

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

Flávio Gabriel Duarte

Um Framework para Alocação de Portfólios Baseada em
Agrupamento Particional

São Paulo
2019

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

Flávio Gabriel Duarte

Um Framework para Alocação de Portfólios Baseada em
Agrupamento Particional

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica e Computação.

Orientador: Prof. Dr. Leandro Nunes de Castro

São Paulo
2019

D812f

Duarte, Flávio Gabriel

Um framework para alocação de portfólios baseada em agrupamento particional. / Flávio Gabriel Duarte- São Paulo, 2019.

74 f. : il., 30 cm.

Dissertação (Mestrado em Engenharia Elétrica e Computação) - Universidade Presbiteriana Mackenzie - São Paulo, 2019.

Orientador: Prof. Dr. Leandro Nunes de Castro

Bibliografia: f. 51-54

1.Alocação de ativos 2. Ações 3. Seleção de portfólios 4. Aprendizado de máquina 5. Finanças quantitativas 6. Framework
I. Silva, Leandro Nunes de Castro, *orientador*. II.Título.

CDD 005.3

Bibliotecária Responsável: Maria Gabriela Brandi Teixeira – CRB 8/ 6339

MODELO DE FOLHA DE APROVAÇÃO

FLÁVIO GABRIEL DUARTE

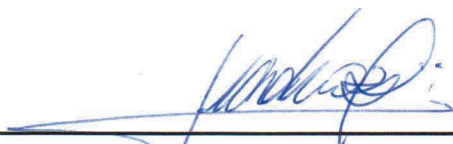
Um Framework para Alocação de Portfólios Baseada em Agrupamento
Particional

Dissertação de Mestrado apresentada
ao Programa de Pós-Graduação em
Engenharia Elétrica e Computação da
Universidade Presbiteriana Mackenzie,
como parte dos requisitos para a
obtenção do título de Mestre em
Engenharia Elétrica e Computação.

Orientador: Prof. Dr. Leandro Nunes de
Castro

Aprovado em

BANCA EXAMINADORA



Prof. Dr. Leandro Nunes de Castro
Universidade Presbiteriana Mackenzie



Prof. Dr. Ismar Frango
Universidade Presbiteriana Mackenzie



Prof. Dr. Fabrício Olivetti de França
Universidade Federal do ABC

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida e pela sabedoria.

Agradeço à Helena, minha esposa, por estar ao meu lado e ter sido compreensível no período do mestrado.

Aos meus pais, Hélio e Ivonete, pela criação e educação que hoje me trazem até aqui.

A minha irmã, Morgana, pela amizade de vida.

Ao meu orientador, Professor Dr. Leandro Nunes de Castro, por aceitar orientar um trabalho com aplicação em um tema diferente de pesquisa do laboratório e pela orientação valiosa ao longo do trabalho.

Aos amigos da Rio Bravo Investimentos, Paulo Bilyk, Rafael Rodrigues, Evandro Buccini e Gustavo Rodrigues.

Aos colegas do Laboratório de Computação Natural (LCoN) pelas contribuições ao longo do curso.

À FAPESP e ao MackPesquisa pelo apoio financeiro aos estudos.

RESUMO

Essa dissertação propõe um *framework* baseado em agrupamento particional para a alocação de recursos em um conjunto de ativos a partir da correlação entre eles. O objetivo do *framework* é fornecer uma técnica de alocação para ajudar investidores a melhorarem seu processo de investimento, sugerindo a alocação usando as informações sobre os grupos gerados a partir de um algoritmo de agrupamento. Este trabalho é diferente das abordagens já propostas na literatura, que utilizam essencialmente algoritmos de agrupamento hierárquico, pois ele está embasado em algoritmos particionais com diferentes propostas para alocação intragrupo e intergrupo. Para determinar o percentual de alocação dos ativos em cada grupo (intragrupo) foram propostas duas técnicas: o método da roleta e o método nebuloso. Para a alocação intergrupo foram utilizados um método baseado na roleta e um método equilibrado. Experimentos foram realizados utilizando dados da Bolsa de Valores Brasileira e os ativos elegíveis para entrarem na alocação foram os que faziam parte do índice Ibovespa no momento do rebalanceamento do portfólio. Os resultados foram comparados com outros métodos de alocação e com o próprio índice Ibovespa. O *framework* proposto ilustra o potencial das técnicas de aprendizagem de máquina na otimização do portfólio de ações.

Palavras-chave: *alocação de ativos, ações, seleção de portfólios, aprendizado de máquina, finanças quantitativas, framework.*

ABSTRACT

This dissertation proposes a framework based on partitional clustering for the asset allocation in a set of assets from the correlation between them. The goal of the framework is to provide an asset allocation method to help investors improve their investment process, suggesting the allocation using information about the groups found with the clustering algorithm. This work is different from the approaches already proposed in the literature, which essentially use hierarchical clustering algorithms, since it is based on partitional algorithms with different proposals for intragroup and intergroup allocation. To determine the percentage of asset allocation in each group (intragroup), two techniques were proposed: a roulette wheel and a fuzzy method. For the intergroup allocation, a roulette wheel and a balanced method were used. The method was tested using data from the Brazilian Stock Exchange and the assets eligible to enter the allocation were those that were part of the Ibovespa index at the time of portfolio rebalancing. The results were compared with other allocation methods and with the Ibovespa index. The proposed framework illustrates the potential of machine learning techniques in portfolio optimization.

Keywords: *asset allocation, stocks, portfolio selection, machine learning, quantitative finance, framework.*

Lista de Figuras

1	Gráfico da Fronteira Eficiente. Fonte: (Fama; French, 2007).	8
2	Desempenho das ações ITUB4 e RADL3 em 2017.	10
3	Desempenho do portfólio das ações ITUB4 e RADL3 em 2017.	10
4	Framework proposto para alocação de ativos utilizando agrupamento particional.	17
5	Evolução mensal dos percentuais máximos de cada portfólio.	44

Lista de Tabelas

1	Comparação das estratégias utilizando janela de correlação de 12 meses. . .	29
2	Comparação das estratégias utilizando janela de correlação de 24 meses. . .	30
3	Comparação das estratégias utilizando janela de correlação de 12 meses . .	32
4	Comparação das estratégias utilizando janela de correlação de 24 meses. . .	33
5	Comparação das estratégias utilizando janela de correlação de 12 meses . .	35
6	Comparação das estratégias utilizando janela de correlação de 24 meses . .	36
7	Comparação das estratégias utilizando janela de correlação de 12 meses . .	37
8	Comparação das estratégias utilizando janela de correlação de 24 meses . .	38
9	Comparação das estratégias utilizando janela de correlação de 12 meses . .	40
10	Comparação das estratégias utilizando janela de correlação de 24 meses . .	41
11	Melhores resultados para cada configuração utilizando janela de correlação de 12 meses	42
12	Melhores resultados para cada configuração utilizando janela de correlação de 24 meses	42
13	Carteira dos portfólios	45
13	Carteira dos portfólios	46

Sumário

1	Introdução	1
1.1	Motivação e Justificativa	3
1.2	Objetivos	3
1.3	Organização do Documento	4
2	Referencial Teórico	5
2.1	Conceitos Básicos sobre Finanças e Portfólios	5
2.2	Teoria de Alocação de Portfólios	6
2.2.1	Benefício da Combinação de Ativos	9
2.3	Agrupamento	11
2.3.1	Algoritmo k -Médias	11
2.3.2	Algoritmo k -Medoides	12
2.3.3	Agrupamento Nebuloso	13
2.4	Índice Ibovespa	14
3	Um Framework para Alocação de Portfólios Baseado em Agrupamento Particional	16
3.1	Fluxo do Processo de Alocação	17
3.2	Pré-Processamento: Cálculo da Matriz de Correlação e Distância	18
3.3	Agrupamento	18
3.4	Alocação Intragrupo	20
3.4.1	Método Nebuloso	20
3.4.2	Método da Roleta	20
3.5	Alocação Intergrupo	21
3.5.1	Alocação Equilibrada de Capital	21
3.5.2	Alocação Baseada na Roleta	22
3.6	Avaliação de Desempenho	23
3.7	Trabalhos Correlatos	24
4	Avaliação de Desempenho	26
4.1	Metodologia Experimental	26
4.2	Resultados e Discussão	28
4.2.1	Configuração 1: Seleção aleatória dos medoides iniciais, alocação nebulosa intragrupo e alocação equilibrada intergrupo	28

4.2.2	Configuração 2: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação equilibrada intergrupo	31
4.2.3	Configuração 3: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação intergrupo pelo método da roleta	34
4.2.4	Configuração 4: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação equilibrada intergrupo	37
4.2.5	Configuração 5: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação intergrupo também pelo método da roleta	39
4.2.6	Comparação das Carteiras dos Portfólios	42
5	Conclusão e Trabalhos Futuros	47
	Referências Bibliográficas	54
	Apêndices	55
A	Carteira teórica do índice Ibovespa	55

1 Introdução

Com os avanços da computação e a evolução dos computadores, juntamente com a popularização dos processos baseados em dados, tornou-se comum o uso de algoritmos na gestão de recursos por parte de gestores profissionais. Segundo a revista *Institutional Investor* (2018), cinco dos seis maiores fundos de investimentos do mundo utilizam computadores na maior parte ou em todo o processo decisório. O rápido desenvolvimento e absorção das técnicas de **Inteligência Artificial** (IA) tem criado mecanismos cada vez mais eficientes de **seleção** ou **otimização de portfólios**, o que corresponde ao processo de selecionar um conjunto de **ativos** (por exemplo, **ações**) a partir dos ativos disponíveis, de forma a otimizar (maximizar ou minimizar) um ou mais objetivos (Markowitz, 1952), (Ponsich; Jaimes; Coello, 2013). Os objetivos geralmente incluem o **retorno esperado**, que deve ser maximizado, e o **risco**, que deve ser minimizado.

O trabalho de Markowitz (1952) é considerado pioneiro na introdução de um modelo média-variância para a otimização de portfólios de ações, partindo do princípio de que a diversificação de ativos permite aumentar o retorno do portfólio e reduzir o risco. Ao longo do tempo, várias técnicas foram sendo introduzidas para aprimorar o processo de alocação de portfólios, como os modelos propostos por Black e Litterman (1992) e Michaud (1989), inspirados na proposta de otimização de Markowitz (1952).

A solução do problema de otimização de portfólios requer conhecimentos de Finanças, Matemática e Computação. Os **algoritmos de otimização de portfólio** normalmente partem da premissa de que é necessário incorporar os retornos futuros e riscos esperados no processo de otimização. Uma estimativa errada de retorno e risco futuros leva a um portfólio não ótimo como resultado da otimização (Michaud, 1989).

A existência de mais de um objetivo a ser otimizado simultaneamente, assim como o desconhecimento a priori da superfície de decisão do problema, torna a aplicação de técnicas de IA bastante conveniente para essa tarefa. Dentre essas, os **algoritmos evolutivos multiobjetivos** têm recebido muita atenção da comunidade como ferramentas eficientes para a otimização de portfólios (Ponsich; Jaimes; Coello, 2013).

Os métodos acima possuem deficiências, pois dependem dos valores futuros de retorno e risco, o que pode comprometer o resultado da otimização se as estimativas estiverem

erradas. Também existe uma restrição algébrica em que o aumento do tamanho da matriz de correlação implica no aumento do número de condicionamento, causando erros numéricos que tornam a matriz instável (López de Prado, 2016). Dados os problemas acima citados, tornaram-se populares as técnicas de alocação de portfólios conhecidas como *risk parity*, nas quais as alocações são realizadas com base no risco previsto sem a necessidade de incorporar os retornos futuros esperados (Clarke; De Silva; Thorley, 2013). Entretanto, López de Prado (2016) nota que, nesses métodos, continua sendo necessária a inversão da matriz positiva definida de covariância que pode levar a erros numéricos e instabilidade.

López de Prado (2016) propõe uma técnica chamada "Hierarchical Risk Parity" para alocação de portfólios, que utiliza um algoritmo de **agrupamento hierárquico** para realizar a alocação. Baseado na técnica proposta de alocação utilizando agrupamentos, propomos um *framework* baseado em um algoritmo de agrupamento particional, conhecido como *k-medoides*, para a solução do problema de alocação de portfólios.

Enquanto um agrupamento hierárquico realiza uma decomposição hierárquica da base de dados, um agrupamento particional realiza uma divisão em partições da base de dados. O *framework* proposto permite flexibilidade na construção do portfólio, pois tanto as informações intragrupo quanto as informações intergrupo, permitem a aplicação de diferentes técnicas para a alocação.

Sendo assim, o objetivo desse trabalho é testar a alocação de portfólios utilizando um algoritmo do tipo *k-medoides*, testando diferentes valores para *k* grupos e diferentes métodos de alocação intragrupos e intergrupos. Os experimentos serão realizados considerando os períodos de 30/12/2009 até 28/12/2017 e o portfólio será sempre rebalanceado mensalmente. Os ativos utilizados serão ações negociadas no mercado brasileiro de ações (B3), e os ativos elegíveis para compor os agrupamentos são os mesmos que compuserem a carteira teórica do índice Ibovespa na respectiva data de rebalanceamento. Todos os ativos que passarem pelo agrupamento serão alocados e o portfólio estará sempre comprado (*long-only*).

1.1 Motivação e Justificativa

A motivação deste trabalho é estudar o problema de alocação de portfólios, normalmente tratado como um problema de otimização, e também investigar métodos de alocação de portfólios que utilizam conceitos de aprendizado de máquina. Ao final, será proposta uma técnica de seleção e alocação de portfólios utilizando aprendizado de máquina e que tem como núcleo o algoritmo particional k -medoides.

Atualmente, os algoritmos de aprendizado de máquina têm sido utilizados cada vez mais em processos de muitas áreas da economia, como gestão empresarial, logística, entre outros. Em âmbito mundial, cinco dos seis maiores fundos de investimentos utilizam técnicas computacionais e modelos quantitativos na maior parte ou em todo o processo decisório Institutional Investor (2018). Porém, no Brasil menos de 1% dos recursos de investidores são alocados utilizando algoritmos e inteligência artificial. Justifica-se, portanto, o estudo de algoritmos de aprendizado de máquina para alocação de portfólio, sendo um assunto que possui cada vez mais relevância no mundo e no Brasil.

1.2 Objetivos

O objetivo geral da pesquisa é propor uma técnica de alocação de portfólio baseada em agrupamento particional de dados. Especificamente, o projeto visa:

- Investigar a influência da inicialização do algoritmo k -medoides aplicado ao problema de alocação de portfólio;
- Propor um método de inicialização para o k -medoides que considera a volatilidade das ações;
- Investigar diferentes técnicas de alocação intra- e intergrupo, incluindo métodos inspirados em algoritmos evolutivos e agrupamento nebuloso;
- Avaliar o desempenho da técnica proposta utilizando dados da B3 no período de 30/12/2009 a 28/12/2017 e comparar com outros algoritmos da literatura.

1.3 Organização do Documento

Essa dissertação está organizada em cinco capítulos de forma a apresentar os conceitos necessários à compreensão de aprendizado de máquina, alocação de portfólios, a técnica proposta e avaliar seu desempenho. Após a introdução, o trabalho está organizado do seguinte modo.

O Capítulo 2 apresenta o referencial teórico da pesquisa, incluindo conceitos básicos sobre finanças e portfólios, a teoria de alocação de portfólios, algoritmos particionais de agrupamento de dados e fala sobre o índice Ibovespa.

