

**CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO  
MESTRADO EM ADMINISTRAÇÃO**

**WENDEL MOREIRA DUARTE**

**MODELO AUTOMATIZADO PARA DECISÕES DE INVESTIMENTO  
USANDO MACHINE LEARNING E A TÉCNICA DE PORTFOLIO  
SELECTION DE BLACK-LITTERMAN**

**BELO HORIZONTE - MG  
2018**

WENDEL MOREIRA DUARTE

MODELO AUTOMATIZADO PARA DECISÕES DE INVESTIMENTO  
USANDO MACHINE LEARNING E A TÉCNICA DE PORTFOLIO  
SELECTION DE BLACK-LITTERMAN

Dissertação de mestrado apresentada ao Centro Federal de Educação Tecnológica de Minas Gerais como parte das exigências do Programa de Pós-Graduação em Administração, área de concentração em Processos e Sistemas Decisórios, para a obtenção título de Mestre.

Orientador

Dr. Rodrigo Tomás Nogueira Cardoso

Coorientador

Dr. Felipe Dias Paiva

BELO HORIZONTE - MG  
2018

D812m Duarte, Wendel Moreira  
Modelo automatizado para decisões de investimento usando machine learning e a técnica de portfólio selection de black-litterman.  
/ Wendel Moreira Duarte. -- Belo Horizonte, 2018.  
76 f. : il.

Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Administração, 2018.

Orientador: Prof. Dr. Rodrigo Tomás Nogueira Cardoso

#### Bibliografia

1. Ações (Finanças). 2. Investimentos. 3. Aprendizado do Computador. I. Cardoso, Rodrigo Tomás Nogueira. II. Centro Federal de Educação Tecnológica de Minas Gerais. III. Título

CDD 332.6



MINISTÉRIO DA EDUCAÇÃO  
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO DO CEFET-MG  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO EM ADMINISTRAÇÃO do(a) Senhor(a) Wendel Moreira Duarte. No dia 02 de julho de 2018, às 14h45min, reuniu-se no Campus II do Centro Federal de Educação Tecnológica de Minas Gerais – CEFET-MG, a Banca Examinadora de dissertação designada pelo Colegiado do Programa de Pós-graduação em Administração do CEFET-MG para julgar o trabalho final intitulado “**Modelo Automatizado para Decisões de Investimento Usando Machine Learning e a Técnica de Portfolio Selection de Black-Litterman**”, requisito para a obtenção do **Grau de Mestre em Administração**, linha de pesquisa: **Processos e Sistemas Decisórios em Arranjos Organizacionais**. Abrindo a sessão, o(a) Senhor(a) Presidente da Banca, Prof.(a) Dr.(a) Felipe Dias Paiva, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra ao(à) aluno(a) para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do(a) aluno(a). Logo após, a Banca se reuniu, sem a presença do aluno e do público, para julgamento e expedição do seguinte resultado final:

Aprovação.

Aprovação com recomendação de aperfeiçoamento, condicionada à satisfação das exigências feitas pela banca examinadora.

Recomendação de reapresentação.

Reprovação.

O resultado final foi comunicado publicamente ao(à) aluno(a) pelo(a) Senhor(a) Presidente da Banca. Nada mais havendo a tratar, o(a) Senhor(a) Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Banca Examinadora.

Belo Horizonte, 02 de julho de 2018.

**Assinaturas:**

Prof. Dr. Felipe Dias Paiva (Coorientador – PPGA-CEFET-MG)

Prof. Dr. Adriano César Machado Pereira (UFMG)

Prof. Dr. Juliano Lima Pinheiro (PPGA CEFET-MG)

Prof. Dr. Rogério Martins Gomes (CEFET-MG)

## AGRADECIMENTOS

Ao Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), agradeço a oportunidade de realização do Mestrado e o suporte financeiro e estrutural concedido em todas etapas do trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) sou grato pelo suporte financeiro concedido.

Aos meus orientadores, Prof. Dr. Rodrigo Tomás Nogueira Cardoso e Prof. Dr. Felipe Dias Paiva, agradeço a amizade, atenção, suporte, tutoria e a oportunidade de aprender o que é ser um professor.

Aos professores e funcionários do PPGA agradeço toda a ajuda, orientação e suporte.

Às colegas de classe, com suas risadas, conversas, companheirismo e ajuda nos momentos mais complicados que todos enfrentamos. Não foram poucos!

Aos amigos e colegas, que em algum momento compartilharam ideias, conversas e viveram comigo momentos cruciais durante essa etapa.

Em especial, à minha família, Tarcísio, Célia, Wilian e Riva, que acompanharam de perto todo esse processo de crescimento, trabalho e aprendizado. Agradeço a presença, compreensão, ombros e colos.

*O conteúdo é sempre concreto e, por conseguinte, imprevisível; há sempre invenção.*

*A única coisa que importa é saber se a invenção que se faz se faz em nome da liberdade.*

*Jean-Paul Sartre*

## RESUMO

O presente trabalho tem como objetivo propor modelos baseados em técnicas de *machine learning*, especificamente, usando redes neurais artificiais (RNA) e *support vector machine* (SVM), para realizar a classificação de ativos financeiros com potencial de atingimento de *target* específico e então compor carteiras utilizando o modelo de Black-Litterman (BL). Quanto aos objetivos específicos, a proposta é de construir um modelo baseado em RNA e outro modelo baseado em SVM para realizar a classificação de ativos financeiros com potencial de atingimento de *target* específico; em seguida, otimizar carteiras utilizando o modelo BL, tendo como parâmetro de *views* as previsões da RNA e SVM; e, por fim, comparar os resultados das carteiras obtidas com os valores de referência (índice Ibovespa e uma carteira elaborada pelo método proposto por Markowitz). A amostra de dados foi composta de ativos do índice Ibovespa, entre Janeiro/2006 e Julho/2017, com periodicidade diária. Os modelos de RNA e SVM trabalharam como classificadores, onde foram gerados diariamente indicativos de atingimento de determinado *target*. Essas previsões foram acopladas ao modelo BL, sendo os *targets* em questão considerados como *views* e a medida de *precision* usada para representar confiança relativa àquela previsão. Para efeito comparativo, foram geradas estratégias alternativas que serviram como base de comparação para as avaliações de performance. Foram compostas (i) uma carteira com geração de sinais aleatórios e posterior combinação com o modelo BL, (ii) carteiras que valeram-se dos sinais gerados pelos classificadores, mas em posterior otimização, (iii) carteiras que valeram-se dos sinais gerados pelos classificadores, mas com posterior acoplamento ao modelo MV, (iv) uma aplicação simples do modelo MV, e por fim, (v) uma carteira com estratégia *buy-and-hold* considerando o índice Ibovespa (IBOV). Os resultados indicam melhor performance do classificador baseado em SVM quando comparado a uma estratégia ingênua, superioridade em termos de cardinalidade, volume de operações e retorno da operação para a estratégia que SVM+BL, tanto em termos totais quanto em análises por biênio. Estes resultados atestam o potencial do uso de técnicas de *soft computing* no processo de formação de portfólios, o que possibilita que o *decision maker* tenha indicativos para montar uma operação mais enxuta, entrando em menos operações com mais assertividade e conseguindo melhores resultados como um todo. As contribuições do trabalho incluem a proposição de um modelo para suportar o *decision maker* no processo de tomada de decisão de investimento, o desenvolvimento de um classificador com assertividade estatisticamente superior a uma solução ingênua, o uso carteira de mínima variância para obtenção da distribuição de equilíbrio do mercado e o acoplamento das informações obtidas via técnicas de *machine learning* ao modelo BL, fazendo uso da medida de *precision* como parâmetro  $\Omega$ .

Palavras-chave: Teoria de seleção de portfólios, Investimentos, Modelo de Black-Litterman, Redes Neurais Artificiais, *Support Vector Machine*, Ibovespa.

## ABSTRACT

This work aims to propose models using machine learning techniques, specifically using artificial neural networks (ANN) and support vector machine (SVM), to predict financial assets which have potential to achieve specific targets and then compose portfolios using the Black-Litterman model (BL). Regarding the specific aims, the propose is to build a model using ANN and another one using SVM to classify financial assets which have potential to achieve specific targets, and then, optimize portfolios using BL model, taking as view parameter the ANN and SVM predictions, and, finally, compare the portfolios with the baselines (Ibovespa index and a portfolio built using the Markowitz's model). The sample was composed by assets in Ibovespa index, from January 2006 to July 2017, considering daily data. The ANN and SVM models operated as classifiers, where was daily generated achievement indicators from specific targets. This predictions was coupled in BL model using the targets as a view parameter and the precision measure used to represent the confidence regarding this view. To compare the performance of the proposed model five other models were configured: (i) a model using signals from a random process coupled in BL model, (ii) a model using signals from the classifiers, but without optimization, (iii) a model using signals from the classifiers coupled in MV model, (iv) the MV model in simple application, and (v) a portfolio using a buy-and-hold strategy based in Ibovespa index. The results indicated higher performance from the classifier based on SVM when compared with a naive approach, superiority in terms of cardinality, operation volume and return for SVM+BL strategy, in a total period and by biennium analysis. This results reinforce the potential of using the soft computing techniques in a portfolio selection process, providing insights for the decision maker to elaborate more assertive strategies, performing a more accuracy operation and achieving better results. The contributions of this work include the proposition from a model to support the decision maker in a process of investment decision, the development of a classifier which has statistical superior assertivity if compared to a naive approach, the use of the minimum variance portfolio to estimate the market equilibrium distribution, and the coupled of the predictions from machine learning techniques with the BL model, using the precision measure as  $\Omega$  parameter.

Keywords: Portfolio selection, Investments, Black-Litterman model, Artificial neural networks, Support vector machine, Ibovespa index.

## LISTA DE FIGURAS

1	Fronteira eficiente . . . . .	18
2	Escolha do portfólio usando a curva de utilidade . . . . .	19
3	<i>Capital market line</i> . . . . .	21
4	Representação esquemática do modelo de Black-Litterman . . . . .	29
5	Modelo de um neurônio . . . . .	32
7	Representação esquemática de uma rede MLP . . . . .	34
8	Hiperplano ótimo de separação . . . . .	35
10	Transformação para dados linearmente não separáveis . . . . .	36
11	Evolução da composição do índice Ibovespa . . . . .	38
12	Desenho do experimento . . . . .	39
13	Classificador baseado em <i>machine Learning</i> . . . . .	41
14	Janelas de aprendizado e simulação da RNA . . . . .	41
15	Combinações de parâmetros para modelo de RNA . . . . .	42
17	RNA - Precision do classificador por biênio . . . . .	47
18	RNA - Ordens de compra e venda por biênio . . . . .	48
19	RNA - Cardinalidade média dos modelos . . . . .	48
20	RNA - Retorno acumulado sem custos de transação . . . . .	49
21	RNA - Retorno bienal por modelo . . . . .	51
24	SVM - Precision do classificador por biênio . . . . .	54
25	SVM - Ordens de compra e venda por biênio . . . . .	55
26	SVM - Cardinalidade média dos modelos . . . . .	56
27	SVM - Retorno acumulado sem custos de transação . . . . .	57
28	SVM - Retorno bienal por modelo . . . . .	58
30	Geral - Ordens de compra e venda por biênio . . . . .	61
31	Geral - Cardinalidade média dos modelos . . . . .	62
32	Geral - Retorno acumulado sem custos de transação . . . . .	63
33	Geral - Retorno bienal por modelo . . . . .	65
35	Geral - Retorno x Risco - sem custos de transação . . . . .	67

## LISTA DE TABELAS

1	RNA - Desempenho do classificador (%) . . . . .	46
2	RNA - Trades (compra e venda) por modelo . . . . .	47
3	RNA - Teste de Dunn para cardinalidade . . . . .	49
4	RNA - Retorno médio mensal . . . . .	49
5	RNA - Estatísticas adicionais . . . . .	50
6	RNA - Retorno acumulado por biênio (%) . . . . .	50
7	RNA - Volatilidade por biênio (%) . . . . .	51
8	SVM - Desempenho do classificador (%) . . . . .	54
9	SVM - Trades (compra e venda) por modelo . . . . .	55
10	SVM - Teste de Dunn para cardinalidade . . . . .	56
11	SVM - Retorno médio mensal . . . . .	56
12	SVM - Teste de Dunn para retorno mensal . . . . .	57
13	SVM - Estatísticas adicionais . . . . .	58
14	SVM - Retorno acumulado por biênio (%) . . . . .	58
15	SVM - Volatilidade por biênio (%) . . . . .	59
16	Trades (compra e venda) por modelo . . . . .	61
17	Teste de Dunn para cardinalidade . . . . .	62
18	Retorno médio mensal . . . . .	62
19	Teste de Dunn para retorno mensal . . . . .	63
20	Estatísticas adicionais . . . . .	64
21	Retorno acumulado por biênio (%) . . . . .	64
22	Volatilidade por biênio (%) . . . . .	65
23	Teste de Dunn para volatilidade . . . . .	66

## LISTA DE QUADROS

1	Atributos para modelagem . . . . .	40
2	Empresas consideradas na amostra . . . . .	76

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>14</b>
2.1	Teoria de seleção de portfólios: o modelo de Markowitz	14
2.1.1	Fronteira eficiente e escolha do portfólio ótimo	17
2.2	<i>Capital asset pricing model</i>	19
2.2.1	<i>Capital market line</i>	21
2.2.2	<i>Security market line</i>	23
2.2.3	Implicações gerais sobre o modelo CAPM	23
2.3	O modelo de Black-Litterman	24
2.3.1	Equilíbrio de mercado	25
2.3.2	Escolha da <i>proxy</i> de mercado	26
2.3.3	Visões do investidor	27
2.3.4	Parâmetros de ajuste	29
2.4	<i>Machine learning</i>	30
2.4.1	Redes neurais artificiais	31
2.4.2	<i>Support vector machine</i>	34
<b>3</b>	<b>METODOLOGIA</b>	<b>38</b>
3.1	Conjunto amostral	38
3.2	Desenho do experimento	39
3.2.1	Tratamento dos dados e cálculo dos indicadores	39
3.2.2	Classificação dos ativos com potencial para atingimento do <i>target</i>	40
3.2.3	Otimização das carteiras usando o modelo BL	42
3.2.4	Comparação com as <i>baselines</i> propostas	43
3.2.4.1	Naive+BL	43
3.2.4.2	RNA+1/N e SVM+1/N	43
3.2.4.3	RNA+MV e SVM+MV	44
3.2.4.4	MV	44
3.2.5	Modelo de operação	44
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>45</b>
4.1	Resultados das redes neurais artificiais	45
4.2	Resultados da <i>support vector machine</i>	52
4.3	Comparações com as <i>baselines</i>	60
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>69</b>
	<b>REFERÊNCIAS</b>	<b>71</b>
<b>A</b>	<b>APÊNDICE</b>	<b>76</b>

## 1 INTRODUÇÃO

Decisões de seleção de portfólio ótimo mantêm-se em alta como objeto de discussão no campo de finanças. Definir em qual carteira colocar seus recursos, quantos ativos deve ter essa carteira e quais ativos devem compô-la são alguns dos questionamentos que os investidores fazem ao se confrontarem com a necessidade de tomar uma decisão de investimento. Entretanto, esta é uma tarefa árdua, visto que o mercado financeiro é extremamente complexo e dinâmico, absorvendo diferentes formas de estímulos dos participantes do mercado, como também os acontecimentos das esferas política e econômica em escala global.

Harry Markowitz, reconhecido como o pai da moderna teoria de finanças, publicou no ano de 1952 um dos mais importantes e revolucionários artigos<sup>1</sup> no âmbito das finanças. Neste seminal trabalho ele defendeu que um investidor poderia maximizar ganhos e minimizar riscos de seus investimentos por meio da adoção de um portfólio diversificado. Tal ideia balizou-se num modelo de otimização, hoje considerado simples e perspicaz, mas que para os padrões computacionais da época foi bastante arrojado e de difícil compreensão quanto à sua aplicabilidade<sup>2</sup>.

A essência do modelo de Markowitz repousa na premissa de que os investidores, maximizadores de utilidade, buscam, ao mesmo tempo, o máximo retorno para um dado nível de risco, e o mínimo risco para um dado nível de retorno. Para lidar com esses objetivos discordantes, ele propôs que se fizessem estimativas do retorno, variância e correlação entre os ativos, para que então, valendo-se da solução de um problema de otimização multiobjetivo, o investidor seja capaz de escolher uma combinação de risco-retorno desejada.

Contudo, desde a divulgação do modelo de Markowitz novas discussões têm sido suscitadas em diferentes frentes de estudo, sejam relacionadas às estratégias de investimento e financiamento ou até mesmo referentes às políticas operacionais. Destacam-se três críticas que são centrais quando se trata desse modelo: o uso da média como estimativa dos retornos esperados e o uso da variância como medida de risco.

Quanto à primeira crítica, alguns autores<sup>3</sup> sustentam o argumento de que a média é um estimador de retorno que possui muitos erros, e que também devem ser consideradas outras informações sobre os ativos além de apenas métricas estatísticas. Esses erros de estimativa levam a portfólios não diversificados e contraintuitivos, que muitas vezes assumem posições extremadas, chamados de portfólios “mal-comportados”. Já a segunda crítica diz respeito à forma como o investidor percebe o risco. Assumindo a variância como medida de risco, considera-se que as variações nas duas caudas são indesejadas, quando, de fato, o investidor quer abster-se apenas das variações na cauda das perdas.

Para mitigar tanto essas críticas quanto problemas de ordem prática, no início dos anos de 1990, Fischer Black e Robert Litterman propuseram uma abordagem<sup>4</sup> baseada na

<sup>1</sup>MARKOWITZ, H. M. Portfolio selection. **The Journal of Finance**, v. 7, n. 1, p. 77–91, 1952.

<sup>2</sup>“... when I defended my dissertation as a student in the Economics Department of the University of Chicago, Professor Milton Friedman argued that portfolio theory was not Economics, and that they could not award me a Ph.D. degree in Economics for a dissertation which was not in Economics.” (MARKOWITZ, 1991, p. 286)

<sup>3</sup>Destacam-se os trabalhos de DeMiguel, Garlappi e Uppal (2009) e Tu e Zhou (2011)

<sup>4</sup>BLACK, F.; LITTERMAN, R. B. Asset allocation: combining investor views with market equilibrium. **The Journal of Fixed Income**, v. 1, n. 2, p. 7–18, 1991; BLACK, F.; LITTERMAN, R. B.

combinação entre o equilíbrio de mercado e as perspectivas individuais do investidor, criando uma metodologia que permite gerar portfólios diversificados e intuitivos. Nesse trabalho, as visões do investidor representam os seus sentimentos subjetivos em relação a mercados ou ativos específicos, sendo incorporados ao modelo por meio de técnicas de estatística bayesiana.

Utilizando essa metodologia, o investidor pode expressar suas visões em relação a um ponto de referência, que no modelo é definido como o equilíbrio de mercado. Destaca-se que o investidor não precisa necessariamente se posicionar sobre todos os ativos do conjunto, mas apenas sobre aqueles que possuem alguma inclinação de fato. Esse conjunto de perspectivas está ligado a uma variável que expressa a confiança da visão para cada ativo, de forma a definir quais visões são mais relevantes. Essa variável pode ser usada junto a um escalar, que representa a incerteza na estimativa dos retornos, de forma a realizar um ajuste refinado do modelo. Por meio desse ajuste a carteira pode tender à distribuição de equilíbrio ou às visões do investidor, de acordo com as condições contingenciais.

O uso dessas variáveis de refino traz ao modelo muita versatilidade, contudo, esses trabalhos iniciais que apresentam a metodologia não elucidam de forma clara como esses parâmetros devem ser calculados, e em função disso muito se discute na literatura sobre a melhor forma de proceder com essas estimativas. Além disso, o fato de o modelo fazer uso das perspectivas do investidor e suas visões subjetivas abre um pressuposto de que melhores desempenhos do portfólio estão intimamente ligados à expertise do investidor e sua capacidade de compreensão do mercado.

Deduz-se, então, que o desempenho do portfólio está limitado pela capacidade de compreensão, previsão e acesso à informação de mercado do investidor. O mercado financeiro, tal qual foi supracitado, é um ambiente complexo e dinâmico, onde existe um grande número de informações circulando em tempo real, causando movimentos de diversas naturezas. Em razão disso, torna-se infactível uma completa compreensão e manejo de todas as variáveis, considerando a capacidade cognitiva de um simples indivíduo.

Dada essa limitação, destacam-se os avanços dos estudos em inteligência artificial, que buscam desenvolver técnicas que possibilitem o processamento de um grande volume de informação, em vistas de ampliar em diversos pontos a capacidade de acesso e compreensão do cérebro humano. Um dos campos de estudo em inteligência artificial, denominado de *machine learning*, trata do estudo de mecanismos adaptativos que permitem que computadores aprendam por meio de exemplos, analogias e experiências, de forma que os sistemas possam melhorar sua performance com o passar do tempo.

As redes neurais artificiais e o *support vector machine* são exemplos de abordagens estudadas em *machine learning*, onde por meio de arquiteturas organizadas para propiciar uma ampla capacidade de processamento é possível fazer estimativas de eventos futuros, com base em dados atuais e passados, com um notável nível de assertividade. A vasta versatilidade das técnicas e a capacidade de aprender por meio de experiência e generalizar as informações trazem grande potencial de previsão assertiva e soluções de problemas relacionados a um ambiente tão complexo e dinâmico como o mercado financeiro.

Diante do exposto, o presente trabalho busca responder a seguinte questão: como integrar

previsões de ativos financeiros baseados em *machine learning* ao processo de otimização de portfólio fundamentado no modelo de Black-Litterman? Portanto, o objetivo geral deste trabalho é propor modelos baseados em *machine learning* para realizar previsões de retornos de ativos financeiros que serão usadas para compor carteiras utilizando o modelo de Black-Litterman.

Para responder ao questionamento exposto, o trabalho busca atingir os seguintes objetivos específicos:

- i Construir um modelo baseado em redes neurais artificiais para realizar previsões de ativos financeiros.
- ii Construir um modelo baseado em *support vector machine* para realizar previsões de ativos financeiros.
- iii Otimizar carteiras utilizando o modelo de Black-Litterman, tendo como parâmetro de *views* as previsões da RNA e SVM.
- iv Comparar os resultados das carteiras obtidas com os valores de referência (índice Ibovespa e uma carteira elaborada pelo método proposto por Markowitz).

Dito isso, destaca-se que este trabalho é relevante, pois traz como contribuição o uso do modelo de Black-Litterman, uma metodologia recente e muito discutida na literatura de finanças, associado a técnicas oriundas da ciência da computação utilizadas amplamente em diversos campos dada a sua versatilidade. A proposição de modelos de tomada de decisões de investimento que fazem o uso de técnicas de *machine learning* busca explorar o potencial dessas ferramentas para lidar com sistemas complexos, que apresentam dados variados, volumosos e heterogêneos.

Resultados promissores dessa combinação podem ser incorporados como ferramentas de suporte a indivíduos ou organizações que necessitem tomar decisões de investimento. Ademais, o trabalho ampliará a discussão de conceitos básicos dentro da Teoria de Seleção de Portfólios, bem como apresentará um modelo criado a partir de uma associação ainda não contemplada na literatura do campo.

O trabalho está dividido da seguinte forma: a seção 2 apresenta os pressupostos teóricos que fundamentam o trabalho, e a seção 3 apresenta o método utilizado no trabalho, os critérios para seleção e tratamento dos dados, bem como faz as especificações do modelo a ser utilizado. A seção 4 apresenta os resultados obtidos e as comparações propostas. Por fim, a seção 5 traz as considerações finais e perspectivas de trabalhos futuros.

## 2 REFERENCIAL TEÓRICO

Nesta seção são apresentados os pressupostos teóricos que respaldam o presente estudo. O primeiro item trata do modelo de média-variância para seleção de portfólios desenvolvido por Markowitz. Em seguida, faz-se uma breve discussão do modelo de precificação de ativos financeiros (CAPM) desenvolvido a partir da teoria de portfólios de Markowitz. Tal vereda se faz necessária para o entendimento da abordagem baseada em equilíbrio adotada por Black e Litterman na proposição do seu modelo, tratada em detalhes no terceiro item desta seção. Por fim, são apresentados os conceitos básicos de redes neurais artificiais e *support vector machine*, com um destaque especial para as tipologias usadas no presente trabalho.

### 2.1 Teoria de seleção de portfólios: o modelo de Markowitz

Um marco nos estudos em finanças, o trabalho de Markowitz (1952) é reconhecido como precursor da moderna teoria de finanças. O seminal trabalho apresentou uma proposta matemática para lidar com o *trade off* da maximização do retorno esperado e a minimização do risco para uma cesta de investimento. Conforme Fabozzi e Markowitz (2011), o modelo é delineado como uma teoria normativa, ou seja, é uma teoria que descreve um padrão ou norma de comportamento na qual assume-se que os investidores seguem durante o processo de construção de um portfólio, em vez de concentrar-se em uma previsão sobre o comportamento real individual dos mesmos.

Assim, o modelo de média-variância (MV) proposto por Markowitz tem como ponto de partida a premissa de que *“the investor does (or should) consider expected return a desirable thing and variance of return a undesirable thing”* (MARKOWITZ, 1952, p. 77). De acordo com Medeiros, Passos e Vasconcelos (2014), o modelo MV visa minimizar a volatilidade condicional esperada do portfólio, dado um determinado retorno condicional esperado e a covariância dos ativos. Desta forma, o investidor tem dois objetivos conflitantes e guiadores do processo de otimização que irão culminar na tomada de decisão final: maximizar o retorno esperado e minimizar o risco da operação.

