

**Universidade Federal de Pernambuco  
Centro de Ciências Sociais Aplicadas  
Departamento de Ciências Administrativas  
Programa de Pós-Graduação em Administração – PROPAD**

**José Márcio Arcanjo dos Santos**

**Otimização dos Retornos dos Ativos e a Técnica de  
*Stacked Generalization*: uma aplicação na B3**

**Recife, 2024**

UNIVERSIDADE FEDERAL DE PERNAMBUCO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

CLASSIFICAÇÃO DE ACESSO A TESES E DISSERTAÇÕES

Considerando a natureza das informações e compromissos assumidos com suas fontes, o acesso a monografia do Programa de Pós-Graduação em Administração da Universidade Federal de Pernambuco é definido em três graus:

- “Grau 1”: livre (sem prejuízo das referências ordinárias em citações diretas e indiretas);
- “Grau 2”: com vedação a cópias, no todo ou em parte, sendo, em consequência, restrita a consulta em ambientes de biblioteca com saída controlada;
- “Grau 3”: apenas com autorização expressa do autor, por escrito, devendo, por isso, o texto se confiado a bibliotecas que assegurem a restrição, ser mantido em local sob chave ou custódia.

A classificação desta dissertação se encontra, abaixo, definida por seu autor.

Solicita-se aos depositários e usuários sua fiel observância, a fim de que preservem as condições éticas e operacionais da pesquisa científica na área da administração.

---

Título da Dissertação: Otimização dos Retornos dos Ativos e a Técnica de Stacked Generalization: uma aplicação na B3

Nome do Autor: José Márcio Arcanjo dos Santos

Data da Aprovação: 29 de maio de 2024.

Classificação, conforme especificação acima:

Grau 1

Grau 2

Grau 3

Recife, 29 de maio de 2024.

---

Assinatura do autor

# **Otimização dos Retornos dos Ativos e a Técnica de *Stacked Generalization*: uma aplicação na B3**

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Joséte Florencio dos Santos

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Administração, na área de concentração Gestão Organizacional, com ênfase em Finanças, do Programa de Pós-graduação em Administração da Universidade Federal de Pernambuco.

Recife, 2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Santos, José Márcio Arcanjo dos.

Otimização dos Retornos dos Ativos e a Técnica de Stacked Generalization: uma aplicação na B3 / José Márcio Arcanjo Dos Santos. - Recife, 2024.  
154f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Administração, 2024.

Orientação: Joséte Florencio dos Santos.

Inclui referências e apêndices.

1. Otimização de retornos; 2. Precificação de ativos; 3. Stacked Generalization; 4. Machine Learning. I. Santos, Joséte Florencio dos. II. Título.

UFPE-Biblioteca Central

José Márcio Arcanjo dos Santos

## **Otimização dos Retornos dos Ativos e a Técnica de *Stacked Generalization*: uma aplicação na B3**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Administração, na área de concentração Gestão Organizacional, com ênfase em Finanças, do Programa de Pós-graduação em Administração da Universidade Federal de Pernambuco.

Apresentado em: 29/05/2024

### **BANCA EXAMINADORA**

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Joséte Florencio dos Santos (Orientadora)  
Universidade Federal de Pernambuco

---

Prof. Dr. Kliver Lamarthine Alves Confessor (Examinador Externo)  
Universidade Federal do Vale do São Francisco

---

Prof. Dr. Moisés Araújo Almeida (Examinador Externo)  
Universidade Federal de Paraíba

*Dedico este trabalho à minha querida mãe, Maria Vitória da Silva,  
falecida durante a minha jornada no Mestrado.  
Independentemente de onde estejas,  
sei que estás feliz por mim.*

## **Agradecimentos**

A minha mãe, Maria Vitória (*in memoriam*), que sempre lutou para me proporcionar acesso a uma educação que ela nunca teve.

A minha esposa e parceira, Barbara Arcanjo, pela parceria na vida, no trabalho, no amor e nos estudos.

A minha orientadora, Prof.<sup>a</sup> Dr.<sup>a</sup> Joséte Florencio, pelas enormes orientações, apoio e paciência durante o desenvolvimento do trabalho.

Aos professores e professoras do PROPAD, que durante todo o período do curso compartilharam seus conhecimentos com muito prazer e humildade intelectual.

Aos meus queridos(as) amigos e amigas, que me apoiaram nos momentos difíceis, compreenderam os longos períodos de ausência e nunca desistiram de mim.

Aos meus colegas de turma, que mesmo durante o isolamento social gerado pela pandemia de Covid-19 proporcionaram momentos inesquecíveis de aprendizado e convivência.

*“Quem tem por que viver pode suportar quase qualquer coisa.”*

(Friedrich Nietzsche)

## Resumo

Esta pesquisa tem como objetivo investigar se os retornos das ações dos modelos de precificação de ativos são potencializados quando utilizada a técnica de *stacked generalization* em *machine learning* nos ativos da B3. A amostra foi composta por 202 empresas com informações disponíveis no período. Os dados foram coletados por meio de fontes secundárias, com informações provenientes de fontes públicas, como o Banco Central do Brasil (BACEN), a Bolsa de Valores brasileira (B3) e o Núcleo de Pesquisa em Economia Financeira da Universidade de São Paulo (NEFIN). Para aplicação da técnica de *stacked generalization* foram considerados cinco modelos de precificação de ativos recorrentes na literatura: *Capital Asset Pricing Model* (CAPM), *Learning CAPM*, *Downside CAPM*, *Conditional CAPM* e a Regressão *Cross-section* de Fama e French (1996), e quatro algoritmos de *machine learning*: *Random Forest* (RF), *Gradient Boosted Regression Trees* (GBRT), *Neural Networks* (NN) e *Linear Regression* (LR). Para atingir os objetivos da pesquisa, a amostra foi dividida em três conjuntos de dados em um esquema de janelas de expansão, na qual os dados são segmentados em partes, seguindo um formato de estimativas recursivas. A eficácia das previsões foi avaliada por meio dos ganhos estatísticos – mensurados pelo  $R^2$  fora da amostra de Campbell e Thompson (2008), e medidas de previsões de sinal, como acurácia, sensibilidade e especificidade. Os resultados evidenciam que a aplicação da técnica de *stacked generalization* para combinar modelos de precificação de ativos pode melhorar significativamente os resultados das previsões em diferentes contextos e condições de mercado. Além disso, treinar os meta-modelos de *stacked generalization* com a seleção dos melhores modelos base pode melhorar a previsão de acurácia negativa, especialmente em períodos de queda. Todavia, as evidências indicam que os modelos estudados foram incapazes de superar a média histórica dos retornos das ações de forma significativa durante um choque exógeno.

**Palavras-chave:** otimização de retornos, precificação de ativos, *stacked generalization*, *machine learning*.

## Abstract

This research aims to investigate whether stock returns from asset pricing models are enhanced when the stacked generalization technique is used in machine learning on B3 assets. The sample consisted of 202 companies with information available during the period. The data were collected through secondary sources, with information from public sources, such as the Central Bank of Brazil (BACEN), the Brazilian Stock Exchange (B3) and the Center for Research in Financial Economics of the University of São Paulo (NEFIN). To apply the stacked generalization technique, five asset pricing models that are common in the literature were considered: Capital Asset Pricing Model (CAPM), Learning CAPM, Downside CAPM, Conditional CAPM and the Cross-section Regression of Fama and French (1996), and four machine learning algorithms: Random Forest (RF), Gradient Boosted Regression Trees (GBRT), Neural Networks (NN) and Linear Regression (LR). To achieve the research objectives, the sample was divided into three data sets using an expansion window scheme, in which the data are segmented into parts, following a recursive estimation format. The effectiveness of the forecasts was assessed through statistical gains – measured by the out-of-sample  $R^2$  of Campbell and Thompson (2008), and measures of signal forecasts, such as accuracy, sensitivity and specificity. The results show that the application of the stacked generalization technique to combine asset pricing models can significantly improve forecast results in different contexts and market conditions. In addition, training the stacked generalization meta-models by selecting the best base models can improve the forecast accuracy of negative returns, especially in periods of decline. However, the evidence indicates that the models studied were unable to significantly outperform the historical average of stock returns during an exogenous shock.

**Keywords:** return optimization, asset pricing, stacked generalization, machine learning.

## Lista de Figuras

Figura 1 - Estrutura do <i>Stacked Generalization</i> .....	55
Figura 2 – Matriz de confusão da previsão de sinais.....	71
Figura 3 – Série Temporal das Ações por Ano .....	75
Figura 4 – Séries Temporais das Carteiras de Ativos .....	77
Figura 5 – Retorno Previsto das Carteiras no Conjunto de Teste .....	84
Figura 6 – <i>Boxplot</i> Anual do Retorno Logaritmo Mensal do Ibovespa.....	87
Figura 7 – <i>Boxplot</i> anual do <i>DrawDown</i> Mensal do Ibovespa .....	88
Figura 8 – Retorno Previsto das Ações durante a Crise Político-Econômica de 2016 .....	92
Figura 9 – Exemplo de <i>Random Forest</i> (RF).....	108
Figura 10 – Exemplo de <i>Gradient Boosted Regression Trees</i> (GBRT) .....	109
Figura 11 – Exemplo de <i>Neural Networks</i> (NN).....	110

## **Lista de Quadros**

<b>Quadro 1 - Hipóteses de Pesquisa e Objetivos Específicos .....</b>	<b>73</b>
---	-----------

## **Lista de Tabelas**

<b>Tabela 1 – Estatística Descritiva das Carteiras baseadas no Valor de Mercado.....</b>	<b>76</b>
<b>Tabela 2 – R-quadrado fora da amostra de Campbell e Thompson (2008).....</b>	<b>78</b>
<b>Tabela 3 – Resultados de Previsão no Conjunto de Teste.....</b>	<b>80</b>
<b>Tabela 4 – Resultados do Teste de Wilcoxon na Previsão dos Meta-modelos .....</b>	<b>85</b>
<b>Tabela 5 – Resultado de Previsão durante a Crise Político-Econômica de 2016.....</b>	<b>90</b>
<b>Tabela 6 – Resultado de Previsão durante a Pandemia de Covid-19 em 2020 .....</b>	<b>94</b>

## Lista de Abreviaturas e Siglas

AMEX	<i>American Stock Exchange</i>
APT	<i>Arbitrage Pricing Model</i>
ARIMA	<i>Autoregressive Integrated Moving Average</i>
B/M	<i>Book-to-Market</i>
B3	Brasil, Bolsa, Balcão
BACEN	Banco Central do Brasil
CAPM	<i>Capital Asset Pricing Model</i>
C-CAPM	<i>Conditional CAPM</i>
CDI	Certificado de Depósito Interbancário
CMA	<i>Conservative Minus Aggressive</i>
D-CAPM	<i>Downside CAPM</i>
DD	<i>Drawdown</i>
DGP	Dados de Comutação de Regime
EUA	Estados Unidos da América
F&F	Fama e French
FMS	<i>Fast Minus Slow</i>
FN	<i>False Negatives</i>
FP	<i>False Positives</i>
GBRT	<i>Gradient Boosted Regression Trees</i>
HML	<i>High Minus Low</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IPO	<i>Initial Public Offering</i>
L-CAPM	<i>Learning CAPM</i>
LR	<i>Linear Regression</i>
MDD	<i>Maximum Drawdown</i>
ME	Tamanho da empresa
MH	Média Histórica
ML	<i>Machine Learning</i>
MQO	Mínimos Quadrados Ordinários
MSE	<i>Mean Squared Error</i>
MSFE	<i>Mean Squared Forecast Error</i>
NASDAQ	<i>National Association of Securities Dealers Automated Quotations</i>
NEFIN	Núcleo de Pesquisa em Economia Financeira da Universidade de São Paulo
NN	<i>Neural Networks</i>
NYSE	<i>New York Stock Exchange</i>
ON	Ações Ordinárias
P/L	Preço-Lucro
PN	Ações Preferenciais Nominativas
RF	<i>Random Forest</i>
RMW	<i>Robust Minus Weak</i>
SEO	<i>Seasoned Equity Offering</i>
SMB	<i>Small Minus Big</i>
TN	<i>True Negatives</i>
TP	<i>True Positives</i>
WML	<i>Winners Minus Losers</i>

## Lista de Equações

<b>Equação 01 (2)</b> .....	<b>30</b>
<b>Equação 02 (2)</b> .....	<b>31</b>
<b>Equação 03 (2)</b> .....	<b>33</b>
<b>Equação 04 (2)</b> .....	<b>37</b>
<b>Equação 05 (2)</b> .....	<b>40</b>
<b>Equação 06 (2)</b> .....	<b>42</b>
<b>Equação 07 (2)</b> .....	<b>42</b>
<b>Equação 08 (2)</b> .....	<b>43</b>
<b>Equação 09 (2)</b> .....	<b>44</b>
<b>Equação 10 (2)</b> .....	<b>47</b>
<b>Equação 11 (2)</b> .....	<b>47</b>
<b>Equação 12 (2)</b> .....	<b>48</b>
<b>Equação 13 (2)</b> .....	<b>49</b>
<b>Equação 14 (2)</b> .....	<b>50</b>
<b>Equação 15 (2)</b> .....	<b>52</b>
<b>Equação 16 (2)</b> .....	<b>50</b>
<b>Equação 17 (2)</b> .....	<b>52</b>
<b>Equação 18 (2)</b> .....	<b>53</b>
<b>Equação 19 (2)</b> .....	<b>53</b>
<b>Equação 20 (3)</b> .....	<b>67</b>
<b>Equação 21 (3)</b> .....	<b>68</b>
<b>Equação 22 (3)</b> .....	<b>68</b>
<b>Equação 23 (3)</b> .....	<b>68</b>
<b>Equação 24 (3)</b> .....	<b>69</b>
<b>Equação 25 (3)</b> .....	<b>69</b>
<b>Equação 26 (4)</b> .....	<b>88</b>
<b>Equação 27 (4)</b> .....	<b>88</b>

## Sumário

<b>1.</b>	<b>INTRODUÇÃO .....</b>	<b>18</b>
1.1	CONTEXTUALIZAÇÃO DO TEMA .....	18
1.2	OBJETIVO GERAL.....	23
1.3	OBJETIVOS ESPECÍFICOS .....	23
1.4	JUSTIFICATIVA .....	24
<b>2.</b>	<b>REFERENCIAL TEÓRICO .....</b>	<b>27</b>
2.1	PRECIFICAÇÃO DE ATIVOS DE CAPITAL .....	27
<b>2.1.1</b>	<b><i>Portfolio Selection e Capital Asset Pricing Model (CAPM)</i></b> .....	<b>27</b>
<b>2.1.2</b>	<b><i>Arbitrage Pricing Theory (APT)</i></b> .....	<b>32</b>
<b>2.1.3</b>	<b>Modelos de Três e Quatro Fatores</b> .....	<b>36</b>
<b>2.1.4</b>	<b>Modelo de Cinco Fatores</b> .....	<b>42</b>
<b>2.1.5</b>	<b>Modelos para Otimização</b> .....	<b>45</b>
2.2	COMBINAÇÃO DE PREVISÕES .....	49
<b>2.2.1</b>	<b>Métodos para Combinar Previsões</b> .....	<b>49</b>
<b>2.2.2</b>	<b>Estudos Empíricos sobre Combinação de Previsões no Retornos de Ações</b> .	<b>57</b>
<b>3.</b>	<b>PROCEDIMENTOS METODOLÓGICOS .....</b>	<b>64</b>
3.1	DELINEAMENTO DA PESQUISA.....	64
3.2	DADOS, POPULAÇÃO E AMOSTRA DA PESQUISA .....	65
3.3	<i>STACKED GENERALIZATION</i> APLICADA A PRECIFICAÇÃO DE ATIVOS 66	
3.4	HIPÓTESES DE PESQUISA.....	72
<b>4.</b>	<b>RESULTADOS E DISCUSSÕES .....</b>	<b>75</b>
4.1	ESTATÍSTICA DESCRITIVA .....	75
4.2	RESULTADOS DA TÉCNICA DE <i>STACKED</i> .....	77
<b>4.2.1</b>	<b>Desempenho de previsão no conjunto de teste</b> .....	<b>78</b>
<b>4.2.2</b>	<b>Desempenho de previsão em condições extremas de mercado</b> .....	<b>86</b>

<b>5.</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>95</b>
5.1	LIMITAÇÕES DA PESQUISA E SUGESTÕES DE ESTUDOS FUTUROS ...	98
	<b>REFERÊNCIAS .....</b>	<b>100</b>
	<b>APÊNDICE A – FLUXOGRAMA DE <i>STACKED GENERALIZATION</i> APLICADO AOS MODELOS DE PRECIFICAÇÃO DE ATIVOS.....</b>	<b>107</b>
	<b>APÊNDICE B – INTRODUÇÃO AOS META-MODELOS.....</b>	<b>108</b>
	<b>APÊNDICE C – CÓDIGO-FONTE PARA AQUISIÇÃO DA BASE DE DADOS DOS RETORNOS DOS ATIVOS .....</b>	<b>111</b>
	<b>APÊNDICE D – CÓDIGO-FONTE PARA CALCULAR O QUINTIL DA VARIÁVEL TAMANHO (ME) .....</b>	<b>112</b>
	<b>APÊNDICE E – CÓDIGO-FONTE PARA CALCULAR O VALOR DA CARTEIRA TAMANHO (ME).....</b>	<b>114</b>
	<b>APÊNDICE F – CÓDIGO-FONTE PARA CRIAR OS DATAFRAMES EM QUE SERÃO REALIZADAS AS PREVISÕES DAS CARTEIRAS.....</b>	<b>117</b>
	<b>APÊNDICE G – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O COEFICIENTE BETA DA CARTEIRA .....</b>	<b>121</b>
	<b>APÊNDICE H – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO CAPM.....</b>	<b>122</b>
	<b>APÊNDICE I – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O FILTRO DE KALMAN .....</b>	<b>124</b>
	<b>APÊNDICE J – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO L-CAPM.....</b>	<b>125</b>
	<b>APÊNDICE K – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O BETA <i>DOWNSIDE</i> .....</b>	<b>127</b>
	<b>APÊNDICE L – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO D-CAPM.....</b>	<b>128</b>
	<b>APÊNDICE M – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO C-CAPM.....</b>	<b>130</b>
	<b>APÊNDICE N – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO F&amp;F (1996) .....</b>	<b>133</b>

<b>APÊNDICE O – CÓDIGO-FONTE PARA CRIAR A MATRIZ DE PREVISÕES E CALCULAR O R-QUADRADO FORA DA AMOSTRA DAS CARTEIRAS .....</b>	<b>138</b>
<b>APÊNDICE P – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DOS META-MODELOS ST-LR, ST-RF, ST-GBRT E ST-NN.....</b>	<b>146</b>

# 1. INTRODUÇÃO

---

## 1.1 CONTEXTUALIZAÇÃO DO TEMA

Há um longo debate sobre se os preços das ações são previsíveis (Jing, Wu e Wang, 2021). O tema é importante porque desempenha um papel fundamental em diversas áreas de finanças, como precificação de ativos, alocação de portfólio e avaliação de desempenho de gestores de investimentos (Giovannelli, Massacci e Soccorsi, 2021). No entanto, a previsibilidade do retorno das ações ainda é controversa e a especificação do melhor modelo permanece uma questão em aberto. Nos últimos 50 anos, uma quantidade significativa de pesquisas foi realizada sobre o tema, com destaque para duas categorias: (i) estudos que buscam encontrar o melhor modelo de previsão individual, fundamentados nos princípios delineados por Markowitz (1952); e (ii) estudos que se concentram em combinar previsões, respaldados pelo artigo seminal de Bates e Granger (1969).

Na primeira categoria de estudos, encontram-se pesquisas que utilizam modelos de previsão individuais. O *Capital Asset Pricing Model* – CAPM desenvolvido por Sharpe (1964) e Lintner (1965), o *Arbitrage Pricing Model* – APT de Ross (1976), e os modelos de Três Fatores de Fama e French (1993), Quatro Fatores de Carhart (1997) e Cinco Fatores de Fama e French (2015), desempenham um papel crucial na explicação de grande parte da variação dos retornos das ações. Essa compreensão apenas foi possível devido à contribuição fundamental de Markowitz (1952), cuja teoria de seleção de portfólios fundamenta-se na intrínseca relação entre risco e retorno, postulando a existência de uma carteira eficiente.

Markowitz (1952) destaca que a relação entre risco e retorno não apenas implica em diversificação, mas o tipo certo de diversificação pela razão certa. Para o autor, ao tentar reduzir a variância não basta investir em muitos títulos. É necessário investir em indústrias com características econômicas diferentes, visto que elas têm covariâncias mais baixas do que as empresas dentro de uma mesma indústria. As principais conclusões de Markowitz (1952) consideram que através de uma carteira eficiente é possível obter o retorno esperado do ativo assumindo um determinado nível de risco. Todavia, a diversificação não impede que uma carteira eficiente encontre o maior retorno esperado dado o menor nível de risco possível. A partir desta contribuição, outros autores estabeleceram as bases para o desenvolvimento do CAPM que fornece uma relação entre o retorno esperado de um ativo e

seu risco sistemático.

O CAPM é um modelo fundamental e tradicional para a precificação de ativos de capital, sendo a base para diversos outros modelos. Desenvolvido a partir da contribuição de vários autores na década de 1960, como Sharpe (1964), Lintner (1965), Mossin (1966) e Black (1972), o modelo estabelece uma relação linear entre risco e retorno, com a ideia central de uma carteira eficiente e a influência do beta na equação de precificação. O coeficiente beta, representando a sensibilidade do ativo em relação ao mercado, desempenha um papel crucial no modelo, sendo calculado pela covariância do retorno do ativo e do mercado, dividida pela variância do retorno do mercado.

Apesar das premissas restritivas do CAPM, que incluem aversão ao risco, livre competição de mercado e uma taxa livre de risco, o modelo tem sido amplamente aceito pelas implicações que oferece. No entanto, estudos empíricos questionaram sua eficácia, especialmente em relação ao comportamento real dos retornos de títulos. Pesquisas de Friend e Blume (1970), Jensen, Black e Scholes (1972), e Fama e French (2004) revelaram inconsistências entre os retornos observados e os previstos pelo CAPM, levando à rejeição do modelo em certos cenários. Diante dessas limitações, surgiram modelos alternativos, como o APT, buscando superar as simplificações do CAPM e oferecer uma abordagem mais robusta para a precificação de ativos.

Desenvolvido por Ross (1976), o APT se destaca como uma abordagem alternativa ao CAPM, abraçando uma perspectiva mais flexível para a precificação de ativos. O autor reconhece a necessidade de um modelo que supere as premissas restritivas do CAPM, proporcionando uma estrutura que possa lidar com a complexidade do mercado financeiro. Ao contrário do CAPM, que se baseia em uma única equação para relacionar risco e retorno, o APT adota um enfoque multifatorial. Este modelo permite a inclusão de diversos fatores de risco que influenciam os retornos dos ativos, refletindo a diversidade de influências presentes no mercado.

A inovação incorporada pelo modelo apresentado por Ross (1976) está na consideração de que múltiplos fatores de risco podem afetar os preços dos ativos. Esses fatores podem incluir variáveis macroeconômicas, eventos específicos do setor ou outras condições que impactam os mercados financeiros. Assim, o APT oferece uma estrutura mais adaptável e abrangente para entender as fontes de variação nos retornos dos ativos em comparação com o CAPM. A transição do CAPM para o APT reflete a evolução contínua na busca por uma compreensão mais precisa do comportamento dos preços dos ativos. Ambos os modelos compartilham o objetivo fundamental de fornecer um arcabouço teórico para a

precificação de ativos, mas o APT se destaca como uma abordagem mais sofisticada e adaptável às nuances do mercado financeiro, superando as limitações percebidas do CAPM.

Entretanto, o APT recebeu críticas significativas devido a sua falta de especificidade em relação aos fatores de risco a serem considerados. Além disso, os resultados dos testes empíricos geraram ambiguidades, levando pesquisadores a conclusões diversas sobre a validade do modelo. Tais limitações instigaram uma busca por abordagens mais robustas. Autores como Fama e French (1993), Carhart (1997) e Fama e French (2015) responderam a essas críticas apresentando soluções inovadoras baseadas em modelos multifatoriais. Essas abordagens buscam superar as lacunas identificadas no APT, incorporando um conjunto específico de fatores de risco com capacidade para melhorar a explicação dos retornos dos ativos.

Fama e French (1993) propuseram o Modelo de Três Fatores em que incorporaram outras duas variáveis ao prêmio de mercado: o tamanho da empresa e o *book-to-market*, proporcionando uma explicação mais abrangente para os retornos das ações nos mercados financeiros dos EUA. A inclusão desses fatores permitiu uma melhor compreensão das variações nos retornos, especialmente para empresas pequenas e com diferentes características contábeis. Em consequência disto, Carhart (1997) introduziu o Modelo de Quatro Fatores, ao ampliar o modelo de três fatores incorporando um fator baseado na performance passada das ações, o fator “momento”, desenvolvido por Jegadeesh e Titman (1993). Essa adição visou capturar padrões observados nos retornos das ações que não eram explicados pelo CAPM ou pelo modelo de três fatores. A inclusão do fator momento melhorou a capacidade do modelo em explicar anomalias e persistências de desempenho em fundos de investimento.

Posteriormente, Fama e French (2015) apresentaram o Modelo de Cinco Fatores, em que os fatores de rentabilidade e investimento foram adicionados ao modelo original de três fatores de Fama e French (1993) com a finalidade de aumentar a explicação da previsão do retorno das ações. Essa extensão foi baseada na relação entre os dividendos esperados, rentabilidade e investimento, formulando uma abordagem que conecta o retorno esperado das ações a esses fatores. O modelo de cinco fatores proporciona uma explicação mais abrangente dos retornos médios das ações, particularmente ao considerar as relações entre rentabilidade, investimento e outras características contábeis.

Em síntese, os modelos de precificação de ativos, como o CAPM, APT e os modelos de Três, Quatro e Cinco Fatores, representam avanços notáveis na teoria de precificação de ativos. Contudo, diversos estudiosos como Wei (1988), Chen, Roll e Ross (1986), Rapach,

Strauss e Zhou (2010), Rapach e Zhou (2020) e Zhao e Cheng (2022) argumentam que a combinação de diferentes previsões oferece uma explicação mais robusta para os retornos médios das ações, prêmio do risco de capital e retorno de títulos corporativos. Essa abordagem integrada considera uma variedade de fatores e perspectivas, visando proporcionar uma compreensão mais abrangente e precisa do comportamento do mercado financeiro e dos ativos. Assim, a convergência de múltiplos modelos pode oferecer uma visão mais completa e refinada, superando as limitações inerentes aos enfoques individuais.

A segunda categoria de estudos que merecem destaque são os que se concentram em combinar previsões. Zhao e Cheng (2022) propõe que as combinações de previsões são frequentemente consideradas melhores do que as previsões individuais em estudos empíricos e geralmente são mais precisas, à medida que as previsões individuais combinadas forem independentes e diversificadas, porque os erros aleatórios são calculados em média. Além disso, as combinações geralmente estabilizam as previsões, de maneira semelhante à diversificação entre ativos individuais que reduz a variância de um portfólio. O argumento clássico mais recente para combinar previsões é fornecido pelo artigo seminal de Bates e Granger (1969), ampliado por Granger e Ramanathan (1984).

Bates e Granger (1969) são considerados os primeiros a sistematizar um método para combinar os resultados dos modelos de previsão mais eficientes com a finalidade de agregar as características mais relevantes de cada um dos modelos utilizados. Os autores perceberam que quando pesquisadores e pessoas de negócios realizavam duas (ou mais) previsões individuais para um mesmo evento, era natural que aceitassem o melhor modelo e descartassem o outro. O problema é que a previsão descartada possui algumas informações independentes que podem ser bastante úteis. Como quando uma previsão é baseada em variáveis ou informações que a outra previsão não considerou, ou mesmo quando a previsão faz uma suposição diferente sobre a forma de relacionamento entre as variáveis. Neste sentido, combinar as previsões pode aumentar a acurácia e o desempenho esperado.

O método proposto por Bates e Granger (1969) é considerado relativamente simples e eficaz. Simples porque sugere um esquema de combinação linear, que envolve a atribuição de pesos às previsões de cada modelo com base em seu desempenho passado. Os pesos são calculados utilizando a matriz de covariância de erros de suas respectivas previsões individuais e a previsão que obtiver os menores erros quadráticos recebe o maior peso. Eficaz porque a combinação de previsões pode produzir resultados mais precisos do que qualquer modelo individual, pois cada modelo tem suas próprias forças e fraquezas e a combinação deles pode reduzir os erros médios causados por estas limitações.

A partir dos marcos estabelecidos pelos trabalhos destes autores, os métodos de combinação de previsões ampliaram seu escopo para diversas áreas do conhecimento, abrangendo não apenas a estatística, mas também a economia e as finanças. Desde o início da década de 1970, a literatura financeira tem registrado uma variedade de modelos de combinação de previsões na análise e explicação dos retornos das ações. Esses modelos incluem abordagens como o método de combinação de previsões probabilísticas em Von Holstein (1972), combinação de previsões adaptativas em Timmermann (2008), combinação simples em Rapach, Strauss e Zhou (2010), combinação iterada em Lin, Wu e Zhou (2018) e estratégias que incorporam técnicas de *machine learning* (ML) como a combinação simples em Rapach e Zhou (2020) e a generalização empilhada (*stacked generalization*) em Zhao e Cheng (2022).

O método de *stacked generalization* em *machine learning*, investigado por Zhao e Cheng (2022), revela-se uma estratégia promissora para aprimorar a compreensão do comportamento dos retornos das ações. Por meio de testes empíricos, o método demonstrou não apenas a capacidade de superar o desempenho histórico médio do prêmio de risco das ações, mas também de superar abordagens tradicionais em ambientes de elevada complexidade e incerteza no mercado dos EUA. Intuitivamente, essa abordagem utiliza um algoritmo que ajusta um modelo (algoritmo de *machine learning*) de nível superior, chamado de meta-modelo, a um grupo de modelos base para produzir uma previsão refinada.

Zhao e Cheng (2022) descobriram que realizar o *stacked generalization* com um meta-modelo um pouco mais complexo pode superar o *benchmark* médio histórico tradicional em termos de desempenho dentro e fora da amostra. Além disso, o *stacked generalization* geralmente pode ter um desempenho melhor do que vários outros modelos em termos de previsão. Os autores também descobriram que os ganhos de previsão fora da amostra aumentam em mercados de desaceleração extrema.

Para os autores, a principal diferença entre o *stacked generalization* e os métodos tradicionais de combinação de previsões está em dois aspectos: (i) o meta-modelo pode ser uma função complicada, enquanto os métodos tradicionais usam uma função simples como média ou mediana; e (ii) os modelos base podem ser modelos lineares ou não lineares, enquanto as abordagens tradicionais usam modelos lineares de forma intensiva. A utilização do *stacked generalization* é mais adequada no problema de previsão com um ruído, não estacionário, e processo de geração de dados de comutação de regime (DGP). Um DGP é um modelo estatístico que descreve a geração de dados onde as características dos dados podem mudar repentinamente ao longo do tempo ou em diferentes condições, muito comum nos

mercados de capitais. Isso significa que as relações entre as variáveis que geram os dados podem mudar de forma inesperada, e essas mudanças podem afetar a precisão dos modelos de previsão.

O argumento principal para combinar previsões na precificação de ativos de capital é a crença de que a previsão de retorno das ações não pode ser descrita eficientemente por um único modelo, e que modelos de *machine learning*, como o *stacked generalization*, são capazes de melhorar significativamente a explicação de retorno das ações por meio da combinação de diferentes modelos que utilizam algoritmos sofisticados (Granger; Ramanathan, 1984; Timmermann; 2008; Rapach; Strauss; Zhou, 2010). A metodologia de *stacked generalization*, conforme delineada por Zhao e Cheng (2022), não apenas se alinha a essa tendência emergente, mas também fornece evidências concretas de resultados superiores em comparação com diversos outros métodos de combinação de previsões existentes.

Diante deste cenário de inovações e avanços tecnológicos surge a seguinte pergunta de pesquisa: **As previsões dos retornos das ações dos modelos de precificação de ativos são potencializadas quando a técnica de *stacked generalization* em *machine learning* é aplicada nos ativos da B3?**

## 1.2 OBJETIVO GERAL

Esta pesquisa tem como objetivo geral investigar se os retornos das ações dos modelos de precificação de ativos são potencializados quando utilizada a técnica de *stacked generalization* em *machine learning* nos ativos da B3.

## 1.3 OBJETIVOS ESPECÍFICOS

Para atingir o objetivo geral, foram determinados os seguintes objetivos específicos:

- a) Analisar a eficácia dos modelos de precificação de ativos na estimativa do retorno das ações na B3 ao longo do período de 2003 a 2022.
- b) Avaliar se a técnica de *stacked generalization* em *machine learning* aplicada aos modelos de precificação de ativos tem potencial para otimizar o retorno das ações na B3 ao longo do período de 2003 a 2022.

## 1.4 JUSTIFICATIVA

Esta pesquisa é motivada por uma dualidade de fatores que abrange aspectos empíricos e teóricos. Primeiramente, ela se fundamenta nas divergências e limitações observadas em estudos prévios sobre os principais modelos de precificação de ativos. A existência de discrepâncias e desafios identificados nesse corpo de pesquisa sinaliza a necessidade de uma investigação mais aprofundada e crítica para contribuir com avanços significativos. Em segundo lugar, ela se ancora no progresso constante na área de inteligência artificial e no surgimento de novas abordagens em *machine learning* aplicadas às previsões do mercado de capitais. A rápida evolução dessas tecnologias oferece um cenário propício para explorar inovações e possíveis aprimoramentos na previsão de retornos, abrindo espaço para a consideração de métodos mais avançados e adaptáveis.

Os resultados de estudos empíricos dos principais modelos de precificação ativos como CAPM, APT e os modelos multifatoriais (três, quatro e cinco fatores) revelam divergências significativas nas conclusões sobre a validade destes modelos para explicação dos retornos. Enquanto o CAPM enfrenta críticas de que suas premissas simplificadoras não refletem adequadamente a complexidade do mercado, o APT enfrenta desafios relacionados ao número e estabilidade dos fatores de risco, questionando sua aplicabilidade prática. Os modelos de três, quatro e cinco fatores, se apresentam como alternativas para superar estas limitações, no entanto produzem divergências significativas na interpretação dos resultados. Por outra perspectiva, o cenário financeiro global tem se caracterizado por uma constante busca por técnicas de previsão mais eficazes que possam oferecer evidências confiáveis para os investidores e tomadores de decisão. A volatilidade dos mercados de capitais, especialmente no contexto do mercado acionário, apresenta desafios significativos para pesquisadores e investidores, uma vez que a eficácia das previsões desempenha um papel crucial na tomada de decisões estratégicas. Todavia, à medida que a capacidade computacional evoluiu e as técnicas de *machine learning* foram aprimoradas, surgiram oportunidades para explorar abordagens mais avançadas e inovadoras na previsão financeira.

Nos últimos anos, observa-se um expressivo crescimento no número de publicações que integram métodos de previsão financeira e *machine learning* (Kumbure *et al.*, 2022). Essa combinação tem se revelado altamente promissora, dada sua capacidade de analisar, de maneira ágil e precisa, dados históricos e padrões complexos, superando, assim, a eficiência

dos modelos tradicionais. A convergência da aprimorada capacidade de processamento e da sofisticação das técnicas de *machine learning* proporciona a pesquisadores e profissionais ferramentas robustas para tomada de decisões e otimização de modelos de previsão em um cenário econômico cada vez mais dinâmico. Kumbure *et al.* (2022) realizaram uma revisão sistemática da literatura sobre aplicação dos métodos de *machine learning* para previsão do mercado de ações. Os autores revisaram 138 artigos de periódicos publicados durante os anos 2000 e 2019 e descobriram que a maioria dos estudos encontrados haviam sido publicados nos últimos cinco anos. Mais de 50% dos artigos foram publicados entre os anos de 2015 e 2019, enquanto todo o percentual restante foi publicado nos 15 anos anteriores a 2015. Para os autores, os resultados parecem indicar um crescimento exponencial no número de estudos que se concentraram em realizar previsões do mercado de ações usando *machine learning*. Nesse contexto, o avanço crescente na aplicação de métodos de *machine learning* para previsões financeiras, juntamente com as discrepâncias observadas em estudos empíricos que empregaram métodos tradicionais de precificação de ativos, constituem os pilares fundamentais que motivam o desenvolvimento desta pesquisa.

Outros dois fatores preponderantes são fundamentais para justificar a importância de investigar o método de *stacked generalization* na previsão de retorno das ações no contexto do mercado brasileiro: (i) a carência de estudos empíricos que se utilizam de métodos de combinação de previsões no mercado de ações brasileiro – com exceção das pesquisas realizadas por Cavaleri e Ribeiro (2011) e Rosa (2016) que estudaram a combinação de previsões aplicadas a volatilidade, a produção brasileira sobre o tema é virtualmente inexistente; e (ii) o recente crescimento nas publicações de estudos que empregam abordagens de *machine learning* na previsão do comportamento das ações no cenário nacional – nos últimos anos a publicação de estudos com esta abordagem tem aumentado perceptivelmente. Autores como Santos (2020), Abreu (2021), Torres (2021), Bluvol (2022) e Silva (2023) publicaram dissertações com o objetivo de entender a relação entre a utilização de uma variedade de modelos de *machine learning* e redes neurais na previsão do comportamento do mercado de ações brasileiro. Este argumento corrobora com os resultados da pesquisa de Kumbure *et al.* (2022).

Neste sentido, esta pesquisa busca atingir seu objetivo ao introduzir uma inovação em pesquisas sobre o mercado de ações brasileiro: a abordagem de *stacked generalization*. Estudar o mercado acionário brasileiro é relevante porque o Brasil é um país com uma economia em desenvolvimento, e entender melhor a previsão de retornos das ações neste contexto pode fornecer *insights* valiosos para investidores e pesquisadores interessados em

economias emergentes. Por fim, a comparação entre a abordagem de *stacking* e modelos de precificação de ativos pode fornecer uma contribuição acadêmica significativa para o campo de previsão de retorno das ações. Esta pesquisa pode fornecer evidências sobre a eficácia de diferentes métodos e ajudar a avançar o conhecimento na área, beneficiando outros pesquisadores e estudiosos. Ademais, os resultados do estudo podem ter implicações práticas para investidores, profissionais do mercado financeiro e instituições financeiras. Ao demonstrar resultados superiores em comparação aos modelos de precificação de ativos tradicionais, a abordagem de *stacked generalization* pode sugerir uma alternativa eficaz para melhorar a tomada de decisões de investimento e gerenciamento de riscos em cenários reais.

## 2. REFERENCIAL TEÓRICO

---

Este capítulo oferece uma visão abrangente dos conceitos fundamentais relacionados à precificação de ativos de capital e à combinação de previsões no contexto financeiro. Inicialmente, são abordadas as principais teorias de precificação de ativos, como *Portfolio Selection*, *Capital Asset Pricing Model*, *Arbitrage Pricing Theory* e os modelos multifatoriais, com especial destaque para os Modelos de Três, Quatro Fatores e Cinco Fatores. Além disso, são apresentados os modelos de precificação de ativos usados para comparação com o *stacked generalization*. Em seguida, exploram-se os conceitos fundamentais dos principais métodos de combinação de previsões, encerrando com uma revisão de estudos empíricos que investigaram a eficácia de diferentes métodos de combinação na previsão do retorno do prêmio de capital, retorno de ações e retorno de títulos corporativos. O intuito deste capítulo é proporcionar uma compreensão abrangente desses temas, consolidando as bases para a compreensão e aplicação dos métodos discutidos ao longo deste trabalho.

### 2.1 PRECIFICAÇÃO DE ATIVOS DE CAPITAL

#### 2.1.1 *Portfolio Selection e Capital Asset Pricing Model (CAPM)*

Até o início dos anos 1950 o processo de seleção de portfólio para ativos de investimento era realizado em duas etapas. A primeira começava com uma observação baseada na experiência do investidor e finalizava com uma perspectiva sobre o desempenho futuro dos ativos disponíveis, enquanto a segunda iniciava com a crença sobre o desempenho futuro dos ativos disponíveis e terminava com a seleção do portfólio. Neste processo, algumas hipóteses e regras amplamente aceitas consideravam que os investidores além de maximizar o retorno descontado deveriam diversificar e maximizar o retorno esperado de um título (Markowitz, 1952).