O Capítulo 3 introduz o *framework* proposto, incluindo seu fluxo de processo e principais fases. São apresentadas as etapas de pré-processamento dos dados, a aplicação do algoritmo de agrupamento seguida da alocação intra e intergrupo, as medidas de avaliação de desempenho e uma breve revisão da literatura sobre trabalhos correlatos.

O desempenho do *framework* proposto é avaliado no Capítulo 4, no qual são investigadas cinco configurações distintas, cada uma com uma configuração diferente do *framework*. Uma discussão mais profunda sobre o *framework* é apresentada ao final do capítulo.

A pesquisa é concluída no Capítulo 5 com uma discussão geral e perspectivas de trabalhos futuros.

2 Referencial Teórico

Essa dissertação envolve conceitos sobre finanças, portfólios de ações, teoria de alocação de portfólios e mineração de dados, mais especificamente agrupamento de dados. Este capítulo faz uma revisão geral sobre os conceitos necessários ao bom entendimento da pesquisa e do método proposto para alocação de portfólio de ações.

2.1 Conceitos Básicos sobre Finanças e Portfólios

Um **ativo** pode ser definido como todos os bens e direitos de um indivíduo (IFRS Foundation, 2010). Porém, ao tratar de investimentos financeiros, objeto deste trabalho, de acordo com as normas internacionais de contabilidade (IFRS Foundation, 2009), **ativos financeiros** podem ser dinheiro, ações e um direito contratual de receber algum ativo financeiro no futuro.

Um comportamento observado das **ações**, no caso o ativo financeiro objeto de pesquisa desse trabalho, é a variação de preços, geralmente em período diário. A partir da formação da série temporal de preços, é possível analisar propriedades dessa série (Tsay, 2005).

Ao conduzir estudos em finanças, é utilizado o **retorno do ativo**, ao invés do preço, para análise. Estatisticamente, as propriedades do retorno são mais tratáveis do que as propriedades do preço e também o retorno é a medida livre de escala e completa do ativo (Tsay, 2005). Para obter o **retorno líquido** simples em um período de tempo t é utilizada a Equação 1:

$$r(t) = \frac{p(t) - p(t-1)}{p(t-1)}, \quad (1)$$

sendo $r(t)$ o retorno do ativo no tempo t , $p(t)$ o preço do ativo no tempo t e $p(t-1)$ o preço do ativo no tempo $t-1$.

Outra propriedade observada nas séries temporais dos retornos é a **volatilidade**, também chamada de **risco** neste trabalho. Segundo Magalhães de Pinho, de Camargos e Marques Figueiredo (2017), a volatilidade pode ser determinada a partir de três formas de cálculo:

- **Estocástico:** por meio de modelos econométricos a partir de dados históricos;
- **Implícito:** calculado a partir do preço de mercado com o preço modelado pelo modelo de Black e Scholes (1973);
- **Histórico:** obtido a partir de uma média de uma função de n retornos de uma série ou pelo desvio padrão da mesma.

Quando um investidor decide investir em uma ação, seu patrimônio fica exposto a oscilação de preço da ação. Porém, quando o investidor decide investir em mais ações, ele fica exposto a oscilação das ações combinadas e essa combinação de ativos passa a ser chamada de **portfólio**. Markowitz (1952) foi o pioneiro nos estudos de portfólios. A próxima seção apresenta os conceitos sobre portfólios e como são calculados o retorno e o risco deles.

2.2 Teoria de Alocação de Portfólios

A seleção de portfólios é uma atividade que vem sendo amplamente pesquisada, principalmente entre profissionais de investimentos e acadêmicos (Kolm; Tütüncü; Fabozzi, 2014). É um tema de pesquisa que possui intersecção das áreas de Finanças, Matemática e Computação (Mansini; Ogryczak; Speranza, 2014; Ponsich; Jaimes; Coello, 2013).

O primeiro trabalho envolvendo a seleção de portfólios foi publicado por Markowitz (1952), que propôs um modelo matemático para a seleção de portfólios de ativos contendo a ideia de diversificação em investimentos. Através da diversificação é possível diminuir o risco sem diminuir o retorno esperado. Markowitz propôs que um investidor deve buscar maximizar o retorno do portfólio enquanto a variância, que é considerada o risco, deve ser minimizada (Rubinstein, 2002). As ideias propostas por Markowitz se tornaram o que veio a ser chamado de **Teoria Moderna de Portfólios** e o modelo de Markowitz também é conhecido como modelo de **média-variância** (Fabozzi; Gupta; Markowitz, 2002).

Fabozzi, Gupta e Markowitz (2002) comentam que um dos aspectos principais do trabalho de Markowitz (1952) não é avaliar o risco do ativo isoladamente, porém avaliar a contribuição que esse ativo causa no risco do portfólio.

O trabalho sobre média-variância de Markowitz (1952) considera um conjunto de N

ativos, onde X_i é o **percentual** do ativo i no portfólio. O modelo de **média-variância** apresenta duas restrições:

$$X_i \geq 0, \quad (2)$$

$$\sum_{i=1}^N X_i = 1. \quad (3)$$

A Equação 2 estabelece que o portfólio só possui ativos com percentual não-negativo na carteira e a Equação 3 define que a soma de todos os ativos do portfólio deve ser 1.

O **retorno esperado** (E) do portfólio é calculado a partir da soma dos retornos esperados dos ativos ponderados pelo percentual do ativo no portfólio:

$$E = \sum_{i=1}^N X_i \mu_i, \quad (4)$$

sendo X_i o percentual do ativo no portfólio, μ_i o retorno do ativo i , e E o retorno esperado do portfólio.

O **risco esperado** (V) do portfólio é dado a partir da variância dos ativos e da correlação entre eles:

$$V = \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij}, \quad (5)$$

sendo X_i e X_j os percentuais dos ativos i e j no portfólio, σ_{ij} a covariância entre X_i e X_j , e V a variância do portfólio.

A principal tese defendida por Markowitz é a diversificação em portfólios e os seus benefícios (Markowitz, 1952). Com o modelo proposto acima, o investidor tem um conjunto de combinações de E e V , onde cada portfólio possui valores esperados diferentes, de acordo com o percentual de cada ativo em cada portfólio.

Dado um conjunto de combinações de portfólios (E, V) , o investidor deve escolher o portfólio que possui a melhor relação entre E e V , que é o maior E para o menor V possível.

Fabozzi, Gupta e Markowitz (2002) destacam que a Teoria Moderna de Portfólios é

uma teoria normativa, descrevendo padrões e normas de comportamentos que os investidores devem perseguir ao alocar portfólios. De acordo com Fabozzi, Gupta e Markowitz (2002), a **teoria de precificação de ativos** (Sharpe, 1964) formaliza o relacionamento entre risco e retorno que deveria existir entre os ativos. Fabozzi, Gupta e Markowitz (2002) destacam que o modelo de **média-variância** e o modelo de **precificação de ativos** formam um *framework* robusto de seleção de portfólios. Os autores também destacam que o modelo de **média-variância** é um modelo independente de qualquer teoria de precificação de ativos.

A Figura 1, apresentada por Fama e French (2007), ilustra as oportunidades de portfólios. O eixo X mostra o risco medido pelo desvio padrão dos retornos e o eixo Y mostra o retorno esperado dos portfólios. Na Figura 1 é possível observar o que é chamado na literatura de fronteira de variância mínima, onde são representadas combinações de retorno esperado e risco dos portfólios e o nível de retorno esperado diminui quando o risco dos portfólios é minimizado.

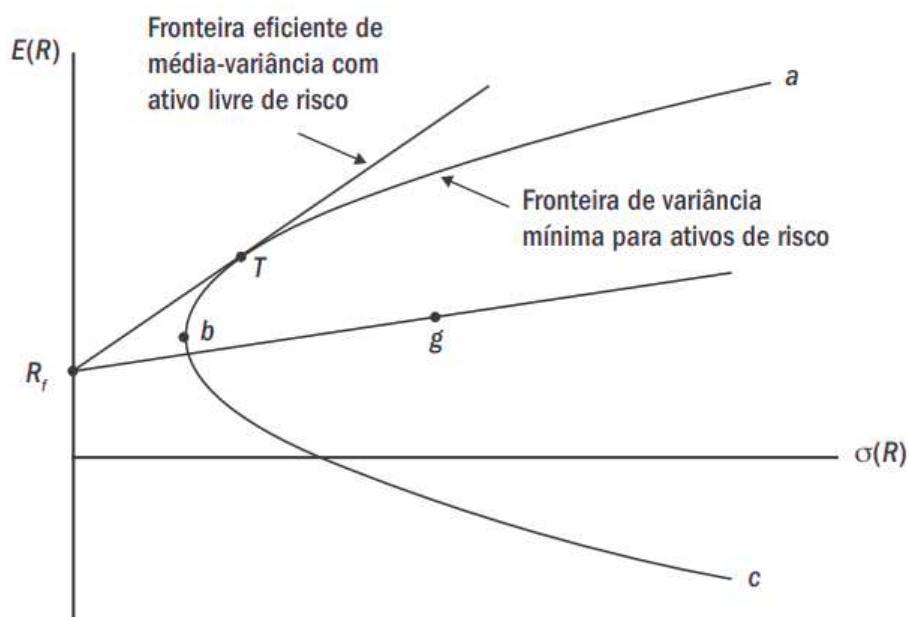


Figura 1: Gráfico da Fronteira Eficiente. Fonte: (Fama; French, 2007).

Fama e French (2007) comentam sobre a dinâmica de seleção de portfólios na fronteira de variância e é evidente no gráfico a relação entre risco e retorno. O investidor que deseja um nível de retorno elevado, poderá investir no portfólio *a*, porém esse portfólio

apresentará mais volatilidade. O ponto c é o ponto contrário do a , sendo um portfólio com alta volatilidade e com retorno esperado negativo. O ponto T representa o portfólio eficiente em média-variância, ou seja, a melhor relação entre risco e retorno com retorno intermediário com menor volatilidade. Ainda segundo Fama e French (2007), apenas os portfólios ao longo da curva abc acima do ponto b são eficientes em média-variância, pois maximizam o retorno esperado, dada a variância dos retornos.

O portfólio eficiente é encontrado traçando-se uma reta em diagonal a partir do ponto R_f , que representa o risco e o retorno do ativo livre de risco. O portfólio da tangente da reta é o portfólio eficiente (Fama; French, 2007). Esse portfólio deve ser perseguido pelos investidores que buscam a melhor relação entre risco e retorno.

López de Prado (2016) comenta que Markowitz (1952) compreendeu que para vários níveis de risco, existem portfólios ótimos em termos de retornos ajustados ao risco, por isso o conceito de fronteira eficiente.

Para avaliar o desempenho de um portfólio e conseguir encontrar alguma razão que avalie o desempenho, Sharpe (1966) propôs o uso de um indicador, que foi chamado de *reward-to-variability ratio*, e ficou conhecido como *Sharpe Ratio*. O indicador de desempenho do portfólio é calculado conforme a Equação 6:

$$SR = \frac{\mu - r_f}{\sigma}, \quad (6)$$

onde μ é o retorno da série no período, r_f é o retorno do ativo livre de risco e σ é o desvio padrão da série analisada.

2.2.1 Benefício da Combinação de Ativos

Para exemplificar a dinâmica de formação de portfólio e o resultado da volatilidade do portfólio a partir da combinação de ações, foi construído um portfólio com as ações ITUB4 e RADL3, onde inicialmente ambas ações possuíam o mesmo percentual no portfólio (50%) e com período de avaliação no ano de 2017. Na Figura 2 é possível observar o desempenho individual das ações no período e na Figura 3 é possível observar o desempenho das ações juntamente com o portfólio resultante da alocação entre elas.



Figura 2: Desempenho das ações ITUB4 e RADL3 em 2017.



Figura 3: Desempenho do portfólio das ações ITUB4 e RADL3 em 2017.

Para o período avaliado, a ação ITUB4 teve um desempenho de 30,6% com uma

volatilidade de 24,8% ao ano, já a ação RADL3 apresentou um desempenho de 51,3% e volatilidade de 23,8%. O portfólio resultante apresentou um desempenho de 42,2% no período com uma volatilidade de 20,3%, sendo possível observar a queda da volatilidade causada pela combinação dos ativos.

2.3 Agrupamento

Em Mineração de Dados, os algoritmos de agrupamento podem ser divididos em: particionais, hierárquicos, baseados em densidade e baseados em grafos (de Castro; Ferrari, 2016).

Os algoritmos hierárquicos mais conhecidos são o *single-link* e o *complete-link* e ambos possuem variações de implementação. O DBSCAN é uma implementação de algoritmos baseados em densidade e a Árvore Geradora de Custo Mínimo (*Minimum Spanning Tree*) é uma implementação utilizada para agrupamento com grafos (de Castro; Ferrari, 2016).

Os algoritmos particionais mais utilizados são o *k*-médias (*k-means*), o *k*-medoides (*k-medoids*) e variações de ambos (de Castro; Ferrari, 2016). Este trabalho utilizará um algoritmo particional do tipo *k*-medoides, na sua variação conhecida como **particionamento em torno dos medoides** (Kaufman; Rousseeuw, 1987). A escolha do *k*-medoides deve-se ao fato de que a matriz de distâncias é obtida a partir da correlação entre as séries temporais das ações. Dessa forma, a distância entre os objetos da base (ativos) e um protótipo que não seja outro objeto da base não existe e, portanto, o protótipo deve ser um objeto da base (medoide). Como o algoritmo *k*-medoides opera recebendo como entrada uma matriz de distâncias já calculada, ele mostrou-se apropriado para ser aplicado a esse problema.

2.3.1 Algoritmo *k*-Médias

O algoritmo *k*-médias recebe como entrada o parâmetro *k*, que representa o número de grupos em que se deseja dividir a base de dados. Dessa forma, o algoritmo divide a base de dados em *k* conjuntos disjuntos de objetos. O resultado esperado do agrupamento é que a similaridade **intragrupo** seja alta e a similaridade **intergrupo** seja baixa (de Castro; Ferrari, 2016).

O algoritmo é inicializado definindo k protótipos (centroides) de forma aleatória ou a partir de objetos da base de dados. Um protótipo é um objeto que representa um grupo (*cluster*) de objetos. Após a inicialização, cada objeto é atribuído ao grupo cujo centroide esteja mais próximo. Os centroides têm suas posições atualizadas de acordo com a média da distância de todos os objetos do grupo. Após o reposicionamento dos centroides, o processo de atribuir os objetos aos grupos e recalculando o posicionamento do centroide é realizado de forma iterativa até que o algoritmo tenha convergido e não haja mais atualizações dos centroides e dos grupos (de Castro; Ferrari, 2016).

2.3.2 Algoritmo k -Medoides

O algoritmo k -medoides é um método de agrupamento similar ao k -médias, mas cujo protótipo é sempre um objeto da base. Enquanto o k -médias é sensível a objetos ruidosos e discrepantes da base de dados, por causa da média que o algoritmo utiliza para a definição e atualização dos grupos, o k -medoides é mais robusto a tais problemas, pois ao invés de utilizar a média das distâncias para representar o centro do grupo, o k -medoides utiliza um objeto do próprio grupo para representar o centro. O objeto do centro do grupo no k -medoides se chama **medoide** e representa o objeto mais central do grupo com base no mínimo da soma das distâncias para os outros objetos do grupo (Jin; Han, 2010).