Dessa forma, o modelo de média-variância pode ser expresso por meio de uma formulação multiobjetiva de otimização, que buscará um ótimo local com o intuito de minimizar o risco e maximizar o retorno esperado. Essa geração de um conjunto de soluções ótimas é denominada de fronteira eficiente de investimentos, assunto que será tratado com mais detalhes adiante. Cabe, agora, apresentar o modelo formal, descrito nas equações 1 a 4:

$$\min_{w_1, \dots, w_n} \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \quad (1)$$

$$\max_{w_1, \dots, w_n} \sum_{i=1}^n w_i \mu_i \quad (2)$$

Sujeito a:

$$\sum_{i=1}^n w_i = 1 \quad (3)$$

$$0 \leq w_i \leq 1, \quad \forall i = 1, \dots, n \quad (4)$$

Onde:

$w_i$  = proporção do valor inicial investido no portfólio para o ativo i;

$w_j$  = proporção do valor inicial investido no portfólio para o ativo j;

$\sigma_{ij}$  = covariância entre os ativos i e j;

$\mu_i$  = retorno esperado do ativo i.

O mesmo conjunto ótimo de ativos também pode ser alcançado por meio de uma formulação mono-objetiva. Para isso, introduz-se ao modelo uma variável que representa a aversão ao risco do investidor, fator que descreverá o comportamento do investidor frente às opções de investimento de risco. Essa formulação está representada nas equações 5 a 7:

$$\min_{w_1, \dots, w_n} \lambda \left[ \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \right] - (1 - \lambda) \left[ \sum_{i=1}^n w_i \mu_i \right] \quad (5)$$

Sujeito a:

$$\sum_{i=1}^n w_i = 1 \quad (6)$$

$$0 \leq w_i \leq 1, \quad \forall i = 1, \dots, n \quad (7)$$

Onde:

$\lambda$  = coeficiente de aversão ao risco;

$w_i$  = proporção do valor inicial investido no portfólio para o ativo i;

$w_j$  = proporção do valor inicial investido no portfólio para o ativo j;

$\sigma_{ij}$  = covariância entre os ativos i e j;

$\mu_i$  = retorno esperado do ativo i.

Quanto à métrica de retorno de um portfólio em um determinado período de tempo, a estimativa é feita por meio da média dos retornos que o compõem, ponderado pela proporção de cada ativo na carteira. De acordo Fabozzi e Markowitz (2011), num escopo de gestão de portfólios, os investidores querem saber antecipadamente o retorno de um portfólio, ou, em outros termos, querem saber qual o retorno esperado de um portfólio.

Em se tratando dessa estimativa, acredita-se que

*If there is no reason to believe that future long-run returns should differ significantly from historical long-run returns, then probabilities assigned to different return outcomes based on the historical long-run performance of an uncertain investment could be a reasonable estimate for the probability distribution (FABOZZI; MARKOWITZ, 2011, p. 51).*

Sendo assim, o retorno esperado de um conjunto de ativos pode ser descrito como “um conjunto de resultados e com probabilidades de ocorrência associadas a cada um” (ELTON; GRUBER; BROWN, 2012, p. 46). Nessa condição em que os possíveis retornos da carteira são associados a uma distribuição de probabilidade, o retorno esperado seria

a medida de tendência central desses retornos. Desta forma, o retorno de um portfólio específico pode ser calculado a partir da média ponderada pelo peso de cada ativo no portfólio. O cálculo do retorno de um portfólio é apresentado na equação 8:

$$\bar{R}_p = \sum_{i=1}^n w_i \bar{R}_i \quad (8)$$

Onde:

$\bar{R}_p$  = retorno esperado do portfólio;

$w_i$  = proporção do valor inicial investido no portfólio para o ativo  $i$ ;

$\bar{R}_i$  = retorno esperado para o ativo  $i$ .

Alguns autores expõem que devido à incerteza e ao conseqüente erro na estimativa do retorno esperado, a estratégia de otimização MV não conduzirá ao almejado portfólio ótimo, sendo preferível então adotar uma estratégia tida como ingênua e que exige menor custo computacional (DEMIGUEL; GARLAPPI; UPPAL, 2009; TU; ZHOU, 2011). DeMiguel, Garlappi e Uppal (2009) sugerem que para que o uso do modelo de MV seja efetivo é necessário que se minimize os erros no estimador de retorno esperado, por meio não apenas de estimativas estatísticas, mas também de outras informações disponíveis sobre os ativos.

Nessa linha de raciocínio, a literatura discorre sobre o delineamento de estratégias de otimização que levam em conta apenas as medidas de risco, dada a dificuldade de trabalhar com os estimadores de retorno. Faz-se então necessária a discussão dos conceitos de risco e diversificação para tratar de abordagens que seguem essa linha de raciocínio.

Em vista disso, ainda em um uma concepção de distribuição de probabilidades, o risco de um portfólio é uma medida de dispersão em torno da média, e para tanto tem-se o desvio padrão ou a variância (ELTON; GRUBER; BROWN, 2012). Assim, a variância associada à distribuição de retornos mede a proximidade da distribuição em torno da média, ou, no caso, do retorno esperado (FABOZZI; MARKOWITZ, 2011). Desta forma, quanto menor a variância de um ativo, menor é o risco associado a ele, sendo que um ativo livre de risco tem uma dispersão em torno do retorno esperado igual a zero, ou seja, o retorno desse ativo é certo.

Contudo, diferente do retorno esperado, o risco de um portfólio não irá depender apenas da parcela de cada ativo na composição total do portfólio, mas também da forma como os ativos se movem em conjunto, movimento que é estudado por meio da covariância desses ativos. Uma medida de covariância positiva entre os ativos indicará que os retornos dos ativos irão se mover no mesmo sentido, em uma determinada proporção. Analogamente, uma medida de covariância negativa entre os ativos indicará que os retornos dos ativos irão se mover em sentido contrário, em uma determinada proporção. Isto posto, faz-se necessária a discussão do processo de diversificação e da forma como esse processo é tratado pelo modelo MV.

A ideia de diversificação se fundamenta na afirmação de que a alocação de recurso em mais de um ativo pode gerar um risco menor que o dos ativos avaliados separadamente. Rubinstein (2002, p. 1043) afirma que:

*the decision to hold a security should not be made simply by comparing its expected return and variance to others, but rather the decision to hold any security would depend on what other securities the investor wants to hold. Securities could not be properly evaluated in isolation, but only as a group.*

Conforme Fabozzi e Markowitz (2011), o processo de diversificação proposto pelo modelo MV busca combinar ativos com retornos com correlação negativa, num esforço de reduzir o risco dos ativos em conjunto sem abrir mão do retorno esperado, sendo este o fator que torna o modelo mais eficiente que uma abordagem de diversificação ingênua.

Com efeito, deve-se ressaltar que mesmo com um processo eficiente de diversificação não se reduz totalmente o risco, sendo este limitado à possível redução a um risco equivalente à covariância entre os ativos que compõem o conjunto (ELTON; GRUBER; BROWN, 2012; MARKOWITZ, 1952). Assim, o cálculo da variância de um portfólio é apresentado na equação 9:

$$\sigma_P^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \quad (9)$$

Onde:

$\sigma_P^2$  = variância do portfólio;

$w_i$  = proporção do valor inicial investido no portfólio para o ativo i;

$w_j$  = proporção do valor inicial investido no portfólio para o ativo j;

$\sigma_{ij}$  = covariância dos ativos i e j.

Uma crítica bem comum ao modelo MV diz respeito ao uso da variância como medida do risco. Isso se deve ao fato de que uma análise baseada na variância elimina ambas as caudas da distribuição e, por conseguinte, ela previne contra grandes perdas e grandes ganhos. Não obstante, os investidores percebem o risco como variações no sentido das perdas, dado que variações anormais no sentido dos ganhos seria algo desejado. Em vista disso, surgem as chamadas medidas de *downside risk*, medidas que trabalham apenas a cauda das perdas, o que faz com que os resultados tendam a apresentar-se melhores (MARKOWITZ, 1959).

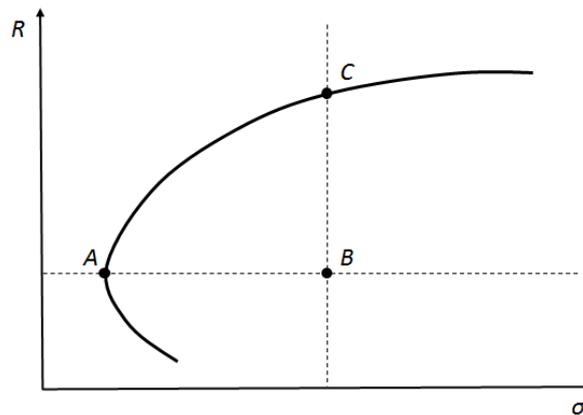
Ademais, existem diversas outras opções entre medidas de risco, ficando a decisão por uma ou outra condicionada às propriedades da distribuição de retornos. Entretanto, os estimadores de risco alternativos geralmente são mais difíceis de se estimar corretamente, ficando a cargo do investidor ou do analista ponderar entre um possível erro de estimativa e uma medida de risco conceitualmente mais adequada às particularidades da análise em questão (MICHAUD; MICHAUD, 2008).

### 2.1.1 Fronteira eficiente e escolha do portfólio ótimo

A par dos conceitos de retorno e risco, o investidor pode valer-se dessas duas medidas para tomar sua decisão de investimento, onde, em uma representação geométrica, obtém-se a chamada fronteira eficiente com as possíveis soluções ótimas de risco e retorno. Por uma solução dominante deve-se entender como aquela solução que maximiza o retorno a

um nível de risco, ou que minimiza o risco dado um nível de retorno (FABOZZI; MARKOWITZ, 2011). Assim, pode-se afirmar que a fronteira eficiente é uma curva formada pelos pontos dominantes dada uma relação de risco e retorno.

A Figura 1 ilustra esse conceito. Observa-se que entre os pontos A e B tem-se o mesmo nível de retorno esperado para diferentes exposições ao risco, logo, pode-se dizer que o ponto A é dominante em relação ao ponto B, pois um investidor racional irá preferir o ponto A. Analogamente, entre os pontos B e C tem-se o mesmo nível de exposição ao risco para diferentes níveis de retorno esperado, logo, pode-se dizer que o ponto C é dominante em relação ao ponto B, pois um investidor racional irá preferir o ponto C.

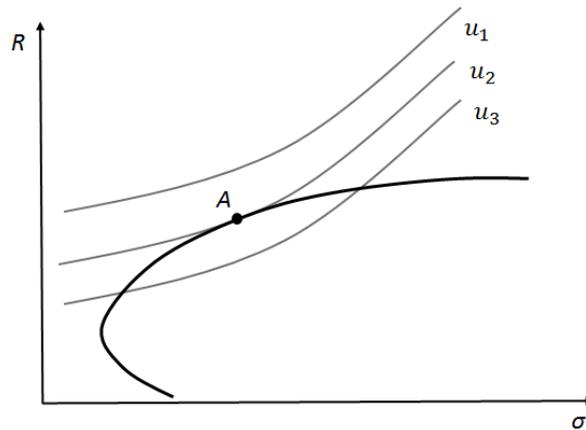


**Figura 1: Fronteira eficiente**

Fonte: adaptado de Bodie, Kane e Marcus (2015).

Com efeito, os portfólios que compõem a fronteira eficiente representam o *trade off* entre retorno e risco citado anteriormente, de forma que, dentro do espaço de soluções dominantes, escolher um risco menor implica abrir mão de certa quantidade de retorno, e escolher um retorno maior implica assumir mais risco. Assim, de posse desse conjunto de soluções, a escolha de um ponto específico que indique o portfólio ótimo cabe ao investidor, que fará a escolha de acordo com suas características como tomador ou não de riscos e com seu comportamento em relação a esse *trade off* (FABOZZI; MARKOWITZ, 2011).

Uma maneira de encontrar esse ponto é valendo-se das curvas de indiferença, que representam a função utilidade do investidor. Uma curva de indiferença indica as combinações de risco e retorno esperado, dado o mesmo nível de utilidade. A escolha, então, é definida no ponto tangente da curva de indiferença e a fronteira eficiente (COPELAND; WESTON; SHASTRI, 2005; FABOZZI; MARKOWITZ, 2011), representado pelo ponto A na Figura 2.



**Figura 2: Escolha do portfólio usando a curva de utilidade**

Fonte: adaptado de Fabozzi e Markowitz (2011).

Alternativamente, retomando as discussões no âmbito das dificuldades na estimativa do retorno esperado, a escolha pode ser dada sem se levar em conta o retorno esperado, observando-se apenas a questão do risco na escolha do portfólio ótimo. Esse é um caso especial do modelo de Markowitz denominado de carteira de mínima variância, em que o parâmetro de aversão ao risco é considerado infinito (SANTOS; TESSARI, 2012) e a resolução do problema pode ser dada conforme a equação 10:

$$\min_{w_1, \dots, w_n} \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \quad (10)$$

Onde:

$w_i$  = proporção do valor inicial investido no portfólio para o ativo  $i$ ;

$w_j$  = proporção do valor inicial investido no portfólio para o ativo  $j$ ;

$\sigma_{ij}$  = covariância dos ativos  $i$  e  $j$ .

Outra forma muito comum na literatura de definir esse ponto é a escolha da melhor relação entre retorno esperado por unidade de risco assumida. Essa abordagem foi proposta por Sharpe (1964) e Sharpe (1966), e é conhecida como índice de Sharpe (IS). Assim, a escolha nessa concepção será dada pelo ponto em que se obtém a maior quantidade de retorno por unidade de risco tomada, assumindo uma taxa livre de risco na operação. O modelo para essa operação é apresentado na equação 11:

$$\max IS = \frac{R_P - R_f}{\sigma_P} \quad (11)$$

Onde:

$R_P$  = retorno esperado do portfólio;

$R_f$  = taxa livre de risco;

$\sigma_P$  = desvio padrão do portfólio.

## 2.2 Capital asset pricing model

Após a divulgação do trabalho de Markowitz, três trabalhos foram desenvolvidos de forma independente por Sharpe (1964), Lintner (1965) e Mossin (1966), apresentando um modelo

de precificação de ativos financeiros (CAPM - *capital asset pricing model*), com o intuito de desenvolver uma teoria que descrevesse o relacionamento do mercado com ativos de risco em condições de equilíbrio. Destaca-se que o modelo veio a se tornar uma peça central da economia financeira moderna (BODIE; KANE; MARCUS, 2015), e que a principal contribuição do modelo é mostrar que o retorno esperado de um ativo está relacionado com sua medida de risco, conhecida como beta (ALEXANDER; SHARPE; BAILEY, 2001).

Para tal feito, o modelo CAPM, como qualquer outro modelo, faz uso de premissas básicas que servem para simplificar uma realidade complexa que se deseja estudar, permitindo o nível de abstração necessário e focando apenas nos elementos mais importantes (ALEXANDER; SHARPE; BAILEY, 2001). Estipular as premissas é um processo que exige ponderação, pois a falta dessas deixa o problema complexo a ponto de inviabilizar a análise, e o excesso pode resultar em um modelo insustentável quando contraposto ao ambiente real.

Isso posto, as premissas do modelo CAPM (ALEXANDER; SHARPE; BAILEY, 2001), são:

- i Os investidores tomam suas decisões a respeito de seus portfólios com base no retorno esperado em seu desvio padrão, considerando um horizonte de investimentos de um período.
- ii Investidores são insaciáveis, logo, dada a escolha entre dois portfólios com o mesmo desvio padrão, eles irão optar pelo portfólio de maior retorno esperado.
- iii Investidores possuem aversão ao risco, logo, dada a escolha entre dois portfólios com o mesmo retorno esperado, eles irão optar pelo portfólio de menor desvio padrão.
- iv Os ativos são infinitamente divisíveis, o que implica que o investidor pode comprar uma fração de um ativo se assim desejar.
- v Existe uma taxa livre de risco sob a qual pode-se emprestar ou tomar emprestado qualquer valor monetário.
- vi Impostos e custos de transação são irrelevantes.
- vii Todos os investidores consideram o mesmo horizonte de investimento, no caso, um período.
- viii A taxa livre de risco é a mesma para todos os investidores.
- ix Não existe assimetria de informações, ou seja, todas as informações relevantes estão disponíveis para todos os investidores de forma gratuita e instantânea.
- x Investidores possuem expectativas homogêneas, ou seja, eles possuem as mesmas percepções sobre retorno esperado, desvio padrão e covariância dos ativos.

Percebe-se que os pressupostos do CAPM reduzem a situação a um caso extremo, criando um ambiente em que todos os investidores analisam a situação da mesma forma em um

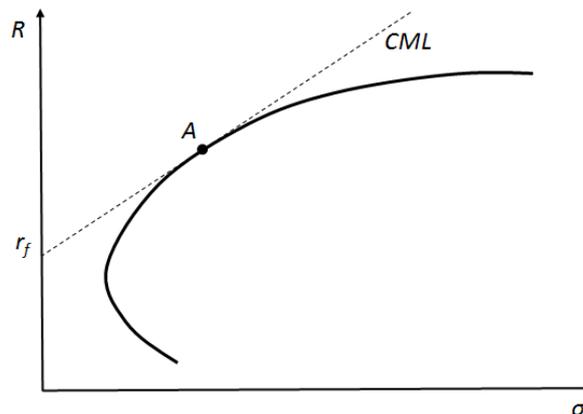
mercado em que não existem potenciais impedimentos para a negociação dos ativos. Elton, Gruber e Brown (2012) afirmam que existe o questionamento de que, dadas premissas tão restritivas, o modelo se tornaria insustentável. Contudo, os autores destacam que a questão relevante é avaliar o quanto da realidade é distorcida quando são feitas essas suposições e quais são as conclusões sobre os mercados de capitais que elas trazem.

De acordo com Alexander, Sharpe e Bailey (2001, p. 191),

*This approach allows the focus to shift from how an individual should invest to what would happen to securities prices if everyone invest in a similar manner. Examining the collective behavior of all investors in the marketplace enables one to develop the resulting equilibrium relationship between each security's risk and return.*

### 2.2.1 *Capital market line*

Dadas as premissas discutidas anteriormente, passa-se à análise das implicações do modelo, ou seja, à discussão do que aconteceria se todos os investidores compartilhassem um universo idêntico de investimentos possíveis e utilizassem a mesma lista de dados para traçar a fronteira eficiente (BODIE; KANE; MARCUS, 2015). É evidente que, dadas as mesmas contingências e as mesmas expectativas, os investidores irão obter em equilíbrio o mesmo portfólio, localizado no ponto em que a taxa livre de risco tangencia a fronteira eficiente, representando o ponto com a melhor relação risco-retorno (ALEXANDER; SHARPE; BAILEY, 2001). A Figura 3 ilustra esse conceito.



**Figura 3:** *Capital market line*

Fonte: adaptado de Bodie, Kane e Marcus (2015).

Assim sendo, todas as escolhas dos investidores serão dadas por uma relação linear da combinação entre a carteira eficiente e a taxa livre de risco. Os investidores podem tomar emprestado para investir mais na carteira de risco ou emprestar uma parcela do seu capital e investir o restante na carteira de risco, escolhas essas deliberadas em função das suas características como tomadores de risco. Esse cenário implica que, somadas todas as carteiras de risco de todos os investidores individuais, a concessão e a contratação de

crédito se anulam e o valor da carteira de risco agregada será igual à riqueza total da economia, sendo esta a carteira de mercado (BODIE; KANE; MARCUS, 2015).

Ressalta-se que, em equilíbrio, todos os ativos têm uma proporção diferente de zero na carteira de mercado. Caso contrário, se um ativo tivesse proporção zero na carteira seu preço iria cair o que aumentaria seu retorno esperado, e faria com que os investidores comprassem esse ativo, fazendo seu preço subir novamente até uma condição de equilíbrio. Assim, cada ativo é mantido na carteira na proporção do seu valor de mercado em relação ao valor de mercado total de todos os ativos de risco (ALEXANDER; SHARPE; BAILEY, 2001; BODIE; KANE; MARCUS, 2015; ELTON; GRUBER; BROWN, 2012).

Uma segunda implicação é a de que qualquer investidor estará satisfeito com aplicações em um fundo de mercado mais a capacidade de tomar emprestado ou emprestar a uma taxa livre de risco, afirmação conhecida como teorema dos fundos mútuos (ELTON; GRUBER; BROWN, 2012). Neste sentido, Alexander, Sharpe e Bailey (2001, p. 192) afirmam que *“the optimal combination of risky assets for an investor can be determined without any knowledge of the investor’s preference toward risk and return”*. Ou seja, independentemente da posição de tomar ou emprestar capital a uma taxa livre de risco, todos os investidores irão investir na mesma proporção em ativos de risco.

Um terceira implicação é a de que todos os investidores vão acabar tendo uma carteira em algum ponto da reta traçada entre o ponto da taxa livre de risco e a carteira de mercado, denominada de *Capital Market Line* (CML) (ELTON; GRUBER; BROWN, 2012). Conforme Alexander, Sharpe e Bailey (2001), a inclinação da CML é dada pela diferença entre o retorno da carteira de mercado e o retorno da taxa livre de risco divididos pela diferença entre os respectivos riscos. Como o intercepto vertical da CML é a taxa livre de risco, a equação da CML é dada conforme a equação 12:

$$R_P = R_f + \left( \frac{R_M - R_f}{\sigma_M} \right) \sigma_P \quad (12)$$

Onde:

$R_P$  = retorno esperado do portfólio;

$R_f$  = taxa livre de risco;

$R_M$  = retorno da carteira de mercado;

$\sigma_P$  = desvio padrão do portfólio;

$\sigma_M$  = desvio padrão da carteira de mercado.

Desta forma, o equilíbrio no mercado de capitais pode ser caracterizado por dois valores-chave: em primeiro lugar, o intercepto vertical da CML, que é a taxa livre de risco, também chamado de *reward for waiting* ou *price of time*, que representa o retorno requerido por adiar um potencial consumo por um período dada a certeza do fluxo de caixa futuro (ALEXANDER; SHARPE; BAILEY, 2001; ELTON; GRUBER; BROWN, 2012). Em segundo lugar, a inclinação da CML, também chamada de *reward per risk borne* ou *price of risk*, que representa o preço de mercado do risco para todas as carteiras eficientes, ou seja, representa o retorno extra que se pode ganhar ao aumentar o risco da carteira eficiente em uma unidade (ALEXANDER; SHARPE; BAILEY, 2001; ELTON; GRUBER; BROWN, 2012).

### 2.2.2 *Security market line*

A CML representa o equilíbrio, dada a relação risco e retorno para portfólios eficientes. Ativos considerados individualmente estarão sempre abaixo da CML, visto que parte do risco de um ativo pode ser eliminado por meio da diversificação do portfólio. Como a carteira de mercado é uma carteira muito bem diversificada, não interessa ao investidor o risco específico de cada ativo, dado que essa parcela pode ser eliminada por meio da diversificação. Assim, o mercado só remunera o investidor com relação à parte do risco não diversificável que ele assume (ELTON; GRUBER; BROWN, 2012).

Com efeito, a covariância dos ativos - parte do risco que não é diversificável - é de vital importância para o investidor, pois alterações na covariância de um ativo podem causar alterações na inclinação da CML (ALEXANDER; SHARPE; BAILEY, 2001). Assim, a contribuição de cada ativo para o desvio padrão do portfólio está ligada à covariância desse ativo. Desta forma, a medida de risco relevante para um ativo é sua covariância com relação à carteira de mercado. Essa medida é denominada pela letra grega beta ( $\beta$ ), e representa a contribuição de um ativo para a variância da carteira de mercado como fração da variância total dessa carteira, descrita conforme a equação 13:

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2} \quad (13)$$

Onde:

$\sigma_{iM}$  = covariância entre ativo i e a carteira de mercado;

$\sigma_M^2$  = variância da carteira de mercado;

Pode-se representar a relação entre retorno esperado e o beta graficamente, traçando uma reta denominada de *security market line* (SML). A SML pode ser descrita conforme a equação 14, onde, dado que o beta da carteira de mercado é igual a 1, a inclinação da SML é igual ao prêmio de risco da carteira de mercado (BODIE; KANE; MARCUS, 2015). Ainda pode-se inferir que, tendo como referência o beta da carteira de mercado que é igual a 1, os ativos com beta maior que 1 serão considerados agressivos, pois apresentam uma sensibilidade acima da média do mercado, e ativos com valores de beta menor que 1 serão considerados defensivos (BODIE; KANE; MARCUS, 2015).

$$R_i = R_f + (R_M - R_f) \beta_i \quad (14)$$

Onde:

$R_i$  = retorno esperado de um ativo i;

$R_f$  = taxa livre de risco;

$R_M$  = retorno da carteira de mercado;

$\beta_i$  = beta do ativo i;

### 2.2.3 *Implicações gerais sobre o modelo CAPM*

A carteira de mercado é um assunto muito discutido nos estudos em finanças, visto que é peça central do modelo CAPM que alicerça toda a discussão sobre o relacionamento do mercado com ativos de risco em condições de equilíbrio. Ademais, a carteira de mercado serve como um padrão de performance universal em que os investidores comparam a performance de seus fundos com os retornos da carteira de mercado (ALEXANDER;

SHARPE; BAILEY, 2001).

Contudo, a forma de compor essa carteira se mostra verdadeiramente complexa na prática, o que gera grandes debates na literatura. Em suma, a carteira de mercado é composta de todos os ativos ponderados pelo seu valor de mercado, entretanto, quando se considera um cenário global, essa tarefa se torna extremamente complexa e inviável. Em função disso, Alexander, Sharpe e Bailey (2001) relatam que na prática são usadas *proxies* de mercado, sendo escolhidos índices amplos que possam representar de fato o comportamento do mercado de uma forma geral.

Entretanto, em seu célebre trabalho, Roll (1977) apresenta críticas em relação ao modelo CAPM, argumentando que dada a inviabilidade na estimação, a carteira de mercado é inobservável. Sem ser possível estimar uma carteira que contemple todos os ativos possíveis, não é possível testar o modelo CAPM, e também não é possível constatar sua validade. Além disso, ele afirma que se usadas *proxies* para a carteira de mercado, testar o modelo CAPM se torna equivalente a testar se um portfólio é eficiente em termos de média-variância. Desta forma o CAPM é tautológico, dada a premissa de os investidores tomarem suas decisões a respeito de seus portfólios com base no retorno esperado e no desvio padrão.

