Carteira Otimizada	32
Figura 14 – VaR e CVaR em diferentes níveis de confiança	33
Figura 15 – Simulação de Monte Carlo com cenários de estresse	34
Figura 16 – Retorno Acumulado Final Carteira Risk Parity e Selic	35
Figura 17 – Retorno Acumulado Final Carteira Dividendos e Selic	37
Figura 18 – Retorno Acumulado Final Carteira Momentum e Selic	38
Figura 19 – Previsão de resultados para a próxima semana da Carteira Otimizada	40
Figura 20 – Comparação rebalanceamento da carteira otimizada	43
Figura 21 – Processo de rebalanceamento da Carteira Otimizada	44
Figura 22 – Análise de Sensibilidade NTCO3: Peso vs. Retorno/Volatilidade	45
Figura 23 – Análise de Sensibilidade PETR4: Peso vs. Retorno/Volatilidade	46

LISTA DE ABREVIATURAS E SIGLAS

B3	Bolsa de Valores
BCB	Banco do Brasil
DY	Dividend Yields
ETFs	Exchange-traded Funds
FIIs	Fundos de Investimento Imobiliário
IMPA	Instituto de Matemática Pura e Aplicada
IBM	International Business Machines Corporation
MPT	Modern Portfolio Theory
MSE	Erro Quadrático Médio
SBM	Sociedade Brasileira de Matemática
SI	Sistema Internacional
TMP	Teoria Moderna do Portfólio
VaR	Value at Risk
CVaR	Conditional Value at Risk

SUMÁRIO

1	INTRODUÇÃO	10
1.1	OBJETIVO	11
1.2	JUSTIFICATIVA	11
1.3	ESTRUTURA DO TRABALHO	11
2	CONCEITOS FUNDAMENTAIS	12
2.1	TEORIA MODERNA DO PORTFÓLIO	12
2.2	ANÁLISE DE CORRELAÇÃO	13
2.3	ESTATÍSTICA E MÉTODOS QUANTITATIVOS	14
2.4	DESEMPENHO DE ATIVOS	15
2.5	SHARPE RATIO	16
2.6	RISCO LIVRE	16
2.7	SIMULAÇÃO DE MONTE CARLO	17
2.8	RANDOM FOREST	18
3	TRABALHOS RELACIONADOS	19
3.1	PROJETOS RELACIONADOS A OTIMIZAÇÃO DE PORTFÓLIOS POR ABORDAGENS TRADICIONAIS	19
3.2	PROJETOS RELACIONADOS A OTIMIZAÇÃO DE PORTFÓLIOS POR ABORDAGENS MODERNAS	21
4	METODOLOGIA	23
4.1	DEFINIÇÃO DO PROBLEMA E OBJETIVOS	23
4.2	COLETA E PREPARAÇÃO DOS DADOS	24
4.3	CÁLCULO DE MÉTRICAS FINANCEIRAS	25
4.4	DESENVOLVIMENTO DE SOFTWARE	26
4.5	SIMULAÇÕES E BACKTESTING	27
4.5.1	CARTEIRA OTIMIZADA	29
4.5.2	PERFORMANCE RISK PARTY	34
4.5.3	PERFORMANCE CARTEIRA DE DIVIDENDOS	36
4.5.4	PERFORMANCE CARTEIRA DE MOMENTUM	37

4.5.5	PREVISÃO DE RESULTADOS	39
4.5.6	REBALANCEAMENTO DA CARTEIRA	41
5	RESULTADOS	43
5.1	Integrando as soluções	43
6	CONSIDERAÇÃO FINAIS	47
7	CONCLUSÃO	49
	REFERÊNCIAS	51

1 INTRODUÇÃO

A educação financeira é uma realidade que poucos têm acesso no Brasil. Dessa forma, poupar dinheiro não é uma realidade para a maioria dos brasileiros e o investimento no mercado financeiro muitas vezes soa como algo elitizado.

Além disso, saber onde investir o montante poupado pode se tornar um verdadeiro desafio. O investimento é tão importante quanto a poupança em si, pois sem uma aplicação adequada, todo o esforço para economizar pode ser desperdiçado em um investimento mal-sucedido.

O mercado financeiro é um ambiente dinâmico e complexo, onde a análise de ativos desempenha um papel crucial na tomada de decisões estratégicas de investimento. Segundo Markowitz (1952), ao propor a Teoria Moderna do Portfólio, a diversificação de ativos pode reduzir o risco sem sacrificar o retorno esperado. Esse conceito é particularmente relevante no contexto da Bolsa de Valores Brasileira (B3), onde uma compreensão profunda das relações entre ativos e seus desempenhos pode proporcionar vantagens competitivas a investidores e analistas.

Ao longo das últimas décadas, a análise quantitativa ganhou força, com contribuições significativas de Fama e French (1992) ao descreverem os fatores de risco que impactam o desempenho dos ativos. A correlação entre diferentes ativos, conforme destacado por Bodie, Kane e Marcus (2014), não só revela padrões de comportamento, como também auxilia na mitigação de riscos através da construção de portfólios mais eficientes. Assim, compreender essas relações se torna essencial para uma alocação de recursos mais estratégica.

Com o avanço da tecnologia e o crescimento do volume de dados disponíveis, ferramentas mais sofisticadas têm sido desenvolvidas para a análise de ativos financeiros, como ressaltado por Fabozzi et al. (2014) em seus estudos sobre a integração de ciência de dados e finanças. Nesse contexto, este trabalho propõe o desenvolvimento de um sistema de análise que visa identificar correlações entre ativos listados na B3, além de comparar seus desempenhos ao longo de períodos específicos e prever seu comportamento. A aplicação de métodos quantitativos, como regressões, permitirá a criação de um sistema acessível para todos que buscam insights mais profundos no mercado financeiro.

Portanto, ao longo deste estudo, serão explorados alguns fatores que influenciam a correlação entre ativos e o comportamento destes ativos podem ser usados para otimizar a construção de portfólios.

1.1 OBJETIVO

Desenvolver um software capaz de auxiliar desde investidores até analistas de mercado a fazer uma análise mais prática e precisa, afim de facilitar tomadas de decisão no momento de investir e aprimorar a criação de portfólio lucrativos com investimento em diferentes classes de ativos para reduzir o risco total, aliando conceitos teóricos com tecnologias modernas de dados.

1.2 JUSTIFICATIVA

O mercado financeiro brasileiro, representado pela B3, é caracterizado por sua volatilidade e pela diversidade de ativos disponíveis para negociação. Para investidores e analistas, a tarefa de selecionar ativos adequados, equilibrando risco e retorno, é desafiadora. Nesse contexto, a compreensão das correlações entre ativos e a otimização de carteiras se tornam fundamentais para a tomada de decisões estratégicas. Embora essa abordagem seja amplamente conhecida no meio acadêmico e entre investidores, a disponibilidade de sistemas práticos e acessíveis que facilitem essas análises ainda é limitada, especialmente no contexto brasileiro.

1.3 ESTRUTURA DO TRABALHO

O trabalho será dividido por seções descritas abaixo:

Ao Capítulo 2 serão apresentados conceitos utilizados nos estudos das bases de dados.

Logo após, no Capítulo 3 serão citados trabalhos relevantes anteriores relacionados ao tema do projeto, com objetivo de demonstrar técnicas modernas e tradicionais.

No capítulo 4 será detalhada a metodologia que foi utilizada para o desenvolvimento do software, explicitando a definição dos objetivos, desenvolvimento e testes, coleta de dados e cálculos realizados.

E seguindo a lógica, no Capítulo 5 serão apresentados os resultados finais obtidos após estudo e desenvolvimento.

O Capítulo 6 contará brevemente sobre as considerações finais destes resultados. O que conecta ao Capítulo 7, onde será apresentada a conclusão deste trabalho.

2 CONCEITOS FUNDAMENTAIS

2.1 TEORIA MODERNA DO PORTFÓLIO

Introduzida por Harry Markowitz em 1952, a Teoria Moderna do Portfólio (TMP) propõe a ideia de que investidores podem maximizar o retorno esperado de um portfólio para um dado nível de risco ao diversificar os ativos. A diversificação permite que os riscos específicos de ativos individuais sejam mitigados, o que resulta em uma relação mais eficiente entre risco e retorno. Essa abordagem desafiou a visão tradicional, que muitas vezes considerava a análise de ativos de forma isolada, sem levar em conta a interação entre eles.

Em relação ao risco, a TMP enfatiza que o risco de um portfólio não é simplesmente a soma dos riscos individuais dos ativos. O risco de portfólio é influenciado pela covariância entre os ativos, o que significa que seu comportamento interage para afetar o risco global. Mesmo que um ativo individual tenha alta volatilidade, se ele apresentar uma correlação baixa com outros ativos, pode ajudar a reduzir o risco geral do portfólio. Portanto, a diversificação eficaz, de acordo com a TMP, não se baseia apenas na quantidade de ativos, mas em sua escolha, cujas correlações minimizem o risco total do portfólio. Markowitz e sua teoria revolucionaram o mundo das finanças, alterando radicalmente a forma como os investidores abordam a alocação e a diversificação de ativos.

Outro conceito importante e inovador na TMP é a Fronteira Eficiente, ilustrada por um gráfico que exibe as combinações ótimas de risco e retorno. Esta mostra todos os portfólios que oferecem o maior retorno possível para um dado nível de risco. Assim, qualquer portfólio fora da Fronteira Eficiente é considerado subótimo, pois existe uma combinação de ativos que poderia gerar um retorno maior para o mesmo nível de risco ou reduzi-lo para um retorno esperado equivalente.

No contexto do presente projeto, a TMP será aplicada para construir portfólios que maximizem o retorno esperado de acordo com um nível de risco predefinido. Para isso, será realizada uma análise cuidadosa das correlações, dos retornos esperados e dos riscos de cada ativo, a fim de otimizar a alocação de recursos. A combinação de ativos será feita de forma a situar o portfólio na Fronteira Eficiente, garantindo que o retorno máximo possível seja alcançado dentro dos limites de risco estabelecidos. A aplicação da TMP no projeto permite que os investidores possam tomar decisões informadas sobre como alocar seus recursos de maneira a equilibrar o

risco e o retorno de maneira eficaz, ao mesmo tempo em que busca a diversificação e a redução da exposição a riscos sistemáticos.

2.2 ANÁLISE DE CORRELAÇÃO

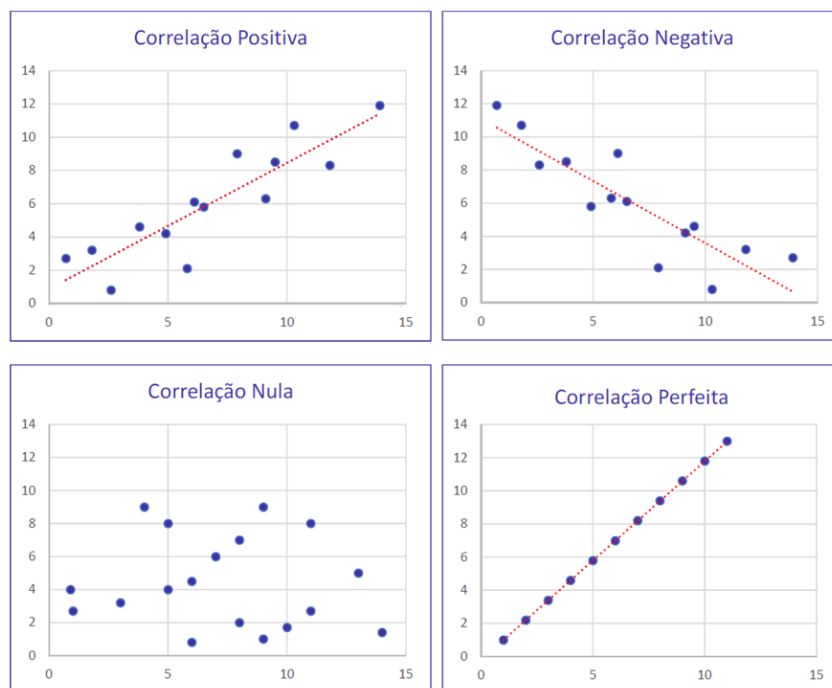
A correlação é uma medida estatística amplamente utilizada para indicar o grau e a direção da relação linear entre dois ativos financeiros. A correlação mede a relação entre dois ou mais ativos financeiros, indicando como seus preços se movem em relação um ao outro. No contexto de investimentos, a correlação é uma ferramenta essencial para compreender o comportamento dos preços de diferentes ativos e como eles interagem entre si. Se dois deles têm uma correlação positiva, isso significa que seus preços tendem a se mover na mesma direção. Por exemplo, uma correlação de $+0,8$ entre duas ações indica que, em média, quando o preço de uma ação sobe, a outra também tende a subir.

Uma correlação negativa, por sua vez, indica que os preços dos ativos tendem a se mover em direções opostas. Por exemplo, uma correlação de $-0,8$ entre dois ativos sugere que, quando o preço de um sobe, o outro tende a cair. Esse tipo de correlação é extremamente útil para a construção de portfólios diversificados, já que ativos com correlação negativa podem ajudar a mitigar os riscos, uma vez que as perdas de um ativo podem ser compensadas pelos ganhos de outro. A diversificação de um portfólio com ativos negativamente correlacionados costumam reduzir a volatilidade, promovendo um equilíbrio maior entre risco e retorno (Elton e Gruber, 1997).

Já uma correlação nula, ou próxima de zero, significa que não há uma relação linear clara entre os movimentos de preço dos dois ativos. Isso implica que possam se mover de maneira independente, sem que um influencie diretamente o outro. Embora essa falta de correlação não forneça um benefício claro em termos de diversificação, ela pode ser interessante para investidores que buscam ativos independentes, ou que desejam estratégias baseadas em movimentos de preços específicos de cada ativo, sem dependência dos outros no portfólio (Figura 1).