Uma implementação comum encontrada na literatura para o k -medoides é chamado de **particionamento em torno dos medoides** (do inglês *partitioning around medoids*, PAM), proposto por Kaufman e Rousseeuw (1987). O PAM consiste em selecionar k objetos da base de dados para serem os medoides iniciais. Em seguida, cada objeto é associado ao grupo cujo medoide é o mais próximo. O reposicionamento dos medoides é realizado minimizando a função objetivo, que é a soma das distâncias dos objetos do grupo ao medoide. Após o reposicionamento dos medoides, o processo de atribuir os objetos aos grupos e recalculando o posicionamento do medoide é realizado de forma iterativa até que o algoritmo tenha convergido e não haja mais atualizações dos medoides e dos grupos (Jin; Han, 2010). O pseudocódigo do algoritmo k -medoides é apresentado no Algoritmo 1. O PAM opera de forma a minimizar a distância intragrupo (Krishnapuram; Joshi; Yi, 1999) e, de acordo com de Castro e Ferrari (2016), a maior parte dos algoritmos de agrupamento particional visam minimizar o erro quadrático entre os objetos de um grupo

e seu protótipo (distância intragrupo).

Reynolds, Richards e Rayward-Smith (2004) destacam que no algoritmo k -medoides não há a necessidade de recalculá-la a distância entre os objetos e o protótipo a cada interação, pois as distâncias já estão calculadas na matriz de distâncias, reduzindo o custo computacional quando comparado ao k -médias.

Algorithm 1 Pseudocódigo do algoritmo k -medoides PAM

```
def kmedoides(k, data):
    medoides = selecionar_medoides_aleatorios(k, data)
    medoides_copia = medoides
    while not medoides == medoides_copia:
        grupos = calcular_grupos(data, medoides, k)

        medoides_copia = atualizar_medoides_grupo(
            grupos, medoides
        )

        medoides = medoides_copia
    return grupos
```

2.3.3 Agrupamento Nebuloso

O Agrupamento Nebuloso (do inglês *fuzzy clustering*) é uma técnica que pode ser aplicada aos algoritmos k -médias e k -medoides, cujo objetivo é atribuir um grau de pertinência de cada objeto em relação a cada um dos grupos. Enquanto o agrupamento sem o método nebuloso, denominado *crisp*, indica se um objeto pertence ou não a um grupo, no método nebuloso todo objeto pertence a todos os grupos, porém com diferentes graus de pertinência (de Castro; Ferrari, 2016).

Uma característica do agrupamento nebuloso é que para cada objeto i do grupo j , existe um valor $\mu_{i,j}$ que corresponde ao grau de pertinência do objeto aos grupos. A soma dos graus de pertinência de cada objeto da base a todos os grupos deve ser igual a 1, que também é o valor máximo de pertinência de um objeto a um grupo (de Castro; Ferrari,

2016):

$$\sum_{j=1}^k \mu_{i,j} = 1 \quad (7)$$

Para o algoritmo k -médias, o posicionamento do centroide é calculado considerando o grau de pertinência dos objetos ao grupo. No entanto, para o algoritmo k -medoides, cujo protótipo é um objeto da base, o grau de pertinência não interfere na definição do centro do grupo, dado que o grau de pertinência é determinado após os medoides serem calculados. Para calcular o grau de pertinência dos objetos de um agrupamento realizado com o k -medoides, é utilizada a Equação 8 (Bezdek, 1981; Krishnapuram; Joshi; Yi, 1999):

$$\mu_{ij} = \frac{\left(\frac{1}{r(x_i, v_j)}\right)^{1/(m-1)}}{\sum_{k=1}^c \left(\frac{1}{r(x_i, v_k)}\right)^{1/(m-1)}}, \quad (8)$$

sendo x_i o objeto cuja pertinência está sendo calculada, $m > 1$ o parâmetro de **fuzzy-ficação** (um coeficiente que pondera o grau de pertinência), $r(., .)$ uma medida de distância entre os ativos, v_j o medoide do grupo do objeto x_i , e v_k os outros medoides.

2.4 Índice Ibovespa

Segundo a B3 (2015), o índice Ibovespa é o resultado de uma carteira teórica elaborada de acordo com uma metodologia e que tem por objetivo ser um indicador do desempenho médio dos ativos de maior representatividade e negociabilidade do mercado acionário brasileiro. O índice Ibovespa é reavaliado e rebalanceado a cada quatro meses. Para um ativo ser incluído no índice Ibovespa, ele deve, no período de vigência de três carteiras anteriores, atender aos seguintes critérios:

- Estar entre os 85% maiores ativos no índice de negociabilidade.
- Ter presença de 95% nos pregões.
- Ter participação de volume financeiro maior que 0,1% do total do mercado a vista.
- Não ter cotação inferior a R\$ 1,00 (um real) por ação.

A ponderação do ativo no índice Ibovespa é feita pelo valor de mercado das ações disponíveis para negociação das empresas, com limite de participação baseada na liquidez do ativo.

3 Um Framework para Alocação de Portfólios Baseado em Agrupamento Particional

López de Prado (2016) propôs o uso de **aprendizagem de máquina** para alocação de portfólios utilizando uma técnica chamada de **Construção Hierárquica de Portfólio** (do inglês *Hierarchical Portfolio Construction*), que utiliza um algoritmo de agrupamento hierárquico para realizar a alocação com base no resultado do agrupamento.

Sharpe (1964) nota que, embora conceitualmente a otimização de portfólio seja simples, envolvendo a criação de fronteiras eficientes, as soluções são complexas. Normalmente esse problema é resolvido com programação quadrática paramétrica, que recorre ao uso de matrizes inversas para sua solução (Markowitz, 1956 apud Sharpe, 1964).

López de Prado (2016) nota que nos otimizadores de portfólios, conforme ativos correlacionados são adicionados a essa matriz utilizada para a otimização de um portfólio, o número de condicionamento da matriz aumenta, tornando o cálculo da inversa instável. O método proposto de alocação de portfólios a partir de agrupamento particional, não necessita calcular a matriz inversa, pois o processo de alocação consiste nos seguintes passos:

1. Realizar o agrupamento a partir da matriz de correlação entre os ativos;
2. Realizar a alocação intergrupo; e
3. Realizar a alocação intragrupo.

Inspirado na ideia de realizar alocações de portfólios utilizando algoritmos de agrupamento e considerando diferentes formas de estruturar o processo de alocação de portfólios, esse trabalho propõe um *framework* de alocação de portfólios de ações usando algoritmos de agrupamento particional. O *framework* corresponde a um processo sequencial de aplicação de diferentes blocos construtivos de forma que, ao final do processo, obtenha-se uma alocação de portfólios.

3.1 Fluxo do Processo de Alocação

O *framework* inicia calculando a matriz de correlação entre os ativos e depois a matriz de distâncias entre eles a partir da correlação. Em seguida é aplicado um algoritmo de agrupamento particional, que segmenta os ativos em diferentes grupos (*clusters*) de acordo com a similaridade, ou dissimilaridade, entre eles. Uma vez segmentados, é necessário realizar a alocação dentro de cada grupo (intragrupo) e, por fim, a alocação entre os grupos (intergrupo), onde será feita a alocação final do portfólio levando em consideração as alocações de cada grupo. Após a alocação, o desempenho do portfólio gerado é avaliado. Esse fluxo está resumido na Figura 4 e cada um dos blocos construtivos será descrito separadamente nas seções seguintes.

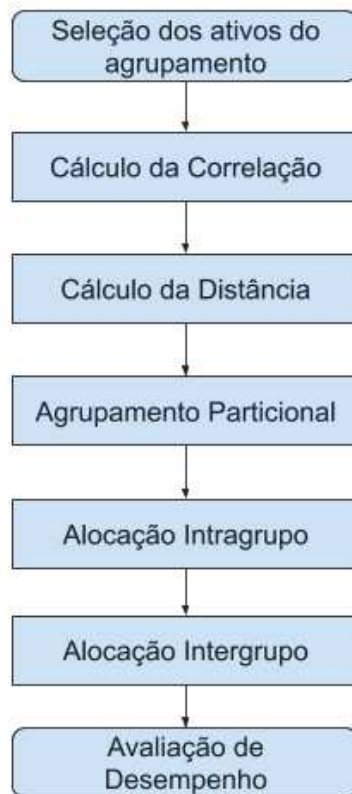


Figura 4: Framework proposto para alocação de ativos utilizando agrupamento particional.

3.2 Pré-Processamento: Cálculo da Matriz de Correlação e Distância

Para realizar o agrupamento a partir de séries temporais financeiras, é necessário realizar algumas transformações com as séries. Segundo Marti et al. (2017), a técnica para cálculo de agrupamento em séries temporais foi definida por Mantegna (1999) para agrupamentos hierárquicos e tem sido amplamente adotada desde então. A técnica pode ser definida como a seguir:

- Seja N o número de ativos que serão agrupados.
- Seja $P_i(t)$ o preço do ativo i no tempo t , onde $1 \leq i \leq N$.
- Seja $r_i(t)$ o retorno no tempo t do ativo i , onde:

$$r_i(t) = \log P_i(t) - \log P_i(t-1). \quad (9)$$

- Para cada par de ativos i, j , é calculada a correlação $\rho_{i,j}$:

$$\rho_{i,j} = \frac{\sum (r_i - \bar{r}_i)(r_j - \bar{r}_j)}{\sqrt{\sum (r_i - \bar{r}_i)^2} \sqrt{\sum (r_j - \bar{r}_j)^2}} \quad (10)$$

- Os coeficientes de correlação $\rho_{i,j}$ são convertidos em distância $d_{i,j}$:

$$d_{i,j} = \sqrt{2(1 - \rho_{i,j})} \quad (11)$$

3.3 Agrupamento

Após as etapas de pré-processamento que geram a matriz de distâncias entre as correlações dos ativos, Mantegna (1999) aplica o algoritmo de cálculo da **árvore geradora mínima** para encontrar grupos de forma hierárquica. Os mesmos passos são utilizados por López de Prado (2016) e serão empregados no *framework* com a aplicação de um algoritmo de agrupamento particional.

Neste trabalho, em vez de utilizar o retorno em log, é utilizado o retorno líquido simples, e outra diferença é que a equação da distância utilizada é a mesma adotada por López de Prado (2016), conforme a Equação 12. Como os resultados do agrupamento particional serão comparados com o portfólio gerado proposto pela técnica de López de

Prado (2016), foi utilizado o mesmo pré-processamento nos dados a fim de ter uma base de comparação equivalente.

$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})} \quad (12)$$

Após o cálculo da distância, é realizado o agrupamento utilizando um algoritmo de agrupamento particional, como o k -medoides (de Castro; Ferrari, 2016).

De maneira geral, o algoritmo k -medoides é sensível aos medoides iniciais, ou seja, diferentes escolhas dos medoides iniciais resultam em diferentes particionamentos da base de dados.

A partir dos medoides iniciais, o algoritmo k -medoides é executado, verificando a qual grupo cada objeto da base pertence, sendo o grupo formado pelos objetos que apresentarem a menor distância ao medoide do grupo.

Após o primeiro particionamento dos grupos, todos os objetos em cada grupo são verificados a fim de encontrar o objeto que apresente a menor distância para todos os objetos do grupo. Esse processo permite o recálculo do posicionamento do medoide.

Como o algoritmo é iterativo, a associação ao grupo e recálculo dos medoides é executada até que o algoritmo tenha convergido e não existam mais atualizações a serem feitas.

O resultado do agrupamento do algoritmo k -medoides é sensível à escolha dos medoides iniciais. Dessa forma, o *framework* permite o uso de diferentes métodos de inicialização dos protótipos (medoides) iniciais. Serão consideradas duas formas de inicializar os medoides: um método tradicional de inicialização aleatória; e é proposto um método heurístico que toma como medoides iniciais os ativos que possuem a menor volatilidade observada no conjunto. O uso do método heurístico é motivado a fim de obter o mesmo resultado a cada execução. Seguindo essa heurística, os medoides iniciais serão os ativos com menor desvio padrão da série passada de preços. Isso faz com que os ativos com menor volatilidade tenham maior peso no portfólio e seja potencialmente produzido um portfólio com menor volatilidade. Segundo Baker, Bradley e Wurgler (2011), portfólios compostos por ações de baixa volatilidade apresentam retornos maiores e *drawdowns* menores.

3.4 Alocação Intragrupo

Uma vez identificados os grupos e os objetos pertencentes a cada um deles, é necessário realizar a **alocação intragrupo**, que tem como objetivo alocar os recursos em cada um dos ativos que compõem o grupo.

Nessa dissertação são propostos dois métodos de alocação intragrupos: o **método nebuloso** (do inglês *fuzzy*), que calcula o grau de pertinência de cada objeto a cada grupo; e o **método da roleta**, que faz a alocação proporcional a um indicador de desempenho, dando maior peso aos objetos com maiores indicadores, e vice-versa.

3.4.1 Método Nebuloso

Em tarefas de agrupamento de dados, um processo nebuloso (*fuzzy*) pode ser utilizado para calcular a pertinência de cada objeto a cada grupo. Esse método é utilizado aplicando a Equação 8 para encontrar o grau de pertinência do objeto i ao grupo j . O grau de pertinência indica que quanto mais próximo o objeto do medoide, maior a pertinência dele àquele grupo, enquanto quanto mais longe do medoide, menor a pertinência ao grupo. Como o medoide é o objeto que possui a menor distância média a todos os objetos do grupo, entende-se que ele é o objeto que melhor representa o grupo e, portanto, seu grau de pertinência ao grupo é máximo. Após a aplicação da equação, teremos o grau de pertinência para todos os objetos de todos os grupos, onde $\mu_{i,j}$ representa a informação de pertinência do objeto i ao grupo que o objeto pertence, representado por j .

3.4.2 Método da Roleta

O mecanismo de seleção conhecido como **Método da Roleta** nos Algoritmos Genéticos (Golberg, 1989) é apropriado para definir a alocação a ser realizada em cada grupo no *framework*. O método da roleta considera um indicador de desempenho (do inglês *fitness*) dos objetos (ativos) e pondera a relevância entre todos os objetos de maneira proporcional a esse indicador. No caso da alocação intragrupo, é considerado o desempenho dos ativos no período analisado, fazendo com que os ativos que tiveram melhor desempenho tenham maior relevância (alocação) no grupo, enquanto os que tiveram pior desempenho tenham relevância menor no grupo.

A Equação 13 representa o funcionamento do método da roleta (Lipowski; Lipowska, 2012):

$$\mu_{i,j} = \frac{f_{i,j}}{\sum_{c=1}^n f_{c,j}}, \quad (13)$$

sendo $\mu_{i,j}$ a representação da importância do objeto i ao grupo j , de acordo com métrica f utilizada para calcular e ponderar a relevância de cada objeto do grupo. No caso deste trabalho, a métrica f utilizada foi o desempenho passado dos ativos. No somatório, n é o número de ativos do grupo j .

3.5 Alocação Intergrupo

Uma vez identificados os grupos, os objetos pertencentes a eles e realizada a alocação intragrupo, é necessário realizar a alocação intergrupo. O resultado da alocação Intergrupo será a alocação final do portfólio. A alocação intergrupo considera as informações presentes em cada objeto de cada grupo, encontradas a partir da utilização dos métodos anteriormente apresentados, que são:

- **método nebuloso:** cada objeto i possui um $\mu_{i,j}$, que representa o grau de pertinência ao grupo j .
- **método da roleta:** cada objeto i possui um $\mu_{i,j}$, que apresenta a relevância ao grupo j , e a essa relevância é encontrada a partir de uma métrica definida a priori.