### 2.3 O modelo de Black-Litterman

Dados os pontos de discussão sobre o modelo de Markowitz apresentados anteriormente, Black e Litterman (1991) e Black e Litterman (1992) propuseram uma abordagem baseada na combinação entre o equilíbrio de mercado, composto por meio de um processo de otimização reversa que infere o vetor de retornos esperados implícitos de equilíbrio, e as perspectivas individuais do investidor, utilizando uma abordagem bayesiana, resultando num novo vetor que permite gerar portfólios diversificados e intuitivos (IDZOREK, 2007).

O modelo de Black-Litterman (BL) parte da premissa de que, em função de fatores de instabilidade, os mercados oscilam entre uma tendência de aproximação e de afastamento em relação a uma condição de equilíbrio. Assume-se que, mesmo com essa dinâmica, os valores não irão desviar para tão longe deste equilíbrio, o que sugere que os investidores podem combinar suas visões e previsões com as informações de preço e retorno de equilíbrio para poderem lucrar (BLACK; LITTERMAN, 1992). Caso o investidor não tenha nenhuma perspectiva que divirja da informação de equilíbrio, assume-se que a melhor opção para ele seja assumir o portfólio de equilíbrio como melhor escolha.

Diante disso, o equilíbrio de mercado fornece um ponto de referência neutro, de forma que o investidor possa expressar suas perspectivas com base nessa referência sem precisar apresentar um conjunto de retornos esperados para cada visão alternativa que ele tenha do mercado (BLACK; LITTERMAN, 1992). O investidor pode realizar inferências sobre alguns ativos do conjunto e abster-se de fazer previsões para os quais ele se sinta menos confortável. Além disso, a confiabilidade das previsões pode ser incorporada individualmente ao modelo, o que permite distinguir estimativas qualificadas de simples palpites no processo de composição do portfólio (BESSLER; OPFER; WOLFF, 2017). Os itens seguintes tratam dos conceitos centrais do modelo, sua especificação, principais aplicações e críticas.

### 2.3.1 Equilíbrio de mercado

Como visto acima, o modelo BL usa o equilíbrio como um ponto de partida neutro (IDZOREK, 2007), sendo este definido como um conjunto de retornos esperados que equilibram a demanda por títulos no mercado *hedge* (BLACK; LITTERMAN, 1991; BLACK; LITTERMAN, 1992). Dado que o mercado financeiro é complexo, caótico e fortuito por natureza, espera-se que ocorram movimentos fazendo com que os preços oscilem em torno dessa condição de equilíbrio. Esses afastamentos da condição de equilíbrio constituem momentos de oportunidade, nos quais os investidores podem auferir vantagens e ganhos (LITTERMAN, 2003).

Conforme proposto pelos seus fundadores, o modelo BL original considera o CAPM como modelo de equilíbrio em contexto doméstico, e sua extensão proposta por Black (1989) para um ambiente global. Faz-se uma ressalva de que essa escolha não indica que o mercado sempre irá se comportar de acordo com essa prescrição, mas esse é um bom ponto de partida para a análise (LITTERMAN, 2003). Mais adiante será feita uma discussão mais específica sobre essa escolha.

Posto isso, passa-se a um processo de engenharia reversa (FISHER, 1975; HEROLD, 2005; SHARPE, 1974) para se obter um vetor de retornos esperados implícitos de equilíbrio, que são um conjunto de retornos esperados que, se considerados como entrada de um otimizador, irão retornar a carteira de equilíbrio de mercado (HEROLD, 2005). O método de otimização reversa usado para obtenção do vetor de retornos esperados implícitos de equilíbrio é descrito conforme a equação 15:

$$\Pi = \lambda \Sigma w_{mkt} \quad (15)$$

Onde:

$\Pi$  = vetor de retorno esperado implícitos de equilíbrio;

$\lambda$  = coeficiente de aversão ao risco;

$\Sigma$  = matriz de covariância dos retornos (N x N);

$w_{mkt}$  = peso do valor de mercado dos ativos (N x 1).

N = número de ativos considerados.

Segundo Idzorek (2007), coeficiente de aversão ao risco ( $\lambda$ ) é a taxa pela qual o investidor irá abdicar de um maior retorno em função de um menor risco. Esse coeficiente atua na função de otimização reversa como um escalar dos retornos esperados, onde um maior lambda implica maiores retornos esperados. Em uma aplicação num cenário global, G. He e Litterman (2002) estipulam uma taxa de aversão média igual a 2,5. Entretanto, P. W. He, Grant e Fabre (2013) optaram por usar a abordagem de Best e Grauer (1991), calculando o valor de ( $\lambda$ ) em função da *proxy* do mercado, procedimento descrito na equação 16.

$$\lambda = \frac{E(r)_m - r_f}{\sigma_m^2} \quad (16)$$

Onde:

$\lambda$  = coeficiente de aversão ao risco;

$E(r)_m$  = retorno esperado do índice de mercado;

$r_f$  = taxa livre de risco;

$\sigma_m^2$  = variância do índice de mercado.

O item seguinte fará uma discussão sobre a escolha da *proxy* de mercado, apresentando alguns pressupostos e implicações.

### 2.3.2 Escolha da *proxy* de mercado

Conforme supracitado, o modelo BL tem como ponto de partida uma abordagem baseada no equilíbrio do mercado, ou seja, o princípio básico do modelo é de que o mercado tende a se manter em equilíbrio, sendo apresentados desvios em relação a essa referência em função de diversos fatores, mas que são rapidamente absorvidos, voltando o mercado sempre à sua condição de equilíbrio. A existência desses desvios possibilita ao investidor, dada a sua capacidade de análise e previsão de ocorrência desses desvios, auferir ganhos adicionais ao delinear uma estratégia coerente.

Seguindo esse conceito, Litterman (2003) assume que, dada a complexidade do sistema financeiro e as diversas variáveis envolvidas, ainda não se tem totalmente definido o conceito da melhor representação do equilíbrio de mercado. De fato, dentro do escopo proposto pelo modelo, apenas uma representação deste se faz necessária como referência, sendo o CAPM um bom ponto de partida, mas que, ainda que não seja, necessariamente, a única e melhor opção.

Os modelos de precificação de ativos financeiros, tal qual o CAPM, apresentam uma série de pressupostos que simplificam o ambiente real para possibilitar a representação e análise, mas que também afastam o modelo da realidade. Célebres autores, como Peters (1996) e Mandelbrot e Hudson (2004), levantaram críticas e observações sobre os pressupostos da hipótese de mercados eficientes, cerne do pensamento dominante dos estudos em finanças e ideia base dos modelos de precificação. Mesmo em modelos propostos recentemente, como o X-CAPM de Barberis et al. (2015) e o modelo de cinco fatores de Fama e French (2016), as limitações em termos de pressupostos ainda é algo presente, visto que essa é uma questão inerente à modelagem em si.

Destaca-se que os pressupostos citados já foram levantados na seção anterior, de forma que não cabe aqui a posição apenas de crítica, mas uma ressalva que endossa a proposição supracitada pelo autor do modelo BL, de que a escolha da *proxy* de mercado não é uma opção única, desde que feitas as devidas considerações.

Especificamente, em aplicações do modelo BL, alguns trabalhos divergem da escolha pelo CAPM como, por exemplo, Bessler, Opfer e Wolff (2017), que trabalham tendo como ponto de partida carteiras de ponderação ingênua e carteiras com ponderação baseadas em risco. Essa escolha anui com o trabalho de Ardia e Boudt (2015), que compara os retornos implícitos de mercado obtidos pelo CAPM com outros vetores derivados de outros portfólios. Esses têm um pequeno ganho na questão da previsão, apresentam-se mais estáveis e menos dispersos. Posto isso, passa-se à questão das visões do investidor, discutida na sequência.

### 2.3.3 Visões do investidor

A partir do processo de otimização reversa obtém-se o vetor de retornos esperados implícitos de equilíbrio, a partir do qual pode-se definir o conjunto de visões do investidor em relação a essa referência, gerando um vetor de retornos esperados que resultará no portfólio ótimo para o investidor (BLACK; LITTERMAN, 1991).

Destaca-se que no modelo BL as visões do investidor representam os sentimentos subjetivos do investidor em relação a mercados específicos. Caso ele não tenha uma posição sobre algum mercado ou ativo, ele não precisa apresentar uma, ou ainda, caso ele possua diferentes intensidades de percepções sobre uma visão e outra, ele precisa estar apto a expressar essas diferenças (BLACK; LITTERMAN, 1992).

Cabe aqui destacar um dos pressupostos centrais da HME, que defende investidores aversos ao risco, com um comportamento racional em um universo de média e variância. Esse pressuposto implica que os investidores tomem suas decisões sobre ativos com base em todas as informações disponíveis sobre eles, sempre num sentido de maximização de utilidade.

Entretanto, trabalhos como os de Kahneman e Tversky (1979) e Kahneman (2003) mostram que o *decision maker* toma decisões de forma distinta em situações de risco, apontando comportamentos distintos do que se esperava dentro do pressuposto de racionalidade quando se apresentam possibilidades de ganho certo ou perda certa. Além disso, também foram encontrados diversos vieses cognitivos, nos quais os investidores incorporam crenças subjetivas sobre determinado cenário futuro, o que altera o sentido da decisão, e o momento em que ela ocorre.

Acrescentam-se as discussões feitas por Simon (1971), com destaque às limitações na tomada de decisão do indivíduo dadas pela sua capacidade de cognição, que, ao impor um limite ao acesso e processamento da informação, restringe a capacidade de previsão e de decisão. Desta forma, o processo de concepção das visões, tratado especificamente de forma distante pelos autores do modelo BL, apresenta-se como uma lacuna a ser trabalhada pela literatura.

Discutidos tais pontos, volta-se à formulação do modelo, na qual as visões podem ser expressas de forma relativa, com a expectativa de que um ativo irá superar outro numa dada proporção, ou de forma absoluta, esperando que um ativo atinja determinado patamar, subindo ou caindo em relação ao seu estado de equilíbrio.

A par dessas considerações, o novo vetor que combina o vetor de retornos esperados implícitos de equilíbrio com o conjunto de visões do investidor é obtido por meio da equação 17:

$$E[R] = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1} [(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q] \quad (17)$$

Onde:

N = número de ativos considerados;

K = número de visões do investidor consideradas;

$E[R]$  = novo vetor combinado (N x 1);

$\tau$  = escalar que representa a incerteza dos retornos;

$\Sigma$  = matriz de covariância dos retornos (N x N);  
 $P$  = matriz que identifica os ativos envolvidos no conjunto visões do investidor (K x N);  
 $\Omega$  = matriz diagonal que representa a incerteza de cada visão do investidor (K x K);  
 $\Pi$  = vetor de retornos esperados implícitos de equilíbrio;  
 $Q$  = conjunto de visões do investidor (K x 1).

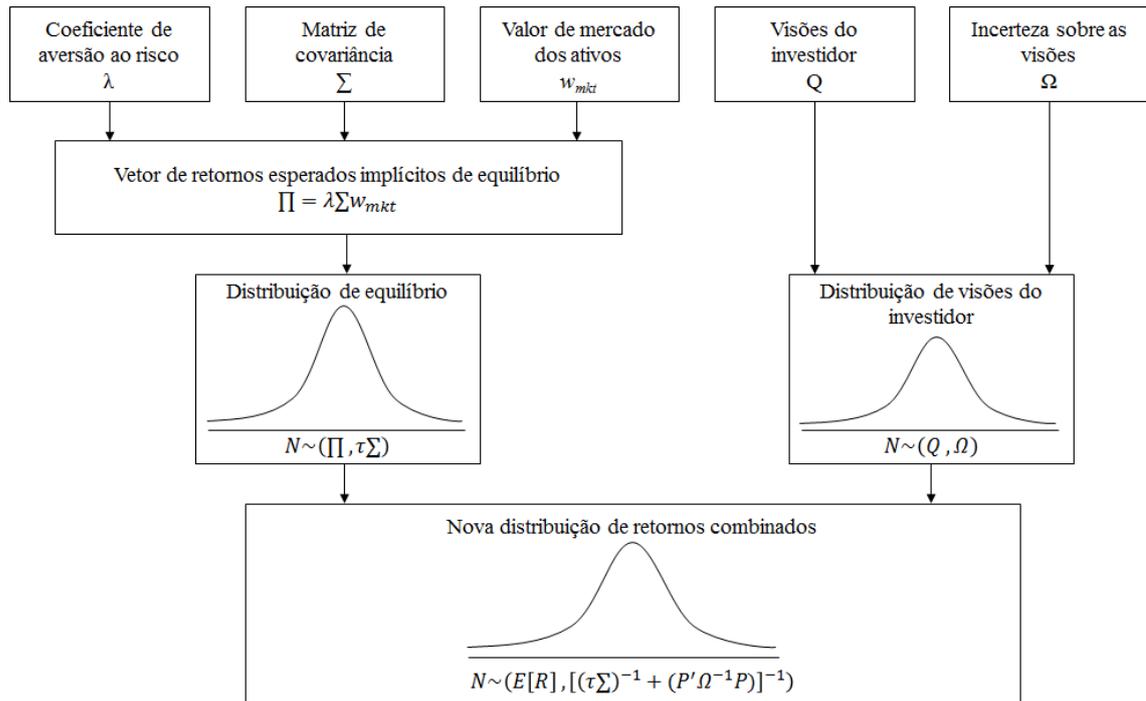
Dada essa nova distribuição de retornos, que agora representa os retornos de equilíbrio associados às visões do investidor, faz-se necessário calcular também a nova matriz de covariância, que também é alterada em função das visões do investidor. A equação 18, proposta por Satchell e Scowcroft (2007), apresenta o cálculo da nova matriz de covariância para o modelo BL:

$$\Sigma_{BL} = \Sigma + [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1} \quad (18)$$

Onde:

$\Sigma_{BL}$  = matriz de covariância do novo conjunto de retornos;  
 $\tau$  = é um escalar que representa a incerteza dos retornos;  
 $\Sigma$  = é a matriz de covariância dos retornos (N x N);  
 $P$  = matriz que identifica os ativos envolvidos no conjunto visões do investidor (K x N);  
 $\Omega$  = matriz diagonal que representa a incerteza de cada visão do investidor (K x K).

A Figura 4 apresenta um resumo esquemático do modelo BL com os pontos apresentados até então. Em suma, a distribuição a priori tem como medida de tendência central os retornos implícitos de equilíbrio e parâmetro de dispersão a variância dos dados, sendo esta associada a um fator de incerteza de estimativa dos retornos. Em um universo bayesiano, os dados observados são representados pelo conjunto de visões subjetivas do investidor, que é centrado nas visões propriamente ditas e tem parâmetro de dispersão igual à incerteza expressa sobre a realização dessas visões. A nova distribuição de retornos combinados é dada em função do novo vetor de retornos obtido pela equação 17, e uma nova variância, obtida em função da equação 18. Existem, ao longo de todo esse processo, elementos de refino e ajuste do modelo, que são os elementos tau ( $\tau$ ) e ômega ( $\Omega$ ), discutidos na seção seguinte.



**Figura 4: Representação esquemática do modelo de Black-Litterman**

Fonte: adaptado de Idzorek (2007).

### 2.3.4 Parâmetros de ajuste

O modelo BL possui ferramentas que possibilitam um ajuste fino, onde alteram-se os parâmetros  $\tau$  e  $\Omega$  de acordo com as características apresentadas pelo novo vetor de retornos combinados (IDZOREK, 2007). Esses ajustes são feitos de acordo com a experiência do gestor e permitem que sejam gerados portfólios possíveis de se investir, minimizando os efeitos de carteiras com grandes posições tomadas, tal qual ocorre em carteiras obtidas com o modelo MV (MICHAUD; ESCHA; MICHAUD, 2013).

Sobre o parâmetro  $\tau$ , ele é um escalar que diz respeito ao ajuste da incerteza em relação à estimativa dos retornos e da matriz de covariância da distribuição a priori. Para valores pequenos de  $\tau$ , a distribuição a posteriori tende a se aproximar da distribuição do equilíbrio. Para grandes valores do parâmetro, a distribuição tende a se aproximar da distribuição de visões do investidor, indicando uma grande certeza com relação a essas expectativas ou um grande receio com relação às estimativas dos retornos (BESSLER; OPFER; WOLFF, 2017; MICHAUD; ESCHA; MICHAUD, 2013).

Entre as aplicações encontradas na literatura, Satchell e Scowcroft (2007) trabalharam com o  $\tau$  igual a 1 e Firoozye e Kumar (2003) utilizaram um  $\tau$  igual a 1 dividido pelo número de pontos observados. Bessler, Opfer e Wolff (2017) testaram variações desse parâmetro e os resultados corroboraram os valores mais comumente usados na literatura, variando entre 0.025 e 1.000 (BLACK; LITTELMAN, 1992; G. HE; LITTELMAN, 2002; IDZOREK, 2007).

Quanto ao parâmetro  $\Omega$ , trata-se de uma matriz diagonal que representa a incerteza de cada visão do investidor, e é considerado o elemento mais abstrato do modelo BL (IDZOREK, 2007). Idzorek (2007) propôs um método para estimar o parâmetro, de

forma a se estipular um nível de confiança por meio de um algoritmo que trabalha cada visão em questão, chegando a uma estimativa final que dispensa o uso do parâmetro  $\tau$ . Bessler, Opfer e Wolff (2017) trabalharam com a abordagem proposta por Meucci (2010), na qual a confiabilidade de cada visão do investidor é baseada em erros históricos de estimativa. G. He e Litterman (2002) e Walters (2011) igualam a incerteza das visões do investidor aos pesos e à variância da distribuição de equilíbrio, sendo esta a abordagem mais utilizada na literatura (WALTERS, 2011). A equação 19 apresenta esse cálculo:

$$\Omega = \text{diag} [P (\tau \Sigma) P'] \quad (19)$$

Onde:

$\Omega$  = matriz diagonal que representa a incerteza de cada visão do investidor (K x K);

$\tau$  = escalar que representa a incerteza dos retornos;

$\Sigma$  = matriz de covariância dos retornos (N x N);

$P$  = matriz que identifica os ativos envolvidos no conjunto visões do investidor (K x N).

Por fim, Litterman (2003) afirma que a melhor forma de se estimar o parâmetro ômega é uma pergunta sem uma resposta universal, sendo que a melhor maneira de especificá-lo depende do contexto de análise. Percebe-se que a calibragem do modelo se dá de maneira *ad hoc*, sendo necessário que o analista seja coeso e consistente dado o contexto e as informações disponíveis. Litterman (2003, p.87) usa um exemplo que de fato se assemelha à abordagem de Bessler, Opfer e Wolff (2017) e Meucci (2010), ao argumentar que:

*When the views are the product of quantitative modeling, for example, the expected returns might be a function of historical performance, the degree of confidence might be set proportional to the amount of data supporting the view, and correlations between views might be assumed to be equal to the historical correlations between view portfolio returns.*

## 2.4 Machine learning

*Machine learning* é um campo de estudo da inteligência artificial que trata do estudo de mecanismos adaptativos que permitem que computadores aprendam por meio de exemplos, analogias e experiências, a fim de que os sistemas possam melhorar sua performance com o passar do tempo. A ideia central é o aprendizado por meio de exemplos de um conjunto particular de dados, de forma a ser extrapolado para a realização de previsões ou classificação.

Segundo Lorena e Carvalho (2007), as técnicas de aprendizado empregam um princípio de inferência denominado de indução, no qual obtém-se conclusões genéricas a partir de um conjunto particular de exemplos. Em um trabalho específico sobre redes neurais artificiais (RNA), Braga, Carvalho e Ludermir (2007, p. 12) definem aprendizado como o “processo pelo qual os parâmetros livres de uma rede neural são ajustados por meio de uma forma continuada de estímulo pelo ambiente externo, sendo o tipo específico de aprendizado definido pela maneira particular como ocorrem os ajustes dos parâmetros livres”.

Assim, existem dois tipos de aprendizado: primeiro, o aprendizado não supervisionado, em que não existe a presença de um supervisor externo e são fornecidas apenas informações de

entrada para a rede, que irá identificar recorrências e assim definir as saídas. Essas técnicas são utilizadas principalmente quando o objetivo for encontrar padrões ou tendências que auxiliem no entendimento dos dados (LORENA; CARVALHO, 2007).

O segundo tipo é o aprendizado supervisionado, que, como já diz o nome, implica a presença de um supervisor externo, responsável por estimular o aprendizado fornecendo pares de entradas e saídas, observando a saída gerada pela rede e comparando com os valores desejados. O que se almeja é a obtenção de uma função que gere imagens corretas para novos dados não apresentados anteriormente. As aplicações feitas neste trabalho irão valer-se desse segundo tipo de aprendizado.

Outro ponto de grande destaque das técnicas baseadas em *machine learning* é a capacidade de lidar com dados ruidosos. Destaca-se a robustez na presença de *outliers* durante o processo de aprendizado e obtenção da função de classificação ou regressão (LORENA; CARVALHO, 2007).

Ademais, as técnicas baseadas em *machine learning* possuem grande versatilidade e vêm sendo empregadas nos estudos em finanças devido à sua capacidade de lidar com estruturas de dados complexos, imprecisos e volumosos, bem como à não sujeição dos modelos ao grande número de pressupostos, característicos dos sistemas especialistas, que são tradicionalmente usados nesse campo (PAIVA; DUARTE et al., 2017).

As seções seguintes tratam dos conceitos centrais das redes neurais artificiais e *support vector machine*, as duas técnicas de *machine learning* aplicadas no presente estudo.

#### 2.4.1 Redes neurais artificiais

Segundo Haykin (2001), o trabalho em redes neurais artificiais (RNA) tem sido motivado pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro humano é um sistema de processamento de informação altamente complexo, responsável por funções cognitivas básicas e controle de funções motoras. Ademais, ele ainda tem capacidade de reconhecer padrões, relacioná-los, usar e armazenar conhecimentos por experiência, além de interpretar observações (BRAGA; CARVALHO; LUDERMIR, 2007).

Para realizar tais funções o cérebro humano conta com cerca de  $10^{11}$  neurônios que trabalham simultaneamente, fazendo com que funcione muito mais rápido que qualquer computador convencional (NEGNEVITSKY, 2005; HAYKIN, 2001). Essa rede neural complexa possui uma propriedade denominada de plasticidade, o que significa que na presença de um padrão de estímulos fornecidos pelo ambiente os neurônios podem criar ligações com outros neurônios ou mudar os pesos entre as conexões existentes, permitindo que a rede se adapte a esse novo contexto, sendo essa a base do processo de aprendizado no cérebro.

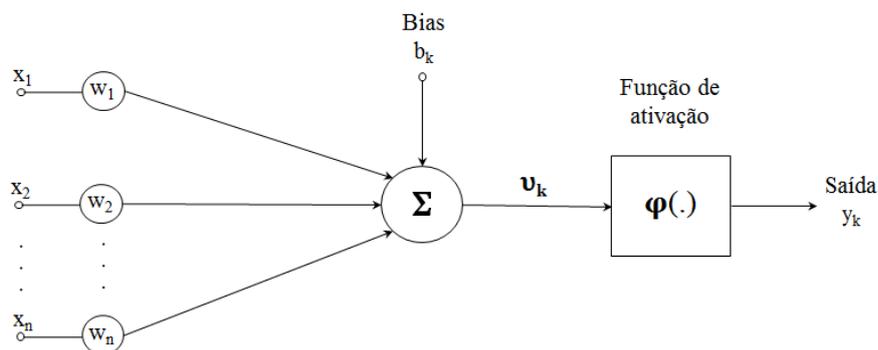
Inspirados em toda essa potencialidade, as RNA buscam reproduzir as funções das redes biológicas, a fim de implementarem seu comportamento funcional e sua dinâmica (BRAGA; CARVALHO; LUDERMIR, 2007). Desta forma, Haykin (2001, p. 28) define as redes neurais como:

Um processador maciçamente paralelamente distribuído constituído de

unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro humano em dois aspectos: (I) O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem. (II) Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Tais semelhanças conferem às RNA capacidade de generalização, considerando que a rede pode produzir saídas adequadas para entradas que não estavam presentes durante o treinamento, e adaptabilidade, pois uma RNA pode adaptar seus pesos sinápticos a modificações no ambiente, seja em tempo real ou em um novo processo de treinamento. Essas duas propriedades tornam factível o tratamento de problemas complexos e dinâmicos em grande escala (HAYKIN, 2001).

Uma RNA é composta de um conjunto de processadores simples e interconectados denominados de neurônios. Um neurônio é uma unidade básica de processamento de informação de uma RNA e uma simplificação do que se sabe sobre o funcionamento do neurônio biológico. A Figura 5 apresenta uma representação esquemática de um neurônio.



**Figura 5: Modelo de um neurônio**

Fonte: adaptado de Haykin (2001).

