



Universidade de Brasília
FACE - Faculdade de Administração Economia e Contabilidade
Departamento de Administração

Previsão da direção de índices da BOVESPA por intermédio de Máquina de Suporte Vetorial

Felipe Dutra Calainho

Brasília
2015

Felipe Dutra Calainho

Previsão da direção de índices da BOVESPA por
intermédio de Máquina de Suporte Vetorial

Monografia apresentada ao Departamento
de Administração da Universidade de Bra-
sília como requisito parcial para a obtenção
do título de Bacharel em Administração.

Orientador: Prof. Dr. Pedro Henrique
Melo Albuquerque

Brasília
2015

CALAINHO, Felipe Dutra

Previsão da direção de índices da BOVESPA por intermédio de Máquina de Suporte Vetorial

96 páginas

Monografia - Faculdade de Economia, Administração e Contabilidade da Universidade de Brasília. Departamento de Administração.

1. Máquinas de suporte vetorial
2. Previsão
3. BOVESPA

I. Universidade de Brasília. Faculdade de Economia, Administração e Contabilidade. Departamento de Administração.

Comissão Julgadora:

Prof. Dr. Vinicius Amorim Sobreiro

Prof. Dr. Herbert Kimura

Prof. Dr. Pedro Henrique Melo Albuquerque
Orientador

“Nenhuma nação pode florescer e ser feliz enquanto grande parte de seus membros for formada de pobres e miseráveis.”

- Adam Smith

Agradecimentos

Primeiramente a meus pais que fazem de tudo por mim, sem eles esse trabalho nunca seria possível. À Marina, minha companheira e meu grande amor. Ao meu orientador Pedro, que me motivou e explorou minhas capacidades ao longo desse período. Ao Yuri Maluf, meu professor e amigo que me expôs constantemente a novos desafios e me tirou várias dúvidas. Aos meus colegas de curso, em especial ao Pedro Alexandre Moura Barros Henrique, que teve muita paciência e me explicou várias questões de programação.

Resumo

Esta pesquisa tem por objetivo analisar a aplicação de Máquinas de Suporte Vetorial com o intuito de prever o movimento de índices de ações da BOVESPA. Os dados da pesquisa abrangem o período de 22/01/2001 até 22/09/2015. Os dados de entrada da máquina são os Log-Retornos dos índices e dois indicadores de análise técnica - Índice de Força Relativa e Médias Móveis Convergentes Divergentes. Esses dados são utilizados para determinar o movimento do índice (subir ou descer) e a probabilidade de ocorrência da previsão. Uma validação cruzada (*k-fold*) é realizada para a escolha dos melhores parâmetros, onde o melhor desempenho da máquina é uma acurácia de 70% na previsão.

Palavras-chave: Máquinas de Suporte Vetorial, Previsão, BOVESPA.

Abstract

This research aims to examine the application of Support Vector Machines in order to predict the movement of the Bovespa stock index. This survey data cover the period from 01/22/2001 to 09/22/2015. Machine input data is the log-returns of the indices and two technical analysis indicators - Relative Strength Index and Moving Average Convergence Divergence. These data are used to determine the movement (up or down) of the indices and the probability of the forecast. A cross-validation (*k-fold*) is performed to choose the best parameters, where the best machine performance in forecasting is a hit ratio of 70%.

Keywords: Support Vector Machines, Forecasting, BOVESPA

Lista de Figuras

3.1	Fluxograma do método da pesquisa	18
3.2	Separador ideal	32
3.3	Separador de máxima margem	33
3.4	Suportes vetoriais	34
3.5	Truque Kernel	38
3.6	Truque Kernel 2	41
3.7	<i>K-fold</i> com $k=3$	43
3.8	Organização dos <i>inputs</i> da MSV	44
3.9	Forma final dos <i>inputs</i> da MSV	45
4.1	Resultados 1	48
4.2	Resultados 2	49
4.3	Hiperplanos 1	52
4.4	Hiperplanos 2	53
4.5	Hiperplanos 3	54
4.6	Hiperplanos 4	55

Lista de Tabelas

3.1	Tabela de Kernels do pacote " <i>kernelab</i> "	42
3.2	Números da base de dados	44
4.1	Intervalos da base de dados	47
4.2	Melhores parâmetros de C e σ para H-1	50
4.3	Melhores parâmetros de C e σ para H-5	50
4.4	Matriz de confusão da predição da MSV	50

Lista de abreviaturas e siglas

ACP Análise de Componente Principal.

AT Análise Técnica.

BM&FBOVESPA Bolsa de Valores, Mercadorias e Futuros de São Paulo.

BOVESPA Bolsa de Valores de São Paulo.