Inicialmente serão propostos dois métodos de alocação intergrupo: **alocação equilibrada de capital;** e **alocação baseada na roleta.**

3.5.1 Alocação Equilibrada de Capital

A alocação equilibrada de capital intergrupo é aquela que divide o valor investido de forma equilibrada em todos os grupos, ou seja, cada grupo recebe a mesma quantidade de capital. DeMiguel, Garlappi e Uppal (2007) apresentam essa estratégia como sendo uma regra de diversificação ingênua, mas que tende a ter uma boa performance apresentando um bom *sharpe ratio* e retorno quando comparada a uma estratégia tradicional de

otimização de alocação de capital.

$$C_j = \frac{C_t}{k}, \quad (14)$$

sendo C_j o capital para cada grupo, C_t o capital total e k o número de grupos. O método proposto por DeMiguel, Garlappi e Uppal (2007) é expresso pela Equação 14, porém a alocação final do portfólio deve totalizar todo o capital disponível para investimento, então a normalização conforme a Equação 15.

$$w_{i,j} = \frac{\mu_{i,j} * \frac{1}{k}}{\sum_{c=1}^m \sum_{g=1}^n \mu_{c,g} * \frac{1}{k}}, \quad (15)$$

sendo $w_{i,j}$ o percentual do ativo no portfólio final, $\mu_{i,j}$ obtido a partir da alocação intergrupo, k a quantidade de grupos, e $\mu_{c,g}$ utilizado para encontrar a soma de todos os valores de todos os grupos.

3.5.2 Alocação Baseada na Roleta

O **método da roleta**, utilizado no contexto da alocação intragrupo, também será utilizado no contexto da alocação intergrupo. Enquanto na alocação intragrupo esse método era aplicado para dar relevância a objetos dentro dos grupos, no caso da alocação intergrupo esse método é utilizado para dar relevância aos grupos. No caso da alocação intergrupo, é considerado o desempenho de todos os ativos do grupo no período analisado, fazendo com que os grupos que tiveram melhor desempenho, tenham maior relevância na alocação final, enquanto os que tiveram pior desempenho tenham relevância menor no grupo. A Alocação Baseada na Roleta, é realizada conforme a Equação 16 e da mesma forma como na Alocação Equilibrada de Capital, este método normaliza as alocações entre os grupos para que o investimento final totalize todo o capital disponível para investimento.

$$w_{i,j} = \frac{\mu_{i,j}}{\sum_c \mu_{c,j}} * \frac{f_j}{\sum_g f_g}, \quad (16)$$

sendo $w_{i,j}$ o percentual do ativo no portfólio final, $\mu_{i,j}$ obtido a partir da alocação intergrupo, $\mu_{c,j}$ utilizado para encontrar a soma dos valores do grupo j , f_j a métrica utilizada

para aplicar a roleta no grupo j e f_g é utilizado para encontrar a soma das métricas de todos os grupos.

Após a utilização da Alocação Equilibrada de Capital e da Alocação Baseada na Roleta, o portfólio final está composto e apto para o investimento.

3.6 Avaliação de Desempenho

Para a avaliação das estratégias geradas, serão utilizados os indicadores abaixo:

- **Retorno do Período:** é o indicador da variação da estratégia no período.

$$R = \frac{p(t)}{p(i)} - 1 \quad (17)$$

onde $p(t)$ é o valor acumulado da estratégia no instante observado e $p(i)$ é o valor inicial da estratégia.

- **Volatilidade Anualizada:** é a medida da dispersão dos retornos da série.

$$\sigma_p = \sigma_{diario} * \sqrt{252} \quad (18)$$

onde σ_p é a volatilidade da série no período, σ_{diario} é o desvio padrão da série de retornos diários no período e $\sqrt{252}$ é a anualização do desvio padrão. Essa medida informa a volatilidade realizada do portfólio.

- **Sharpe Ratio:** é uma medida que indica o retorno ajustado pelo risco tomado no investimento Sharpe (1966). O indicador é obtido conforme a Equação 6.
- **Turnover Ratio:** é a medida que indica a mudança do portfólio a cada rebalanceamento.

$$TR = \sum_{i=1}^N |w_i(t) - w_i(t-1)| \quad (19)$$

onde $w_i(t)$ é o peso do ativo no mês atual e $w_i(t-1)$ é o peso do ativo no mês anterior.

- **Maximum Drawdown:** é o número que indica em percentual a maior queda do pico ao vale da série de retornos. A equação abaixo, apresentada por Magdon-Ismail et

al. (2004), é utilizada para encontrar o valor de *Maximum Drawdown*

$$MDD(t) = \sup_{t \in [0, T]} [\sup_{s \in [0, t]} X(s) - X(t)], \quad (20)$$

onde $X(t)$ representa o valor do portfólio no instante t e $X(s)$ o maior valor do portfólio no passado do instante t .

3.7 Trabalhos Correlatos

Ao observar a otimização proposta por Markowitz (1952) e a fronteira criada pelo método, naturalmente propõe-se a aplicação de algoritmos multiobjetivos. Na pesquisa por contribuições realizadas a essa área, é possível observar que é um assunto amplamente estudado conforme a revisão por Ponsich, Jaimes e Coello (2013).

A tarefa de agrupamento pode ser considerada um problema de otimização (Grötschel; Wakabayashi, 1989), e a seleção de portfólios, sendo originalmente uma tarefa de otimização, também pode ser resolvida utilizando algoritmos de agrupamento de dados.

A partir dos trabalhos de López de Prado (2016) e Raffinot (2017), que aplicam algoritmos de agrupamento hierárquico para realizar alocação de portfólios, foi proposto o uso de algoritmos particionais nebulosos para alocação na nossa pesquisa. A ideia de usar o agrupamento nebuloso era de considerar a pertinência aos grupos como uma medida da alocação de cada ativo.

López de Prado (2016) e Raffinot (2017) utilizam como pré-processamento do agrupamento, os conceitos propostos em (Mantegna, 1999) e utilizados desde então. Mantegna (1999) foi o primeiro a propor e aplicar o uso de algoritmos de agrupamento em séries temporais de ações a partir da matriz de correlação. O método proposto demonstrou ser um método robusto para encontrar quais grupos de ações são mais expostas e afetadas por variáveis econômicas.

Seguindo o conceito de agrupamento hierárquico, López de Prado (2016) propôs um método de alocação hierárquica utilizando informações contidas na matriz de covariância e criar portfólios mais diversificados do que os criados com métodos tradicionais a partir dessas informações. Os resultados obtidos foram portfólios mais diversificados e com

menor volatilidade.

A partir do trabalho de López de Prado (2016), Raffinot (2017) testou os seguintes métodos de agrupamento hierárquico: *Simple Linkage*, *Complete Linkage*, *Average Linkage*, *Ward's Method* e *Directed Bubble Hierarchical Tree*. O autor concluiu que os portfólios criados a partir de agrupamento hierárquico são mais diversificados e possuem melhores retornos ajustados ao risco.

Marti et al. (2017) revisam sobre agrupamentos no mercado financeiro comentando sobre o trabalho de Mantegna (1999) e como a técnica definida por ele se tornou base de agrupamento de séries temporais. A revisão também comenta sobre aplicações de agrupamentos para a seleção de portfólios, *trading*, gerenciamento de risco e elaboração de políticas financeiras.

Nanda, Mahanty e Tiwari (2010) aplicaram os algoritmos k -médias, mapa auto-organizável e *fuzzy C-means* em ações. Os autores aplicam esses algoritmos não apenas nas séries temporais a partir da matriz de correlação, como é o objetivo desse trabalho, mas aplicam esses algoritmos com indicadores utilizados para realizar uma análise fundamentalista e de balanço de uma empresa, além de utilizar dados obtidos dos preços. O resultado encontrado foi que o algoritmo k -médias gerou agrupamentos mais compactos, e quando comparado com o índice do mercado, apresentou resultado melhor.

Tola et al. (2008) aplicaram análise de agrupamento para otimização de portfólios. A partir de conceitos da *Random Matrix Theory*, utilizaram os algoritmos hierárquicos *Single linkage* e *Average linkage* para realizar o agrupamento. Também contribuíram com uma técnica para indicar qual o tamanho ideal de um portfólio. Os resultados obtidos são comentados pelos autores na questão da diferença de técnica de filtro de cada algoritmo. Os testes apresentaram resultados consistentes no período de avaliação em comparação com o período de testes.

Dose e Cincotti (2005) aplicaram algoritmos de agrupamento hierárquico para construção de portfólios que têm por objetivo refletir o mesmo desempenho de um índice de mercado, onde o resultado do agrupamento é um portfólio otimizado que consegue replicar o resultado do índice em questão.

4 Avaliação de Desempenho

O *framework* proposto nessa dissertação possui sete fases sequenciais, a saber:

1. Selecionar os ativos;
2. Calcular a correlação entre os ativos;
3. Calcular a matriz de distâncias;
4. Aplicar o algoritmo de agrupamento particional;
5. Realizar a alocação intragrupo;
6. Realizar a alocação intergrupo;
7. Avaliar o desempenho.

Para avaliar a aplicação do *framework* proposto, na fase (1) serão utilizadas ações negociadas na Bolsa de Valores Brasileira (B3) e que fazem parte do índice Ibovespa no respectivo período de avaliação da alocação. Os ativos que fazem parte do índice Ibovespa encontram-se no Apêndice A. Os testes serão realizados considerando o período de 29/12/2009 até 30/12/2017.

Na Fase (2), será adotado o cálculo da correlação, conforme Equação 10, e na Fase (3) a distância será calculada de acordo com a Equação 12. O algoritmo de agrupamento particional (Fase (4)) selecionado foi o *k*-medoides com duas formas distintas de inicializar os protótipos. A alocação intragrupo (Fase (5)) utilizou os métodos nebuloso e pela roleta, e a alocação intergrupo (Fase (6)) utilizou os métodos de alocação equilibrada e pela roleta. O desempenho do *framework* (Fase (7)) considerou as medidas apresentadas na Seção 3.6.

4.1 Metodologia Experimental

Como as Fases (4), (5) e (6) do *framework* permitem algumas variações, diferentes combinações para eles resultaram em diferentes configurações de avaliação, conforme apresentado a seguir:

- Configuração 1: Seleção aleatória dos medoides iniciais, alocação nebulosa intragrupo e alocação equilibrada intergrupo.
- Configuração 2: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação equilibrada intergrupo.
- Configuração 3: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação intergrupo pelo método da roleta.
- Configuração 4: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação equilibrada intergrupo.
- Configuração 5: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação intergrupo também pelo método da roleta.

O processo de alocação de portfólio utilizando agrupamento particional é realizado mensalmente, com rebalanceamento no primeiro dia útil de cada mês. Ou seja, o algoritmo de agrupamento é executado no primeiro dia do mês e mudanças no portfólio são realizadas.

Para a execução do algoritmo de agrupamento, no momento do rebalanceamento serão consideradas as informações de preço de fechamento em um período anterior de 12 ou 24 meses. A partir dessas informações desses períodos, a correlação é calculada. Como não é possível saber a priori a quantidade ideal de grupos em um agrupamento particional, serão executadas alocações de portfólio utilizando valores de k entre 3 e 20.

Em cada rebalanceamento é necessário um conjunto de ativos para calcular a correlação entre os ativos e então realizar os agrupamentos particionais. Os ativos que farão parte de cada rebalanceamento são os que fazem parte da carteira teórica do índice Ibovespa no respectivo mês de avaliação do rebalanceamento. A composição da carteira teórica a cada rebalanceamento do índice está descrita no Apêndice A.

4.2 Resultados e Discussão

4.2.1 Configuração 1: Seleção aleatória dos medoides iniciais, alocação nebulosa intragrupo e alocação equilibrada intergrupo

Após a execução dos testes conforme a metodologia descrita, os resultados foram organizados em uma tabela considerando cada valor de k .

A Tabela 1 apresenta os resultados para as estratégias que utilizam 12 meses de informação de preço dos ativos para calcular a correlação e a Tabela 2 apresenta o resultado para a mesma configuração, porém as estratégias foram calculadas utilizando 24 meses de informação de preço para o cálculo da correlação. Como o algoritmo k -medoides é estocástico e sensível à condição inicial, foram executadas 10 simulações para cada valor de k com a inicialização aleatória dos medoides.

As Tabelas 1 e 2 apresentam as informações por coluna divulgando os dados de média, pior e melhor, utilizando o formato: Média [Pior, Melhor]. A informação de Média, apresenta o resultado médio das 10 execuções com inicialização aleatória dos medoides. A informação de Pior, apresenta o pior resultado obtido no teste com inicialização aleatória dos medoides. A informação de Melhor, apresenta o melhor resultado obtido no teste com inicialização aleatória dos medoides.

Para efeito de comparação, é apresentado o resultado do índice Ibovespa no período, o desempenho de um portfólio utilizando o modelo de **média-variância** (MVP) proposto por Markowitz (1952) e também o modelo *Hierarchical Risk Parity* (HRP) proposto por López de Prado (2016). Os portfólios foram construídos analisando janelas de 12 e 24 meses de correlação e rebalanceamento mensal.

Tabela 1: Comparação das estratégias utilizando janela de correlação de 12 meses.

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	58.1% [32.0%, 88.1%]	22.0% [22.4%, 21.7%]	76.9% [81.7%, 77.0%]	-47.5% [-53.1%, -49.7%]	-3.1 [-4.2, -1.7]
4	59.5% [13.2%, 103.4%]	22.0% [22.0%, 22.0%]	86.8% [89.4%, 84.0%]	-44.3% [-56.0%, -40.8%]	-3.0 [-5.1, -1.0]
5	54.8% [-4.5%, 113.6%]	22.0% [22.0%, 22.2%]	91.6% [92.1%, 91.1%]	-48.1% [-49.4%, -36.8%]	-3.2 [-5.9, -0.5]
6	62.2% [28.0%, 114.9%]	21.9% [21.3%, 22.3%]	97.2% [100.5%, 99.1%]	-45.6% [-54.7%, -49.3%]	-2.9 [-4.6, -0.5]
7	47.4% [9.9%, 75.0%]	22.2% [21.9%, 22.4%]	101.4% [103.1%, 101.3%]	-48.4% [-54.0%, -40.9%]	-3.5 [-5.3, -2.3]
8	37.2% [-5.0%, 116.9%]	21.8% [21.9%, 22.2%]	103.6% [102.5%, 101.4%]	-47.1% [-49.4%, -32.6%]	-4.1 [-6.0, -0.4]
9	55.6% [22.6%, 79.6%]	21.7% [21.7%, 21.4%]	106.9% [106.0%, 109.0%]	-43.7% [-49.3%, -43.3%]	-3.2 [-4.8, -2.2]
10	38.5% [11.9%, 71.9%]	21.7% [22.0%, 22.0%]	107.4% [107.5%, 109.1%]	-48.4% [-51.5%, -44.7%]	-4.0 [-5.2, -2.4]
11	33.7% [-9.4%, 88.6%]	21.5% [21.6%, 21.5%]	108.3% [108.4%, 108.4%]	-49.6% [-60.7%, -39.6%]	-4.3 [-6.2, -1.7]
12	39.5% [-4.8%, 83.4%]	21.8% [21.6%, 21.8%]	109.4% [108.9%, 114.8%]	-47.4% [-61.7%, -40.4%]	-3.9 [-6.1, -1.9]
13	48.8% [17.6%, 104.0%]	21.8% [21.8%, 22.2%]	109.4% [109.8%, 109.2%]	-46.3% [-47.1%, -32.6%]	-3.5 [-5.0, -1.0]
14	54.7% [26.2%, 115.2%]	21.7% [21.7%, 21.8%]	108.6% [109.2%, 108.3%]	-42.7% [-50.3%, -35.7%]	-3.3 [-4.6, -0.5]
15	50.2% [18.0%, 77.5%]	21.8% [22.0%, 21.7%]	108.4% [111.7%, 112.2%]	-47.0% [-58.5%, -44.3%]	-3.5 [-4.9, -2.2]
16	41.0% [18.4%, 67.1%]	21.8% [22.3%, 22.2%]	108.2% [109.3%, 109.1%]	-48.3% [-47.7%, -49.7%]	-3.9 [-4.8, -2.6]
17	53.9% [30.7%, 91.5%]	21.5% [21.0%, 21.6%]	108.1% [111.5%, 108.3%]	-45.2% [-52.4%, -30.6%]	-3.3 [-4.5, -1.6]
18	43.6% [10.8%, 97.5%]	21.6% [21.3%, 21.7%]	106.4% [105.5%, 106.7%]	-45.6% [-46.7%, -35.0%]	-3.8 [-5.4, -1.3]
19	44.0% [31.2%, 64.8%]	21.6% [21.8%, 21.7%]	105.6% [106.6%, 106.0%]	-45.8% [-52.5%, -35.7%]	-3.8 [-4.3, -2.8]
20	52.3% [29.5%, 79.0%]	21.8% [22.2%, 21.5%]	105.0% [107.0%, 104.1%]	-44.5% [-52.4%, -35.6%]	-3.4 [-4.3, -2.2]
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	151,9%	14,2%	31,4%	-16,9%	1,8
HRP	69,0%	17,1%	21,6%	-30,0%	-3,3

Tabela 2: Comparação das estratégias utilizando janela de correlação de 24 meses.