Esta compreensão foi modificada quando Harry Markowitz publicou em 1952 um artigo seminal intitulado “*Portfolio Selection*”. Para o autor, ao considerar as imperfeições de mercado, a hipótese de que o investidor deveria maximizar o retorno descontado deve ser rejeitada, pois não implica na existência de uma carteira diversificada que seja preferível entre todas as carteiras diversificadas. O problema consiste no fato de que

independentemente da forma como os retornos esperados são formados - se são utilizadas taxas de desconto iguais ou diferentes para títulos diferentes, ou se as taxas variam ao longo do tempo - esta hipótese implica que o investidor coloque todos os seus fundos no título com o maior valor descontado. Se dois ou mais títulos tiverem o mesmo valor, qualquer um deles, ou qualquer variação deles será tão boa quanto qualquer outro. Da mesma forma, a regra de que o investidor deve diversificar e maximizar o retorno esperado também deve ser rejeitada. Para Markowitz (1952), o pressuposto de que existe uma carteira que proporciona tanto o retorno esperado máximo como a variância mínima, quanto a ideia de que lei dos grandes números garantirá que o rendimento real da carteira será igual ao rendimento esperado, não pode ser aceita, visto que os retornos dos títulos estão demasiado correlacionados e a diversificação não pode eliminar toda a variância.

Neste sentido, a solução proposta por Markowitz (1952) foi a criação da regra *expected returns-variance* (E-V) – “retornos esperados-variância”. A regra é baseada na relação entre risco e retorno e conjectura a existência de uma carteira eficiente. A carteira com o máximo retorno esperado não é necessariamente aquela com a variância mínima, além disso existe uma taxa na qual o investidor pode maximizar o retorno esperado assumindo uma determinada variância ou reduzir a variância dado um determinado nível de retorno. Markowitz considera que a relação entre risco e retorno não apenas implica em diversificação, mas o tipo certo de diversificação pela razão certa. Para o autor, ao tentar reduzir a variância não basta investir em muitos títulos. É necessário investir em indústrias com características econômicas diferentes, visto que elas têm covariâncias mais baixas do que as empresas dentro de uma mesma indústria.

As principais conclusões de Markowitz (1952) consideram que através de uma carteira eficiente é possível obter o retorno esperado do ativo assumindo um determinado nível de risco. Todavia, a diversificação não impede que uma carteira eficiente encontre o maior retorno esperado dado o menor nível de risco possível. O autor também defende que é necessário que a covariância entre os ativos seja a menor possível, e que isto pode ser alcançado investindo-se em empresas de segmentos diferentes; e finalmente, a diversificação não pode eliminar totalmente o risco da carteira. A partir destas conclusões uma quantidade significativa de pesquisas sobre o comportamento do mercado de capitais foram desenvolvidas nos anos seguintes, com destaque para os trabalhos de Tobin (1958), Sharpe (1964), Lintner (1965), Mossin (1966) e Black (1972) que tiveram como objetivo estudar a relação entre risco e retorno e estabeleceram as bases para o desenvolvimento do *Capital Asset Pricing Model* (CAPM).

O CAPM é considerado o principal e o mais tradicional modelo para precificação de ativos de capital, fornecendo apoio para vários outros modelos. O seu desenvolvimento é atribuído a vários autores que apresentaram o modelo de forma independente na década de 1960. Sharpe (1964) foi um dos primeiros a discutir o tema ao publicar o artigo "*Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*" em que abordou os princípios para construção de uma teoria de equilíbrio de mercado sob condições de risco. O seu objetivo era lançar uma luz sobre a relação entre o preço de um ativo e os seus vários componentes de risco. Para o autor, há uma relação linear entre o risco e o retorno de um ativo, de tal forma que, um incremento no retorno esperado apenas pode ocorrer quando o mesmo incorre em mais risco para o investidor.

No ano seguinte, Lintner (1965) publicou um artigo seminal que abordou as questões de seleção de portfólios de investimento e avaliação de ativos de risco sob ótica de um equilíbrio de mercado. Dentre as suas principais conclusões, ele estabeleceu que existe uma correlação positiva entre o risco e o retorno esperado em ativos financeiros. Isso significa que os investidores esperam, ou pelo menos deveriam esperar, receber um retorno esperado mais alto por assumir mais riscos. Semelhantemente, Mossin (1966) investigou as propriedades de ativos de risco com base em um modelo de equilíbrio geral onde os investidores individuais procuram maximizar as funções de referência sobre o rendimento esperado e a variância do rendimento de suas carteiras. Dentre as suas principais contribuições para o CAPM o autor apresenta resultados semelhantes aos de Lintner (1965), encontrando uma correlação positiva entre o risco e o retorno esperado dos ativos, mesmo usando uma formulação diferente, em que o retorno esperado é estimado com base em seu risco e no retorno esperado do mercado como um todo. Mossin (1966) também delineou uma teoria dos prêmios de risco de mercado e apontou que o equilíbrio geral implica a existência de uma chamada linha de mercado, que relaciona o preço do rendimento esperado e o desvio padrão do rendimento.

Por sua vez Black (1972) investigou o comportamento do CAPM em mercados com condições de empréstimos restritos. Ao considerar que uma das premissas fundamentais do CAPM é garantir que os investidores possam emprestar ou tomar fundos emprestados livremente, o autor decidiu analisar se os resultados do modelo eram válidos em mercados com restrição de empréstimos. Os principais resultados indicaram que o CAPM é válido mesmo quando os investidores não possam emprestar livremente, e que o modelo também pode ser usado para selecionar carteiras ótimas de ativos nesses mercados. As descobertas de Black apontaram que o CAPM era um modelo mais versátil do que se pensava

anteriormente e poderia ser usado para estimar o retorno esperado de ativos mesmo em mercados que não são perfeitamente eficientes.

Durante a década de 1960, vários autores contribuíram para o desenvolvimento do CAPM. Com base no modelo, respeitando certos pressupostos, o retorno esperado de qualquer ativo durante um único período será satisfeito pela seguinte equação (Friend e Blume, 1970; Black, 1972):

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \quad (1)$$

Onde:

$E(R_i)$  = retorno esperado do ativo  $i$

$R_f$  = retorno do ativo livre de risco

$\beta_i$  = sensibilidade do ativo  $i$  em relação ao mercado

$E(R_m)$  = retorno esperado do mercado

$E(R_m) - R_f$  = prêmio do risco de mercado

A equação do CAPM expressa que o retorno esperado de um ativo é igual à taxa livre de risco  $R_f$  mais uma compensação pelo risco sistemático  $\beta_i[E(R_m) - R_f]$ .

Ao desenvolverem o CAPM os autores destacaram a existência de dois tipos de riscos que impactam os retornos dos ativos. O primeiro é o risco não sistemático, ou diversificável – é a volatilidade que pode ser reduzida por meio da diversificação de uma carteira. Considera fatores internos de uma empresa ou de um setor, como governança, concorrência, mão de obra qualificada ou eventos específicos que não impactam o mercado como um todo. Para Markowitz (1952), a diversificação pode mitigar o risco não sistemático, enquanto permite aos investidores obterem um maior retorno esperado. No entanto, ela não pode eliminar toda a variância. O segundo é o risco sistemático, ou risco de mercado – é a volatilidade associada a fatores que afetam o mercado financeiro como um todo. Ele não pode ser solucionado por meio da diversificação, pois está relacionado a outros componentes de riscos, como eventos macroeconômicos, mudanças na taxa de juros, instabilidade política, crises financeiras globais e desastres naturais. Sharpe (1964) apresentou as diretrizes para calcular o risco sistemático de cada ativo e ajudou a estabelecer as bases para inserção da sensibilidade do ativo em relação ao mercado  $\beta_i$  na equação do CAPM.

A sensibilidade do ativo em relação ao mercado, também conhecida como coeficiente beta do ativo, é definida pela seguinte equação:

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \quad (2)$$

Desta forma, o coeficiente beta do ativo  $\beta_i$  é expresso pela razão entre a covariância do retorno do ativo  $R_i$  e o retorno de mercado  $R_m$  e variância do retorno de mercado  $R_m$  (Black, 1972).

As premissas e os pressupostos geralmente usadas na derivação da equação do CAPM são as seguintes: (i) preferência do investidor – o investidor pensa nos resultados de qualquer investimento em termos de alguma distribuição de probabilidade e ao avaliar a conveniência de um investimento está disposto a agir com base em apenas dois parâmetros: valor esperado e desvio padrão; (ii) maximização da utilidade esperada – o investidor escolhe a carteira que maximiza a utilidade esperada da riqueza no final do período; (iii) aversão ao risco – o investidor é avesso ao risco e opta por investimentos que ofereçam um desvio padrão menor; (iv) livre competição de mercado – todo investidor é capaz de contrair empréstimos ou emprestar fundos em igualdade de condições; (v) mercado puramente competitivo – o mercado é livre de custos de transação e impostos; e (vi) taxa livre de risco – todo investidor pode negociar a uma taxa livre de risco em condições de igualdade (Sharpe, 1964; Lintner, 1965; Black, 1972).

Para Sharpe (1964) mesmo que as suposições do CAPM pareçam ser altamente restritivas e irrealistas, a formulação apresentada está longe de ser rejeitada. Ele defende que o teste de uma teoria não é o realismo de seus pressupostos, mas a aceitabilidade de suas implicações. Ademais, as premissas apresentadas implicam condições de equilíbrio de mercado que constituem uma parte importante da doutrina financeira clássica. Especialmente, em virtude da escassez de modelos alternativos que conduzam a resultados semelhantes. Todavia, o argumento de Sharpe não poupou o CAPM de ser fortemente criticado por autores como Black (1972), Jensen, Black e Scholes (1972) e Fama e French (2004).

Estudos empíricos sugeriram que o retorno dos títulos não se comporta como previsto pelo CAPM. Friend e Blume (1970) utilizam uma regressão transversal entre o desempenho ajustado ao risco e o risco para testar empiricamente 200 carteiras aleatórias no mercado dos EUA no período de 1960 a 1968. Eles observaram que as carteiras de alto risco parecem ter um desempenho fraco, enquanto as carteiras de baixo risco têm um bom desempenho. Os autores assumem que há algum viés em seu teste, mas afirmam

alternadamente que o viés é tão pequeno que pode ser ignorado e que explica metade do efeito que observam. No entanto, Black (1972) defende que esse viés é grave, e mesmo após uma análise extensiva de sua natureza as correções indicam que existe uma relação negativa entre o risco e o desempenho das carteiras.

Jensen, Black e Scholes (1972) testaram empiricamente os retornos das carteiras de ações em diferentes níveis de risco no período de 1926 a 1966 e encontraram resultados semelhantes aos de Friend e Blume (1970). Eles descobriram que os retornos médios das carteiras não são consistentes com o CAPM, especialmente no período pós-guerra de 1946 a 1966. Os resultados indicam que o retorno esperado das carteiras não é estritamente proporcional ao beta. As estimativas de retorno com níveis mais baixos de beta foram consistentemente mais altas do que o previsto pelo CAPM, enquanto as estimativas de retorno com níveis mais altos de beta foram consistentemente mais baixas do que o previsto pelo modelo. Os autores concluíram que as evidências encontradas são suficientemente fortes para justificar a rejeição da forma tradicional do modelo.

Na visão de Fama e French (2004), as premissas do modelo são simplificações demasiadas da realidade e os seus resultados não refletem o desempenho do mercado. Para os autores, os testes empíricos do CAPM apresentaram resultados ruins o bastante para invalidar a forma como o modelo é utilizado. Os problemas empíricos poderiam refletir falhas teóricas - resultado de muitas suposições simplificadoras - ou serem causados por dificuldades na implementação de testes do modelo. Em todo o caso, eles veem que a falha do CAPM em testes empíricos implica que a maioria das aplicações do modelo seriam inválidas. Diante destas limitações, outros modelos como o *Arbitrage Pricing Model* – APT, começaram a surgir como alternativa ao CAPM.

### ***2.1.2 Arbitrage Pricing Theory (APT)***

O APT foi apresentado de forma consolidada pela primeira vez por Ross (1976), com base em suas próprias contribuições realizadas durante a década de 1970. O modelo foi proposto como uma alternativa ao CAPM, que havia se tornado a principal ferramenta analítica para explicar fenômenos observados nos mercados de capitais para ativos de risco. Diferentemente do CAPM, o APT presume que o preço de um ativo não pode ser explicado por um único fator, mas por uma série de fatores de riscos (Ross, 1976). Estes fatores podem incluir taxas de juros, taxas de inflação, mudanças nas taxas de câmbio, produção industrial,

dentre outros (Chen, Roll e Ross, 1986). O APT pode ser considerado o primeiro modelo multifatorial da moderna teoria financeira, pois estabeleceu precedentes de que o retorno dos ativos pode ser explicados por uma combinação de fatores de risco além do risco sistemático representado pelo beta.

Megginson (1997) aponta que o modelo proposto por Ross é baseado em cinco suposições: (i) o mercado de capitais é perfeitamente competitivo e sem fricções; (ii) os investidores têm expectativas homogêneas sobre os retornos esperados de ativos individuais; (iii) há uma relação linear entre o retorno esperado dos ativos e os fatores de risco; (iv) os fatores de risco do modelo capturam todo o risco sistemático ao qual os ativos estão expostos; e (v) o número de fatores de risco deve ser menor do que o número de ativos. O APT considera que o retorno esperado de um conjunto de ativos pode ser expresso pela seguinte equação (Ross, 1976):

$$R_i = E(R_i) + b_{i1}\delta_1 + \dots + b_{in}\delta_n + \epsilon_i \quad (3)$$

Onde:

$R_i$  = retorno do ativo  $i$

$E(R_i)$  = retorno esperado do ativo  $i$

$b_{ij}$  = sensibilidade do ativo  $i$  em relação ao fator  $j$

$\delta_j$  = fator zero médio comum em relação a  $j$

$\epsilon_i$  = valor residual de média zero

Percebe-se que em consequência da relação linear entre os fatores de risco e o retorno do ativo, o APT constitui uma teoria relacionada ao CAPM. No entanto, Ross (1976), Huberman (1982) e Megginson (1997) apontam distinções tanto do ponto de vista teórico quanto do ponto de vista empírico que tornam os modelos substancialmente diferentes.

O APT conceitua que à medida em que o número de ativos em um modelo aumenta, o termo de erro tende a ser zero. Para Ross (1976) isso é extremamente importante, porque significa que os termos de erro desaparecerão efetivamente à medida que um grande número de ativos for adicionado a uma carteira. Um ponto fraco deste argumento é que conforme o número de ativos aumenta, a riqueza também aumentará, e o aumento da riqueza, porém, pode levar ao aumento da aversão ao risco por parte de alguns investidores.

Megginson (1997) indica que o processo de arbitragem real proposto pelo APT explora o fato de que um investidor pode usar o dinheiro obtido pela compra ou venda de

alguns títulos para comprar outros. Desta forma, o investidor pode, por meio de um processo de arbitragem, construir uma carteira que não exija nenhum compromisso de riqueza e ausente de riscos sistemáticos e não sistemáticos. A lógica deste argumento é que o investidor pode construir uma carteira de arbitragem combinando um grupo de ativos para que a soma do valor de seus investimentos em ativos individuais seja igual a zero. Para o autor, se essa carteira for construída utilizando muitos ativos, o risco intrínseco será muito próximo de zero, e se a carteira também for construída de tal forma que todo o risco sistemático seja eliminado, será uma carteira livre de risco.

Outra característica relevante do modelo é a eliminação da necessidade de alguns pressupostos estabelecidos pelo CAPM, como a existência de uma distribuição de probabilidade de retorno dos ativos e a hipótese de que o mercado está em equilíbrio. Para os autores, o APT é válido não apenas em situações de equilíbrio, mas em todos os tipos de desequilíbrio, desde que não haja oportunidades de arbitragem. O modelo de arbitragem também descarta a necessidade de um portfólio de mercado ao deixar de lado a discussão sobre como identificá-lo e sobre como tratá-lo em estudos empíricos. A principal justificativa de Ross (1976) é que a carteira de mercado não desempenha nenhum papel especial no modelo.

Tais distinções contribuíram para que estudos empíricos sobre a aplicação do modelo de arbitragem como forma de prever o retorno do preço de ativos financeiros fossem publicados a partir da década de 1980 com destaque para os trabalhos de Roll e Ross (1980), Dhrymes, Friend e Gultekin (1984), Chen, Roll e Ross (1986) e Wei (1988).

Os primeiros a testar empiricamente o APT foram Roll e Ross (1980) que propuseram estimar estatisticamente quantos fatores seriam necessários para explicar o retorno das ações individuais no mercado de ações dos EUA no período de 1962 a 1972. Os autores encontraram evidências de que os retornos de ações individuais podem ser explicados por pelo menos três e possivelmente quatro fatores de risco sistemáticos. Alguns destes fatores incluem: (i) o retorno do mercado - o retorno esperado de uma ação é positivamente correlacionado com o retorno esperado do mercado; (ii) o tamanho da empresa - as ações de empresas maiores tendem a ter retornos mais altos do que as ações de empresas menores; (iii) a relação *book-to-market* - as ações de empresas com uma relação *book-to-market* alta tendem a ter retornos mais altos do que as ações de empresas com uma relação *book-to-market* baixa; e (iv) o impulso - as ações que tiveram retornos acima da média no passado tendem a continuar tendo retornos acima da média no futuro. Eles concluíram que o APT é uma teoria mais geral e flexível do que o CAPM, e que pode ser usada para melhor explicar

os retornos de ações individuais, através de múltiplos fatores de risco.

No entanto, Dhrymes, Friend e Gultekin (1984) decidiram reexaminar as evidências empíricas apresentadas por Roll e Ross (1980) e apontaram falhas na metodologia empregada por eles. Os autores encontraram evidências de que o número de fatores necessários no modelo aumenta de acordo com o número de ativos sendo avaliados, o que é inconsistente com o APT, que prevê que o número de fatores seja constante. Também indicaram que o modelo não é capaz de determinar quais ou quantos fatores de risco são necessários para explicar o retorno dos ativos, e que estes fatores são instáveis ao longo do tempo. Mesmo após controlá-los estatisticamente, há uma parcela significativa da variação dos retornos dos ativos que não é explicada. Eles concluíram que a metodologia utilizada por Roll e Ross (1980) e por outros trabalhos empíricos são falhas, e que os resultados desses trabalhos não podem ser interpretados como evidência de que o APT é um modelo válido para explicar os preços dos ativos.

Tais conclusões reforçaram as críticas de que o modelo original do APT não é capaz de determinar quais são os fatores que conseguem explicar o retorno dos ativos e nem quantos fatores são necessários. No entanto, Chen, Roll e Ross (1986) afirmam que ao considerar a perspectiva dos mercados eficientes e as teorias de precificação, os preços dos ativos dependem de sua exposição às variáveis macroeconômicas. Desta forma, eles expandiram a análise realizada por Roll e Ross (1980) e exploraram um conjunto de variáveis macroeconômicas e as suas respectivas influências no retorno do mercado de ações. O estudo foi realizado no mercado dos EUA no período de 1953 a 1983 e considerou variáveis como produção industrial, inflação, taxa de juros, prêmio de risco, consumo e preço do barril de petróleo. Os resultados mostraram que as variáveis que mais influenciam significativamente nos retornos das ações são a produção industrial, o prêmio de risco, as mudanças na curva de rendimentos e, de forma um pouco mais fraca, as medidas de inflação. Por outro lado, foram encontradas evidências de que a variável consumo e o preço do barril do petróleo não exercem influência significativa no retorno das ações. Com base nestes resultados os autores lançaram luz sobre a influência dos fatores de risco no preço dos ativos. Eles concluíram que os retornos das ações estão significativamente expostos ao comportamento das variáveis macroeconômicas, e em decorrência disso, são precificados com base nestas exposições.

Wei (1988) propôs uma abordagem em que tanto os aspectos positivos do modelo APT quanto do modelo CAPM são combinados para criar um modelo unificado - APT/CAPM. Os resultados demonstraram que ao adicionar a carteira de mercado como um fator extra ao APT obtém-se um melhor poder explicativo dos retornos dos ativos. Ele

também explorou as implicações empíricas sobre a determinação do número de fatores de risco necessários para explicar os retornos dos ativos, e descobriu que, se todos os fatores forem omitidos, o novo modelo é reduzido ao CAPM. Por outra perspectiva, se nenhum dos fatores for omitido, o novo modelo torna-se ou o APT numa economia infinita ou a teoria unificada de precificação de ativos numa economia finita. Outra conclusão importante deste estudo é que o papel desempenhado pela carteira de mercado depende do número de fatores omitidos, se tornando mais importante à medida que o número de fatores explicativos é reduzido.

Os resultados dos estudos empíricos sobre as aplicações do APT e CAPM levaram pesquisadores a diferentes conclusões sobre a validade dos modelos. Se por um lado o risco sistemático proposto pelo CAPM parecia não explicar o retorno médio das ações, por outro, o APT não conseguia especificar quais fatores de risco influenciavam os preços dos ativos. Neste sentido, Fama e French (1993) propuseram a utilização de um modelo alternativo, o modelo de três fatores, que leva em consideração fatores adicionais ao risco sistemático, como o tamanho da empresa e a relação entre o valor contábil e o valor de mercado (*book-to-market*). Os resultados dos estudos empíricos sobre o modelo de três fatores trouxeram uma nova perspectiva para compreensão do retorno das ações e inspirou o desenvolvimento de diversos outros modelos multifatoriais, como o modelo de quatro fatores de Carhart (1997) e o modelo de cinco fatores de Fama e French (2015).

### **2.1.3 Modelos de Três e Quatro Fatores**

O modelo de três fatores foi desenvolvido por Fama e French (1993). Eles observaram que no mercado dos EUA o beta possuía pouca relação positiva com o retorno médio das ações. No entanto, estudos anteriores indicavam que, além do prêmio de mercado, outras variáveis pouco exploradas pelas teorias de precificação de ativos afetavam diretamente os mercados de títulos e de ações. Após testarem a existência de correlação entre o retorno das ações e estes fatores de risco, os autores concluíram que a combinação de três fatores poderiam explicar o retorno do mercado de ações melhor do que o beta. As variáveis originais do modelo de três fatores construído por Fama e French são: (i) o prêmio de risco de mercado, como indicado por Sharpe (1964) e Lintner (1965); (ii) o tamanho da empresa, definido pelo preço das ações vezes o número de ações, e; (iii) o *book-to-market*, definido pela relação entre o valor de mercado e o valor do patrimônio líquido.

Para determinar a variável tamanho, os autores classificaram as empresas em dois grupos: grandes (*big*) e pequenas (*small*), com base na mediana das ações negociadas em junho de cada ano. A variável *book-to-market* foi calculada considerando a razão entre o valor contábil do ano t-1 e o valor de mercado da empresa em dezembro do ano t-1. Com base nisto, as empresas foram divididas em três grupos de acordo com o resultado do indicador: as 30% inferiores (*low*), as 40% intermediárias (*medium*) e as 30% mais altas (*high*). A decisão de dividir as variáveis tamanho e *book-to-market* em grupos baseou-se no estudo de Fama e French (1992) que indicava que esta formulação apresentava maior poder explicativo em relação ao retorno das ações.

Desta forma, as empresas foram classificadas em 6 carteiras distintas:

- SL = empresas pequenas (*small*) com baixo (*low*) *book-to-market*.
- SM = empresas pequenas (*small*) com médio (*medium*) *book-to-market*.
- SH = empresas pequenas (*small*) com alto (*high*) *book-to-market*.
- BL = empresas grandes (*big*) com baixo (*low*) *book-to-market*.
- BM = empresas grandes (*big*) com médio (*medium*) *book-to-market*.
- BH = empresas grandes (*big*) com alto (*high*) *book-to-market*.

Para calcular o retorno das carteiras, foi utilizada a média aritmética de todas as carteiras que compõe o grupo de empresas que se deseja analisar. Portanto, o retorno das carteiras das empresas *small* é dado por  $(SL + SM + SH) \div 3$ , enquanto o retorno das carteiras *big* é igual a  $(BL + BM + BH) \div 3$ . A mesma regra é aplicada aos retornos das carteiras com alto, médio e baixo *book-to-market*.

No intuito de minimizar a correlação entre as variáveis explicativas do modelo, os autores propuseram que o tamanho das empresas deve ser determinado pelo excesso de retorno das empresas pequenas sobre as empresas grandes (*Small Minus Big*, ou SMB). O *book-to-market* deve considerar o excesso de retorno das empresas com alto *book-to-market* sobre as empresas com baixo *book-to-market* (*High Minus low*, ou HML), enquanto o prêmio de risco será calculado pela diferença entre a taxa livre de risco e o retorno de mercado.

O modelo de três fatores de Fama e French (1993) pode ser expresso pela seguinte equação:

$$R_{it} - R_{Ft} = \alpha_i + b_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + \epsilon_{it} \quad (4)$$

Onde:

$R_{it}$  = retorno do ativo  $i$  no tempo  $t$

$R_{Ft}$  = retorno do ativo livre de risco no tempo  $t$

$R_{Mt}$  = retorno da carteira de mercado no tempo  $t$

$SMB_t$  = retorno da carteira *small minus big* no tempo  $t$

$HML_t$  = retorno da carteira *high minus low* no tempo  $t$

$\alpha_i$  = intercepto para todas as carteiras  $i$

$b_i, s_i, h_i$  = coeficientes de sensibilidade do ativo  $i$

$\epsilon_{it}$  = valor residual de média zero

A capacidade do modelo de três fatores em aumentar o poder explicativo do retorno dos ativos em comparação ao CAPM, levou os autores a conduzirem diversos estudos empíricos sobre o tema. Um ano antes à publicação do artigo que apresentou o modelo consolidado, Fama e French (1992) publicaram um estudo no qual buscavam avaliar a existência de correlação entre os fatores tamanho, alavancagem, preço-lucro (P/L) e *book-to-market* e o retorno das ações da *American Stock Exchange* (Amex), *National Association of Securities Dealers Automated Quotations* (NASDAQ) e *New York Stock Exchange* (NYSE) no período de 1963 a 1990. Eles concluíram que a combinação dos fatores tamanho e *book-to-market* eram suficientes para melhor explicar o retorno das ações, pois demonstraram ser capazes de absorver a contribuição trazida pelos fatores alavancagem e P/L, e com isso, estabeleceram as bases para o modelo.

No ano seguinte, após testar os três fatores, Fama e French (1993) descobriram que: em relação ao fator tamanho, as empresas menores mostraram uma maior sensibilidade ao fator SMB. Os autores perceberam que enquanto as empresas do maior quintil de valor de mercado apresentaram sensibilidade negativa ao fator SMB, os demais quintis apresentaram sensibilidade positiva. Em relação ao *book-to-market*, as empresas com maior *book-to-market* mostraram maior sensibilidade ao fator HML. Eles descobriram que as empresas de menor quintil apresentaram sensibilidades negativas, enquanto as empresas dos demais quintis apresentaram uma variação positiva do indicador, de tal forma que, quanto maior o *book-to-market*, maior a sensibilidade ao fator HML. Em relação ao fator mercado, o resultado foi significativo com um  $R^2$  de 0,83 em todas as regressões, sendo superior a 0,9 na maioria dos casos, enquanto o intercepto se aproximava de zero. Neste contexto, eles acreditam que os resultados empíricos evidenciam que o modelo de três fatores é capaz de

capturar uma grande parte da variação no retorno das carteiras para o período de 1963 a 1990.

Banz (1981), Lakonishok e Shapiro (1986), Jaffe, Keim e Westerfield (1989) e Fama e French (1992) apresentaram evidências empíricas de que o retorno médio das ações está correlacionado com diferentes fatores de risco, além do beta. Como estes padrões nos retornos não são aparentemente explicados pelo CAPM, são chamados de anomalias. Fama e French (1996) apresentaram evidências empíricas que apoiam o modelo de três fatores como uma explicação para as anomalias de precificação de ativos. Eles descobriram que, com exceção do fator momento, o modelo captura a maior parte das anomalias discutidas na literatura financeira. O modelo é capaz de explicar o retorno superior das ações de pequenas empresas em relação às ações de grandes empresas, o retorno superior das ações com baixa relação P/L em relação às ações com alta relação P/L, e o retorno superior das ações com alta volatilidade em relação às ações com baixa volatilidade. Eles concluíram que mesmo sendo uma explicação consistente para a relação entre o retorno dos ativos e as características das empresas, algumas evidências indicam que o modelo de três fatores não é capaz de explicar todas as anomalias de precificação de ativos, como o fator momento.

O fator momento foi desenvolvido por Jegadeesh e Titman (1993). Os autores descobriram que a compra ou vendas de ações com base no desempenho dos últimos doze meses, tendem a produzir retornos anormais positivos durante o ano subsequente. O fator momento indica que se os investidores detiverem as ações com bom desempenho recente tenderão a obter retornos anormais por um período maior do que os investidores com desempenho recente ruim. As evidências empíricas de Jegadeesh e Titman (1993) foram corroboradas por Rouwenhorst (1998, 1999) e Jegadeesh e Titman (2001). No entanto, foi Carhart (1997) que primeiro introduziu o fator momento ao modelo original de três fatores de Fama e French (1993), dando origem ao que ficou conhecido como modelo de quatro fatores.

O estudo de Carhart (1997) analisou os retornos das ações de 1.892 fundos de investimento no mercado dos EUA para o período de 1962 a 1993. O autor utilizou uma metodologia que controlou para o viés de sobrevivência, que ocorre quando, os fundos com melhor desempenho no passado têm maior probabilidade de permanecer no mercado no futuro. Ele construiu um modelo de quatro fatores ao incluir o fator momento de Jegadeesh e Titman (1993) ao modelo de três fatores de Fama e French (1993). O autor considera que o modelo de quatro fatores é consistente com um modelo de equilíbrio de mercado com quatro fatores de risco. Isso significa que o modelo assume que os retornos das ações são determinados pelo risco e não por fatores não observáveis, como habilidades ou informações

privilegiadas dos gestores de fundos. Os resultados indicaram que o fator momento é uma variável significativa que explica a persistência do desempenho dos fundos de investimento. Os fundos que tiveram um bom desempenho no ano anterior apresentaram um desempenho superior ao mercado no ano seguinte. Essa persistência foi significativa mesmo após o controle para o viés de sobrevivência.

Alternativamente, o modelo de quatro fatores pode ser interpretado como um modelo de atribuição de desempenho. Nesse caso, os coeficientes e prêmios nas carteiras que imitam fatores indicam a proporção do retorno médio atribuível a quatro estratégias elementares: (i) ações de beta alto versus baixo – ações com betas mais altos têm retornos médios mais altos do que ações com betas mais baixos; (ii) ações de grande capitalização de mercado *versus* ações de pequena capitalização – ações de grande capitalização têm retornos médios mais altos do que ações de pequena capitalização; (iii) valor versus crescimento – ações com valores de mercado em relação ao seu valor contábil mais baixos têm retornos médios mais altos do que ações com valores de mercado mais altos; (iv) fator momento *versus* ações contrárias – ações que tiveram um bom desempenho no ano anterior têm retornos médios mais altos do que ações que tiveram um desempenho ruim no ano anterior.

O modelo de quatro fatores de Carhart (1997) pode ser expresso pela seguinte equação:

$$R_{it} - R_{Ft} = \alpha_i + b_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + w_i WML_t + \epsilon_{it} \quad (5)$$

Onde o  $WML_t$  representa o retorno da carteira *winner minus losers* no tempo  $t$ . Brav, Geczy e Gompers (2000), utilizam o modelo de quatro fatores para avaliar se o retorno anormal após emissões de ações em ofertas públicas iniciais, *initial public offering* (IPO) e secundárias, *seasoned equity offering* (SEO) é uma anomalia para o período de 1975 a 1992 no mercado dos EUA. Os resultados indicam que, em geral, o retorno anormal após as emissões de ações é negativo. Todavia, apenas o menor tercil de SEOs apresenta desempenho inferior ao modelo de três e quatro fatores. Os autores também identificaram que os retornos anormais são ainda mais negativos em empresas com menor valor de mercado e maior valor contábil, assim como torna-se mais negativo em empresas que realizam uma IPO do que em empresas que realizaram uma SEO. Eles concluem que o retorno anormal após as emissões de ações não é uma anomalia. O retorno negativo pode ser explicado pelo fato de que as empresas que realizam emissões de ações tendem a ser menores, com maior valor de mercado contábil e com maior risco de mercado. Os autores também sugerem que o retorno anormal

após as emissões de ações pode ser explicado por fatores como a assimetria de informação entre investidores e gestores das empresas e pelas expectativas de que as empresas que realizaram as emissões estão buscando capital para financiar projetos de investimento lucrativos ou enfrentando dificuldades financeiras.

Jegadeesh (2000) utilizou o modelo de quatro fatores de Carhart (1997) como *benchmark* para investigar o desempenho de longo prazo das ofertas públicas de ações (SEOs) de 1980 a 1997 no mercado dos EUA. As evidências empíricas indicam que as empresas tiveram um desempenho significativamente inferior ao *benchmark* do modelo de quatro fatores nos cinco anos seguintes às emissões de ações, e que o retorno ajustado ao risco é negativo em todos os períodos analisados, independentemente do tamanho da empresa, do crescimento das vendas ou da avaliação. Estes resultados são divergentes dos encontrados por Bray, Geczy e Gompers (2000). Para o autor, uma possível explicação para tal divergência, é que talvez, o *benchmark* do modelo de fator utilizado pelos autores tenha sido especificado incorretamente. Em conclusão, ele sugere que o modelo de quatro fatores não é adequado para explicar o desempenho de longo prazo das SEOs, pois, o desempenho inferior das SEOs não pode ser atribuído a fatores sistemáticos, como risco de mercado, tamanho, valor e fator momento.

Com o objetivo de esclarecer se os modelos empíricos de precificação de ativos capturam os padrões e dinâmicas nos retornos médios internacionais, e entender até que ponto a precificação de ativos está integrada nestes mercados, Fama e French (2012) investigaram a validade dos modelos de três e quatro fatores em 23 países nas regiões da América do Norte, Europa, Japão e Ásia-Pacífico no período de 1989 a 2011. As evidências indicam que o modelo de quatro fatores é válido para explicar o retorno das ações em mercados internacionais. Os resultados sugerem que os fatores tamanho, *book-to-market* e mercado apresentaram magnitudes semelhantes às observadas em mercados desenvolvidos, no entanto, a relevância do fator de momento é maior em mercados emergentes.

Os resultados de Fama e French (2012) são consistentes com os de outros estudos que investigaram a validade do modelo de quatro fatores em mercados internacionais. Esses estudos sugerem que a relevância do fator momento é maior em mercados emergentes, e que os interceptos das regressões dos modelos são significativos, fornecendo indícios de que há outros fatores que não estão sendo capturados pelos modelos. Estas descobertas levaram outros pesquisadores a buscarem novos fatores explicativos para os retornos médios das ações. Fama e French (2015), por exemplo, incluíram os fatores de lucratividade e investimento ao modelo original de três fatores de Fama e French (1993) e deram origem ao

modelo de cinco fatores. Este modelo tem sido amplamente utilizado para demonstrar melhoria na capacidade explicativa no retorno médio das ações.

## 2.1.4 Modelo de Cinco Fatores

O modelo de cinco fatores de Fama e French (2015) acrescentou os fatores lucratividade e investimento aos fatores mercado, tamanho e *book-to-market* do modelo original de três fatores de Fama e French (1993). Em seu estudo, Fama e French (2015) destacam que há muitas evidências de que os retornos médios das ações estão relacionados com o índice *book-to-market*, e que a lucratividade e o investimento são fatores importantes para determinação dos retornos médios fornecidos pelo *book-to-market*. Nesta investigação, eles decidiram usar o modelo de desconto de dividendos para explicar porque estas variáveis estão relacionadas aos retornos médios. O modelo propõe que o valor de mercado de uma ação consiste no valor descontado dos dividendos esperados por ação, e pode ser descrito da seguinte forma:

$$m_t = \sum_{\tau=1}^{\infty} \frac{E(d_{t+\tau})}{(1+r)^\tau} \quad (6)$$

Onde:

$m_t$  = preço da ação no tempo  $t$

$E(d_{t+\tau})$  = dividendo esperado por ação no tempo  $t + \tau$

$r$  = taxa interna de retorno sobre os dividendos esperados

A Equação 6 afirma que se no tempo  $t$  as ações de duas empresas tiverem os mesmos dividendos esperados, mas preços diferentes, a ação com um preço mais baixo terá um retorno esperado maior. Se a precificação for racional, os dividendos futuros da ação com o preço mais baixo devem ter maior risco. Fama e French (2015) extraíram desta equação implicações para as relações entre retorno esperado e rentabilidade esperada, investimento esperado e *book-to-market*. Primeiramente, eles calcularam o valor de mercado total das ações da empresa, conforme Miller e Modigliani (1961):

$$M_t = \sum_{\tau=1}^{\infty} \frac{E(Y_{t+\tau} - dB_{t+\tau})}{(1+r)^\tau} \quad (7)$$

Onde:

$M_t$  = valor total das ações da empresa no tempo  $t$

$Y_{t+\tau}$  = lucro total no tempo  $t + \tau$

$dB_{t+\tau} = B_{t+\tau} - B_{t+\tau-1}$  = variação no valor patrimonial

Em seguida, dividiram a equação de Miller e Modigliani (1961) pelo valor patrimonial da empresa no tempo  $t$ :

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau}) / (1+r)^\tau}{B_t} \quad (8)$$

Onde:

$M_t$  = preço da ação no tempo  $t$

$B_t$  = retorno do ativo livre de risco no tempo  $t$

$Y_{t+\tau}$  = dividendo esperado por ação no tempo  $t + \tau$

$dB_{t+\tau} = B_{t+\tau} - B_{t+\tau-1}$  = variação no patrimônio líquido total

$r$  = taxa interna de retorno sobre os dividendos esperados

As implicações extraídas do modelo de desconto indicam que ele pode fornecer uma relação entre os fatores de rentabilidade e investimento e o retorno esperado das ações. A relação entre o fator rentabilidade e retorno das ações é identificada quando, ao manter as demais variáveis constantes, o aumento nos dividendos esperados levam ao aumento do retorno esperado das ações. Já a relação entre retorno das ações e o fator investimento consiste no fato de que, ao manter as demais variáveis constantes, o crescimento dos investimentos implica em menores índices de rentabilidade. Fama e French (2015) argumentam que ambos os fatores também se relacionam com o *book-to-market*, pois o valor de mercado da ação responde às variações em previsões de rentabilidade e investimentos divulgados pela empresa, e empresas com maior valor de mercado tendem a apresentar retornos menores no preço das ações.

A relação entre o retorno das ações e o fator rentabilidade foi identificada empiricamente por Novy-Marx (2013). O autor apresentou evidências de que as ações de empresas com maior rentabilidade tendem a gerar retornos superiores às empresas com menor rentabilidade. Ele sugere que o fator rentabilidade pode explicar o valor das ações de forma diferente do que é sugerido pelas teorias tradicionais de precificação de ativos. Quase simultaneamente, Aharoni, Grundy e Zeng (2013) testaram empiricamente a relação entre o retorno das ações e o fator investimento. Eles argumentam que empresas com um alto fator investimento apresentam maiores retornos no valor das ações e menores fluxos de caixas

operacionais por unidade de investimento.

Diante destas evidências, Fama e French (2015) adicionaram o fator rentabilidade (*Robust Minus Weak* – RMW), calculado pela diferença entre os retornos das carteiras diversificadas de ações com rentabilidade robusta e fraca, e o fator investimento (*Conservative Minus Aggressive* – CMA) calculado pela diferença entre os retornos das carteiras diversificadas das ações de empresas de baixo e alto investimento, chamadas de conservadoras e agressivas, ao modelo original de três fatores de Fama e French (1993).

O modelo de cinco fatores de Fama e French (2015) pode ser expresso pela seguinte equação:

$$R_{it} - R_{Ft} = \alpha_i + b_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \epsilon_{it} \quad (9)$$

Onde:

$R_{it}$  = retorno do ativo  $i$  no tempo  $t$

$R_{Ft}$  = retorno do ativo livre de risco no tempo  $t$

$R_{Mt}$  = retorno da carteira de mercado no tempo  $t$

$SMB_t$  = retorno da carteira *small minus big* no tempo  $t$

$HML_t$  = retorno da carteira *high minus low* no tempo  $t$

$RMW_t$  = retorno da carteira *robust minus weak* no tempo  $t$

$CMA_t$  = retorno da carteira *conservative minus aggressive* no tempo  $t$

$\alpha_i$  = intercepto para todas as carteiras  $i$

$b_i, s_i, h_i, r_i, c_i$  = coeficientes de sensibilidade do ativo  $i$

$\epsilon_{it}$  = valor residual de média zero

Fama e French (2015) analisaram empiricamente se o modelo de cinco fatores e os modelos que incluem subconjuntos de seus fatores explicam os retornos médios das carteiras formadas para produzir grandes *spreads* em tamanho, *book-to-market*, lucratividade e investimento. Os resultados evidenciam a existência de um padrão no retorno médio das ações em relação ao fator tamanho, e um maior poder explicativo da variável valor. O retorno médio das ações aumenta à medida que o *book-to-market* aumenta, principalmente após a inclusão dos fatores de rentabilidade e investimento. Os autores concluem que o modelo de cinco fatores é importante para explicar os retornos das ações no mercado dos EUA no período de 1963 a 2013. Os fatores tamanho e valor são significativos. Já os fatores rentabilidade e investimento também

são significativos, mas em magnitude menor.