A análise dessa métrica é uma das etapas cruciais no processo de otimização de portfólios, pois permite que investidores encontrem ativos que se complementem em termos de movimento de preço.

Figura 1 – Exemplo Análise de Correlação



Fonte: Oliveira, K.P et al. (2023) Ciência de Dados na Educação Pública.

2.3 ESTATÍSTICA E MÉTODOS QUANTITATIVOS

As ferramentas estatísticas e os métodos quantitativos desempenham um papel crucial na transformação de dados financeiros brutos em informações mensuráveis, que podem ser utilizadas para otimizar a alocação de ativos e melhorar a tomada de decisões no mercado financeiro. Métodos como regressão linear, análise de clusters e análise de séries temporais são amplamente aplicados em finanças para prever tendências, identificar padrões e otimizar portfólios de investimentos.

A regressão linear é um exemplo de técnica estatística fundamental para prever o retorno de um ativo com base em uma ou mais variáveis. No contexto de investimentos, investidores podem utilizar para tentar explicar o retorno de um ativo com base em fatores como o comportamento geral do mercado, taxas de juros ou outros indicadores econômicos. A regressão linear é amplamente utilizada para modelar relações entre variáveis e para prever como as mudanças em uma variável explicativa (como o preço do petróleo) podem impactar o retorno de um ativo específico (Granger e Newbold, 2014).

A análise de clusters é outra ferramenta valiosa, que agrupa ativos com características semelhantes. Essa também é uma técnica útil para a construção de portfólios diversificados, pois permite ao investidor identificar grupos de ativos que podem se comportar de maneira similar em

diferentes condições de mercado, auxiliando na eficiência da diversificação (Everitt et al., 2011).

Neste projeto, a integração desses métodos proporcionará uma boa base para a otimização de portfólios. A regressão linear permitirá prever os retornos dos ativos, enquanto a análise de clusters ajudará a identificar grupos que compartilhem características de risco semelhantes, facilitando sua diversificação e gestão.

2.4 DESEMPENHO DE ATIVOS

O desempenho de ativos é uma avaliação crucial para determinar o sucesso de um investimento, pois fornece uma medida de quão eficaz um ativo ou portfólio tem sido em gerar retornos ao longo do tempo. O desempenho é normalmente medido pelos retornos gerados durante um período específico, mas há várias formas de calcular esses retornos, dependendo do tipo de análise e das variáveis que se deseja considerar.

O retorno nominal é a medida mais simples, refletindo o ganho bruto sobre o investimento, sem levar em conta outros fatores, como a inflação ou o risco. Esse retorno é simplesmente a diferença entre o valor final do investimento e seu valor inicial, expresso como uma porcentagem. Embora o retorno nominal forneça uma boa visão inicial do desempenho de um ativo ou carteira, pode ser enganoso caso não considere o impacto da inflação, que pode reduzir o poder de compra do retorno obtido (Bodie, Kane e Marcus, 2014).

O retorno ajustado pela inflação considera a inflação ao calcular o ganho real de poder de compra e permite uma avaliação mais precisa do quanto o investimento contribuiu para aumentar o poder de compra do investidor, ao invés de simplesmente aumentar o valor nominal do portfólio (Campbell e Viceira, 2002).

Uma abordagem mais completa é o retorno ajustado ao risco, que é fundamental para a avaliação do desempenho de ativos quando se leva em conta o risco envolvido. O Sharpe Ratio é uma das métricas mais comuns utilizadas para esse fim. O Sharpe Ratio mede o retorno em excesso de um ativo em relação ao ativo livre de risco (geralmente títulos do governo) e divide esse retorno pelo desvio padrão dos retornos, ajustando o desempenho para o nível de risco associado. Quanto maior o Sharpe Ratio, melhor é o retorno do ativo em relação ao risco assumido. Essa métrica ajuda os investidores a comparar ativos de diferentes perfis de risco de forma mais refinada, permitindo a seleção de ativos que ofereçam o melhor retorno ajustado ao risco (Sharpe, 1966).

2.5 SHARPE RATIO

Desenvolvido por William F. Sharpe em 1966, o Sharpe Ratio é uma métrica amplamente utilizada no mundo das finanças para avaliar o desempenho de um ativo, levando em conta o risco envolvido.

O Sharpe Ratio é calculado pela fórmula (1) :

$$sharperatio = R_p - R_f / \sigma_p \quad (1)$$

Onde:

- R_p é o retorno esperado do portfólio ou investimento.
- R_f é a taxa livre de risco (geralmente o retorno de um ativo de baixo risco, como títulos públicos).
- σ_p é o desvio padrão (volatilidade) dos retornos do portfólio, que é uma medida de risco.

A interpretação do Sharpe Ratio é a seguinte: um valor superior a 1 indica um bom retorno ajustado ao risco, enquanto um valor igual a 1 significa que o retorno é adequado em relação ao risco. Já um valor inferior a 1 sugere que o retorno não é proporcional ao risco, e um valor negativo indica que o investimento tem desempenho pior que a taxa livre de risco. Entre suas vantagens, destacam-se a simplicidade, pois é fácil de calcular e entender, e a possibilidade de comparar diferentes investimentos com base no retorno ajustado ao risco.

Em termos de aplicações, o Sharpe Ratio é útil para a seleção de investimentos, ajudando a comparar ativos ou portfólios, e também na gestão de portfólios, ao auxiliar na otimização da alocação de ativos.

2.6 RISCO LIVRE

O risco livre, ou taxa livre de risco, é o rendimento de um investimento considerado totalmente seguro, sem risco de perda. Normalmente, esse rendimento vem de ativos como títulos do governo de baixo risco, como os Treasuries nos Estados Unidos ou o Tesouro Nacional no Brasil. Em um portfólio de investimentos, a taxa livre de risco serve como uma base de comparação, já que qualquer ativo de risco deve gerar um retorno superior ao risco livre para compensar a incerteza e volatilidade envolvidas.

Quando você está construindo um portfólio de investimento, uma das premissas fundamentais é que você precisa avaliar o trade-off entre risco e retorno. O risco livre é um ponto de referência para esse trade-off. Além disso, incluir ativos de risco livre pode ser uma excelente estratégia para equilibrar a volatilidade do portfólio.

No contexto de simulações de Monte Carlo ela pode ser extremamente relevante para calcular a performance ajustada ao risco, pois possibilita ver quantas das suas simulações ultrapassam a taxa livre de risco, ou quantas ficam abaixo dessa taxa, o que indicaria o impacto do risco no portfólio

2.7 SIMULAÇÃO DE MONTE CARLO

A simulação de Monte Carlo é uma técnica estatística empregada para modelar a incerteza e estimar possíveis resultados futuros de processos ou sistemas complexos, levando em consideração diferentes cenários e variáveis aleatórias. Essa abordagem é amplamente utilizada em diversas áreas, especialmente neste contexto, onde serve para avaliar riscos e simular o comportamento de portfólios de investimentos ao longo do tempo.

No contexto de um portfólio de investimentos, a simulação de Monte Carlo tem início com os dados históricos de retornos dos ativos, como ações ou títulos. Esses dados são utilizados para calcular a média, a volatilidade e as correlações entre os ativos. Com base nesses parâmetros, a técnica gera milhares de cenários possíveis para os retornos futuros, assumindo que os retornos dos ativos seguem uma distribuição estatística normal. Cada cenário simula a evolução dos preços dos ativos ao longo do tempo, considerando tanto a variabilidade dos retornos quanto as interações entre os ativos do portfólio.

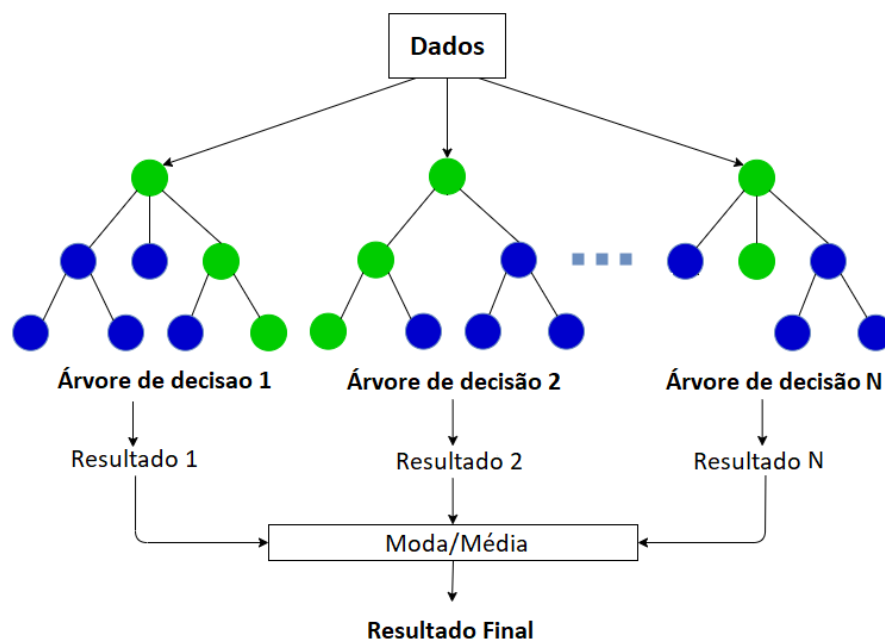
A partir das milhares de simulações geradas, o valor final do portfólio é calculado em cada cenário, e isso permite ao investidor obter uma distribuição dos possíveis resultados. Dessa forma, pode oferecer uma visão mais detalhada mostrando como o portfólio pode se comportar sob diferentes condições de mercado, levando em consideração os pesos atribuídos a cada ativo. Além disso, a simulação pode ser usada para testar cenários de estresse, como uma queda repentina nos mercados financeiros, para ver como o portfólio se comportaria sob condições extremas.

2.8 RANDOM FOREST

De acordo com a IBM (2024) o Random Forest é um algoritmo de aprendizado de máquina que utiliza um conjunto de árvores de decisão para melhorar a precisão e a robustez de previsões, sendo aplicado em tarefas de classificação e regressão. Como ilustrado na Figura 2, o método cria diversas árvores a partir de subconjuntos aleatórios dos dados de treinamento e das características disponíveis, promovendo a diversidade no modelo e mitigando o risco de overfitting. A combinação dessas árvores é feita por votação majoritária no caso de classificação, ou pela média no caso de regressão, o que garante uma maior capacidade de generalização para dados novos. Dessa forma, no caso do presente projeto o retorno futuro da carteira será previsto com base na média das previsões das árvores da floresta.

Ainda de acordo com a IBM, uma das grandes vantagens do Random Forest é que ele é resistente ao overfitting, ou seja, ele tem uma capacidade maior de generalizar para novos dados, mesmo quando as árvores de decisão individuais são complexas. Isso ocorre porque a combinação de várias árvores independentes ajuda a minimizar o impacto de ruídos ou dados específicos do conjunto de treinamento. O Random Forest também é muito versátil, podendo ser usado tanto para problemas de classificação quanto para regressão, e pode lidar bem com dados faltantes e variáveis categóricas sem exigir pré-processamento complexo.

Figura 2 – Exemplo funcionamento Random Forest



Fonte: Adilane Ribeiro da Silva (2021) Modelos de Machine Learning utilizando o pacote Caret.

3 TRABALHOS RELACIONADOS

Este capítulo visa citar artigos relacionados com o tema deste trabalho, além de ressaltar os diferenciais do mesmo em relação ao que já foi feito. Para tal, consideramos relevante separar pelas seguintes subseções:

- **Projetos relacionados a otimização de portfólios por abordagens tradicionais:** Neste tópico serão apresentados referências que serviram de base para as análises de risco neste trabalho.
- **Projetos relacionados a otimização de portfólios por abordagens modernas:** Serão apresentados referências que buscam transcender dos métodos 'tradicionais' e trazem uma aplicação mais moderna, isto é, utilizando abordagens com aplicações computacionais e com técnicas de simulações variadas.

3.1 PROJETOS RELACIONADOS A OTIMIZAÇÃO DE PORTFÓLIOS POR ABORDAGENS TRADICIONAIS

Artigos que possuem abordagens 'tradicionais' geralmente utilizam como base a métodos consolidados na teoria moderna de finanças, especificamente o modelo de Harry Markowitz (1952) em Modern Portfolio Theory (MPT)(Markowitz, 1952). Este modelo é fundamental para a construção de carteiras eficientes, que buscam otimizar o equilíbrio entre retorno e risco. A principal premissa dessa teoria é que um portfólio de investimentos não deve ser avaliado com base no desempenho de seus ativos individuais, mas sim em sua combinação, levando em consideração a correlação entre eles. A abordagem clássica por média-variância é central nesse modelo e utiliza a relação entre o retorno esperado e o risco, que é representado pela variância ou desvio-padrão dos retornos dos ativos, para construir carteiras eficientes.

Projetos nessa linha costumam determinar as expectativas de retorno dos ativos e calcular a matriz de covariância entre eles para ter uma medida de como se comportam em relação uns aos outros. A partir disso, são aplicados alguns métodos matemáticos visando encontrar a combinação de ativos que maximize o retorno esperado para um dado nível de risco ou o oposto disso, minimizando o risco para um nível de retorno esperado. Além disso, as análises comumente consideram restrições, como a imposição de pesos para os ativos ou a definição de

limites máximos e mínimos para alocação em cada um, de modo a garantir a diversificação e evitar concentrações excessivas em ativos específicos (Sharpe, 1964). Esses princípios formam a base da implementação de ferramentas como o pacote PyPortfolioOpt, utilizado neste projeto para aplicar o modelo de forma programática.