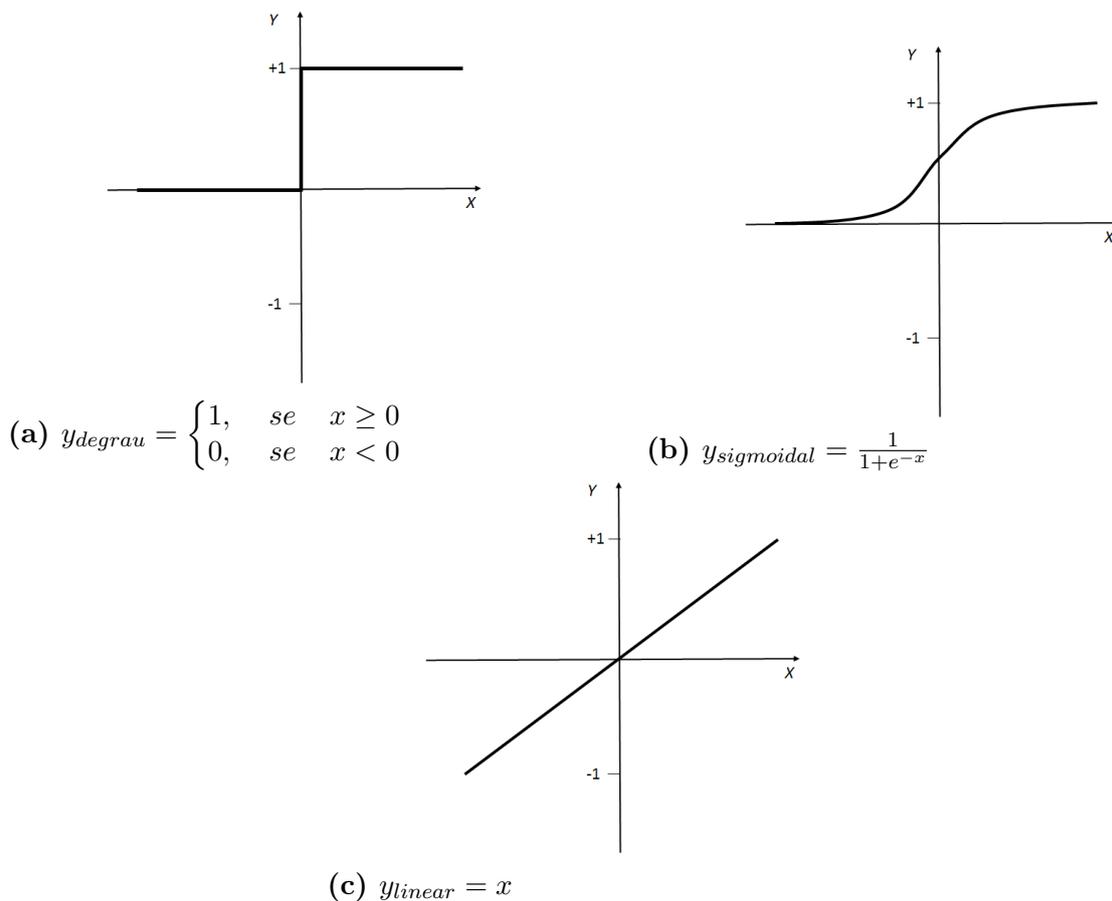
Destacam-se aqui três estruturas básicas para a composição do neurônio: inicialmente, tem-se um conjunto de sinapses, que possui pesos acoplados ( $w_n$ ), cujos valores serão multiplicados pelos valores de entrada ( $x_n$ ) em cada unidade. Em seguida tem-se um somador ( $\Sigma$ ), que irá realizar a soma das entradas ponderadas pelos pesos atribuídos. Nesta etapa existe um parâmetro denominado de bias ( $b_k$ ), que permite a calibragem do sinal de saída. E por fim, uma função de ativação  $\varphi(\cdot)$ , que restringe o intervalo permissível de amplitude do sinal de saída a um valor finito.

Haykin (2001) descreve o funcionamento de um neurônio matematicamente, conforme as equações 20 e 21:

$$v_k = \sum_{j=1}^n w_j x_j \quad (20)$$

$$y_k = \varphi(v_k + b_k) \quad (21)$$

Sobre a função de ativação supracitada, ela é responsável por gerar a saída ( $y_k$ ) do neurônio em termos do campo local induzido ( $\nu_k$ ). A literatura vigente apresenta alguns exemplos básicos, sendo que escolheu-se apresentar os exemplos comuns entre as principais referências usadas no presente trabalho (BRAGA; CARVALHO; LUDERMIR, 2007; HAYKIN, 2001; NEGNEVITSKY, 2005), representados na Figura 6.



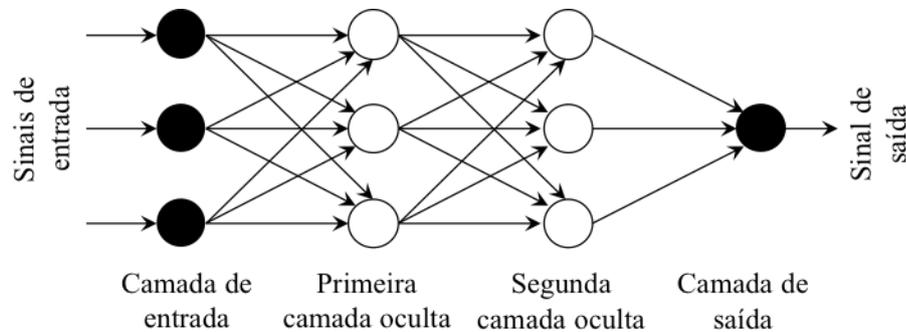
**Figura 6: Funções de ativação**

Fonte: adaptado de Negnevitsky (2005).

Braga, Carvalho e Ludermir (2007) ressaltam que independente da função de ativação escolhida, os neurônios individuais possuem capacidade de computação limitada, sendo que o grande poder computacional das RNA se deve à sua estrutura maciça e paralelamente distribuída. Assim, considerando um conjunto de neurônios, é vital definir qual tipo de arquitetura é o mais adequado. Essa escolha é dada em função de características do problema, tais como complexidade, representatividade dos dados, características dinâmicas ou estáticas e conhecimento a priori sobre o problema.

Em função da natureza do problema, o presente trabalho adota o uso de redes *perceptron* de múltiplas camadas (MLP), que é uma rede composta de uma camada de entrada, uma ou mais camadas ocultas, e uma camada de saída, onde a informação é passada adiante camada por camada (HAYKIN, 2001; NEGNEVITSKY, 2005). Desta forma, a resposta observada na camada mais externa da rede corresponde à composição das

respostas dos neurônios das camadas anteriores (BRAGA; CARVALHO; LUDERMIR, 2007). Negnevitsky (2005) afirma que, com uma camada oculta, uma rede MLP pode representar qualquer função contínua, e com mais de uma camada oculta, qualquer função pode ser representada. A Figura 7 apresenta de forma esquemática um exemplo de uma rede MLP.



**Figura 7: Representação esquemática de uma rede MLP**

Fonte: adaptado de Negnevitsky (2005).

Haykin (2001) destaca que as redes MLP têm três características distintivas: (I) o modelo de cada neurônio da rede inclui uma função de ativação não linear, sendo que a não linearidade é suave (diferenciável em qualquer ponto); (II) as redes possuem uma ou mais camadas ocultas de neurônios que não são parte da entrada ou saída da rede; (III) as redes possuem alto grau de conectividade, determinado pelas sinapses da rede, o que faz com que uma modificação em uma conectividade de rede demande uma mudança na população das conexões sinápticas ou de seus pesos.

A combinação dessas características confere às redes MLP versatilidade e potencial de aplicação, qualidades que vêm acompanhadas de algumas deficiências e de mais complexidade para a aplicação. As características de não linearidade e alta conectividade da rede tornam difícil a sua análise teórica, e o uso de neurônios em camadas ocultas torna mais difícil o processo de aprendizado da rede (BRAGA; CARVALHO; LUDERMIR, 2007; HAYKIN, 2001; NEGNEVITSKY, 2005).

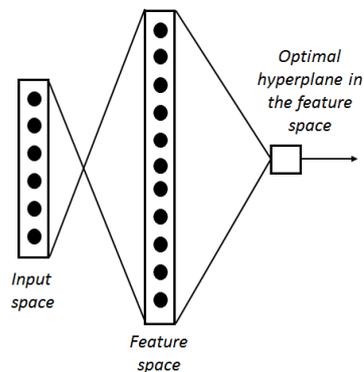
#### 2.4.2 *Support vector machine*

Seguindo os conceitos da Teoria de Aprendizado Estatístico, surge a técnica de *Support vector machine* (SVM), tendo como expoente os trabalhos de Vladimir N. Vapnik. Desde a sua proposição na década de 1990, a SVM aparece como uma técnica de destaque devido à sua potencialidade para aplicação em diversos campos de estudo.

A SVM faz uma aplicação da Teoria de Aprendizado Estatístico para obtenção de classificadores com boa generalização, ou seja, estruturas com capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu, podendo assim resolver problemas de reconhecimento de padrões, estimativa de regressão, séries temporais e de classificação, propriamente ditos (PAIVA; DUARTE et al., 2017; VAPNIK, 2013; LORENA; CARVALHO, 2007).

Vapnik (2013) elucida que a ideia básica da SVM é a criação de um hiperplano ótimo de separação, gerado em um ambiente em que mapeou-se os vetores de entrada em um

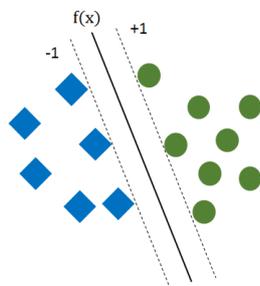
espaço de grande dimensão, usando um elemento de mapeamento não linear. A Figura 22 ilustra a ideia.



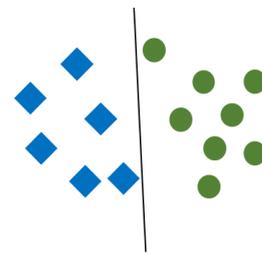
**Figura 8: Hiperplano ótimo de separação**

Fonte: adaptado de Vapnik (2013).

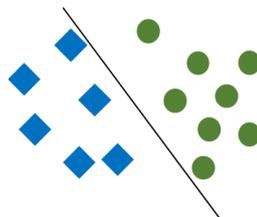
Quanto à questão da separação, dado que o conjunto dos dados pode ser linearmente separável, conforme exemplo ilustrativo na Figura 9, o algoritmo busca a maximização das margens de separação, auferindo ao classificador maior robustez e capacidade de generalização (MOHRI; ROSTAMIZADEH; TALWALKAR, 2012). Observa-se que as Figuras 9b e 9c apresentam possíveis separações entre os dados, mas sem maximização do espaço de fronteira entre os dados. Já na Figura 9a, o espaço de fronteira entre os dados está sendo maximizado, com destaque para as linhas  $+1$  e  $-1$ , que são tangentes ao conjunto de dados e são chamadas de vetores de suporte.



(a) separação otimizada



(b) separação ingênua



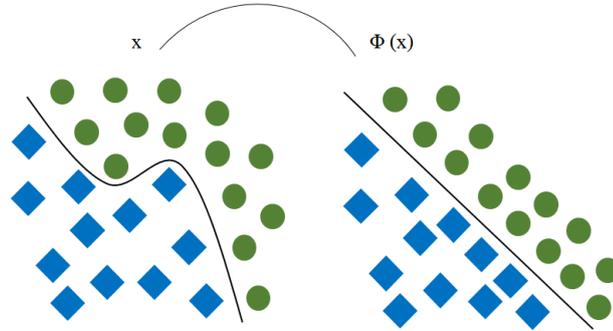
(c) separação ingênua

**Figura 9: Dados linearmente separáveis**

Fonte: adaptado de Mohri, Rostamizadeh e Talwalkar (2012).

Para casos onde não é possível a separação dos dados por meio de uma função linear,

como acontece na maioria dos casos de aplicações em situações reais, utiliza-se de um procedimento de mapeamento dos dados no *input space*, valendo-se de uma função específica, fazendo com os que dados passem a ser linearmente separáveis no *feature space*. A Figura 10 elucidada o conceito.



**Figura 10: Transformação para dados linearmente não separáveis**

Fonte: adaptado de Cristianini e Scholkopf (2002).

Dito isto, a formulação apresentada segue Ding, Song e Zen (2008), L. Ni, Z. Ni e Gao (2011) e Luo e Chen (2013), em que, supondo um conjunto de pontos para treinamento,  $D = \{x_i, y_i\}_{i=1}^N$ , onde o vetor de *inputs*  $x_i = (x_i^1, \dots, x_i^n) \in R^n$  e os *outputs*  $y_i \in \{0, 1\}$ , são satisfeitos pela condição descrita na equação 22, onde  $w$  representa o vetor peso e  $b$  a variável *bias*:

$$y_i (w^T \phi(x_i) + b \geq 1) \quad \forall i = 1, \dots, n \quad (22)$$

A função não linear  $\phi(\cdot) : R^n \rightarrow R^{nk}$  mapeia os inputs no espaço em alta dimensão. Assim, analisando a equação acima, pode-se perceber que há várias soluções para separar os pontos de treinamento. De uma perspectiva de generalização, é melhor escolher dois hiperplanos limitadores em lados opostos de um hiperplano de separação  $w^T \phi(x) + b = 0$  com maior margem  $2 / (\|w\|^2)$ . Entretanto, como vários problemas de classificação não são linearmente separáveis, faz-se necessária a introdução de uma variável de folga para permitir a classificação errada. Desta forma, considerando  $C$  como o parâmetro de erro, o problema de otimização pode ser descrito nas equações 23 a 25:

$$\min_{w, b, \xi} \left( \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \right) \quad (23)$$

Sujeito a:

$$y_i (w^T \phi(x) + b \geq 1 - \xi_i) \quad \forall i = 1, \dots, N \quad (24)$$

$$\xi_i \geq 0 \quad (25)$$

A solução do problema primário é obtida após a construção Lagrangiana. Em seguida, o problema primal pode ser convertido no QP-problema descrito nas equações 26 a 28,

sendo  $\alpha_i$  um multiplicador Lagrangiano,  $Q_{ij} = y_i y_j \phi(x_i)^T \phi(x_j)$ :

$$\max_{\alpha} \left( \alpha^T \alpha - \frac{1}{2} \alpha^T Q \alpha \right) \quad (26)$$

Sujeito a:

$$0 \leq \alpha_i \leq C, i = 1, \dots, N \quad (27)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (28)$$

Em razão do esforço computacional exigido, o produto interno é substituído pela função *kernel*, que satisfaz a condição de Mercer,  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ . Assim, chega-se à função de decisão não linear no espaço primal para o caso linearmente não separável, conforme a equação 29:

$$y(x) = \text{sgn} \left( \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) \quad (29)$$

Destacam-se as funções *kernel* linear ( $K(x_i, x_j) = x_i^T x_j$ ), polinomial ( $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$ ), radial ( $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ ) e sigmoid ( $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$ ).

Feitas essas considerações, passa-se à seção seguinte, que trata das questões metodológicas do trabalho.

### 3 METODOLOGIA

Escolheu-se uma abordagem quantitativa para delineamento da pesquisa, na qual se utiliza de dados secundários para se estudar a aplicação do modelo proposto. Conforme Gil (2008), a presente pesquisa pode ser classificada quanto aos fins como descritiva, visto que busca explicar o relacionamento entre variáveis de determinada população. Quanto aos meios, a pesquisa se caracteriza como *ex-post-facto*, tipo de pesquisa sistemática e empírica na qual, dada a prévia ocorrência e manifestação das variáveis independentes, o pesquisador não tem controle direto sobre essas variações (GIL, 2008).

#### 3.1 Conjunto amostral

A pesquisa foi desenvolvida a partir de ativos do mercado acionário brasileiro, sendo a amostra composta de ativos do índice Ibovespa. Foram coletados dados de volume (Vol), preço de abertura (*Open*), preço máximo (*High*), preço mínimo (*Minimum*) e preço de fechamento (*Close*), sendo essas informações insumos para cálculo dos indicadores usados no desenvolvimento da pesquisa.

A amostra contempla dados de Janeiro/2006 a Julho/2017. Nesse período, 137 ativos compuseram o índice, sendo o máximo de 73 ativos e o mínimo de 54 ativos presentes concomitantemente no Ibovespa. A listagem com os ativos que compuseram a amostra consta no apêndice A e a evolução da quantidade de ativos na composição do índice pode ser vista na Figura 11.



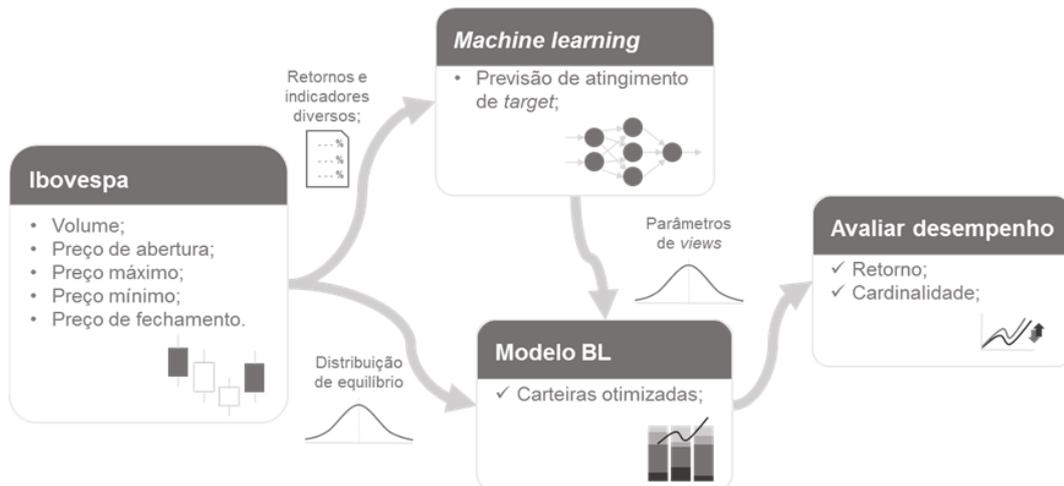
**Figura 11: Evolução da composição do índice Ibovespa**

Fonte: dados da pesquisa (2018).

Os dados de cotação diária utilizados foram ajustados para dividendos, bonificações, *split* e *insplit* e foram coletados por meio da base de dados *Bloomberg*. Posteriormente, foram tratados e manipulados a partir dos softwares *Microsoft Excel* e *Phyton*.

### 3.2 Desenho do experimento

A Figura 12 apresenta de forma esquemática o desenho do experimento, que foi composto das etapas de (i) tratamento dos dados e cálculo dos indicadores, (ii) classificação dos ativos com potencial para atingimento do *target*, (iii) otimização das carteiras usando o modelo BL e (iv) comparação com as *baselines* propostas.



**Figura 12: Desenho do experimento**

Fonte: dados da pesquisa (2018).

#### 3.2.1 Tratamento dos dados e cálculo dos indicadores

Na primeira etapa foi feita a coleta e tratamento dos dados, e, de posse do conjunto amostral descrito no item anterior, foram calculados os atributos utilizados para modelagem, descritos no Quadro 1. Foram calculados 22 indicadores, escolhidos de acordo com Paiva, Duarte et al. (2017), Paiva (2014) e Paiva, Cardoso et al. (2011), e montados em três grupos distintos que compuseram a instância de entradas do processo multiparâmetro.

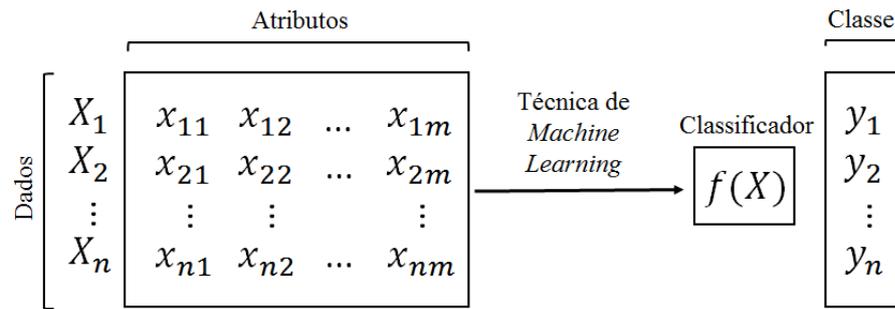
Quadro 1: Atributos para modelagem

n	Atributo	n	Atributo
1	$r_1 = \ln \left( \frac{close_i}{close_{i-1}} \right)$	12	$r_{12} = \ln \left( \frac{low_i}{open_i} \right)$
2	$r_2 = \ln \left( \frac{close_{i-1}}{close_{i-2}} \right)$	13	$r_{13} = \ln \left( \frac{low_{i-1}}{open_{i-1}} \right)$
3	$r_2 = \ln \left( \frac{close_{i-2}}{close_{i-3}} \right)$	14	$r_{14} = \ln \left( \frac{low_{i-2}}{open_{i-2}} \right)$
4	$r_4 = \ln \left( \frac{close_{i-3}}{close_{i-4}} \right)$	15	$r_{15} = \ln \left( \frac{low_{i-3}}{open_{i-3}} \right)$
5	$r_5 = \ln \left( \frac{high_i}{open_i} \right)$	16	Momentum (close, per. = 10)
6	$r_6 = \ln \left( \frac{high_i}{open_{i-1}} \right)$	17	RSI (close, per. = 14)
7	$r_7 = \ln \left( \frac{high_i}{open_{i-2}} \right)$	18	SAR Parab.(high, low, $\alpha=0$ , PE=0)
8	$r_8 = \ln \left( \frac{high_i}{open_{i-3}} \right)$	19	ATR (high, low, close, per.=14)
9	$r_9 = \ln \left( \frac{high_{i-1}}{open_{i-1}} \right)$	20	True range (high, low, close)
10	$r_{10} = \ln \left( \frac{high_{i-2}}{open_{i-2}} \right)$	21	Chaikin A/D (high, low, close, vol)
11	$r_{11} = \ln \left( \frac{high_{i-3}}{open_{i-3}} \right)$	22	On balance volume (close, vol)

### 3.2.2 Classificação dos ativos com potencial para atingimento do *target*

O passo seguinte consistiu em classificar os ativos com potencial de atingimento de determinados *targets* valendo-se das técnicas de *machine learning*. Para tal, foram escolhidas as técnicas de RNA e SVM, nas quais, em ambos os casos, os modelos trabalharam como classificadores. Neste tipo de modelagem, informa-se um conjunto de entradas de diversos ativos ao modelo para que ele retorne com uma matriz binária indicando quais ativos irão atingir determinado resultado. A Figura 13 elucida o conceito.

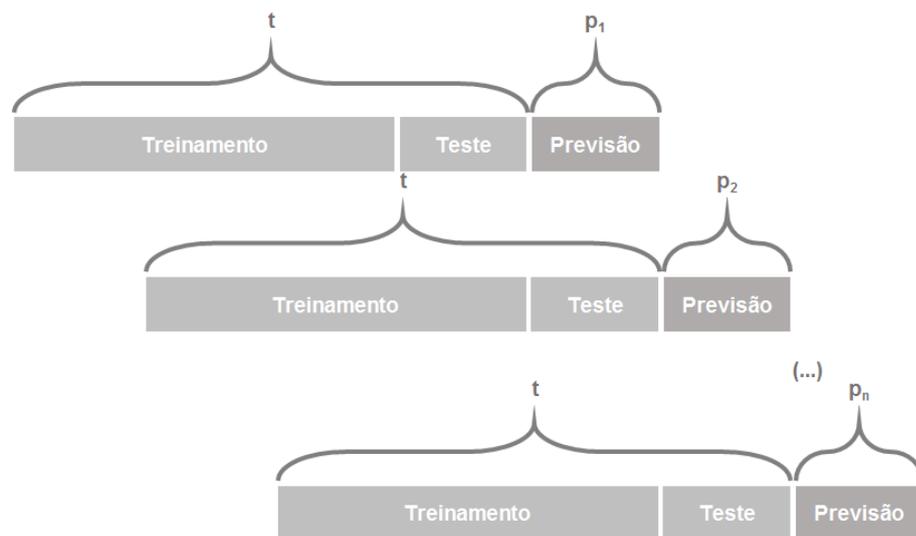
No caso da aplicação feita no trabalho, os dados (representados na Figura 13 por  $X_1$  à  $X_n$ ) correspondem aos indicadores de um ativo em um janela que varia durante o tempo, os atributos são os indicadores propriamente ditos, e as classes (representadas na Figura 13 por  $y_1$  à  $y_n$ ) são um vetor binário com o indicativo de atingimento do *target* esperado.



**Figura 13: Classificador baseado em *machine Learning***

Fonte: adaptado de Lorena e Carvalho (2007).

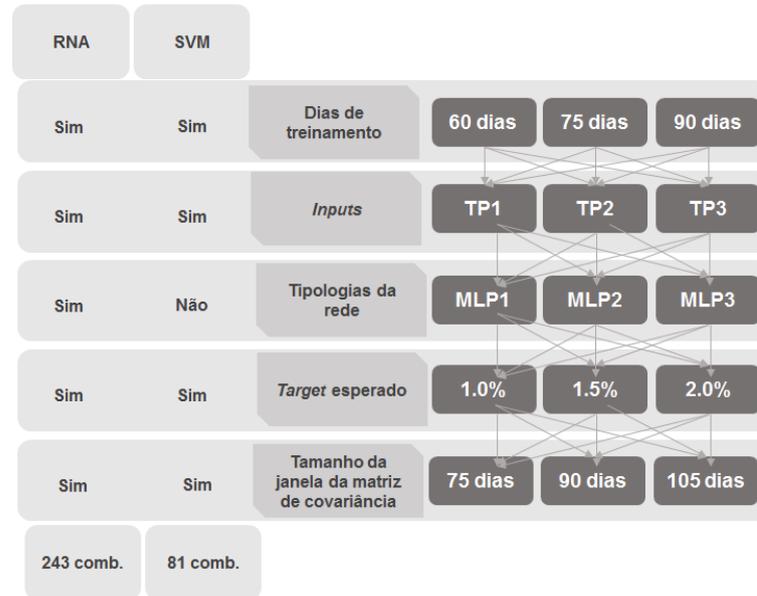
Ambos os modelos trabalharam considerando um esquema de janelas deslizantes para treinamento, teste e previsão, representado na Figura 14. Nesse esquema, as informações de um intervalo de tempo passado de tamanho  $t$  serão utilizadas para se obter previsões num intervalo de tempo futuro de tamanho  $p$ . Para as previsões no próximo intervalo de tempo futuro, descartam-se os  $p$  pontos mais antigos do intervalo de tempo passado e acrescenta-se  $p$  novos pontos. Desta forma, serão feitas previsões para um novo intervalo futuro de tamanho  $p$ , considerando-se sempre o intervalo de tempo passado de tamanho  $t$  mais recente.



**Figura 14: Janelas de aprendizado e simulação da RNA**

Fonte: dados da pesquisa (2018).