Lin (2017) testou empiricamente o modelo de cinco fatores na China para o período de 1997 a 2015. Ele concluiu que o modelo supera consistentemente o modelo de três fatores no mercado de ações chinês. Todavia, em contraste com as conclusões de Fama e French (2015), tanto os fatores de valor como de rentabilidade são importantes, enquanto o fator de investimento é considerado redundante para descrever os retornos médios na amostra. Guo *et al.* (2017) também testaram o modelo de cinco fatores no mercado de ações chinês. Eles encontraram padrões fortes de tamanho, valor e rentabilidade nos retornos médios das ações, mas padrões fracos no fator de investimento. Semelhantemente a Lin (2017), eles descobriram que o fator rentabilidade melhora significativamente a descrição do retorno médio das ações, mas o fator investimento faz contribuições marginais. Os testes de abrangência provaram que o fator de investimento é redundante para o mercado de ações chinês. Para os autores, a maior contribuição de seu trabalho é que o modelo de cinco fatores passou nos testes de eficiência de Gibbons, Ross e Shanken (1989) para a maioria dos portfólios testados.

Ali, Khurram e Jiang (2019) utilizaram uma amostra do mercado de ações paquistanês durante o período de 2003 a 2016 para avaliar e comparar o desempenho dos principais modelos de precificação de ativos baseado em fatores. Eles descobriram que o modelo de cinco fatores sempre supera o modelo de três fatores do Fama-French (1993). Equitativamente aos resultados encontrados no mercado chinês, eles observaram que a relação entre o retorno médio das ações e os fatores de valor e rentabilidade no mercado paquistanês é forte, mas fraca para os padrões de investimento. Após testarem os resultados do modelo de cinco fatores com os resultados dos modelos de quatro fatores de Carhart (1997) e os resultados do modelo de seis fatores, que acrescenta o fator impulso ao modelo de cinco fatores de Fama e French (2015), os autores concluíram que o modelo de cinco fatores supera todos os modelos analisados, incluindo o modelo de seis fatores em todas as métricas.

## 2.1.5 Modelos para Otimização

Além do CAPM, outros quatro modelos de precificação de ativos são usados nesta pesquisa como forma de otimizar os retornos dos ativos por meio da técnica de *stacked generalization*: (i) *Learning CAPM*, conforme proposto por Adrian e Franzoni (2009); (ii) *Downside CAPM*, baseado em Rutkowska-Ziarko *et al.* (2022); (iii) *Conditional CAPM* de Campbell e Vuolteenaho (2004); e (iv) Modelo de Regressão *Cross-section* (F&F) de Fama

e French (1996).

A seleção destes modelos foi realizada por meio de uma pesquisa de levantamento realizada na base de dados da Scopus. Primeiramente, conduziu-se uma busca criteriosa, empregando termos relacionados a previsão de retorno das ações, precificação de ativos e métodos de previsão. Em seguida, foram selecionados artigos com base em três critérios: (i) artigos contendo modelos de precificação de ativos com número relevante de citações na literatura financeira dos últimos anos; (ii) artigos publicados em periódicos pertencentes aos Qualis mais altos, A1 e A2, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); e (iii) artigos compreendidos por modelos estatísticos possíveis de realizar previsões de retornos das ações. Um breve resumo destes modelos é apresentado a seguir.

O modelo *Learning* CAPM (L-CAPM) proposto por Adrian e Franzoni (2009) é uma extensão do CAPM. Enquanto o CAPM tradicional prevê que os retornos esperados de um ativo são proporcionais ao seu beta, o L-CAPM reconhece que os investidores podem não ter acesso a informações perfeitas sobre os parâmetros do modelo, como o beta de um ativo específico. Em vez disso, os investidores formam suas expectativas com base em informações históricas e outras informações disponíveis no momento. Assim, o L-CAPM incorpora um processo de aprendizado pelos investidores ao longo do tempo. Mazzeu, Costa Junior e Santos (2013) apresentam uma aplicação do L-CAPM ao mercado de ações brasileiro, em que o CAPM tradicional é estimado por Mínimos Quadrados Ordinários (MQO) e o beta é dado pelo coeficiente do retorno do mercado estimado sobre a amostra inteira. Enquanto no L-CAPM, o beta é o obtido pela média da série das previsões do beta de um período à frente construídas a partir do filtro de Kalman.

O filtro de Kalman consiste em um conjunto de equações matemáticas que fornecem uma solução recursiva ótima para o método de mínimos quadrados. O objetivo dessa solução consiste em calcular um estimador linear, não viesado e ótimo do estado de um sistema em  $t$  com base na informação disponível em  $t - 1$ , e atualizar, com a informação disponível em  $t$ , tais estimativas. O filtro trabalha assumindo que o sistema pode ser descrito por meio de um modelo estocástico linear, em que tanto o erro associado ao sistema quanto a informação adicional que se incorpora no modelo têm uma distribuição normal com média zero e variância constante (Mazzeu, Costa Junior e Santos, 2013).

O modelo L-CAPM pode ser descrito pela equação do CAPM, em que o beta é calculado a partir do filtro de Kalman:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \quad (10)$$

Onde:

$E(R_i)$  = retorno esperado do ativo  $i$

$R_f$  = retorno do ativo livre de risco

$\beta_i$  = sensibilidade do ativo  $i$  construídas a partir do filtro de Kalman

$E(R_m)$  = retorno esperado do mercado

$E(R_m) - R_f$  = prêmio do risco de mercado

O modelo *Downside* CAPM (D-CAPM) foi desenvolvido por Estrada (2000) como uma alternativa ao CAPM. A principal diferença entre o D-CAPM e o CAPM tradicional é a medida de sensibilidade, denominada de *downside* beta. O D-CAPM pode ser calculado pela razão entre o semidesvio dos retornos do ativo e o semidesvio dos retornos do mercado, ou seja, pela co-semivariância dividida pela semivariância dos retornos do mercado. Para Estrada (2000), essa medida de *downside risk* possui um maior poder de explicação dos retornos dos ativos em mercados emergentes que o tradicional beta do CAPM.

O D-CAPM baseado na pesquisa de Rutkowska-Ziarko et al. (2022) se concentra especificamente na consideração dos riscos de queda. Neste modelo, os investidores ajustam suas expectativas e avaliações de risco levando em consideração não apenas o risco total de um ativo, mas também o risco de queda associado a ele. Faria e Santos (2010) realizaram a aplicação do modelo D-CAPM no mercado brasileiro. Para os autores, o modelo consegue estimar melhor o retorno que o investidor efetivamente espera por estar investindo seu capital em ativos de um país emergente.

O modelo D-CAPM baseado na pesquisa de Rutkowska-Ziarko et al. (2022) pode ser descrito pela seguinte equação:

$$R_i = R_f + (R_m - R_f) \beta^D \quad (11)$$

$$\beta^D = \frac{E(\text{Min}((r_i - \bar{r}_i), 0)) \times E(\text{Min}((r_M - \bar{r}_M), 0))}{E(\text{Min}((r_i - \bar{r}_i), 0)^2)}$$

Onde:

$R_i$  = retorno esperado do ativo

$R_f$  = retorno do ativo livre de risco

$R_m$  = retorno do mercado

$R_m - R_f =$  prêmio do risco de mercado

$\beta^D =$  beta *downside* do ativo

O modelo *Conditional CAPM* (C-CAPM) de Campbell e Vuolteenaho (2004) amplia o CAPM tradicional, incluindo a relação entre os retornos dos ativos e a dinâmica das variáveis econômicas fundamentais. No C-CAPM, os retornos dos ativos são modelados como uma função não apenas dos retornos do mercado, mas também de outras variáveis macroeconômicas relevantes. Essas variáveis macroeconômicas podem incluir taxas de juros, inflação, crescimento do PIB e outras métricas econômicas fundamentais. A principal contribuição do C-CAPM é reconhecer que os retornos dos ativos podem ser influenciados por fatores macroeconômicos específicos, além dos movimentos gerais do mercado.

Mazzeu, Costa Junior e Santos (2013) testaram o C-CAPM de Campbell e Vuolteenaho (2004) no mercado de ações brasileiro. Eles selecionaram como variáveis macroeconômicas condicionantes, o excesso de retorno do mercado (PRM), o *spread* de juros (SPREAD) e o *spread* de valor (HML). O excesso de retorno do mercado foi construído pela diferença entre o retorno do Ibovespa e o retorno da taxa Selic, utilizada como *proxy* do ativo livre de risco. O *spread* de juros foi construído pela diferença entre a taxa Selic e a taxa CDI. Finalmente, o *spread* de valor foi construído com base na diferença entre o retorno de uma carteira formada por ações de valor e o retorno de uma carteira formada por ações de crescimento. Esta variável é comumente conhecida na literatura financeira pelo fator de risco HML (*high minus low*), introduzido por Fama e French (1993).

O modelo C-CAPM utilizado nesta pesquisa pode ser descrito pela seguinte equação:

$$R^i_{t+1} = \alpha^i + \beta^{1i} R^M_{t+1} + \beta^{2i} R^M_{t+1} R_m - R_f_t + \beta^{3i} R^M_{t+1} HML_t + \beta^{4i} R^M_{t+1} Spread_t + \epsilon_{t+1} \quad (12)$$

Onde:

$R^i_{t+1}$  = retorno esperado do ativo  $i$

$\beta^i$  = beta condicional do ativo  $i$

$R^M_{t+1}$  = retorno do mercado

$PRM_t$  = prêmio de risco do mercado

$HML_t$  = *spread* de valor

$Spread_t$  = *spread* de juros

$\epsilon_{t+1}$  = erro aleatório

O modelo de regressão *cross-section* de Fama e French (F&F) utilizado nesta pesquisa também utiliza o modelo CAPM como referência. Em uma análise do mercado de ações brasileiro Varga e Brito (2016) realizaram uma regressão *cross-section*, mês a mês, dos retornos das ações em variáveis explicativas candidatas, como tamanho da empresa (ME), razão valor contábil/valor de mercado (BM), alavancagem e retorno acumulado. A análise dos autores buscou entender se características individuais como tamanho, valor contábil/valor de mercado e momento são fatores de risco fundamentais que explicam a variação transversal dos retornos médios das ações

O modelo F&F baseado na regressão *cross-section* de Fama & French (1996) utilizado nesta pesquisa pode ser descrito pela seguinte equação:

$$R_{i,t+\tau} = \lambda_{t+\tau}^0 + \lambda_{t+\tau}^M \beta_{i,t-1} + \lambda_{t+\tau}^{ME} \ln(ME_{i,t-1}) + \lambda_{t+\tau}^{BM} \ln(BM_{i,t-1}) + v_{i,t+\tau}, \forall i, \text{ para cada } (t + \tau) \quad (13)$$

Onde:

$R_{i,t+\tau}$  = retorno esperado do ativo

$t, t - 1$  = meses de julho e junho, respectivamente

$\tau = 0, 1, \dots, 11$

$t + \tau$  = período de julho  $t$  a junho  $t - 1$  do ano seguinte

$\lambda_{t+\tau}^0$  = taxa livre de risco

$\beta_{i,t-1}$  = sensibilidade do ativo em relação ao mercado

$ME_{i,t-1}$  = valor de mercado

$BM_{i,t-1} = (BE_{i,t-1} / ME_{i,t-1})$  = valor do patrimônio líquido ao valor de mercado

$\bar{\lambda}^M = (T^{-1} \sum_{t=1}^T \lambda_t^M) > 0$

## 2.2 COMBINAÇÃO DE PREVISÕES

### 2.2.1 Métodos para Combinar Previsões

O artigo seminal de Bates e Granger (1969) sobre combinação de previsões é um dos trabalhos mais influentes sobre métodos de previsão já publicado. No artigo, os autores

apresentam provas matemáticas de que a previsão combinada é sempre mais precisa do que qualquer previsão individual. Os autores são considerados os primeiros a sistematizar um método para combinar previsões. Eles sugerem um esquema de combinação linear, cujos coeficientes ou pesos são calculados utilizando a matriz de covariância de erros de previsões individuais. O método baseia-se na ideia de que a previsão combinada deve ser a média ponderada das previsões dos diferentes modelos, e seus pesos, determinados pelos erros quadráticos médios, de modo a maximizar a precisão da previsão combinada. A equação a seguir fornece a combinação de duas previsões  $F_c$ .

$$F_c = wF_1 + (1 - w)F_2 \quad (14)$$

Onde:

$F_c$  = valor previsto das previsões combinadas

$F_1$  e  $F_2$  = previsões individuais a serem combinadas

$w$  = peso proporcional dado a previsão 1

$1 - w$  = peso proporcional dado a previsão 2

O valor ótimo de  $w$  é encontrado pela minimização da variância dos erros da combinação da previsão  $F_c$ , que é dada pela Equação (11):

$$\sigma_c^2 = w^2\sigma_1^2 + (1 - w)^2\sigma_2^2 + 2\rho w\sigma_1(1 - w)\sigma_2 \quad (15)$$

Onde:

$\sigma_1^2$  e  $\sigma_2^2$  = variâncias dos erros das previsões a serem combinadas

$\rho$  = coeficiente de correlação dos erros das previsões

Essa técnica ficou conhecida como técnica da variância mínima. O valor mínimo de  $\sigma_c^2$  ocorre quando  $w$  assume:

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (16)$$

Nos casos em que os erros não são correlacionados  $\rho = 0$ ,  $w$  é reduzido a:  $w = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$ .

O método da variância mínima proposto por Bates e Granger (1969) forneceu as bases para combinar previsões em diferentes áreas do conhecimento (Clemen, 1989). Em finanças, o método tem sido utilizado, dentre outras aplicações, na precificação de ativos e previsão do

prêmio de capital. Para os autores, as principais contribuições do método incluem: (i) melhoria da precisão da previsão – a combinação das previsões ajuda a reduzir a variância, levando a estimativas mais precisas e, em muitos casos, mais acuradas; e (ii) ponderação ótima – determinação adequada dos pesos para cada previsão individual ao combiná-las. A ponderação ótima, considera o histórico de desempenho de cada previsão para atribuir pesos relativos.

Todavia, Clemen e Winkler (1986) e Armstrong (2001) argumentam que o método apresenta algumas limitações que incluem, mas não se limitam a

- a) Pressupostos da independência – o método assume que as previsões feitas pelos diferentes modelos são independentes entre si. No entanto, na prática, os modelos muitas vezes estão correlacionados, o que pode afetar os resultados da combinação.
- b) Ponderação ótima não garantida – a abordagem de atribuir pesos ótimos para cada previsão individual baseia-se em históricos de desempenho passados. Entretanto, o desempenho passado pode não ser um indicador confiável do desempenho futuro, e a escolha dos pesos pode ser subjetiva e sensível aos dados disponíveis.
- c) Diversidade limitada de modelos – o método pressupõe que existem diferentes modelos independentes disponíveis para combinar. Na prática, pode ser desafiador obter uma ampla diversidade de modelos relevantes e independentes para combinação.
- d) Sensibilidade às previsões individuais – a eficácia do método pode ser altamente sensível à precisão das previsões individuais. Se algumas previsões individuais forem extremamente imprecisas ou enviesadas, elas podem ter um impacto desproporcional nos resultados combinados.
- e) Mudança de estrutura do modelo – caso a estrutura subjacente dos modelos mude ao longo do tempo, a combinação de previsões pode perder sua eficácia, pois os modelos podem passar a capturar diferentes fenômenos.
- f) Requerimento de histórico de previsões – o método depende de históricos de previsões passadas para atribuir pesos ótimos. Isso pode ser um problema quando os modelos são novos ou quando não há histórico suficiente de previsões passadas.

Newbold e Granger (1974) ampliaram o número de modelos a serem combinados, passando de dois para  $n$ , mas mantiveram todas as suposições feitas inicialmente por Bates e Granger (1969). Assim como no artigo original, Newbold e Granger (1974) concluíram que as correlações poderiam ser ignoradas na estimação dos pesos, pois elas têm pouca influência nos resultados. Winkler e Makridakis (1983) corroboraram esta ideia.

Granger e Ramanathan (1984) também contribuíram diretamente para ampliar o argumento clássico para combinar previsões. Os autores apresentaram uma revisão das

principais técnicas de combinação de previsões em que discutem os diferentes tipos de dados que podem ser combinados, incluindo dados de séries temporais, dados de regressão e dados de julgamentos de especialista. Os autores propuseram diferentes métodos para escolher os pesos de combinação, que incluem métodos baseados na precisão histórica das previsões, na correlação entre as previsões e na estrutura de dependência das séries temporais. Eles apresentaram uma série de estudos empíricos que corroboram a ideia de que a combinação de previsões pode melhorar significativamente a precisão da previsão final, especialmente quando os métodos individuais têm diferentes pontos fortes e fracos. As principais conclusões de Granger e Ramanathan (1984) foram que: (i) a combinação de previsões pode melhorar a precisão da previsão final; (ii) existem diferentes técnicas de combinação de previsões, incluindo técnicas lineares e não lineares; e (iii) o método de combinação mais adequado depende dos objetivos da combinação e do tipo de dados que estão sendo combinados.

Após a disseminação do método de variância mínima proposto por Bates e Granger (1969), outros métodos começaram a ganhar destaque na literatura como o método de combinação simples de Gupta e Wilton (1977); o método de combinação linear por Mínimos Quadrados Ordinários (MQO) de Granger e Ramanathan (1984), o método de combinação por pesos móveis de Terui e Van Dijk (2002), e o método de combinação por *stacked generalization* de Wolpert (1992).

No tocante ao método de combinação simples, Gupta e Wilton (1977) propõe a utilização de uma média aritmética simples para combinar as previsões, em que para  $n$  previsões de modelos individuais a previsão combinada será:

$$F_c = \frac{F_1 + F_2 + \dots + F_n}{n} \quad (17)$$

Por ser considerado de fácil aplicação e obter resultados empíricos tão bons quanto outras técnicas de combinação complexas, o método de combinação simples tem sido defendido por diversos pesquisadores como Winkler e Makridakis (1983), Clemer e Winkler (1986) e Armstrong e Collopy (1998). Para Winkler (1989), uma vantagem, é que o método não requer a estimação de parâmetros e, por esse motivo, não apresenta problemas de estimação. Winkler e Makridakis (1983) argumentam que utilizar a média aritmética simples das previsões é sistematicamente melhor do que utilizar um modelo ruim ou uma técnica de previsão pobre, pois apresenta consideráveis benefícios práticos em termos de melhoria de acurácia e diminuição da variabilidade.

Quanto ao método de combinação linear por MQO, Granger e Ramanathan (1984) propõe uma combinação linear sem restrição aos pesos, em que os parâmetros dos pesos são estimados por MQO, e uma constante diferente de zero é adicionada ao modelo. Neste método, a previsão combinada é obtida por meio da seguinte equação:

$$F_c = w_1 F_1 + w_2 F_2 + \dots w_n F_n + k \quad (18)$$

Onde:

$F_c$  = valor previsto das previsões combinadas

$F_n$  = previsões individuais a serem combinadas

$w_n$  = peso proporcional (estimados por MQO) das previsões individuais

$k$  = constante diferente de zero

Granger e Ramanathan (1984) apresentaram evidências empíricas de que as combinações de previsões produzem, individualmente, um resultado melhor do qualquer outro modelo de previsão original. Eles concluíram que o método de combinação linear por MQO produz uma previsão combinada com menor erro quadrático médio que os modelos individuais. Estes resultados foram corroborados por Holmen (1987), que testou empiricamente a abordagem de combinação de previsões lineares por MQO. O autor observou a importância de combinar previsões de lucros de curto prazo e corroborou as evidências empíricas de Granger e Ramanathan (1984) após simular dados de diferentes modelos.

Terui e Van Dijk (2002) propuseram um método de combinação por pesos móveis, em que os coeficientes variam ao longo do tempo. Para eles, o método obtém vantagem ao modelar tanto modelos individuais lineares quanto não lineares. Tal como Granger e Ramanathan (1984), os autores introduzem um termo constante à equação e não impõe restrições aos parâmetros, conforme as equações a seguir:

$$\begin{aligned} F_c &= \alpha_t + \alpha_t^1 F_t^1 + \alpha_t^2 F_t^2 + \dots \alpha_t^n F_t^n + \epsilon_t; & \epsilon_t &\sim N(0, \eta^2) \\ \alpha_t &= \alpha_{t-1} + \mu_t; & \mu_t &\sim N(0, \Sigma) \end{aligned} \quad (19)$$

Onde:

$F_c$  = valor previsto das previsões combinadas

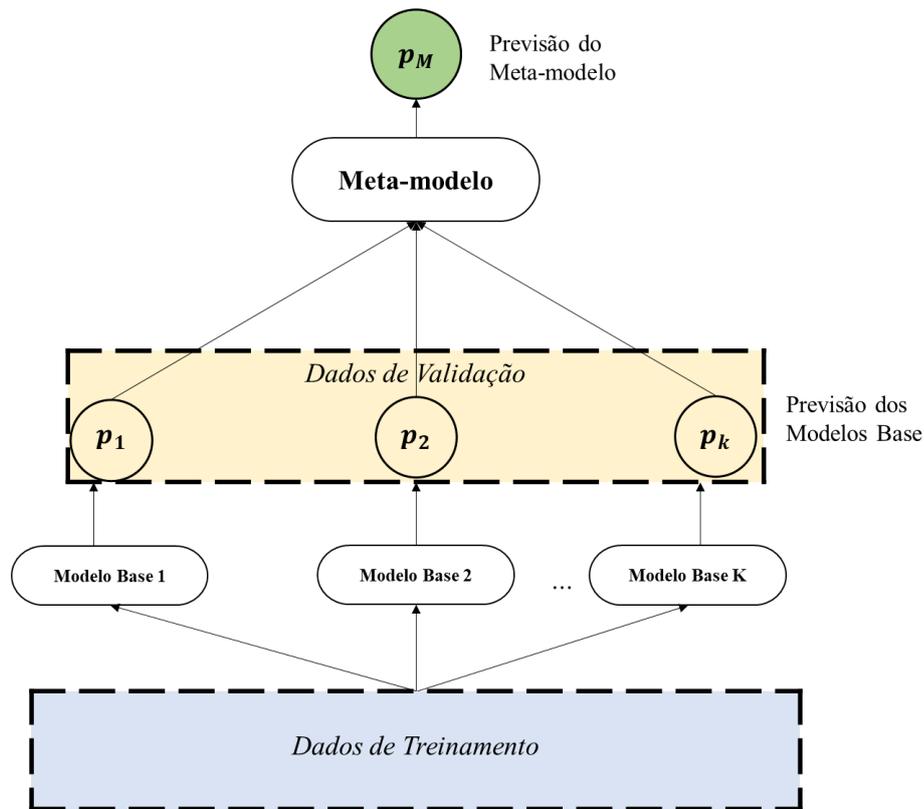
$F_t, F_t^1 \dots F_t^n$  = previsões gerados pelos modelos individuais

$\alpha_t, \alpha_t^1 \dots \alpha_t^n$  = vetores da regressão

O modelo de Terui e Van Dijk (2002) pode ser interpretado como um modelo de espaço de estados, em que a Equação (15) é a equação de medida que define a distribuição de  $F_t$ ,  $t \geq 1$ ; e a Equação (16) é a equação de estado que define a distribuição de  $\alpha_t$  para cada  $t$ ,  $t \geq 1$ . Para Wang *et al.* (2023) ao ser interpretado como um modelo de espaço de estados, o modelo generalizado pode ser estimado usando a atualização do filtro de Kalman.

O método de combinação por *stacked generalization* foi desenvolvido por Wolpert (1992) e pertence à família de técnicas de aprendizado conjunto (*ensemble learning*) no âmbito do *machine learning*. O autor apresentou o método como uma técnica para melhorar o desempenho preditivo de modelos de *machine learning*, combinando as previsões de vários modelos em um nível mais alto, permitindo a criação de um novo modelo que aprende a combinar as previsões de modelos base e aproveita as forças individuais de cada modelo. Para o autor, a estrutura ideal do *stacked generalization* precisa ter dois níveis. No primeiro nível, vários modelos base são treinados com os dados de treinamento. No segundo nível, um meta-modelo é treinado usando as previsões dos modelos base como entrada. O objetivo é que o meta-modelo aprenda a ponderar as previsões dos modelos base de maneira a otimizar a performance geral.

Wolpert (1992) acredita que o método pode ajudar a reduzir tanto o viés quanto a variância dos modelos individuais, pois modelos diferentes podem ter diferentes fontes de erro e o *stacked generalization* pode combinar suas previsões para reduzir esses erros, levando a uma previsão final mais precisa e geralmente com menor erro total. O autor enfatizou que os modelos base devem ser independentes e produzir previsões diferentes para diferentes exemplos. Se os modelos base forem correlacionados ou altamente dependentes, a técnica pode não ser tão eficaz. No *stacked generalization*, vários algoritmos de *machine learning* são usados em conjunto com o objetivo de alcançar um desempenho preditivo superior, que não poderia ser alcançado por um método de previsão individual. O método agrega diversas previsões de modelos de base em um meta-modelo para obter maior precisão preditiva. A ideia básica é que, ao invés de usar funções elementares, como a média aritmética simples para combinar as previsões de todos os preditores, um modelo generalizado seja treinado para realizar essa agregação. A Figura 1 apresenta a estrutura geral do *stacked generalization* proposta por Zhao e Cheng (2022).



**Figura 1 - Estrutura do Stacked Generalization**

Fonte: Adaptado de Zhao e Cheng (2022)

Em uma análise da proposta de Wolpert (1992), Zhao e Cheng (2022) observaram que o *stacked generalization* pode ser dividido em quatro etapas principais: (i) treinamento dos modelos base – diferentes algoritmos de *machine learning* são selecionados como modelos base. Isso pode incluir regressão linear, árvores de decisão, redes neurais, dentre outros. Cada modelo é treinado em um conjunto de dados de treinamento; (ii) validação dos modelos base – os modelos base são avaliados e ajustados usando um conjunto de dados de validação. Isso ajuda a otimizar os hiperparâmetros e a selecionar os melhores modelos base para inclusão no processo de empilhamento; (iii) treinamento do meta-modelo – além dos modelos base, um meta-modelo ou modelo de nível superior é criado. Esse meta-modelo é treinado usando as previsões realizadas pelos modelos base como entradas. Pode-se usar uma variedade de algoritmos para o meta-modelo, como regressão linear ou regressão logística; (iv) combinação e previsão – uma vez que o meta-modelo é treinado, ele é utilizado para fazer previsões sobre um conjunto de dados de teste. As previsões dos modelos base também são combinadas pelo meta-modelo para gerar uma previsão final. Essa combinação pode envolver a ponderação das previsões dos modelos base de acordo com as habilidades de previsão do pesquisador.

Para os autores, esse procedimento pode ser visto como uma versão mais sofisticada da validação cruzada, em que o conjunto de dados disponível é dividido em partes distintas de treinamento e teste, permitindo que o modelo seja treinado em uma fração do conjunto e testado em outra. Na validação cruzada, esse procedimento é repetido várias vezes, alternando as frações utilizadas para treinamento e teste, o que resulta em uma avaliação mais robusta e confiável do desempenho do modelo (Laan, Dudoit 2003, e Vaart, 2006). A validação cruzada é especialmente valiosa em situações em que a disponibilidade de dados é limitada, além de auxiliar na detecção e mitigação de problemas como o *overfitting*.

Pesquisadores têm descoberto que um dos maiores desafios em realizar previsões com abordagens computacionais mais sofisticadas, como o *stacked generalization*, é o aumento do risco de *overfitting*. Rasekhschaffe e Jones (2019) consideram que o *overfitting* ocorre quando um modelo captura ruídos em vez de sinais. Neste caso, o modelo com sobreajuste têm bom desempenho na amostra, mas pouca previsibilidade quando aplicado a dados nunca vistos. Na concepção dos autores, embora as técnicas de *machine learning* possam revelar padrões sutis em dados anteriores, o *overfitting* apresenta um grande desafio. Na precificação de ativos, por exemplo, as relações entre os fatores de risco e os retornos dos ativos são frequentemente ruidosas e existem muitos fatores potenciais, o que aumenta a dimensionalidade do problema. Outros autores como Zhao e Cheng (2022) defendem que combinação de previsões e os processos de validação cruzada são os dois maiores aliados na luta contra o *overfitting*.

A importância do processo de validação da amostra foi estudada por alguns autores. Laan, Dudoit e Vaart (2006) forneceram fundamentos teóricos do uso de validação cruzada na seleção de um modelo ótimo entre muitos modelos candidatos. Shahhosseini, Hu e Pham (2022) propuseram um algoritmo baseado em validação cruzada para o ajuste de hiperparâmetros e ponderação de modelos. Zhao e Cheng (2022) utilizaram um modelo de validação cruzada comum, composto pelas etapas de treinamento, validação e teste, baseado em Wolpert (1992), para prever o prêmio de capital fora da amostra para uma variedade de modelos.

Em síntese, desde o artigo seminal de Bates e Granger (1969) uma variedade de métodos para combinar previsões foram publicados na literatura. Wang *et al.* (2023) realizaram uma extensa revisão da literatura publicada sobre o tema no período de 1969 a 2021. Os autores identificaram uma grande variedade de métodos, que além dos citados anteriormente, incluem, mas não se limitam a: combinações qualitativas baseadas em opinião de especialistas (Winkler e Makridakis, 1983); combinações simples de previsão pontual (Bunn, 1985; Clemen e Winkler, 1986; Stock e Watson, 2003); combinações com pesos

baseados em critérios de informação (Akaike, 1974); combinações com pesos bayesianos (Öller, 1978; Bunn, 1975) e combinações não lineares (Donaldson e Kamstra, 1996). A seguir são apresentados alguns estudos empíricos que se dedicaram a investigar a utilização dos métodos de combinação de previsões no retorno de ações.

## **2.2.2 Estudos Empíricos sobre Combinação de Previsões no Retornos de Ações**

Von Holstein (1972) foi um dos primeiros a testar empiricamente a relevância dos métodos de combinação de previsões no retorno médio das ações. Ele investigou o efeito da combinação de previsões após diferentes participantes quantificarem, em termos probabilísticos, suas crenças relacionadas ao desempenho do mercado de ações. O experimento foi realizado com uma amostra de 100 participantes, incluindo banqueiros, especialistas do mercado de ações, estatísticos, professores e estudantes de negócios. Os resultados mostraram que, em geral, a combinação de previsões produz resultados melhores do que as previsões individuais. As evidências indicaram que as pessoas foram capazes de quantificar suas crenças de forma razoável, embora houvesse uma grande variação no desempenho individual, ora mitigada pelas previsões combinadas. O autor chegou a duas conclusões principais: a primeira é que a combinação de previsões pode levar em consideração diferentes perspectivas e experiências dos participantes. A segunda, é que o método pode ajudar a reduzir a variabilidade das previsões individuais por meio de uma previsão agregada, corroborando as ideias propostas no artigo seminal de Bates e Granger (1969).

A abordagem de Von Holstein (1972) ganhou força, e Virtanen e Yli-Olli (1987) utilizaram um modelo composto de previsões lineares e ARIMA (*Autoregressive Integrated Moving Average*) para estudar o comportamento dos preços das ações na bolsa de valor de Helsinque na Finlândia para o período de 1975 a 1984. Eles testaram empiricamente se o índice geral de preços, mensal e trimestral, em um mercado de baixa liquidez, como o finlandês, poderia ser previsível. O principal argumento deles é que em mercados de baixa liquidez as previsões baseadas em modelos tradicionais, como o modelo de médias móveis, podem ser imprecisas. Pois, os preços das ações são mais propensos a flutuar devido a eventos inesperados, como notícias ou mudanças na política econômica.

Para resolver esse problema, Virtanen e Yli-Olli (1987) propuseram a combinação linear entre o modelo ARIMA e outros três modelos lineares baseados em MQO propostos por

Granger e Ramanathan (1984). Os três modelos consistiam em: (i) sem termo constante e com pesos restritos; (ii) sem termo constante e com pesos irrestritos; e (iii) com termo constante e pesos irrestritos. Os resultados mostraram que as previsões mensais e trimestrais de quase todos os modelos individuais apresentaram um bom desempenho, com as previsões mensais obtendo uma leve vantagem. Como nenhuma medida de erro médio foi estatisticamente significativa, os autores concluíram que todas as previsões individuais foram imparciais. No entanto, as evidências revelaram que as previsões combinadas produziram resultados superiores com estatísticas de erros ainda mais baixas. Os autores sugerem que o método de combinação de previsões pode ser uma ferramenta eficaz para melhorar a precisão da previsão de preços de ações em mercados de baixa liquidez, como o finlandês.

Anos depois, Timmermann (2008) estudou a previsão de retorno de ações do mercado dos EUA para o período de 1959 a 2005. Em sua análise, ele propôs uma abordagem de combinação de previsões adaptativa, em que considerou um conjunto de modelos que incluem o retorno sobre as carteiras de valor e a diferença entre o retorno de pequenas e grandes ações – SMB, estudada por Fama e French (1992, 1993). O autor argumenta que a previsibilidade de retornos de ações é um fenômeno elusivo, que pode ser observada apenas por períodos relativamente curtos. Os resultados empíricos dos modelos de previsão individuais sugerem que, nenhum modelo supera consistentemente a média histórica durante longos períodos e que qualquer previsibilidade de retorno é, na melhor das hipóteses, de curta duração e provavelmente deteriorar-se-á bastante rapidamente. Para o autor, as abordagens de previsão que utilizam sempre o mesmo modelo têm, portanto, pouca probabilidade de serem bem-sucedidas. Neste sentido, se faz necessário utilizar uma abordagem combinada de previsões.

Para comprovar a inconsistência do retorno das ações no longo prazo, Timmermann (2008) utilizou a validação cruzada, dividindo o extenso período de sua amostra em duas partes. A primeira foi considerada como período de estimativa inicial, de dezembro de 1959 a dezembro de 1969. A segunda foi o período de avaliação fora da amostra, de janeiro de 1970 a dezembro de 2005. Ele também considerou duas abordagens de estimativa distintas: (i) janela de expansão – que utilizou a estimativa recursiva com dados de dezembro de 1959 até o momento da previsão, para gerar uma série de previsões um passo à frente. Desta forma, a primeira previsão foi gerada para janeiro 1970, utilizando dados de dezembro de 1959 a dezembro de 1969. No mês seguinte, a janela de dados é expandida para incluir também janeiro de 1970, e os parâmetros dos modelos de previsão são novamente estimados, e depois usados para prever os retornos das ações para fevereiro de 1970, e assim por diante até o final da amostra; e (ii) janela contínua – essa abordagem utilizou uma janela de comprimento fixo dos

dez anos de dados mais recentes (120 observações mensais) para estimar os parâmetros dos modelos de previsão e, em seguida, prever retornos para o próximo período, condicionados a essas estimativas de parâmetros.

Os resultados empíricos do modelo de combinação adaptativa de Timmermann (2008) mostram que, no geral, a combinação de previsões pode melhorar significativamente a precisão da previsão individual, especialmente em períodos de baixa previsibilidade. Na maioria dos casos, a abordagem de combinação adaptativa é capaz de produzir previsões com valores de erro quadrático médio inferiores à média histórica. Os melhores resultados foram gerados nas janelas de curta duração, de 18 ou 36 meses. O autor conclui que os resultados indicam o potencial para abordagens de combinação adaptativa, com base no princípio de que a previsibilidade do retorno será provavelmente de curta duração. Para ele, diferentes abordagens poderiam ser utilizadas, mas ainda seriam necessárias mais evidências empíricas.

Essa visão foi corroborada por Rapach, Strauss e Zhou (2010), que investigaram a utilização da combinação de previsões no prêmio de capital fora da amostra no mercado dos EUA para o período de 1947 a 2005. O estudo deles, e posteriormente diversos outros autores, foi baseado no trabalho de Welch e Goyal (2008), que estudaram o desempenho de um conjunto de variáveis macroeconômicas sugeridas pela literatura acadêmica como boas preditores do prêmio de ações. Os métodos de combinação utilizados por Rapach, Strauss e Zhou (2010) foram organizados em duas classes: a primeira utiliza esquemas de médias simples, como média aritmética e mediana. A segunda baseia-se em Stock e Watson (2004), em que os pesos da combinação formados no tempo  $t$ , são funções do desempenho histórico de previsão dos modelos individuais durante o período de validação fora da amostra. Os autores também consideraram outros métodos mais complexos, em que os pesos combinados são selecionados de forma mais elaborada usando o ajuste do modelo na amostra, no entanto, estes modelos tiveram um desempenho ruim em comparação com os esquemas mais simples. Para os autores, este princípio está de acordo com a literatura de previsão, que indica que os métodos de combinação simples normalmente superam os métodos mais complicados (Timmermann, 2006).

As principais descobertas de Rapach, Strauss e Zhou (2010) reforçam as conclusões de Welch e Goyal (2008), ao demonstrar que é muito difícil identificar variáveis econômicas individuais capazes de gerar previsões confiáveis fora da amostra do prêmio de capital. Durante a análise não foi possível identificar uma única variável dentre as 15 estudadas que produzisse um  $R^2$  positivo ao longo de cada um dos períodos examinados fora da amostra. No entanto, a combinação de previsões se mostrou um método muito eficaz para prever o prêmio de capital,

ao superar a média histórica por margens estatisticamente e economicamente significativas para uma variedade de períodos fora da amostra. Para os autores, há duas explicações empíricas para os benefícios da combinação de previsões: (i) a combinação de previsões incorpora informações de numerosas variáveis econômicas, ao mesmo tempo que reduz substancialmente a volatilidade das previsões; (ii) as previsões combinadas estão ligadas à economia real.

Lin, Wu e Zhou (2018) propuseram uma abordagem de combinação iterada para prever o retorno dos títulos corporativos dentro e fora da amostra no mercado dos EUA para o período de 1973 a 1998. A primeira grande questão abordada pelos autores foi saber quais variáveis macroeconômicas teriam maior poder preditivo para o retorno dos títulos corporativos. Para resolver tal questão, eles utilizaram um conjunto abrangente de 27 variáveis macroeconômicas típicas na literatura. A segunda questão foi determinar uma forma sobre como combinar este grande conjunto de variáveis macroeconômicas para obter ótimas previsões de retornos. Eles sugeriram um modelo de combinação iterada, em que tanto o método de combinação com médias simples de Rapach, Strauss e Zhou (2010) quanto o método de variância mínima de Bates e Granger (1969) são combinados com a previsão média da amostra histórica para produzir uma nova previsão combinada com o objetivo de melhorar significativamente as previsões originais. Para os autores, do ponto de vista econométrico, as novas previsões combinadas são um caso especial do quadro geral de combinação estabelecido por Granger e Ramanathan (1984).

Os resultados empíricos de Lin, Wu e Zhou (2018) apontam que, no geral, o método de combinação iterada melhora a previsibilidade dos retornos de títulos corporativos em termos de precisão da previsão, ganhos estatísticos e ganhos econômicos. Os autores descobriram que os métodos de combinação por média simples e variância mínima parecem ter o melhor desempenho entre os diferentes métodos de combinação analisados. No entanto, os métodos que utilizam a combinação iterada apresentaram resultados ainda mais surpreendentes. Cada uma das combinações iteradas apresentaram um menor erro quadrático médio (*mean squared error* - MSE) em comparação aos métodos subjacentes; um  $R^2$  substancialmente mais elevado do que a sua respectiva contraparte combinada; ganhos de utilidade muito maiores do que outros modelos, e consideravelmente mais elevados do que os que foram relatado por Rapach, Strauss e Zhou (2010) para o mercado de ações, sugerindo que existe um valor econômico substancial na utilização de um grande conjunto de preditores e na metodologia proposta para previsões de retorno dos títulos.