Outrossim, diversos artigos que seguem essa vertente sobre otimização de portfólios também discutem a incorporação de métricas alternativas de risco, visando expandir os aspectos considerados do risco financeiro. Isso se dá pelo fato de que a variância, embora amplamente utilizada, não captura todos os aspectos do risco financeiro, especialmente em cenários de perdas extremas. Um dos conceitos frequentemente abordados é o Value at Risk (VaR), que estima a perda máxima esperada em um intervalo de confiança específico, como 1% ou 5%, ao longo de um horizonte de tempo determinado (Jorion, 2007). O Conditional Value at Risk (CVaR) é uma extensão do VaR, que calcula a média das perdas que excedem o VaR em situações de eventos extremos, sendo uma medida mais sensível para riscos de cauda causado por eventos improváveis, que ocorrem nas extremidades das distribuições de probabilidade (Rockafellar e Uryasev, 2002). Outra métrica considerada relevante e comumente utilizada é o drawdown máximo, que mede a maior perda acumulada de um pico a um vale, indicando a profundidade da queda do valor de um portfólio antes de sua recuperação. Essas métricas são úteis para avaliar o desempenho e a robustez das carteiras, especialmente em mercados voláteis ou durante crises financeiras (Jorion, 2007). Neste projeto, essas métricas são integradas para fornecer uma visão mais detalhada do risco e da resiliência da carteira em possíveis cenários adversos.

A diversificação é outro conceito central abordado em projetos sobre otimização de portfólios. A ideia de reduzir o risco idiossincrático, ou seja, o risco específico de um ativo, ao combinar ativos que se comportam de maneira descorrelacionada é extensamente explorada. A exploração desse conceito se dá, em grande parte, devido ao fato de que diversificação entre ativos de diferentes classes e setores pode reduzir significativamente o risco total do portfólio sem comprometer o retorno esperado (Markowitz, 1952).

A análise da fronteira eficiente também é frequentemente utilizada para identificar carteiras ótimas, ou seja, carteiras que oferecem o melhor equilíbrio entre risco e retorno. É válido lembrar que o uso de benchmarks como índices de mercado é uma forma estimada pelos pesquisadores para avaliar a performance relativa de uma carteira. Uma das métricas utilizadas nesse contexto é o Tracking Error, que serve para avaliar se a carteira está seguindo o caminho esperado, mostrando o quanto a carteira 'desvia' do seu benchmark. Ao analisar

a volatilidade dessa diferença, podemos avaliar a capacidade da carteira em acompanhar as oscilações do mercado e verificar se o portfólio está atingindo os objetivos de risco-retorno de forma eficiente (Fama e French, 1993).

Os conceitos tradicionais de otimização enfrentam algumas limitações que acabam afetando os resultados obtidos. Michael J. Brennan discute a dificuldade de lidar com a incerteza nos mercados financeiros, especialmente em relação à avaliação de ativos e à determinação de retornos futuros. Ele reconhece que os mercados são dinâmicos e que a volatilidade torna as estimativas de retorno e a matriz de covariância dos ativos altamente incertas, o que é um desafio central na construção de carteiras eficientes. Uma das principais dificuldades reside na necessidade de estimativas precisas tanto para o retorno esperado quanto para a matriz de covariância dos ativos, o que pode ser problemático devido à natureza incerta e volátil dos mercados financeiros. Além disso, esses modelos são altamente sensíveis às mudanças nos dados de entrada, o que pode resultar em carteiras ineficazes quando os parâmetros se alteram. Para superar essas limitações, o presente projeto incorporou técnicas de simulação, como simulações de Monte Carlo e o uso de cenários históricos, visando gerar análises mais dinâmicas que não dependem exclusivamente de parâmetros estáticos.

Portanto, os conceitos tradicionais de otimização de portfólios, embora muito poderosos, sofrem com limitações que exigem a adoção de métodos complementares para aumentar a precisão e a robustez das soluções. O desenvolvimento deste projeto busca combinar essas abordagens clássicas com novas técnicas de avaliação de risco e simulação, oferecendo uma análise mais abrangente e adaptável dos portfólios financeiros.

3.2 PROJETOS RELACIONADOS A OTIMIZAÇÃO DE PORTFÓLIOS POR ABORDAGENS MODERNAS

Os artigos sobre otimização de portfólios que adotam abordagens 'modernas' se diferenciam da Teoria Moderna de Portfólios de Markowitz ao incorporar inovações tecnológicas, dados mais complexos e técnicas computacionais avançadas (Dempster e Leemans, 2007). Essas abordagens buscam utilizar ferramentas mais flexíveis e personalizáveis, como algoritmos heurísticos, inteligência artificial e métodos probabilísticos, para criar soluções mais dinâmicas e adaptáveis no processo de otimização de portfólios. As metodologias contemporâneas enfatizam a utilização de dados mais ricos e a exploração de novas técnicas computacionais, como Machine Learning, para melhorar os resultados das estratégias de investimento.

Um dos principais avanços no campo da otimização de portfólios modernos é o uso de Machine Learning. He e Liao investigam como o uso de modelos preditivos podem ser utilizados para prever retornos futuros, o que está diretamente relacionado à busca por melhores decisões de investimento. Eles discutem como técnicas como redes neurais e regressão com base em características financeiras podem ser aplicadas para estimar os retornos de maneira mais precisa do que abordagens tradicionais. Além disso, é comumente enfatizado em artigos que seguem essa abordagem a importância de utilizar grandes volumes de dados financeiros (dados 'ricos') e aplicar métodos de Machine Learning para identificar padrões e tendências que podem não ser evidentes com técnicas tradicionais.

A inclusão de simulações de cenários extremos também tem se tornado uma prática comum, utilizando dados históricos e simulações de Monte Carlo para incluir eventos disruptivos, como choques de mercado, pandemias ou crises financeiras, no processo de decisão. No presente projeto, as simulações de Monte Carlo foram incorporadas para analisar a robustez das carteiras diante da incerteza, ampliando as análises com base em diversos cenários. Isso proporcionou uma visão mais dinâmica e realista das possíveis reações das carteiras a choques imprevistos, visando proteger a carteira do risco de cauda (Bertsimas et al., 2004).

Portanto, as abordagens modernas de otimização de portfólios buscam superar as limitações sofridas pelos modelos tradicionais, incorporando tecnologias mais sofisticadas e métodos computacionais avançados. O desenvolvimento deste projeto buscou refletir essas tendências, buscando oferecer uma solução mais dinâmica e adaptável.

4 METODOLOGIA

A metodologia deste projeto seguiu uma abordagem estruturada, dividida em etapas interconectadas que englobaram desde a definição do problema até a implementação de uma solução prática e adaptável para otimização de portfólios.

4.1 DEFINIÇÃO DO PROBLEMA E OBJETIVOS

O projeto foi iniciado com a identificação dos principais desafios enfrentados por investidores no mercado financeiro brasileiro, com ênfase na B3, a bolsa de valores do Brasil. Esses desafios incluem a complexidade na alocação de ativos, a gestão de riscos e a necessidade de ferramentas adequadas para otimizar o desempenho dos investimentos. A partir dessa análise, os objetivos principais do projeto foram definidos para atender a essas demandas específicas do mercado.

O principal objetivo foi desenvolver um software capaz de criar portfólios otimizados, utilizando conceitos avançados de finanças, como a Teoria Moderna do Portfólio aliado a técnicas modernas de análise de dados. O software foi projetado para realizar otimizações que maximizam o retorno esperado para um dado nível de risco ou minimizam o risco para um nível de retorno pré-definido. Para isso, o software incorpora métodos quantitativos como regressão linear, análise de séries temporais e análise de clusters, que permitem uma análise mais aprofundada do comportamento dos ativos e a construção de portfólios mais eficientes.

Além disso, o projeto visou incorporar métricas de risco e retorno além das tradicionais, com o objetivo de fornecer uma visão mais completa e robusta sobre os investimentos. Entre as métricas selecionadas, destacam-se o Value at Risk (VaR), que estima a perda máxima esperada para um dado nível de confiança, e o Conditional Value at Risk (CVaR), que vai além do VaR ao avaliar a média das perdas que superam o limite do VaR. Essas métricas são fundamentais para medir os riscos de cauda, que são eventos extremos no mercado, como crises financeiras ou quedas abruptas de preços.

O último objetivo foi implementar formas de comparar a performance da carteira otimizada desenvolvida pelo software com outras estratégias de investimento já conhecidas, além de mantermos uma comparação com o índice de inflação.

4.2 COLETA E PREPARAÇÃO DOS DADOS

Os dados financeiros necessários para o projeto foram coletados por meio da API do Yahoo Finance, que oferece acesso a uma ampla gama de informações financeiras, incluindo o histórico de preços das ações listadas na B3. Essa abordagem permitiu uma coleta de dados eficiente e abrangente, essencial para o desenvolvimento do projeto.

A primeira etapa da preparação foi a limpeza de dados, que envolveu a remoção de ativos com alta incidência de valores ausentes (Figura 3). A presença de dados faltantes é um problema comum em séries temporais financeiras, podendo distorcer os resultados das análises se não forem tratados adequadamente. Além disso, foi realizado o preenchimento de lacunas nas séries temporais, o que é essencial para garantir que as sequências de dados de preços estejam completas e consistentes ao longo do tempo. A limpeza e o tratamento adequado dos dados são etapas cruciais para garantir a integridade das análises e dos modelos financeiros desenvolvidos, pois a precisão dos resultados depende diretamente da qualidade dos dados utilizados (Pritsker, 2013).

Figura 3 – Limpeza inicial dos dados

```
# Limpar os dados removendo colunas com muitos NaNs e preenchendo os restantes
df_fechamento = df_fechamento.dropna(axis=1, thresh=int(0.8 * len(df_fechamento))) # Remove ações com mais de 20% de NaNs
df_fechamento = df_fechamento.fillna(method='ffill').fillna(method='bfill') # Preenche os NaNs restantes
```

Fonte: Autores.

A etapa seguinte foi a transformação dos dados, que incluiu o cálculo dos retornos diários e anuais para todos os ativos selecionados (Figura 4). O cálculo dos retornos é uma prática fundamental em finanças, pois permite avaliar o desempenho dos ativos ao longo do tempo. Os retornos diários foram calculados com base na variação percentual dos preços de fechamento dos ativos entre dois dias consecutivos, enquanto os retornos anuais foram obtidos pela soma dos retornos diários durante o ano. Essa transformação é importante para que os dados estejam prontos para a análise de risco e otimização de portfólios, além de facilitar a comparação entre diferentes ativos ao longo do tempo.

Figura 4 – Função para cálculos dos retornos diários

```
# Função para calcular o retorno esperado (média dos retornos diários)
def calcular_retorno_esperado(df):
    retornos_diarios = df.pct_change(fill_method=None).mean() * 252 # Anu
    return retornos_diarios
```

Fonte: Autores.

Por fim, foi realizada a estruturação dos dados, com a criação de um DataFrame consolidado para análise. Esse DataFrame inclui as informações de preços ajustados e os retornos esperados de todos os ativos selecionados. A utilização de um DataFrame estruturado é fundamental para facilitar a análise e a manipulação dos dados, além de ser a base para a construção de modelos quantitativos de otimização de portfólios. Esses processos garantiram a qualidade e a consistência dos dados, permitindo realizar as simulações e otimizações necessárias para a construção de portfólios eficientes.

4.3 CÁLCULO DE MÉTRICAS FINANCEIRAS

Para a análise dos ativos e a construção do portfólio, foram calculadas tanto as métricas básicas quanto as avançadas, a fim de garantir uma avaliação abrangente do desempenho e risco da carteira. Essas métricas fornecem as informações necessárias para otimizar a alocação de ativos e melhorar a tomada de decisões de investimento.

Os retornos esperados foram calculados utilizando a média histórica dos retornos dos ativos (Figura 4). A média histórica é uma das abordagens mais comuns para estimar os retornos futuros, já que se baseia no desempenho passado dos ativos.

A matriz de covariância foi implementada para estimar a variabilidade conjunta entre os ativos (Figura 5). A covariância mede o grau de variação simultânea entre dois ativos, ou seja, se eles tendem a se mover na mesma direção ou em direções opostas. Quando os ativos têm baixa covariância entre si, a diversificação tende a reduzir o risco da carteira como um todo.

Figura 5 – Cálculos dos retorno esperado e matriz de covariância

```
# Calcular os retornos esperados e a matriz de covariância usando PyPortfolio
retornos_esperados = expected_returns.mean_historical_return(df_fechamento)
matriz_cov = risk_models.sample_cov(df_fechamento)
```

Fonte: Autores.

Além das métricas básicas, o projeto também incluiu métricas de risco avançado, essenciais para analisar a robustez da carteira em diferentes cenários de mercado. O Value at Risk (VaR) foi calculado para estimar a perda máxima esperada de um ativo ou portfólio dentro de um intervalo de confiança especificado. O VaR é uma medida importante para entender o risco de perdas extremas, sendo amplamente utilizado para o gerenciamento de risco financeiro (Jorion, 2007). Em complemento, o Conditional Value at Risk (CVaR) foi utilizado para medir a média das perdas que excedem o VaR, proporcionando uma avaliação mais detalhada dos riscos de

cauda e eventos extremos (Figura 6). O Drawdown Máximo, por sua vez, foi calculado para identificar a maior perda acumulada da carteira a partir de um pico até um vale, uma métrica importante para entender a magnitude de quedas em períodos de estresse no mercado (Figura 7).