HME Hipótese dos Mercados Eficientes.

IAT Indicadores de Análise Técnica.

IBOVESPA Índice da Bolsa de Valores de São Paulo.

IBrX-50 Índice Brasil 50.

IEE Índice Energia Elétrica.

IFNC Índice Financeiro.

IFR Índice de Força Relativa.

INDX Índice Industrial.

LR Log-Retorno.

MMCD Médias Móveis Convergentes Divergentes.

MSV Máquina de Suporte Vetorial.

SMLL Índice Small Cap.

Sumário

1	Introdução	1
1.1	Formulação do problema	3
1.2	Objetivo geral	3
1.3	Objetivos específicos	3
1.4	Justificativa	4
2	Referencial teórico	7
2.1	Hipótese dos Mercados Eficientes	7
2.2	HME e sua dinâmica com a AT	10
2.2.1	Análise Técnica	11
2.3	MSV no contexto de finanças	14
3	Metodologia	17
3.1	Métodos e técnicas de pesquisa	17
3.2	Tipo e descrição geral da pesquisa	17
3.3	Caracterização da organização, setor ou área	18
3.4	População e amostra	19
3.4.1	Cálculo dos índices	19
3.5	Caracterização dos instrumentos de pesquisa	29
3.6	Procedimentos de coleta e de análise dos dados	29

3.6.1	LR, IFR e MMCD	30
3.6.2	Metodologia da MSV	32
3.6.3	MSV com Margem Suave	35
3.6.4	Método Kernel	37
3.6.5	MSV Não-Linear com Margem Suave	38
3.7	Parâmetros do SVM	41
3.7.1	Kernel e seus parâmetros	41
3.7.2	Parâmetro C	42
3.8	Validação Cruzada	43
3.9	Dados de entrada da MSV	43
4	Resultados	47
4.1	<i>Performance</i> da MSV	47
4.2	Análise de Componente Principal e Hiperplano Separador	51
5	Considerações finais	57
	Referências Bibliográficas	60
A	Composição da carteira teórica IBOVESPA	67
B	Composição da carteira teórica IBrX-50	71
C	Composição da carteira teórica IEE	74
D	Composição da carteira teórica IFNC	76
E	Composição da carteira teórica INDX	78
F	Composição da carteira teórica SMLL	81

G	Programação VBA	85
G.1	Programação VBA RSI	85
G.2	Programação VBA MMCD	87
H	Programação R	89
H.1	MSV de previsão	89
H.2	ACP e MSV	93

Capítulo 1

Introdução

A bolsa de valores é um mecanismo que, ao longo dos anos, tem ajudado no desenvolvimento da economia e da sociedade, seja captando recursos para o desenvolvimento de atividades econômicas ou gerando lucros para seus investidores. A globalização tem grande impacto nas bolsas de valores pois amplia as fontes de captação de investimento para empresas que tem ativos negociados na bolsa e aumenta o número de ativos pelos quais o investidor pode optar. O agente financeiro que opera no mercado de capitais busca ferramentas e teorias que tentem explicar o funcionamento desse complexo sistema. Há décadas a bolsa de valores vem sendo pesquisada por estudiosos como Harry Markowitz, Merton Miller, William F. Sharpe e Fama, que buscam explicar e compreender melhor o seu funcionamento.

A BOVESPA, que é a única bolsa de valores no Brasil, vem crescendo em valor de mercado e em número de investidores: em 1994 a BOVESPA valia R\$160,3 bilhões em valores atualizados, já em 2015 seu valor de mercado é de R\$2,2553 trilhões; em 2010 o número médio de investidores era de 42,3 mil, quatro anos depois, em 2014, esse número foi para 118,7 mil (BM&FBOVESPA, 2015). Esses dados demonstram a importância do estudo na área de finanças e mercado de capitais, bem como do desenvolvimento de sistemas que auxiliem os investidores a operar nesse mercado.

Capítulo 1. Introdução

Um dos objetivos da avaliação de ativos do mercado financeiro é determinar o seu preço no futuro. Saber como um ativo irá se comportar no futuro traz vantagens para a organização e para o investidor. Prever o movimento semanal dos índices ajuda no gerenciamento de curto-prazo de portfólios, permitindo que investidores se protejam de perdas a longo prazo (Ince e Trafalis, 2008). No entanto, essa tarefa não é trivial visto que séries temporais financeiras possuem ruídos, são deterministicamente caóticas e não-estacionárias (Abu-Mostafa e Atiya, 1996). Outro aspecto relevante é a Hipótese de Mercados Eficientes (HME) de Fama (1970), que diz que os preços de mercado seguem um passeio aleatório e não podem ser previstos a partir de seu comportamento passado. Complementarmente, existem eventos independentes que podem afetar o mercado, a exemplo das condições macroeconômicas, das expectativas do *trader* e de eventos políticos. Essas características fazem com que seja difícil prever os movimentos futuros usando modelos determinísticos (paramétricos) (Ince e Trafalis, 2008).