O enfoque de combinação iterada de Lin, Wu e Zhou (2018) despertou o interesse de Zhang *et al.* (2020), que exploraram a eficácia de métodos de combinação de previsões para o

retorno das ações no mercado dos EUA no período de 1926 a 2016. Para tornar seus resultados comparáveis com estudos anteriores, eles utilizaram os dados mensais atualizados de Welch e Goyal (2008). Os autores compararam vários modelos sofisticados e populares no contexto da previsão de retorno de ações, e introduziram um novo método de combinação que incorpora a incerteza do modelo, a instabilidade dos parâmetros e a contração simultaneamente. Para eles, esse novo método introduz três inovações: (i) utiliza o método de combinação iterada de Lin, Wu e Zhou (2018), mas em vez de usar uma combinação simples de modelos univariados, utiliza uma combinação da média histórica com vários modelos sofisticados; (ii) utilizam pesos iguais ou pesos baseados no desempenho das próprias combinações. Isso torna a abordagem uma extensão natural do método de combinação simples Rapach, Strauss e Zhou (2010); e (iii) para acomodar a instabilidade dos parâmetros, eles estimaram os modelos sofisticados usando o método de previsão de janelas médias proposto por Pesaran e Timmermann (2007), em vez do tradicional esquema de janelas de expansão recursiva.

As evidências empíricas sugerem que os métodos de combinação de previsões podem melhorar significativamente a precisão das previsões de retorno das ações. Zhang *et al.* (2020) descobriram que, quando os modelos são estimados tradicionalmente, usando janelas de expansão recursiva e sem levar em consideração a instabilidade dos parâmetros, nenhum dos modelos sofisticados ofereceu um  $R^2$  fora da amostra significativamente positivo. Os resultados basicamente confirmaram que a combinação simples de preditores individuais é superior. Além disso, quando foi incorporada a instabilidade dos parâmetros usando o método de previsão de janelas médias, todos os modelos sofisticados reduziram modestamente seus erros quadráticos médios de previsão (*mean squared forecast error* – MSFE), com exceção do LASSO e do *Elastic Net*. Por outro lado, os modelos mais sofisticados melhoraram significativamente quando foram combinados com a média histórica. Uma combinação ponderada igual da média histórica com a regressão preditiva multivariada padrão estimada usando o método de janelas médias, por exemplo, é significativamente mais precisa do que o modelo multivariado preditivo individual em períodos de instabilidade dos parâmetros, atingindo um  $R^2$  fora da amostra estatisticamente significativo de 1,10% e ganhos de utilidade anuais de 2,34%.

Rapach e Zhou (2020) introduziram o *C-Enet* (*Combination Elastic Net*), que consiste numa abordagem para previsão de retorno de ações que combina os métodos *Elastic Net* e a previsão de combinação simples. O *C-Enet* foi introduzido com o objetivo de refinar a previsão de retorno de combinação simples de Rapach, Strauss e Zhou (2010). O argumento principal é que os métodos de *machine learning* podem ser mais eficazes do que os métodos tradicionais de previsão de séries temporais, pois são capazes de capturar as relações lineares e não lineares

entre as variáveis. A operacionalização do *C-Enet* pode ser realizada em três etapas: (i) implementar as regressões lineares univariadas e coletar os resultados das previsões um passo à frente; (ii) ajustar um modelo Enet com essas previsões no conjunto de validação e selecionar os modelos univariados com coeficientes Enet diferentes de zero; e (iii) as previsões selecionadas pela Enet são simplesmente combinadas, ou seja, a média das previsões selecionadas é obtida e usada como previsão final.

Eles aplicaram o *C-Enet* a um conjunto de dados de retornos de ações dos EUA para o período de 1927 a 2018, e descobriram que além de melhorar substancialmente a previsão de combinação simples, o método é capaz de melhorar o desempenho de outros métodos de previsão tradicionais de duas maneiras: (i) capturando as relações lineares entre as variáveis, e; (ii) evitando o *overfitting*, que é considerado pela literatura como um dos grandes problemas do *machine learning*. As evidências sugerem que a utilização do *C-Enet* gerou previsões de retorno de ações mais precisas. O modelo proporcionou uma melhoria substancial do desempenho do portfólio, produzindo maiores retornos excedentes anualizados, menor volatilidade e índices de Sharpe significativamente melhorados.

Em uma abordagem mais recente, Zhao e Cheng (2022) estenderam o emprego de métodos de combinação de previsões com *machine learning* na previsão de retornos das ações. Eles empregaram uma abordagem de *stacked generalization* para refinar uma variedade de modelos de previsão de retornos de ações individuais, lineares e não lineares, na previsão de retorno do prêmio de ações no mercado dos EUA para o período de 1927 a 2018.

Adicionalmente à aplicação da abordagem convencional de *stacked generalization*, conforme sugerido por Wolpert (1992), na qual o meta-modelo é sobreposto aos modelos base, os autores voltaram seu foco para a significância da validação no processo de escolha de modelos. Ao invés de descartar indiscriminadamente todos os modelos base no meta-modelo, eles conduziram uma avaliação do desempenho dos modelos base em um conjunto de validação não visível, selecionando posteriormente vários dos modelos mais eficientes. Em seguida, optaram por diferentes abordagens de complexidade crescente como candidatas para o meta-modelo. Formaram um conjunto de hiperparâmetros e escolheram o conjunto mais eficaz, analisando o desempenho do meta-modelo correspondente no conjunto de validação. Após a definição dos hiperparâmetros, recalibraram o meta-modelo com os melhores valores e realizaram previsões fora da amostra. Esse processo tornou o modelo adaptável às constantes mudanças na série de retornos de ações.

Eles descobriram que, com uma estrutura simples, o *stacked generalization* pode superar a média histórica do prêmio de ações e de vários outros métodos populares em termos

de previsão e medidas de desempenho dentro e fora da amostra, de forma consistente. Mais importante, eles descobriram que os ganhos de previsão fora da amostra do *stacked generalization* são especialmente evidentes durante movimentos de mercado de desaceleração extrema. Os autores concluíram que, no geral, o *stacked generalization* pode gerar melhorias substanciais na previsibilidade de retorno do prêmio de ações. Os autores resumiram as contribuições de seu artigo da seguinte forma: (i) o *stacked generalization* ajuda a melhorar as previsões de combinação tradicionais, permitindo que os componentes da combinação sejam formas funcionais flexíveis; (ii) o *stacked generalization* funciona melhor do que a média simples e muitos outros métodos populares existentes; (iii) o *stacked generalization* enfatiza a importância de encontrar modelos apropriados pela abordagem de validação; e (iv) os resultados encontrados contribuem para uma longa linha de pesquisa sobre a previsibilidade fora da amostra de retornos de ações e demonstra as vantagens de modelos de *machine learning* e procedimentos de validação.

Em síntese, o desenvolvimento contínuo e a diversificação de abordagens nos métodos de combinação de previsões para o retorno de ações revelam uma evolução consistente no campo da previsão financeira. Desde as análises pioneiras de Von Holstein (1972) até as inovações contemporâneas como o *C-Enet* de Rapach e Zhou (2020) e o *stacked generalization* de Zhao e Cheng (2022), a pesquisa tem buscado constantemente aprimorar a precisão das previsões, reconhecendo a complexidade e a variabilidade inerentes aos mercados. A incorporação de técnicas avançadas, como *machine learning* e métodos iterativos, destaca o esforço em lidar com desafios dinâmicos. Esses avanços reforçam a noção de que, apesar da natureza elusiva da previsibilidade de retornos, abordagens combinadas adaptativas e diversificadas continuam a desempenhar um papel crucial na busca por melhores estratégias de previsão no ambiente financeiro.

### 3. PROCEDIMENTOS METODOLÓGICOS

---

Este capítulo oferece uma análise abrangente do desenho e execução da pesquisa, destacando os procedimentos metodológicos adotados. Inicialmente, é apresentado o delineamento em torno do paradigma de pesquisa, incluindo a classificação quanto aos objetivos e os métodos selecionados. Em seguida, são discutidos os dados, a população e a amostra da pesquisa. Na sequência, são detalhados os procedimentos de operacionalização da técnica de *stacked generalization* nos modelos de precificação de ativos, abrangendo as janelas de dados, os modelos e meta-modelos utilizados, além das metodologias de avaliação das previsões. Por fim, são apresentadas as hipóteses do estudo.

#### 3.1 DELINEAMENTO DA PESQUISA

Saccol *et al.* (2012) esclarecem que as pesquisas científicas devem ser delineadas em torno de três elementos centrais: (i) o paradigma de pesquisa – qualitativo (fenomenológico) ou quantitativo (positivista); (ii) a classificação da pesquisa quanto aos objetivos – explorar, descrever ou explicar, e; (iii) os métodos ou estratégias de pesquisa escolhidos.

No que se refere ao paradigma, esta pesquisa é de natureza quantitativa, pois adota um paradigma positivista ao empregar técnicas estatísticas para fornecer uma investigação robusta da relação entre a técnica de *stacked generalization*, os modelos de precificação de ativos, e os retornos das ações no mercado brasileiro. Conforme Lakatos e Marconi (2010), este tipo de estudo contém hipóteses explícitas, derivadas da teoria, e que devem ser verificadas.

Em relação aos objetivos, a pesquisa classifica-se como exploratória e descritiva. Exploratória porque estuda a combinação de previsões por *stacked generalization* – um tema ainda pouco explorado na literatura financeira sobre previsão de retorno das ações. Descritiva porque se propõe a compreender a relação comparativa entre técnica de *stacked generalization* e os métodos de precificação de ativos de capital. Gil (2002) afirma que as pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar uma visão geral acerca de determinado fato. Normalmente, quando o tema escolhido é pouco explorado. Enquanto as pesquisas descritivas têm como objetivo principal, dentre outras finalidades, o

estabelecimento de relações entre variáveis.

No que diz respeito aos métodos, a pesquisa caracteriza-se como *ex post facto*, tendo em vista que a análise das variáveis foi realizada com dados secundários derivados de fatos já ocorridos. A tradução literal da expressão *ex post facto* é “a partir do fato passado”. Isso significa que, neste tipo de pesquisa, o estudo foi realizado após a ocorrência de alterações na variável dependente, no curso natural dos acontecimentos (Gil, 2002).

## 3.2 DADOS, POPULAÇÃO E AMOSTRA DA PESQUISA

A população da pesquisa compreende todas as empresas com ações listadas na bolsa de valores brasileira (B3) no período de janeiro de 2003 a dezembro de 2022, portanto, conforme Varga e Brito (2016), não há viés de sobrevivência. No Brasil, há diferentes classificações de ações: ordinárias nominativas (ON), ações com direito a voto nas assembleias de acionistas; além das preferenciais nominativas (PN), que possuem preferência sobre o pagamento de dividendos, e as suas distintas classes: A (PNA), B (PNB), C (PNC) e D (PND), respectivamente. Existem também os recibos de subscrição de ações (RCTB), as UNITs que são compostas por ações ON e PN, os fundos de índice ETFs, os fundos de investimento imobiliário FIIs, e os certificados de depósito de valores mobiliários BDRs.

Nesta pesquisa, seguindo Varga e Brito (2016), foram mantidas apenas uma ação para cada empresa listada no período, de preferência a PN, por ser geralmente a mais negociada. Além disso, foram realizadas as seguintes exclusões da amostra, conforme Rizzi (2012), Portela (2016) e Varga e Brito (2016):

- a) Ações de empresas estrangeiras negociadas via BDR;
- b) Ações de bancos devido a sua alavancagem;
- c) Ações de empresas com valor contábil negativo no período em que isso ocorreu;
- d) Ações de empresas com valor de mercado ausente, e;
- e) Ativos negociados via UNITs, ETFs e FIIs.

Foram mantidas as empresas de serviços como corretoras de valores, seguradoras e *holdings*, por não terem alavancagem tão alta. A amostra final contém 202 empresas em todo o período amostral, composto por 240 observações mensais.

A variável dependente é o retorno da carteira, calculado a partir dos retornos das ações das empresas que compõem cada um dos grupos. Já as variáveis independentes são os fatores de risco pertencentes aos modelos de precificação de ativos observados na Seção 2.1.5. Para formar as carteiras, classificamos os ativos em ordem crescente com base na variável tamanho (ME) mensurada pelo valor de mercado do ativo. Os ativos foram distribuídos em quintis, dando origem a cinco carteiras distintas. O menor quintil abrange os 20% de ativos com os valores de mercado mais baixos, ao passo que o maior quintil inclui os 20% com os valores de mercado mais altos.

Os dados para calcular os modelos de precificação de ativos foram coletados por meio de fontes secundárias, com informações provenientes de fontes públicas. Os dados relacionados ao retorno de mercado e retorno das ações usados na formação de carteiras foram obtidos em consulta a B3. Os dados relacionados a taxa livre de risco, usada para calcular o prêmio de risco de mercado, e o Certificado de Depósito Interbancário (CDI) usado para calcular o *spread* de juros foram coletados no *site* do Banco Central do Brasil (BACEN). Os dados relacionados as variáveis tamanho e *book-to-market* foram obtidos por meio da base de dados da plataforma ComDinheiro, e o fator de risco HML foi obtido na base de dados do Núcleo de Pesquisa em Economia Financeira da Universidade de São Paulo (NEFIN). Os dados coletados foram organizados e processados por meio do *software* de planilha eletrônica Microsoft Excel, versão do Office 365 e as estimativas das regressões, os testes econométricos e a implementação do algoritmo de *machine learning* da técnica de *stacked generalization* foi realizado por meio da linguagem Python, com o desenvolvimento de um código de programação específico para essa finalidade, elaborado pelo próprio autor.

### 3.3 *STACKED GENERALIZATION* APLICADA A PRECIFICAÇÃO DE ATIVOS

Para atingir os objetivos específicos da pesquisa, primeiramente a amostra foi dividida em três conjuntos de dados: (i) treinamento, (ii) validação e (iii) teste. Optou-se pela abordagem de janela de expansão proposta por Timmermann (2008), na qual os dados são segmentados em partes, seguindo um formato de estimativas recursivas. As janelas de estimativas recursivas utilizadas foram determinadas com base nos estudos recentes de Rapach e Zhou (2020) e Zhao e Cheng (2022).

A amostra inicial de treinamento abrange o período de janeiro de 2003 a dezembro

de 2005, totalizando 36 meses (15%). Todavia, ao considerar a restrita disponibilidade de dados macroeconômicos brasileiros, que limita o tamanho da amostra a uma pequena fração dos estudos internacionais realizados sobre o tema, optou-se por estender o período de validação a fim de assegurar que os modelos de *machine learning* pudessem realizar uma validação robusta da amostra. Análises preliminares indicaram que períodos muito curtos de treinamento e validação resultam em *underfitting*, uma situação na qual as previsões são subestimadas, destacando a importância de janelas de dados mais amplas. Esse ajuste visa otimizar a capacidade dos modelos em aprender padrões complexos e melhorar a generalização para períodos futuros. Assim, o conjunto de validação abrange o intervalo de janeiro de 2006 a dezembro de 2012, totalizando 84 meses (35%), enquanto o período de teste se estende de janeiro de 2013 a dezembro de 2022, com uma duração de 120 meses (50%).

Em seguida, utilizou-se o conjunto de dados de treinamento para empregar a janela de expansão de Timmermann (2008), visando realizar previsões um passo à frente para os modelos base no conjunto de dados de validação. Desta forma, a primeira previsão foi gerada para janeiro de 2006, utilizando dados de janeiro de 2003 a dezembro de 2005. No mês seguinte, a janela de dados de treinamento é expandida para incluir também janeiro de 2006, e os parâmetros dos modelos de previsão são novamente estimados, e depois usados para prever os retornos das ações para fevereiro de 2006, e assim por diante até o final da amostra de validação. Este mesmo procedimento foi aplicado ao conjunto de dados de teste com relação ao conjunto de dados de validação, e os resultados das previsões dos modelos de precificação de ativos foram armazenados em uma matriz de previsões. O Apêndice A apresenta um fluxograma detalhado do procedimento de *stacked generalization* aplicado aos modelos de precificação de ativos.

Diferentemente de Zhao e Cheng (2022), não foram utilizados algoritmos de *machine learning* como modelos base. Foram considerados cinco modelos de precificação de ativos citados por diversas pesquisas que envolvem a previsão de retorno das ações nos últimos anos. Os cinco modelos utilizados são apresentados a seguir e uma descrição detalhada deles pode ser observada nas Seções 2.1.1 e 2.1.5.

**a) CAPM tradicional:**

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \quad (20)$$

Onde:

$E(R_i)$  = retorno esperado do ativo  $i$

$R_f$  = retorno do ativo livre de risco

$\beta_i$  = sensibilidade do ativo  $i$  em relação ao mercado

$E(R_m)$  = retorno esperado do mercado

$E(R_m) - R_f$  = prêmio do risco de mercado

**b) *Learning* CAPM (L-CAPM):**

A única diferença deste modelo em comparação ao CAPM tradicional é que a sensibilidade do ativo em relação ao mercado que é calculado a partir do filtro de Kalman.

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \quad (21)$$

Onde:

$E(R_i)$  = retorno esperado do ativo  $i$

$R_f$  = retorno do ativo livre de risco

$\beta_i$  = sensibilidade do ativo  $i$  construídas a partir do filtro de Kalman

$E(R_m)$  = retorno esperado do mercado

$E(R_m) - R_f$  = prêmio do risco de mercado

**c) *Downside* CAPM (D-CAPM):**

$$R_i = R_f + (R_m - R_f) \beta^D \quad (22)$$

$$\beta^D = \frac{E(\text{Min}((r_i - \bar{r}_i), 0)) \times E(\text{Min}((r_M - \bar{r}_M), 0))}{E(\text{Min}((r_i - \bar{r}_i), 0)^2)}$$

Onde:

$R_i$  = retorno esperado do ativo

$R_f$  = retorno do ativo livre de risco

$R_m$  = retorno do mercado

$R_m - R_f$  = prêmio do risco de mercado

$\beta^D$  = beta *downside* do ativo

**d) *Conditional* CAPM (C-CAPM):**

$$R_{t+1}^i = \alpha^i + \beta^{1i} R_{t+1}^M + \beta^{2i} R_{t+1}^M R_m - R_{f_t} + \beta^{3i} R_{t+1}^M HML_t + \beta^{4i} R_{t+1}^M Spread_t + \epsilon_{t+1} \quad (23)$$

Onde:

$R_{t+1}^i$  = retorno esperado do ativo  $i$

$\beta^i$  = beta condicional do ativo  $i$

$R_{t+1}^M$  = retorno do mercado

$Rm - Rf_t$  = prêmio de risco do mercado

$HML_t$  = retorno da carteira high minus low

$Spread_t$  = spread de juros

$\epsilon_{t+1}$  = erro aleatório

**e) Regressão *cross-section* de Fama & French (F&F):**

$$R_{i,t+\tau} = \lambda_{t+\tau}^0 + \lambda_{t+\tau}^M \beta_{i,t-1} + \lambda_{t+\tau}^{ME} \ln(ME_{i,t-1}) + \lambda_{t+\tau}^{BM} \ln(BM_{i,t-1}) + v_{i,t+\tau}, \forall i, \text{ para cada } (t + \tau) \quad (24)$$

Onde:

$R_{i,t+\tau}$  = retorno esperado do ativo

$t, t - 1$  = meses de julho e junho, respectivamente

$\tau = 0, 1, \dots, 11$

$t + \tau$  = período de julho  $t$  a junho  $t - 1$  do ano seguinte

$\lambda_{t+\tau}^0$  = taxa livre de risco

$\beta_{i,t-1}$  = sensibilidade do ativo em relação ao mercado

$ME_{i,t-1}$  = valor de mercado

$BM_{i,t-1} = (BE_{i,t-1} / ME_{i,t-1})$  = valor do patrimônio líquido ao valor de mercado

$\bar{\lambda}^M = (T^{-1} \sum_{t=1}^T \lambda_t^M) > 0$

Na sequência, seguindo Zhao e Cheng (2022) avaliou-se a eficácia do desempenho das previsões no conjunto de dados de validação por meio do  $R^2$  fora da amostra ( $R_{OOS}^2$ ) de Campbell e Thompson (2008). Nesta etapa, os três modelos de precificação de ativos com os melhores  $R_{OOS}^2$  foram selecionados para treinar o meta-modelo de *stacked generalization*, enquanto as duas piores previsões foram descartadas.

O  $R_{OOS}^2$  de Campbell e Thompson (2008) é dado por:

$$R_{OOS,\mathbb{A}}^{2,(j)} = 1 - \frac{\sum_{\tau \in \mathbb{A}} (y_{\tau} - \hat{y}_{\tau|\tau-1}^{(j)})^2}{\sum_{\tau \in \mathbb{A}} (y_{\tau} - \bar{y}_{\tau|\tau-1})^2}, \quad (25)$$

Onde  $j$  indica cada um dos modelos em consideração, e  $\mathbb{A}$  é a janela de tempo para comparação.  $y_{\tau}$  é o prêmio de risco das ações observado na data  $\tau$ , e  $\hat{y}_{\tau|\tau-1}^{(j)}$  é a previsão na

data  $\tau - 1$  do modelo  $j$ .  $\bar{y}_{\tau|\tau-1} = \frac{1}{\tau-1} \sum_{i=1}^{\tau-1} y_i$  é a previsão feita pela média histórica (MH).

Neste sentido,  $R_{OOS,\mathbb{A}}^{2,(j)}$  também pode ser escrito como:  $R_{OOS,\mathbb{A}}^{2,(j)} = 1 - \frac{MFSE_{\mathbb{A}}^{(j)}}{MFSE_{\mathbb{A}}^{(MH)}}$ , onde  $MFSE_{\mathbb{A}}^{(j)} = \frac{1}{\#\mathbb{A}} \sum_{\tau \in \mathbb{A}} (y_{\tau} - \hat{y}_{\tau|\tau-1}^{(j)})^2$  e  $MFSE_{\mathbb{A}}^{(MH)} = \frac{1}{\#\mathbb{A}} \sum_{\tau \in \mathbb{A}} (y_{\tau} - \bar{y}_{\tau|\tau-1})^2$  representam os erros de previsão quadráticos médios (*Mean Squared Forecast Error* - MSFE) no conjunto  $\mathbb{A}$  de um determinado modelo  $j$  e MH, respectivamente.

$\#\mathbb{A}$  é a cardinalidade do conjunto  $\mathbb{A}$ . Portanto,  $R_{OOS,\mathbb{A}}^{2,(j)}$  é uma estatística conveniente para comparar MSFE de um modelo concorrente em relação ao MH. Quando  $R_{OOS,\mathbb{A}}^{2,(j)}$  for maior que zero, o modelo concorrente supera a MH em termos de MSFE.

A significância estatística foi avaliada pela estatística ajustada ao MSFE proposta por Clark e West (2007). A hipótese nula é que a MH possui um MSFE menor ou igual em relação ao modelo de previsão concorrente, e a hipótese alternativa é que a MH tem um MSFE maior que o modelo concorrente. Se a hipótese nula for rejeitada, o modelo concorrente tem um desempenho melhor que a MH (Zhao e Cheng, 2022).

Na etapa seguinte, os três modelos de precificação de ativos com melhor  $R_{OOS}^2$  foram utilizados no conjunto de validação para treinar o meta-modelo de *stacked generalization*. Para realizar esta previsão, foram selecionados três meta-modelos com base em Zhao e Cheng (2022): *Random Forest* (St-RF); *Gradient Boosted Regression Trees* (St-GBRT) e *Neural Networks* (St-NN). A escolha destes meta-modelos basearam-se nos resultados obtidos pelos autores que os aplicaram na previsão de retorno do prêmio de risco das ações no mercado dos EUA. Em sua análise eles constataram que dentre os meta-modelos utilizados, o St-RF foi aquele que apresentou uma performance superior. Ademais, com base nos resultados obtidos em testes preliminares decidiu-se incluir um meta-modelo com Regressão Linear (St-LR) baseado em MQO. Na sequência, empregou-se novamente a janela expansão de Timmermann (2008) para realizar previsões um passo à frente no conjunto de dados de teste, da mesma forma realizada com os modelos de precificação de ativos no conjunto de dados de validação. O Apêndice B apresenta uma descrição resumida dos meta-modelos utilizados no estudo.

Para realizar uma análise comparativa entre o desempenho de previsão dos meta-modelos que utilizaram a seleção dos modelos base com os melhores  $R_{OOS}^2$ , e os meta-modelos sem essa seleção, foi realizado o teste de Wilcoxon para dados pareados. Este teste é uma versão não paramétrica do teste  $t$ , e é usado na ausência da suposição de normalidade e quando se assume que os dados são provenientes de uma distribuição simétrica (Barros;

Mazucheli, 2005). A hipótese nula do teste considera que não há diferença significativa entre as duas previsões, enquanto a hipótese alternativa considera que há uma diferença significativa entre as duas previsões.

Por fim, comparou-se os resultados das previsões entre os modelos de precificação de ativos e os meta-modelos de duas formas distintas: (i) ganhos estatísticos e (ii) previsões de sinal. Os ganhos estatísticos foram avaliados através do  $R_{00S}^2$  de Campbell e Thompson (2008). Enquanto a previsão de sinal foi investigada conforme proposto por Zhao e Cheng (2022), em que métricas como acurácias positivas e negativas são utilizadas para avaliar a capacidade do modelo em prever corretamente a direção dos movimentos dos retornos. A previsão de sinal é uma parte importante da análise de retornos das ações, pois determina a capacidade do modelo em prever corretamente a direção dos movimentos do preço dos ativos. Por exemplo, a alocação de ativos por gestores de fundos depende de previsões direcionais, pois sua regra de decisão é entrar comprado nos ativos com os maiores retornos positivos esperados e vendido nos ativos cujos retornos esperados são negativos (Pesaran e Timmermann, 2002). Na previsão de sinal, a ocorrência de um retorno positivo (negativo) é definido como um evento positivo (negativo), que recebe o valor 1 quando positivo e zero nos casos contrários.

A Figura 2 apresenta os possíveis resultados da previsão de sinal em problemas de classificação.

		PREVISÃO	
		Evento Positivo	Evento Negativo
REAL	Evento Positivo	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Evento Negativo	Falso Positivo (FP)	Verdadeiro Negativo (TN)

**Figura 2 – Matriz de confusão da previsão de sinais**

Fonte: Elaborada pelo autor (2024).

- **TP (*True Positives*):** representa o número de casos verdadeiros positivos corretamente previstos pelo modelo, ou seja, os casos em que o modelo previu retorno positivo corretamente.
- **FP (*False Positives*):** representa o número de casos falsos positivos, ou seja, os casos em que o modelo previu retorno negativo, mas na verdade, o retorno foi positivo.
- **FN (*False Negatives*):** representa o número de casos falsos negativos, ou seja, os

casos em que o modelo previu retorno negativo, mas na verdade, o retorno foi positivo.

- **TN (*True Negatives*):** representa o número de casos verdadeiros negativos corretamente previstos pelo modelo, ou seja, os casos em que o modelo previu retorno negativo corretamente.

As três principais medidas de avaliação da previsão de sinal utilizadas nesta pesquisa são a (i) acurácia, (ii) sensibilidade e (iii) especificidade.

- Acurácia (*accuracy*) =  $TP + TN \div (TP + FP + FN + TN)$ :** mensura a proporção de previsões corretas em relação ao número total de previsões. É calculada como a soma das previsões corretas (positivas e negativas) dividida pelo número total de previsões. Essa métrica fornece uma visão geral da eficácia geral do modelo em prever corretamente o retorno do prêmio de capitais.
- Sensibilidade (*recall*) =  $TP \div (TP + FN)$ :** também conhecida como *recall*, taxa de verdadeiros positivos ou acurácia positiva, é uma métrica essencial para avaliar a capacidade de um modelo em identificar corretamente as instâncias positivas. Mensura a proporção de casos positivos corretamente previstos em relação ao número total de casos positivos reais.
- Especificidade (*specificity*) =  $TN \div (TN + FP)$ :** também conhecida como taxa de falsos positivos ou acurácia negativa, é uma métrica importante utilizada para avaliar a capacidade de um modelo em identificar corretamente as instâncias negativas. Mensura a proporção de casos negativos corretamente previstos em relação ao número total de casos negativos reais.

### 3.4 HIPÓTESES DE PESQUISA

A fim de atingir aos objetivos da pesquisa, formularam-se hipóteses destinadas a testar a validade da técnica de *stacked generalization* e dos modelos de precificação de ativos no mercado de ações brasileiro. De acordo com Creswell (2010), as hipóteses em pesquisas quantitativas representam previsões feitas pelo pesquisador sobre a relação esperada entre as

variáveis. Sua avaliação envolve procedimentos estatísticos nos quais inferências são realizadas sobre a população com base na amostra do estudo.

Creswell (2010) destaca duas formas de utilização de hipóteses em uma pesquisa. A primeira, a hipótese nula, adota uma abordagem mais tradicional, prevendo que, na população geral, não existe uma relação ou diferença significativa entre os grupos em uma variável específica. A segunda, a hipótese alternativa, é formulada quando o pesquisador faz uma previsão sobre o resultado esperado, fundamentando essa previsão na literatura e em estudos anteriores sobre o tema, os quais sugerem um resultado potencial.

Nesta pesquisa, as hipóteses nulas, denotadas por  $H_0$ , postulam que a técnica de *stacked generalization* não apresenta uma relação significativa com o retorno das ações no mercado brasileiro e, adicionalmente, gera previsões menos precisas em comparação aos modelos de precificação de ativos selecionados. Por outro lado, as hipóteses alternativas, identificadas como  $H_1$  e  $H_2$ , propõem argumentos contrários às hipóteses nulas, proporcionando uma base para comparação entre ambas. As relações entre as hipóteses e os objetivos específicos da pesquisa são apresentadas no Quadro 1.

### Quadro 1 - Hipóteses de Pesquisa e Objetivos Específicos

Este quadro apresenta as hipóteses de pesquisa, relacionando-as com os objetivos específicos do estudo. Os objetivos específicos podem ser visualizados na seção 1.3, enquanto a descrição do modelo de *stacked generalization* pode ser visualizada na seção 2.2.1 e os modelos de precificação de ativos nas seções 3.5.2.

Objetivo Específico	Hipóteses	Descrição
Analisar a eficácia dos modelos de precificação de ativos na estimativa do retorno das ações na B3 ao longo do período de 2003 a 2022.	$H_0$	Os modelos de precificação de ativos não fornecem previsões significativamente positivas na estimativa do retorno das ações na B3.
	$H_1$	Os modelos de precificação de ativos fornecem previsões significativamente positivas na estimativa do retorno das ações na B3.
Avaliar se a técnica de <i>stacked generalization</i> em <i>machine learning</i> aplicada aos modelos de precificação de ativos tem potencial para otimizar o retorno das ações	$H_0$	A técnica de <i>stacked generalization</i> em <i>machine learning</i> oferece previsões menos precisas que os modelos de precificação de ativos na estimativa do retorno das ações na B3.

na B3 ao longo do período de 2003 a 2022.	H2	A técnica de <i>stacked generalization</i> em <i>machine learning</i> oferece previsões mais precisas que os modelos de precificação de ativos na estimativa do retorno das ações na B3.
---	----	--

Fonte: Elaborado pelo autor (2024).

A hipótese H1 é fundamental para entender se os modelos de precificação de ativos selecionados são capazes de fornecer previsões significativamente positivas na estimativa do retorno das ações. Por outro lado, a hipótese H2 é crucial para determinar se a técnica de *stacked generalization* apresenta vantagens em termos de precisão na previsão dos retornos das ações em comparação aos modelos de precificações de ativos estabelecidos. As validações destas hipóteses sugerem que os modelos de precificação de ativos estudados e a técnica de *stacked generalization* aplicada sobre estes modelos são capazes de gerar previsões de retorno das ações mais eficazes, e podem contribuir positivamente na tomada de decisões de investimento.

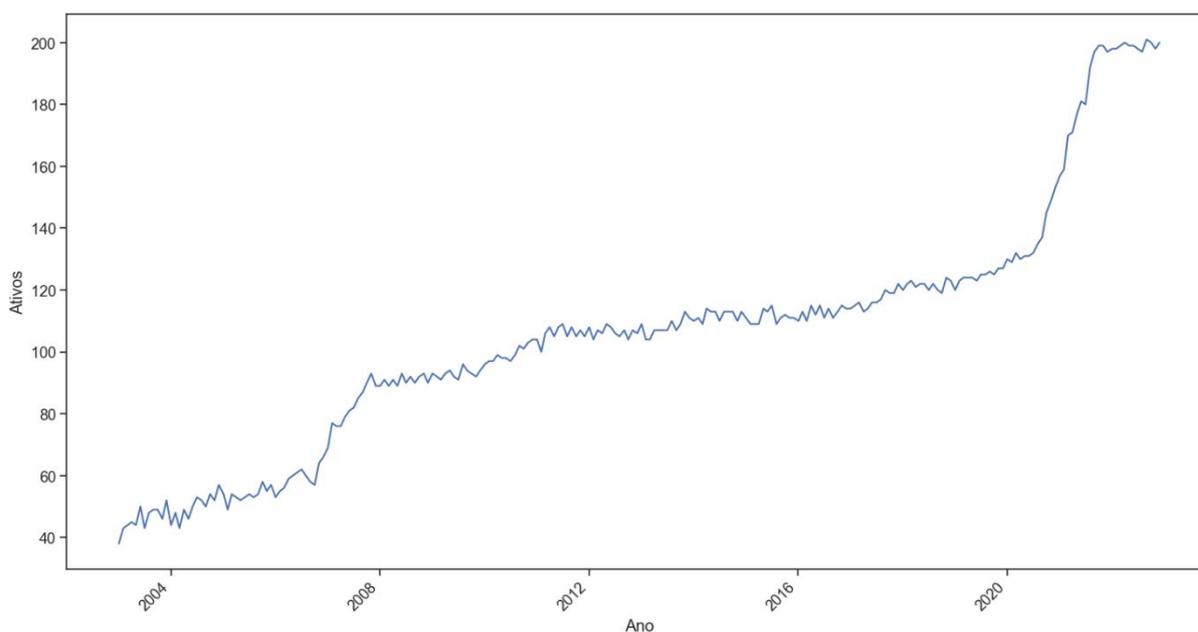
## 4. RESULTADOS E DISCUSSÕES

---

Este capítulo inicia com a apresentação das estatísticas descritivas da pesquisa, oferecendo uma análise detalhada das principais características da amostra dos ativos selecionados. Em seguida, o desempenho de previsão dos modelos de precificação de ativos e meta-modelos de *stacked generalization* tanto com quanto sem a seleção de  $R_{OOS}^2$  são avaliados em todo o conjunto de dados de teste e em condições extremas de mercado.

### 4.1 ESTATÍSTICA DESCRITIVA

Esta seção apresenta as estatísticas descritivas das ações e carteiras de ações utilizadas na amostra estudada. Optou-se pela formação de carteiras mensais dado o retorno mensal dos ativos. Essa escolha tem por finalidade garantir que as previsões de retorno possam acompanhar as possíveis mudanças na composição das carteiras de forma adequada, semelhante às decisões de investimento em um ambiente real. O período utilizado na análise compreende a amostra completa de janeiro de 2003 a dezembro de 2022.



**Figura 3 – Série Temporal das Ações por Ano**

Fonte: Elaborada pelo autor (2024).

Como o processo de formação de carteiras baseou-se em quintis, com 20% dos ativos

alocados em cada carteira mensal, o número de ativos que compõe as carteiras é quase sempre o mesmo, com pequenas alterações a cada nova oferta de ações. Na Figura 3, observamos que no início de 2003 cada carteira de ações era composta em média por 8 ativos, 40 no total. Enquanto no final de 2022 esse número era quatro vezes maior, com um total de aproximadamente 40 ativos por carteira, e 202 ativos no total.

A Tabela 1 apresenta as estatísticas descritivas das carteiras de ações estudadas. Inicialmente, observa-se que o retorno médio mensal das carteiras decai à medida que o tamanho da empresa aumenta. O retorno médio das empresas do terceiro quintil (*Median-ME*) é menor do que o retorno médio das empresas do quarto quintil (*Small-Median-ME*), e este é consequentemente menor do que o retorno médio das empresas do quinto quintil (*Large-ME*). Estas observações corroboram com a teoria de risco e retorno quando analisamos o risco mensurado pelo desvio padrão dos retornos. Os resultados apontam que o desvio padrão diminui à medida em que o tamanho da empresa aumenta. Todavia, ao observar o coeficiente de variação dos retornos, percebemos que o risco condicional da carteira das grandes empresas (*Large-ME*) é duas vezes maior do que o observado nas empresas de médio porte (*Median-ME*). Isto pode indicar que o risco de investir em grandes empresas não diminui na mesma proporção de seus retornos.

**Tabela 1 – Estatística Descritiva das Carteiras baseadas no Valor de Mercado**

Carteiras	Média do Retorno	Desvio Padrão dos Retornos	Coefficiente de Variação dos Retornos	Mínimo Retorno	Máximo Retorno	Valor de Mercado Médio (Milhões R\$)
<i>Small-ME</i>	0,0197	0,0934	4,7416	-0,3870	0,3064	242,84
<i>Small-Median-ME</i>	0,0199	0,0869	4,3723	-0,3476	0,3348	978,29
<i>Median-ME</i>	0,0162	0,0803	4,9663	-0,3615	0,3327	2.554,77
<i>Median-Large-ME</i>	0,0147	0,0735	5,0067	-0,3418	0,3040	5.941,74
<i>Large-ME</i>	0,0080	0,0715	8,8862	-0,4968	0,1891	39.399,89

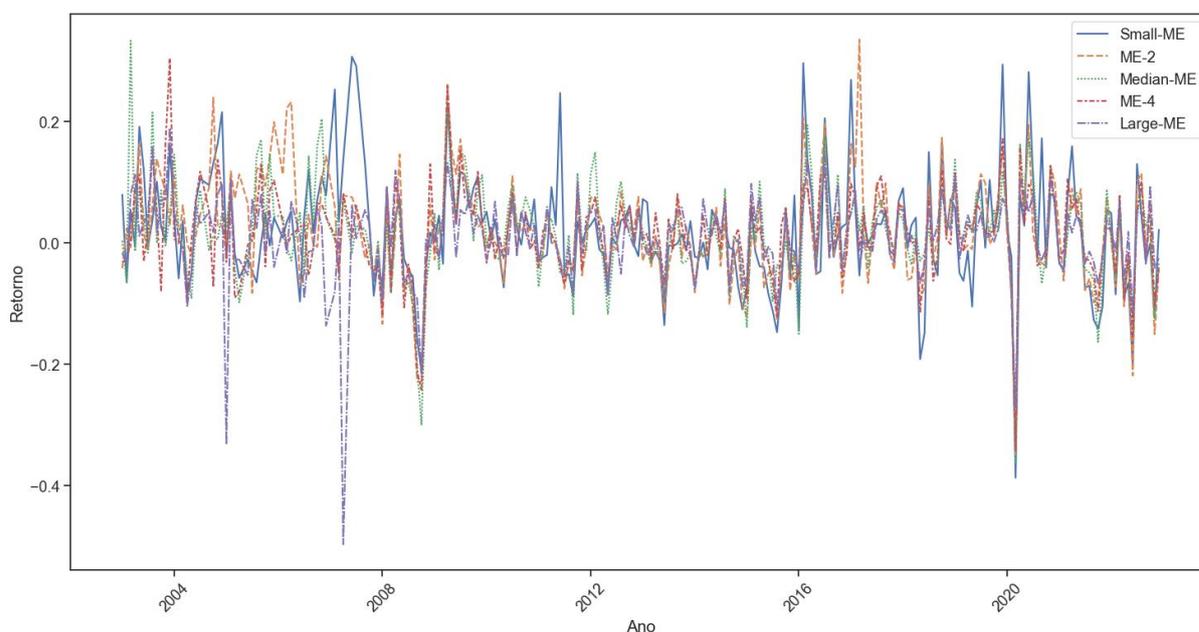
Nota: Esta tabela apresenta as estatísticas descritivas (média dos retornos, desvio padrão dos retornos, coeficiente de variação dos retornos, mínimo retorno, máximo retorno e valor de mercado médio em milhões) para as cinco carteiras de ativos baseadas no tamanho, mensuradas pelo valor de mercado dos ativos. *Small-ME* são os ativos do primeiro quintil; *Small-Median-ME* são os ativos do segundo quintil; *Median-ME* são os ativos do terceiro quintil; *Median-Large-ME* são os ativos do quarto quintil; e *Large-ME* são os ativos do maior quintil.

Fonte: Elaborada pelo autor (2024)

Os valores de retorno mínimo e máximo contribuem com essa análise. A carteira composta pelas grandes empresas obteve uma perda máxima superior ao de todas as outras carteiras (-49,68%), enquanto o retorno máximo mensal foi apenas 2/3 do retorno máximo obtido pelas demais empresas (18,91%). Por fim, as estatísticas indicam que de fato, há uma diferença significativa no tamanho das empresas da amostra. Com as pequenas empresas

valendo em média R\$ 242 milhões, as médias empresas R\$ 2,5 bilhões e as grandes empresas R\$ 39 bilhões.