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	46.2% [9.5%, 111.0%]	22.3% [22.7%, 22.6%]	81.1% [80.3%, 83.9%]	-50.6% [-60.6%, -43.1%]	-3.6 [-5.1, -0.6]
4	54.4% [22.5%, 99.5%]	22.2% [22.3%, 22.7%]	91.5% [91.8%, 94.6%]	-49.4% [-56.1%, -37.2%]	-3.2 [-4.6, -1.2]
5	38.5% [-31.4%, 143.4%]	22.3% [22.4%, 22.0%]	97.5% [96.2%, 100.4%]	-52.4% [-63.4%, -32.9%]	-3.9 [-7.0, 0.8]
6	39.6% [-9.6%, 98.4%]	22.0% [22.7%, 22.8%]	99.6% [96.7%, 97.1%]	-53.4% [-64.4%, -43.3%]	-3.9 [-6.0, -1.2]
7	40.9% [-3.9%, 82.9%]	22.1% [21.7%, 22.1%]	102.1% [101.2%, 104.6%]	-50.3% [-59.2%, -38.7%]	-3.8 [-6.0, -1.9]
8	49.4% [18.3%, 104.8%]	22.0% [21.9%, 22.4%]	105.0% [104.9%, 107.3%]	-50.6% [-50.6%, -48.5%]	-3.5 [-4.9, -0.9]
9	50.1% [-3.0%, 160.5%]	22.0% [21.8%, 22.3%]	106.6% [108.6%, 107.2%]	-49.6% [-63.4%, -30.5%]	-3.4 [-5.9, 1.6]
10	47.1% [16.7%, 114.1%]	21.8% [21.9%, 22.0%]	106.4% [108.2%, 108.0%]	-48.5% [-51.6%, -36.0%]	-3.6 [-5.0, -0.5]
11	34.8% [-7.5%, 84.0%]	21.9% [21.9%, 21.9%]	108.7% [109.9%, 105.1%]	-49.8% [-67.2%, -30.5%]	-4.2 [-6.1, -1.9]
12	51.2% [6.1%, 102.1%]	22.1% [21.8%, 21.6%]	110.5% [110.1%, 108.6%]	-46.4% [-56.9%, -38.3%]	-3.4 [-5.5, -1.1]
13	51.3% [8.4%, 98.1%]	21.9% [22.0%, 22.0%]	109.5% [110.0%, 108.0%]	-46.4% [-60.6%, -46.9%]	-3.4 [-5.3, -1.2]
14	36.2% [-14.1%, 56.4%]	21.8% [21.4%, 21.4%]	109.0% [106.3%, 107.8%]	-48.8% [-60.7%, -48.7%]	-4.1 [-6.5, -3.2]
15	39.3% [8.1%, 93.5%]	21.8% [21.2%, 21.8%]	106.6% [104.1%, 106.1%]	-48.9% [-53.1%, -40.8%]	-4.0 [-5.6, -1.5]
16	46.3% [21.2%, 73.5%]	21.9% [21.5%, 22.1%]	107.2% [107.3%, 107.3%]	-46.5% [-47.6%, -40.2%]	-3.6 [-4.8, -2.4]
17	33.6% [3.2%, 65.4%]	21.9% [22.3%, 21.7%]	107.2% [108.8%, 107.7%]	-50.0% [-61.5%, -46.6%]	-4.2 [-5.5, -2.8]
18	43.5% [2.6%, 80.4%]	22.0% [22.0%, 22.1%]	106.3% [108.3%, 108.7%]	-47.1% [-56.5%, -41.0%]	-3.7 [-5.6, -2.0]
19	42.0% [13.4%, 90.0%]	21.8% [21.9%, 21.4%]	104.8% [106.1%, 102.5%]	-47.7% [-55.1%, -36.5%]	-3.8 [-5.1, -1.7]
20	46.8% [8.6%, 99.9%]	21.7% [21.5%, 21.7%]	104.0% [101.9%, 103.0%]	-47.4% [-52.4%, -38.0%]	-3.6 [-5.4, -1.2]
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	157,4%	14,6%	19,0%	-16,8%	2,1
HRP	78,5%	17,6%	17,9%	-32,5%	-2,7

Os testes foram realizados sempre utilizando um portfólio 100% comprado. É possível observar, a partir da análise dos resultados das estratégias, que o desempenho é bastante impactado pelos medoides iniciais, pois existe uma grande diferença entre o melhor e o pior resultado para o mesmo valor de k .

É possível observar nas Tabelas 1 e 2 que o resultado médio dos testes é melhor que do índice Ibovespa. A volatilidade manteve-se próxima a observada no índice Ibovespa e o *turnover ratio* aumenta conforme o valor de k aumenta, porém o retorno médio é melhor nas estratégias com valor de k maior. O *drawdown* médio é próximo ao do índice Ibovespa e em algumas configurações esse valor é maior.

4.2.2 Configuração 2: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação equilibrada intergrupo

Dada a sensibilidade do algoritmo k -medoides aos medoides iniciais, foi utilizada uma heurística de inicialização dos medoides tomando como medoides iniciais as ações que apresentaram a menor volatilidade na amostra analisada para o período de cálculo da correlação.

A Tabela 3 apresenta resultados para testes que foram realizados utilizando 12 meses de informação de preço para cálculo da correlação. Para alocação intragrupo foi utilizada a alocação equilibrada de capital entre os grupos e para a alocação intergrupo foi utilizado o processo nebuloso. Foram testadas estratégias com valores de k entre 3 e 20.

A Tabela 4 utiliza as mesmas configurações com análise de 24 meses de correlação.

Tabela 3: Comparação das estratégias utilizando janela de correlação de 12 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	36,6%	21,6%	51,4%	-48,9%	-4,1
4	38,2%	21,2%	59,8%	-45,2%	-4,1
5	38,7%	20,4%	66,1%	-44,7%	-4,2
6	54,4%	20,3%	63,2%	-42,6%	-3,5
7	66,7%	19,9%	66,6%	-34,8%	-3,0
8	106,9%	19,3%	66,0%	-28,7%	-1,0
9	97,1%	19,0%	64,6%	-29,7%	-1,5
10	103,3%	18,7%	65,1%	-30,2%	-1,2
11	129,0%	18,2%	65,0%	-26,9%	0,2
12	150,4%	17,6%	55,2%	-27,2%	1,4
13	110,0%	17,6%	51,3%	-33,0%	-0,9
14	123,7%	17,7%	50,0%	-30,9%	-0,1
15	144,5%	17,7%	47,8%	-30,4%	1,1
16	141,1%	17,6%	48,0%	-29,9%	0,9
17	121,2%	17,5%	45,9%	-29,2%	-0,2
18	99,6%	17,6%	44,8%	-29,2%	-1,5
19	93,9%	17,6%	44,3%	-30,0%	-1,8
20	71,8%	17,7%	43,1%	-32,0%	-3,0
12 ¹	150,4%	17,6%	55,2%	-27,2%	1,4
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	151,9%	14,2%	31,4%	-16,9%	1,8
HRP	69,0%	17,1%	21,6%	-30,0%	-3,3

¹Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

Tabela 4: Comparação das estratégias utilizando janela de correlação de 24 meses.

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	-8,5%	21,7%	45,0%	-62,1%	-6,2
4	25,2%	21,2%	50,7%	-55,7%	-4,7
5	38,5%	20,7%	54,1%	-43,2%	-4,2
6	64,6%	20,3%	53,4%	-35,3%	-3,0
7	30,6%	20,2%	57,4%	-44,3%	-4,7
8	34,7%	19,2%	55,3%	-42,4%	-4,7
9	42,0%	18,9%	51,0%	-42,7%	-4,4
10	40,5%	18,3%	47,8%	-43,8%	-4,7
11	56,8%	17,9%	43,2%	-41,7%	-3,8
12	52,9%	17,8%	40,7%	-32,8%	-4,1
13	63,6%	17,8%	38,4%	-29,8%	-3,5
14	59,3%	17,5%	35,5%	-30,9%	-3,8
15	66,7%	17,3%	34,0%	-30,3%	-3,4
16	68,2%	17,4%	30,9%	-30,8%	-3,3
17	56,7%	17,6%	30,4%	-37,3%	-3,9
18	50,9%	17,5%	28,6%	-37,7%	-4,3
19	61,6%	17,5%	27,2%	-36,3%	-3,7
20	72,6%	17,6%	28,1%	-32,5%	-3,0
16 ²	68,2%	17,4%	30,9%	-30,8%	-3,3
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	157,4%	14,6%	19,0%	-16,8%	2,1
HRP	78,5%	17,6%	17,9%	-32,5%	-2,7

É possível observar na Tabela 4 que conforme o valor de k aumenta, o retorno aumenta e o *drawdown* diminui. Também é possível observar que a volatilidade do portfólio é menor quando o valor de k é maior e o *turnover ratio* dos portfólios testados diminui conforme o valor de k aumenta. Quando comparado aos testes da Configuração 1, esse valor é menor.

Quando comparado os resultados dos testes com o índice Ibovespa, os retornos são maiores, a volatilidade é menor e o *drawdown* também é menor, sendo possível observar o fenômeno citado por Baker, Bradley e Wurgler (2011) e a alocação realizando o comportamento esperado conforme a heurística adotada comentada na Seção 3.3.

Entende-se que os portfólios com retorno maior e *drawdown* menor para os maiores

²Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

valores de k são resultado de uma menor alocação máxima em uma única ação. A alocação máxima que uma ação pode ter em um portfólio, em casos que o grupo é composto por apenas uma única ação, é $1/k$, sendo k o valor de grupos. Dessa forma, quanto maior o valor de k , menor a alocação máxima.

Os resultados observados nos testes com janela de 24 meses, na Tabela 4, apresentam também portfólios com *turnover* menor, retorno maior e *drawdown* menor conforme o valor de k aumenta.

4.2.3 Configuração 3: Uso da heurística de volatilidade para inicialização dos medoides, alocação nebulosa intragrupo e alocação intergrupo pelo método da roleta

A Configuração 3 de testes é muito parecido com a Configuração 2, com a diferença de alocação intergrupo. Enquanto na Configuração 2 a alocação era equilibrada em relação à quantidade de grupos, na Configuração 3 a alocação é inspirada nos algoritmos genéticos e a proporção de alocação em cada grupo segue a lógica do operador genético conhecido como seleção de roleta (Golberg, 1989). Nesse caso, é calculado o desempenho do grupo considerando que o capital foi alocado de forma equilibrada em todos os ativos daquele grupo no período anterior utilizado para a análise do agrupamento. Após o cálculo do desempenho dos grupos, o capital é alocado em cada grupo, onde o grupo que teve melhor desempenho recebe mais alocação, enquanto o que teve pior desempenho, recebe menos.

A Tabela 5 apresenta os resultados utilizando 12 meses de informação e preço para cálculo da correlação, enquanto a Tabela 6 apresenta os resultados utilizando 24 meses de correlação.

Tabela 5: Comparação das estratégias utilizando janela de correlação de 12 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	47,3%	21,5%	53,9%	-44,3%	-3,6
4	39,4%	21,0%	61,3%	-40,0%	-4,1
5	63,8%	20,1%	69,4%	-41,0%	-3,1
6	63,0%	19,7%	65,5%	-39,4%	-3,2
7	74,4%	19,4%	68,8%	-29,5%	-2,6
8	100,7%	18,6%	68,1%	-25,9%	-1,3
9	97,2%	18,2%	66,1%	-26,4%	-1,6
10	105,8%	18,0%	65,7%	-25,9%	-1,1
11	148,6%	17,6%	65,6%	-21,6%	1,3
12	171,7%	17,1%	55,1%	-23,5%	2,7
13	132,6%	17,1%	51,9%	-29,1%	0,4
14	143,2%	17,1%	50,5%	-27,8%	1,0
15	159,4%	17,2%	48,3%	-27,6%	2,0
16	154,7%	17,1%	48,6%	-27,0%	1,7
17	126,8%	17,1%	46,0%	-27,4%	0,1
18	120,6%	17,1%	44,7%	-26,7%	-0,3
19	114,6%	17,1%	44,2%	-27,3%	-0,6
20	90,8%	17,2%	43,1%	-27,6%	-2,0
12 ³	171,7%	17,1%	55,1%	-23,5%	2,7
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	151,9%	14,2%	31,4%	-16,9%	1,8
HRP	69,0%	17,1%	21,6%	-30,0%	-3,3

³Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

Tabela 6: Comparação das estratégias utilizando janela de correlação de 24 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	-18,4%	21,5%	47,5%	-63,4%	-6,7
4	22,2%	20,5%	55,9%	-51,3%	-5,1
5	32,2%	19,7%	57,9%	-34,6%	-4,7
6	67,9%	19,3%	57,7%	-26,4%	-3,0
7	50,2%	19,2%	61,0%	-30,0%	-3,9
8	56,6%	18,3%	59,1%	-25,1%	-3,8
9	64,6%	17,9%	53,0%	-27,4%	-3,4
10	55,6%	17,3%	47,9%	-30,6%	-4,1
11	70,2%	17,0%	42,9%	-30,6%	-3,2
12	64,3%	16,9%	39,5%	-26,6%	-3,6
13	82,2%	16,9%	38,1%	-22,3%	-2,6
14	73,4%	16,7%	35,7%	-23,7%	-3,1
15	75,7%	16,5%	35,0%	-23,6%	-3,0
16	85,2%	16,6%	31,5%	-22,0%	-2,4
17	72,3%	16,8%	30,7%	-25,9%	-3,2
18	67,4%	16,7%	28,9%	-26,5%	-3,5
19	74,1%	16,7%	28,3%	-27,5%	-3,1
20	82,3%	16,8%	28,6%	-26,8%	-2,6
16 ⁴	85,2%	16,6%	31,5%	-22,0%	-2,4
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	157,4%	14,6%	19,0%	-16,8%	2,1
HRP	78,5%	17,6%	17,9%	-32,5%	-2,7

Quando observados os resultados dos portfólios da Configuração 3 utilizando janela de correlação de 12 meses, comparado com a Configuração 2, é possível observar que a mudança da alocação intergrupo melhorou os resultados dos portfólios. Os portfólios produzidos possuem retorno maior, volatilidade, *turnover* e *drawdown* menor. Quando comparado com o índice Ibovespa, os resultados dos portfólios gerados são significativamente melhores do ponto de vista de retorno e de risco.