Figura 6 – Cálculo do VaR e CVaR

```
# Definir diferentes níveis de confiança para análise de backtesting
niveis_confianca = [0.90, 0.95, 0.99]

# DataFrame para armazenar VaR e CVaR para cada nível de confiança ao longo do tempo
historico_risco = pd.DataFrame(index=portfolio_retorno_diario.index)

# Calcular VaR e CVaR para cada nível de confiança
for nivel in niveis_confianca:
    VaR_tempo = portfolio_retorno_diario.rolling(window=252).apply(lambda x: np.percentile(x, 100 * (1 - nivel)))
    CVaR_tempo = portfolio_retorno_diario.rolling(window=252).apply(lambda x: x[x <= np.percentile(x, 100 * (1 - nivel))].mean())
    historico_risco[f'VaR_{int(nivel*100)}'] = VaR_tempo
    historico_risco[f'CVaR_{int(nivel*100)}'] = CVaR_tempo
```

Fonte: Autores.

Figura 7 – Cálculo do Drawdown Máximo

```
# Calcular o Drawdown Máximo
retorno_acumulado = (1 + portfolio_retorno_diario).cumprod()
pico_acumulado = retorno_acumulado.cummax()
drawdown = (retorno_acumulado - pico_acumulado) / pico_acumulado
drawdown_maximo = drawdown.min()
print(f"Drawdown Máximo: {drawdown_maximo:.2%}")
```

Fonte: Autores.

Com essas métricas, é possível realizar uma análise mais completa dos ativos, ajustando as alocações de acordo com a tolerância ao risco e os objetivos de investimento do investidor.

4.4 DESENVOLVIMENTO DE SOFTWARE

O desenvolvimento do software foi guiado pelo objetivo de oferecer uma ferramenta flexível, acessível e ao mesmo tempo eficiente e robusta, que fosse capaz de atender tanto a investidores iniciantes quanto profissionais. A arquitetura do software foi estruturada visando facilitar testes, além de facilitar não apenas o desenvolvimento e a manutenção, mas também possibilitar a expansão futura do sistema.

Para garantir a escalabilidade e compatibilidade do software, utilizamos tecnologias amplamente adotadas no mercado. A linguagem Python foi escolhida devido à sua versatilidade e à ampla gama de bibliotecas para ciência de dados e finanças, como Pandas e NumPy, que são essenciais para manipulação e análise de dados (McKinney, 2010). Para os cálculos matemáticos

e as implementações de otimização, foram utilizadas as bibliotecas SciPy e PyPortfolioOpt, que oferecem ferramentas específicas para otimização de portfólios (Clarke et al., 2015). As bibliotecas de visualização, como Matplotlib e Seaborn, foram usadas para gerar gráficos estáticos detalhados, enquanto o Plotly foi utilizado para criar visualizações interativas, mais adaptáveis às necessidades do usuário.

Os resultados são apresentados em gráficos de forma objetiva, sempre destacando as métricas-chave como retorno esperado, volatilidade, VaR e CVaR, e as recomendações de alocação de ativos são feitas de forma clara.

A criação de visualizações informativas também exigiu atenção especial, pois os gráficos precisavam ser intuitivos e personalizáveis para atender aos diversos perfis de usuários. Para isso, incluímos rótulos dinâmicos e opções de filtros para focar em ativos ou períodos específicos, melhorando a interação e a compreensão dos resultados.

Além disso, validamos o software utilizando dados reais de diferentes períodos históricos e cenários de mercado, comparando os resultados obtidos com benchmarks estabelecidos, o que permitiu verificar a robustez e a confiabilidade das soluções propostas.

4.5 SIMULAÇÕES E BACKTESTING

As simulações e o backtesting foram componentes essenciais no desenvolvimento e validação das estratégias de otimização de portfólios. A simulação de Monte Carlo foi utilizada para projetar possíveis cenários futuros e avaliar a robustez das carteiras diante de diferentes condições de mercado. A abordagem de Monte Carlo envolveu a criação de um grande número de cenários (10.000) hipotéticos baseados em distribuições estatísticas dos retornos dos ativos. O principal objetivo dessa metodologia foi avaliar a probabilidade de retornos positivos e identificar os riscos extremos associados às carteiras, além de testar a resiliência das alocações contra variáveis desconhecidas, como mudanças abruptas nos preços dos ativos (Bertsimas et al., 2004).

Para realizar as simulações, utilizamos os retornos diários históricos de cada ativo para derivar parâmetros como média, volatilidade e covariância. A partir desses parâmetros, geramos milhares de caminhos simulados para os retornos diários usando distribuições normais multivariadas. Cada caminho simulou como os preços poderiam evoluir ao longo de um período definido, como por exemplo, um ano. As métricas de risco e retorno foram recalculadas para cada simulação, criando uma distribuição de resultados possíveis. Os resultados das simulações revelaram o intervalo provável de retornos anuais, levando em consideração diferentes níveis

de confiança, como 90%, 95% e 99%. Além disso, cenários extremos foram analisados para identificar potenciais perdas e impactos no portfólio, permitindo uma avaliação mais detalhada da robustez das carteiras (He e Liao, 2016).

Utilizamos o backtesting como ferramenta para avaliar o desempenho de nossas carteiras otimizadas em comparação com outras estratégias de investimento. Ao simular o comportamento das carteiras em diferentes cenários de mercado, incluindo períodos de crise e alta volatilidade, conseguimos mensurar a eficácia do nosso modelo de otimização.

Para evitar distorções e conclusões equivocadas, ao realizar as simulações, consideramos apenas ativos que apresentam consistência tanto na qualidade quanto na quantidade dos dados. Assim, durante o processamento inicial, ao filtrar apenas os ativos que possuem informações completas dentro do período selecionado, já excluimos uma quantidade significativa de ativos (Figura 7). Em seguida, ao aplicar um critério adicional, filtrando aqueles com 70% ou mais dos valores preenchidos, removemos mais alguns ativos (Figura 8).

Figura 8 – Tratamento inicial dos dados

```
for ativo in acoes_b3:
    try:
        dados = yf.download(ativo, start=start_date, end=end_date)
        if not dados.empty:
            dados_historicos[ativo] = dados['Close']
            sucesso += 1
            print(f"Sucesso ao processar {ativo}")
        else:
            erro += 1
            print(f"Erro: {ativo} não retornou dados.")
    except Exception as e:
        erro += 1
        print(f"Erro ao processar {ativo}: {e}")
```

Fonte: Autores.

Figura 9 – Remover valores ausentes dos dados processados

```
# Excluir ativos com mais de 30% de valores ausentes
percentual_ausente = df_fechamento.isnull().mean()
ativos_para_excluir = percentual_ausente[percentual_ausente > 0.3].index
df_fechamento.drop(columns=ativos_para_excluir, inplace=True)
```

Fonte: Autores.

4.5.1 CARTEIRA OTIMIZADA

Simulamos um cenário utilizando diferentes ativos da B3, incluindo 10 ações, 80 fundos imobiliários (FIIs), ETFs e debêntures. Primeiramente, uma lista com os ativos é definida, contendo o código de identificação de 100 ações, 99 fundos imobiliários, ETFs, e debêntures da B3. Baixamos os dados históricos de fechamento desses ativos durante o período de 1º de janeiro de 2019 até 1º de janeiro de 2024, utilizando a biblioteca `yfinance`.

Após realizar um tratamento inicial para preencher valores ausentes e excluir os ativos com mais de 20% de dados faltantes (utilizando a função `dropna(axis=1, thresh=int(0.8 * len(df_fechamento)))`), obtemos um total de 162 ativos que atenderam aos critérios de integridade e completude dos dados.

Uma vez que os dados estão prontos e processados, calculamos a correlação entre os preços de fechamento dos ativos, utilizando o método `corr()` do Pandas, que retorna uma matriz de correlação. Essa matriz mostra como os preços de dois ativos variam em conjunto: uma correlação de 1 significa que os ativos se movem exatamente da mesma forma, enquanto -1 indica que eles se movem em direções opostas.

Em seguida, duas funções são definidas para extrair informações úteis dessa matriz de correlação. A função `maiores_correlações` encontra as 10 maiores correlações, ou seja, os pares de ativos cujos preços se movem de forma mais sincronizada. Da mesma forma, a função `menores_correlações` encontra os 10 pares de ativos cujas correlações são mais baixas, ou seja, que têm a menor sincronização em seus movimentos de preços. Ambas as funções desconsideram a correlação de 1, que representa a correlação de um ativo consigo mesmo (Figura 10).

Os resultados obtidos corroboram as expectativas, especialmente no que diz respeito às maiores correlações. Observamos que os pares com maior grau de correlação incluem ativos com forte relação intrínseca, como fundos imobiliários relacionados ao mercado internacional, por exemplo, IVVB11 e SPXI11, ou ativos pertencentes a uma mesma holding ou grupo econômico, como GGBR4 e GOAU4 (do grupo Gerdau) e ELET6 e ELET3 (da Eletrobras).

Por outro lado, as menores correlações identificadas também fazem sentido dentro do contexto financeiro. Um exemplo claro é o par IMAB11, um produto de renda fixa atrelado à inflação, e TORD11, um fundo imobiliário focado em recebíveis. Além disso, outros pares de baixa correlação incluem combinações de categorias distintas, como ações sendo comparadas a fundos imobiliários, fundos de investimento ou debêntures. Essa diversidade reflete as diferenças nas dinâmicas de mercado e na natureza dos ativos.

Figura 10 – 10 maiores correlações dos ativos processados

As 10 maiores correlações entre os ativos da B3 são:			
	Ativo 1	Ativo 2	Correlação
0	IVVB11.SA	SPXI11.SA	0.999850
2	BOVV11.SA	BOVA11.SA	0.999812
4	BOVA11.SA	GOVE11.SA	0.991121
6	BOVV11.SA	GOVE11.SA	0.990670
8	GOAU4.SA	GGBR4.SA	0.987374
10	VALE3.SA	BRAP4.SA	0.969170
12	XPLG11.SA	KFOF11.SA	0.963770
14	GOLL4.SA	AZUL4.SA	0.961082
16	ELET6.SA	ELET3.SA	0.960113
18	ECOR3.SA	YDUQ3.SA	0.957239

As 10 menores correlações entre os ativos da B3 são:			
	Ativo 1	Ativo 2	Correlação
0	IMAB11.SA	TORD11.SA	-0.920099
2	SLCE3.SA	BBPO11.SA	-0.910401
4	TORD11.SA	VULC3.SA	-0.895466
6	TORD11.SA	PRI03.SA	-0.887784
8	FOFT11.SA	BRSR6.SA	-0.884178
10	IVVB11.SA	ABCP11.SA	-0.881537
12	SPXI11.SA	ABCP11.SA	-0.879790
14	BRSR6.SA	IVVB11.SA	-0.871294
16	BRSR6.SA	SPXI11.SA	-0.869843
18	BRCR11.SA	PRI03.SA	-0.867730

Fonte: Autores.

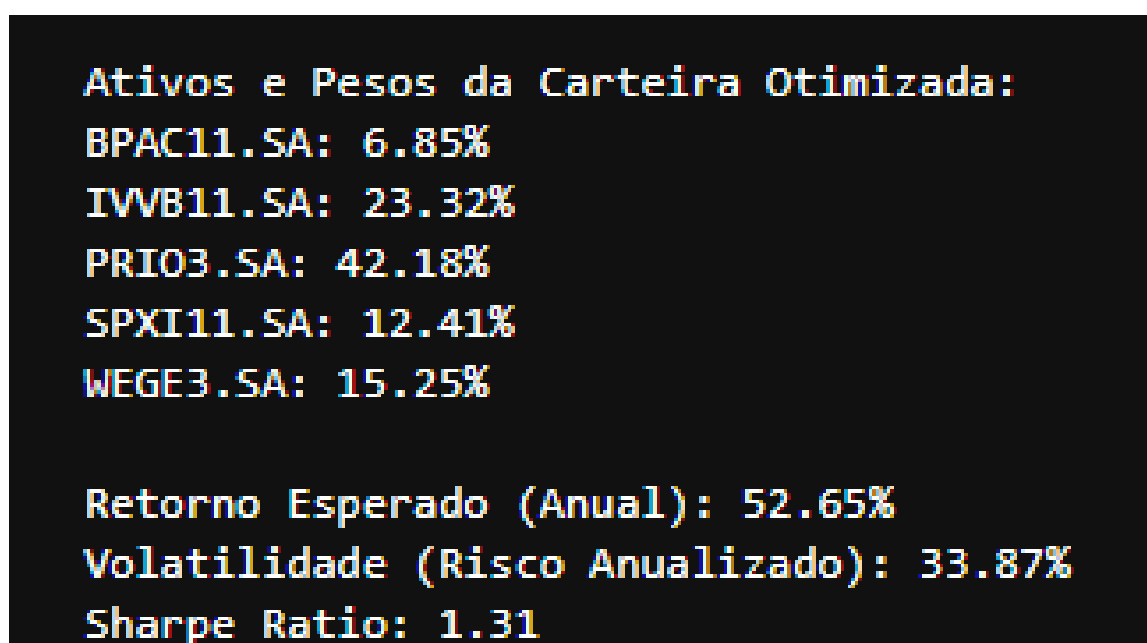
Os pesos dos ativos na carteira otimizada são obtidos com o método `max_sharpe`, que leva em consideração uma taxa de juros livre de risco de 8,3%. Em seguida, a função `clean_weights()` é chamada para ajustar os pesos dos ativos, removendo valores extremamente pequenos e apresentando uma carteira mais limpa.