A cada simulação que foi realizada, os ativos que compuseram a amostra foram escolhidos de acordo com o índice Ibovespa na data em que foi feita a previsão. Para a escolha da parametrização como um todo, optou-se por trabalhar com uma abordagem de multiparâmetros, sendo considerado um período *in-sample*, no qual a escolha pela configuração a ser usada na operação do dia seguinte era feita pela configuração de maior retorno do dia anterior, critério adotado para toda a amostra. A Figura 15 apresenta de forma esquemática a proposição de parametrização para o caso do modelo de RNA.



**Figura 15: Combinações de parâmetros para modelo de RNA**

Fonte: dados da pesquisa (2018).

No caso em questão, tem-se três tamanhos da janela de treinamento, três conjuntos de entradas, três tipologias de rede, três *targets* e três tamanhos de janela de cálculo da matriz de covariância, resultando assim em 243 combinações diferentes. A lógica é a mesma para o modelo de SVM, onde não foi aplicável apenas a questão da tipologia, resultando em 81 combinações diferentes.

Destaca-se alguns detalhes da parametrização do modelo de RNA, onde foram utilizadas redes MLP, sendo testados três arranjos com duas camadas ocultas, com [12, 6], [12, 8] e [10, 8] neurônios em cada camada. Foi utilizada a função de ativação linear e taxa aprendizado igual a 0,001. Para a SVM, foi utilizada a função *kernel* radial, com parâmetro de erro  $C = 1$  e  $\gamma = 1/n_{features}$ .

### 3.2.3 Otimização das carteiras usando o modelo BL

Na terceira etapa da pesquisa foi feita a otimização das carteiras utilizando o modelo BL. Foi considerada como distribuição de equilíbrio do mercado a mínima variância. Na aplicação do modelo MV, foi utilizado para cálculo da matriz de covariância o esquema de janelas deslizantes, sendo testadas diferentes variações no tamanho das janelas.

Justifica-se a escolha desta carteira pelo trabalho supracitado de Ardia e Boudt (2015), que atestou o uso de retornos implícitos derivados de carteiras baseadas no binômio retorno-risco como *proxy* de mercado. Em uma aplicação do modelo BL, Bessler, Opfer e Wolff (2017) usaram esta, dentre outras carteiras, como base para cálculo do vetor de retornos implícitos, reforçando assim a escolha feita no presente trabalho.

Além da coerência com os trabalhos e contribuições supracitados, sugere-se a suposição do amplo conhecimento do modelo MV e sua aplicação pelos praticantes de mercado, sendo a escolha do ponto de menor variância tida como uma posição conservadora e que demanda um baixo custo computacional.

No que tange à questão das visões do investidor, elas foram obtidas por meio dos modelos de RNA e SVM, detalhados acima. Para acoplamento, foi fornecido ao modelo BL uma matriz com a diagonal composta dos sinais do classificador como matriz  $P$ , e um vetor com o *target* em questão como vetor  $Q$ . O  $\Omega$  foi calculado por meio do *precision* histórico do ativo no momento em questão, sendo que esse cálculo também seguiu a lógica das janelas deslizantes.

Os demais parâmetros do modelo seguiram as seguintes definições: o  $\lambda$  foi calculado conforme a equação 16, onde foi confirmado, dentre diversos testes, o valor sugerido pela literatura, ficando fixo em 2,5, e o  $\tau$  foi definido como 0,025.

### 3.2.4 Comparação com as *baselines* propostas

Foram construídos mais quatro modelos que serviram como comparação para as estratégias que usam o modelo BL. A estrutura e as variações estão expostas a seguir.

#### 3.2.4.1 Naive+BL

Esta estratégia utiliza os sinais indicativos de ativos que tem previsão de atingir determinado *target* por meio de um processo randômico, e em seguida esse sinais são acoplados ao modelo BL. A ideia aqui é gerar um linha de comparação para a eficiência dos preditores, eliminando o efeito dos sinais ao serem acoplados ao modelo BL.

Para esse modelo tem-se as três instâncias de *targets* e três de tamanho da janela para cálculo da matriz de covariância, resultando assim em nove combinações diferentes. Desta forma, para cada *target*, foram gerados 31 listas de sinais aleatórios que contemplavam todos os ativos e dias da amostra em questão.

De posse destes dados, foram feitas simulações de operação combinando com as variações do tamanho de janela de cálculo da matriz de covariância. Por fim, para cada combinação de *target* e tamanho de janela de cálculo da matriz de covariância foi selecionada a curva mediana do retorno acumulado, ficando assim com as nove combinações para a operação *in-sample* e *out-of-sample*.

#### 3.2.4.2 RNA+1/N e SVM+1/N

Essa estratégia utiliza os sinais indicativos de ativos que têm previsão de atingir determinado *target* gerados pelos modelos de RNA e SVM, mas sem posterior otimização, sendo os pesos dos ativos escolhidos para compor a carteira ponderados de maneira igualitária. A ideia aqui é gerar um linha de comparação para a eficiência da otimização.

Para esse modelo, no caso da aplicação com RNA, tem-se as três instâncias de tamanho da janela de treinamento, três conjuntos de entradas, três tipologias de rede e três *targets*, resultando assim em 81 combinações diferentes. No caso da aplicação com SVM, tem-se três instâncias de tamanho da janela de treinamento, três conjuntos de entradas e três *targets*, que resultam em 27 combinações diferentes.

### 3.2.4.3 RNA+MV e SVM+MV

Esta estratégia utiliza os sinais indicativos de ativos que têm previsão de atingir determinado *target* gerados pelos modelos de RNA e SVM, com posterior otimização feita por meio do modelo MV. A ideia aqui é gerar um linha de comparação direta do modelo BL *versus* o modelo MV.

Para esse modelo, no caso da aplicação com RNA, têm-se as três instâncias de tamanho da janela de treinamento, três conjuntos de entradas, três tipologias de rede, três *targets* e três tamanhos de janela de cálculo da matriz de covariância, resultando assim em 243 combinações diferentes. No caso da aplicação com SVM, têm-se três instâncias de tamanho da janela de treinamento, três conjuntos de entradas, três *targets* e três tamanhos de janela de cálculo da matriz de covariância, que resultam em 81 combinações diferentes.

### 3.2.4.4 MV

Essa estratégia é uma aplicação simples do modelo MV, escolhendo o ponto de menor variância da curva. A ideia aqui é gerar um linha de comparação com menor complexidade e custos computacionais que os demais modelos. Para esse modelo têm-se as três instâncias de *targets* e três tamanhos da janela para cálculo da matriz de covariância, resultando em nove combinações diferentes.

### 3.2.5 Modelo de operação

Sobre a estratégia de operação adotada, todas as carteiras consideraram periodicidade diária, sendo a compra feita com base no preço de abertura do dia seguinte. Para a venda, quando o ativo atingiu o *target* esperado para aquele dia, a venda foi feita por esse valor, caso contrário, o ativo foi vendido pelo preço de abertura do dia subsequente. Para os casos em que não foram gerados sinais de atingimento de *target* no dia, escolheu-se adotar uma postura conservadora e não fazer nenhuma aplicação com o capital no período em questão.

As simulações não consideram custos de transação, tais como tributos ou emolumentos. Para isso foi realizada uma análise específica, que considera um custo padrão medido em bps, no qual 1 bps representa o valor de 0,01%. Esse valor foi vinculado ao número de ativos envolvidos na operação, sendo o produto destes subtraído da rentabilidade final da carteira em cada ponto da curva. Foram considerados quatro cenários: 1,00 bps, 0,50 bps, 0,10 bps e 0,05 bps.

Dito isso, passa-se à próxima seção, que trata da apresentação e discussão dos resultados.

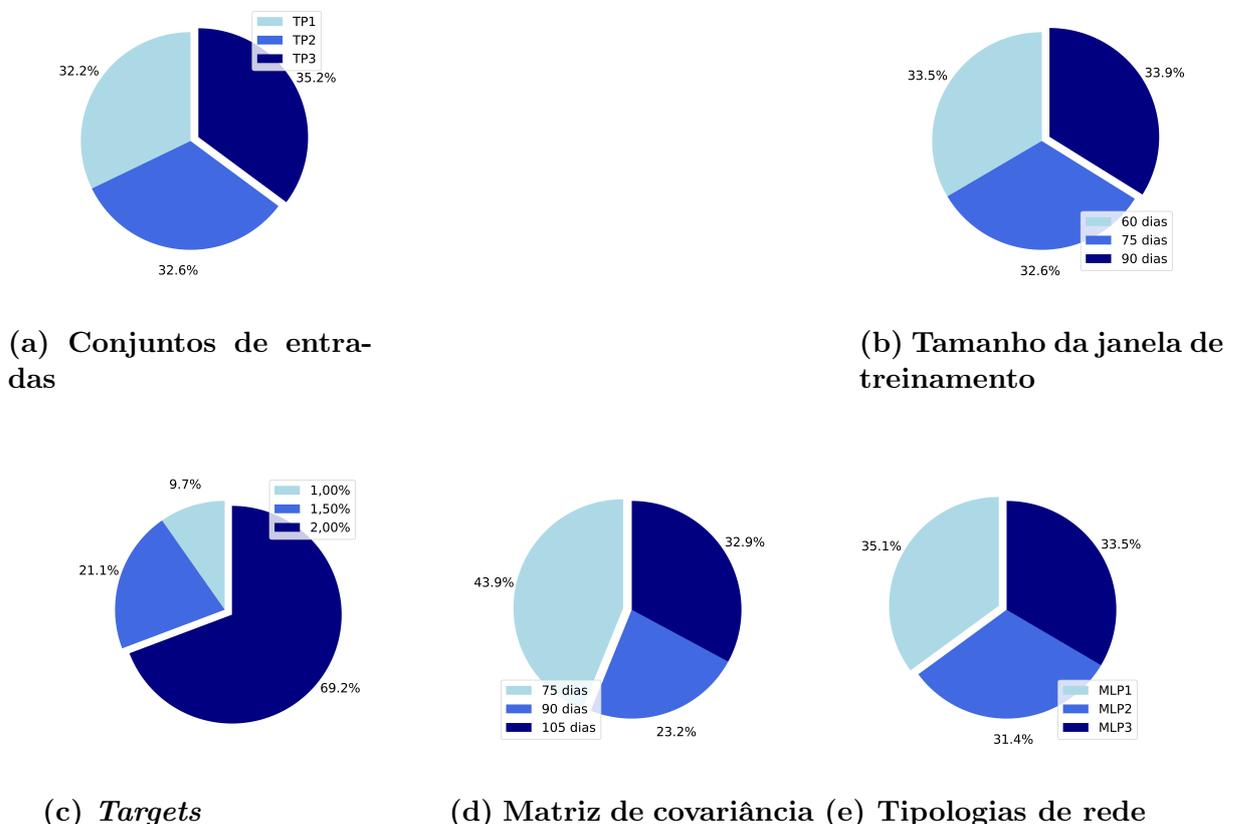
## 4 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados os resultados dos experimentos propostos na pesquisa. Na primeira etapa são apresentados os resultados dos modelos que usam a técnica de RNA, sendo avaliado o desempenho do classificador e a performance das carteiras. Em seguida os resultados são apresentados para o uso da técnica de SVM, valendo-se dos mesmos critérios. Por fim, a melhor estratégia de cada variação das técnicas de *machine learning* foi comparada com as *baselines* propostas.

### 4.1 Resultados das redes neurais artificiais

Uma vez coletadas e calculadas as variáveis que nortearam a pesquisa, iniciou-se a construção de um modelo segundo as técnicas de *machine learning*, baseado na técnica de RNA para classificação de ativos financeiros com potencial para alcançar o *target* de ganho delineado. Desta forma, passa-se a discutir o primeiro objetivo específico proposto.

Conforme descrito no capítulo dos métodos, foram testadas diversas combinações de parâmetros e feita a escolha diária do melhor conjunto para a operação, respeitando o período *in-sample*. A Figura 16 apresenta a frequência de escolha dos parâmetros para o período *out-of-sample*. Observa-se uma uniformidade quanto à escolha dos parâmetros, exceto pelos parâmetros de *target* (Figura 16c) e tamanho da janela da matriz de covariância (Figura 16d), onde predominaram o *target* de 2,0% e a janela de 75 dias, com proporções de 69,3% e 43,9%, respectivamente.



**Figura 16: RNA - Resultado da seleção de parâmetros**

Fonte: dados da pesquisa (2018).

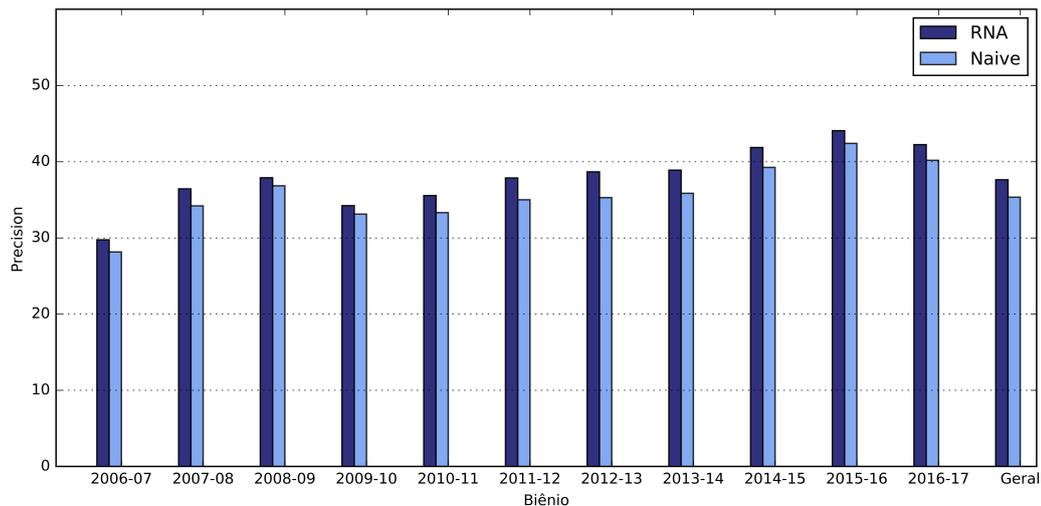
Com relação ao desempenho do classificador, apresenta-se inicialmente a medida de *precision*<sup>5</sup>, conforme Tabela 1. Para melhor entendimento da Tabela 1, para as colunas sob o rótulo RNA, denota-se que a coluna denominada de "Geral" apresenta o *precision* de todos os períodos ponderado pelos *targets* selecionados, e as colunas com os *targets* específicos apresentam a medida para os dias em que aquele *target* foi escolhido, ou seja, a coluna de 1,0% tem os dados referentes à parcela de 9,7% dos dias totais de operação (vide Figura 16c), e assim por diante. A base de comparação é a abordagem ingênua, representada pelas colunas sob o rótulo *Naive*, segundo o qual pressupõem-se que em todos os dias os *targets* serão atingidos, e em seguida verifica-se quais os dias em que de fato houve êxito.

**Tabela 1: RNA - Desempenho do classificador (%)**

Biênio	RNA				Naive			
	Geral	1,0%	1,5%	2,0%	Geral	1,0%	1,5%	2,0%
2006-2007	29,7	42,6	33,5	25,1	28,2	41,5	32,9	24,8
2007-2008	36,4	50,8	37,8	33,6	34,2	50,3	38,0	31,5
2008-2009	37,9	58,4	38,1	35,3	36,8	59,3	39,8	34,3
2009-2010	34,2	54,7	36,4	28,8	33,1	55,1	38,5	28,7
2010-2011	35,5	52,6	37,9	27,8	33,3	53,3	38,3	27,6
2011-2012	37,9	54,8	41,1	30,9	35,0	54,5	41,2	29,6
2012-2013	38,7	57,8	42,5	32,2	35,3	56,6	41,9	30,3
2013-2014	38,9	58,5	40,6	32,8	35,9	58,0	40,3	31,3
2014-2015	41,9	58,0	44,4	36,8	39,2	57,8	44,2	35,0
2015-2016	44,1	59,1	47,2	39,4	42,4	59,5	47,4	38,3
2016-2017	42,2	57,9	44,9	37,5	40,2	59,5	45,3	36,0
<b>Média</b>	<b>37,6</b>	<b>54,2</b>	<b>40,1</b>	<b>32,7</b>	<b>35,3</b>	<b>54,0</b>	<b>40,3</b>	<b>31,2</b>
<b>DP</b>	<b>4,0</b>	<b>4,9</b>	<b>4,1</b>	<b>4,4</b>	<b>3,9</b>	<b>5,3</b>	<b>4,0</b>	<b>4,0</b>
<b>Pearson</b>	<b>10,7</b>	<b>9,1</b>	<b>10,2</b>	<b>13,4</b>	<b>11</b>	<b>9,9</b>	<b>9,9</b>	<b>12,8</b>

Pode-se observar que o classificador com base em RNA não conseguiu apresentar um desempenho expressamente melhor em relação à abordagem ingênua, ficando 6,5% maior no período total. Na análise por *targets*, observa-se que nos *targets* de 1,0% e 1,5% a variação é próxima de zero. Nessas parcelas da amostra, sob o prisma bienal, alterna-se entre ganhos e perdas com relação à *baseline*, o que evidencia o fraco desempenho do classificador. A pequena diferença é evidenciada também na Figura 17. Foi realizado o teste de Mann-Whitney para comparação da medida de tendência central da séries, no qual foram encontradas evidências da similaridade dos desempenhos, uma vez que a hipótese nula do teste foi aceita ( $p\text{-value} = 0,178$ ).

<sup>5</sup> $Precision = \frac{tp}{tp+tn}$ , sendo que *tp* e *tn* indicam *true positive* e *true negative*. *Positive* e *negative* se referem à previsão, e *true* e *false* se referem ao sentido desta em relação ao que de fato ocorreu.



**Figura 17: RNA - Precision do classificador por biênio**

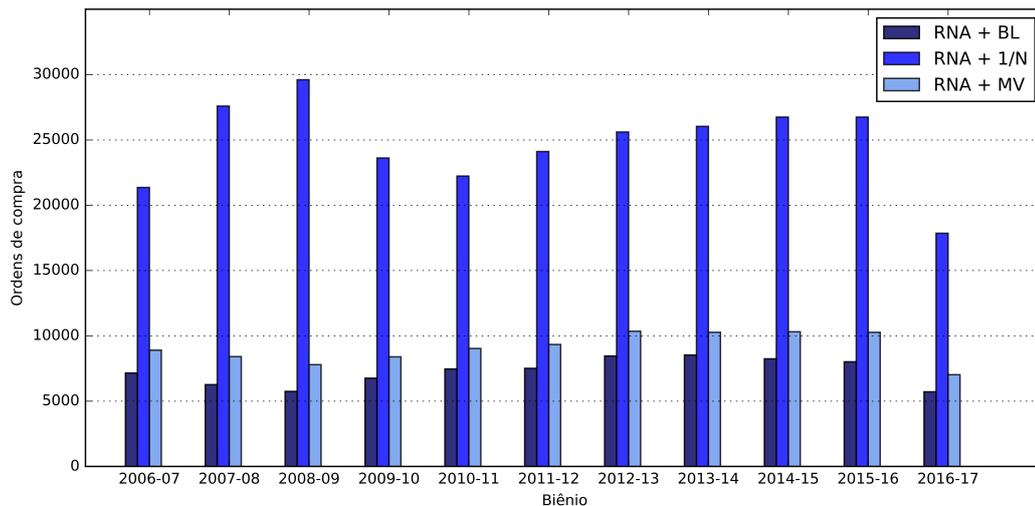
Fonte: dados da pesquisa (2018).

A Tabela 2 resume o número de *trades* (operações de compra e venda) para cada modelo. O modelo RNA+BL apresenta o menor número de operações entre as estratégias indicadas, com um resultado 70,2% menor quando comparado à estratégia RNA+1/N. Destaca-se a importância desse resultado, visto que um número menor de transações é indicativo de menores custos com corretagem.

**Tabela 2: RNA - Trades (compra e venda) por modelo**

	RNA+BL	RNA+1/N	RNA+MV
Total Trades	42.758	143.390	53.398
Média / dia	16	52	20
Desvio padrão	5,08	19,36	5,21
Mínimo / dia	1	4	3
Máximo / dia	80	120	44

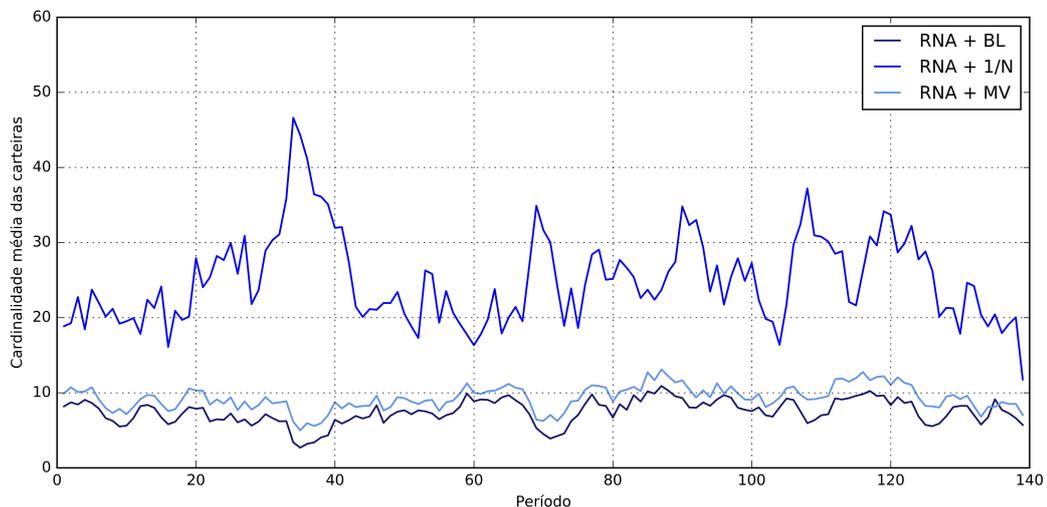
A Figura 18 apresenta uma análise complementar sobre o volume do número de operações dos modelos a partir de um horizonte bienal. Os resultados corroboram a análise geral feita anteriormente, no qual o modelo RNA+BL apresenta um menor número de operações ao longo de todo o período quando comparado aos demais modelos. Fica evidente a diferença entre o volume de ordens dos modelos RNA+BL e RNA+MV e o modelo RNA+1/N (cerca de 1/3 do valor em todos os períodos).



**Figura 18: RNA - Ordens de compra e venda por biênio**

Fonte: dados da pesquisa (2018).

A Figura 19 apresenta a cardinalidade média mensal dos modelos. Os resultados são congruentes com aqueles apontados anteriormente, sendo apenas detalhados para um espectro temporal menor. Pode-se observar que a estratégia RNA+BL apresentou menores valores ao longo de todo o período de análise. A estratégia RNA+1/N foi o modelo com maior volatilidade em função da variação na quantidade de sinais gerados pelo classificador em cada período.



**Figura 19: RNA - Cardinalidade média dos modelos**

Fonte: dados da pesquisa (2018).

Na sequência foram realizados testes estatísticos para se verificar a igualdade da medida de tendência central dos modelos. Como premissa inicial, foi averiguada a condição de normalidade das séries, que não foi confirmada. Assim sendo, procedeu-se o teste de Kruskal-Wallis, onde encontrou-se evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,000$ ). Desta forma, seguiu-se com o teste de Duun para comparações par a par.

Os valores podem ser vistos na Tabela 3, que mostra uma diferença significativa entre os três modelos propostos.

**Tabela 3: RNA - Teste de Dunn para cardinalidade**

	RNA+BL	RNA+1/N	RNA+MV
RNA+BL	-	-	-
RNA+1/N	0,000	-	-
RNA+MV	0,000	0,000	-

No âmbito do retorno da operação, a Tabela 4 apresenta a estatística descritiva básica da série de retornos sem custos de transação. Observa-se que a carteira RNA+MV apresenta a maior média mensal dentre as estratégias, um valor muito próximo ao da estratégia RNA + BL, tendo ambas apresentado um valor mínimo expressivo (-35,26% e -38,04%).

**Tabela 4: RNA - Retorno médio mensal**

	RNA+BL	RNA+1/N	RNA+MV
Média %	0,81	-0,25	0,85
Desvio padrão %	5,69	5,02	5,74
Pearson	7,07	19,71	6,77
Mínimo %	-38,04	-11,66	-35,26
Máximo %	12,55	14,25	22,38

A Figura 20 apresenta o retorno acumulado para cada um dos modelos. Pode-se observar um comportamento muito semelhante dos modelos RNA+BL e RNA+MV, onde as duas séries alternam entre períodos no *ranking* de maior retorno acumulado, mas sem nenhum descolamento entre as séries, com um resultado final próximo (138,47% e 153,74%, respectivamente).



**Figura 20: RNA - Retorno acumulado sem custos de transação**

Fonte: dados da pesquisa (2018).

Na sequência foram realizados testes estatísticos para se verificar a igualdade da medida de tendência central dos modelos. Como premissa inicial, seguiu-se com a averiguação dos pressupostos de normalidade, e constatou-se que não foram atingidos. Em vista disso, procedeu-se com o teste de Kruskal-Wallis, onde não encontrou-se evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,068$ ).

A Tabela 5 traz uma análise complementar para as séries de retorno. O modelo RNA+BL apresenta a maior proporção de dias com retorno positivo (57,04%), bem como a maior média de retorno em operações dessa natureza (0,83%), contudo, são valores muito próximos dos demais modelos. O modelo RNA+MV apresentou a melhor relação perda média por ganho média, em que para cada operação de perda são necessárias 1,19 operações de ganho para compensar o valor.