Os portfólios criados a partir da análise de 24 meses também apresentam resultados melhores quando comparados com a Configuração 2.

⁴Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

4.2.4 Configuração 4: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação equilibrada intergrupo

A Configuração 4 dos testes é parecido com a Configuração 2, ambas utilizam a heurística de seleção inicial de medoides a partir de ações com menor volatilidade e o método de alocação equilibrada intragrupos. Nessa configuração, os portfólios foram criados utilizando o método da roleta intergrupo e o capital é alocado em cada grupo utilizando a Equação 14.

Tabela 7: Comparação das estratégias utilizando janela de correlação de 12 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	57,0%	19,8%	48,3%	-33,8%	-3,5
4	73,5%	19,3%	55,2%	-27,2%	-2,7
5	55,6%	18,7%	58,8%	-33,1%	-3,7
6	78,0%	18,7%	56,4%	-29,8%	-2,6
7	83,0%	18,5%	60,1%	-27,1%	-2,3
8	103,4%	18,1%	57,5%	-23,6%	-1,2
9	108,4%	17,9%	54,8%	-26,3%	-1,0
10	127,0%	17,8%	56,1%	-27,0%	0,1
11	139,3%	17,4%	55,8%	-25,0%	0,8
12	150,4%	17,0%	48,2%	-26,6%	1,5
13	115,4%	17,1%	44,7%	-30,6%	-0,6
14	130,1%	17,2%	43,5%	-29,5%	0,3
15	153,4%	17,2%	41,9%	-27,6%	1,6
16	151,9%	17,1%	40,3%	-27,0%	1,5
17	142,2%	17,1%	38,1%	-25,6%	1,0
18	125,8%	17,2%	38,8%	-25,5%	0,0
19	121,1%	17,2%	37,4%	-25,8%	-0,3
20	105,6%	17,2%	37,2%	-27,5%	-1,2
15 ⁵	153,4%	17,2%	41,9%	-27,6%	1,6
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	151,9%	14,2%	31,4%	-16,9%	1,8
HRP	69,0%	17,1%	21,6%	-30,0%	-3,3

⁵Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

A Tabela 7 apresenta os resultados utilizando 12 meses de informação de preço para cálculo da correlação enquanto a Tabela 8 apresenta os resultados utilizando 24 meses de correlação.

Tabela 8: Comparação das estratégias utilizando janela de correlação de 24 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	21,3%	19,4%	44,5%	-42,1%	-5,4
4	37,4%	18,9%	46,5%	-37,7%	-4,7
5	54,8%	19,0%	50,4%	-28,7%	-3,7
6	68,5%	18,8%	50,0%	-30,0%	-3,0
7	49,2%	18,7%	52,6%	-31,4%	-4,1
8	44,2%	18,0%	50,9%	-33,2%	-4,5
9	61,1%	17,9%	46,9%	-32,7%	-3,6
10	64,4%	17,5%	42,9%	-33,7%	-3,5
11	59,5%	17,3%	38,6%	-36,3%	-3,8
12	55,4%	17,2%	37,9%	-28,5%	-4,1
13	60,4%	17,1%	35,8%	-26,9%	-3,8
14	65,9%	17,0%	33,4%	-27,8%	-3,5
15	75,7%	16,9%	32,0%	-26,4%	-3,0
16	69,9%	16,9%	30,3%	-28,9%	-3,3
17	60,8%	17,1%	29,2%	-33,8%	-3,8
18	59,4%	17,0%	27,8%	-33,3%	-3,9
19	65,4%	17,1%	25,9%	-31,6%	-3,5
20	68,4%	17,2%	26,3%	-31,7%	-3,3
15 ⁶	75,7%	16,9%	32,0%	-26,4%	-3,0
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	157,4%	14,6%	19,0%	-16,8%	2,1
HRP	78,5%	17,6%	17,9%	-32,5%	-2,7

Quando observamos os resultados dos portfólios da Configuração 4 utilizando janela de correlação de 12 meses comparado a Configuração 2, onde ambos utilizam alocação equilibrada intergrupo, é possível observar que os resultados da Configuração 4 são apenas um pouco melhores.

Os portfólios produzidos na Configuração 4 possuem *drawdown* menores, quando comparado a Configuração 2. Em relação ao índice Ibovespa, os resultados dos portfólios

⁶Valor de *k* que apresentou o melhor resultado nos testes para os parâmetros utilizados

gerados são significativamente melhores do ponto de vista de retorno e de risco.

4.2.5 Configuração 5: Uso da heurística de volatilidade para inicialização dos medoides, alocação intragrupo pelo método da roleta e alocação intergrupo também pelo método da roleta

A Configuração 5 dos testes é parecido com a Configuração 3, ambos utilizam a heurística de seleção inicial de medoides a partir de ações com menor volatilidade, também utilizam o método da roleta para alocar capital intragrupos e intergrupos. Neste caso, as ações de cada grupo que tiveram maior retorno na janela analisada, terão maior alocação no portfólio gerado.

A Tabela 9 apresenta os resultados dos portfólios calculados a partir de 12 meses de correlação, enquanto a Tabela 10 apresenta os resultados para portfólios calculados a partir de 24 meses de correlação.

Tabela 9: Comparação das estratégias utilizando janela de correlação de 12 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	68,3%	19,8%	49,8%	-31,2%	-2,9
4	72,5%	19,3%	55,4%	-26,0%	-2,8
5	77,6%	18,6%	60,7%	-31,5%	-2,6
6	83,6%	18,4%	57,2%	-27,9%	-2,3
7	86,3%	18,2%	60,7%	-23,4%	-2,2
8	94,7%	17,6%	58,5%	-22,3%	-1,8
9	103,3%	17,5%	55,3%	-24,8%	-1,3
10	122,3%	17,3%	55,8%	-24,7%	-0,2
11	152,0%	17,0%	55,7%	-21,5%	1,6
12	163,2%	16,6%	47,4%	-24,5%	2,3
13	132,5%	16,7%	44,9%	-28,2%	0,4
14	145,5%	16,8%	43,7%	-27,4%	1,2
15	170,4%	16,8%	42,1%	-25,5%	2,7
16	166,9%	16,8%	40,3%	-24,9%	2,5
17	151,0%	16,8%	37,9%	-24,2%	1,5
18	147,2%	16,8%	38,3%	-23,8%	1,3
19	141,4%	16,8%	36,8%	-23,8%	0,9
20	127,3%	16,8%	36,6%	-24,5%	0,1
15 ⁷	170,4%	16,8%	42,1%	-25,5%	2,7
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	151,9%	14,2%	31,4%	-16,9%	1,8
HRP	69,0%	17,1%	21,6%	-30,0%	-3,3

⁷Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

Tabela 10: Comparação das estratégias utilizando janela de correlação de 24 meses

k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
3	9,3%	19,4%	45,0%	-45,8%	-6,0
4	35,9%	18,6%	49,8%	-36,2%	-4,8
5	48,5%	18,4%	52,6%	-27,8%	-4,2
6	71,0%	18,1%	52,5%	-26,1%	-3,0
7	66,6%	18,1%	54,3%	-26,4%	-3,2
8	65,4%	17,4%	53,3%	-22,4%	-3,5
9	89,5%	17,2%	47,6%	-23,7%	-2,1
10	78,0%	16,8%	42,2%	-25,1%	-2,8
11	73,7%	16,6%	37,6%	-26,4%	-3,1
12	65,8%	16,4%	35,9%	-25,3%	-3,6
13	78,0%	16,4%	35,0%	-22,7%	-2,9
14	81,9%	16,3%	32,5%	-23,4%	-2,7
15	89,1%	16,2%	31,5%	-22,9%	-2,3
16	90,1%	16,2%	29,6%	-22,2%	-2,2
17	74,6%	16,4%	28,3%	-25,5%	-3,1
18	73,9%	16,4%	27,0%	-26,6%	-3,2
19	76,5%	16,5%	26,2%	-26,8%	-3,0
20	79,2%	16,5%	26,4%	-27,9%	-2,8
16 ⁸	90,1%	16,2%	29,6%	-22,2%	-2,2
Ibovespa	11,4%	22,7%	–	-48,6%	-5,0
MVP	157,4%	14,6%	19,0%	-16,8%	2,1
HRP	78,5%	17,6%	17,9%	-32,5%	-2,7

Os resultados encontrados nos portfólios da Configuração 5, para o testes realizados com janela de correlação de 12 e 24 meses, são melhores quando comparados as configurações anteriores. O retorno é maior, a volatilidade, o *turnover* e o *drawdown* são menores, para valores altos de k , apresentando melhor relação de risco e retorno, que pode ser observada com valores de *sharpe ratio* maiores.

As Tabelas 11 e 12 abaixo, apresentam os melhores resultados encontrados para as configurações de 2 a 5, as quais utilizam o mesmo método de inicialização de medoides. A Tabela 11 abaixo, apresenta os melhores resultados que utilizaram 12 meses de informação para o cálculo da correlação.

⁸Valor de k que apresentou o melhor resultado nos testes para os parâmetros utilizados

Tabela 11: Melhores resultados para cada configuração utilizando janela de correlação de 12 meses

Config	k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
Config. 2	12	150,4%	17,6%	55,2%	-27,2%	1,4
Config. 3	12	171,7%	17,1%	55,1%	-23,5%	2,7
Config. 4	15	153,4%	17,2%	41,9%	-27,6%	1,6
Config. 5	15	170,4%	16,8%	42,1%	-25,5%	2,7
Ibovespa		11,4%	22,7%	–	-48,6%	-5,0
MVP		151,9%	14,2%	31,4%	-16,9%	1,8
HRP		69,0%	17,1%	21,6%	-30,0%	-3,3

A Tabela 12 abaixo, apresenta os melhores resultados que utilizaram 24 meses de informação para o cálculo da correlação

Tabela 12: Melhores resultados para cada configuração utilizando janela de correlação de 24 meses

Config	k	Retorno	Volatilidade	Turnover	Drawdown	Sharpe
Config. 2	16	68,2%	17,4%	30,9%	-30,8%	-3,3
Config. 3	16	85,2%	16,6%	31,5%	-22,0%	-2,4
Config. 4	15	75,7%	16,9%	32,0%	-26,4%	-3,0
Config. 5	16	90,1%	16,2%	29,6%	-22,2%	-2,2
Ibovespa		11,4%	22,7%	–	-48,6%	-5,0
MVP		157,4%	14,6%	19,0%	-16,8%	2,1
HRP		78,5%	17,6%	17,9%	-32,5%	-2,7

4.2.6 Comparação das Carteiras dos Portfólios

Nas Configurações 2, 3, 4 e 5 é possível observar que conforme o valor de k aumenta, a volatilidade diminui de forma significativa, fazendo com que esses portfólios tenham um *drawdown* menor. Os portfólios testados com valores de k menores, tendem a ser mais concentrados, não sendo tão diversificados e ficando expostos a oscilações de poucas ações.

A partir da Configuração 2, a heurística de seleção dos medoides iniciais foi sempre a mesma, utilizando ações com menor volatilidade histórica para iniciar a busca dos grupos. Apesar das mudanças dos métodos de alocação intragrupo e intergrupo, que resultam em

portfólios com resultados diferentes nas Configurações 2, 3, 4 e 5, é possível concluir que a heurística de seleção inicial dos medoides causou impacto significativo no desempenho dos portfólios. Quando comparados os resultados dos portfólios da Configuração 1 com os portfólios da Configuração 2 é possível notar a mudança de resultado e do comportamento dos portfólios. Podemos concluir, dessa forma, que os medoides escolhidos para inicializar a busca do algoritmo são muito importantes para a execução da tarefa, influenciando o resultado final do agrupamento e da formação do portfólio.

Quando comparado o resultado do portfólio criado utilizando o modelo de Markowitz (1952) com o resultado dos portfólios em cada configuração, é possível observar para o MVP que o retorno é maior, mas a volatilidade e o *drawdown* são menores. Porém, ao analisar a distribuição de capital nos ativos do portfólio MVP em comparação com o método HRP e com os métodos das configurações de teste, é possível observar que o portfólio MVP concentra a maior parte da alocação em poucas ações, fazendo o portfólio ficar exposto a oscilações de algumas ações.

Na Figura 5 é possível observar a evolução mensal dos percentuais máximos em cada portfólio. Os portfólios das Configurações 2, 3, 4 e 5, utilizados para essa análise, foram criados considerando informações de 12 meses de correlação entre os ativos e 15 grupos como valor de k no agrupamento do k -medoides. Para os portfólios criados nas Configurações 2 e 4, foi utilizado o método da alocação equilibrada intergrupo. Dessa forma, o peso máximo que um ativo pode ter com essa alocação é $1/k$, nos casos em que o agrupamento com o k -medoides encontra grupos com apenas um ativo. Então, é possível ver no gráfico uma linha reta para essas duas configurações, pois em todos os meses que o agrupamento foi executado, foram encontrados grupos com um único ativo.

Nas Configurações 3 e 5 é possível observar que os percentuais máximos mudam a cada execução do agrupamento. Isso acontece por conta da alocação intergrupo baseada na roleta, onde os grupos com melhor desempenho observado no período analisado têm maior relevância na roleta, dessa forma tendo mais relevância na alocação final. Apesar de na alocação intragrupo as duas configurações utilizarem métricas diferentes, o peso máximo apresenta diferenças pequenas entre as duas configurações, pois a métrica utilizada para a alocação intergrupo foi a mesma.

Ao observar o percentual máximo dos portfólios obtidos a partir da otimização pro-

posta por Markowitz (MVP), é possível notar que os portfólios obtidos concentram um percentual muito alto em um único ativo. A média ao longo dos meses de percentual máximo para o MVP é de 22%. A concentração alta no ano de 2017 para o portfólio MVP é explicada pela característica natural do modelo, pois a otimização considera o risco e o retorno. Como nesse portfólio foi utilizada a volatilidade histórica, o modelo de otimização concentrou mais alocação em uma ação que apresentava volatilidade muito baixa por conta de uma dinâmica de fechamento de capital, isso é, saída da ação da bolsa de valores. A alocação HRP, por sua vez, considera a estrutura hierárquica e a volatilidade entre os nós. A mesma empresa por conta de sua volatilidade teve maior relevância na alocação do HRP.