A Figura 4 evidencia o comportamento das séries temporais das cinco carteiras de ativos em todo o período da amostra de janeiro de 2003 a dezembro de 2022. Observa-se que alguns períodos específicos apresentam uma maior volatilidade (discutiremos estes períodos com maiores detalhes mais adiante neste trabalho). A evidência é que, perceptivelmente, estes períodos afetam negativamente em maior escala as ações das grandes empresas, identificadas pelas linhas pontilhadas em roxo, enquanto as ações das pequenas empresas possuem comportamento inverso. Ou seja, os gráficos de séries temporais indicam que uma queda significativa no retorno das ações das grandes empresas é quase sempre precedida ou sucedida pelo aumento no retorno das ações das pequenas empresas.



**Figura 4 – Séries Temporais das Carteiras de Ativos**

Fonte: Elaborada pelo autor (2024).

## 4.2 RESULTADOS DA TÉCNICA DE *STACKED*

Esta seção apresenta os resultados empíricos da técnica de *stacked generalization* aplicada sobre os modelos de precificação de ativos, visando atingir os dois objetivos específicos da pesquisa, que consistem em: (i) analisar a eficácia dos modelos de precificação de ativos na estimativa do retorno das ações na B3; e (ii) avaliar se a técnica de *stacked generalization* em *machine learning* aplicada aos modelos de precificação de ativos tem

potencial para otimizar o retorno das ações na B3 ao longo do período de 2003 a 2022.

### 4.2.1 Desempenho de previsão no conjunto de teste

Primeiramente, utilizou-se toda a amostra de teste cobrindo o período de janeiro de 2013 a dezembro de 2022 para avaliar o desempenho fora da amostra dos (i) modelos de precificação de ativos, (ii) combinação de média simples, (iii) meta-modelos de *stacked generalization* com todos os modelos base; e (iv) meta-modelos de *stacked generalization* com a seleção de modelos base com os melhores  $R_{OOS}^2$  no conjunto de validação para as cinco carteiras de ativos: (i) empresas de pequeno porte (*Small-ME*) – ativos do primeiro quintil; (ii) empresas de porte pequeno-médio (*Small-Median-ME*) – ativos do segundo quintil; (iii) empresas de médio porte (*Median-ME*) – ativos do terceiro quintil; (iv) empresas de porte médio-grande (*Median-Large-ME*) – ativos do quarto quintil; e grandes empresas (*Large-ME*) – ativos do maior quintil.

A Tabela 2 apresenta os resultados do  $R_{OOS}^2$  de Campbell e Thompson (2008) dos modelos de precificação de ativos no conjunto de validação. Seguindo Zhao e Cheng (2022), foram selecionados os três melhores  $R_{OOS}^2$  de cada carteira para realizar o treinamento dos meta-modelos de *stacked generalization* para prever os seus respectivos retornos no conjunto de dados de teste. Na carteira *Small-ME* os modelos base selecionados foram o CAPM, D-CAPM e L-CAPM. Enquanto, nas demais carteiras os modelos com melhores desempenhos foram o F&F (1996), C-CAPM e CAPM.

**Tabela 2 – R-quadrado fora da amostra de Campbell e Thompson (2008)**

Carteiras	CAPM	D-CAPM	L-CAPM	C-CAPM	F&F (1996)
<i>Small-ME</i>	0,1474	-0,0806	-0,1447	-40,0754	-18,6253
<i>Small-Median-ME</i>	-0,2147	-0,2354	-0,2354	-0,1577	-0,1291
<i>Median-ME</i>	0,5567	0,1928	-0,0756	0,5989	0,2217
<i>Median-Large-ME</i>	-0,0932	-0,0949	-0,0962	-0,0644	0,1106
<i>Large-ME</i>	-0,283	-0,4291	-0,4833	-0,2500	-0,4028

Notas: CAPM: *Capital Asset Pricing Model* (modelo de precificação de ativos de capital); D-CAPM: é o modelo *Downside CAPM*; L-CAPM: é o modelo *Learning CAPM*; C-CAPM: é o modelo *Conditional CAPM*; F&F (1996): é o modelo de precificação de ativos de Fama e French (1996).

Fonte: Elaborada pelo autor (2024).

A Tabela 3 apresenta os resultados de previsão de todos os modelos e meta-modelos no conjunto de teste. Os Painéis A, C, E, G e I denotam os resultados dos modelos de precificação de ativos, combinação de média simples, e meta-modelos de *stacked generalization* com todos

os modelos base para cada uma das cinco carteiras de ativos, respectivamente. Por outro lado, os Painéis B, D, F, H e J evidenciam os resultados dos meta-modelos de *stacked generalization* com a seleção de modelos base com os melhores  $R_{OOS}^2$ , conforme resultados da Tabela 2.

Os resultados de  $R_{OOS}^2$  da carteira *Small-ME* mostram que dentre os modelos de precificação de ativos, apenas o modelo D-CAPM foi positivo e significativo, obtendo um  $R_{OOS}^2$  de 0,04 ao nível de significância estatística de 5% no período. Por outro lado, com exceção do St-GBRT, todos os meta-modelos do Painel A foram significativamente melhores do que a MH, com destaque para os meta-modelos St-LR e St-RF, com  $R_{OOS}^2$  de 0,34 e 0,24 ao nível de significância estatística de 1%, respectivamente. Com relação as medidas de previsão de sinal, todos os meta-modelos do Painel A, com exceção do St-GBRT, demonstraram melhor desempenho do que os modelos de precificação de ativos, tanto em termos de acurácia geral, quanto de especificidade. A capacidade dos meta-modelos em prever movimentos de queda (especificidade) pode ser o principal motivo para explicar a diferença significativa de desempenho entre os modelos.

Outrossim, os resultados do Painel B evidenciam que o meta-modelo St-LR possui o melhor desempenho de previsão em termos de  $R_{OOS}^2$  dentre todos os modelos e meta-modelos da carteira *Small-ME*, obtendo um  $R_{OOS}^2$  de 0,36 ao nível de significância estatística de 1%. Além disso, os resultados também consideram que realizar o *stacked generalization* com a seleção dos melhores modelos base pode melhorar a acurácia geral, e principalmente, os resultados das medidas de sensibilidade quando comparados aos resultados obtidos pelos modelos de *stacked* sem a seleção dos modelos base. Em suma, os resultados dos Painéis A e B evidenciam que os meta-modelos de *stacked generalization* melhoram significativamente os resultados de previsão dos modelos de precificação de ativos na carteira das pequenas empresas.

Os resultados dos Painéis C e D, referentes a carteira *Small-Median-ME*, indicam que apenas o modelo de precificação de ativos D-CAPM foi significativo e positivo, com um  $R_{OOS}^2$  de 0,01 ao nível de significância de 1%. Assim como na carteira *Small-ME*, este foi o único modelo base com resultados significativo e positivo. Todavia, cabe ressaltar que o desempenho do D-CAPM não foi muito diferente de zero, em ambas as situações. Em relação as medidas de previsão de sinal, o D-CAPM obteve um desempenho de 100% na previsão das instâncias positivas (sensibilidade). Mas, acertou apenas 4,62% das instâncias negativas (especificidade). É provável que o baixo desempenho em prever as instâncias negativas tenham impactado diretamente a acurácia geral do modelo que gerou uma acurácia geral de apenas 48,33% durante todo o período do conjunto de teste.

Tabela 3 – Resultados de Previsão no Conjunto de Teste

(continua)

janeiro de 2013 a dezembro de 2022					
<b>Painel A: Carteira <i>Small-ME</i></b>	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,45	0,0052	74,17	80,00	68,33
D-CAPM	0,04**	0,0091	58,33	98,33	18,33
L-CAPM	-0,04***	0,0099	50,00	100,00	0,00
C-CAPM	-58,31***	0,5604	55,00	81,67	28,33
F&F (1996)	-11,03	0,1136	41,67	35,00	48,33
MediaM	-1,74***	0,0259	50,83	78,33	23,33
<b>Stacked com todos os modelos base</b>					
St-LR	0,34***	0,0062	69,17	93,33	45,00
St-RF	0,24***	0,0072	62,50	95,00	30,00
St-GBRT	-0,02***	0,0096	62,50	85,00	40,00
St-NN	0,01***	0,0094	66,67	88,33	45,00
<b>Stacked com os modelos CAPM, D-CAPM e L-CAPM</b>					
St-LR	0,36***	0,006	69,17	96,67	41,67
St-RF	0,20***	0,0076	65,00	96,67	33,33
St-GBRT	0,01***	0,0093	65,83	91,67	40,00
St-NN	-0,20***	0,0114	50,00	100,00	0,00
<b>Painel B: Carteira <i>Small-Median-ME</i></b>					
	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	-201,24	1,5489	19,17	18,18	20,00
D-CAPM	0,01***	0,0076	48,33	100,00	4,62
L-CAPM	-0,03***	0,0079	45,83	100,00	0,00
C-CAPM	-819,68***	6,2856	29,17	40,00	20,00
F&F (1996)	-25,12***	0,2	45,83	29,09	60,00
<b>Stacked com todos os modelos base</b>					
MediaM	-68,08***	0,5291	21,67	25,45	18,46
St-LR	-8886,77	68,0716	37,50	29,09	44,62
St-RF	-15996,49*	122,525	61,67	98,18	30,77
St-GBRT	-39339,76*	301,3114	62,50	98,18	32,31
St-NN	-1028,48**	7,8848	39,17	25,45	50,77
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	-7166,32	54,8946	23,33	20,00	26,15
St-RF	-13267,86*	101,6264	64,17	100,00	33,85
St-GBRT	-19106,36*	146,3435	63,33	100,00	32,31
St-NN	-595,71**	4,5702	43,33	38,18	47,69
<b>Painel C: Carteira <i>Median-ME</i></b>					
	$R^2_{00s}$	MSFE	Acu. (%)	Sem. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,71	0,0016	80,83	83,58	77,36
D-CAPM	0,28***	0,0039	80,00	94,03	62,26
L-CAPM	-0,03***	0,0056	55,83	100,00	0,00
C-CAPM	0,69**	0,0017	80,83	91,04	67,92
F&F (1996)	0,07***	0,005	65,83	76,12	52,83
MediaM	0,51***	0,0027	85,00	95,52	71,70
<b>Stacked com todos os modelos base</b>					
St-LR	0,71**	0,0016	84,17	92,54	73,58
St-RF	0,60**	0,0021	78,33	88,06	66,04
St-GBRT	0,52***	0,0026	70,00	86,57	49,06
St-NN	0,58***	0,0023	79,17	98,51	54,72
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,71**	0,0016	84,17	91,04	75,47
St-RF	0,59***	0,0022	78,33	95,52	56,60
St-GBRT	0,52***	0,0026	72,50	89,55	50,94
St-NN	0,50***	0,0027	75,00	62,69	90,57

Tabela 3 – Resultados de Previsão no Conjunto de Teste

(conclusão)

janeiro de 2013 a dezembro de 2022					
Painel D: Carteira <i>Median-Large-ME</i>	$R_{00s}^2$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,72	0,0013	82,50	86,15	78,18
D-CAPM	0,27	0,0034	77,50	96,92	54,55
L-CAPM	-0,05	0,0048	54,17	100,00	0,00
C-CAPM	-0,88***	0,0087	63,33	98,46	21,82
F&F (1996)	-0,63***	0,0075	39,17	52,31	23,64
MediaM	0,43***	0,0026	67,50	98,46	30,91
<b>Stacked com todos os modelos base</b>					
St-LR	-8,03*	0,0416	43,33	41,54	45,45
St-RF	0,60	0,0018	77,50	95,38	56,36
St-GBRT	0,55	0,0021	77,50	92,31	60,00
St-NN	0,45***	0,0025	72,50	89,23	52,73
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	-4,86	0,027	55,00	70,77	36,36
St-RF	0,59	0,0019	80,00	93,85	63,64
St-GBRT	0,58	0,0019	80,00	95,38	61,82
St-NN	0,34***	0,003	75,83	96,92	50,91
<b>Painel E: <i>Large-ME</i></b>					
	$R_{00s}^2$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,78	0,0007	84,17	83,10	85,71
D-CAPM	0,33	0,0021	80,83	95,77	59,18
L-CAPM	-0,04	0,0034	59,17	100,00	0,00
C-CAPM	0,73	0,0009	82,50	83,10	81,63
F&F (1996)	0,04**	0,0031	60,83	73,24	42,86
MediaM	0,67	0,0011	85,83	90,14	79,59
<b>Stacked com todos os modelos base</b>					
St-LR	-4,85***	0,0189	60,83	42,25	87,76
St-RF	-1,97	0,0096	75,00	69,01	83,67
St-GBRT	-2,44	0,0111	78,33	71,83	87,76
St-NN	0,81*	0,0006	86,67	91,55	79,59
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,19***	0,0026	80,00	67,61	97,96
St-RF	-1,16**	0,007	74,17	64,79	87,76
St-GBRT	-0,88***	0,006	77,50	73,24	83,67
St-NN	0,64***	0,0012	75,83	59,15	100,00

Fonte: Elaborada pelo autor (2024). Notas: CAPM: Capital Asset Pricing Model (modelo de precificação de ativos de capital); D-CAPM: é o modelo Downside CAPM; L-CAPM: é o modelo Learning CAPM; C-CAPM: é o modelo Conditional CAPM; F&F: é o modelo de precificação de ativos de Fama e French (1996). Abreviaturas do nome do modelo: St significa stacked. St-LR: é o stacked com o modelo de Regressão Linear; St-RF: é o stacked com o modelo de Random Forest; St-GBRT: é o stacked com o modelo de Gradient Boosted Regression Trees; St-NN: é o stacked com o modelo de Neural Networks; MediaM: é a combinação de média simples dos modelos financeiros. MSFE é o erro quadrático médio de previsão; Acu. (%) é a medida de acurácia; Sen. (%) é a medida de sensibilidade; e Esp. (%) é a medida de especificidade. \*, \*\*, e \*\*\* indicam significância estatística nos níveis de 10%, 5% e 1%, respectivamente.

Fonte: Elaborada pelo autor (2024)

No tocante aos resultados de  $R_{00s}^2$  da carteira Median-ME, todos os modelos de precificação de ativos, com exceção dos modelos CAPM e L-CAPM, foram significativamente melhores do que a MH. Além disso, tanto a combinação de média simples, quanto todos os meta-modelos de stacked generalization com e sem a seleção dos melhores modelos base

também alcançaram resultados positivos e significativos. O modelo de precificação de ativos com melhor desempenho na carteira Median-ME foi o modelo C-CAPM com um  $R_{OOS}^2$  de 0,69 ao nível de significância de 5%, acurácia de 81%, sensibilidade de 91% e especificidade de aproximadamente 68%. Outro destaque foram os resultados da combinação de média simples, com um  $R_{OOS}^2$  de 0,51 ao nível de significância de 1%, acurácia de 85%, sensibilidade de 95% e especificidade de aproximadamente 71%.

Conforme mencionado anteriormente, todos os meta-modelos de *stacked generalization* foram significativamente melhores do que a MH no período. O melhor desempenho na carteira Median-ME foi do meta-modelo St-LR que obteve um  $R_{OOS}^2$  de 0,71 ao nível de significância de 5%, tanto no Painel E com todos os modelos base, quanto no Painel F com os melhores modelos base. Considerando as medidas de previsão de sinal, o meta-modelo St-LR gerou uma acurácia de 84% em ambas as análises, sensibilidade de 93% com todos os modelos base e 91% com os melhores modelos base, e especificidade de aproximadamente 74% na primeira análise e 75% na segunda. Estes resultados superam o desempenho dos modelos de precificação de ativos em todas as dimensões analisadas.

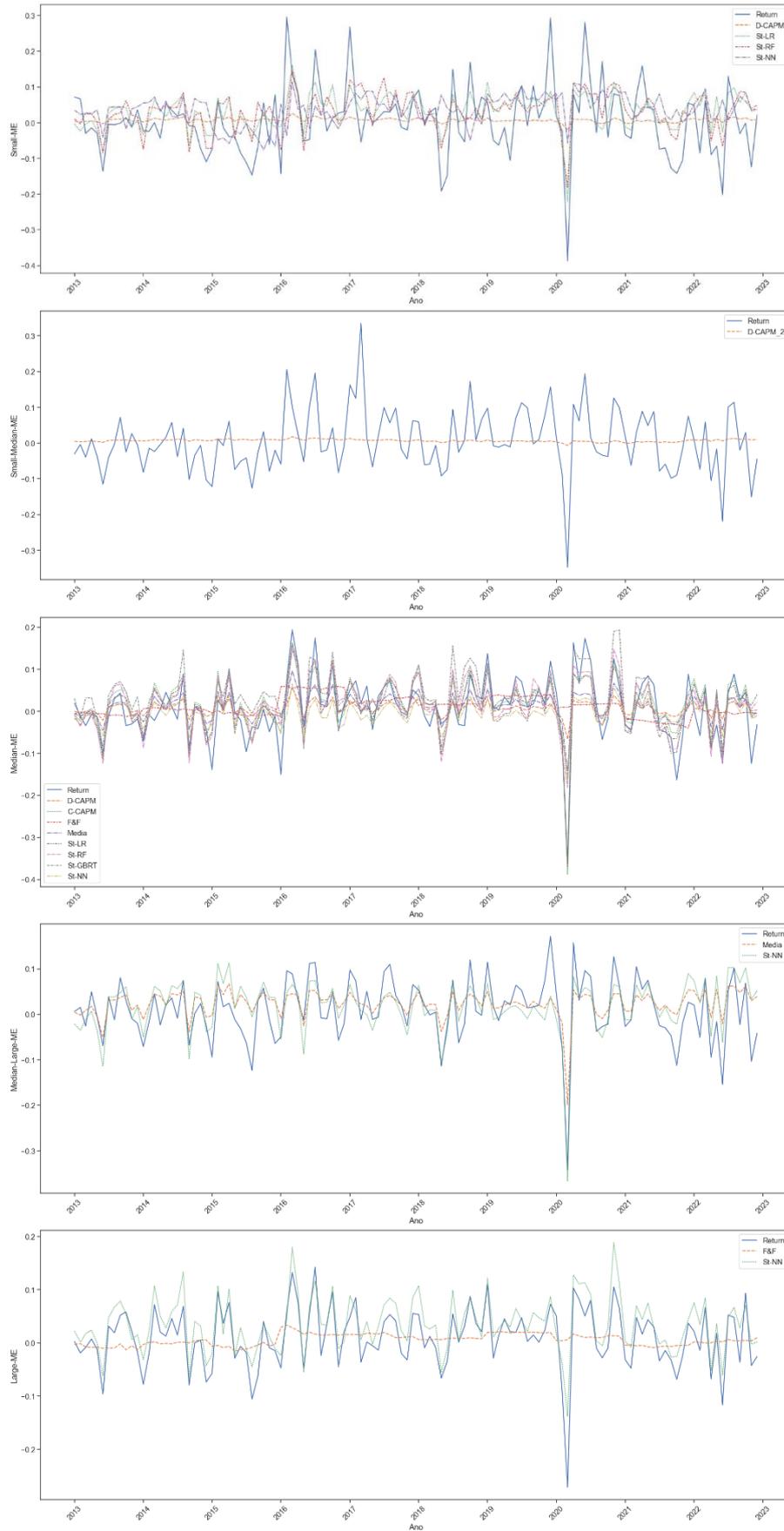
Todavia, ao comparar os resultados entre os meta-modelos com e sem a seleção de  $R_{OOS}^2$ , observa-se que, em geral, os meta-modelos não apresentaram melhorias significativas em termos de  $R_{OOS}^2$ , acurácia e sensibilidade. Na realidade, a maioria dos meta-modelos com seleção dos melhores modelos base por  $R_{OOS}^2$  demonstraram resultados semelhantes ou pior. Por outro lado, as medidas de especificidade que mensuram a acurácia negativa, foram geralmente melhores do que aquelas observadas nos modelos de precificação de ativos e meta-modelos sem a seleção de  $R_{OOS}^2$ , como são os casos dos meta-modelos St-LR, e St-NN que obteve uma especificidade de aproximadamente 91%. Estas evidências são condizentes com os resultados de Zhao e Cheng (2022), que constataram que os meta-modelos com seleção de modelos base por meio do  $R_{OOS}^2$  são capazes de realizar melhores previsões em movimentos extremos de queda.

Os resultados da carteira *Median-Large-ME* apontam que nenhum modelo de precificação de ativos foi significativamente melhor do que a MH no período. Apenas a combinação de média simples e o meta-modelo St-NN foram positivos e significativos no período. No Painel G, o meta-modelo St-NN apresentou o melhor resultado dentre todos os modelos e meta-modelos na carteira *Median-Large-ME*, com um  $R_{OOS}^2$  de 0,45 ao nível de significância de 1%, acurácia de 72,5%, sensibilidade de 89% e especificidade de aproximadamente 53%. Quanto aos meta-modelos do Painel H (com seleção dos melhores

modelos base), o meta-modelo St-NN apenas apresentou melhora no desempenho da medida de sensibilidade, com aproximadamente 97% de acerto, mas ainda abaixo do resultado apresentado pela combinação de média simples com 98,5%.

Por fim, ao considerar os resultados da carteira *Large-ME*, observa-se que apenas o modelo de precificação de ativos F&F (1996) foi significativamente melhor do que a MH no período, com um  $R_{OOS}^2$  de 0,04 ao nível de significância de 5%, acurácia de 61%, sensibilidade de 73% e especificidade de aproximadamente 43%. Todavia, o melhor desempenho na carteira *Large-ME* foi obtido pelo meta-modelo St-NN do Painel I, com um  $R_{OOS}^2$  de 0,81 ao nível de significância de 10%, acurácia de 87%, sensibilidade de 91,5% e especificidade de aproximadamente 80%. No que diz respeito aos meta-modelos do Painel J, os modelos St-LR e St-NN foram significativamente melhores do que a MH e superiores aos modelos de precificação de ativos. No entanto, estes meta-modelos não apresentaram melhorias em relação aos meta-modelos do Painel I, com exceção das medidas de especificidade, durante a mensuração de instâncias negativas, em que foi possível observar desempenhos de 97% e 100%, respectivamente.

Em geral, os resultados no conjunto de testes apontam que os meta-modelos de *stacked generalization* melhoraram o desempenho de todos os modelos de precificação de ativos, em todas as carteiras que apresentaram algum resultado significativamente melhor do que a média histórica no período. Estes resultados são, em média, significativamente melhores do que os encontrados por Zhao & Cheng (2022) no mercado dos EUA. Na ocasião, em várias janelas de intervalo, os melhores resultados alcançados pelos autores foi um  $R_{OOS}^2$  de 1,75 do modelo St-RF. Todavia, as evidências corroboram com os resultados destes autores ao identificar que empilhar com meta-modelos mais simples, como St-LR, tem melhor desempenho do que empilhar como meta-modelos mais complexos em termos de  $R_{OOS}^2$ . Com relação as medidas de previsão de sinal, os meta-modelos demonstraram melhor desempenho em prever instâncias negativas do que os modelos de precificação de ativos. A capacidade dos meta-modelos de *stacked* em prever movimentos de queda pode ser o principal motivo para explicar a diferença de desempenho entre os modelos.



**Figura 5 – Retorno Previsto das Carteiras no Conjunto de Teste**

Fonte: Elaborada pelo autor (2024)

A Figura 5 apresenta os retornos previstos dos principais modelos e meta-modelos em cada uma das carteiras estudadas. O desempenho dos meta-modelos St-LR e St-RF na carteira das pequenas empresas (*Small-ME*) são notáveis. O St-LR (ponto verde) e o St-RF (traço laranja) se sobrepõe com o retorno das ações (linha azul) de forma bastante próxima, mesmo durante alguns saltos, como a crise gerada pela pandemia de Covid-19 durante o ano de 2020. O meta-modelo St-LR é mais suave e mais preciso do que o St-RF e segue muito bem as tendências de crescimento, mas não consegue acompanhar de perto os movimentos de declínio, com exceção dos movimentos extremos de queda. Por outro lado, uma grande variedade de modelos e meta-modelos conseguiram acompanhar de forma muito precisa os movimentos de mercado na carteira das empresas de médio porte (*Median-ME*). Todavia, o mesmo não pode ser observado nas demais carteiras estudadas.

Com base nestas evidências, foi aplicado o teste de Wilcoxon em toda a amostra de teste com o objetivo de identificar se há diferenças significativas entre as previsões dos meta-modelos realizadas com a seleção dos melhores modelos base, através do  $R_{OOS}^2$ , e as previsões realizadas sem essa seleção, com todos os modelos base. A Tabela 4 apresenta os resultados do teste para cada um dos meta-modelos nas três carteiras de ativos estudadas.

**Tabela 4 – Resultados do Teste de Wilcoxon na Previsão dos Meta-modelos**

Modelos	<i>Small-ME</i>	<i>Small-Median-ME</i>	<i>Median-ME</i>	<i>Median-Large-ME</i>	<i>Large-ME</i>
<b>St-LR</b>	0,1067	0,0031***	0,2106	0,0034***	0,0000***
<b>St-RF</b>	0,7139	0,4050	0,5297	0,0297**	0,0008***
<b>St-GBRT</b>	0,0000	0,3947	0,0053***	0,1149	0,033**
<b>St-NN</b>	0,0601*	0,0000***	0,0000***	0,0000***	0,0000***

Notas: St significa *stacked*. St-LR: é o *stacked* com o modelo de Regressão Linear; St-RF: é o *stacked* com modelo de *Random Forest*; St-GBRT: é o *stacked* com o modelo de *Gradient Boosted Regression Trees*; St-NN: é o *stacked* com o modelo de *Neural Networks*. \*, \*\*, e \*\*\* indicam significância estatística nos níveis de 10%, 5% e 1%, respectivamente.

Fonte: Elaborada pelo autor (2024)

As estatísticas do teste de Wilcoxon sugerem que há na carteira *Small-ME* uma diferença significativa entre as previsões do meta-modelo St-NN. Também foram encontradas diferenças significativas entre as previsões dos meta-modelos St-LR e St-NN na carteira *Small-Median-ME*, mesmo esta carteira não apresentando resultados melhores do que a MH em termos de  $R_{OOS}^2$ . Com relação a carteira *Median-ME*, encontraram-se diferenças significativas entre as previsões dos meta-modelos St-GBRT e St-NN. Por fim, os resultados do teste indicam que há diferenças significativas entre as previsões de todos os meta-modelos nas carteiras *Median-Large-ME* e *Large-ME*, com exceção do meta-modelo St-GBRT na carteira *Median-*

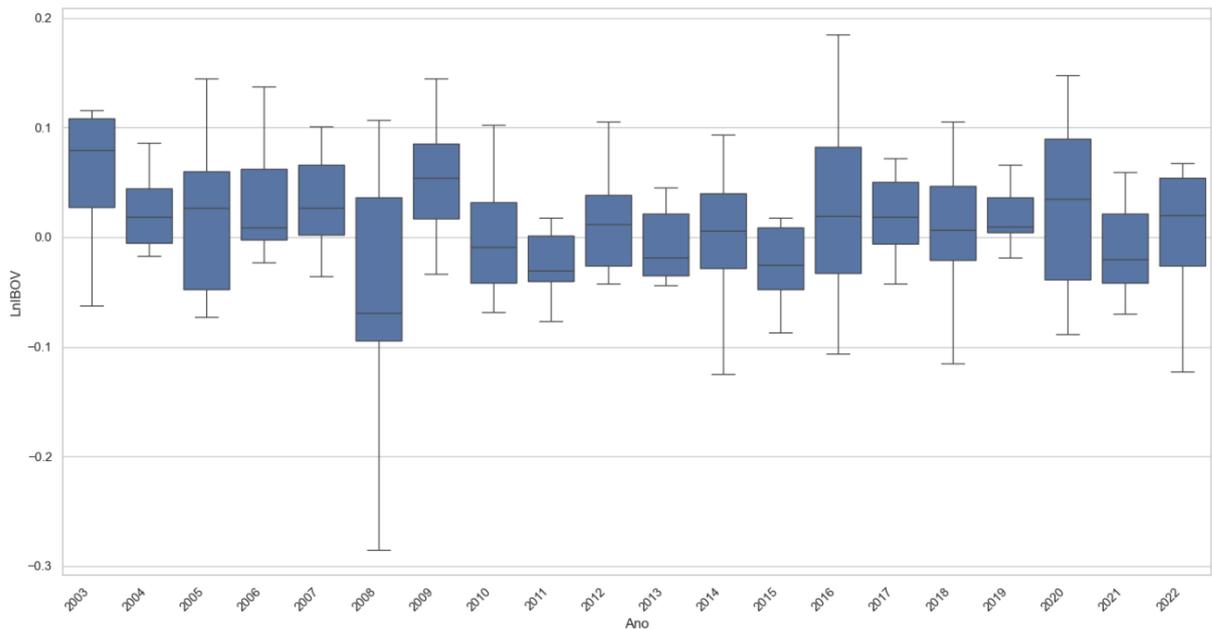
*Large-ME.*

Em geral, os resultados das carteiras indicam que a seleção dos modelos de precificação de ativos com base no desempenho do  $R_{OOS}^2$  no conjunto de dados de validação não produziu resultados mais eficazes do que a previsão dos meta-modelos sem essa seleção. Neste caso, o meta-modelo St-NN foi mais eficaz nos resultados de previsão quando o *stacked generalization* foi aplicado com todos os modelos base. Em relação a carteira das empresas maiores, as previsões de todos os meta-modelos foram significativamente diferentes, com uma única exceção. As evidências sugerem que assim como nas carteiras menores, o meta-modelo St-NN possui melhor desempenho quando os modelos base não são selecionados. Todavia, há uma diferença significativa entre as previsões do meta-modelo St-LR, sendo observada como mais eficaz quando os melhores modelos de precificação de ativos no conjunto de validação são selecionados para treinar os meta-modelos de *stacked generalization*.

## 4.2.2 Desempenho de previsão em condições extremas de mercado

Como o desempenho dos meta-modelos com seleção de  $R_{OOS}^2$  demonstram ser capazes de melhor prever os movimentos de queda, esta seção examina o desempenho das previsões em períodos com condições extremas de mercado, seguindo Zhao e Cheng (2022). Para tal, tornou-se imperativo identificar de forma adequada quais são os períodos com condições extremas de mercado.

Primeiramente, foi analisado o comportamento do mercado, mensurado pelo retorno do logaritmo natural do Ibovespa, ao longo do período estudado. A Figura 6 ilustra o gráfico *boxplot* anual do retorno logaritmo mensal do Ibovespa em todo o período da amostra. Nela, é possível identificar uma oscilação consistente da mediana ao longo do intervalo temporal, sugerindo a possível ausência de uma tendência de crescimento ou declínio. A amplitude interquartil, definida como a diferença entre o terceiro quartil (Q3) e o primeiro quartil (Q1), representada pela caixa-retângulo, evidencia, de maneira análoga à mediana, uma variação consistente, com certa sobreposição entre os períodos analisados.



**Figura 6 – Boxplot Anual do Retorno Logaritmo Mensal do Ibovespa**

Fonte: elaborada pelo autor, 2024.

Todavia, chama atenção o tamanho da extensão e assimetria dos bigodes (linhas que se estendem para fora das caixas) em alguns períodos do intervalo, como os anos de 2008 a 2010; 2014 a 2016; e 2020 a 2022, em vários períodos do intervalo analisado. Estas observações podem indicar condições extremas de mercado.

Neste sentido, foi calculado o *Drawdown* (DD) do retorno mensal do Ibovespa visando identificar períodos em que o mercado apresentou variações extremas. Geronazzo (2019) define o DD como uma representação de perdas significativas no valor de um ativo ou portfólio. O DD é calculado como o retorno negativo observado entre o valor máximo local (pico) e o próximo mínimo local (vale) em uma série temporal. Essencialmente, ele caracteriza a redução percentual máxima entre períodos de desempenho mais alto e os subsequentes declínios. Por sua vez, o *Maximum Drawdown* (MDD), refere-se ao valor máximo de um DD dentro de um intervalo de tempo específico. Em outras palavras, o MDD destaca a maior queda percentual experimentada ao longo de um período determinado, proporcionando uma métrica crucial para avaliar o risco associado a um investimento ou carteira.

Nesta pesquisa, foi calculado o DD anual do retorno mensal do Ibovespa como uma forma eficiente para identificar períodos de movimentos extremos de mercado. Ao considerar um horizonte de tempo fixo  $T \in [0, \infty]$ , o processo de *drawdown*  $D^X = \left\{ D_t^{(X)} \right\}_{t \in [0, T]}$  correspondente a um processo estocástico  $X \in R^\infty$  é definido pela seguinte equação:

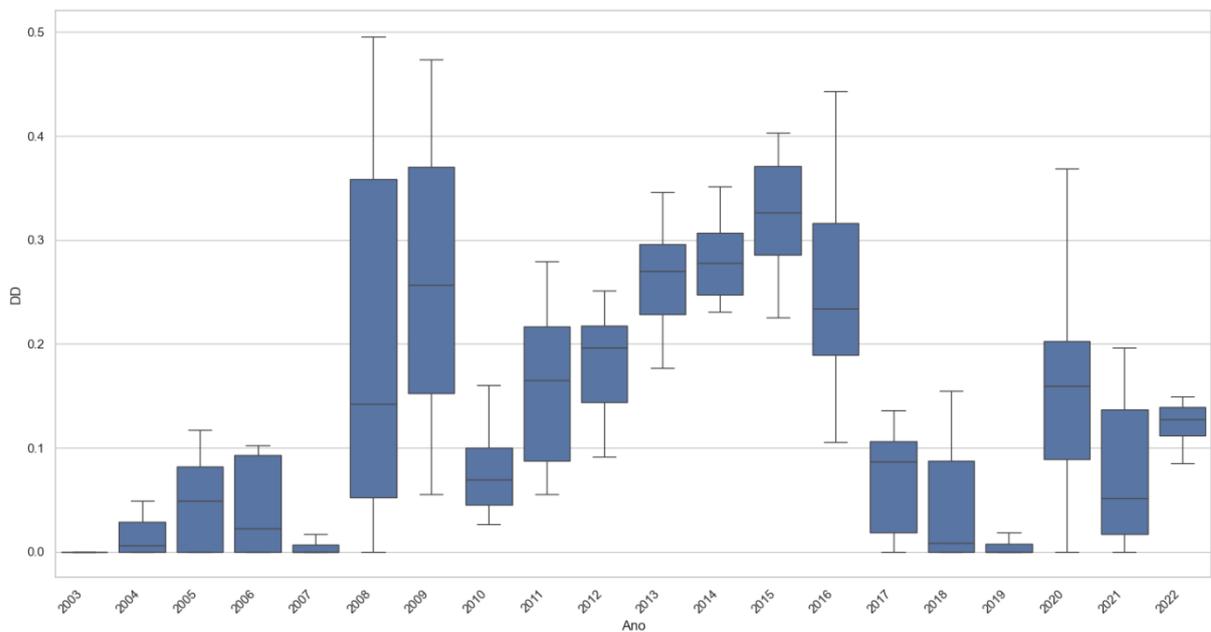
$$D_t^{(X)} = M_t^{(X)} - X_t \quad (26)$$

Onde:

$$M_t^{(X)} = \sup_{\mu \in [0, t]} X_\mu \quad (27)$$

é o máximo local de X até o tempo  $t$ .

A Figura 7 apresenta o gráfico *boxplot* anual do *drawdown* mensal do Ibovespa para o período completo da amostra. Observa-se que os resultados obtidos são condizentes com aqueles percebidos no gráfico anual do retorno logaritmo mensal do Ibovespa. Os anos de 2008 e 2009 denotam a manifestação de flutuações extremas em certos intervalos ao longo do ano, apresentando medianas de natureza moderada, entretanto, destacando-se por picos extraordinariamente pronunciados que alcançam perdas aproximadas de 50% do valor de mercado em determinados momentos, sem evidências de *outliers*. Consequentemente, constata-se uma ascensão progressiva do DD a partir de 2010, culminando em uma perda máxima superior a 40% no ano de 2016. Por fim, após um breve período de aparente estabilidade, observa-se uma nova perda significativa acima de 35% no ano de 2016.



**Figura 7 – Boxplot anual do DrawDown Mensal do Ibovespa**

Fonte: Elaborado pelo autor, 2024.

Interessante notar que os três períodos em que o Ibovespa apresentou perdas significativas mensurada pelo DD foram marcados pela existência de riscos não sistemáticos muito fortes:

- (i) Crise Financeira Global (2008 e 2009) – desencadeada pela crise no mercado imobiliário dos EUA, e que afetou os mercados financeiros em todo o mundo, incluindo a bolsa de valores brasileira e o seu principal índice, o Ibovespa;
- (ii) Crise Político-Econômica (2015 e 2016) – caracterizada por uma combinação de instabilidade política e desafios econômicos, que culminou no *impeachment* da presidente Dilma Rousseff e gerou incertezas quanto à direção política e econômica do país, que contribuiu para a intensificação de perdas na bolsa de valores brasileira; e
- (iii) Pandemia de Covid-19 (2020) – que desencadeou uma crise de saúde global e teve impactos dramáticos nos mercados financeiros. No Brasil, as medidas de *lockdown* e as incertezas em torno da propagação do vírus resultaram em interrupções significativas nas atividades econômicas. A queda na demanda, juntamente com a volatilidade nos preços do petróleo, afetou negativamente as empresas e a economia do país como um todo. Isso se refletiu no Ibovespa que experimentou uma nova perda significativa acima de 35%.

As Tabelas 5 e 6 denotam os resultados de previsão de três das cinco carteiras de ativos (i) empresas de pequeno porte (*Small-ME*) – ativos do primeiro quintil; (ii) empresas de médio porte (*Median-ME*) – ativos do terceiro quintil e; (iii) grandes empresas (*Large-ME*) – ativos do maior quintil, para dois períodos de crise contidos no período de teste: (i) de janeiro de 2015 a dezembro de 2016 – impactado pela crise político-econômica que culminou no *impeachment* da Presidente Dilma Rousseff; e (ii) janeiro de 2020 a junho de 2021 – afetado pela pandemia de Covid-19 que desencadeou uma crise de saúde global com impactos dramáticos nos mercados financeiros. Para treinamento e previsão dos meta-modelos de *stacked generalization* foram considerados os modelos de precificação de ativos com os três melhores  $R_{OOS}^2$  no conjunto de validação, conforme resultados da Tabela 2.

**Tabela 5 – Resultado de Previsão durante a Crise Político-Econômica de 2016**

janeiro de 2015 a dezembro de 2016					
<b>Painel A: Carteira <i>Small</i>-ME</b>	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,21**	0,0078	70,83	72,73	69,23
D-CAPM	-0,04***	0,0102	45,83	100,00	0,00
L-CAPM	-0,08***	0,0106	45,83	100,00	0,00
C-CAPM	-53,18***	0,5324	54,17	90,91	23,08
F&F (1996)	-20,45***	0,2108	29,17	27,27	30,77
MediaM	-9,56***	0,1038	54,17	90,91	23,08
<b>Stacked com os modelos CAPM, D-CAPM e L-CAPM</b>					
St-LR	0,31***	0,0068	79,17	90,91	69,23
St-RF	0,21***	0,0078	54,17	72,73	38,46
St-GBRT	0,06***	0,0093	54,17	63,64	46,15
St-NN	0,02***	0,0096	75,00	90,91	61,54
<b>Painel B: Carteira <i>Median</i>-ME</b>					
	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,71*	0,0018	87,50	84,62	90,91
D-CAPM	0,22***	0,0049	87,50	100,00	72,73
L-CAPM	-0,12***	0,0069	54,17	100,00	0,00
C-CAPM	0,67***	0,002	87,50	92,31	81,82
F&F (1996)	0,00***	0,0062	75,00	76,92	72,73
MediaM	0,62***	0,0024	91,67	92,31	90,91
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,68***	0,0020	91,67	92,31	90,91
St-RF	0,65**	0,0022	95,83	92,31	100,00
St-GBRT	0,64**	0,0023	91,67	92,31	90,91
St-NN	0,42***	0,0036	83,33	100,00	63,64
<b>Painel C: <i>Large</i>-ME</b>					
	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,73	0,001	83,33	83,33	83,33
D-CAPM	0,24	0,0027	75,00	100,00	50,00
L-CAPM	-0,17	0,0042	50,00	100,00	0,00
C-CAPM	0,68	0,0011	79,17	75,00	83,33
F&F (1996)	-0,04	0,0038	70,83	75,00	66,67
MediaM	0,76	0,0008	83,33	83,33	83,33
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,00***	0,0036	75,00	50,00	100,00
St-RF	-1,22	0,008	75,00	50,00	100,00
St-GBRT	-2,55*	0,0128	83,33	75,00	91,67
St-NN	0,81	0,0007	83,33	83,33	83,33

Fonte: Elaborada pelo autor (2024). Notas: CAPM: Capital Asset Pricing Model (modelo de precificação de ativos de capital); D-CAPM: é o modelo Downside CAPM; L-CAPM: é o modelo Learning CAPM; C-CAPM: é o modelo Conditional CAPM; F&F: é o modelo de precificação de ativos de Fama e French (1996). Abreviaturas do nome do modelo: St significa stacked. St-LR: é o stacked com o modelo de Regressão Linear; St-RF: é o stacked com o modelo de Random Forest; St-GBRT: é o stacked com o modelo de Gradient Boosted Regression Trees; St-NN: é o stacked com o modelo de Neural Networks; MediaM: é a combinação de média simples dos modelos financeiros. MSFE é o erro quadrático médio de previsão; Acu. (%) é a medida de acurácia; Sen. (%) é a medida de sensibilidade; e Esp. (%) é a medida de especificidade. \*, \*\*, e \*\*\* indicam significância estatística nos níveis de 10%, 5% e 1%, respectivamente.