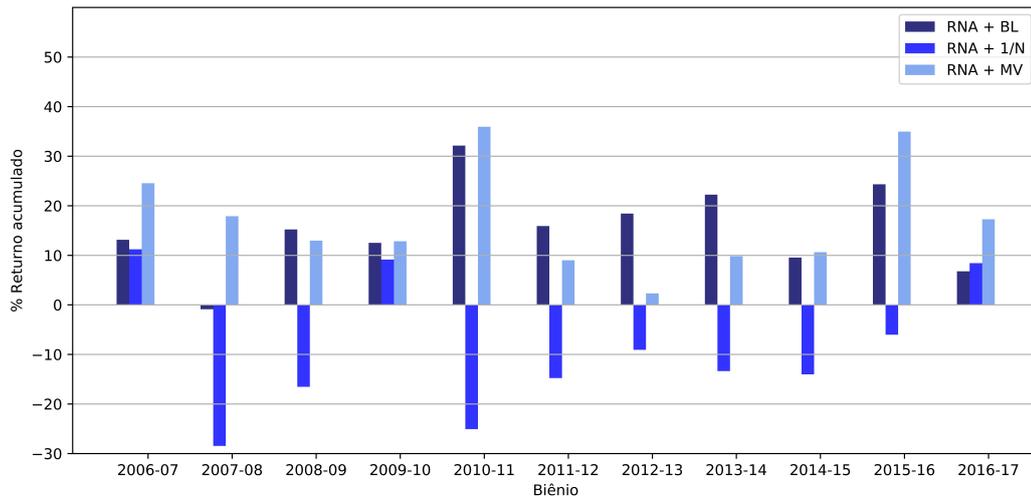
**Tabela 5: RNA - Estatísticas adicionais**

Descrição	RNA+BL	RNA+1/N	RNA+MV
(1) % dias de retorno > 0	57,04	56,79	56,76
(2) % dias de retorno < 0	42,96	43,21	43,24
(3) retorno médio em dias de ganho %	0,83	0,75	0,79
(4) retorno médio em dias de perda %	-1,00	-1,02	-0,94
Ratio (1) / (2)	1,33	1,31	1,31
Ratio (4) / (3)	1,21	1,35	1,19

Para um melhor detalhamento das séries de retorno, a Tabela 6 e a Figura 21 apresentam uma análise bienal por modelo. O modelo RNA+1/N apresenta retorno acumulado menor que pelo menos um dos outros dois modelos em todos os biênios em análise, com retorno negativo em oito dos onze períodos em análise. Os modelos RNA+BL e RNA+MV alternam entre melhor performance comparativa, confirmando as análises feitas anteriormente.

**Tabela 6: RNA - Retorno acumulado por biênio (%)**

Biênio	RNA+BL	RNA+1/N	RNA+MV
2006-2007	13,13	11,20	24,52
2007-2008	-0,89	-28,46	17,88
2008-2009	15,22	-16,54	12,94
2009-2010	12,50	9,14	12,82
2010-2011	32,12	-25,07	35,92
2011-2012	15,89	-14,77	8,99
2012-2013	18,41	-9,08	2,32
2013-2014	22,22	-13,36	9,81
2014-2015	9,54	-14,03	10,63
2015-2016	24,31	-6,02	34,93
2016-2017	6,77	8,42	17,27
<b>Acumulado</b>	<b>138,47</b>	<b>-41,07</b>	<b>153,74</b>



**Figura 21: RNA - Retorno bial por modelo**

Fonte: dados da pesquisa (2018).

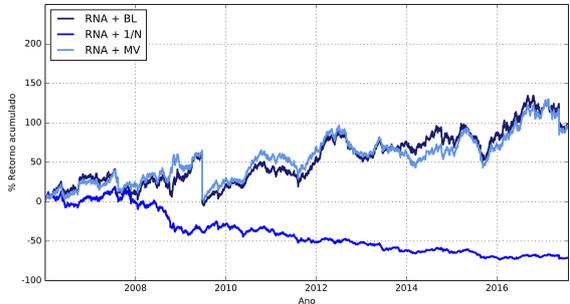
Quanto ao risco de cada operação, a Tabela 7 apresenta uma análise bial da volatilidade por estratégia. Observa-se que a estratégia RNA+BL exibe maior volatilidade que as demais estratégias na análise geral e nos quatro primeiros biênios. Para os demais períodos ela fica à frente da estratégia RNA+1/N. Foram feitos testes estatísticos para comparação da medida de tendência central da séries, sendo que não foram encontradas evidências para rejeitar a hipótese nula para as três comparações.

**Tabela 7: RNA - Volatilidade por biênio (%)**

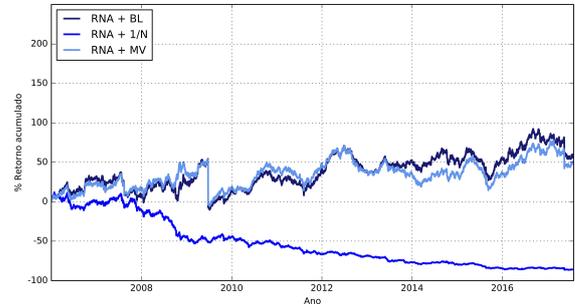
Biênio	RNA+BL	RNA+1/N	RNA+MV
2006-2007	29,44	25,74	27,89
2007-2008	34,42	30,17	31,00
2008-2009	51,21	30,30	47,42
2009-2010	47,25	26,04	44,53
2010-2011	22,91	24,38	22,25
2011-2012	23,30	23,77	22,06
2012-2013	21,18	23,44	20,19
2013-2014	22,17	24,50	21,66
2014-2015	25,04	25,20	24,30
2015-2016	25,12	25,61	23,91
2016-2017	27,55	28,38	26,43
<b>Média</b>	<b>31,32</b>	<b>26,26</b>	<b>29,58</b>
<b>DP</b>	<b>10,27</b>	<b>2,42</b>	<b>9,28</b>
<b>Pearson</b>	<b>32,8</b>	<b>9,22</b>	<b>31,37</b>

Por fim, fez-se uma análise considerando os custos de transação. A Figura 22 apresenta um estudo considerando custo de transação de 0,05 bps, 0,10 bps, 0,50 bps e 1,00 bps por ordem de compra ou venda. Para os valores de 0,05 bps e 0,10 bps, Figuras 22a e 22b, respectivamente, observa-se que os modelos RNA+BL e RNA+MV continuam apresentando um comportamento bem parecido, e mesmo com a redução da rentabilidade

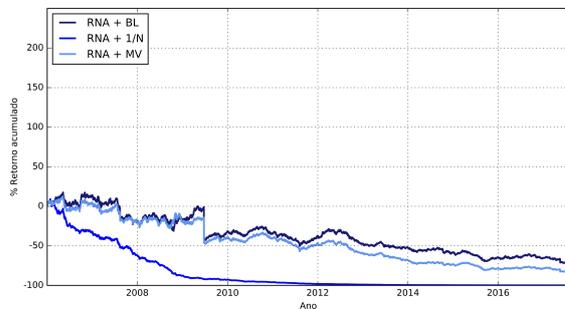
total esperada, eles ainda chegam ao final do período com um saldo final positivo. Quando aumenta-se os valores para 0,50 bps e 1,00 bps, Figuras 22c e 22d, respectivamente, todas as estratégias passam a apresentar saldo negativo, o que inviabiliza a operação.



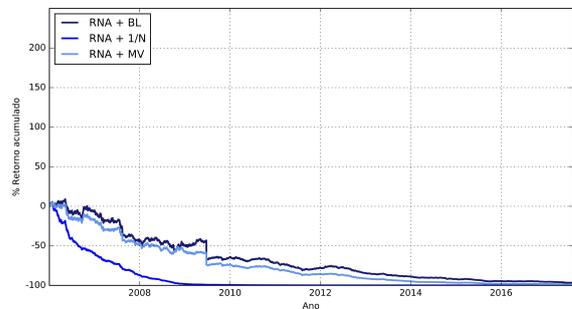
(a) 0,05 bps



(b) 0,10 bps



(c) 0,50 bps



(d) 1,00 bps

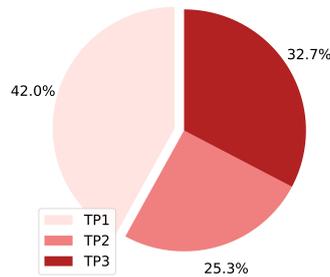
**Figura 22: RNA - Retorno acumulado considerando custos de transação**

Fonte: dados da pesquisa (2018).

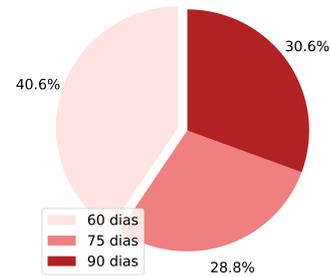
## 4.2 Resultados da *support vector machine*

Dentro do escopo proposto, passa-se à análise do segundo modelo, que valeu-se da técnica de SVM para classificar ativos com potencial o *target* de ganho delineado, conforme o segundo objetivo específico proposto. Nesta seção serão feitas as mesmas análises e testes da seção anterior, a fim de se comparar a aplicação das duas técnicas sob o mesmo espectro.

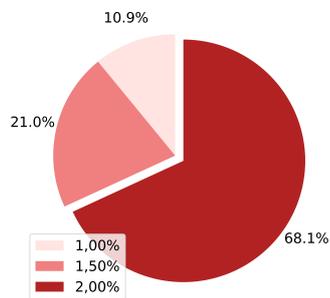
A Figura 23 apresenta a frequência de escolha dos parâmetros para o período *out-of-sample*. Novamente, observa-se um equilíbrio geral na escolha dos parâmetros, exceto pelos parâmetros de *target* (Figura 23c) e tamanho da janela da matriz de covariância (Figura 23d), onde predominaram o *target* de 2,0% e a janela de 75 dias, com proporções de 68,1% e 58,0%, respectivamente.



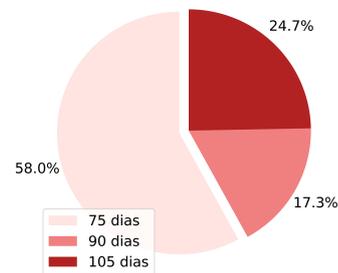
(a) Conjuntos de entradas



(b) Tamanho da janela de treinamento



(c) Targets



(d) Matriz de covariância

**Figura 23: SVM - Resultado da seleção de parâmetros**

Fonte: dados da pesquisa (2018).

Com relação ao desempenho do classificador, o resultado de medida *precision* é apresentado na Tabela 8 e na Figura 24. Para as colunas sob o rótulo SVM, denota-se que a coluna denominada de "Geral" apresenta o *precision* de todos os períodos da análise, e as colunas com os *targets* específicos apresentam a medida para os dias em que aquele *target* foi escolhido, ou seja, a coluna de 1,0% tem os dados referentes à parcela de 10,9% dos dias totais de operação (vide Figura 23, e assim por diante. A base de comparação é a abordagem ingênua, representada pelas colunas sob o rótulo *Naive*, onde pressupõem-se que em todos os dias os *targets* esperados seriam alcançados. Em seguida verificou-se quais os dias de fato em que houve êxito.

Tabela 8: SVM - Desempenho do classificador (%)

Biênio	SVM				Naive			
	Geral	1,0%	1,5%	2,0%	Geral	1,0%	1,5%	2,0%
2006-2007	42,3	41,2	39,8	43,1	27,9	35,4	29,2	26,7
2007-2008	54,2	60,1	55,1	53,1	35,0	44,1	35,8	33,3
2008-2009	54,7	62,5	54,7	53,6	38,0	48,9	39,1	35,9
2009-2010	46,3	55,0	46,4	45,1	33,2	45,7	35,3	30,8
2010-2011	40,8	50,8	37,7	40,2	33,0	45,3	34,4	30,2
2011-2012	46,6	55,8	44,5	46,1	34,9	46,9	36,9	32,2
2012-2013	49,7	55,4	51,7	48,3	35,7	48,6	39,2	32,4
2013-2014	48,2	50,1	51,0	47,2	36,4	47,8	39,9	33,3
2014-2015	48,3	59,0	47,3	47,3	39,2	49,9	40,7	37,2
2015-2016	50,6	60,6	53,5	48,5	42,4	53,0	45,2	40,0
2016-2017	49,3	56,4	52,3	47,1	41,0	50,5	43,0	38,5
<b>Média</b>	<b>49,0</b>	<b>56,3</b>	<b>49,2</b>	<b>47,8</b>	<b>35,6</b>	<b>46,8</b>	<b>37,4</b>	<b>33,3</b>
<b>DP</b>	<b>4,3</b>	<b>6,0</b>	<b>5,9</b>	<b>3,9</b>	<b>4,0</b>	<b>4,6</b>	<b>4,4</b>	<b>3,9</b>
<b>Pearson</b>	<b>8,7</b>	<b>10,7</b>	<b>12</b>	<b>8,1</b>	<b>11,4</b>	<b>9,8</b>	<b>11,8</b>	<b>11,8</b>

Ao contrário da aplicação com RNA, observa-se que o classificador com base na técnica SVM conseguiu apresentar um melhor resultado que se destaca em relação à abordagem ingênua, ficando 37,6% maior no período total. Na análise por *targets*, percebe-se que o *target* de 2,0% foi o que apresentou melhor desempenho do classificador, 43,6% melhor que o *naive*. Seguindo uma análise por biênio, constata-se uma performance superior em todas as comparações com relação à *baseline*. Essa performance é vista também na Figura 24. Foi realizado o teste de Mann-Whitney para comparação da medida de tendência central da séries, cujas evidências para rejeição da hipótese nula atestam a eficiência do classificador ( $p\text{-value} = 0,000$ ).

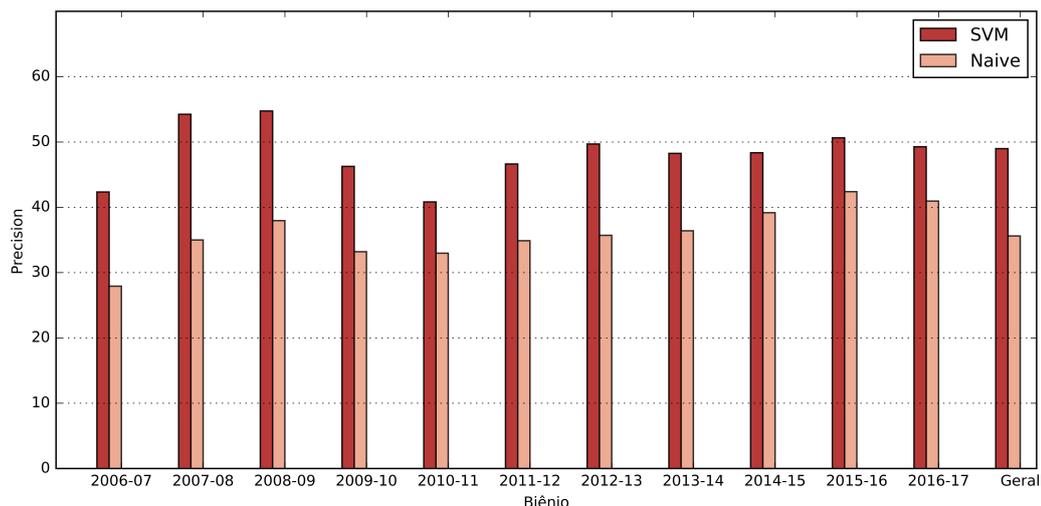


Figura 24: SVM - Precision do classificador por biênio

Fonte: dados da pesquisa (2018).

A Tabela 9 resume o número de *trades* (operações de compra e venda) para cada modelo. O modelo SVM+BL apresenta o menor número de operações entre as estratégias

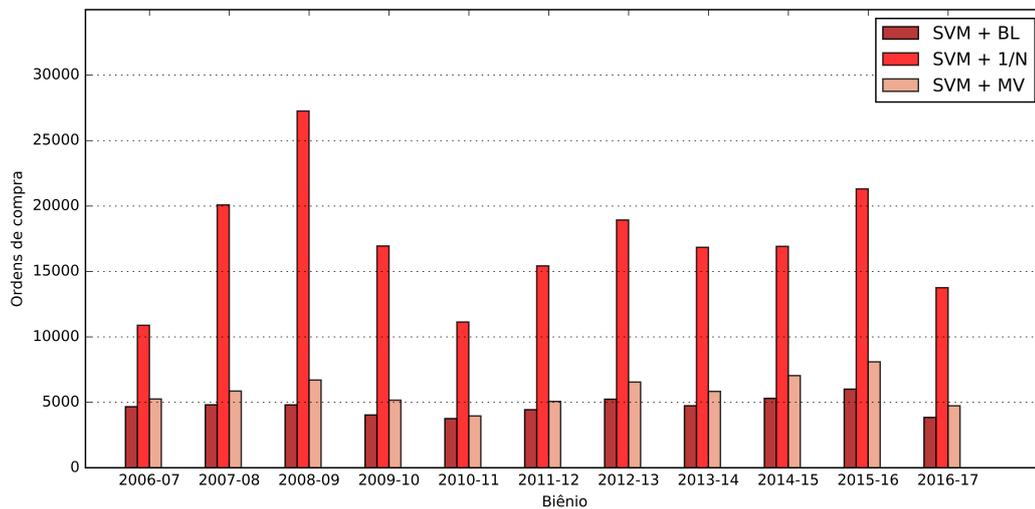
indicadas, com um resultado 72,1% menor quando comparado à estratégia SVM+1/N.

**Tabela 9: SVM - Trades (compra e venda) por modelo**

	SVM+BL	SVM+1/N	SVM+MV
Total Trades	27.596	98.876	34.244
Média / dia	10	36	12
Desvio padrão	5,91	36,3	6,95
Mínimo / dia	0	0	0
Máximo / dia	38	140	42

Em uma comparação entre os modelos SVM+1/N e RNA+1/N, em que pode ser avaliada a quantidade de sinais gerados pelos classificadores, pode-se observar que o modelo com base em SVM gera uma quantidade média menor de sinais (18 sinais em média *versus* 26 sinais em média, que resultaram em 36 ordens do modelo que utiliza a técnica de SVM, incluindo compra e venda, contra 52 ordens do modelo que utiliza a técnica de RNA) e com mais assertividade, conforme visto na análise do *precision*. Novamente, frisa-se a importância desse resultado, visto que um número menor de transações é indicativo de menores custos de corretagem.

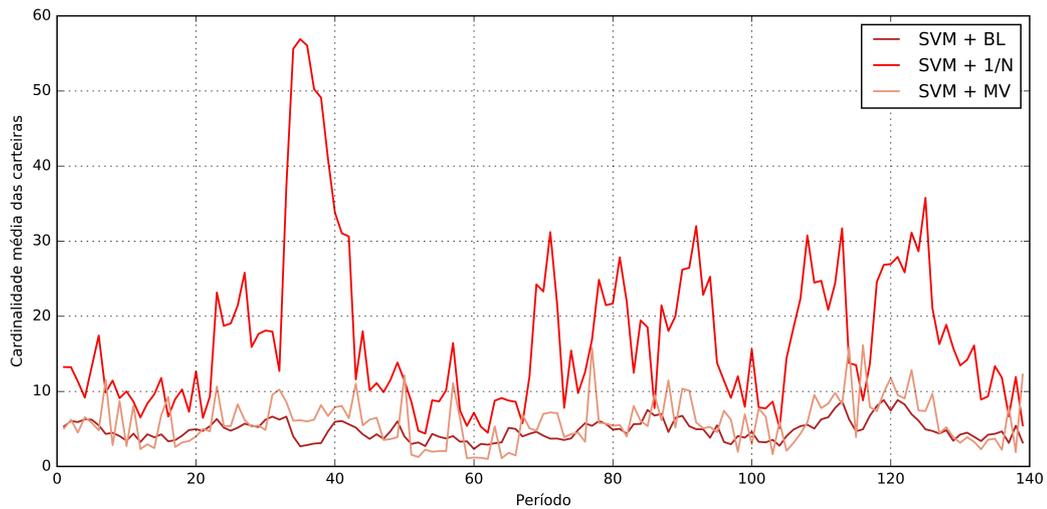
A Figura 25 apresenta uma análise complementar tratando do volume de operações dos modelos em um prisma bienal. Os resultados corroboram a análise geral feita anteriormente, na qual o modelo SVM+BL apresenta um menor número de operações quando comparado aos demais modelos, no geral e em todos os biênios analisados.



**Figura 25: SVM - Ordens de compra e venda por biênio**

Fonte: dados da pesquisa (2018).

A Figura 26 apresenta uma análise sobre o volume do número de operações dos modelos a partir de um prisma bienal. Essa análise reforça os resultados apresentados acima, onde observa-se a estratégia SVM+BL com menores valores durante todo o período de análise. A estratégia SVM+1/N possui uma série com maior volatilidade em função da variação na quantidade de sinais gerados pelo classificador em cada período.



**Figura 26: SVM - Cardinalidade média dos modelos**

Fonte: dados da pesquisa (2018).

Em seguida, foram realizados testes estatísticos para verificar a igualdade da medida de tendência central dos modelos. Como premissa inicial, foi averiguado o pressuposto de normalidade, sendo rejeitado. Em vista disso, procedeu-se com o teste de Kruskal-Wallis, onde encontrou-se evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,000$ ). Desta forma, seguiu-se com o teste de Duun para comparações par a par. Os valores podem ser vistos na Tabela 10, que mostra uma diferença significativa entre os três modelos propostos.

**Tabela 10: SVM - Teste de Dunn para cardinalidade**

	SVM+BL	SVM+1/N	SVM+MV
SVM+BL	-	-	-
SVM+1/N	0,000	-	-
SVM+MV	0,007	0,000	-

No âmbito do retorno médio mensal de cada modelo, a Tabela 11 apresenta a estatística descritiva das séries de retornos sem custos de transação. Observa-se que a carteira SVM+BL apresenta a maior média mensal dentre as estratégias e o menor coeficiente de variação de Pearson, o que implica maior retorno médio mensal e melhor relação proporcional entre variação e retorno médio.

**Tabela 11: SVM - Retorno médio mensal**

	SVM+BL	SVM+1/N	SVM+MV
Média %	2,03	0,33	1,84
Desvio padrão %	6,75	6,67	6,33
Pearson	3,32	20,24	3,44
Mínimo %	-16,14	-14,88	-13,23
Máximo %	22,82	14,54	21,27

A Figura 27 apresenta o retorno acumulado para cada modelo. Pode-se observar um comportamento muito semelhante dos modelos SVM+BL e SVM+MV, mas com evidente superioridade do modelo SVM+BL, fechando com valores finais de 1.113,00% e 868,33%, respectivamente.



**Figura 27: SVM - Retorno acumulado sem custos de transação**

Fonte: dados da pesquisa (2018).

Na sequência, realizou-se testes estatísticos para se verificar a igualdade da medida de tendência central dos modelos. Como premissa inicial, seguiu-se com a averiguação do pressuposto de normalidade, que não foi confirmado. Em vista disso, procedeu-se com o teste de Kruskal-Wallis, onde encontrou-se evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,004$ ). Desta forma, seguiu-se com o teste de Duun para comparações par a par. A Tabela 12 apresenta os valores do teste, no qual não foi encontrada diferença significativa entre os modelos SVM+BL e SVM+MV, mas foi encontrada diferença desses dois modelos quando comparados ao modelo SVM+1/N.

**Tabela 12: SVM - Teste de Dunn para retorno mensal**

	SVM+BL	SVM+1/N	SVM+MV
SVM+BL	-	-	-
SVM+1/N	0,001	-	-
SVM+MV	0,296	0,029	-

A Tabela 13 traz uma análise complementar para as séries de retorno. O modelo SVM+BL apresenta a maior proporção de dias com retorno positivo (59,97%), bem como a maior média de retorno em operações dessa natureza (1,08%), contudo, são valores muito próximos dos demais modelos. O modelo SVM+MV apresentou a melhor relação perda média por ganho média, em que para cada operação de perda são necessárias 1,25 operações de ganho para compensar o valor.

Tabela 13: SVM - Estatísticas adicionais

Descrição	SVM+BL	SVM+1/N	SVM+MV
(1) % dias de retorno > 0	59,97	58,78	59,52
(2) % dias de retorno < 0	40,03	41,22	40,48
(3) retorno médio em dias de ganho %	1,08	0,93	1,03
(4) retorno médio em dias de perda %	-1,37	-1,29	-1,29
Ratio (1) / (2)	1,50	1,43	1,47
Ratio (4) / (3)	1,27	1,38	1,25

Para melhor detalhamento das séries de retorno, a Tabela 14 e a Figura 28 apresentam uma análise por biênio dos modelos. O modelo SVM+BL apresenta retorno acumulado maior que os outros dois modelos em todos os biênios em análise, exceto pelo biênio 2016-17. Destaca-se que os modelos SVM+BL e SVM+MV não apresentam retorno negativo em nenhum dos onze períodos em análise.

Tabela 14: SVM - Retorno acumulado por biênio (%)

Biênio	SVM+BL	SVM+1/N	SVM+MV
2006-2007	47,11	13,18	41,98
2007-2008	28,62	-10,06	28,09
2008-2009	99,73	-14,82	69,76
2009-2010	87,96	14,70	60,31
2010-2011	19,32	-1,88	14,20
2011-2012	80,80	42,36	83,33
2012-2013	91,82	27,25	85,47
2013-2014	47,28	-32,33	30,64
2014-2015	41,63	-5,17	29,44
2015-2016	56,11	33,38	51,94
2016-2017	27,35	1,95	46,54
<b>Acumulado</b>	<b>1113,00</b>	<b>16,38</b>	<b>868,33</b>

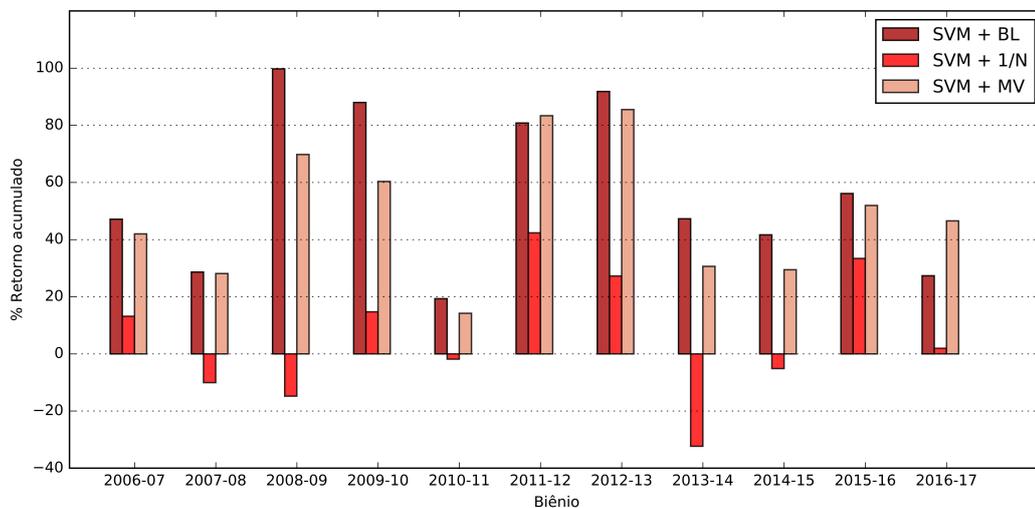


Figura 28: SVM - Retorno bienal por modelo

Fonte: dados da pesquisa (2018).