Para construir uma carteira otimizada, utilizamos a função `mean_historical_return(df_fechamento)` para calcular os retornos médios históricos das ações, enquanto a função `sample_cov(df_fechamento)` calcula a matriz de covariância desses retornos. Com esses dados em mãos, o código utiliza o modelo de Carteira Eficiente da biblioteca `PyPortfolioOpt` para otimizar a alocação de ativos,

maximizando o Sharpe Ratio. Consideramos a taxa livre de risco de 8,32%, que corresponde à média da taxa Selic no período selecionado, conforme dados do Banco Central do Brasil (BCB). Com base nessa taxa, determinamos os pesos ideais com o método `max_sharpe` para alocar em cada ativo da carteira otimizada. Em seguida, usamos a função `clean_weights()` apenas para ajustar os pesos dos ativos, removendo valores extremamente pequenos e apresentando uma carteira mais limpa.

Além de apresentar a evolução da carteira (Figura 10), os ativos e suas respectivas alocações, calculamos a volatilidade anualizada, o retorno esperado anual e o índice de Sharpe, afim de avaliar a eficácia da otimização e a relação risco-retorno da carteira (Figura 11).

Figura 11 – Carteira otimizada



Fonte: Autores.

Para enriquecer ainda mais a análise, apresentamos uma visualização gráfica da performance da carteira otimizada durante o período selecionado (Figura 12) e do drawdown máximo (Figura 13), que representa a queda percentual no valor acumulado da carteira em relação ao seu maior valor anterior (pico). Durante o período da pandemia, observamos um drawdown máximo de -55,54%, o que é um resultado compreensível, considerando o impacto significativo que a B3 sofreu nesse período.

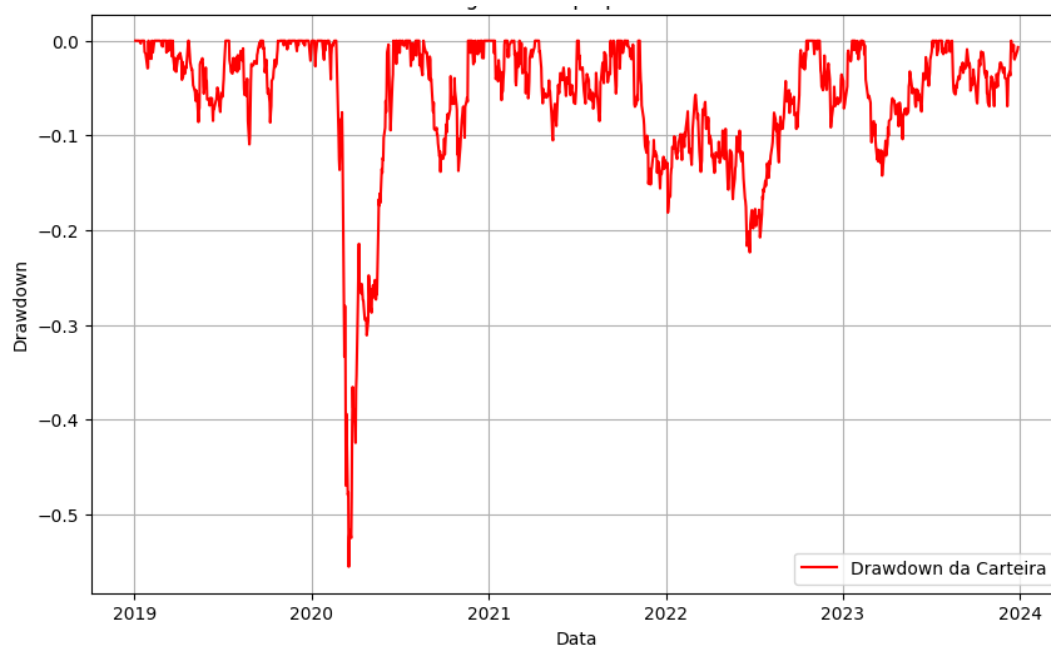
Para aprimorar a análise de risco do portfólio, utilizamos métodos como o cálculo de Value at Risk (VaR), Conditional Value at Risk (CVaR) e simulações de Monte Carlo. Inicialmente, o código calcula os retornos diários dos ativos e os ajusta conforme os pesos otimizados do portfólio. Com esses retornos, são calculados o VaR e o CVaR para diferentes níveis de confiança

Figura 12 – Retorno Acumulado da Carteira Otimizada e Ativos Individuais



Fonte: Autores.

Figura 13 – Drawdown ao longo do tempo para Carteira Otimizada



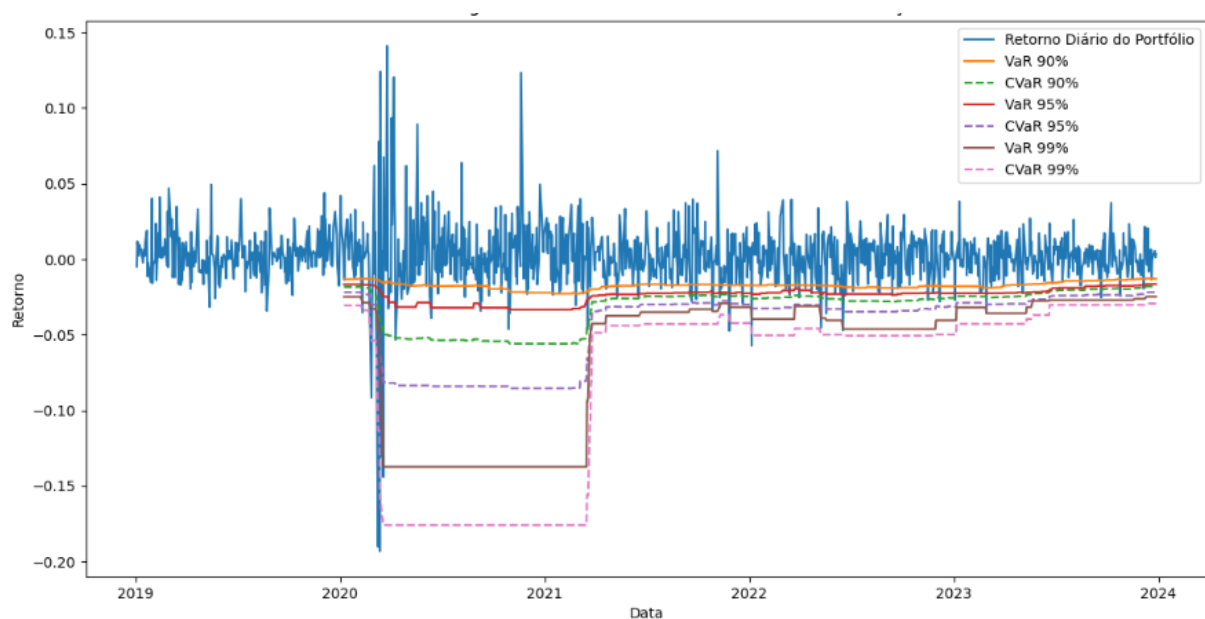
Fonte: Autores.

(90%, 95% e 99%), utilizando uma janela de 252 dias — o que corresponde a um ano de negociação, considerando 252 dias úteis em um mercado financeiro. A cada nova data no gráfico, esses cálculos são atualizados, sempre com base nos últimos 252 dias. Embora estamos lidando com dados de 5 anos, a análise para cada ponto é feita apenas com os dados dos últimos 252 dias.

A variação nos níveis de confiança tem impacto direto nos resultados do VaR e do CVaR.

Quanto maior o nível de confiança, menor será a perda esperada. Por exemplo, ao calcular o VaR de 99%, o valor resultante indica que há uma probabilidade de 1% de o portfólio sofrer uma perda superior a esse valor. Em contrapartida, para níveis de confiança mais baixos (como 90%), o VaR será maior, pois o risco de perdas mais severas é estimado com menor probabilidade. Em outras palavras, o VaR de 95% significa que, em 95% das vezes, o portfólio terá perdas dentro de um valor específico, enquanto o VaR de 99% reflete um cenário mais conservador, com uma perda maior, mas com uma probabilidade ainda mais baixa de ser atingida (Figura 14).

Figura 14 – VaR e CVaR em diferentes níveis de confiança

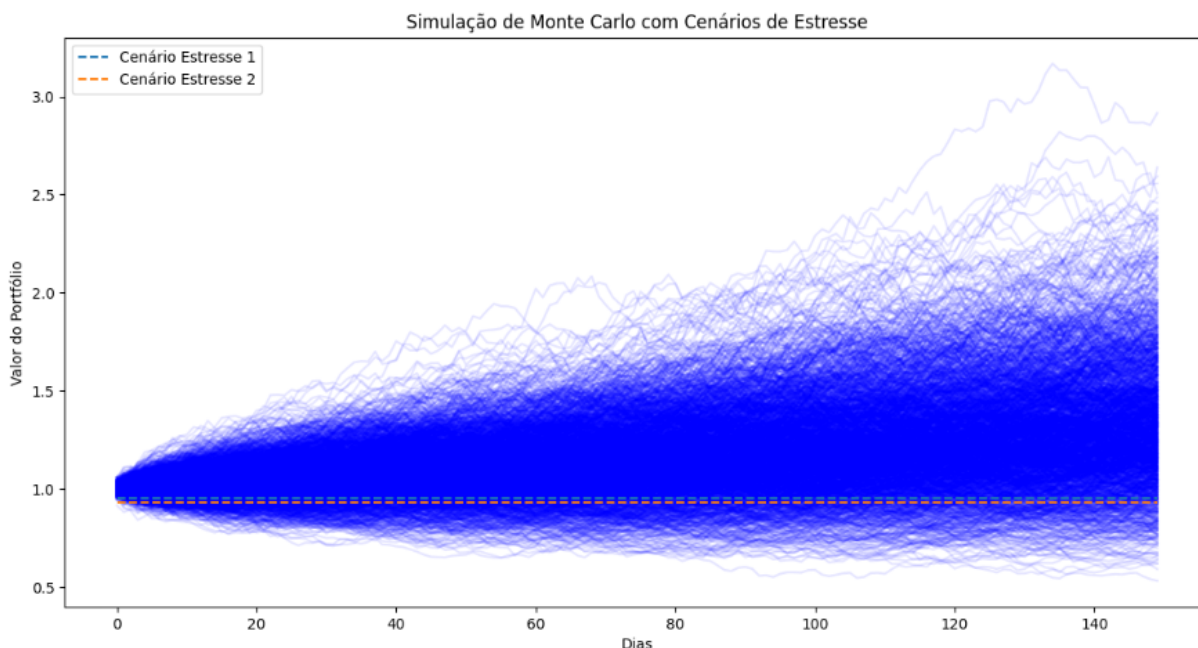


Fonte: Autores.

Além disso, em relação aos cenários de estresse, consideramos duas quedas hipotéticas no valor do portfólio: uma de 5% e outra de 7% (Figura 15). Esses valores são baseados em cenários de eventos extremos do mercado, como crises financeiras ou pandemias. A queda de 5% é considerada significativa, mas não extrema, enquanto a de 7% representa um impacto mais severo, semelhante a um choque de mercado de grande magnitude. Esses cenários ajudam a avaliar o comportamento do portfólio diante de eventos raros e adversos, proporcionando uma visão mais ampla do risco envolvido.

Analisando os resultados do VaR e CVaR, pode-se observar o comportamento que era previsto: durante períodos de alta volatilidade, como em março de 2020, o VaR e o CVaR se distanciam significativamente da linha de retorno diário, evidenciando um aumento dos riscos no portfólio. Após 2021, o gráfico mostra uma estabilização, com menor volatilidade nos retornos diários e uma redução nas estimativas de risco, tanto para VaR quanto para CVaR. A diferença

Figura 15 – Simulação de Monte Carlo com cenários de estresse



Fonte: Autores.

entre o VaR e o CVaR é especialmente relevante, pois reflete a magnitude das perdas extremas. Quando essa diferença é maior, como nos períodos de crise, significa que os eventos adversos tendem a ter um impacto mais significativo no portfólio.

Quanto a Simulação de Monte Carlo, Foram avaliadas duas condições de estresse no portfólio. Para o cenário de estresse de -5%, 4.418 das 10.000 trajetórias simuladas conseguiram superar o impacto, representando 44,18% do total, enquanto 5.582 não superaram, o que equivale a 55,82%. Já no cenário de estresse de -7%, 5.428 trajetórias superaram o impacto, o que corresponde a 54,28% das simulações, e 4.572 não conseguiram, resultando em 45,72% do total.

4.5.2 PERFORMANCE RISK PARTY

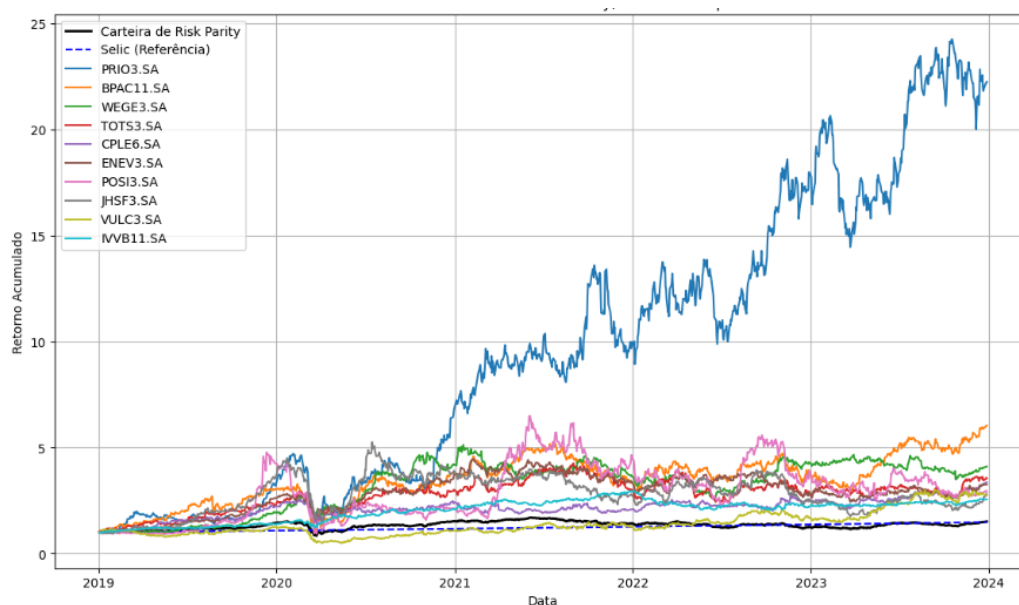
Para entender melhor os resultados obtidos até o momento, comparamos a estratégia adotada até agora — a maximização do Sharpe Ratio e o equilíbrio do trade-off entre risco e retorno — com outras abordagens de construção de portfólio.