Fonte: Elaborada pelo autor (2024)

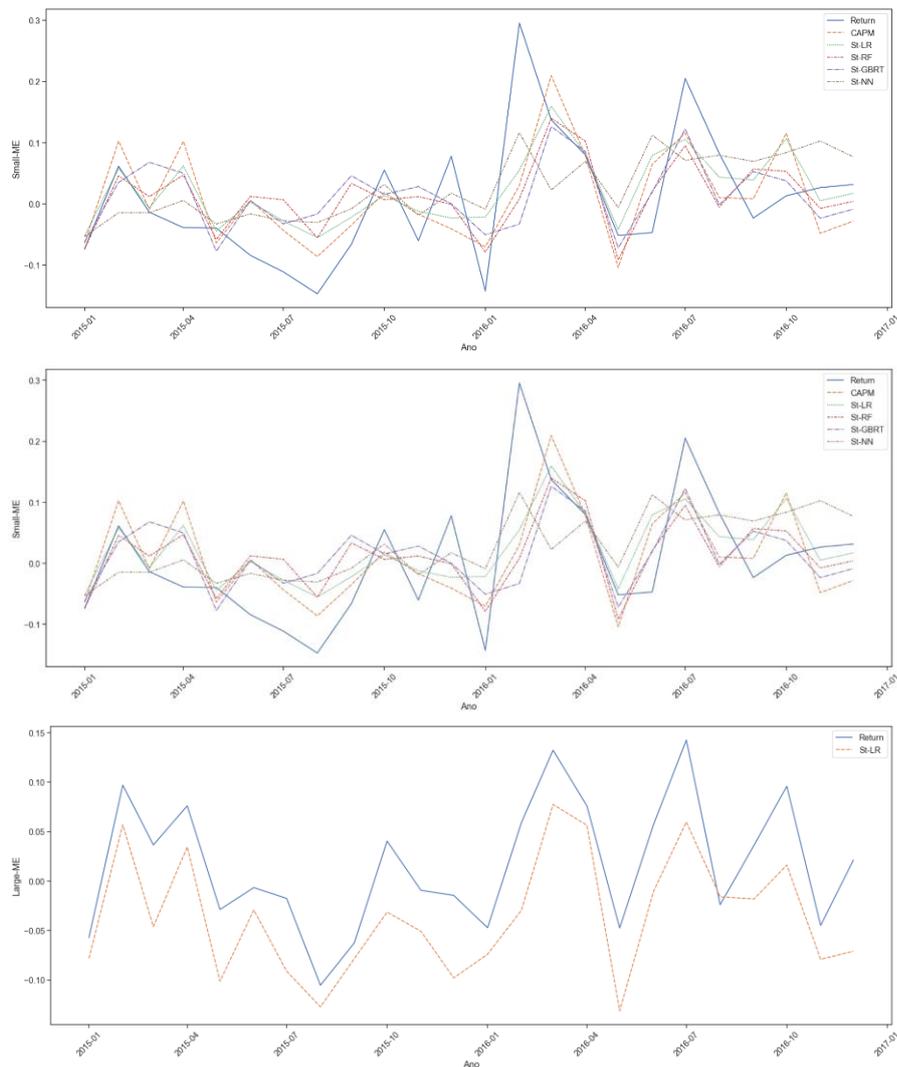
A decisão de analisar apenas três das cinco carteiras de ativos foi tomada com base nos resultados obtidos no conjunto de teste ao considerar os seguintes pontos: (i) o desempenho de  $R_{OOS}^2$  das carteiras *Small-Median-ME*, *Median-Large-ME* e *Large-ME* no conjunto de validação foram, em geral, piores do que a MH. Estes resultados negativos persistiram no conjunto de teste e por este motivo, optou-se por manter apenas a *Large-ME* como referência às empresas de grande porte; e (ii) nenhum dos modelos de precificação de ativos das carteiras *Small-Median-ME* e *Median-Large-ME* foram significativamente melhores do que a MH no conjunto de teste, com exceção do meta-modelo D-CAPM na carteira *Small-Median-ME* com um  $R_{OOS}^2$  de 0,01, não muito diferente de zero, e do meta-modelo St-NN na carteira *Median-Large-ME* que obteve  $R_{OOS}^2$  de 0,45.

Os resultados da Tabela 5 indicam que na carteira *Small-ME* apenas o modelo CAPM é significativamente melhor do que a MH no período com um  $R_{OOS}^2$  de 0,21 a um nível de significância de 5%. Além disso, este modelo apresenta, ao lado o meta-modelo St-LR, o melhor desempenho na previsão de instâncias negativas, com uma especificidade de 69,23%. Por outro lado, todos os meta-modelos de *stacked* foram significativamente melhores do que a MH no período. Assim como no conjunto completo, o destaque foi o meta-modelo St-LR com um  $R_{OOS}^2$  de 0,31 ao nível de significância de 1%, acurácia de 79%, sensibilidade de aproximadamente 91% e especificidade de 69%. Estes resultados indicam que os meta-modelos com a seleção dos melhores  $R_{OOS}^2$  foram capazes de produzir resultados superiores aos modelos de precificação de ativos durante na carteira das empresas de pequeno porte, durante o período analisado.

Semelhantemente aos resultados de previsão realizado em todo o conjunto de teste, observa-se que a carteira *Median-ME* apresenta um desempenho superior ao das pequenas empresas e empresas de grande porte. Com exceção do modelo L-CAPM, todos os modelos e meta-modelos foram significativamente melhores do que a MH do período. Os destaques dos modelos de precificação de ativos em termos de  $R_{OOS}^2$  e especificidade foram o CAPM (0,71 e 90%) e o C-CAPM (0,67 e 82%), respectivamente. Nenhum modelo de precificação de ativos foi significativamente melhor do que a MH na carteira de grandes empresas, e apenas o meta-modelo St-LR obteve um resultado igual a média histórica do período, revelando o desafio em prever o comportamento das ações de grandes empresas durante momentos de crise.

No tocante a técnica de *stacked generalization*, todos os meta-modelos foram significativamente melhores do que a MH na previsão das empresas de pequeno e médio porte. O meta-modelo St-LR obteve o melhor desempenho em termos de  $R_{OOS}^2$  dentre todas as

previsões nas duas carteiras de ativos, com resultados equivalentes a 0,31 na carteira de pequenas empresas e 0,68 na carteira das empresas de médio porte ao nível de significância estatística de 1%. Os meta-modelos também superaram o desempenho dos modelos base em termos de acurácia e especificidade, com maior evidência na carteira das empresas de médio porte, com previsões superiores a 90%. Estes resultados parecem evidenciar que os meta-modelos podem melhorar as previsões dos modelos de precificação de ativos tanto em termos estatísticos quanto em termos de previsão de sinal em momentos de crise.



**Figura 8 – Retorno Previsto das Ações durante a Crise Político-Econômica de 2016**

Fonte: Elaborada pelo autor (2024).

A Figura 8 apresenta a série temporal para os retornos previstos dos principais modelos em cada uma das carteiras estudada. Na carteira das pequenas empresas, os meta-modelos St-LR (ponto verde) e St-RF (traço vermelho) demonstram bom desempenho,

enquanto o modelo CAPM (traço laranja) apresenta menor acurácia geral, mas possui forte inclinação para acompanhar movimentos extremos, o que pode justificar o bom desempenho na previsão de instâncias negativas. Na carteira das empresas de médio porte, todos os modelos e meta-modelos analisados seguem fortemente a direção do retorno dos ativos, com exceção do modelo D-CAPM (ponto verde) que parece mais suave em alguns momentos. Já na carteira das empresas de grande porte, o único modelo ou meta-modelo com desempenho significativo é o St-LR que consegue acompanhar suavemente a movimentação do mercado, mas devido a falta de intensidade, obteve um desempenho literalmente igual a MH.

Os resultados da Tabela 6 indicam que diferentemente dos resultados encontrados na análise anterior, nenhum modelo de precificação de ativos foi significativamente melhor do que MH no período de janeiro de 2020 a junho de 2021, afetado pela pandemia de Covid-19. Esta diferença pode ser explicada pela característica das crises estudadas. Como a crise de 2016 teve um contexto político e econômico, percebemos ao analisar a Figura 8 que o mercado precificou os movimentos de queda, em pelo menos, 24 meses antes do período analisado. Desta forma, os métodos de previsão foram capazes de acompanhar os movimentos de mercado à medida em que a crise se desenrolava. Todavia, a crise de 2020 gerada pela pandemia de Covid-19 pode ser considerada como um choque exógeno. Ou seja, durante este período o mercado não conseguiu prever o movimento de queda antecipadamente. Além disso, as medidas de saúde pública que geraram *lockdowns* e o fechamento das empresas não essenciais, podem ter afetado a confiança do mercado, e, em consequência disto, os modelos de precificação de ativos estudados nesta pesquisa não foram capazes de superar a média histórica dos retornos das ações de forma significativa.

Para treinamento e previsão dos meta-modelos de *stacked generalization*, utilizou-se a mesma metodologia da análise anterior, em que foram considerados os modelos de precificação de ativos com os três melhores  $R_{OOS}^2$  no conjunto de validação. Os resultados apontam que apenas os meta-modelos St-RF e St-GBRT na carteira *Small-ME* e o meta-modelo St-NN na carteira *Median-ME* foram significativamente melhores do que a MH no período analisado. Enquanto na carteira *Large-ME* nenhum meta-modelo obteve um resultado positivo e significativo. Em termos de  $R_{OOS}^2$  o único resultado expressivo foi gerado pelo meta-modelo St-NN na carteira das empresas de médio porte com um  $R_{OOS}^2$  de 0,50 ao nível de significância estatística a 1%, acurácia de 72%, sensibilidade de 100% e especificidade de apenas 17%. Em geral, os resultados da medida de especificidade foram extremamente baixos para todos os modelos significativos, com no máximo 50% para os modelos St-RF e St-GBRT na carteira *Small-ME*.

**Tabela 6 – Resultado de Previsão durante a Pandemia de Covid-19 em 2020**

janeiro de 2020 a junho de 2021					
<b>Painel A: Carteira <i>Small</i>-ME</b>	$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)
<b>Modelos base</b>					
CAPM	0,47	0,0071	88,89	83,33	100,00
D-CAPM	-0,23	0,0167	88,89	91,67	83,33
L-CAPM	-0,43	0,0193	66,67	100,00	0,00
C-CAPM	-67,56	0,9282	61,11	75,00	33,33
F&F (1996)	-1,61***	0,0354	33,33	16,67	66,67
MediaM	-6,10	0,0961	61,11	66,67	50,00
<b>Stacked com os modelos CAPM, D-CAPM e L-CAPM</b>					
St-LR	0,51	0,0066	88,89	91,67	83,33
St-RF	0,07***	0,0126	77,78	91,67	50,00
St-GBRT	0,19**	0,011	77,78	91,67	50,00
St-NN	-0,09**	0,0148	72,22	100,00	16,67
<b>Painel B: Carteira <i>Median</i>-ME</b>					
$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)	
<b>Modelos base</b>					
CAPM	0,80	0,0021	88,89	83,33	100,00
D-CAPM	0,08	0,0096	88,89	83,33	100,00
L-CAPM	-0,39	0,0145	66,67	100,00	0,00
C-CAPM	0,85	0,0015	83,33	83,33	83,33
F&F (1996)	-0,36	0,0142	55,56	66,67	33,33
MediaM	0,66	0,0035	77,78	75,00	83,33
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,79	0,0022	77,78	75,00	83,33
St-RF	0,57	0,0044	83,33	83,33	83,33
St-GBRT	0,61	0,004	83,33	83,33	83,33
St-NN	0,50***	0,0053	72,22	100,00	16,67
<b>Painel C: <i>Large</i>-ME</b>					
$R^2_{00s}$	MSFE	Acu. (%)	Sen. (%)	Esp. (%)	
<b>Modelos base</b>					
CAPM	0,85	0,0008	94,44	90,91	100,00
D-CAPM	0,15	0,0047	94,44	90,91	100,00
L-CAPM	-0,38	0,0076	61,11	100,00	0,00
C-CAPM	0,70	0,0017	88,89	90,91	85,71
F&F (1996)	-0,30	0,0072	50,00	63,64	28,57
MediaM	0,80	0,0011	88,89	90,91	85,71
<b>Stacked com os modelos CAPM, C-CAPM e F&amp;F (1996)</b>					
St-LR	0,55	0,0025	83,33	81,82	85,71
St-RF	0,62	0,0021	94,44	100,00	85,71
St-GBRT	0,86	0,0008	94,44	100,00	85,71
St-NN	0,72	0,0016	88,89	81,82	100,00

Fonte: Elaborada pelo autor (2024). Notas: CAPM: Capital Asset Pricing Model (modelo de precificação de ativos de capital); D-CAPM: é o modelo Downside CAPM; L-CAPM: é o modelo Learning CAPM; C-CAPM: é o modelo Conditional CAPM; F&F: é o modelo de precificação de ativos de Fama e French (1996). Abreviaturas do nome do modelo: St significa stacked. St-LR: é o stacked com o modelo de Regressão Linear; St-RF: é o stacked com o modelo de Random Forest; St-GBRT: é o stacked com o modelo de Gradient Boosted Regression Trees; St-NN: é o stacked com o modelo de Neural Networks; MediaM: é a combinação de média simples dos modelos financeiros. MSFE é o erro quadrático médio de previsão; Acu. (%) é a medida de acurácia; Sen. (%) é a medida de sensibilidade; e Esp. (%) é a medida de especificidade. \*, \*\*, e \*\*\* indicam significância estatística nos níveis de 10%, 5% e 1%, respectivamente.

Fonte: Elaborada pelo autor (2024)

## 5. CONSIDERAÇÕES FINAIS

---

Esta pesquisa teve como objetivo investigar se os retornos das ações dos modelos de precificação de ativos são potencializados quando utilizada a técnica de *stacked generalization* em *machine learning* nos ativos da B3. Os modelos de precificação de ativos como o CAPM desenvolvido por Sharpe (1964) e Lintner (1965) na década de 1970, o modelo APT de Ross (1976) e os modelos multifatoriais de Três Fatores de Fama e French (1993), Quatro Fatores de Carhart (1997) e Cinco Fatores de Fama e French (2015) são cruciais na explicação de grande parte da variação dos retornos das ações. Já a técnica de *stacked generalization* é uma técnica para combinar previsões desenvolvida por Wolpert (1992). Para Zhao e Cheng (2022) as combinações de previsões são frequentemente consideradas melhores do que as previsões individuais em estudos empíricos, pois geralmente são mais precisas e estabilizam as previsões de maneira semelhante à diversificação entre ativos individuais que reduz a variância de um portfólio.

A metodologia utilizada nesta pesquisa baseia-se na proposta de Zhao e Cheng (2022) que empregaram a técnica para refinar uma variedade de modelos de previsão de retornos de ações individuais, lineares e não lineares, na previsão de retorno do prêmio de ações no mercado dos EUA para o período de 1927 a 2018. Para atingir os objetivos específicos da pesquisa, a amostra completa de janeiro de 2003 a dezembro de 2022 foi dividida em três conjuntos de dados: (i) treinamento, (ii) validação e (iii) teste. Em seguida, utilizou-se a abordagem de janela de expansão proposta por Timmermann (2008), na qual os dados são segmentados em partes, seguindo um formato de estimativas recursivas, para realizar as previsões dos modelos. Por fim, as previsões foram avaliadas por meio de duas medidas de eficácia:  $R_{OS}^2$  fora da amostra de Campbell e Thompson (2008) e previsões de sinal, composta pela medidas de acurácia, especificidade e sensibilidade.

Os resultados da pesquisa sugerem que: (i) utilizar a técnica de *stacked generalization* para combinar previsões de modelos de precificação de ativos de capital pode melhorar significativamente o resultado das previsões em diferentes contextos e condições de mercado; (ii) realizar o *stacked generalization* com meta-modelos mais simples, como St-LR, pode produzir resultados melhores do que com meta-modelos mais complexos; (iii) treinar os meta-modelos de *stacked generalization* com a seleção dos melhores modelos base pode melhorar a previsão de acurácia negativa, especialmente em períodos de queda; (iv) não há diferenças

significativas, em termos de  $R_{OOS}^2$ , na comparação entre os resultados dos meta-modelos treinados com a seleção dos melhores modelos base e os meta-modelos treinados sem essa seleção; (v) nenhum modelo ou meta-modelo é capaz de superar a média histórica dos retornos das ações de forma significativa durante choques exógenos; e (vi) embora os modelos de precificação de ativos tenham seu lugar, técnicas *de machine learning* como o *stacked generalization* podem ser mais robustas e consistentes na previsão de retornos de ações em uma variedade de cenários.

Para atender ao primeiro objetivo específico da pesquisa, que consiste em analisar a eficácia dos modelos de precificação de ativos na estimativa do retorno das ações na B3, utilizou-se, inicialmente, toda a amostra de teste cobrindo o período de janeiro de 2013 a dezembro de 2022, para avaliar o desempenho dos modelos em três carteiras de ativos: empresas de pequeno porte (*Small-ME*), empresas de médio porte (*Median-ME*) e grandes empresas (*Large-ME*). Os resultados sugerem que nenhum modelo de precificação de ativos é significativamente melhor do que a média histórica (MH) em termos de  $R_{OOS}^2$  nas três carteiras de ativos, simultaneamente, embora possamos observar alguns destaques individuais como o modelo D-CAPM na carteira de pequeno e médio porte, o modelo C-CAPM na carteira das médias empresas, e o modelo F&F na carteira das médias e grandes empresas. No geral, os modelos de precificação de ativos apresentaram baixa eficácia na previsão do retorno das ações no período, tanto em termos de  $R_{OOS}^2$  quanto em relação as medidas de acurácia, que invariavelmente são melhores em prever corretamente instâncias positivas do que instâncias negativas.

A eficácia dos modelos de precificação de ativos também foi avaliada em condições extremas de mercado, mais especificamente, durante a crise político-econômica que afetou a bolsa de valores brasileira em 2016, e a crise gerada pela pandemia de Covid-19 em 2020. As evidências sugerem que, semelhantemente aos resultados obtidos na amostra de teste, os modelos de precificação de ativos são mais eficazes em prever o comportamento das empresas de médio porte. Em ambos os períodos, apenas o modelo CAPM foi significativamente melhor do que a MH na previsão das pequenas empresas, e nenhum modelo de precificação de ativos obteve resultado positivo e significativo na carteira das grandes empresas, assim como durante a crise pela pandemia de Covid-19. A ineficiência dos modelos de precificação de ativos em prever o comportamento de queda entre janeiro de 2020 e junho de 2021 pode ser explicado pelo choque exógeno gerado pela pandemia de Covid-19. Observamos que durante a crise de 2016 o mercado precificou os movimentos de queda, pelo menos, 24 meses antes do período analisado. No entanto, a pandemia de Covid-19 foi repentina, um choque exógeno, em que o

mercado não precificou o movimento de queda com antecedência, e em consequência disto, os modelos de precificação de ativos não conseguiram obter resultados positivos e significativos.

Em relação ao segundo objetivo específico da pesquisa, que consiste em avaliar se a técnica de *stacked generalization* em *machine learning* aplicada aos modelos de precificação de ativos tem potencial para otimizar o retorno das ações na B3. Utilizou-se inicialmente, toda a amostra de teste cobrindo o período de janeiro de 2013 a dezembro de 2022 para avaliar o desempenho de quatro meta-modelos de *stacked generalization*: *Random Forest* (St-RF); *Gradient Boosted Regression Trees* (St-GBRT), *Neural Networks* (St-NN) e *Linear Regression* (St-LR). Neste primeiro momento, a medida de  $R_{00S}^2$  não foi utilizada para selecionar os melhores modelos base. Em outras palavras, os cinco modelos de precificação de ativos foram utilizados no treinamento dos meta-modelos.

As evidências indicam que os meta-modelos, em particular os modelos St-LR, St-RF e St-NN, demonstraram consistentemente resultados superiores aos modelos de precificação de ativos em todas as três carteiras estudadas. Com exceção do meta-modelo St-GBRT na carteira das pequenas empresas, todos os meta-modelos de *stacked generalization* foram significativamente melhores do que a MH na previsão das empresas de pequeno e médio porte. Todavia, apenas o meta-modelo St-NN obteve bom desempenho na previsão de grandes empresas, ressaltando a dificuldade em prever corretamente o desempenho das empresas maiores. O maior destaque foi o meta-modelo St-LR, que superou todos os demais modelos e meta-modelos na previsão das carteiras das empresas de pequeno e médio porte, com  $R_{00S}^2$  de 0,34 e 0,71, respectivamente. Em termos de previsão de sinal, em média, os meta-modelos superaram o desempenho dos modelos de precificação de ativos, considerando as três medidas de avaliação nas três carteiras de ativos. Além disso, os meta-modelos demonstraram melhor desempenho em prever instâncias negativas do que os modelos de precificação de ativos. A capacidade dos meta-modelos em prever movimentos de queda pode ser o principal motivo para explicar a diferença de desempenho entre os modelos.

Em seguida, analisou-se o desempenho dos meta-modelos seguindo a metodologia utilizada por Zhao e Cheng (2022), em que os modelos base com os melhores  $R_{00S}^2$  no conjunto de validação são selecionados para treinar os meta-modelos. Nesta pesquisa, foram selecionados os três modelos de precificação de ativos com os melhores  $R_{00S}^2$  no conjunto de dados de validação. O objetivo deste método consiste em testar se realizar o treinamento dos meta-modelos com os melhores modelos base pode melhorar os resultados das previsões.

Os resultados apontam que os meta-modelos treinados com os melhores modelos base, melhoram significativamente as previsões dos modelos de precificação de ativos em todas as

três carteiras de ativos estudadas. No entanto, não foram encontradas diferenças significativas entre os meta-modelos treinados com os melhores modelos base e os meta-modelos sem essa seleção. Na realidade, com poucas exceções, a maioria dos meta-modelos apresentaram desempenho semelhante ou pior. Todavia, ao considerar a acurácia negativa, os meta-modelos treinados com os melhores modelos base foram geralmente melhores do que os modelos de precificação de ativos e meta-modelos treinados sem essa seleção.

Por fim, como os meta-modelos com seleção de  $R_{OOS}^2$  apresentaram melhor capacidade de prever eventos negativos, analisou-se o desempenho das previsões durante períodos de condições extremas de mercado, como a crise político-econômica de 2016 e a pandemia de Covid-19 em 2020. As evidências indicam que todos os meta-modelos foram significativamente melhores do que a MH na previsão das empresas de pequeno e médio porte no período de janeiro de 2015 a dezembro de 2016. Os meta-modelos também demonstraram melhor eficácia em termo de especificidade (acurácia negativa), com maior evidência na carteira das empresas de médio porte, com previsões superiores a 90%. Estes resultados parecem evidenciar que os meta-modelos podem melhorar as previsões dos modelos de precificação de ativos tanto em termos estatísticos quanto em termos de previsão de sinal em momentos de crise. Todavia, assim como os modelos de precificação de ativos, em geral, os meta-modelos demonstraram baixa capacidade em prever o comportamento de queda entre janeiro de 2020 e junho de 2021.

## 5.1 LIMITAÇÕES DA PESQUISA E SUGESTÕES DE ESTUDOS FUTUROS

Uma das principais limitações desta pesquisa relaciona-se ao período completo da amostra de dados. Enquanto a pesquisa utilizou dados de janeiro de 2003 a dezembro de 2022 (20 anos). Pesquisas internacionais sobre o tema, principalmente aquelas realizadas em mercados maduros como o mercado dos EUA, comumente utilizam conjuntos de dados ainda mais extensos, com décadas de informações sobre variáveis macroeconômicas, comportamento das empresas e retorno das ações. Alguns exemplos incluem Rapach, Strauss e Zhou (2010) que utilizaram dados de janeiro de 1947 a dezembro de 2005 (59 anos); Rapach e Zhou (2013) com dados de dezembro de 1926 a dezembro de 2010 (84 anos); e Lima e Meng (2017) com dados de dezembro de 1926 a dezembro de 2013 (87 anos).

Outra limitação da pesquisa consiste na metodologia de seleção dos meta-modelos de *stacked generalization*. Nesta pesquisa, foram considerados os meta-modelos utilizados por

Zhao e Cheng (2022). Todavia, os autores não determinaram um método claro pelos quais os algoritmos de *machine learning* foram selecionados. Por outro lado, com relação aos modelos de precificação de ativos usados como modelos base, a principal limitação é que mesmo tendo sido identificados por meio de uma pesquisa de levantamento, estes modelos podem não representar as melhores escolhas para aplicar a técnica de *stacked generalization*, ou mesmo não obterem os melhores desempenhos na previsão de retorno das ações do mercado brasileiro.

Por fim, a maior limitação da pesquisa consiste no fato de que não há evidências na literatura de previsão de retorno das ações, em que a técnica de *stacked generalization* tenha sido utilizada para combinar previsões de modelos que não sejam algoritmos de *machine learning*. Desta forma, esta pesquisa enfrentou vários desafios tecnológicos e científicos durante a sua condução, o que limitou melhorias que poderiam ser realizadas, como as seleções de hiperparâmetros dos modelos de *stacked generalization*, validações da amostra, e variações dos modelos base na composição dos meta-modelos.

Como sugestões para estudos futuros, recomenda-se aumentar o número de modelos base com a inclusão de outros modelos precificação de ativos, como os modelos multifatoriais: Três Fatores de Fama e French (1993), Quatro Fatores de Carhart (1997) e Cinco Fatores de Fama e French (2015). Esta sugestão tem duas finalidades: (i) possibilitar que a técnica de *stacked generalization* capture o maior número possível de relacionamento dos dados; e (ii) permitir que a técnica seja aplicada em um maior número de carteiras. Além disso, recomenda-se testar o maior número possível de composições dos modelos base durante o treinamento dos meta-modelos, variando os parâmetros dos meta-modelos, até encontrar o conjunto de modelos base que apresentem os melhores resultados de previsão no retorno das ações.

## REFERÊNCIAS

---

ABREU, André Fidelis Figueiredo de. **Aplicação de machine learning na pré-seleção de ativos para portfólios de investimento**. 2021. Dissertação de Mestrado.

ADRIAN, Tobias; FRANZONI, Francesco. **Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM**. *Journal of Empirical Finance*, v. 16, n. 4, p. 537-556, 2009.

AHARONI, Gil; GRUNDY, Bruce; ZENG, Qi. **Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis**. *Journal of financial economics*, v. 110, n. 2, p. 347-357, 2013.

AKAIKE, Hirotugu. **A new look at the statistical model identification**. *IEEE transactions on automatic control*, v. 19, n. 6, p. 716-723, 1974.

ALI, Fahad; KHURRAM, Muhammad Usman; JIANG, Yuexiang. **The five-factor asset pricing model tests and profitability and investment premiums: evidence from Pakistan**. *Emerging Markets Finance and Trade*, v. 57, n. 9, p. 2651-2673, 2019.

ARMSTRONG, J. Scott. **Combining forecasts**. Springer US, 2001.

BANZ, Rolf W. **The relationship between return and market value of common stocks**. *Journal of financial economics*, v. 9, n. 1, p. 3-18, 1981.

BARROS, Emílio Augusto Coelho; MAZUCHELI, Josmar. **Um estudo sobre o tamanho e poder dos testes t-Student e Wilcoxon**. *Acta Scientiarum. Technology*, v. 27, n. 1, p. 23-32, 2005.

BATES, John M.; GRANGER, Clive WJ. **The combination of forecasts**. *Journal of the operational research society*, v. 20, n. 4, p. 451-468, 1969.

BLACK, Fischer. **Capital market equilibrium with restricted borrowing**. *Journal of Business*, v. 45, p. 444-455, 1972.

BLUVOL, Leonardo. **Análise de algoritmos de machine learning para previsão de preços de IBOVESPA**. 2022. Dissertação de Mestrado.

BRAV, Alon; GECZY, Christopher; GOMPERS, Paul A. **Is the abnormal return following equity issuances anomalous?**. *Journal of financial economics*, v. 56, n. 2, p. 209-249, 2000.

BREIMAN, Leo. **Random forests**. *Machine learning*, v. 45, p. 5-32, 2001.

- BUNN, Derek W. **A Bayesian approach to the linear combination of forecasts**. Journal of the Operational Research Society, v. 26, n. 2, p. 325-329, 1975.
- BUNN, Derek W. **Statistical efficiency in the linear combination of forecasts**. International Journal of Forecasting, v. 1, n. 2, p. 151-163, 1985.
- CAMPBELL, John Y.; THOMPSON, Samuel B. **Predicting excess stock returns out of sample: Can anything beat the historical average?**. The Review of Financial Studies, v. 21, n. 4, p. 1509-1531, 2008.
- CAMPBELL, John Y.; VUOLTEENAHO, Tuomo. **Bad beta, good beta**. American Economic Review, v. 94, n. 5, p. 1249-1275, 2004.
- CARHART, Mark M. **On persistence in mutual fund performance**. The Journal of finance, v. 52, n. 1, p. 57-82, 1997.
- CAVALERI, Rosangela; RIBEIRO, Eduardo Pontual. **Combinação de previsões de volatilidade: um estudo**. Revista Economia, v. 12, n. 2, p. 239-261, 2011.
- CHEN, Nai-Fu; ROLL, Richard; ROSS, Stephen A. **Economic forces and the stock market**. Journal of business, v. 59, n. 3, p. 383-403, 1986.
- CLARK, Todd E.; WEST, Kenneth D. **Approximately normal tests for equal predictive accuracy in nested models**. Journal of econometrics, v. 138, n. 1, p. 291-311, 2007.
- CLEMEN, Robert T. **Combining forecasts: A review and annotated bibliography**. International journal of forecasting, v. 5, n. 4, p. 559-583, 1989.
- CLEMEN, Robert T.; WINKLER, Robert L. **Combining economic forecasts**. Journal of Business & Economic Statistics, v. 4, n. 1, p. 39-46, 1986.
- CRESWELL, John W. **Projeto de pesquisa: Métodos qualitativo, quantitativo e misto**. 3. ed. - Dados eletrônicos. - Porto Alegre : Artmed, 2010.
- DHRYMES, Phoebus J.; FRIEND, Irwin; GULTEKIN, N. Bulent. **A critical reexamination of the empirical evidence on the arbitrage pricing theory**. The Journal of Finance, v. 39, n. 2, p. 323-346, 1984.
- DONALDSON, R. Glen; KAMSTRA, Mark. **Forecast combining with neural networks**. Journal of Forecasting, v. 15, n. 1, p. 49-61, 1996.
- ESTRADA, Javier. **The cost of equity in emerging markets: a downside risk approach**. 2000.
- FAMA, Eugene F.; FRENCH, Kenneth R. **A five-factor asset pricing model**. Journal of financial economics, v. 116, n. 1, p. 1-22, 2015.

FAMA, Eugene F.; FRENCH, Kenneth R. **Common risk factors in the returns on stocks and bonds**. Journal of financial economics, v. 33, n. 1, p. 3-56, 1993.

FAMA, Eugene F.; FRENCH, Kenneth R. **Multifactor explanations of asset pricing anomalies**. The journal of finance, v. 51, n. 1, p. 55-84, 1996.

FAMA, Eugene F.; FRENCH, Kenneth R. **Size, value, and momentum in international stock returns**. Journal of financial economics, v. 105, n. 3, p. 457-472, 2012.

FAMA, Eugene F.; FRENCH, Kenneth R. **The capital asset pricing model: Theory and evidence**. Journal of economic perspectives, v. 18, n. 3, p. 25-46, 2004.

FAMA, Eugene F.; FRENCH, Kenneth R. **The cross-section of expected stock returns**. The Journal of Finance, v. 47, n. 2, p. 427-465, 1992.

FARIA, Adriana Moreira Bastos de; SANTOS, Lucas Maia dos. **Análise comparativa dos modelos de precificação de ativos Capital Asset Pricing Model e Downside Capital Asset Pricing Model**. Revista da FAE, v. 13, n. 2, 2010.

FRIEND, Irwin; BLUME, Marshall. **Measurement of portfolio performance under uncertainty**. The American economic review, v. 60, n. 4, p. 561-575, 1970.

GERONAZZO, Arthur. **Técnicas de estresse teste de mercado usando maximum drawdown**. 2019. Dissertação de Mestrado.

GIBBONS, Michael R.; ROSS, Stephen A.; SHANKEN, Jay. **A test of the efficiency of a given portfolio**. Econometrica: Journal of the Econometric Society, p. 1121-1152, 1989.

GIL, Antonio Carlos et al. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GIOVANNELLI, Alessandro; MASSACCI, Daniele; SOCCORSI, Stefano. **Forecasting stock returns with large dimensional factor models**. Journal of Empirical Finance, v. 63, p. 252-269, 2021.

GRANGER, Clive WJ; RAMANATHAN, Ramu. **Improved methods of combining forecasts**. Journal of forecasting, v. 3, n. 2, p. 197-204, 1984.

GUO, Bin et al. **The five-factor asset pricing model tests for the Chinese stock market**. Pacific-Basin Finance Journal, v. 43, p. 84-106, 2017.

GUPTA, Sunil; WILTON, Peter C. **Combination of forecasts: an extension**. Management Science, v. 33, n. 3, p. 356-372, 1987.

HOLMEN, Jay S. **A note on the value of combining short-term earnings forecasts: a test of Granger and Ramanathan**. International Journal of Forecasting, v. 3, n. 2, p. 239-243, 1987.

HUBERMAN, Gur. **A simple approach to arbitrage pricing theory**. Journal of Economic Theory, v. 28, n. 1, p. 183-191, 1982.

JAFFE, Jeffrey; KEIM, Donald B.; WESTERFIELD, Randolph. **Earnings yields, market values, and stock returns**. The Journal of Finance, v. 44, n. 1, p. 135-148, 1989.

JEGADEESH, Narasimhan. **Long-term performance of seasoned equity offerings: Benchmark errors and biases in expectations**. Financial Management, p. 5-30, 2000.

JEGADEESH, Narasimhan; TITMAN, Sheridan. **Returns to buying winners and selling losers: Implications for stock market efficiency**. The Journal of finance, v. 48, n. 1, p. 65-91, 1993.

JENSEN, Michael C.; BLACK, Fischer; SCHOLES, Myron S. **The capital asset pricing model: Some empirical tests**. 1972.

JING, Nan; WU, Zhao; WANG, Hefei. **A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction**. Expert Systems with Applications, v. 178, p. 115019, 2021.

KUMBURE, M. M. et al. **Machine learning techniques and data for stock market forecasting: A literature review**. Expert Systems with Applications, v. 197, p. 116659, 2022.

LAAN, Mark J; DUDOIT, Sandrine; VAART, Aad W. van der. **The cross-validated adaptive epsilon-net estimator**. Statistics & Decisions, v. 24, n. 3, p. 373-395, 2006.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos da metodologia científica. In: Fundamentos da metodologia científica**. 2010. p. 320-320.

LAKONISHOK, Josef; SHAPIRO, Alan C. **Systematic risk, total risk and size as determinants of stock market returns**. Journal of Banking & Finance, v. 10, n. 1, p. 115-132, 1986.

LIMA, Luiz Renato; MENG, Fanning. **Out-of-sample return predictability: A quantile combination approach**. Journal of Applied Econometrics, v. 32, n. 4, p. 877-895, 2017.

LIN, Hai; WU, Chunchi; ZHOU, Guofu. **Forecasting corporate bond returns with a large set of predictors: An iterated combination approach**. Management Science, v. 64, n. 9, p. 4218-4238, 2018.

LIN, Qi. **Noisy prices and the Fama–French five-factor asset pricing model in China**. Emerging Markets Review, v. 31, p. 141-163, 2017.

LINTNER, John. **The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets**. Review of Economics and Statistics, v. 47, p. 13-47, 1965.

MARKOWITZ, H. **Portfolio Selection**. The Journal of Finance, v. 7, No. 1, pp. 77-91, 1952.

MAZZEU, João Henrique Gonçalves; COSTA JUNIOR, Newton Carneiro Affonso da; SANTOS, André Alves Portela. **CAPM condicional com aprendizagem aplicado ao mercado brasileiro de ações**. RAM. Revista de Administração Mackenzie, v. 14, p. 143-175, 2013.

MEGGINSON, William L. **Corporate finance theory**. Addison-Wesley, 1997.

MILLER, Merton H.; MODIGLIANI, Franco. **Dividend policy, growth, and the valuation of shares**. the Journal of Business, v. 34, n. 4, p. 411-433, 1961.

MOSSIN, Jan. **Equilibrium in a capital asset market**. Econometrica: Journal of the econometric society, p. 768-783, 1966.

NEWBOLD, Paul; GRANGER, Clive WJ. **Experience with forecasting univariate time series and the combination of forecasts**. Journal of the Royal Statistical Society: Series A (General), v. 137, n. 2, p. 131-146, 1974.

NOVY-MARX, Robert. **The other side of value: The gross profitability premium**. Journal of financial economics, v. 108, n. 1, p. 1-28, 2013.

ÖLLER, L.-E. **A method for pooling forecasts**. Journal of the Operational Research Society, v. 29, n. 1, p. 55-63, 1978.

PESARAN, M. Hashem; TIMMERMANN, Allan. **Market timing and return prediction under model instability**. Journal of Empirical Finance, v. 9, n. 5, p. 495-510, 2002.

PORTELA, Daniel Lucas Martins. **A inclusão do fator de risco macroeconômico para explicar os retornos das carteiras no mercado acionário brasileiro: utilização em modelos multifatoriais**. 2016. Dissertação de Mestrado. Universidade Federal de Pernambuco.

RAPACH, David E.; STRAUSS, Jack K.; ZHOU, Guofu. **Out-of-sample equity premium prediction: Combination forecasts and links to the real economy**. The Review of Financial Studies, v. 23, n. 2, p. 821-862, 2010.

RAPACH, David E.; ZHOU, Guofu. **Time-series and cross-sectional stock return forecasting: New machine learning methods**. Machine learning for asset management: New developments and financial applications, p. 1-33, 2020.

RASEKHSCHAFFE, Keywan Christian; JONES, Robert C. **Machine learning for stock selection**. Financial Analysts Journal, v. 75, n. 3, p. 70-88, 2019.

RIZZI, Luciana Julio. **Análise comparativa de modelos para determinação do custo de capital próprio: CAPM, três fatores de Fama e French (1993) e quatro fatores de Carhart (1997)**. 2012. Tese de Doutorado. Universidade de São Paulo.

ROLL, Richard; ROSS, Stephen A. **An empirical investigation of the arbitrage pricing theory**. The journal of finance, v. 35, n. 5, p. 1073-1103, 1980.

ROSA, Filipe Gabriel Dahmer da. **Combinação de previsão de matrizes de covariância avaliada na aplicação em gestão de portfólios**. 2016. Dissertação de Mestrado.

ROSS, Stephen A. **The arbitrage theory of capital asset pricing**. Journal of Economic Theory, v. 13, p. 341-360, 1976.

ROUWENHORST, K. Geert. **International momentum strategies**. The journal of finance, v. 53, n. 1, p. 267-284, 1998.

ROUWENHORST, K. Geert. **Local return factors and turnover in emerging stock markets**. The journal of finance, v. 54, n. 4, p. 1439-1464, 1999.