No que tange à análise de risco de cada modelo, a Tabela 15 apresenta uma análise por biênio da volatilidade por estratégia. Observa-se que a estratégia SVM+BL apresenta maior volatilidade que as demais estratégias em termos médios e em cada um dos biênios avaliados. Foram feitos testes estatísticos para comparação da medida de tendência central da séries, onde foram encontradas evidências para rejeitar a hipótese nula para a comparação entre as estratégias SVM+BL e SVM+1/N ( $p\text{-value} = 0,026$ ) com um nível de significância de 5%. Para as demais comparações não foi possível rejeitar a hipótese nula (SVM+BL x SVM+MV  $p\text{-value} = 0,061$  e SVM+1/N x SVM+MV  $p\text{-value} = 0,724$ ).

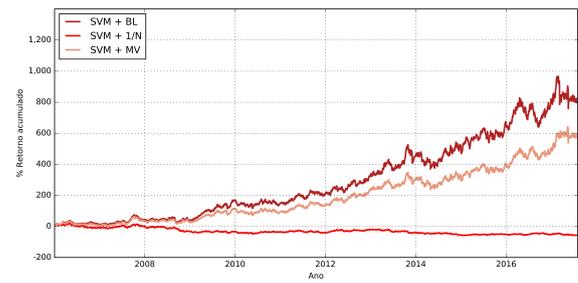
**Tabela 15: SVM - Volatilidade por biênio (%)**

Biênio	SVM+BL	SVM+1/N	SVM+MV
2006-2007	34,84	32,52	33,70
2007-2008	36,56	33,33	34,22
2008-2009	34,97	32,74	32,39
2009-2010	34,99	33,23	33,87
2010-2011	33,65	32,67	33,13
2011-2012	30,96	30,58	29,61
2012-2013	33,29	31,43	30,64
2013-2014	36,75	34,93	35,11
2014-2015	34,59	33,01	33,78
2015-2016	31,75	30,62	29,95
2016-2017	37,91	35,36	33,53
<b>Média</b>	<b>34,77</b>	<b>32,86</b>	<b>32,83</b>
<b>DP</b>	<b>2,09</b>	<b>1,52</b>	<b>1,84</b>
<b>Pearson</b>	<b>6,01</b>	<b>4,64</b>	<b>5,62</b>

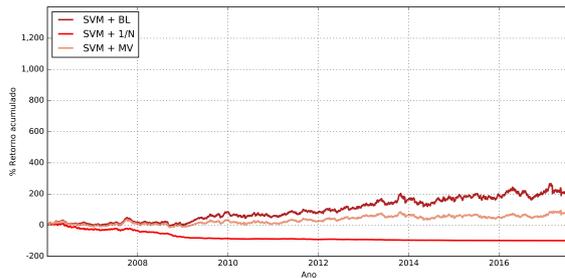
Por fim, foi feita uma análise considerando os custos de transação, cujos critérios são os mesmos utilizados anteriormente. A Figura 29 apresenta uma análise considerando custos de transação de 0,05 bps, 0,10 bps, 0,50 bps e 1,00 bps. Para os valores de 0,05 bps e 0,10 bps, Figuras 29a e 29b, respectivamente, observa-se que os modelos SVM+BL e SVM+MV continuam apresentando um comportamento bem parecido, e mesmo com a redução da rentabilidade total esperada eles ainda chegam ao final do período com um saldo final positivo. Quando aumentam-se os valores para 0,50 bps, Figura 29c, o modelo SVM+BL ainda continua com um saldo final positivo e o modelo SVM+MV fica com um resultado muito próximo de zero. Para 1,00 bps, Figura 29d, todas as estratégias passam a apresentar saldo negativo, o que inviabiliza a operação.



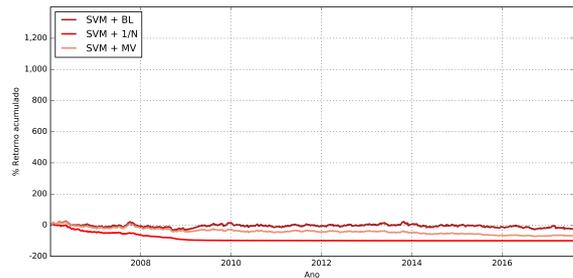
(a) 0,05 bps



(b) 0,10 bps



(c) 0,50 bps



(d) 1,00 bps

**Figura 29: SVM - Retorno acumulado considerando custos de transação**

Fonte: dados da pesquisa (2018).

### 4.3 Comparações com as *baselines*

Nesta subseção são feitas comparações entre os modelos propostos, que valeram-se das técnicas de *machine learning*, e as carteiras alternativas que foram a base de comparação, conforme explicitado na seção de metodologia. Ademais, salienta-se que essa subseção trata do terceiro objetivo específico do trabalho.

Foram selecionados os modelos com maior retorno acumulado em cada uma das técnicas. Para os modelos de RNA, não houve diferença estatisticamente significativa entre os retornos dos modelos RNA+BL e RNA+MV, e estes alternaram em vários momentos a melhor performance dentro da análise por biênio. Assim, optou-se pelo modelo RNA+BL, visto que ele apresentou melhores resultados em termos de cardinalidade e volume de ordens de operação. Para os modelos de SVM, foi escolhido o modelo SVM+BL que apresentou os melhores resultados em termos de retorno e ordens de operação em todas as análises.

Nas esferas de *trades* e cardinalidade, a Tabela 16 resume as operações de compra e venda para cada modelo. Observa-se que a estratégia SVM+BL apresenta os menores valores, com um total de operações 62,7% menor que a estratégia MV, e também tem os menores valores de média diária e máximo de *trades*. Essa performance se confirma em um espectro por biênio (Figura 30), sendo que a estratégia SVM+BL apresenta superioridade em todos os subperíodos analisados.

Tabela 16: Trades (compra e venda) por modelo

	SVM+BL	RNA+BL	Naive+BL	MV
Total Trades	27.596	42.758	47.856	73.950
Média / dia	10	16	18	26
Desvio padrão	5,91	5,08	6,92	8,93
Mínimo / dia	0	1	1	1
Máximo / dia	38	80	72	52

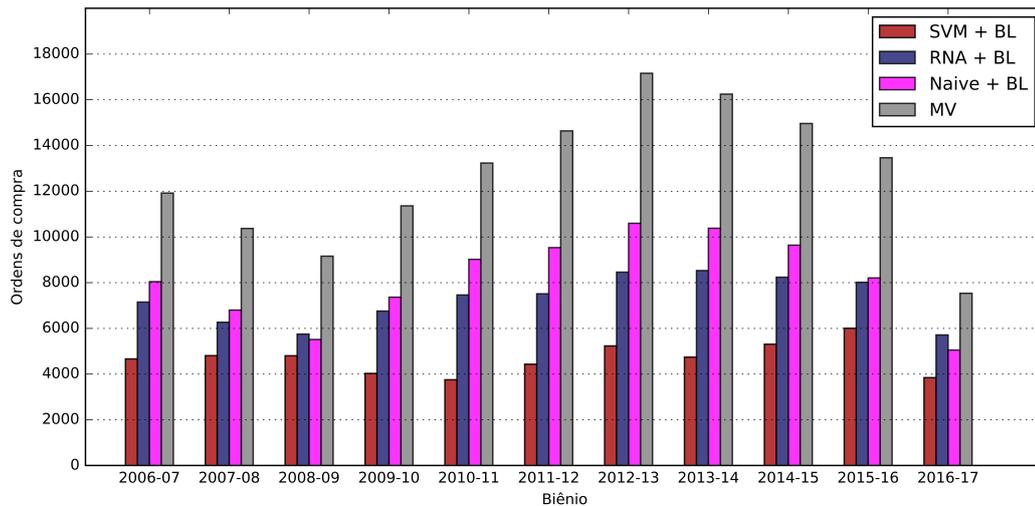
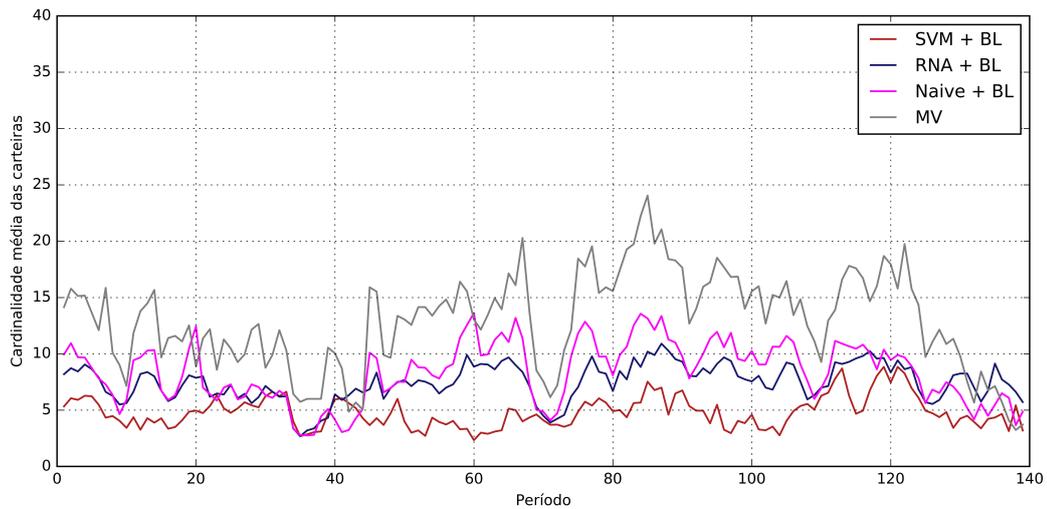


Figura 30: Geral - Ordens de compra e venda por biênio

Fonte: dados da pesquisa (2018).

Em seguida, a Figura 31 apresenta a cardinalidade média mensal dos modelos. Essa análise corrobora os resultados apresentados até o momento, onde observa-se a estratégia SVM+BL com menores valores médios de cardinalidade (janela mensal) durante todo o período de análise. A estratégia MV possui os valores mais altos. As estratégias RNA+BL e Naive+BL apresentam comportamentos semelhantes, ficando em todos os momentos bem próximas num patamar intermediário.



**Figura 31: Geral - Cardinalidade média dos modelos**

Fonte: dados da pesquisa (2018).

Na sequência, realizou-se testes estatísticos para verificar a igualdade da medida de tendência central dos modelos. Uma vez não atendido o pressuposto de normalidade das séries, procedeu-se com o teste de Kruskal-Wallis, onde foram encontradas evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,000$ ). Realizou-se em seguida o teste de Duun para comparações par a par, cujos valores podem ser vistos na Tabela 17. Foi encontrada diferença significativa entre todos os modelos propostos, exceto entre os modelos RNA+BL e Naive+BL.

**Tabela 17: Teste de Duun para cardinalidade**

	SVM+BL	RNA+BL	Naive+BL	MV
SVM+BL	-	-	-	-
RNA+BL	0,000	-	-	-
Naive+BL	0,000	0,060	-	-
MV	0,000	0,000	0,000	-

Sobre o retorno médio mensal por estratégia, a Tabela 18 apresenta a estatística descritiva para as séries de retornos sem custos de transação. Observa-se que a carteira SVM+BL apresenta a maior média mensal dentre as estratégias, valor 61,1% maior que a estratégia MV, que apresenta o segundo maior retorno. Os menores valores foram os das estratégias Naive+BL e IBOV, que ficaram muito próximas de zero.

**Tabela 18: Retorno médio mensal**

	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
Média %	2.03	0.81	-0.02	1.26	0.40
Desvio padrão %	6.75	5.69	7.42	3.65	6.61
Pearson	3.32	7.07	371.03	2.91	16.59
Mínimo %	-16.14	-38.04	-25.75	-16.20	-28.57
Máximo %	22.82	12.55	18.78	11.62	16.20

A Figura 32 apresenta o retorno acumulado para cada um dos modelos. É evidente o destaque e o comportamento crescente da curva de retornos da estratégia SVM+BL, fechando com valor final de 1.113,00%, 166,5% maior que a curva do modelo MV. As demais estratégias não apresentam distinção visual evidente, ficando sempre próximas, com alternância entre melhores e piores resultados.



**Figura 32: Geral - Retorno acumulado sem custos de transação**

Fonte: dados da pesquisa (2018).

Adicionalmente, foram feitos testes estatísticos para verificar a igualdade da medida de tendência central dos modelos. Como premissa inicial, seguiu-se com a averiguação dos pressupostos de normalidade, que não foram atingidos. Em vista disso, procedeu-se com o teste de Kruskal-Wallis, no qual encontrou-se evidências suficientes para rejeitar a hipótese nula ( $p\text{-value} = 0,000$ ). Desta forma, seguiu-se com o teste de Duun para comparações par a par. A Tabela 19 apresenta os valores do teste, nos quais destaca-se a diferença significativa entre o modelo SVM+BL e todas as demais comparações. Para os demais modelos não é possível rejeitar a hipótese de igualdade das médias em pelo menos uma das comparações.

**Tabela 19: Teste de Dunn para retorno mensal**

	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
SVM+BL	-	-	-	-	-
RNA+BL	0,000	-	-	-	-
Naive+BL	0,001	0,001	-	-	-
MV	0,000	0,111	0,000	-	-
IBOV	0,000	0,066	0,000	0,809	-

A Tabela 20 traz uma análise complementar para as séries de retorno. O modelo SVM+BL apresenta a maior proporção de dias com retorno positivo (59,97%), e a segunda maior média de retorno em operações dessa natureza (1,08%). O modelo IBOV apresentou a melhor relação perda média por ganho média e maior média de retorno em operações de

ganho, contudo, esse resultado é acompanhado da pior proporção de operações de ganho, o que faz com que a estratégia não consiga se destacar das demais.

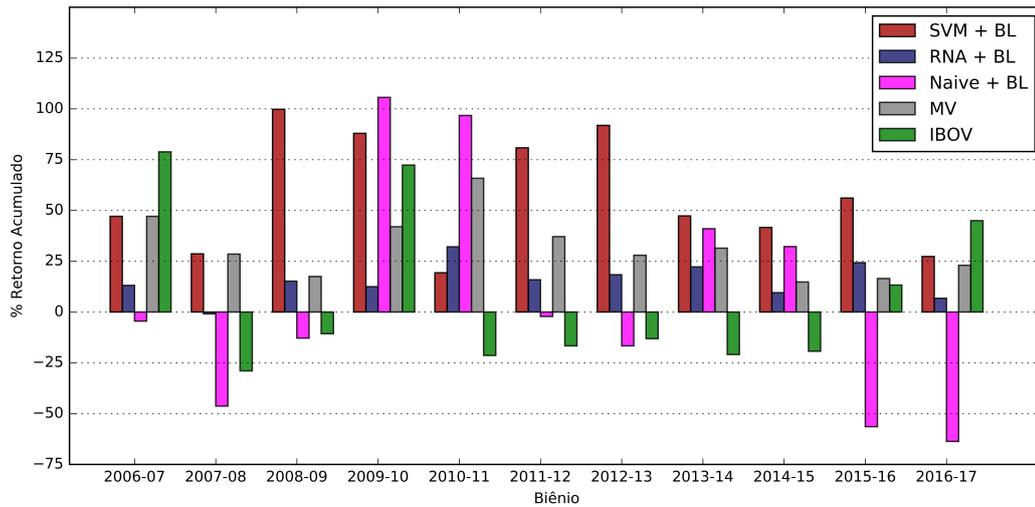
**Tabela 20: Estatísticas adicionais**

	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
(1) % dias retorno > 0	59.97	57.04	57.7	58.75	51.73
(2) % dias retorno < 0	40.03	42.96	42.3	41.25	48.27
(3) ret. médio dias ganho %	1.08	0.83	1.06	0.61	1.24
(4) ret. médio dias perda %	-1.37	-1.00	-1.44	-0.72	-1.28
Ratio (1) / (2)	1.50	1.33	1.36	1.42	1.07
Ratio (4) / (3)	1.27	1.21	1.36	1.18	1.03

A Tabela 21 e a Figura 33 apresentam uma análise bienal por modelo. O modelo SVM+BL apresenta retorno acumulado maior que os outros modelos em sete dos onze biênios em análise, sendo que esse modelo não apresenta retorno negativo em nenhum dos onze períodos em análise. O modelo MV também não apresenta nenhum ponto negativo na análise, mas é inferior ao modelo SVM+BL em oito dos onze biênios.

**Tabela 21: Retorno acumulado por biênio (%)**

	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
2006-2007	47.11	13.13	-4.48	47.11	78.83
2007-2008	28.62	-0.89	-46.27	28.54	-28.95
2008-2009	99.73	15.22	-12.82	17.47	-10.66
2009-2010	87.96	12.50	105.61	42.02	72.33
2010-2011	19.32	32.12	96.74	65.84	-21.37
2011-2012	80.80	15.89	-2.17	37.11	-16.61
2012-2013	91.82	18.41	-16.61	27.95	-13.10
2013-2014	47.28	22.22	41.00	31.43	-20.88
2014-2015	41.63	9.54	32.21	14.79	-19.28
2015-2016	56.11	24.31	-56.45	16.50	13.29
2016-2017	27.35	6.77	-63.67	22.99	44.96
<b>Acumulado</b>	<b>1113,00</b>	<b>138,47</b>	<b>-34,38</b>	<b>417,69</b>	<b>27,74</b>



**Figura 33: Geral - Retorno bianal por modelo**

Fonte: dados da pesquisa (2018).

No que diz respeito à análise de risco de cada modelo, a Tabela 22 apresenta uma análise bianal da volatilidade por estratégia. A estratégia MV apresenta menor volatilidade que as demais estratégias na análise geral e em todos biênios. Observa-se que a estratégia SVM+BL fica em um patamar intermediário, mas em todos os momentos conserva-se abaixo da estratégia MV. Os testes estatísticos para se verificar a igualdade da medida de tendência central das séries estão apresentados na Tabela 23. Foram encontradas evidências para rejeitar a hipótese nula para a estratégia MV em todas as comparações. Para a estratégia SVM+BL, destaque nas análises anteriores, foi possível rejeitar a hipótese nula para as comparações com a estratégia RNA+BL e MV, e para as demais comparações não foi possível rejeitar a hipótese nula.

**Tabela 22: Volatilidade por biênio (%)**

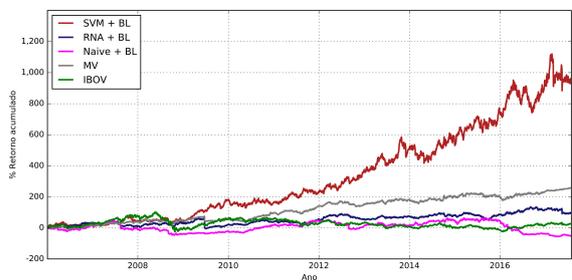
	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
2006-2007	34,84	29,44	43,95	23,06	36,60
2007-2008	36,56	34,42	48,02	27,04	59,30
2008-2009	34,97	51,21	39,96	29,89	61,23
2009-2010	34,99	47,25	27,14	25,02	37,45
2010-2011	33,65	22,91	28,53	17,58	32,11
2011-2012	30,96	23,30	32,49	16,39	32,87
2012-2013	33,29	21,18	32,80	15,49	29,79
2013-2014	36,75	22,17	30,95	17,63	32,50
2014-2015	34,59	25,04	36,40	18,81	34,26
2015-2016	31,75	25,12	44,65	18,07	35,37
2016-2017	37,91	27,55	45,19	14,58	34,94
<b>Média</b>	<b>34,77</b>	<b>31,32</b>	<b>38,03</b>	<b>20,75</b>	<b>39,73</b>
<b>DP</b>	<b>2,09</b>	<b>10,27</b>	<b>7,40</b>	<b>5,10</b>	<b>10,85</b>
<b>Pearson</b>	<b>6,01</b>	<b>32,80</b>	<b>19,45</b>	<b>24,59</b>	<b>27,31</b>

**Tabela 23: Teste de Dunn para volatilidade**

	SVM+BL	RNA+BL	Naive+BL	MV	IBOV
SVM+BL	-	-	-	-	-
RNA+BL	0,049	-	-	-	-
Naive+BL	0,894	0,035	-	-	-
MV	0,000	0,040	0,000	-	-
IBOV	0,790	0,025	0,894	0,000	-

Aqui é apresentada uma análise considerando os custos de transação para as estratégias avaliadas, mantidos os critérios utilizados anteriormente. A Figura 34 apresenta uma análise considerando custos de transação de 0,05 bps, 0,10 bps, 0,50 bps e 1,00 bps. Ressalta-se que a estratégia IBOV não sofre alterações, visto que esta representa um possível fundo ou aplicação com base no índice de mercado.

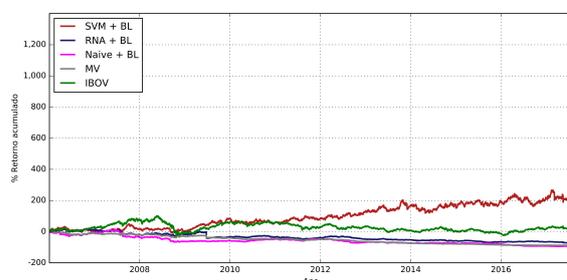
Observa-se que a estratégia MV, que figurava como a segunda melhor estratégia quando não eram considerados os custos, em função da alta cardinalidade, já se aproxima muito das curvas inferiores nas análises com 0,05 bps e 0,10 bps, e passa a perder para a estratégia RNA+BL na análise com 0,50 bps. Tal efeito ocorre em função do maior número de operações da estratégia MV quando comparada às estratégias em que os classificadores precederam a otimização. O modelo SVM+BL é a única estratégia que se mantém com um saldo positivo em todas as análises, ficando com resultado inferior, mas bem próximo ao índice de mercado, mesmo no cenário com 1,00 bps (Figura 34d).



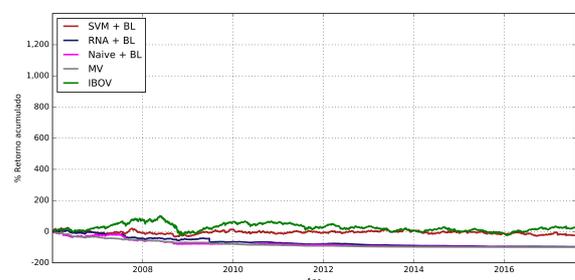
(a) 0,05 bps



(b) 0,10 bps



(c) 0,50 bps



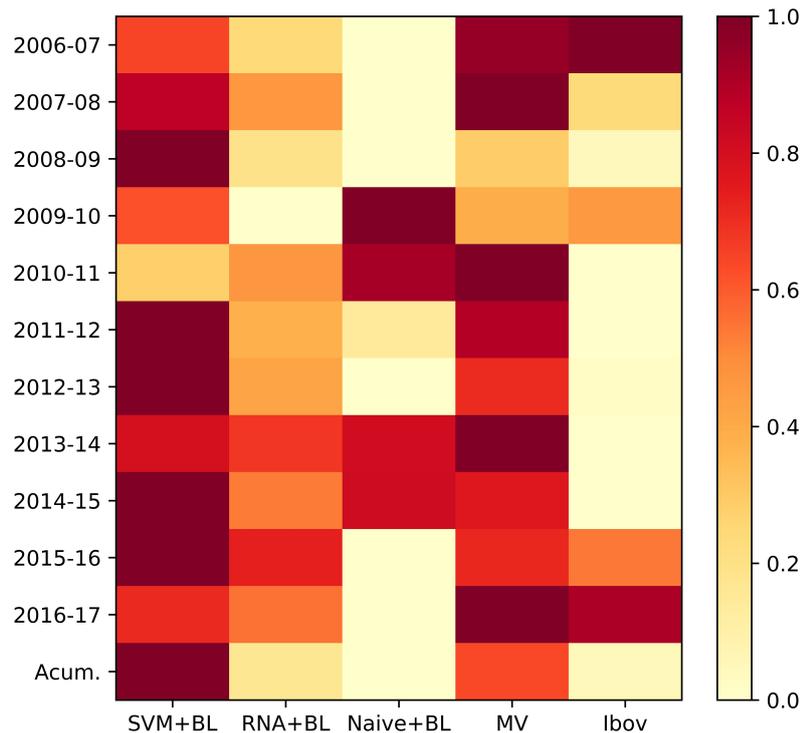
(d) 1,00 bps

**Figura 34: Geral - Retorno acumulado considerando custos de transação**

Fonte: dados da pesquisa (2018).

Por fim, apresenta-se uma avaliação adicional com a relação retorno e risco das estratégias

apresentadas nesta seção. O índice apresentado foi obtido pela razão entre o retorno acumulado e a volatilidade nos períodos analisados, sendo desejável a obtenção de valores quanto maiores possíveis para essa métrica. A Figura 35 apresenta o mapa de calor considerando o cenário sem custos de transação, onde cada linha representa a comparação entre as cinco estratégias por biênio, exceto pela linha final que apresenta a comparação para o período total.