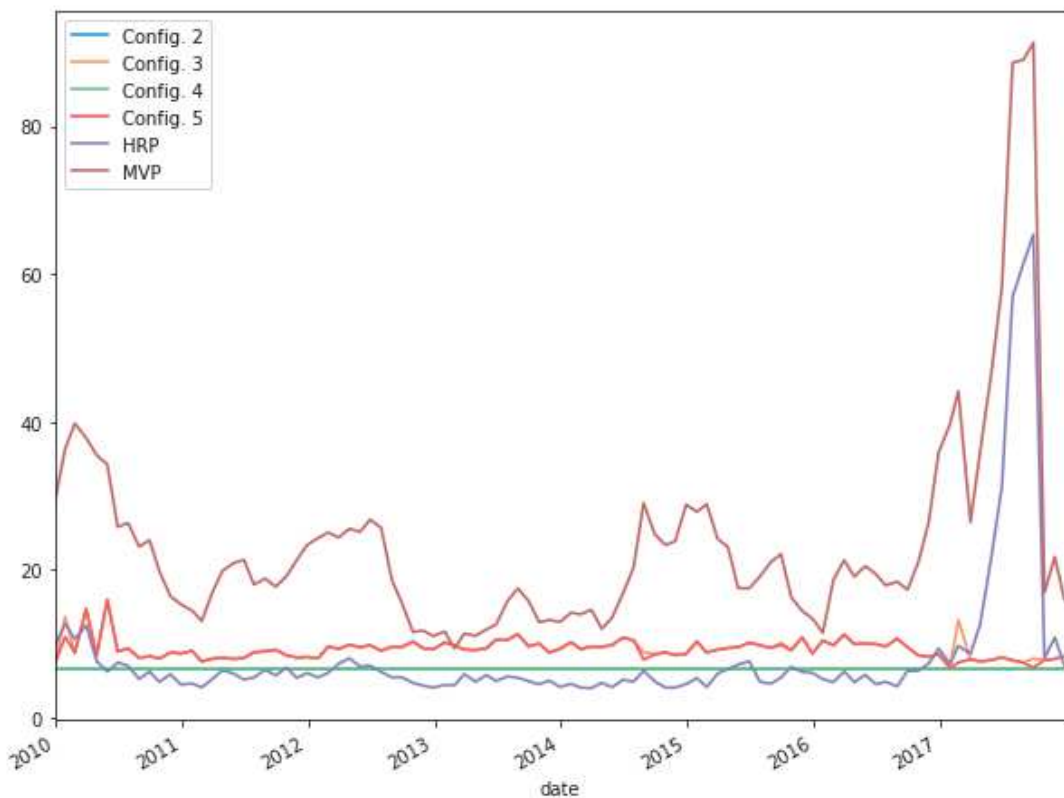


Figura 5: Evolução mensal dos percentuais máximos de cada portfólio.

A Tabela 13 apresenta a alocação de um mês específico para efeito de comparação entre os métodos. Os portfólios das Configurações 2, 3, 4 e 5 foram gerados utilizando 12 meses de correlação e 15 grupos, conforme o caso da Figura 5. A data dos portfólios é referente a agosto de 2016. No portfólio MVP, a soma do percentual das 5 maiores posições é de 67,3%, enquanto no HRP é 22%; nas Configurações 2, 3, 4 e 5 esses valores

são 33,3%, 42,2%, 33,3% e 42,3%, respectivamente.

Tabela 13: Carteira dos portfólios

Ticker	MVP	HRP	Config. 2	Config. 3	Config. 4	Config. 5
ABEV3	17.9%	4.2%	6.1%	5.3%	3.2%	2.7%
B3SA3	0.0%	1.1%	0.2%	0.2%	0.5%	0.5%
BBAS3	0.0%	0.4%	0.3%	0.3%	0.3%	0.3%
BBDC3	0.0%	0.8%	0.4%	0.4%	0.3%	0.3%
BBDC4	0.0%	0.6%	0.5%	0.5%	0.3%	0.3%
BBSE3	0.0%	0.6%	0.2%	0.2%	0.2%	0.3%
BRAP4	0.0%	0.6%	1.1%	1.1%	0.9%	0.9%
BRFS3	5.3%	3.8%	6.2%	3.9%	3.4%	2.1%
BRKM5	3.6%	3.2%	6.7%	8.6%	6.7%	8.6%
BRML3	0.0%	1.6%	0.2%	0.2%	0.4%	0.4%
CCRO3	0.0%	1.1%	0.2%	0.2%	0.3%	0.3%
CESP6	0.1%	1.4%	0.4%	0.4%	0.7%	0.7%
CIEL3	0.0%	1.5%	6.0%	5.7%	3.0%	2.8%
CMIG4	0.0%	0.5%	0.2%	0.2%	0.3%	0.3%
CPFE3	0.0%	1.2%	0.5%	0.5%	1.2%	1.1%
CPLE6	0.0%	0.8%	0.2%	0.2%	0.2%	0.2%
CSAN3	0.1%	1.7%	6.7%	9.7%	6.7%	9.7%
CSNA3	0.0%	0.2%	0.3%	0.3%	1.9%	2.0%
CTIP3	13.4%	4.7%	5.8%	6.1%	2.2%	2.3%
CYRE3	0.2%	1.3%	0.2%	0.2%	0.3%	0.3%
ECOR3	0.0%	0.7%	0.2%	0.2%	0.3%	0.3%
EGIE3	11.3%	3.3%	4.6%	4.5%	1.1%	1.1%
EMBR3	2.0%	1.7%	0.4%	0.2%	1.7%	0.9%
ENBR3	0.0%	1.6%	6.7%	6.5%	6.7%	6.5%
EQTL3	14.4%	4.9%	6.7%	9.0%	6.7%	9.0%
ESTC3	0.2%	1.0%	0.3%	0.3%	1.3%	1.2%
FIBR3	10.3%	3.1%	1.2%	0.6%	1.2%	0.6%
GGBR4	0.0%	0.4%	0.4%	0.4%	1.0%	1.1%
GOAU4	0.0%	0.3%	0.3%	0.3%	0.6%	0.6%
HYPE3	5.0%	4.5%	0.5%	0.5%	2.4%	2.5%
ITSA4	0.0%	1.1%	0.7%	0.7%	0.3%	0.3%
ITUB4	0.0%	0.9%	1.9%	2.0%	0.3%	0.3%
JBSS3	0.4%	1.4%	0.5%	0.3%	3.3%	2.0%
KLBN11	0.0%	2.3%	0.4%	0.2%	2.2%	1.1%
KROT3	0.0%	1.1%	0.2%	0.2%	0.4%	0.4%
LAME4	0.0%	1.1%	0.2%	0.2%	0.3%	0.4%
LREN3	0.0%	1.2%	0.6%	0.6%	3.7%	3.5%
MRFG3	4.9%	3.3%	0.4%	0.3%	1.0%	1.0%
MRVE3	0.0%	2.4%	0.2%	0.2%	0.5%	0.5%
MULT3	1.8%	1.6%	0.2%	0.2%	0.3%	0.4%
NATU3	0.0%	1.5%	0.3%	0.3%	1.3%	1.2%

Tabela 13: Carteira dos portfólios

Ticker	MVP	HRP	Config. 2	Config. 3	Config. 4	Config. 5
PCAR4	0.0%	1.6%	0.3%	0.3%	0.7%	0.6%
PETR3	0.0%	0.3%	1.7%	1.6%	1.2%	1.1%
PETR4	0.0%	0.3%	3.6%	3.3%	1.1%	1.0%
QUAL3	1.7%	2.6%	0.4%	0.5%	2.1%	2.2%
RADL3	3.4%	2.9%	6.7%	8.5%	6.7%	8.5%
RENT3	0.0%	1.7%	0.2%	0.2%	0.4%	0.4%
RUMO3	0.0%	0.7%	0.1%	0.1%	0.2%	0.2%
SANB11	0.0%	1.0%	0.3%	0.3%	0.4%	0.4%
SBSP3	0.0%	1.5%	0.4%	0.4%	1.7%	1.6%
SMLS3	0.0%	1.2%	0.4%	0.4%	1.0%	0.9%
SUZB5	3.3%	3.6%	4.6%	2.4%	1.6%	0.8%
TIMP3	0.0%	1.6%	0.2%	0.2%	0.7%	0.7%
UGPA3	0.0%	3.3%	0.5%	0.4%	3.5%	3.0%
USIM5	0.0%	0.2%	0.4%	0.3%	1.0%	0.9%
VALE3	0.0%	0.3%	1.6%	1.6%	0.8%	0.8%
VALE5	0.0%	0.4%	2.8%	2.9%	0.8%	0.8%
VIVT4	0.2%	3.3%	0.2%	0.2%	0.3%	0.3%
WEGE3	0.1%	2.9%	6.7%	4.6%	6.7%	4.6%

É possível observar na Tabela 13, que no portfólio MVP muitas ações não são alocadas, ou seja, possuem 0,0% de peso na alocação final do portfólio. Quando comparado com os outros portfólios, esse portfólio é menos diversificado. Apesar de no final do período de teste analisado o retorno do portfólio MVP ser satisfatório, a concentração em poucas ações, por causa da otimização, pode levar o portfólio a ter grandes perdas caso as ações com maiores relevâncias tenham perdas acentuadas de valores.

5 Conclusão e Trabalhos Futuros

A partir da aplicação de agrupamento hierárquico para seleção de portfólios e da natureza de otimização deste problema, essa dissertação propôs o uso de agrupamento particional organizando o processo em um *framework* para seleção de portfólios. O *framework* está dividido em sete fases sequenciais, sendo que as três primeiras implicam em selecionar os ativos e preparar a base para o agrupamento, encontrando a matriz de distâncias a partir da matriz de correlação entre o valor dos ativos. Feito isso, é executado o agrupamento utilizando o algoritmo k -medoides com diferentes métodos de inicialização dos protótipos. A partir do resultado do agrupamento é realizada a alocação intragrupo e intergrupo e os portfólios gerados são avaliados usando medidas específicas.

Neste trabalho, apresentamos o método tradicional e a teoria de seleção de portfólios proposta por Markowitz (1952) e também os avanços mais recentes na área utilizando algoritmos de aprendizado de máquina para realizar a tarefa de alocação de portfólios, mais especificamente, a contribuição de López de Prado (2016) ao utilizar agrupamento hierárquico para realizar seleção de portfólios.

O *framework* resultante é flexível em seus processos, podendo ficar a cargo do pesquisador ou usuário a decisão de utilizar o mesmo algoritmo de agrupamento e os mesmos métodos de alocação intragrupos e intergrupos aqui sugeridos, ou então fazer alterações nessas etapas do *framework* utilizando outros algoritmos de agrupamento particional ou outras técnicas e métricas nas alocações intragrupos e intergrupos.

Na metodologia experimental, foi sugerido um processo sistemático de testes do *framework* com diferentes configurações. O algoritmo de agrupamento utilizado (k -medoides) é sensível a seleção dos protótipos iniciais e propusemos um método que considera a volatilidade dos ativos para inicializar os protótipos. Outra questão pertinente ao uso dos algoritmos de agrupamento particional, é a necessidade de informar ao algoritmo a quantidade de grupos a serem encontrados na base de dados.

Para entender como os medoides iniciais afetavam o agrupamento e a alocação, foi testada uma configuração com a inicialização aleatória dos medoides e outra configuração com uma heurística de seleção dos medoides iniciais. A heurística sugerida foi utilizar ações com baixa volatilidade histórica como sendo os medoides iniciais para a busca dos

grupos. Com os resultados encontrados para as duas configurações, foi possível concluir como a seleção dos medoides iniciais influencia no portfólio final obtido.

A implementação do algoritmo k -medoides utilizado neste trabalho, conhecido como **particionamento em torno dos medoides** (Kaufman; Rousseeuw, 1987), tende a encontrar grupos ao redor dos medoides utilizados na inicialização da busca. Com a heurística proposta de utilizar ações de baixa volatilidade histórica, entende-se que os grupos encontrados terão as próprias ações utilizadas na inicialização como medoides iniciais ou ações que tenham comportamento muito parecido com elas, já que o objetivo do algoritmo de agrupamento particional é minimizar a distância intragrupo utilizando a matriz de distância encontrada a partir da matriz de correlação.

Para a alocação intragrupo, foram propostos e utilizados os métodos de alocação nebuloso e da roleta. O método de alocação nebuloso foi inspirado nos agrupamentos particionais *fuzzy* (nebuloso), no qual é levado em conta o grau de pertinência dos objetos a todos os grupos (de Castro; Ferrari, 2016) e a alocação utilizando o método da roleta foi inspirada no mecanismo de seleção dos algoritmos genéticos (Golberg, 1989). Para a alocação intergrupo foram utilizados os métodos de alocação equilibrada e da roleta. O método de alocação equilibrada divide em partes iguais o capital entre os grupos (DeMiguel; Garlappi; Uppal, 2007), e o método da roleta possui o mesmo funcionamento que o utilizado na alocação intragrupo.

Na avaliação de desempenho, os testes do *framework* foram realizados utilizando sempre informações realizadas, tanto para cálculo da correlação como da volatilidade, e os testes foram realizados executando rebalanceamento mensal, isto é, a cada mês do teste o *framework* era executado e o portfólio alterado. A mesma configuração de teste foi utilizada para os portfólios criados utilizando os modelos de Markowitz (1952) e também para a alocação hierárquica de López de Prado (2016), que foram utilizados como métricas de comparação com o *framework*.

Quando comparados os resultados das configurações de teste do *framework*, é possível observar o efeito que a seleção dos medoides iniciais causa no portfólio final e, por consequência, no desempenho e no comportamento do portfólio. Quando comparados os resultados obtidos na Configurações 1 dos testes, com as outras configurações, é possível observar a diferença de desempenho dos portfólio. As configurações nas quais a heurística

de inicialização foi utilizada apresentaram resultados superiores quando comparadas a configuração onde a inicialização dos medoides foi aleatória. Baker, Bradley e Wurgler (2011) abordam sobre portfólios criados com ações que apresentam baixa volatilidade e mostram os resultados desses portfólios, que apresentam resultados satisfatórios no longo prazo, baixa volatilidade e *drawdowns* menores. Entende-se que portfólios encontrados a partir da busca se beneficiem desse efeito.

Dentre os portfólios inicializados com a heurística, é possível observar que os métodos de alocação intragrupos e intergrupos contribuem para a melhora do desempenho das alocações. Dentre as combinações dos métodos, a alocação que apresentou o melhor resultado, foi a que utilizou a heurística de seleção dos medoides iniciais a partir das ações com menor volatilidade e que utilizou o método da roleta para a alocação intragrupo e intergrupo. No método da roleta, o indicador utilizado para dar a maior relevância as ações, no caso da alocação intragrupo, e maior relevância aos grupos, no caso da alocação intergrupo, foi o desempenho passado da ação, comparado com o respectivo grupo, ou do desempenho agregado das ações do grupo, comparado com os outros grupos. As ações e os grupos que tiveram melhor desempenho no período comparado, tiveram maior peso na alocação final, por causa da dinâmica de funcionamento da roleta. Jegadeesh e Titman (1993) comentam sobre estratégias de investimentos, formadas a partir de ações que tiveram boa performance no passado, e que essas estratégias apresentam bom desempenho futuro. A formação do portfólio utilizando o método da roleta, com a métrica do desempenho passado, pode estar se beneficiando desse efeito.

Quando comparados os resultados das Configurações 2 a 5 com o índice Ibovespa, é possível observar a superioridade do método de alocação no período, tendo o *framework* criado portfólios que tiveram desempenhos melhores e volatilidades menores. Quando comparado com o portfólio otimizado com o método de Markowitz (1952), a maioria dos portfólios criados utilizando o *framework*, não apresentam desempenho superior, porém, conforme discutido na Seção 4.2.6, os portfólios otimizados com o método de Markowitz não são tão diversificados como os portfólios criados pelo *framework*. É possível observar, conforme apresentado na Figura 5, que a alocação máxima em uma única ação utilizando o método de Markowitz é muito maior quando comparada às alocações do *framework*. As alocações concentradas podem levar o portfólio a ter grandes perdas caso as ações apresentem algum revés.