A primeira estratégia utilizada como base de comparação é a Risk Parity. O conceito central dessa abordagem é que, em vez de dividir o capital igualmente entre os ativos, a alocação é feita com base na contribuição de risco de cada ativo. Isso significa que ativos mais voláteis recebem uma menor alocação de capital, enquanto ativos menos voláteis têm uma maior alocação, de forma a equilibrar o risco total da carteira.

Em uma carteira tradicional, a alocação é geralmente determinada pela expectativa de retorno, comumente seguindo uma distribuição fixa, como 60% em ações e 40% em títulos, por exemplo. Já na estratégia Risk Parity, a alocação é definida com base na volatilidade de cada ativo ou classe de ativos, visando distribuir o risco de maneira mais equilibrada.

Para tal, utilizamos o algoritmo SLSQP, que resolve um problema de otimização quadrática com restrições, como a soma dos pesos igual a 1 e limites entre 0 e 1 para cada peso. Depois de otimizar os pesos, calculamos o retorno diário do portfólio de acordo com os pesos encontrados e o retorno acumulado tanto do portfólio quanto de cada ativo individualmente, utilizando a fórmula de juros compostos. Após isso, gera-se uma exibição gráfica do desempenho acumulado ao longo do tempo realizando um benchmark com a Selic (Figura 16).

Figura 16 – Retorno Acumulado Final Carteira Risk Parity e Selic



Fonte: Autores.

Para calcular a Selic, anualizamos a taxa média do período utilizando a fórmula de juros compostos. A partir do retorno diário, obtemos o acumulado da Selic durante o mesmo intervalo de datas da carteira.

A principal característica do modelo Risk Parity é equilibrar a contribuição de risco de cada ativo, o que não implica necessariamente na maximização do retorno. Os pesos são ajustados para minimizar as diferenças de volatilidade e risco relativo, resultando frequentemente em alocações mais conservadoras para ativos mais voláteis. Assim, quando um ativo de alto retorno também apresenta alta volatilidade, o modelo tende a atribuir-lhe um peso menor, o que reduz a exposição a essas oportunidades de maior retorno. Embora esse equilíbrio reduza a

volatilidade, ele pode limitar o retorno da carteira em comparação aos ativos individuais com melhor desempenho. No caso da estratégia de Risk Parity, o retorno acumulado foi de 51,08%, enquanto a Selic obteve um retorno acumulado de 48,23%.

4.5.3 PERFORMANCE CARTEIRA DE DIVIDENDOS

Uma carteira de dividendos é composta por ações de empresas que distribuem regularmente uma parte de seus lucros na forma de dividendos aos acionistas. O principal objetivo dessa estratégia é gerar renda passiva para o investidor, por meio dos pagamentos periódicos de dividendos, além da possibilidade de valorização das ações ao longo do tempo.

Empresas consolidadas e com boa saúde financeira, especialmente aquelas em setores mais estáveis, como energia, telecomunicações e bancos, costumam ser as mais atrativas para essas carteiras. Isso porque elas têm um histórico consistente de pagamento de dividendos. A estratégia envolve selecionar ações que, além de oferecerem bons dividendos, também tenham potencial de valorização no longo prazo.

Essa abordagem é popular entre investidores que buscam fluxo de caixa estável, sem a necessidade de vender ativos, e que desejam se proteger contra a volatilidade do mercado. Além disso, pode ser uma maneira eficaz de garantir uma fonte de rendimento consistente, especialmente em tempos de incerteza econômica.

Para capturarmos a performance da estratégia de construção de carteira por dividendos, identificamos ações com altos valores de Dividend Yields (DY) no período especificado, calculamos o retorno acumulado de uma carteira composta por essas ações e comparamos com o desempenho individual dos ativos. O Dividend Yield médio anual para cada ação foi obtido dividindo os dividendos totais anuais pelo preço médio anual da ação.

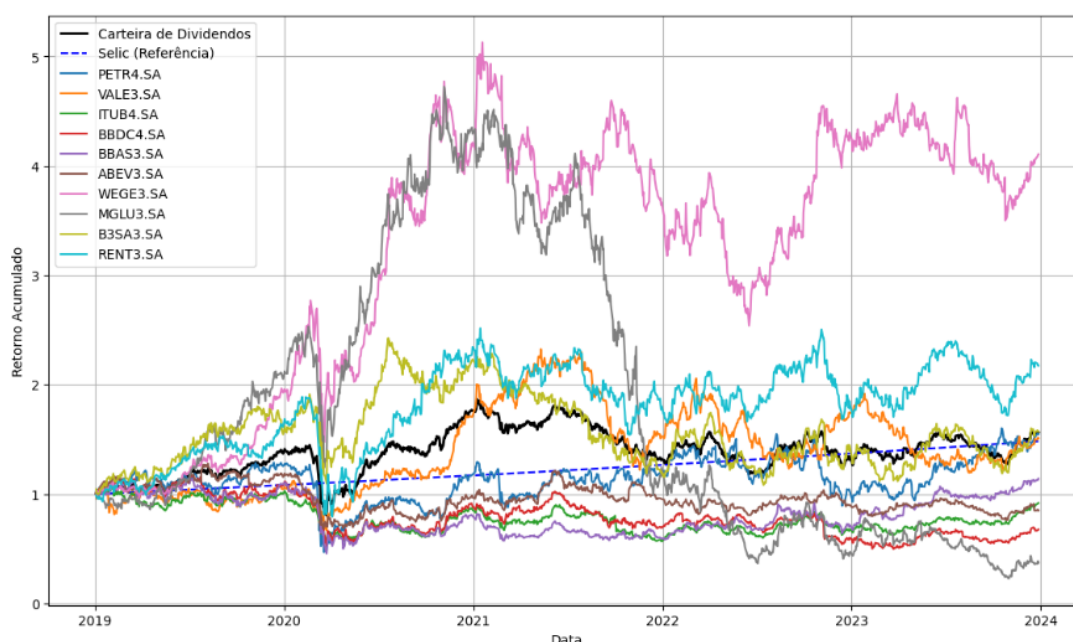
As 10 ações com os maiores Dividend Yields são selecionadas para formar uma carteira igualmente ponderada. Em seguida, são calculados os retornos diários dessa carteira com base nos preços das ações selecionadas. Esses retornos são acumulados para gerar a curva de retorno da carteira. Também calculamos os retornos acumulados individuais de cada uma das 10 ações selecionadas e a Selic para fins de comparação (Figura 17).

Foi observado um retorno acumulado de 56,38%, superando tanto a estratégia de Risk Parity quanto a Selic. No entanto, é fundamental destacar que basear uma carteira exclusivamente em dividendos pode ser arriscado, pois depende fortemente da sustentabilidade dos pagamentos, que pode ser afetada por mudanças nas condições econômicas ou dificuldades financeiras das

empresas. Além disso, essa estratégia pode resultar em falta de diversificação, já que muitas empresas que pagam dividendos elevados estão concentradas em setores específicos, o que aumenta a exposição ao risco de flutuações em um único segmento da economia.

Embora muitas empresas apresentem um histórico consistente de pagamento de dividendos, esses pagamentos não são garantidos. Em períodos de recessão ou instabilidade financeira, as empresas podem decidir cortar ou até suspender os dividendos para preservar caixa, o que impactaria diretamente a renda passiva do investidor. Esse risco torna a estratégia vulnerável a mudanças econômicas inesperadas, exigindo que o investidor esteja atento à sustentabilidade dos dividendos ao longo do tempo.

Figura 17 – Retorno Acumulado Final Carteira Dividendos e Selic



Fonte: Autores.

4.5.4 PERFORMANCE CARTEIRA DE MOMENTUM

A estratégia de momentum é uma abordagem de investimento que se baseia na premissa de que ativos que apresentam bom desempenho no passado recente tendem a continuar com resultados positivos no futuro, enquanto aqueles com desempenho negativo provavelmente seguirão em queda. O foco dessa estratégia é identificar e investir em ativos com tendências de alta, enquanto se evita ou se vende aqueles com tendências de baixa. O cálculo do momentum geralmente leva em consideração os retornos passados dos ativos, que são avaliados em períodos que podem variar de semanas a meses, com o objetivo de capturar os ganhos de ativos que estão

em forte tendência de valorização.

A carteira de momentum é composta por ativos com os melhores retornos recentes, e os investidores ajustam suas posições periodicamente, comprando os ativos em ascensão e vendendo aqueles que perderam força. Embora essa estratégia possa gerar bons retornos durante tendências fortes, ela também envolve riscos, principalmente em relação à reversão de tendência. Se um ativo que estava em alta começar a cair, o investidor pode sofrer perdas rápidas, especialmente se a mudança de tendência ocorrer de forma abrupta.

Na B3, a alta volatilidade e os ciclos econômicos do Brasil podem amplificar os riscos dessa estratégia, especialmente em períodos de incerteza política ou econômica. Portanto, embora a estratégia de momentum possa ser rentável ao aproveitar tendências de curto prazo, ela exige maior atenção e agilidade para lidar com mudanças repentinas nas condições do mercado, um fator particularmente relevante na bolsa brasileira.

Para aplicar a estratégia, calculamos os retornos acumulados dos últimos 6 meses de cada ativo antes de cada data de rebalanceamento. Em seguida, selecionamos os 10 ativos com o melhor desempenho em termos de momentum durante esse período. O rebalanceamento ocorre mensalmente, com base na performance dos últimos 6 meses, garantindo que a carteira esteja sempre alocada nos ativos mais fortes, conforme a estratégia. A cada rebalanceamento, os pesos são redistribuídos igualmente entre os ativos selecionados.

Figura 18 – Retorno Acumulado Final Carteira Momentum e Selic



Fonte: Autores.

A visualização gráfica exhibe o desempenho acumulado da carteira, omitindo os ativos

selecionados em cada rebalanceamento para evitar sobrecarga visual. Além disso, a linha da Selic foi incluída novamente como benchmark, oferecendo uma referência adicional para comparação (Figura 18).

4.5.5 PREVISÃO DE RESULTADOS

A carteira otimizada apresenta bons resultados, mas oferece informações limitadas sobre o desempenho futuro dos ativos. Para suprir essa lacuna, utilizamos um modelo de aprendizado de máquina, especificamente o Random Forest, com o objetivo de prever o comportamento futuro desses ativos. Com base nessas previsões, podemos determinar quando será necessário realizar um rebalanceamento periódico da carteira.

Dessa forma, definimos uma função que utiliza o modelo de Random Forest para prever os retornos futuros das ações na carteira. Essa função começa calculando os retornos logarítmicos diários para a ação selecionada. Em seguida, ela cria variáveis explicativas simples baseadas nos retornos passados, como os retornos de 5, 50 e 500 dias atrás. A partir dessas variáveis, o modelo divide os dados históricos de retornos acumulados em duas partes: uma para treinamento e outra para teste. Ele usa os dados de treinamento para aprender a relação entre variáveis explicativas (features) e a variável de interesse (retorno acumulado no período escolhido, como uma semana ou um mês). Essa divisão é feita pela função `train_test_split`, configurada para alocar 80% dos dados para o treinamento e 20% para o teste.

Durante o treinamento, o modelo Random Forest é ajustado para minimizar a diferença entre os valores previstos e os valores reais da variável de retorno, otimizando seu desempenho. Após o treinamento, o modelo é aplicado aos dados de teste, que são compostos de variáveis explicativas derivadas do mesmo período de retornos históricos, mas que o modelo ainda não viu. Isso permite avaliar sua capacidade de generalização, ou seja, quão bem ele prevê dados fora do conjunto usado para ajustar seus parâmetros.

Os retornos reais do conjunto de teste são comparados às previsões do modelo utilizando o erro quadrático médio (MSE), que serve como métrica para avaliar a precisão das estimativas. Gráficos são gerados para cada ativo e horizonte de tempo analisado (como 20 e 120 dias, correspondentes a aproximadamente 1 e 3 meses, respectivamente), permitindo a visualização dos valores reais e previstos ao longo do período.

Nos gráficos, o eixo x representa as datas, enquanto o eixo y exibe os valores de retorno acumulado. Esse processo proporciona uma análise detalhada do desempenho do modelo sob

diferentes condições de mercado, destacando padrões que ele captura bem e eventuais desvios que podem ser aprimorados.

Figura 19 – Previsão de resultados para a próxima semana da Carteira Otimizada

```
Ativos e Pesos da Carteira Otimizada:
BPAC11.SA: 6.85%
IVVB11.SA: 23.32%
PRIO3.SA: 42.18%
SPXI11.SA: 12.41%
WEGE3.SA: 15.25%
BPAC11.SA - Previsão de Retorno para a próxima semana: 0.24%, MSE: 0.000398
IVVB11.SA - Previsão de Retorno para a próxima semana: 0.99%, MSE: 0.000070
PRIO3.SA - Previsão de Retorno para a próxima semana: -0.59%, MSE: 0.000564
SPXI11.SA - Previsão de Retorno para a próxima semana: 0.25%, MSE: 0.000075
WEGE3.SA - Previsão de Retorno para a próxima semana: -0.11%, MSE: 0.000371
```

Fonte: Autores.