RUTKOWSKA-ZIARKO, Anna et al. **Conventional and downside CAPM: The case of London stock exchange**. Global Finance Journal, v. 54, p. 100759, 2022.

SACCOL, Amarolinda et al. **Metodologia de pesquisa em administração: uma abordagem prática**. São Leopoldo: Unisinos, 2012.

SANTOS, Gustavo Carvalho et al. **Algoritmos de machine learning para previsão de ações da B3**. 2020. Dissertação de Mestrado.

SHAHHOSSEINI, Mohsen; HU, Guiping; PHAM, Hieu. **Optimizing ensemble weights and hyperparameters of machine learning models for regression problems**. Machine Learning with Applications, v. 7, p. 100251, 2022.

SHARPE, William F. **Capital asset prices: A theory of market equilibrium under conditions of risk**. The journal of finance, v. 19, n. 3, p. 425-442, 1964.

SILVA, Marcus Oliveira da. **Precificação empírica das ações brasileiras com o uso de machine learning e gradient boosted trees: aplicação no período 2007-2022**. 2023. Dissertação de Mestrado.

STOCK, James H.; WATSON, Mark W. **How did leading indicator forecasts perform during the 2001 recession?**. FRB Richmond Economic Quarterly, v. 89, n. 3, p. 71-90, 2003.

TERUI, Nobuhiko; VAN DIJK, Herman K. **Combined forecasts from linear and nonlinear time series models**. International Journal of Forecasting, v. 18, n. 3, p. 421-438, 2002.

TIMMERMANN, Allan. **Elusive return predictability**. International Journal of Forecasting, v. 24, n. 1, p. 1-18, 2008.

TOBIN, James. **Liquidity preference as behavior towards risk**. The review of economic studies, v. 25, n. 2, p. 65-86, 1958.

TORRES, Sérgio. **Previsão do preço de ações brasileiras utilizando redes neurais artificiais**. 2021. Dissertação de Mestrado.

VARGA, Gyorgy; BRITO, Ricardo D. **The cross-section of expected stock returns in brazil**. Revista Brasileira de Finanças, 2016.

VIRTANEN, Ilkka; YLI-OLLI, Paavo. **Forecasting stock market prices in a thin security market**. Omega, v. 15, n. 2, p. 145-155, 1987.

VON HOLSTEIN, Carl-Axel S. Staël. **Probabilistic forecasting: An experiment related to the stock market**. Organizational Behavior and Human Performance, v. 8, n. 1, p. 139-158, 1972.

WANG, X. et al. **Forecast combinations: an over 50-year review**. International Journal of Forecasting, 2022.

WEI, KC John. **An asset-pricing theory unifying the capm and apt**. The Journal of Finance, v. 43, n. 4, p. 881-892, 1988.

WELCH, Ivo; GOYAL, Amit. **A comprehensive look at the empirical performance of equity premium prediction**. The Review of Financial Studies, v. 21, n. 4, p. 1455-1508, 2008.

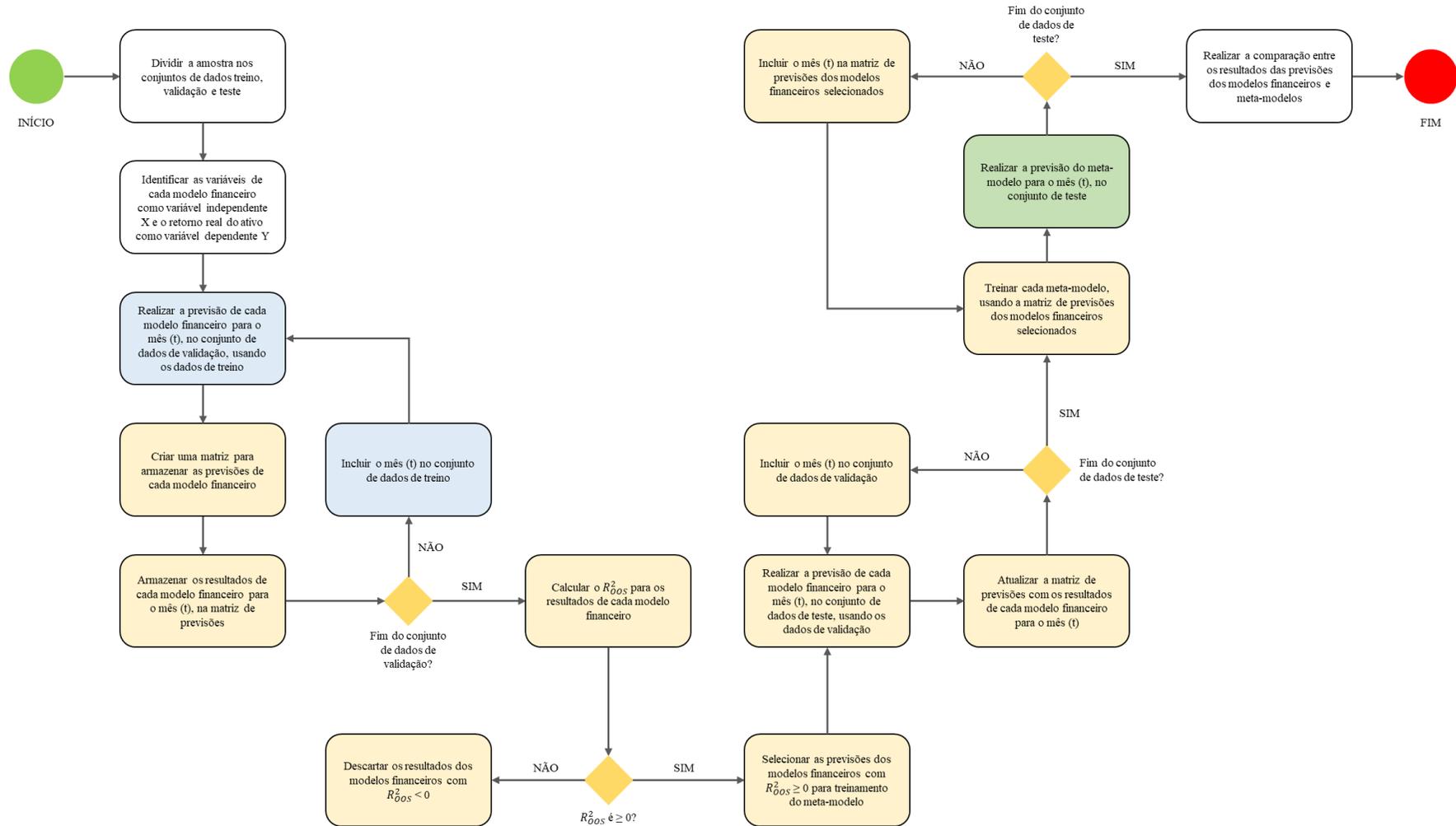
WINKLER, Robert L.; MAKRIDAKIS, Spyros. **The combination of forecasts**. Journal of the Royal Statistical Society: Series A (General), v. 146, n. 2, p. 150-157, 1983.

WOLPERT, David H. **Stacked generalization**. Neural networks, v. 5, n. 2, p. 241-259, 1992.

ZHANG, Hongwei et al. **Forecasting stock returns with model uncertainty and parameter instability**. Journal of Applied Econometrics, v. 35, n. 5, p. 629-644, 2020.

ZHAO, Albert Bo; CHENG, Tingting. **Stock return prediction: Stacking a variety of models**. Journal of Empirical Finance, v. 67, p. 288-317, 2022.

# APÊNDICE A – FLUXOGRAMA DE *STACKED GENERALIZATION* APLICADO AOS MODELOS DE PRECIFICAÇÃO DE ATIVOS

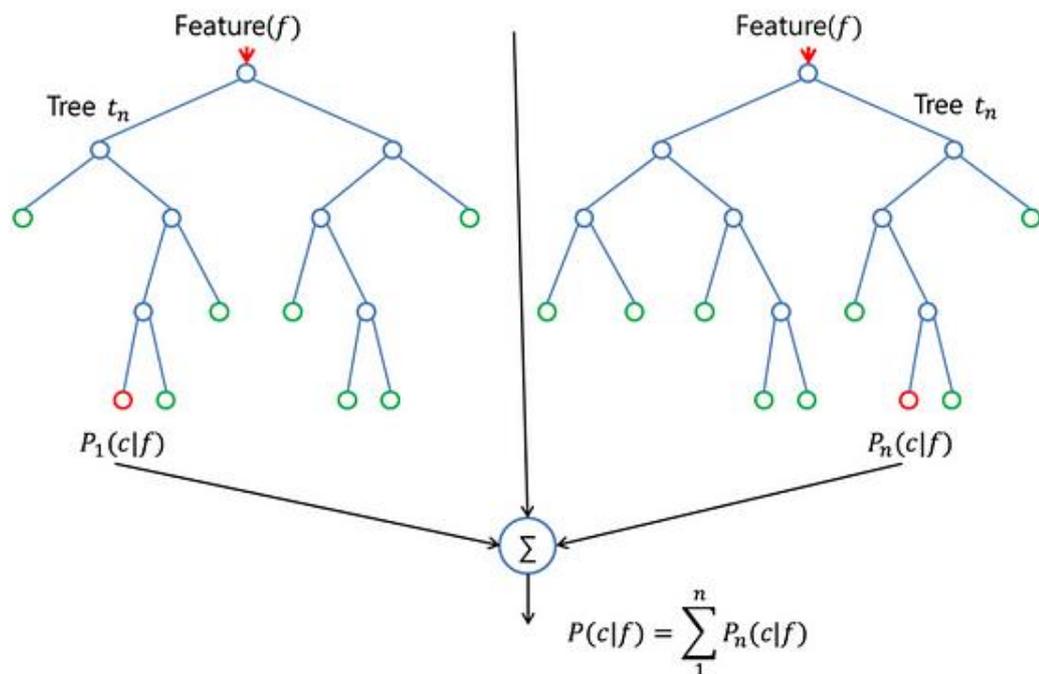


Fonte: Elaborada pelo autor (2024)

## APÊNDICE B – INTRODUÇÃO AOS META-MODELOS

### B.1 *Random Forest* (RF)

A Floresta Aleatória, *Random Forest* (RF), foi proposta formalmente por Breiman (2001). Como o nome sugere, é uma técnica usada para combinar previsões ou classificadores de um certo número de modelos conhecidos como Árvore de Decisão (*Decision Trees* – DT). A DT é um modelo intuitivo que faz perguntas simples sequencialmente. Uma pergunta geralmente é chamada de “nó”, e quando um nó não possui nós filhos, é chamado de nó folha, que encerra o processo de perguntar. A previsão é realizada pela média dos valores das amostras que caem no mesmo nó. As perguntas e as ordens dessas perguntas feitas nas DT's são essenciais para o desempenho dos modelos. Uma desvantagem da DT é que, embora a cada passo ele encontre o melhor par para dividir os dados, ele não necessariamente encontra a sequência ótima geral. Outra desvantagem desse modelo é que ele é instável. Uma pequena alteração nos dados geralmente resulta em uma alteração significativa do modelo, tornando o modelo menos propenso a generalizar bem para novos dados.



**Figura 9 – Exemplo de *Random Forest* (RF)**

Fonte: [www.medium.com/machina-sapiens](http://www.medium.com/machina-sapiens)

A RF é frequentemente utilizada para superar a instabilidade da DT, pois calcula a média das previsões sobre muitas árvores. Por exemplo, para uma amostra  $X_i$  ser previsto, se a primeira árvore de decisão prever o valor 0,01, a segunda árvore prever o valor 0,02 e a terceira árvore de decisão prever o valor 0,025, então a decisão final da RF com essas três árvores seria 0,0183, a média dessas três previsões. A aleatoriedade na RF geralmente se refere à reamostragem das amostras de treinamento para cada árvore (ZHAO e CHENG, 2022). Uma grande vantagem da técnica de RF é que ela pode ser utilizada tanto para tarefas de classificação quanto para regressão. A Figura 10, acima, apresenta uma RF com duas árvores.

## B.2 Gradient Boosted Regression Trees (GBRT)

Árvores de Regressão Impulsionadas por Gradiente, *Gradient Boosted Regression Trees* (GBRT), é um algoritmo de *ensemble learning* em ocorrem a combinação de diversas árvores de decisão. Resumidamente, o GBRT funciona organizando as DT's de maneira sequencial, fazendo com que as próximas árvores corrijam os erros residuais das anteriores. Esse método é diferente da *Random Forest*, pois as árvores não são paralelas, mas uma após a outra. Um pseudocódigo simples, elaborado por Zhao e Cheng (2022), é apresentado na Figura 11. Para uma introdução mais detalhada, ver Gu, Kelly e Xiu (2020).

---

### GBRT example

---

**Tree 1:**

```
tree1 = DecisionTreeRegressor(max_depth = 3)
tree1.fit(X, y)
resid1 = y - tree1.predict(X)
```

**Tree 2:**

```
tree2 = DecisionTreeRegressor(max_depth = 2)
tree2.fit(X, resid1)
resid2 = resid1 - tree2.predict(X)
```

**Tree 3:**

```
tree3 = DecisionTreeRegressor(max_depth = 2)
tree3.fit(X, resid2)
```

**Combine:**

```
y_pred = tree1.predict(X_new) + tree2.predict(X_new) + tree3.predict(X_new)
```

---

**Figura 10 – Exemplo de Gradient Boosted Regression Trees (GBRT)**

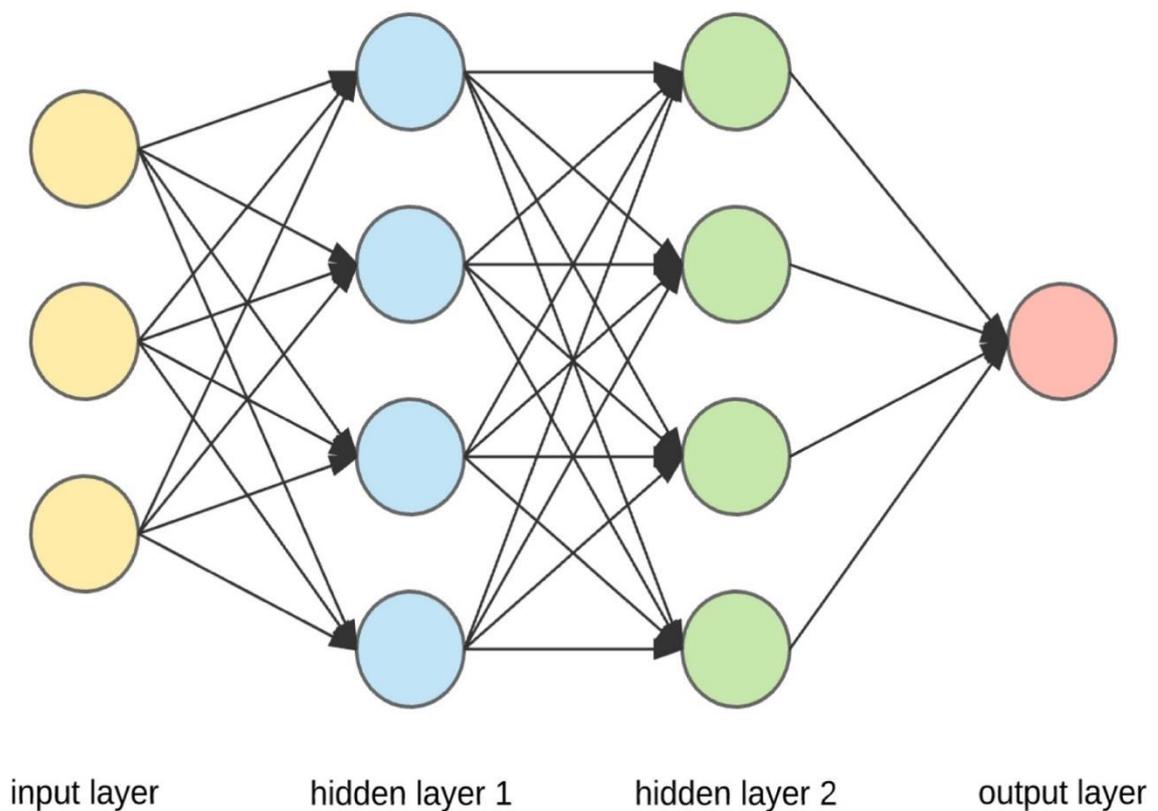
Fonte: Zhao e Cheng (2020)

## B.3 Neural Networks (NN)

As rede neurais, *Neural Networks* (NN), como o nome sugere, tem suas origens nas tentativas de encontrar representações matemáticas do processamento de informações em

sistemas biológicos (SANTOS, 2020). As NN's são modelos inspirados na estrutura do cérebro humano, e consistem em várias camadas de neurônios interconectados, cada neurônio recebendo entradas, realizando operações e passando os resultados destas operações para os neurônios na camada seguinte.

A ideia central por trás de um algoritmo de NN é que a rede neural aprende uma função que mapeia as entradas para as saídas desejadas. Durante o treinamento, os pesos das conexões entre os neurônios são ajustados iterativamente com base nos dados de treinamento, de modo a minimizar uma função de perda que quantifica o quão distantes as saídas previstas da rede estão dos valores reais. Uma NN pode consistir de múltiplas camadas de neurônios, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada neurônio em uma camada está conectado a todos os neurônios na camada subsequente. A Figura 12 apresenta um exemplo de NN denso com uma camada de entrada, duas camadas ocultas e uma camada de saída.



**Figura 11 – Exemplo de *Neural Networks* (NN)**

Fonte: Zhao e Cheng (2020)

## APÊNDICE C – CÓDIGO-FONTE PARA AQUISIÇÃO DA BASE DE DADOS DOS RETORNOS DOS ATIVOS

```
# Importar as Bibliotecas

import pandas as pd
import numpy as np
import yfinance as yf

# Importar o DataFrame com a carteira de Ativos
df = pd.read_excel('data-base-b3.xlsx')

# Criar uma lista única de ativos e informar as
tickers = df['Tickers'].values.tolist()
start_date = '2001-01-01' # Data de Início
end_date = '2024-03-31' # Data de Término

# Importar a função calculate_cumulative_returns
from def_return import calculate_cumulative_returns

# Aplicar a função para recuperar os retornos acumulados mensais dos
ativos
cumulative_returns = calculate_cumulative_returns(tickers,
start=start_date, end=end_date)

# Salvar o resultado em um DataFrame

# Converte os resultados em um DataFrame
df_return = pd.DataFrame(cumulative_returns)

# Iterar sobre os nomes das colunas e substituir ".SA" por uma string
vazia
df_return.columns = df_return.columns.str.replace('.SA', '')

df_return.to_excel('var_returns-monthly.xlsx', index=False)
df_return
```

## APÊNDICE D – CÓDIGO-FONTE PARA CALCULAR O QUINTIL DA VARIÁVEL TAMANHO (ME)

```

# Digite o nome do arquivo que contém o portfólio que deseja gerar

# Escolher apenas um e comentar os outros

name_portfolio = 'var_me-tamanho'
risk_factor = 'ME'

"""
name_portfolio = 'var_me-tamanho'
risk_factor = 'ME'

name_portfolio = 'var_book-to-market'
risk_factor = 'BM'

name_portfolio = 'var_returns-monthly'
risk_factor = 'Return'

"""

# Importar as Bibliotecas
import pandas as pd

# Importar o DataFrame
df = pd.read_excel(f'{name_portfolio}.xlsx')

# Aplicar a função melt para unificar os ativo em única coluna
df_melted = df.melt(id_vars=['Data'], var_name='Ativo',
value_name=risk_factor)

# Ordenar o DataFrame resultante por Data e na sequência por Valor, com
valores em branco primeiro
df_melted.sort_values(by=['Data', risk_factor], na_position='first',
inplace=True)

# Resetar os índices
df_melted.reset_index(drop=True, inplace=True)

# Função para classificar por quintis reiniciando para cada data
def quintile_rank_reset(x):
    return pd.qcut(x, 5, labels=False, duplicates='drop')

# Aplicando a função de classificação por quintis reiniciando para cada
data

```

```

df_melted[f'{risk_factor}_Qnt1'] =
df_melted.groupby('Data')[risk_factor].transform(quintile_rank_reset)

# Resetar os índices
df_melted.reset_index(drop=True, inplace=True)

# Converter a coluna 'Data' para DateTime e adicionar colunas de mês e
ano
df_melted['Data'] = pd.to_datetime(df_melted['Data'])
df_melted['Ano'] = df_melted['Data'].dt.year
df_melted['Mes'] = df_melted['Data'].dt.month

# Salvando o DataFrame em um arquivo Excel com o nome da variável
df_melted.to_excel(f'{name_portfolio}_Qnt1.xlsx', index=False)
print(df_melted)

# Calculando a média de cada um dos quintis para cada data
portfolio = df_melted.groupby(['Data',
f'{risk_factor}_Qnt1'])[risk_factor].mean().unstack()

# Renomeando as colunas para "Carteira 1", "Carteira 2", ..., "Carteira
5"
portfolio.columns = [f"Carteira {i+1}" for i in range(5)]

# Resetar o índice para que 'Data' se torne uma coluna novamente
portfolio.reset_index(inplace=True)

# Salvando o DataFrame em um arquivo Excel com o nome da variável
portfolio.to_excel(f'{name_portfolio}_portfolio.xlsx', index=False)
# Exibindo o DataFrame resultante
print(portfolio)

```

## APÊNDICE E – CÓDIGO-FONTE PARA CALCULAR O VALOR DA CARTEIRA TAMANHO (ME)

```

# Importar as Bibliotecas
import pandas as pd

# Importar os DataFrames
df1 = pd.read_excel('var_me-tamanho_Qntl.xlsx')
df2 = pd.read_excel('var_book-to-market_Qntl.xlsx')
df3 = pd.read_excel('var_returns-monthly_Qntl.xlsx')

# Excluir as colunas Data e Return_Qntl dos DataFrames 2 e 3
df2 = df2.drop(columns=['Data'])
df2 = df2.drop(columns=['BM_Qntl'])
df3 = df3.drop(columns=['Data'])
df3 = df3.drop(columns=['Return_Qntl'])

# Mesclar os DataFrames usando as colunas 'Mes', 'Ano' e 'Ativo' como
chaves de junção
df = pd.merge(df1, df2, on=['Mes', 'Ano', 'Ativo'])
df = pd.merge(df, df3, on=['Mes', 'Ano', 'Ativo'])

# Trocar o nome da coluna ME_Qntl para Carteiras
df.rename(columns={'ME_Qntl': 'Carteiras'}, inplace=True)

# Salvar a fusão do DataFrame em um arquivo Excel
df.to_excel('data-base_portfolio-single.xlsx', index=False)
# Exibir o DataFrame resultante
print(df)

# Remover linhas com valores em branco
df = df.dropna()

# Excluir as linhas com Return igual a zero
indices_zero_return = df[df['Return'] == 0].index
df.drop(indices_zero_return, inplace=True) # Usar o método drop() para
remover as linhas identificadas

# Excluir os ativos com Patrimônio Líquido Negativo
df = df.loc[df['BM'] >= 0]

# Exibir o DataFrame resultante
print(df)

# Calcular a média do fator de risco Tamanho (ME) para cada carteira

```

```

# Calculando a média de cada um dos quintis para cada data
portfolio1 = df.groupby(['Data', 'Carteiras'])['ME'].mean().unstack()

# Renomeando as colunas para "Carteira 1", "Carteira 2", ..., "Carteira
5"
portfolio1.columns = [f"Carteira {i+1}" for i in range(5)]
# Resetando o índice para que 'Data' se torne uma coluna novamente
portfolio1.reset_index(inplace=True)

# Salvar o resultado em um DataFrame
pd.DataFrame(portfolio1)

# Salvar o DataFrame em um arquivo Excel com o nome da variável
portfolio1.to_excel('portfolio-single-tamanho.xlsx', index=False)
# Exibindo o DataFrame resultante
print(portfolio1)

# Calcular a média do fator de risco Book-to_Market (BM) para cada
carteira

# Calculando a média de cada um dos quintis para cada data
portfolio2 = df.groupby(['Data', 'Carteiras'])['BM'].mean().unstack()
# Renomeando as colunas para "Carteira 1", "Carteira 2", ..., "Carteira
5"

portfolio2.columns = [f"Carteira {i+1}" for i in range(5)]
# Resetando o índice para que 'Data' se torne uma coluna novamente
portfolio2.reset_index(inplace=True)

# Salvar o resultado em um DataFrame
pd.DataFrame(portfolio2)

# Salvar o DataFrame em um arquivo Excel com o nome da variável
portfolio2.to_excel('portfolio-single-book-to-market.xlsx', index=False)
# Exibindo o DataFrame resultante
print(portfolio2)

# Calcular a média do fator do retorno para cada carteira

# Calculando a média de cada um dos quintis para cada data
portfolio3 = df.groupby(['Data', 'Carteiras'])['Return'].mean().unstack()
# Renomeando as colunas para "Carteira 1", "Carteira 2", ..., "Carteira
5"

portfolio3.columns = [f"Carteira {i+1}" for i in range(5)]
# Resetando o índice para que 'Data' se torne uma coluna novamente
portfolio3.reset_index(inplace=True)

# Salvar o resultado em um DataFrame

```

```
pd.DataFrame(portfolio3)

# Salvar o DataFrame em um arquivo Excel com o nome da variável
portfolio3.to_excel('portfolio-single-return.xlsx', index=False)

# Exibindo o DataFrame resultante
print(portfolio3)
```

## APÊNDICE F – CÓDIGO-FONTE PARA CRIAR OS DATAFRAMES EM QUE SERÃO REALIZADAS AS PREVISÕES DAS CARTEIRAS

```

# Importar as Bibliotecas
import pandas as pd
import os

# Importar os DataFrames
df_data_base = pd.read_excel('data_base_variables.xlsx')
df_tamanho = pd.read_excel('portfolio-single-tamanho.xlsx')
df_book_to_market = pd.read_excel('portfolio-single-book-to-market.xlsx')
df_return = pd.read_excel('portfolio-single-return.xlsx')

# Converter a coluna 'Data' para DateTime e adicionar colunas de mês e
ano
df_tamanho['Data'] = pd.to_datetime(df_tamanho['Data'])
df_tamanho['Ano'] = df_tamanho['Data'].dt.year
df_tamanho['Mes'] = df_tamanho['Data'].dt.month

# Converter a coluna 'Data' para DateTime e adicionar colunas de mês e
ano
df_book_to_market['Data'] = pd.to_datetime(df_book_to_market['Data'])
df_book_to_market['Ano'] = df_book_to_market['Data'].dt.year
df_book_to_market['Mes'] = df_book_to_market['Data'].dt.month

# Converter a coluna 'Data' para DateTime e adicionar colunas de mês e
ano
df_return['Data'] = pd.to_datetime(df_return['Data'])
df_return['Ano'] = df_return['Data'].dt.year
df_return['Mes'] = df_return['Data'].dt.month

# Carteira 1

# Tamanho
df1 = pd.merge(df_data_base, df_tamanho[['Mes', 'Ano', 'Carteira 1']],
on=['Mes', 'Ano'], how='inner')
df1.rename(columns={'Carteira 1': 'ME'}, inplace=True) # Trocar o nome da
coluna Carteira

# Book-to-market
df1 = pd.merge(df1, df_book_to_market[['Mes', 'Ano', 'Carteira 1']],
on=['Mes', 'Ano'], how='inner')
df1.rename(columns={'Carteira 1': 'BM'}, inplace=True) # Trocar o nome da
coluna Carteira

```

```

# Retorno
df1 = pd.merge(df1, df_return[['Mes', 'Ano', 'Carteira 1']], on=['Mes',
'Ano'], how='inner')
df1.rename(columns={'Carteira 1': 'Return'}, inplace=True) # Trocar o
nome da coluna Carteira
df1

# Carteira 2

# Tamanho
df2 = pd.merge(df_data_base, df_tamanho[['Mes', 'Ano', 'Carteira 2']],
on=['Mes', 'Ano'], how='inner')
df2.rename(columns={'Carteira 2': 'ME'}, inplace=True) # Trocar o nome da
coluna Carteira

# Book-to-market
df2 = pd.merge(df2, df_book_to_market[['Mes', 'Ano', 'Carteira 2']],
on=['Mes', 'Ano'], how='inner')
df2.rename(columns={'Carteira 2': 'BM'}, inplace=True) # Trocar o nome da
coluna Carteira

# Retorno
df2 = pd.merge(df2, df_return[['Mes', 'Ano', 'Carteira 2']], on=['Mes',
'Ano'], how='inner')
df2.rename(columns={'Carteira 2': 'Return'}, inplace=True) # Trocar o
nome da coluna Carteira
df2

# Carteira 3

# Tamanho
df3 = pd.merge(df_data_base, df_tamanho[['Mes', 'Ano', 'Carteira 3']],
on=['Mes', 'Ano'], how='inner')
df3.rename(columns={'Carteira 3': 'ME'}, inplace=True) # Trocar o nome da
coluna Carteira

# Book-to-market
df3 = pd.merge(df3, df_book_to_market[['Mes', 'Ano', 'Carteira 3']],
on=['Mes', 'Ano'], how='inner')
df3.rename(columns={'Carteira 3': 'BM'}, inplace=True) # Trocar o nome da
coluna Carteira

# Retorno
df3 = pd.merge(df3, df_return[['Mes', 'Ano', 'Carteira 3']], on=['Mes',
'Ano'], how='inner')
df3.rename(columns={'Carteira 3': 'Return'}, inplace=True) # Trocar o
nome da coluna Carteira
df3

```

```

# Carteira 4

# Tamanho
df4 = pd.merge(df_data_base, df_tamanho[['Mes', 'Ano', 'Carteira 4']],
on=['Mes', 'Ano'], how='inner')
df4.rename(columns={'Carteira 4': 'ME'}, inplace=True) # Trocar o nome da
coluna Carteira

# Book-to-market
df4 = pd.merge(df4, df_book_to_market[['Mes', 'Ano', 'Carteira 4']],
on=['Mes', 'Ano'], how='inner')
df4.rename(columns={'Carteira 4': 'BM'}, inplace=True) # Trocar o nome da
coluna Carteira

# Retorno
df4 = pd.merge(df4, df_return[['Mes', 'Ano', 'Carteira 4']], on=['Mes',
'Ano'], how='inner')
df4.rename(columns={'Carteira 4': 'Return'}, inplace=True) # Trocar o
nome da coluna Carteira
df4

# Carteira 5

# Tamanho
df5 = pd.merge(df_data_base, df_tamanho[['Mes', 'Ano', 'Carteira 5']],
on=['Mes', 'Ano'], how='inner')
df5.rename(columns={'Carteira 5': 'ME'}, inplace=True) # Trocar o nome da
coluna Carteira
# Book-to-market
df5 = pd.merge(df5, df_book_to_market[['Mes', 'Ano', 'Carteira 5']],
on=['Mes', 'Ano'], how='inner')
df5.rename(columns={'Carteira 5': 'BM'}, inplace=True) # Trocar o nome da
coluna Carteira
# Retorno
df5 = pd.merge(df5, df_return[['Mes', 'Ano', 'Carteira 5']], on=['Mes',
'Ano'], how='inner')
df5.rename(columns={'Carteira 5': 'Return'}, inplace=True) # Trocar o
nome da coluna Carteira
df5

.....
"""# Criar uma pasta no diretório atual"""

def criar_pasta():
    # Nome da nova pasta
    nova_pasta = 'Portfolios'
    # Caminho para o diretório de trabalho atual
    diretorio_atual = os.getcwd()
    # Caminho para a nova pasta

```

```
    caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)
    # Verifica se a pasta já existe
    if os.path.exists(caminho_nova_pasta):
        print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
        return
    # Cria a pasta
    os.makedirs(caminho_nova_pasta)
    print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")
# Chamada da função
criar_pasta()

.....

"""# Salvar os arquivos na pasta correspondente"""

# Salvar o DataFrame em um arquivo Excel com o nome da variável
df1.to_excel('Carteira_1.xlsx', index=False)
df2.to_excel('Carteira_2.xlsx', index=False)
df3.to_excel('Carteira_3.xlsx', index=False)
df4.to_excel('Carteira_4.xlsx', index=False)
df5.to_excel('Carteira_5.xlsx', index=False)
```

## APÊNDICE G – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O COEFICIENTE BETA DA CARTEIRA

```

import pandas as pd
import numpy as np

"""
Equação para calcular o coeficiente beta de um ativo em relação ao
mercado:
 $\beta_i = \text{Cov}(R_i, R_m) / \text{Var}(R_m)$ 

Onde:
Ri = retorno do ativo
Rm = retorno de mercado
Cov(Ri, Rm) = covariância entre o retorno do ativo e o retorno de
mercado
Var(Rm) = variância do retorno de mercado

Função para calcular o coeficiente beta de um ativo em relação ao
mercado ( $\beta_i$ )
:param Ri: Vetor de retornos do ativo.
:param Rm: Vetor de retornos do mercado.
:return: O coeficiente beta do ativo.

"""

def calculate_beta(asset_returns, market_returns):

    covariance = np.cov(asset_returns, market_returns)[0, 1] #
Cov(Ri, Rm)
    market_variance = np.var(market_returns) # Var(Rm)
    beta = covariance / market_variance #  $\beta$ 
    return beta

```

## APÊNDICE H – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO CAPM

```

# Importar as Bibliotecas
import pandas as pd
import numpy as np
import os

# Informar o nome da carteira que deseja analisar
# Lembrar de trocar o número da carteira antes de realizar a previsão
portfolio = 1

# Importar e Estruturar o DataFrame
data = pd.read_excel(f'Carteira_{portfolio}.xlsx')
df = pd.DataFrame(data)

# Criar as variáveis para equação do capm:  $R_i = R_f + \beta_i (R_m - R_f)$ 
Ri = df['Return'] # Vetor de retorno do ativo.
Rm = df['IBOV'] # Vetor de retorno do mercado.
Rf = df['SELIC'] # Vetor do atico livre de risco
beta = []

# Importar a função "beta"
from def_beta import calculate_beta

# Aplicar a função beta
beta = calculate_beta(Ri, Rm)
print("Coeficiente beta do ativo:", beta)

# Importar a função "capm"
from def_capm import calculate_capm
capm = calculate_capm(Rm, Rf, beta) # Aplicar a função capm

"""# Criar uma pasta no diretório atual"""

def criar_pasta():
    # Nome da nova pasta
    nova_pasta = 'forecast_finance'
    # Caminho para o diretório de trabalho atual
    diretorio_atual = os.getcwd()
    # Caminho para a nova pasta
    caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)
    # Verifica se a pasta já existe
    if os.path.exists(caminho_nova_pasta):
        print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
    return

```

```
# Cria a pasta
os.makedirs(caminho_nova_pasta)
print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")

# Chamada da função
criar_pasta()

"""# Salvar o arquivo na pasta correspondente"""

# Salvar o resultado da função CAPM em uma coluna
df[f'CAPM_{portfolio}'] = capm

# Salvar o arquivo no formato excel
nome_arquivo = f'capm_Carteira_{portfolio}.xlsx'

# Nome da pasta onde o arquivo será salvo
caminho_pasta = 'forecast_finance'

# Caminho completo para o arquivo
caminho_arquivo = os.path.join(os.getcwd(), caminho_pasta, nome_arquivo)

# Salvar no formato csv
df.to_excel(caminho_arquivo, index=False)
df
```

## APÊNDICE I – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O FILTRO DE KALMAN

```
import numpy as np
from pykalman import KalmanFilter

def kalman_filter(asset_returns, market_returns):

    # Criar a matriz de observação com o retorno do ativo e uma sequência
    # de valores de unidade
    matriz = np.expand_dims(
        np.vstack([asset_returns, np.ones(len(asset_returns))]).T, axis=1
    )

    # Inicializar o filtro de Kalman
    kf = KalmanFilter(
        initial_state_mean=[0, 0], # Estado inicial do intercepto e
        # coeficiente zero
        initial_state_covariance=np.ones(
            (2, 2)
        ), # Estado inicial da covariância (matriz de identidade 2x2)
        transition_matrices=np.eye(
            2
        ), # Matriz de transição do estado de t para t + 1 (matriz
        # identidade 2x2)
        observation_matrices=matriz, # Matriz das observação
    )

    # Aplicar o filtro de Kalman
    state_means, state_covs = kf.filter(
        market_returns.values
    ) # state_means (coeficientes) e state_covs (covariâncias)
    return state_means, state_covs
```

## APÊNDICE J – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO L-CAPM

```

# Importar as Bibliotecas
import pandas as pd
import numpy as np
import os
from pykalman import KalmanFilter

# Informar o nome da carteira que deseja analisar
# Lembrar de trocar o número da carteira antes de realizar a previsão
portfolio = 5

# Importar e Estruturar o DataFrame
data = pd.read_excel(f'Carteira_{portfolio}.xlsx')
df = pd.DataFrame(data)

# Criar as variáveis para equação do capm:  $R_i = R_f + \beta_i (R_m - R_f)$ 

Ri = df['Return'] # Vetor de retorno do ativo.
Rm = df['IBOV'] # Vetor de retorno do mercado.
Rf = df['SELIC'] # Vetor do atico livre de risco
beta = []

# Importar a função "kalman_filter"
from def_kalman_filter import kalman_filter

# Aplicar a função kalman_filter
state_means, state_covs = kalman_filter(Ri, Rm)

# Criar a coluna "Beta" e atribui-la a variável beta
df['Beta'] = state_means[:, 1]
beta = df['Beta']
df.drop('Beta', axis=1, inplace=True)

# Importar a função "capm"
from def_capm import calculate_capm
capm = calculate_capm(Rm, Rf, beta) # Aplicar a função capm

"""# Criar uma pasta no diretório atual"""

def criar_pasta():

    # Nome da nova pasta
    nova_pasta = 'forecast_finance'

    # Caminho para o diretório de trabalho atual

```

```
diretorio_atual = os.getcwd()

# Caminho para a nova pasta
caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)

# Verifica se a pasta já existe
if os.path.exists(caminho_nova_pasta):
    print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
    return

# Cria a pasta
os.makedirs(caminho_nova_pasta)
print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")
# Chamada da função
criar_pasta()

"""# Salvar o arquivo na pasta correspondente"""

# Salvar o resultado da função CAPM em uma coluna
df[f'L-CAPM_{portfolio}'] = capm

# Salvar o arquivo no formato excel
nome_arquivo = f'capm-learning_Carteira_{portfolio}.xlsx'

# Nome da pasta onde o arquivo será salvo
caminho_pasta = 'forecast_finance'
# Caminho completo para o arquivo

caminho_arquivo = os.path.join(os.getcwd(), caminho_pasta, nome_arquivo)
# Salvar no formato csv
df.to_excel(caminho_arquivo, index=False)
df
```

## APÊNDICE K – CÓDIGO-FONTE QUE DEFINE UMA FUNÇÃO PARA CALCULAR O BETA *DOWNSIDE*

```

# Importar as Bibliotecas
import numpy as np
import pandas as pd

"""
Equação para calcular o Downside Beta:

$$\beta^D = (E(\text{Min}((r_i - r_i^-), 0)) \times E(\text{Min}((r_M - r_M^-), 0))) / E(\text{Min}((r_i - r_i^-), 0)^2)$$


Onde:
 $\beta^D$  é o downside beta
 $r_i$  é o retorno do ativo i
 $r_i^-$  é a média dos retornos do ativo i
 $r_M$  é o retorno do mercado
 $r_M^-$  é a média dos retornos do mercado

E() é o operador de expectativa ou valor esperado, representando a média
Min(a, b) retorna o menor valor entre a e b
"""

def downside_beta(asset_returns, market_returns):
    asset_mean = np.mean(asset_returns) # Média dos retornos do ativo
    market_mean = np.mean(market_returns) # Média dos retornos do
mercado

    downside_asset = np.minimum(asset_returns - asset_mean, 0) #
Downside do ativo
    downside_market = np.minimum(market_returns - market_mean, 0) #
Downside do mercado

    E_downside_asset = np.mean(downside_asset) # Termos esperados de
downside do ativo
    E_downside_market = np.mean(downside_market) # Termos esperados de
downside do mercado

    downside_asset_squared = np.square(downside_asset) # Termo de
downside ao quadrado
    E_downside_asset_squared = np.mean(downside_asset_squared) # Termo
esperado de downside ao quadrado
    downside_beta = (E_downside_asset * E_downside_market) /
E_downside_asset_squared # Downside beta

    return downside_beta