Para trabalhos futuros, é possível explorar o uso de métricas de avaliação de algoritmos de aprendizado de máquina, tanto na avaliação da qualidade dos grupos criados em relação ao desempenho do portfólio consequente, como na aplicação das métricas para realizar a alocação intragrupo e intergrupos. A partir dos trabalhos correlatos, é possível criar uma revisão sobre o estado da arte da pesquisa de alocação de portfólios utilizando algoritmos de agrupamento. Outra possibilidade de trabalho futuro, é a aplicação do *framework* proposto por Marti et al. (2016) para analisar qual o período mais adequado da série temporal para a aplicação do agrupamento particional. A aplicação do *framework* proposto por Marti et al. (2016), possibilita e facilita testar outras configurações do *framework* de alocação, e a partir desses testes, é possível comparar os resultados obtidos com outras alocações realizadas utilizando algoritmos de alocação que representam o estado da arte de alocação de portfólios.

Por conta da flexibilidade que o *framework* apresenta, também é possível testar outros fatores de investimento, incorporando os indicadores nas alocações intragrupos e intergrupos. Dentre os fatores mais conhecidos e que não foram utilizados como métrica de inicialização do algoritmo, estão os de Valor, Crescimento e Qualidade. O fator de Valor (*Value*) é caracterizado pelo portfólio criado utilizando ações subvalorizadas (Fama; French, 1998), o fator de Crescimento (*Growth*) é caracterizado pelo portfólio criado utilizando ações que apresentam potencial para crescimento das receitas (Fama; French, 1998), e o fator de Qualidade (*Quality*) é caracterizado pelo portfólio criado com ações que, comparadas com as outras, os investidores aceitam pagar preços maiores por elas (Asness; Frazzini; Pedersen, 2019).

Referências Bibliográficas

- Asness, C. S.; Frazzini, A.; Pedersen, L. H. Quality minus junk. *Review of Accounting Studies*, Springer, v. 24, n. 1, p. 34–112, 2019.
- B3. *Metodologia do Índice Ibovespa*. 2015. Disponível em: <<http://www.b3.com.br/data/files/1C/56/F7/D5/96E615107623A41592D828A8/IBOV-Metodologia-pt-br.pdf>>.
- Baker, M.; Bradley, B.; Wurgler, J. Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, Taylor & Francis, v. 67, n. 1, p. 40–54, 2011.
- Bezdek, J. C. Objective function clustering. In: *Pattern recognition with fuzzy objective function algorithms*. [S.l.]: Springer, 1981. p. 43–93.
- Black, F.; Litterman, R. Global portfolio optimization. *Financial analysts journal*, CFA Institute, v. 48, n. 5, p. 28–43, 1992.
- Black, F.; Scholes, M. The pricing of options and corporate liabilities. *Journal of political economy*, The University of Chicago Press, v. 81, n. 3, p. 637–654, 1973.
- Clarke, R.; De Silva, H.; Thorley, S. Risk parity, maximum diversification, and minimum variance: An analytic perspective. *The Journal of Portfolio Management*, Institutional Investor Journals Umbrella, v. 39, n. 3, p. 39–53, 2013.
- de Castro, L. N.; Ferrari, D. G. *Introdução à Mineração de Dados: conceitos básicos, algoritmos e aplicações*. [S.l.: s.n.], 2016. 116–119 p.
- DeMiguel, V.; Garlappi, L.; Uppal, R. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, Oxford University Press, v. 22, n. 5, p. 1915–1953, 2007.
- Dose, C.; Cincotti, S. Clustering of financial time series with application to index and enhanced index tracking portfolio. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 355, n. 1, p. 145–151, 2005.
- Fabozzi, F. J.; Gupta, F.; Markowitz, H. M. The legacy of modern portfolio theory. *Journal of Investing*, INSTITUTIONAL INVESTOR INC., v. 11, n. 3, p. 7–22, 2002.

Fama, E. F.; French, K. R. Value versus growth: The international evidence. *The journal of finance*, Wiley Online Library, v. 53, n. 6, p. 1975–1999, 1998.

Fama, E. F.; French, K. R. O modelo de precificação de ativos de capital: teoria e evidências. *Revista de Administração de Empresas*, SciELO Brasil, v. 47, n. 2, p. 103–118, 2007.

Golberg, D. E. Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, v. 1989, n. 102, p. 36, 1989.

Grötschel, M.; Wakabayashi, Y. A cutting plane algorithm for a clustering problem. *Mathematical Programming*, Springer, v. 45, n. 1-3, p. 59–96, 1989.

IFRS Foundation. *International Financial Reporting Standard 9 Financial Instruments (IFRS 9)*. 2009. Disponível em: <<http://eifrs.ifrs.org/eifrs/bnstandards/en/IFRS9.pdf>>.

IFRS Foundation. *Conceptual Framework for Financial Reporting*. 2010. Disponível em: <<http://eifrs.ifrs.org/eifrs/bnstandards/en/framework.pdf>>.

Institutional Investor. *The Hedge Fund 100: Quant Funds Excel — at Asset Gathering*. 2018. Disponível em: <<https://www.institutionalinvestor.com/article/b19dxvr2pvgjq2-/The-Hedge-Fund-100-Quant-Funds-Excel-at-Asset-Gathering>>.

Jegadeesh, N.; Titman, S. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, Wiley Online Library, v. 48, n. 1, p. 65–91, 1993.

Jin, X.; Han, J. K-medoids clustering. In: _____. *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010. p. 564–565. ISBN 978-0-387-30164-8. Disponível em: <https://doi.org/10.1007/978-0-387-30164-8_426>.

Kaufman, L.; Rousseeuw, P. *Clustering by means of medoids*. [S.l.]: North-Holland, 1987.

Kolm, P. N.; Tütüncü, R.; Fabozzi, F. J. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, Elsevier, v. 234, n. 2, p. 356–371, 2014.

Krishnapuram, R.; Joshi, A.; Yi, L. A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering. In: IEEE. *Fuzzy Systems*

Conference Proceedings, 1999. FUZZ-IEEE'99. 1999 IEEE International. [S.l.], 1999. v. 3, p. 1281–1286.

Lipowski, A.; Lipowska, D. Roulette-wheel selection via stochastic acceptance. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 391, n. 6, p. 2193–2196, 2012.

López de Prado, M. Building diversified portfolios that outperform out-of-sample. 2016.

Magalhães de Pinho, F.; de Camargos, M. A.; Marques Figueiredo, J. Uma revisão da literatura sobre modelos de volatilidade em estudos brasileiros. *Revista de Administração FACES Journal*, Universidade FUMEC, v. 16, n. 1, 2017.

Magdon-Ismail, M. et al. On the maximum drawdown of a brownian motion. *Journal of applied probability*, Cambridge University Press, v. 41, n. 1, p. 147–161, 2004.

Mansini, R.; Ogryczak, W.; Speranza, M. G. Twenty years of linear programming based portfolio optimization. *European Journal of Operational Research*, Elsevier, v. 234, n. 2, p. 518–535, 2014.

Mantegna, R. N. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 11, n. 1, p. 193–197, 1999.

Markowitz, H. Portfolio selection. *The journal of finance*, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.

Markowitz, H. The optimization of a quadratic function subject to linear constraints. *Naval research logistics Quarterly*, Wiley Online Library, v. 3, n. 1-2, p. 111–133, 1956.

Marti, G. et al. Clustering financial time series: How long is enough? *arXiv preprint arXiv:1603.04017*, 2016.

Marti, G. et al. A review of two decades of correlations, hierarchies, networks and clustering in financial markets. *arXiv preprint arXiv:1703.00485*, 2017.

Michaud, R. O. The markowitz optimization enigma: Is ‘optimized’ optimal? *Financial Analysts Journal*, CFA Institute, v. 45, n. 1, p. 31–42, 1989.

- Nanda, S.; Mahanty, B.; Tiwari, M. Clustering indian stock market data for portfolio management. *Expert Systems with Applications*, Elsevier, v. 37, n. 12, p. 8793–8798, 2010.
- Ponsich, A.; Jaimes, A. L.; Coello, C. A. C. A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 17, n. 3, p. 321–344, 2013.
- Raffinot, T. Hierarchical clustering based asset allocation. *Available at SSRN 2840729*, 2017.
- Reynolds, A. P.; Richards, G.; Rayward-Smith, V. J. The application of k-medoids and pam to the clustering of rules. In: SPRINGER. *International Conference on Intelligent Data Engineering and Automated Learning*. [S.l.], 2004. p. 173–178.
- Rubinstein, M. Markowitz’s “portfolio selection”: A fifty-year retrospective. *The Journal of finance*, Wiley Online Library, v. 57, n. 3, p. 1041–1045, 2002.
- Sharpe, W. F. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, Wiley Online Library, v. 19, n. 3, p. 425–442, 1964.
- Sharpe, W. F. Mutual fund performance. *The Journal of business*, JSTOR, v. 39, n. 1, p. 119–138, 1966.
- Tola, V. et al. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, Elsevier, v. 32, n. 1, p. 235–258, 2008.
- Tsay, R. S. *Analysis of financial time series*. [S.l.]: John wiley & sons, 2005.

Apêndices

A Carteira teórica do índice Ibovespa

Data	Composição da Carteira
(2009-10-01, 2010-01-01]	ABEV3, ALLL11, ARCZ6, B3SA3, BBAS3, BBDC4, BRAP4, BRFS3, BRKM5, BRTP4, BTOW3, CCRO3, CESP6, CGAS5, CLSC4, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DURA4, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, NATU3, NETC4, OIBR4, PCAR4, PETR3, PETR4, RDCD3, RSID3, SBSP3, TAMM4, TCSL4, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4
(2010-01-01, 2010-04-01]	ABEV3, ALLL11, B3SA3, BBAS3, BBDC4, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRVE3, NATU3, NETC4, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SBSP3, TAMM4, TCSL4, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4
(2010-04-01, 2010-07-01]	ABEV3, ALLL11, B3SA3, BBAS3, BBDC4, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRVE3, NATU3, NETC4, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SBSP3, TAMM4, TCSL4, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4

Data	Composição da Carteira
(2010-07-01, 2010-10-01]	ABEV3, ALLL11, B3SA3, BBAS3, BBDC4, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRVE3, NATU3, NETC4, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SBSP3, TAMM4, TCSSL4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4
(2010-10-01, 2011-01-01]	ABEV3, ALLL11, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, NETC4, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SANB11, SBSP3, TAMM4, TCSSL4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4
(2011-01-01, 2011-04-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, PRTX3, RDCD3, RSID3, SANB11, SBSP3, TAMM4, TCSSL4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4

Data	Composição da Carteira
(2011-04-01, 2011-07-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, PRTX3, RDCD3, RSID3, SANB11, SBSP3, TAMM4, TCSSL4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVO4, VIVT4
(2011-07-01, 2011-10-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SANB11, SBSP3, TAMM4, TCSSL4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA4, USIM3, USIM5, VALE3, VALE5, VIVT4
(2011-10-01, 2012-01-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RSID3, SANB11, SBSP3, TAMM4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2012-01-01, 2012-04-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RENT3, RSID3, SANB11, SBSP3, TAMM4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2012-04-01, 2012-07-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RENT3, RSID3, SANB11, SBSP3, TAMM4, TESA3, TIMP3, TMAR5, TNLP3, TNLP4, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2012-07-01, 2012-10-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RDCD3, RENT3, RSID3, SANB11, SBSP3, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2012-10-01, 2013-01-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2013-01-01, 2013-04-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2013-04-01, 2013-07-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2013-07-01, 2013-10-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2013-10-01, 2014-01-01]	ABEV3, AEDU3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, KROT3, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OGXP3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4
(2014-01-01, 2014-04-01]	ABEV3, AEDU3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, BTOW3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN4, KROT3, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OIBR3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TESA3, TIMP3, TRPL4, UGPA3, USIM3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2014-04-01, 2014-07-01]	ABEV3, AEDU3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DASA3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, ESTC3, EVEN3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KLBN4, KROT3, LAME4, LIGT3, LREN3, MRFG3, MRVE3, NATU3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, PRML3, QUAL3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4
(2014-07-01, 2014-10-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BISA3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, ESTC3, EVEN3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LIGT3, LREN3, MMXM3, MRFG3, MRVE3, NATU3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, QUAL3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4
(2014-10-01, 2015-01-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, ESTC3, EVEN3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LIGT3, LREN3, MRFG3, MRVE3, NATU3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, POMO4, QUAL3, RENT3, RSID3, SANB11, SBSP3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2015-01-01, 2015-04-01]	ABEV3, ALLL3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, ELPL4, EMBR3, ENBR3, ESTC3, EVEN3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LIGT3, LREN3, MRFG3, MRVE3, NATU3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, POMO4, QUAL3, RENT3, RLOG3, RSID3, SANB11, SBSP3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4
(2015-04-01, 2015-07-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, EMBR3, ENBR3, ESTC3, EVEN3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LIGT3, LREN3, MRFG3, MRVE3, MULT3, NATU3, OIBR4, PCAR4, PDGR3, PETR3, PETR4, POMO4, QUAL3, RENT3, RUMO3, SANB11, SBSP3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4
(2015-07-01, 2015-10-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, DTEX3, ECOR3, EGIE3, ELET3, ELET6, EMBR3, ENBR3, ESTC3, FIBR3, GFSA3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, OIBR4, PCAR4, PETR3, PETR4, POMO4, QUAL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4

Data	Composição da Carteira
(2015-10-01, 2016-01-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, BRPR3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CRUZ3, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, ELET3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, GOLL4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, OIBR4, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4
(2016-01-01, 2016-04-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, OIBR3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3
(2016-04-01, 2016-07-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HGTX3, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, OIBR3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3

Data	Composição da Carteira
(2016-07-01, 2016-10-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CESP6, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3
(2016-10-01, 2017-01-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3
(2017-01-01, 2017-04-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CTIP3, CYRE3, ECOR3, EGIE3, ELET3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RENT3, RUMO3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3

Data	Composição da Carteira
(2017-04-01, 2017-07-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CYRE3, ECOR3, EGIE3, ELET3, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RAIL3, RENT3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3
(2017-07-01, 2017-10-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CYRE3, ECOR3, EGIE3, ELET3, ELET6, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RAIL3, RENT3, SANB11, SBSP3, SMLS3, SUZB5, TIMP3, UGPA3, USIM5, VALE3, VALE5, VIVT4, WEGE3
(2017-10-01, 2018-01-01]	ABEV3, B3SA3, BBAS3, BBDC3, BBDC4, BBSE3, BRAP4, BRFS3, BRKM5, BRML3, CCRO3, CIEL3, CMIG4, CPFE3, CPLE6, CSAN3, CSNA3, CYRE3, ECOR3, EGIE3, ELET3, ELET6, EMBR3, ENBR3, EQTL3, ESTC3, FIBR3, GGBR4, GOAU4, HYPE3, ITSA4, ITUB4, JBSS3, KLBN11, KROT3, LAME4, LREN3, MRFG3, MRVE3, MULT3, NATU3, PCAR4, PETR3, PETR4, QUAL3, RADL3, RAIL3, RENT3, SANB11, SBSP3, SMLS3, SUZB5, TAEE11, TIMP3, UGPA3, USIM5, VALE3, VIVT4, WEGE3