Para interpretar o resultado do MSE, é importante relembrar o seu conceito. O Erro Quadrático Médio (MSE) é uma métrica usada para avaliar a precisão de modelos de previsão, especialmente em problemas de regressão. Ele calcula a média dos quadrados das diferenças entre os valores reais e as previsões feitas pelo modelo. Ele basicamente mede o quanto as previsões do modelo estão distantes dos valores reais, com um valor menor indicando um modelo mais preciso.

O MSE é sempre um valor positivo, e quanto menor for o seu valor, melhor será o desempenho do modelo, pois isso indica que a diferença entre os retornos reais e os previstos é menor. Por exemplo, um MSE elevado (como 0,5) sugeriria que o modelo está tendo dificuldades em prever os retornos de forma eficaz, enquanto um MSE baixo (como 0,001) indicaria que as previsões estão bem alinhadas com os dados reais. O MSE é especialmente útil para comparar o desempenho de diferentes modelos ou configurações, ajudando a identificar aquele que apresenta os melhores resultados na tarefa de previsão.

Assim, pode-se concluir que as previsões geradas pelo modelo foram bastante positivas, oferecendo uma boa confiança na análise. Isso nos dá uma base sólida para determinar a necessidade de rebalanceamento da carteira, proporcionando maior segurança nas decisões sobre ajustes nos ativos.

4.5.6 REBALANCEAMENTO DA CARTEIRA

A função `rebalancear_carteira` foi criada para permitir que o usuário substitua os ativos com previsões negativas de retorno, ao mesmo tempo em que garante que os pesos da carteira sejam recalculados com base nos novos ativos e suas características. O processo começa com a identificação dos ativos com previsões negativas, o que permite ao usuário substituí-los por outros de sua escolha, retirados de uma lista de ativos disponíveis. O código verifica se o ativo escolhido está na lista de ativos disponíveis, garantindo que a substituição seja válida. Em seguida, o retorno esperado para o novo ativo é calculado utilizando o modelo de Random Forest, e a carteira passa a ter seus pesos recalculados. Para isso, uma nova matriz de covariância e um novo vetor de retornos esperados são criados com base nos ativos restantes e nas novas previsões. A carteira é, então, otimizada novamente usando o método `max_sharpe` da biblioteca `PyPortfolioOpt`, redistribuindo os pesos de forma a maximizar o retorno ajustado ao risco dos novos ativos.

Além disso, foi criada a função `plot_comparacao`, que gera um gráfico de barras comparando as previsões de retorno antes e depois do rebalanceamento. Esse gráfico facilita a visualização das mudanças nos retornos esperados dos ativos da carteira após a substituição. Para cada ativo, o gráfico exibe duas barras: uma representando a previsão de retorno antes do rebalanceamento e outra, após a substituição dos ativos negativos. As barras antes do rebalanceamento são coloridas de azul, e as barras após a substituição são coloridas de laranja, tornando a comparação visual clara e intuitiva.

Essa estratégia de substituição de ativos permite um controle manual completo sobre a composição da carteira, em que o usuário pode decidir quais ativos substituir com base nas previsões de retorno e, depois de realizar a substituição, recalcular os pesos da carteira para garantir que ela continue otimizada. Após a substituição, o código recalcula automaticamente os pesos da carteira, levando em consideração a nova composição de ativos.

O software oferece flexibilidade total ao usuário na escolha dos ativos a serem substituídos, permitindo que ele ajuste a carteira conforme suas preferências e tolerância ao risco. Após a substituição, os pesos da carteira são recalculados de forma proporcional, garantindo que a alocação de ativos seja mantida de forma otimizada com base nos retornos esperados e na matriz de covariância dos novos ativos. O gráfico gerado fornece um feedback imediato sobre o impacto do rebalanceamento na carteira, ajudando o usuário a visualizar como as substituições afetam o desempenho esperado da carteira ao longo do tempo.

Ao rodar o código, o usuário verá a previsão de retorno de cada ativo na carteira. Para aqueles com previsões negativas, ele será solicitado a escolher um novo ativo da lista disponível. Após a substituição, os pesos otimizados serão recalculados e a nova composição da carteira será exibida. O gráfico comparativo também será gerado, permitindo uma visualização clara do impacto do rebalanceamento. Essa abordagem é especialmente útil para quem deseja realizar rebalanceamentos manuais de carteira, aproveitando previsões de retorno e otimizando a alocação de ativos com base em dados históricos e de mercado.

Ademais, o usuário pode testar diferentes substituições de ativos e ajustar o parâmetro `ativos_disp` para incluir uma lista personalizada de ativos a serem considerados na otimização. Isso pode ser feito de maneira dinâmica, conforme a disponibilidade dos ativos e as preferências de investimento

5 RESULTADOS

Este capítulo apresentará os resultados obtidos pela integração das previsões realizada pelo Random Forest com o rebalanceamento de carteira.

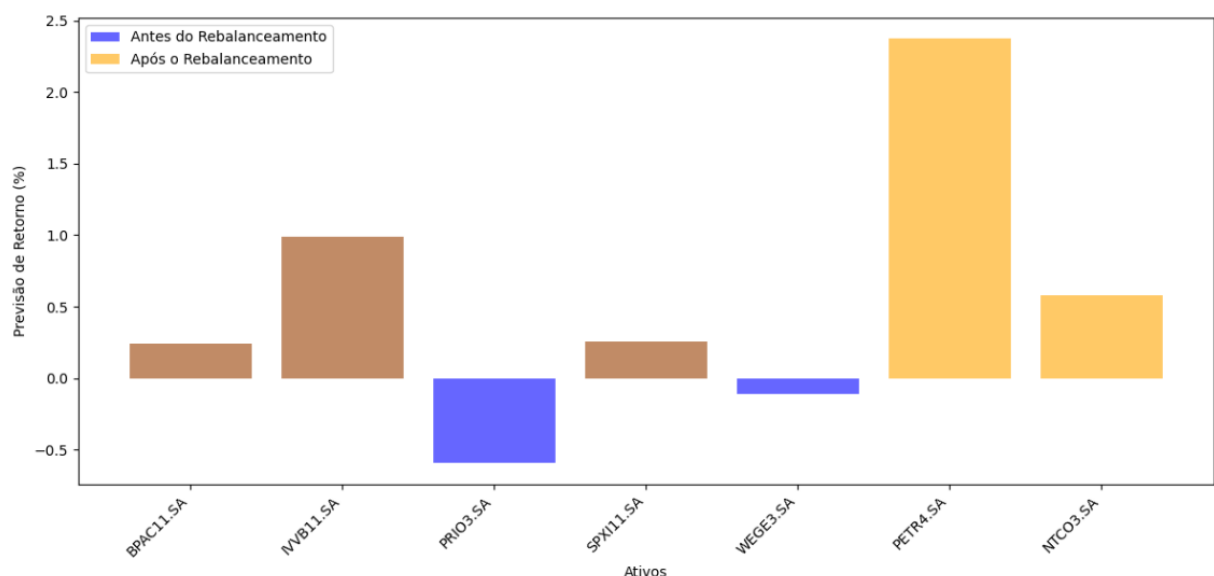
5.1 Integrando as soluções

Neste estágio, a carteira otimizada já foi definida, e o modelo de previsão de retorno para a próxima semana, baseado em Random Forest, está implementado. A novidade é que, após calcular as previsões de retorno para os ativos da carteira otimizada, o software permite ao usuário realizar um rebalanceamento manual sempre que algum ativo apresentar uma previsão de retorno negativa. Nesses casos, o usuário pode substituir o ativo com previsão negativa por outro da lista de opções disponíveis (ativos_disp).

Após o rebalanceamento, o software exibe as novas previsões de retorno para cada ativo na carteira otimizada, levando em conta as substituições realizadas. Esse processo é interativo, permitindo ao usuário ajustar a carteira conforme necessário, com o objetivo de maximizar o retorno ajustado ao risco. A atualização das previsões e dos pesos dos ativos é apresentada visualmente, facilitando o acompanhamento das mudanças ao longo do processo.

Além disso, um gráfico é gerado para comparar o desempenho da carteira antes e após o rebalanceamento, proporcionando uma visão clara do impacto das substituições feitas.

Figura 20 – Comparação rebalanceamento da carteira otimizada



Fonte: Autores.

Figura 21 – Processo de rebalanceamento da Carteira Otimizada

```

Ativos e Pesos da Carteira Otimizada:
BPAC11.SA: 6.85%
IVVB11.SA: 23.32%
PRI03.SA: 42.18%
SPXI11.SA: 12.41%
WEGE3.SA: 15.25%
BPAC11.SA - Previsão de Retorno para a próxima semana: 0.24%, MSE: 0.000398
IVVB11.SA - Previsão de Retorno para a próxima semana: 0.99%, MSE: 0.000070
PRI03.SA - Previsão de Retorno para a próxima semana: -0.59%, MSE: 0.000564
SPXI11.SA - Previsão de Retorno para a próxima semana: 0.25%, MSE: 0.000075
WEGE3.SA - Previsão de Retorno para a próxima semana: -0.11%, MSE: 0.000371

Carteira Atual:
Ativos com Previsão de Retorno Negativo:
PRI03.SA - Previsão de Retorno: -0.59%
WEGE3.SA - Previsão de Retorno: -0.11%

Você pode substituir o ativo PRI03.SA por outro ativo.
Escolha um ativo para substituir PRI03.SA: PETR4.SA
Substituindo PRI03.SA por PETR4.SA.

Você pode substituir o ativo WEGE3.SA por outro ativo.
Escolha um ativo para substituir WEGE3.SA: NTCO3.SA
Substituindo WEGE3.SA por NTCO3.SA.

Novo status da carteira com substituições:
BPAC11.SA - Previsão de Retorno para a próxima semana: 0.24%
IVVB11.SA - Previsão de Retorno para a próxima semana: 0.99%
SPXI11.SA - Previsão de Retorno para a próxima semana: 0.25%
PETR4.SA - Previsão de Retorno para a próxima semana: 2.37%
NTCO3.SA - Previsão de Retorno para a próxima semana: 0.58%

```

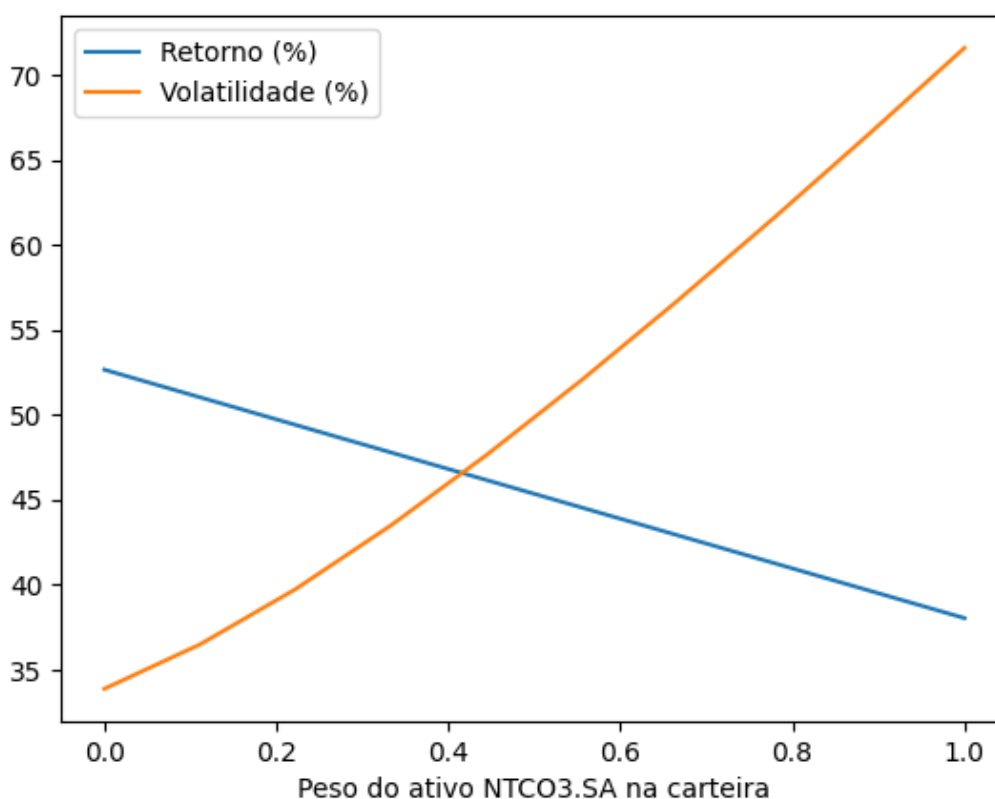
Fonte: Autores.

Em seguida, realizamos uma análise de sensibilidade para investigar como diferentes valores para o peso de um ativo específico, no caso "NTCO3.SA"(Natura), podem afetar o retorno e a volatilidade da carteira otimizada. Para isso, o código gera uma série de 10 valores igualmente distribuídos entre 0 e 1, representando o peso do ativo "NTCO3.SA"na carteira, variando de 0% a 100%. Em seguida, uma lista é criada para armazenar os resultados dessa análise.