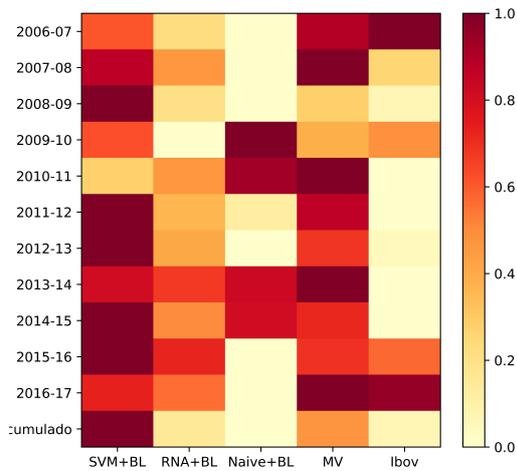


**Figura 35: Geral - Retorno x Risco - sem custos de transação**

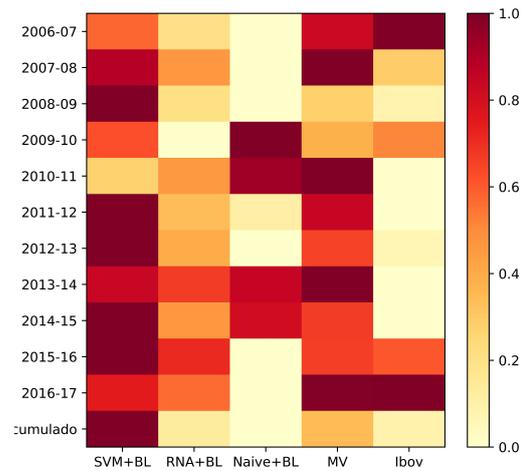
Fonte: dados da pesquisa (2018).

Observa-se que a estratégia SVM+BL apresenta os melhores índices em oito dos onze biênios avaliados e, também, no acumulado do período. As estratégias Naive+BL e Ibov apresentaram os piores valores na maioria dos casos, exceto pelos biênios 2006-07 e 2009-10, onde o Ibov apresentou pontualmente melhor performance.

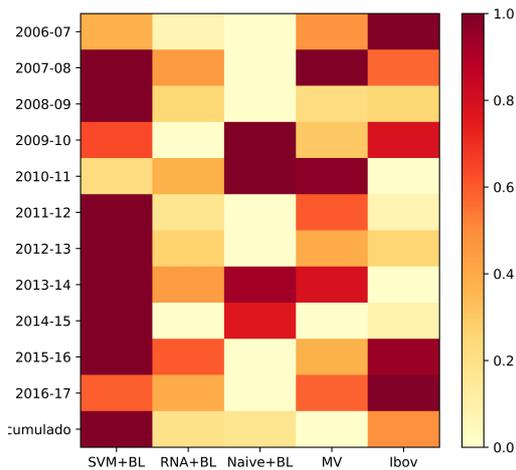
Adicionalmente, a Figura 36 apresenta a avaliação da relação retorno e risco considerando os custos de transação. Observa-se que conforme vão se fazendo os incrementos na taxa base de custos, a estratégia MV tem sua performance impactada de maneira mais intensa que a estratégia SVM+BL. Tal efeito se dá pelo maior número de ativos médio da carteira MV e, conseqüentemente, maior volume no número de ordens de compra e venda, efeitos já avaliados nos tópicos anteriores.



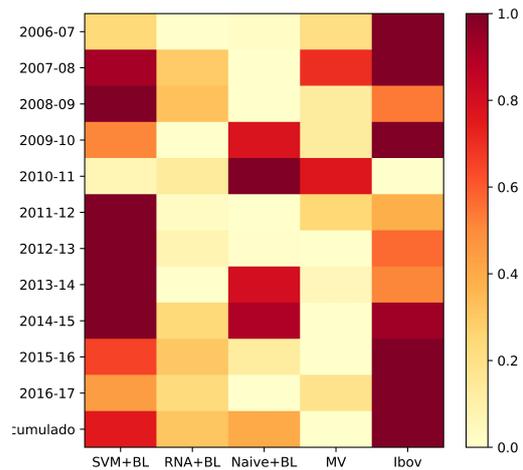
(a) 0,05 bps



(b) 0,10 bps



(c) 0,50 bps



(d) 1,00 bps

**Figura 36: Geral - Retorno x Risco - considerando custos de transação**

Fonte: dados da pesquisa (2018).

Destaca-se ainda que mesmo nos cenários com 0,50bps e 1,00bps, Figuras 36c e 36d, a estratégia SVM+BL consegue alternar períodos com melhor relação retorno x risco que as demais estratégias, sendo a melhor estratégia no acumulado do cenário com 0,50 bps e a segunda melhor no cenário com 1,00 bps.

Posto isto, encerra-se esta etapa de apresentação e discussão de resultados e passa-se às considerações finais.

## 5 CONSIDERAÇÕES FINAIS

O presente trabalho visou propor modelos baseados em *machine learning* para realizar previsões de retornos de ativos financeiros que foram usadas para compor carteiras utilizando o modelo de Black-Litterman.

Desta forma, o primeiro passo da pesquisa foi a construção de modelos baseados nas duas técnicas de *machine learning* escolhidas, redes neurais artificiais e *support vector machine*, para realizar previsões de ativos financeiros. Os modelos trabalharam como classificadores, onde eles retornavam um indicador binário que indicava os ativos com potencial de atingimento de determinados *targets*, no caso, retornos diários de 1,0%, 1,5% e 2,0%.

Neste cenário, os resultados apontaram que o classificador baseado em RNA não apresentou performance estatisticamente significativa quando comparado a uma proposição ingênua. Já para o classificador baseado em SVM, os resultados indicaram uma performance significativamente melhor em todos os espectros quando comparados a uma estratégia ingênua. Os modelos baseados em SVM geraram uma quantidade menor de sinais com melhor taxa de *precision*, o que possibilita que o *decision maker* tenha indicativos para estruturar uma operação mais enxuta, com menos custos de transação, entrando em menos operações, porém com mais assertividade, resultado que se apresenta como a primeira contribuição do trabalho.

Em seguida, no segundo passo do trabalho, buscou-se otimizar carteiras utilizando o modelo de Black-Litterman, tendo como parâmetro de *views* as previsões da RNA e SVM. Além disso, foi fornecido ao modelo BL uma matriz com a diagonal composta dos sinais do classificador como matriz  $P$ , e um vetor com o *target* em questão como vetor  $Q$ . O  $\Omega$  foi calculado por meio do *precision* histórico do ativo no momento em questão, sendo que esse cálculo também seguiu a lógica das janelas deslizantes.

Esse acoplamento figura como a segunda e principal contribuição feita pelo presente trabalho para a literatura em finanças aplicadas e os estudos sobre tomada de decisão em investimentos, visto que a junção de um classificador baseado em *machine learning* dentro da proposta do modelo BL ainda não tinha sido contemplado em estudos e aplicações anteriores. Tal contribuição pode ser dividida em três pontos essenciais: (i) proposta de uma ferramenta de suporte ao *decision maker*, que conta com a ampliação do seu potencial de previsão e aplicação dessa previsão mais assertiva ao processo de composição de portfólio, (ii) o uso carteira de mínima variância para obtenção da distribuição de equilíbrio do mercado, e, por fim, (iii) a proposta da medida de *precision* como parâmetro  $\Omega$  do modelo.

A comparação dos resultados das carteiras obtidas com os valores de referência (índice Ibovespa e uma carteira elaborada pelo método proposto por Markowitz) compôs a etapa final da pesquisa. O cumprimento desse objetivo valida as etapas anteriores e atesta que de fato a proposição do experimento foi bem sucedida, fazendo avaliações em diversos prismas, de modo a garantir a validade e robustez do modelo.

Em termos de cardinalidade e volume de ordens de operação, os modelos que acoplaram as técnicas de *machine learning* ao modelo BL apresentaram melhoria significativa em relação ao modelo MV em uma aplicação simples. Destaque para o modelo SVM+BL, que

apresentou os melhores valores dentre todas as estratégias. Esse resultado é importante porque um número menor de ativos envolvidos na operação é um indicativo de menores gastos com custos e emolumentos.

Em termos de retorno, numa análise sem custos de transação, observou-se que a estratégia SVM+BL alcançou os melhores resultados, ficando bem destacada das demais opções, fechando com um valor final quase que três vezes maior que uma aplicação simples do modelo MV. Essa superioridade também foi confirmada num espectro bienal, onde pode ser vista uma performance superior em todos os subperíodos, o que mostra a consistência do modelo na questão temporal.

Em uma análise em que se considera os custos de transação, observa-se que a estratégia SVM+BL apresenta-se superior às demais estratégias em todos os cenários. Mesmo no cenário de 1bps, essa estratégia ainda se mantém com saldo final positivo, apesar de o resultado ficar muito próximo de zero e da estratégia representativa do índice de mercado. Estes resultados são reflexos da performance supracitada, o que mostra o potencial da aplicação da estratégia em um cenário real, com a incidência de custos efetivos.

Por fim, em uma análise da relação retorno *versus* risco, observa-se que a estratégia SVM+BL tem melhor performance que as demais estratégias, tanto na análise sem custos de transação quanto quando os custos são considerados. À medida que vão se fazendo os incrementos na taxa base de custos, a estratégia MV tem sua performance impactada de maneira mais intensa que a estratégia SVM+BL. Tal efeito se dá pelo maior número de ativos médio da carteira MV e conseqüente maior volume no número de ordens de compra e venda.

Como proposta para trabalhos futuros, sugere-se o teste de uma gama maior de arquiteturas de RNA para compor o conjunto de parametrização, o uso de outras técnicas de *machine learning* como classificadores, o teste desta aplicação em outros mercados e o teste em segmentos de econômicos específicos do mercado brasileiro.

## REFERÊNCIAS

- ALEXANDER, G. J.; SHARPE, W. F.; BAILEY, J. V. **Fundamentals of investments**. 3. ed. New Jersey: Prentice Hall, 2001.
- ARDIA, D.; BOUDT, K. Implied expected returns and the choice of a mean–variance efficient portfolio proxy. **The Journal of Portfolio Management**, v. 41, n. 4, p. 68–81, 2015.
- BARBERIS, N. et al. X-CAPM: An extrapolative capital asset pricing model. **Journal of Financial Economics**, Elsevier, v. 115, n. 1, p. 1–24, 2015.
- BESSLER, W.; OPFER, H.; WOLFF, D. Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naive diversification approaches. **The European Journal of Finance**, v. 23, n. 1, p. 1–30, 2017.
- BLACK, F. Universal hedging: Optimizing currency risk and reward in international equity portfolios. **Financial Analysts Journal**, v. 45, n. 4, p. 16–22, 1989.
- BLACK, F.; LITTERMAN, R. B. Asset allocation: combining investor views with market equilibrium. **The Journal of Fixed Income**, v. 1, n. 2, p. 7–18, 1991.
- \_\_\_\_\_. Global portfolio optimization. **Financial Analysts Journal**, v. 48, n. 5, p. 28–43, 1992.
- BODIE, Z.; KANE, A.; MARCUS, A. J. **Investimentos**. 10. ed. Porto Alegre: AMGH, 2015.
- BRAGA, A. d. P.; CARVALHO, A. C. P. d. L. F. d.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007.
- COPELAND, T. E.; WESTON, J. F.; SHASTRI, K. **Financial theory and corporate policy**. New York: Pearson Education, 2005.
- CRISTIANINI, N.; SCHOLKOPF, B. Support vector machines and kernel methods: the new generation of learning machines. **Ai Magazine**, v. 23, n. 3, p. 31–42, 2002.
- DEMIGUEL, V.; GARLAPPI, L.; UPPAL, R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? **Review of Financial Studies**, v. 22, n. 5, p. 1915–1953, 2009.
- DING, Y.; SONG, X.; ZEN, Y. Forecasting financial condition of Chinese listed companies based on support vector machine. **Expert Systems with Applications**, v. 34, n. 4, p. 3081–3089, 2008.
- ELTON, E.; GRUBER, M.; BROWN, S. **Moderna teoria de carteiras e análise de investimentos**. Rio de Janeiro: Elsevier Brasil, 2012.
- FABOZZI, F. J.; MARKOWITZ, H. M. **The Theory and Practice of Investment Management: Asset Allocation, Valuation, Portfolio Construction, and Strategies**. 2. ed. New Jersey: John Wiley & Sons, 2011.
- FAMA, E. F.; FRENCH, K. R. Dissecting anomalies with a five-factor model. **The Review of Financial Studies**, Oxford University Press, v. 29, n. 1, p. 69–103, 2016.
- FIROOZYE, N.; KUMAR, M. Asset allocation model. **Global Markets Research: Fixed Income Research**, December, p. 99–106, 2003.
- FISHER, L. Using modern portfolio theory to maintain an efficiently diversified portfolio. **Financial Analysts Journal**, v. 31, n. 3, p. 73–85, 1975.

- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.
- HAYKIN, S. **Redes Neurais: Princípios e práticas**. 2. ed. Porto Alegre: Bookman, 2001.
- HE, G.; LITTERMAN, R. B. The intuition behind Black-Litterman model portfolios. **SSRN Electronic Journal**, 2002.
- HE, P. W.; GRANT, A.; FABRE, J. Economic value of analyst recommendations in Australia: an application of the Black–Litterman asset allocation model. **Accounting & Finance**, v. 53, n. 2, p. 441–470, 2013.
- HEROLD, U. Computing implied returns in a meaningful way. **Journal of Asset Management**, v. 6, n. 1, p. 53–64, 2005.
- IDZOREK, T. M. A step-by-step guide to the Black-Litterman model. In: **Forecasting expected returns in the financial markets**. Edição: Stephen Satchell. San Diego: Academic Press, 2007. p. 17–38.
- KAHNEMAN, D. A perspective on judgment and choice: mapping bounded rationality. **American psychologist**, American Psychological Association, v. 58, n. 9, p. 697, 2003.
- KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. **Econometrica**, v. 47, n. 2, p. 263–291, 1979.
- LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The Review of Economics and Statistics**, v. 47, n. 1, p. 13–37, 1965.
- LITTERMAN, R. B. (Ed.). **Modern investment management: an equilibrium approach**. New Jersey: John Wiley & Sons, 2003.
- LORENA, A. C.; CARVALHO, A. C. P. d. L. F. d. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.
- LUO, L.; CHEN, X. Integrating piecewise linear representation and weighted support vector machine for stock trading signal prediction. **Applied Soft Computing**, v. 13, n. 2, p. 806–816, 2013.
- MANDELBROT, B.; HUDSON, R. L. **The (mis)behavior of markets**. New York: Basic Books, 2004.
- MARKOWITZ, H. M. Foundations of portfolio theory. **The Journal of Finance**, Wiley Online Library, v. 46, n. 2, p. 469–477, 1991.
- \_\_\_\_\_. Portfolio selection. **The Journal of Finance**, v. 7, n. 1, p. 77–91, 1952.
- \_\_\_\_\_. **Portfolio Selection, Efficient Diversification of Investments**. New York: J. Wiley, 1959.
- MEDEIROS, M. C.; PASSOS, A. M.; VASCONCELOS, G. F. Parametric Portfolio Selection: Evaluating and Comparing to Markowitz Portfolios/(Seleção Paramétrica de Portfólios: Avaliação e Comparação com Portfólios de Markowitz). **Revista Brasileira de Finanças**, v. 12, n. 2, p. 257, 2014.
- MEUCCI, A. The Black-Litterman approach: original model and extensions. **SSRN Electronic Journal**, 2010.
- MICHAUD, R. O.; ESCHA, D. N.; MICHAUD, R. O. Deconstructing Black Litterman: How To Get the Portfolio You Already Knew You Wanted. **Journal Of Investment Management**, v. 11, n. 1, p. 6–20, 2013.

- MICHAUD, R. O.; MICHAUD, R. O. **Efficient asset management**: a practical guide to stock portfolio optimization and asset allocation. New York: Oxford University Press, 2008.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of machine learning**. [S.l.]: MIT press, 2012.
- MOSSIN, J. Equilibrium in a capital asset market. **Econometrica**, v. 34, n. 4, p. 768–783, 1966.
- NEGNEVITSKY, M. **Artificial intelligence**: a guide to intelligent systems. 2. ed. Harlow: Pearson Education, 2005.
- NI, L.; NI, Z.; GAO, Y.-Z. Stock trend prediction based on fractal feature selection and support vector machine. **Expert Systems with Applications**, v. 38, n. 5, p. 5569–5576, 2011.
- PAIVA, F. D. **Redes neurais para decisões no mercado de ações brasileiro**. Tese (Doutorado em Administração) – Universidade Federal de Lavras, Lavras, 2014.
- PAIVA, F. D.; CARDOSO, R. T. N. et al. Precificação de um ativo financeiro baseado em redes neurais. In: XIV SEMEAD, 2011, São Paulo. **Anais do XIV Semead**. São Paulo: USP, 2011.
- PAIVA, F. D.; DUARTE, W. M. et al. Tomada de decisão de investimento em ativos da bolsa de valores de São Paulo baseado na técnica support vector machine. In: ENANPAD 2017, 2017, São Paulo. **Enanpad 2017**. São Paulo: Anpad, 2017.
- PETERS, E. E. **Chaos and order in the capital markets**: a new view of cycles, prices, and market volatility. [S.l.]: John Wiley & Sons, 1996. v. 1.
- ROLL, R. A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. **Journal of Financial Economics**, v. 4, n. 2, p. 129–176, 1977.
- RUBINSTEIN, M. Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. **The Journal of Finance**, v. 57, n. 3, p. 1041–1045, 2002.
- SANTOS, A. A. P.; TESSARI, C. Técnicas Quantitativas de Otimização de Carteiras Aplicadas ao Mercado de Ações Brasileiro (Quantitative Portfolio Optimization Techniques Applied to the Brazilian Stock Market). **Revista Brasileira de Finanças**, v. 10, n. 3, p. 369, 2012.
- SATCHELL, S.; SCOWCROFT, A. A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction. In: **Forecasting expected returns in the financial markets**. Edição: Stephen Satchell. San Diego: Academic Press, 2007. p. 39–53.
- SHARPE, W. 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.
- \_\_\_\_\_. Imputing expected security returns from portfolio composition. **Journal of Financial and Quantitative Analysis**, v. 9, n. 03, p. 463–472, 1974.
- \_\_\_\_\_. Mutual fund performance. **The Journal of Business**, v. 39, n. 1, p. 119–138, 1966.
- SIMON, H. A. **Comportamento administrativo: estudo dos processos decisórios nas organizações administrativas**. 2. ed. Rio de Janeiro: FGV, 1971.

TU, J.; ZHOU, G. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. **Journal of Financial Economics**, v. 99, n. 1, p. 204–215, 2011.

VAPNIK, V. N. **The nature of statistical learning theory**. [S.l.]: Springer science & business media, 2013.

WALTERS, J. The Black-Litterman model in detail. **SSRN Electronic Journal**, 2011.



## A APÊNDICE

Quadro 2: Empresas consideradas na amostra

Ticker	Empresa	Ticker	Empresa
ABEV3	AMBEV SA	HGTX3	CIA HERING
ACES4	ARCELOR BRASIL SA	HYPE3	HYPERA SA
AEDU3	ANHANGUERA EDUCACIONAL PARTI	INEP4	INEPAR SA INDUSTRIA E C-PREF
AELP3	AES ELPA SA	ITSA4	ITAU
AGEI3	AGRE EMPREENDIMENTOS IMOBILI	ITUB4	ITAU
ALLL11	ALL AMERICA LATINA LOGISTICA	JBSS3	JBS SA
ALLL3	ALL AMERICA LATINA LOGISTICA	KLBN11	KLABIN SA
ARCE3	ARCELOR BRASIL SA	KLBN4	KLABIN SA
ARCZ6	ARACRUZ CELULOSE SA-PREF B	KROT3	KROTON EDUCACIONAL SA
BBAS3	BANCO DO BRASIL S.A.	LAME4	LOJAS AMERICANAS SA-PREF
BBAS4	BANCO DO BRASIL S.A.	LIGH3	LIGHT SA
BBDC3	BANCO BRADESCO S.A.	LIGT3	LIGHT SA
BBDC4	BANCO BRADESCO S.A.	LREN3	LOJAS RENNER S.A.
BBSE3	BB SEGURIDADE PARTICIPACOES	MMXM3	MMX MINERACAO E METALICOS SA
BISA3	BROOKFIELD INCORPORACOES SA	MRFG3	MARFRIG GLOBAL FOODS SA
BNCA3	BANCO NOSSA CAIXA S.A.	MRVE3	MRV ENGENHARIA
BRAP4	BRADESPAR SA -PREF	MULT3	MULTIPLAN EMPREENDIMENTOS
BRDT4	PETROBRAS DISTRIBUIDORA-PREF	NATU3	NATURA COSMETICOS SA
BRFS3	BRF SA	NETC4	NET SERVICOS DE COMUNI-PREF
BRKM5	BRASKEM SA-PREF A	OGXP3	OLEO E GAS PARTICIPACOES SA
BRML3	BR MALLS PARTICIPACOES SA	OIBR3	OI SA
BRPR3	BR PROPERTIES SA	OIBR4	OI SA
BRTP3	BRASIL TELECOM PART SA	PCAR4	CIA BRASILEIRA DE DIS-PREF
BRTP4	BRASIL TELECOM PART SA	PDGR3	PDG REALTY SA
BTOW3	B2W CIA DIGITAL	PETR3	PETROBRAS
BVMF3	B3 SA-BRASIL BOLSA BALCAO	PETR4	PETROBRAS
CCPR3	CYRELA COMMERCIAL PROPERTIES	POMO4	MARCOPOLO SA-PREF
CCRO3	CCR SA	PRML3	PRUMO LOGISTICA SA
CESP5	CIA ENERGETICA DE SP	PRTX3	PORTX OPERACOES PORT SA
CESP6	CIA ENERGETICA DE SP	PTIP4	CIA BRASIL PETRO IPIRANGA-PR
CGAS5	CIA DE GAS DE SAO PAULO-PR A	QUAL3	QUALICORP SA
CIEL3	CIELO SA	RADL3	RAIA DROGASIL SA
CLSC4	CENTRAIS ELETRICAS SANTA CAT	RAIL3	RUMO SA
CMET4	CAEMI - PREF	RDCD3	REDECARD SA
CMIG3	CEMIG	RENT3	LOCALIZA RENT A CAR
CMIG4	CEMIG	RLOG3	COSAN LOGISTICA SA
CPFE3	CPFL ENERGIA SA	RSID3	ROSSI RESIDENCIAL SA
CPLE6	CIA PARANAENSE DE ENERGI-PFB	RUMO3	RUMO SA
CRTP5	CELULAR CRT PARTICIP-PREF A	SANB11	BANCO SANTANDER BRASIL-UNIT
CRUZ3	SOUZA CRUZ LTDA	SBSP3	CIA SANEAMENTO BASICO DE SP
CSAN3	COSAN SA INDUSTRIA COMERCIO	SDIA4	SADIA SA-PREF
CSNA3	CIA SIDERURGICA NACIONAL SA	SMLE3	SMILES SA
CSTB4	CIA SIDERURGICA TUBARAO-PREF	SUZB5	SUZANO PAPEL E CELULO-PREF A
CTAX3	CONTAX PARTICIPACOES SA	TAEE11	TRANSMISSORA ALIANCA DE-UNIT
CTAX4	CONTAX PARTICIPACOES SA	TAMM4	TAM SA-PREF
CTIP3	CETIP SA-MERCADOS ORGANIZADO	TCOC4	TELE CENTRO OESTE CEL-PREF
CYRE3	CYRELA BRAZIL REALTY SA EMP	TCSL4	TIM PARTICIPACOES SA-PREF
DASA3	DIAGNOSTICOS DA AMERICA SA	TDBH4	TELEFONICA DATA BRASIL-PREF
DTEX3	DURATEX SA	TESA3	TERRA SANTA AGRO SA
DURA4	DURATEX SA	TIMP3	TIM PARTICIPACOES SA
EBTP3	EMBRATEL PARTICIPACOES SA	TLCP4	TELE LESTE CELULAR PART-PREF
EBTP4	EMBRATEL PARTICIPACOES SA	TMAR5	TELEMAR NORTE LESTE SA-PR A
ECOR3	ECORODOVIAS INFRA E LOG SA	TMCP4	TELEMIG CELULAR PARTICI-PREF
EGIE3	ENGIE BRASIL ENERGIA SA	TNEP4	TELE NORDESTE CEL PART-PREF
ELET3	CENTRAIS ELETRICAS BRAS.	TNLP3	TELE NORTE LESTE
ELET6	CENTRAIS ELETRICAS BRAS.	TNLP4	TELE NORTE LESTE
ELPL4	ELETROPAULO	TRPL4	CIA DE TRANSMISSAO DE ENE-PF
ELPL5	ELETROPAULO	UBBR11	UNIBANCO-UNITS
EMBR3	EMBRAER SA	UGPA3	ULTRAPAR PARTICIPACOES SA
EMBR4	EMBRAER SA	UGPA4	ULTRAPAR PARTICIPACOES SA
ENBR3	EDP - ENERGIAS DO BRASIL SA	USIM3	USIMINAS
EQTL3	EQUATORIAL ENERGIA SA - ORD	USIM5	USIMINAS
ESTC3	ESTACIO PARTICIPACOES SA	VALE3	VALE
EVEN3	EVEN CONSTRUTORA E INCORPORA	VALE5	VALE
FIBR3	FIBRIA CELULOSE SA	VCPA4	VOTORANTIM CELULOSE PAP-PREF
GFSA3	GAFISA SA	VIVO4	VIVO
GGBR4	GERDAU SA	VIVT4	VIVO
GOAU4	GERDAU SA	WEGE3	WEG SA
GOLL4	GOL LINHAS AEREAS INT SA-PRE		