```

## APÊNDICE L – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO D-CAPM

```

# Importar as Bibliotecas
import pandas as pd
import numpy as np
import os

# Informar o nome da carteira que deseja analisar

# Lembrar de trocar o número da carteira antes de realizar a previsão
portfolio = 1

# Importar e Estruturar o DataFrame
data = pd.read_excel(f'Carteira_{portfolio}.xlsx')
df = pd.DataFrame(data)

# Criar as variáveis para equação do capm:  $R_i = R_f + \beta_i (R_m - R_f)$ 

Ri = df['Return'] # Vetor de retorno do ativo.
Rm = df['IBOV'] # Vetor de retorno do mercado.
Rf = df['SELIC'] # Vetor do atico livre de risco
beta = []

# Importar a função downside_beta
from def_downside_beta import downside_beta

# Aplicar a função downside beta
beta = downside_beta(Ri, Rm)
print("Downside Beta:", beta)

# Importar a função "capm"
from def_capm import calculate_capm
capm = calculate_capm(Rm, Rf, beta) # Aplicar a função capm

"""# Criar uma pasta no diretório atual"""

def criar_pasta():

    # Nome da nova pasta
    nova_pasta = 'forecast_finance'

    # Caminho para o diretório de trabalho atual
    diretorio_atual = os.getcwd()

```

```
# Caminho para a nova pasta
caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)

# Verifica se a pasta já existe
if os.path.exists(caminho_nova_pasta):
    print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
    return

# Cria a pasta
os.makedirs(caminho_nova_pasta)

print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")

# Chamada da função
criar_pasta()

"""# Salvar o arquivo na pasta correspondente"""

# Salvar o resultado da função CAPM em uma coluna
df[f'D-CAPM_{portfolio}'] = capm

# Salvar o arquivo no formato excel
nome_arquivo = f'capm-downside_Carteira_{portfolio}.xlsx'

# Nome da pasta onde o arquivo será salvo
caminho_pasta = 'forecast_finance'

# Caminho completo para o arquivo
caminho_arquivo = os.path.join(os.getcwd(), caminho_pasta, nome_arquivo)

# Salvar no formato csv
df.to_excel(caminho_arquivo, index=False)
df
```

## APÊNDICE M – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO C-CAPM

```

# Importar as Bibliotecas
import pandas as pd
import numpy as np
import statsmodels.api as sm
import os
import warnings
warnings.filterwarnings('ignore')

# Informar o nome da carteira que deseja analisar
# Lembrar de trocar o número da carteira antes de realizar a previsão
portfolio = 1

# Importar e Estruturar o DataFrame
data = pd.read_excel(f'Carteira_{portfolio}.xlsx')
df = pd.DataFrame(data)
df

# Criar as variáveis para equação do capm:  $R_i = R_f + \beta_i (R_m - R_f)$ 
Ri = df['Return'] # Vetor de retorno do ativo.
Rm = df['IBOV'] # Vetor de retorno do mercado.
Rf = df['SELIC'] # Vetor do atico livre de risco
Rm_Rf = df['Rm-Rf']
hml = df['HML']
spread = df['Spread']
beta = []

# Importar a função "beta"
from def_beta import calculate_beta

# Aplicar a função beta
beta = calculate_beta(Ri, Rm)
print("Coeficiente beta do ativo:", beta)

# Selecionar as variáveis independentes
X = df[['Rm-Rf', 'HML', 'Spread']]

# Adicionar interações entre 'Rm' e as outras variáveis
X['Rm*Rm-Rf'] = Rm * Rm-Rf
X['Rm*HML'] = Rm * hml
X['Rm*Spread'] = Rm * spread

# Adicionar uma constante ao modelo (o termo  $\alpha^i$ )
X = sm.add_constant(X)

```

```

# Selecionar a variável dependente
y = Ri

# Criar o modelo de regressão
modelo = sm.OLS(y, X)
resultado = modelo.fit() # Ajustar o modelo aos dados
c_capm = resultado.predict(X) # Fazer previsões
c_capm

""""# Criar uma pasta no diretório atual""""

def criar_pasta():

    # Nome da nova pasta
    nova_pasta = 'forecast_finance'
    # Caminho para o diretório de trabalho atual
    diretorio_atual = os.getcwd()

    # Caminho para a nova pasta
    caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)

    # Verifica se a pasta já existe
    if os.path.exists(caminho_nova_pasta):
        print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
        return

    # Cria a pasta
    os.makedirs(caminho_nova_pasta)

    print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")

# Chamada da função
criar_pasta()

""""# Salvar o arquivo na pasta correspondente""""

# Salvar o resultado da função CAPM em uma coluna
df[f'C-CAPM_{portfolio}'] = c_capm

# Salvar o arquivo no formato excel
nome_arquivo = f'capm-conditional_Carteira_{portfolio}.xlsx'

# Nome da pasta onde o arquivo será salvo
caminho_pasta = 'forecast_finance'

# Caminho completo para o arquivo
caminho_arquivo = os.path.join(os.getcwd(), caminho_pasta, nome_arquivo)

```

```
# Salvar no formato csv
df.to_excel(caminho_arquivo, index=False)
df
```

## APÊNDICE N – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DO MODELO F&F (1996)

```

# Importar as Bibliotecas
import pandas as pd
import numpy as np
import statsmodels.api as sm
import os
import warnings
warnings.filterwarnings('ignore')

# Informar o nome da carteira que deseja analisar
# Lembrar de trocar o número da carteira antes de realizar a previsão
portfolio = 1

# Importar e Estruturar o DataFrame
data = pd.read_excel(f'Carteira_{portfolio}.xlsx')
df = pd.DataFrame(data)

# Aplicar o logaritmo natural às variáveis tamanho (ME) e book-to-market (BM)
df['ln_ME'] = np.log(df['ME'])
df['ln_BM'] = np.log(df['BM'])
df

# Criar as variáveis
Ri = df['Return'] # Vetor de retorno do ativo.
Rm = df['IBOV'] # Vetor de retorno do mercado.
Rf = df['SELIC'] # Vetor do ativo livre de risco
Rm_Rf = df['Rm-Rf']
hml = df['HML']
spread = df['Spread']
me = df['ln_ME']
bm = df['ln_BM']
beta = []
coeficiente_rm = []
coeficiente_me = []
coeficiente_bm = []

# Calcular o Beta Anual do ativo em relação ao mercado

# Importar a função "beta"
from def_beta import calculate_beta

# Lista para armazenar todos os betas calculados
all_betas = []

```

```

# Iterar sobre as observações em grupos de 12
for i in range(0, len(df), 12):
    # Lista para armazenar os betas calculados para o grupo atual
    betas_group = []

    # Extrair os retornos mensais do ativo e do mercado para o grupo
    atual
    asset_returns_group = df.loc[i:i+11, 'Return']
    market_returns_group = df.loc[i:i+11, 'IBOV']

    # Calcular o beta usando a função calculate_beta
    beta = calculate_beta(asset_returns_group, market_returns_group)

    # Adicionar o beta calculado à lista de betas do grupo
    betas_group.extend([beta]*12)

    # Adicionar os betas do grupo atual à lista de todos os betas
    calculados
    all_betas.extend(betas_group)

# Adicionar os betas calculados como uma nova coluna ao DataFrame
beta = all_betas
df['Beta'] = beta
df

# Calcular a média dos coeficientes das variáveis Rm, ME e BM

# Número de meses em um ano
meses_por_ano = 12

# Lista para armazenar os coeficientes calculados
coeficientes_rm = []
coeficientes_me = []
coeficientes_bm = []

# Iterar sobre os anos no DataFrame
for i in range(len(df) // meses_por_ano):
    # Calcular o índice inicial e final para o ano atual
    inicio_ano = i * meses_por_ano
    fim_ano = (i + 1) * meses_por_ano

    mercado_rm = df.iloc[inicio_ano:fim_ano]['IBOV'] # Extrair os
    retornos de mercado para o ano atual
    coeficiente_rm = mercado_rm.mean() # Calcular a média do retorno de
    mercado para o ano atual
    coeficientes_rm.extend([coeficiente_rm] * meses_por_ano) # Adicionar
    o coeficiente do ano atual à lista

```

```

# Garantir que os coeficientes tenham o mesmo comprimento que o DataFrame
original
coeficiente_rm = coeficientes_rm[:len(df)]

# Adicionar os coeficientes calculados como uma nova coluna ao DataFrame
df['λ_Rm'] = coeficiente_rm

# Iterar sobre os anos no DataFrame
for i in range(len(df) // meses_por_ano):
    # Calcular o índice inicial e final para o ano atual
    inicio_ano = i * meses_por_ano
    fim_ano = (i + 1) * meses_por_ano

    mercado_me = df.iloc[inicio_ano:fim_ano]['ME'] # Extrair os retornos
de mercado para o ano atual
    coeficiente_me = mercado_me.mean() # Calcular a média do retorno de
mercado para o ano atual
    coeficientes_me.extend([coeficiente_me] * meses_por_ano) # Adicionar
o coeficiente do ano atual à lista

# Garantir que os coeficientes tenham o mesmo comprimento que o DataFrame
original
coeficiente_me = coeficientes_me[:len(df)]

# Adicionar os coeficientes calculados como uma nova coluna ao DataFrame
df['λ_ME'] = coeficiente_me

# Iterar sobre os anos no DataFrame
for i in range(len(df) // meses_por_ano):
    # Calcular o índice inicial e final para o ano atual
    inicio_ano = i * meses_por_ano
    fim_ano = (i + 1) * meses_por_ano

    mercado_bm = df.iloc[inicio_ano:fim_ano]['BM'] # Extrair os retornos
de mercado para o ano atual
    coeficiente_bm = mercado_bm.mean() # Calcular a média do retorno de
mercado para o ano atual
    coeficientes_bm.extend([coeficiente_bm] * meses_por_ano) # Adicionar
o coeficiente do ano atual à lista

# Garantir que os coeficientes tenham o mesmo comprimento que o DataFrame
original
coeficiente_bm = coeficientes_bm[:len(df)]

# Adicionar os coeficientes calculados como uma nova coluna ao DataFrame
df['λ_BM'] = coeficiente_bm
df

```

```

# Selecionar as variáveis independentes
X = df[['SELIC']]

# Adicionar outras variáveis independentes
X['λ_Rm*Beta'] = df['λ_Rm'] * df['Beta']
X['λ_ME*ln_ME'] = df['λ_ME'] * df['ln_ME']
X['λ_BM*ln_BM'] = df['λ_BM'] * df['ln_BM']

# Adicionar uma constante ao modelo (o termo α^i)
X = sm.add_constant(X)

# Selecionar a variável dependente
y = Ri

# Criar o modelo de regressão
modelo = sm.OLS(y, X)
resultado = modelo.fit() # Ajustar o modelo aos dados
ff_model = resultado.predict(X) # Fazer previsões
ff_model

df = df.drop(columns=['ln_ME', 'ln_BM', 'Beta', 'λ_Rm', 'λ_ME', 'λ_BM'])
df

"""# Criar uma pasta no diretório atual"""

def criar_pasta():

    # Nome da nova pasta
    nova_pasta = 'forecast_finance'

    # Caminho para o diretório de trabalho atual
    diretorio_atual = os.getcwd()

    # Caminho para a nova pasta
    caminho_nova_pasta = os.path.join(diretorio_atual, nova_pasta)

    # Verifica se a pasta já existe
    if os.path.exists(caminho_nova_pasta):
        print(f"A pasta '{nova_pasta}' já existe em {diretorio_atual}!")
        return

    # Cria a pasta
    os.makedirs(caminho_nova_pasta)
    print(f"Pasta '{nova_pasta}' criada com sucesso em
{diretorio_atual}!")

# Chamada da função
criar_pasta()

```

```
"""# Salvar o arquivo na pasta correspondente"""

# Salvar o resultado da função CAPM em uma coluna
df[f'F&F_{portfolio}'] = ff_model

# Salvar o arquivo no formato excel
nome_arquivo = f'f&f-model_Carteira_{portfolio}.xlsx'

# Nome da pasta onde o arquivo será salvo
caminho_pasta = 'forecast_finance'

# Caminho completo para o arquivo
caminho_arquivo = os.path.join(os.getcwd(), caminho_pasta, nome_arquivo)

# Salvar no formato csv
df.to_excel(caminho_arquivo, index=False)
df
```

## APÊNDICE O – CÓDIGO-FONTE PARA CRIAR A MATRIZ DE PREVISÕES E CALCULAR O R-QUADRADO FORA DA AMOSTRA DAS CARTEIRAS

```

# Importar as Bibliotecas
import pandas as pd
import warnings
warnings.filterwarnings('ignore')
import scipy.stats
from scipy.stats import ttest_rel

# Métricas de Avaliação
from sklearn.metrics import accuracy_score, confusion_matrix,
precision_score, recall_score, f1_score
from sklearn.metrics import log_loss, r2_score, mean_squared_error,
mean_absolute_error

""" # Carteiras """

# Lembrar de trocar o número da carteira
carteira = 1

# Importar os DataFrames das Carteiras 1
df_capm = pd.read_excel(f'capm_Carteira_{carteira}.xlsx')
df_d_capm = pd.read_excel(f'capm-downside_Carteira_{carteira}.xlsx')
df_l_capm = pd.read_excel(f'capm-learning_Carteira_{carteira}.xlsx')
df_c_capm = pd.read_excel(f'capm-conditional_Carteira_{carteira}.xlsx')
df_ff_model = pd.read_excel(f'f&f-model_Carteira_{carteira}.xlsx')

# Mesclar as previsões das Carteiras 1
df = pd.merge(df_capm, df_d_capm[['Mes', 'Ano', f'D-CAPM_{carteira}']],
on=['Mes', 'Ano'], how='inner')
df = pd.merge(df, df_l_capm[['Mes', 'Ano', f'L-CAPM_{carteira}']],
on=['Mes', 'Ano'], how='inner')
df = pd.merge(df, df_c_capm[['Mes', 'Ano', f'C-CAPM_{carteira}']],
on=['Mes', 'Ano'], how='inner')
df = pd.merge(df, df_ff_model[['Mes', 'Ano', f'F&F_{carteira}']],
on=['Mes', 'Ano'], how='inner')

# Criar a coluna com as médias das previsões
df['Media'] = df[[f'CAPM_{carteira}',
                  f'D-CAPM_{carteira}',
                  f'L-CAPM_{carteira}',
                  f'C-CAPM_{carteira}'],

```

```

        f'F&F_{carteira}']].mean(axis=1)
df

""" # Salvar a Matriz de Previsões da Carteira em um arquivo em Excel """

df_matrix = df[['Data',
                'Return',
                f'CAPM_{carteira}',
                f'D-CAPM_{carteira}',
                f'L-CAPM_{carteira}',
                f'C-CAPM_{carteira}',
                f'F&F_{carteira}',
                'Media']]

df_matrix.to_excel(f'matrix_Carteira_{carteira}.xlsx', index = False)
df_matrix

""" # Dividir o DataFrame nos conjunto de Treino, Validação e Teste """

linhas = len(df) # obter o tamanho do DataFrame
# ':' após a vírgula significa a seleção de todas as colunas do
dataframe.
df_treino = df.iloc[0:36, :] # 01/2003 até 12/2005
df_validacao = df.iloc[36:120, :] # 01/2006 até 12/2012
df_teste = df.iloc[120:240, :] # 01/2013 até 12/2022

"""# Calcular o R2 fora da amostra (R2OSS) baseado na fórmula de Campbell
e Thompson (2008)."""

# DataFrame de Validação

# Atribuir variáveis às previsões dos modelos financeiros
capm = df_validacao[f'CAPM_{carteira}']
d_capm = df_validacao[f'D-CAPM_{carteira}']
l_capm = df_validacao[f'L-CAPM_{carteira}']
c_capm = df_validacao[f'C-CAPM_{carteira}']
ff_model = df_validacao[f'F&F_{carteira}']
mean = df_validacao['Media']

# Criar a coluna HM (média histórica) no dataframe da Carteira 1
df_validacao['HM'] = df_validacao['Return'].expanding().mean() # HM é a
média histórica (y_mean) do retorno da carteira
y_mean = df_validacao['HM'] # "y_mean" recebe a média histórica da
carteira
y = df_validacao['Return'] # "y" recebe o retorno da carteira

```

```

# Calcular o MSE da média histórica
mse_mh = mean_squared_error(y, y_mean)

# CAPM
capm_mse = mean_squared_error(y, capm)
capm_r2 = 1 - (capm_mse / mse_mh)

# D-CAPM
d_capm_mse = mean_squared_error(y, d_capm)
d_capm_r2 = 1 - (d_capm_mse / mse_mh)

# L-CAPM
l_capm_mse = mean_squared_error(y, l_capm)
l_capm_r2 = 1 - (l_capm_mse / mse_mh)

# C-CAPM
c_capm_mse = mean_squared_error(y, c_capm)
c_capm_r2 = 1 - (c_capm_mse / mse_mh)

# F&F
ff_model_mse = mean_squared_error(y, ff_model)
ff_model_r2 = 1 - (ff_model_mse / mse_mh)

# Media - Media das Previsões
media_mse = mean_squared_error(y, mean)
media_r2 = 1 - (media_mse / mse_mh)

"""# Criar o DataFrame com Resultados do R2OOS e MSE."""

# DataFrame de Validação

# Listar todos os modelos
nomes_modelos = [f'CAPM_{carteira}', f'D-CAPM_{carteira}', f'L-
CAPM_{carteira}', f'C-CAPM_{carteira}', f'F&F_{carteira}', 'Media']

# Criar os dicionários de R2 e MSE
dict_r2 = {
    f"CAPM_{carteira}": [capm_r2],
    f"D-CAPM_{carteira}": [d_capm_r2],
    f"L-CAPM_{carteira}": [l_capm_r2],
    f"C-CAPM_{carteira}": [c_capm_r2],
    f"F&F_{carteira}": [ff_model_r2],
    "Media": [media_r2],
}

dict_mse = {
    f"CAPM_{carteira}": [capm_mse],
    f"D-CAPM_{carteira}": [d_capm_mse],

```

```

f"L-CAPM_{carteira}": [l_capm_mse],
f"C-CAPM_{carteira}": [c_capm_mse],
f"F&F_{carteira}": [ff_model_mse],
"Media": [media_mse],
}

# Criar DataFrames separados para R2 e MSE
df_r2 = pd.DataFrame({'Modelo': nomes_modelos, 'R2':
[float(dict_r2[model][0]) for model in nomes_modelos]})
df_mse = pd.DataFrame({'Modelo': nomes_modelos, 'MSE':
[float(dict_mse[model][0]) for model in nomes_modelos]})

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_mse = pd.merge(df_r2, df_mse, on='Modelo')

print(df_r2_mse)

"""# Calcular o p-value."""

# DataFrame de Validação

# Listar todos os modelos
modelos_p_value = [f'CAPM_{carteira}',
f'D-CAPM_{carteira}',
f'L-CAPM_{carteira}',
f'C-CAPM_{carteira}',
f'F&F_{carteira}',
'Media'
]

# Dicionário para armazenar os resultados
resultados = {'Modelo': [], 'P-Value': [], 'Significativo_01': [],
'Significativo_05': [], 'Significativo_10': []}

# Loop para calcular os testes e preencher os resultados
for modelo in modelos_p_value:

# Calcular o teste t-pareado entre o modelo e a média histórica
t_stat, p_value = ttest_rel(df_validacao[modelo], y_mean)

# Avaliar a significância
significativo_01 = p_value < 0.01
significativo_05 = p_value < 0.05
significativo_10 = p_value < 0.10

# Armazenar os resultados no dicionário
resultados['Modelo'].append(modelo)
resultados['P-Value'].append(p_value)
resultados['Significativo_01'].append(significativo_01)

```

```

    resultados['Significativo_05'].append(significativo_05)
    resultados['Significativo_10'].append(significativo_10)

# Criar um DataFrame a partir do dicionário
p_value = pd.DataFrame(resultados)

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_p_value = pd.merge(df_r2_mse, p_value, on='Modelo')

# Salvar o DataFrame de Validação no arquivo em excel
df_r2_p_value.to_excel(f'r2_validacao_carteira_{carteira}.xlsx',
index=False)

print(df_r2_p_value)

"""# Calcular o R2 fora da amostra (R2OSS) baseado na fórmula de Campbell
e Thompson (2008)."""

# DataFrame de Teste

# Atribuir variáveis às previsões dos modelos financeiros
capm = df_teste[f'CAPM_{carteira}']
d_capm = df_teste[f'D-CAPM_{carteira}']
l_capm = df_teste[f'L-CAPM_{carteira}']
c_capm = df_teste[f'C-CAPM_{carteira}']
ff_model = df_teste[f'F&F_{carteira}']
mean = df_teste['Media']

# Criar a coluna HM (média histórica) no dataframe da Carteira 1
df_teste['HM'] = df_teste['Return'].expanding().mean() # HM é a média
histórica (y_mean) do retorno da carteira
y_mean = df_teste['HM'] # "y_mean" recebe a média histórica da carteira
y = df_teste['Return'] # "y" recebe o retorno da carteira

# Calcular o MSE da média histórica
mse_mh = mean_squared_error(y, y_mean)

# CAPM
capm_mse = mean_squared_error(y, capm)
capm_r2 = 1 - (capm_mse / mse_mh)

# D-CAPM
d_capm_mse = mean_squared_error(y, d_capm)
d_capm_r2 = 1 - (d_capm_mse / mse_mh)

# L-CAPM
l_capm_mse = mean_squared_error(y, l_capm)
l_capm_r2 = 1 - (l_capm_mse / mse_mh)

```

```

# C-CAPM
c_capm_mse = mean_squared_error(y, c_capm)
c_capm_r2 = 1 - (c_capm_mse / mse_mh)

# F&F
ff_model_mse = mean_squared_error(y, ff_model)
ff_model_r2 = 1 - (ff_model_mse / mse_mh)

# Media - Media das Previsões
media_mse = mean_squared_error(y, mean)
media_r2 = 1 - (media_mse / mse_mh)

"""# Criar o DataFrame com Resultados do R2OOS e MSE."""

# DataFrame de Teste

# Listar todos os modelos
nomes_modelos = [f'CAPM_{carteira}', f'D-CAPM_{carteira}', f'L-
CAPM_{carteira}', f'C-CAPM_{carteira}', f'F&F_{carteira}', 'Media']

# Criar os dicionários de R2 e MSE
dict_r2 = {
    f"CAPM_{carteira}": [capm_r2],
    f"D-CAPM_{carteira}": [d_capm_r2],
    f"L-CAPM_{carteira}": [l_capm_r2],
    f"C-CAPM_{carteira}": [c_capm_r2],
    f"F&F_{carteira}": [ff_model_r2],
    "Media": [media_r2],
}

dict_mse = {
    f"CAPM_{carteira}": [capm_mse],
    f"D-CAPM_{carteira}": [d_capm_mse],
    f"L-CAPM_{carteira}": [l_capm_mse],
    f"C-CAPM_{carteira}": [c_capm_mse],
    f"F&F_{carteira}": [ff_model_mse],
    "Media": [media_mse],
}

# Criar DataFrames separados para R2 e MSE
df_r2 = pd.DataFrame({'Modelo': nomes_modelos, 'R2':
[float(dict_r2[model][0]) for model in nomes_modelos]})
df_mse = pd.DataFrame({'Modelo': nomes_modelos, 'MSE':
[float(dict_mse[model][0]) for model in nomes_modelos]})

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_mse = pd.merge(df_r2, df_mse, on='Modelo')

print(df_r2_mse)

```

```

"""# Calcular o p-value."""

# DataFrame de Teste

# Listar todos os modelos
modelos_p_value = [f'CAPM_{carteira}',
                   f'D-CAPM_{carteira}',
                   f'L-CAPM_{carteira}',
                   f'C-CAPM_{carteira}',
                   f'F&F_{carteira}',
                   'Media'
                   ]

# Dicionário para armazenar os resultados
resultados = {'Modelo': [], 'P-Value': [], 'Significativo_01': [],
              'Significativo_05': [], 'Significativo_10': []}

# Loop para calcular os testes e preencher os resultados
for modelo in modelos_p_value:

    # Calcular o teste t-pareado entre o modelo e a média histórica
    t_stat, p_value = ttest_rel(df_teste[modelo], y_mean)

    # Avaliar a significância
    significativo_01 = p_value < 0.01
    significativo_05 = p_value < 0.05
    significativo_10 = p_value < 0.10

    # Armazenar os resultados no dicionário
    resultados['Modelo'].append(modelo)
    resultados['P-Value'].append(p_value)
    resultados['Significativo_01'].append(significativo_01)
    resultados['Significativo_05'].append(significativo_05)
    resultados['Significativo_10'].append(significativo_10)

# Criar um DataFrame a partir do dicionário
p_value = pd.DataFrame(resultados)

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_p_value = pd.merge(df_r2_mse, p_value, on='Modelo')

# Salvar o DataFrame de Validação no arquivo em excel
df_r2_p_value.to_excel(f'r2_teste_carteira_{carteira}.xlsx', index=False)

print(df_r2_p_value)

```



## APÊNDICE P – CÓDIGO-FONTE PARA CALCULAR AS PREVISÕES DOS META-MODELOS ST-LR, ST-RF, ST-GBRT E ST-NN

```

# Importar as Bibliotecas Principais
import pandas as pd
import numpy as np
import os
import warnings
warnings.filterwarnings('ignore')
import scipy.stats
from scipy.stats import ttest_rel

# Modelos de Previsão
from sklearn.linear_model import LinearRegression # LR - Regressão Linear
from statsmodels.robust.robust_linear_model import RLM # HR - Regressão
de Huber
from sklearn.kernel_ridge import KernelRidge # KR - Regressão de Kernel
from sklearn.linear_model import Lasso # LASSO - Regressão de Lasso
from sklearn.linear_model import ElasticNet # Enet - Regressão
ElasticNet
from sklearn.ensemble import GradientBoostingRegressor # GBRT - Árvores
de Regressão Impulsionadas por Gradiente
from sklearn.ensemble import RandomForestRegressor # RF - Florestas
Aleatórias
from sklearn.feature_selection import RFE # CSR - Regressão Completa de
Subconjunto
from mlxtend.feature_selection import SequentialFeatureSelector # MMA -
Média do Modelo de Mallows
from sklearn.decomposition import PCA # PCA - Análise de Componentes
Principais
from sklearn.neural_network import MLPRegressor # NN - Redes Neurais
Regressor

# Métricas de Avaliação
from sklearn.metrics import accuracy_score, confusion_matrix,
precision_score, recall_score, f1_score
from sklearn.metrics import log_loss, r2_score, mean_squared_error,
mean_absolute_error

""" # Impotar e Estruturar os DataFrames """

# Lembrar de trocar o número da carteira
carteira = 1

```

```

# Caminho dos arquivos contendo as Matrizes de Previsões
nome_do_arquivo = f'matrix_Carteira_{carteira}.xlsx'
caminho_arquivo = fr'C:\Users\Geral\OneDrive\Pessoal\Mestrado UFPE\0
Dissertação\01 Dissertação\Códigos de
Programação\stacking_finance\forecast_finance\{nome_do_arquivo}'
df = pd.read_excel(caminho_arquivo)

# Transformar a Data em Index e depois excluí-la do DataFrame
df = df.set_index(pd.DatetimeIndex(df['Data'].values))
df = df.drop(columns=['Data'])
df

""" # Selecionar os modelos financeiros com os melhores R2OOS """

# Excluir os modelos não selecionados
#df = df.drop(columns=[f'CAPM_{carteira}'])
#df = df.drop(columns=[f'D-CAPM_{carteira}'])
#df = df.drop(columns=[f'L-CAPM_{carteira}'])
#df = df.drop(columns=[f'C-CAPM_{carteira}'])
#df = df.drop(columns=[f'F&F_{carteira}'])
#df = df.drop(columns=['Media'])

"""
# Criar a coluna com a nova média das previsões
df['Media'] = df[[f'CAPM_{carteira}',
                  f'D-CAPM_{carteira}',
                  f'L-CAPM_{carteira}',
                  f'C-CAPM_{carteira}',
                  f'F&F_{carteira}'],
                ].mean(axis=1)

"""
df

""" # Dividir o DataFrame nos conjunto de Treino, Validação e Teste """

linhas = len(df) # obter o tamanho do DataFrame
# ':' após a vírgula significa a seleção de todas as colunas do
dataframe.
df_treino = df.iloc[0:36, :] # 01/2003 até 12/2005
df_validacao = df.iloc[36:120, :] # 01/2006 até 12/2012
df_teste = df.iloc[120:240, :] # 01/2013 até 12/2022

"""# Identificar as variáveis independentes X e a variável dependente
Y"""

```

```

# Selecionar as colunas que possuam algum tipo de número do numpy e
excluir a variável y do dataframe x
X_train = df_treino.select_dtypes(include=np.number).drop(['Return'],
axis=1).fillna(0) #axis=1 é a coluna
y_train = df_treino['Return']

X_val = df_validacao.select_dtypes(include=np.number).drop(['Return'],
axis=1).fillna(0) #axis=1 é a coluna
y_val = df_validacao['Return']

X_test = df_teste.select_dtypes(include=np.number).drop(['Return'],
axis=1).fillna(0) #axis=1 é a coluna
y_test = df_teste['Return']

""" # Realizar Treinamento do Meta-modelo no Conjunto de Validação """

# LR - Regressão Linear
st_lr = LinearRegression()
st_lr.fit(X_val, y_val)

# RF - Florestas Aleatórias
st_rf = RandomForestRegressor(n_estimators=150, n_jobs=6,
random_state=70)
st_rf.fit(X_val, y_val)

# GBRT - Árvores de Regressão Impulsionadas por Gradiente
st_gbrt = GradientBoostingRegressor()
st_gbrt.fit(X_val, y_val)

# NN - Redes Neurais
st_nn = MLPRegressor()
st_nn.fit(X_val, y_val)

""" # Realizar Previsão do Meta-modelo no Conjunto de Teste """

# Criar o DataFrame do Metamodelo apenas com as previsões dos modelos
financeiros
df_meta_modelo = df_teste.drop(columns=['Return'])

# Iterar sobre as linhas de teste para obter as previsões do meta-modelo
for i in range(0, len(df_meta_modelo)):

    # Fazer previsões dos meta-modelos para uma instância de cada vez
    pst_lr = st_lr.predict(df_meta_modelo.iloc[:i+1, :].fillna(0))
    pst_rf = st_rf.predict(df_meta_modelo.iloc[:i+1, :].fillna(0))
    pst_gbrt = st_gbrt.predict(df_meta_modelo.iloc[:i+1, :].fillna(0))

```

```

    pst_nn = st_nn.predict(df_meta_modelo.iloc[:i+1, :].fillna(0))

# Armazenar os resultados das previsões binárias dos Meta-modelos
pst_lr_bin = (pst_lr > 0).astype(int)
pst_rf_bin = (pst_rf > 0).astype(int)
pst_gbrt_bin = (pst_gbrt > 0).astype(int)
pst_nn_bin = (pst_nn > 0).astype(int)
y_bin = []

""" # Atualizar o DataFrame de Teste com os resultados dos Meta-modelos
"""

df_teste['St-LR'] = pst_lr
df_teste['St-RF'] = pst_rf
df_teste['St-GBRT'] = pst_gbrt
df_teste['St-NN'] = pst_nn

# Salvar o DataFrame no arquivo em excel
df_teste.to_excel(f'stacking_matrix_complete_{carteira}.xlsx')
df_teste

"""# Calcular o R2 fora da amostra (R2OSS) baseado na fórmula de Campbell
e Thompson (2008)."""

# DataFrame de Teste

# Atribuir variáveis às previsões dos modelos financeiros
capm = df_teste[f'CAPM_{carteira}']
d_capm = df_teste[f'D-CAPM_{carteira}']
l_capm = df_teste[f'L-CAPM_{carteira}']
c_capm = df_teste[f'C-CAPM_{carteira}']
ff_model = df_teste[f'F&F_{carteira}']
mean = df_teste['Media']

# Criar a coluna HM (média histórica) no dataframe da Carteira 1
df_teste['HM'] = df_teste['Return'].expanding().mean() # HM é a média
histórica (y_mean) do retorno da carteira
y_mean = df_teste['HM'] # "y_mean" recebe a média histórica da carteira
y = df_teste['Return'] # "y" recebe o retorno da carteira
y_bin = (y > 0).astype(int)

# Calcular o MSE da média histórica
mse_mh = mean_squared_error(y, y_mean)

# Calcular o MSE e R2OOS dos modelos financeiros

```

```

# CAPM
capm_mse = mean_squared_error(y, capm)
capm_r2 = 1 - (capm_mse / mse_mh)

# D-CAPM
d_capm_mse = mean_squared_error(y, d_capm)
d_capm_r2 = 1 - (d_capm_mse / mse_mh)

# L-CAPM
l_capm_mse = mean_squared_error(y, l_capm)
l_capm_r2 = 1 - (l_capm_mse / mse_mh)

# C-CAPM
c_capm_mse = mean_squared_error(y, c_capm)
c_capm_r2 = 1 - (c_capm_mse / mse_mh)

# F&F
ff_model_mse = mean_squared_error(y, ff_model)
ff_model_r2 = 1 - (ff_model_mse / mse_mh)

# Media - Media das Previsões
media_mse = mean_squared_error(y, mean)
media_r2 = 1 - (media_mse / mse_mh)

# Calcular o MSE e R2OOS das previsões dos Meta-modelos
# LR - Reressão Linear
pst_lr_mse = mean_squared_error(y, pst_lr)
pst_lr_r2 = 1 - (pst_lr_mse / mse_mh)

# RF - Florestas Aleatórias
pst_rf_mse = mean_squared_error(y, pst_rf)
pst_rf_r2 = 1 - (pst_rf_mse / mse_mh)

# GBRT - Árvores de Regressão Impulsionadas por Gradiente
pst_gbrt_mse = mean_squared_error(y, pst_gbrt)
pst_gbrt_r2 = 1 - (pst_gbrt_mse / mse_mh)

# NN - Redes Neurais
pst_nn_mse = mean_squared_error(y, pst_nn)
pst_nn_r2 = 1 - (pst_nn_mse / mse_mh)

""""# Criar o DataFrame com Resultados do R2OOS e MSE.""""

# DataFrame de Teste

# Listar todos os modelos
nomes_modelos = [f'CAPM {carteira}',

```

```

        f'D-CAPM_{carteira}',
        f'L-CAPM_{carteira}',
        f'C-CAPM_{carteira}',
        f'F&F_{carteira}',
        'Media',
        'St-LR',
        'St-RF',
        'St-GBRT',
        'St-NN']

# Criar os dicionários de R2 e MSE
dict_r2 = {
    f"CAPM_{carteira}": [capm_r2],
    f"D-CAPM_{carteira}": [d_capm_r2],
    f"L-CAPM_{carteira}": [l_capm_r2],
    f"C-CAPM_{carteira}": [c_capm_r2],
    f"F&F_{carteira}": [ff_model_r2],
    "Media": [media_r2],
    "St-LR": [pst_lr_r2],
    "St-RF": [pst_rf_r2],
    "St-GBRT": [pst_gbrt_r2],
    "St-NN": [pst_nn_r2]
}

dict_mse = {
    f"CAPM_{carteira}": [capm_mse],
    f"D-CAPM_{carteira}": [d_capm_mse],
    f"L-CAPM_{carteira}": [l_capm_mse],
    f"C-CAPM_{carteira}": [c_capm_mse],
    f"F&F_{carteira}": [ff_model_mse],
    "Media": [media_mse],
    "St-LR": [pst_lr_mse],
    "St-RF": [pst_rf_mse],
    "St-GBRT": [pst_gbrt_mse],
    "St-NN": [pst_nn_mse]
}

# Criar DataFrames separados para R2 e MSE
df_r2 = pd.DataFrame({'Modelo': nomes_modelos, 'R2':
    [float(dict_r2[model][0]) for model in nomes_modelos]})
df_mse = pd.DataFrame({'Modelo': nomes_modelos, 'MSE':
    [float(dict_mse[model][0]) for model in nomes_modelos]})

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_mse = pd.merge(df_r2, df_mse, on='Modelo')
print(df_r2_mse)

"""# Calcular o p-value."""

```

```

# DataFrame de Teste

# Listar todos os modelos
modelos_p_value = [f'CAPM_{carteira}',
                   f'D-CAPM_{carteira}',
                   f'L-CAPM_{carteira}',
                   f'C-CAPM_{carteira}',
                   f'F&F_{carteira}',
                   'Media',
                   'St-LR',
                   'St-RF',
                   'St-GBRT',
                   'St-NN'
                   ]

# Dicionário para armazenar os resultados
resultados = {'Modelo': [], 'P-Value': [], 'Significativo_01': [],
              'Significativo_05': [], 'Significativo_10': []}

# Loop para calcular os testes e preencher os resultados
for modelo in modelos_p_value:

    # Calcular o teste t-pareado entre o modelo e a média histórica
    t_stat, p_value = ttest_rel(df_teste[modelo], y_mean)

    # Avaliar a significância
    significativo_01 = p_value < 0.01
    significativo_05 = p_value < 0.05
    significativo_10 = p_value < 0.10

    # Armazenar os resultados no dicionário
    resultados['Modelo'].append(modelo)
    resultados['P-Value'].append(p_value)
    resultados['Significativo_01'].append(significativo_01)
    resultados['Significativo_05'].append(significativo_05)
    resultados['Significativo_10'].append(significativo_10)

# Criar um DataFrame a partir do dicionário
p_value = pd.DataFrame(resultados)

# Mesclar os DataFrames usando o nome do modelo como chave
df_r2_p_value = pd.merge(df_r2_mse, p_value, on='Modelo')
print(df_r2_p_value)

"""# Criar o DataFrame binário para calcular a Acurácia."""

# Criar um DataFrame vazio

```

```

df_teste_bin = pd.DataFrame()

# Retorno da Carteira, Modelos Financeiros e Média
df_teste_bin['Return'] = (y > 0).astype(int)
df_teste_bin[f'CAPM_{carteira}'] = (capm > 0).astype(int)
df_teste_bin[f'D-CAPM_{carteira}'] = (d_capm > 0).astype(int)
df_teste_bin[f'L-CAPM_{carteira}'] = (l_capm > 0).astype(int)
df_teste_bin[f'C-CAPM_{carteira}'] = (c_capm > 0).astype(int)
df_teste_bin[f'F&F_{carteira}'] = (ff_model > 0).astype(int)
df_teste_bin['Media'] = (mean > 0).astype(int)

# Meta-modelos
df_teste_bin['St-LR'] = (pst_lr > 0).astype(int)
df_teste_bin['St-RF'] = (pst_rf > 0).astype(int)
df_teste_bin['St-GBRT'] = (pst_gbrt > 0).astype(int)
df_teste_bin['St-NN'] = (pst_nn > 0).astype(int)

"""# Calcular a Acurácia, Sensibilidade e Especificidade."""

# Listar todos os modelos
modelos_acc = [f'CAPM_{carteira}',
               f'D-CAPM_{carteira}',
               f'L-CAPM_{carteira}',
               f'C-CAPM_{carteira}',
               f'F&F_{carteira}',
               'Media',
               'St-LR',
               'St-RF',
               'St-GBRT',
               'St-NN'
               ]

# Dicionário para armazenar os resultados
resultados_acc = {'Modelo': [],
                  'Acurácia': [],
                  'Sensibilidade (recall)': [],
                  'Especificidade (specificity)': [],
                  'Precision Positive': [],
                  'Precision Negative': []
                  }

# Loop para calcular os testes e preencher os resultados
for modelo in modelos_acc:
    accuracy = accuracy_score(y_bin, df_teste_bin[modelo]) # Calcular a
    acurácia
    conf_matrix = confusion_matrix(y_bin, df_teste_bin[modelo])
    tp = conf_matrix[1][1]

```

```

tn = conf_matrix[0][0]
fp = conf_matrix[0][1]
fn = conf_matrix[1][0]
recall = tp / (tp + fn) # Sensibilidade (recall) Pos-Acc
specificity = tn / (tn + fp) # Especificidade (precision) Neg-Acc
positive = tp / (tp + fp)
negative = tn / (tn + fn)

# Armazenar os resultados no dicionário
resultados_acc['Modelo'].append(modelo)
resultados_acc['Acurácia'].append(accuracy)
resultados_acc['Sensibilidade (recall)'].append(recall) #
Sensibilidade (recall) Pos-Acc
resultados_acc['Especificidade (specificity)'].append(specificity) #
Especificidade (precision) Neg-Acc
resultados_acc['Precision Positive'].append(positive)
resultados_acc['Precision Negative'].append(negative)

# Criar um DataFrame a partir do dicionário
accuracy = pd.DataFrame(resultados_acc)

# Mesclar os DataFrames usando o nome do modelo como chave
df_final = pd.merge(df_r2_p_value, accuracy, on='Modelo')

# Salvar o DataFrame de Validação no arquivo em excel
df_final.to_excel(f'stacking_complete_carteira_{carteira}.xlsx',
index=False)
print(df_final)

```