Para cada valor de peso, o código faz uma cópia dos pesos ótimos da carteira original e altera o peso do ativo "NTCO3.SA"conforme o valor específico da iteração. Após isso, uma instância do modelo EfficientFrontier é criada, que otimiza a carteira com base nos retornos esperados e na matriz de covariância dos ativos. A nova configuração de pesos é aplicada à carteira, e o método portfolio_performance() calcula o retorno e a volatilidade da carteira com os pesos ajustados. Esses valores são multiplicados por 100 para transformá-los em porcentagens e

armazenados na lista de resultados, junto com o peso do ativo "NTCO3.SA" correspondente.

Figura 22 – Análise de Sensibilidade NTCO3: Peso vs. Retorno/Volatilidade



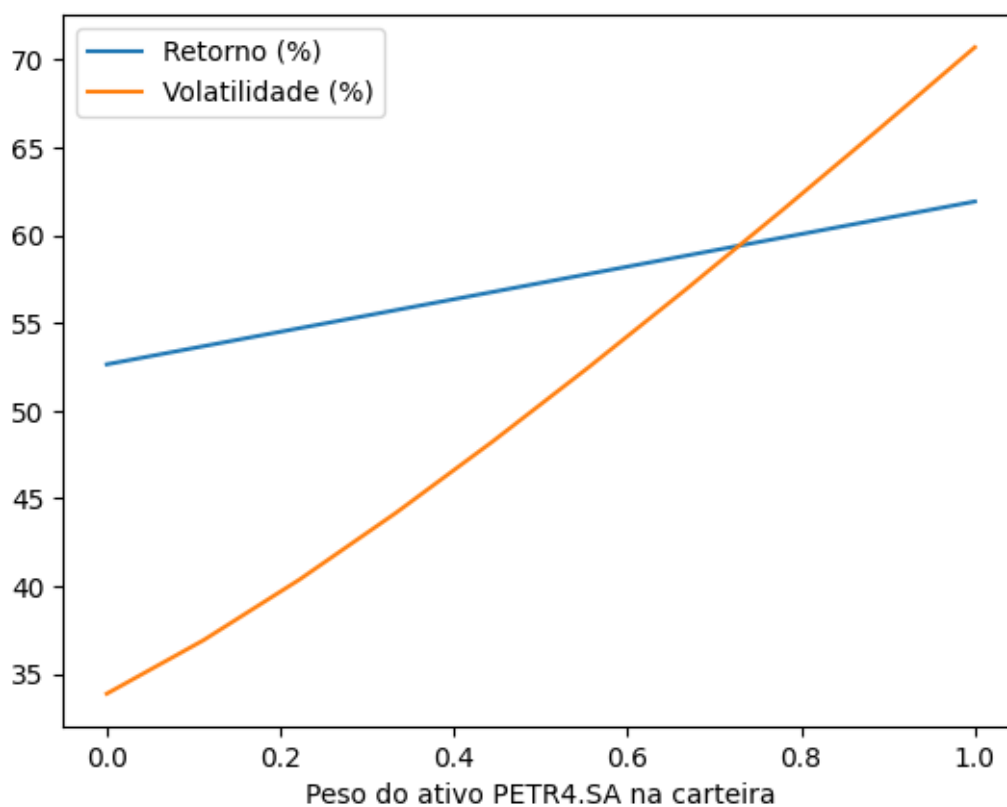
Fonte: Autores.

Após a execução do loop, os resultados são extraídos e a biblioteca matplotlib é utilizada para gerar um gráfico de linha (Figura 22). O gráfico mostra a variação do retorno e da volatilidade em função do peso do ativo "NTCO3.SA" na carteira. Cada linha do gráfico representa uma das métricas (retorno ou volatilidade), permitindo ao usuário visualizar como essas variáveis de desempenho da carteira mudam à medida que o peso do ativo é ajustado. O gráfico é então exibido, oferecendo uma representação visual da análise de sensibilidade entre o peso do ativo e os indicadores financeiros da carteira.

Realizamos a mesma análise para o ativo "PETR4.SA" (Petrobras), que também foi cotado para integrar a carteira otimizada, e observamos um retorno melhor, embora com maior volatilidade (Figura 23).

Mais resultados podem ser alcançados ajustando as variáveis de entrada, como a lista de ativos iniciais, os ativos disponíveis para troca, o período de tempo de referência, a taxa livre de risco (Selic) ou parâmetros técnicos do código. Entre esses ajustes estão a alteração do tamanho do dataset de testes, o número de simulações e árvores, as variáveis explicativas, ou mesmo a remoção da fixação do parâmetro `random_state`. Além disso, todos os re-

Figura 23 – Análise de Sensibilidade PETR4: Peso vs. Retorno/Volatilidade



Fonte: Autores.

sultados obtidos ao longo deste projeto, bem como testes adicionais, estão disponíveis em: https://github.com/bieltanaka11/sistema_analise_b3. Acesso em: 25 nov. 2024

6 CONSIDERAÇÃO FINAIS

As considerações finais deste estudo sintetizam os principais achados sobre as estratégias de investimento analisadas, incluindo dividendos, momentum, risk parity e a carteira otimizada. A pesquisa alcançou seus objetivos ao comparar essas abordagens e demonstrar como a otimização de portfólio, que integra risco e retorno, pode proporcionar vantagens substanciais em relação a estratégias mais específicas ou tradicionais, como as baseadas em dividendos ou momentum. A principal contribuição prática deste estudo é a evidência de que uma abordagem otimizada, que leva em consideração diversas métricas e simula cenários futuros, oferece uma solução mais robusta e adaptável às condições dinâmicas do mercado.

Além disso, este trabalho abriu caminho para novas possibilidades, como a personalização de carteiras com base no perfil de risco do investidor e a integração de tecnologias avançadas, como machine learning e análise de dados em tempo real. A implementação dessas inovações pode aprimorar ainda mais as estratégias de investimento, permitindo ajustes dinâmicos e mais eficientes às condições voláteis do mercado financeiro.

Porém, algumas limitações foram identificadas. A complexidade dos modelos utilizados exige um elevado poder computacional e conhecimento técnico para sua implementação eficaz. Além disso, a análise de risco e retorno pode ser impactada por fatores externos difíceis de prever, o que pode limitar a precisão das simulações e das previsões.

Apesar dessas limitações, as possibilidades abertas pelo projeto são vastas. Destacam-se a utilização de inteligência artificial para otimizar ainda mais as carteiras e a incorporação de novos indicadores de risco, considerando uma gama mais ampla de variáveis econômicas e financeiras. A construção de uma ferramenta que permita a análise dinâmica e em tempo real de carteiras otimizadas poderia ser expandida para integrar diferentes classes de ativos, como criptomoedas, o que ampliaria ainda mais as opções de diversificação. Tais lacunas representam oportunidades significativas para o avanço contínuo no campo das finanças quantitativas, estimulando o desenvolvimento de metodologias mais sofisticadas e aplicáveis.

A análise do risco de cauda e a incorporação de dados alternativos, como indicadores macroeconômicos e sentiment analysis, também são áreas promissoras para aprimorar a robustez das estratégias de otimização. O futuro do projeto pode incluir a exploração de métodos de aprendizado de máquina e inteligência artificial para refinar as previsões de retorno e risco,

proporcionando recomendações ainda mais precisas aos investidores. A pesquisa também pode se expandir para incluir a análise de mercados internacionais e a consideração de variáveis econômicas globais, com o objetivo de construir uma abordagem verdadeiramente global, adaptável às rápidas mudanças do mercado financeiro.

Em suma, este estudo contribui para a melhoria das práticas de alocação de ativos no mercado financeiro, não apenas por meio de uma análise crítica das diferentes estratégias, mas também ao sugerir caminhos para o desenvolvimento de ferramentas mais sofisticadas e adaptáveis às necessidades dos investidores modernos. A construção de carteiras otimizadas, especialmente, é um campo fértil para o avanço da teoria financeira e suas aplicações práticas, abrindo novas perspectivas para uma gestão de portfólio mais eficiente e personalizada.

7 CONCLUSÃO

O objetivo principal deste projeto foi desenvolver um software que auxilie investidores a enfrentar os desafios relacionados à criação, análise e otimização de investimentos no mercado financeiro. Para isso, foram incorporadas diversas métricas financeiras que permitem uma análise aprofundada e precisa. O trabalho integrou a fundamentação acadêmica com uma abordagem prática, combinando metodologias clássicas e modernas, destacando a relevância da Ciência de Dados e da programação na resolução de problemas reais no contexto financeiro.

Ao longo do desenvolvimento, buscamos abordar de maneira inovadora e prática o desafio da construção e análise de carteiras de investimento otimizadas. Para isso, utilizamos tanto abordagens tradicionais quanto técnicas mais recentes para avaliação de risco e retorno. Partindo de uma base teórica sólida, exploramos técnicas consolidadas de otimização de portfólio, como a alocação média-variância e métricas clássicas, como o Índice de Sharpe. Simultaneamente, ampliamos a análise ao integrar métricas avançadas, como Value at Risk (VaR), Conditional Value at Risk (CVaR) e drawdown máximo, oferecendo uma visão mais completa do risco financeiro.

Durante o desenvolvimento, enfatizamos o uso de dados históricos robustos e ferramentas computacionais avançadas, incluindo bibliotecas específicas para finanças e otimização de portfólios. Simulações de Monte Carlo e backtesting foram essenciais para validar as estratégias, demonstrando sua resiliência em cenários adversos e sua consistência em relação a benchmarks de mercado, como a taxa Selic.

Adicionalmente, a análise de performance dos ativos individuais dentro da carteira otimizada permitiu identificar o impacto específico de cada ativo no risco total e no retorno esperado. A integração de visualizações gráficas e métricas quantitativas ajudou a transformar dados complexos em insights claros e acionáveis, facilitando tomadas de decisão.

Ao incorporar elementos de inovação, como o uso de metodologias modernas de otimização e análise de cenários extremos, o trabalho não apenas ampliou os horizontes da análise de portfólios, mas também destacou a importância de construir modelos personalizáveis, adaptáveis aos objetivos e restrições dos investidores. O diferencial do projeto reside na integração de diversas abordagens, sendo capaz de aplicar técnicas de otimização e modelagem matemática com o suporte de linguagens de programação e bibliotecas específicas, demonstrando a capacidade

de transformar problemas abstratos em soluções práticas e aplicáveis. Esse aspecto, aliado à personalização e à adaptabilidade, coloca o software como uma ferramenta robusta para atender às necessidades de diferentes perfis de investidores.

Em conclusão, este trabalho representa uma contribuição significativa para a análise de investimentos e gestão de portfólios. Ele não apenas reafirma a importância da diversificação, do uso de métricas avançadas e da validação empírica como pilares para a tomada de decisões financeiras bem informadas, como demonstra a importância da interdisciplinaridade entre ciência da computação e finanças. Com uma abordagem centrada na inovação e no rigor técnico, o projeto fornece uma base sólida para o avanço das aplicações práticas na área de finanças quantitativas e reforça a relevância da computação na resolução de desafios do mundo real.

REFERÊNCIAS

- CARVALHO, Gabriel Tanaka de; BENVENUTI, Enzo H. S.; GUERRA, Leonardo Modesto. Sistema de análise de ativos listados na B3. GitHub, 2024. Disponível em: <https://github.com/bieltanaka11/sistema_analise_b3>. Acesso em: 25 nov. 2024.
- MARKOWITZ, Harry. Portfolio Selection. *Journal of Finance*, v. 7, n. 1, 1952.
- SHARPE, William F. The Sharpe Ratio. *Journal of Portfolio Management*, v. 21, n. 1, 1966.
- JORION, P. *Value at Risk: The New Benchmark for Managing Financial Risk*. 3. ed. New York: McGraw-Hill, 2007.
- FAMA, Eugene F.; FRENCH, Kenneth R. The Cross-Section of Expected Stock Returns. *The Journal of Finance*, v. 47, n. 2, 1992.
- CLARKE, Joel. PyPortfolioOpt: Implementing Modern Portfolio Theory Using Python. *Journal of Open Source Software*, v. 4, n. 38, 2019.
- IMAMURA, Yuri; KOSAPONG, Benyanee. Portfolio optimization with conditional Value-at-Risk under CEV model. 2023.
- ARAÚJO, Alcides Carlos de; MONTINI, Alessandra de Ávila. Técnicas de Big Data e projeção de risco de mercado utilizando dados em alta frequência. *Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP)*, Brasil.
- KENETT, Dror Y. et al. Partial correlation analysis: Applications for financial markets. *Center for Polymer Studies and Department of Physics, Boston University*, 2014.
- ALBUQUERQUE, Marcelo Masera de. Análise de risco de carteira de ativos usando decomposição wavelets e correlação multivariada de cópulas: determinação do Value-at-Risk e de proporção de falhas. 2019.
- DALCIN, Mariliane Machado. Aplicação da teoria de Markowitz para alocação de ativos financeiros. *Universidade Federal do Pampa (UNIPAMPA)*, 2021.
- OLI, Marxia S. et al. Big data analytics - Application of artificial neural network in forecasting stock price trends in India. *Bharathidasan University*, 2018.

CHIANG, Thomas C.; JEON, Bang Nam; LI, Huimin. Dynamic correlation analysis of financial contagion: Evidence from Asian markets. 2007.

BANCO CENTRAL DO BRASIL. Histórico de taxas de juros. Disponível em: <<https://www.bcb.gov.br/controleinflacao/historicotaxasjuros>>. Acesso em: 25 nov. 2024.

BODIE, Z.; KANE, A.; MARCUS, A. J. *Investments*. 10. ed. New York: McGraw-Hill Education, 2014.

ROCKAFELLAR, R. T.; URYASEV, S. Optimization of Conditional Value-at-Risk. *Journal of Risk*, v. 2, n. 3, 2000.

MCKINNEY, Wes. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. 2. ed. Sebastopol: O'Reilly Media, 2017.