No âmbito da Análise Técnica (AT) existem os Indicadores Técnicos (IAT) que guiam o investidor a respeito de quando ele deve ou não comprar ou vender suas ações. De 30% a 40% dos praticantes dessa técnica aparentam acreditar que a AT é um fator importante para determinar o movimento do preço em recortes temporais curtos de até seis meses (Park e Irwin, 2004). A AT é criticada por dois motivos: 1) porque muitos dos críticos adotam a premissa de que o mercado segue a HME e, portanto, é impossível prever o comportamento dos preços; 2) porque algumas partes da AT, como as figuras, são altamente subjetivas.

A AT foca na dinâmica de comportamento do preço e volume do mercado para a predição do preço futuro, ignorando a HME (Ince e Trafalis, 2004). Segundo Park e Irwin (2004), de um total de 92 estudos recentes, 58 indicaram resultados positivos na utilização da AT, 24, resultados negativos e 10, resultados mistos. Apesar das evidências favoráveis, os autores destacam a existência de indícios de que alguns desses estudos estão sujeitos a problemas em seus procedimentos de teste, a exemplo de *data snooping*,

1.1. Formulação do problema

seleção de regras de negociação *ex post facto*, e dificuldades em estimar risco e custos de transação.

As Máquinas de Suporte Vetorial (MSV), desenvolvidas por [Boser et al. \(1992\)](#), são um tipo de máquina de aprendizado que funciona como um classificador não-paramétrico de máxima margem, operando de forma diferente das outras máquinas de aprendizado, como redes neurais. As MSVs adotam o princípio de minimização do risco empírico e aplicam o princípio da minimização do risco estrutural, buscando minimizar o limite superior do erro de generalização ao invés de minimizar o erro de treinamento da máquina. Isso resulta em um grande poder de generalização ([Cao e Tay, 2001](#)).

1.1 Formulação do problema

Este estudo visa responder à seguinte pergunta:

Máquinas de Suporte Vetorial que utilizam Índice de Força Relativa, Médias Móveis Convergentes Divergentes e o Log-Retornos dos índices são capazes de prever o movimento dos índices da BOVESPA com alguns dias de antecedência?

1.2 Objetivo geral

Medir o desempenho da Máquina de Suporte Vetorial, com o uso de Índice de Força Relativa, Médias Móveis Convergentes Divergentes e Log-Retorno, na previsão do movimento de índices da BOVESPA, utilizando como base de dados os valores diários do IBOVESPA, IBrX-50, IEE, IFNC, INDX e SMLL.

1.3 Objetivos específicos

A fim de atingir o objetivo geral, formulam-se os seguintes objetivos específicos:

Capítulo 1. Introdução

1. Apresentar os conceitos básicos de Hipótese de Mercados Eficientes e sua dinâmica com Indicadores de Análise Técnica no mercado de capitais;
2. Apresentar os conceitos básicos de Máquinas de Suporte Vetorial no contexto de finanças;
3. Descrever o conceito e o cálculo de: Médias Móveis Convergentes Divergentes, Índice de Força Relativa e Log-Retorno;
4. Aplicar o algoritmo de Máquina de Suporte Vetorial na amostra;
5. Medir o desempenho de previsão do movimento de índices da Máquina de Suporte Vetorial.

1.4 Justificativa

Segundo [Jain \(2005\)](#), o avanço tecnológico nas telecomunicações e na internet proporcionou uma mudança no modelo básico de negócios na bolsa de valores. Bolsas ao redor do mundo introduziram sistemas de negociação eletrônicos totalmente automatizados e transparentes, substituindo os pregões onde os corretores efetuavam as ordens manualmente. Isso dá um grande poder ao investidor que pode, rapidamente, pesquisar grandes bases de dados em tempo real para selecionar as ações que melhor se enquadram em um conjunto de critérios, obtendo o resultado em segundos ([Gold e Lebowitz, 1999](#)).

Os idealizadores da MSV foram Vladimir N. Vapnik e Alexey Y. Chervonenkis (1963), mas o desenvolvimento e a aplicação desse algoritmo só foi possível após a imigração de Vapnik da Rússia para os Estados Unidos da América em 1990. A primeira publicação utilizando esse algoritmo foi em 1992 por Bernhard E. Boser, Isabelle M. Guyon e Vapnik, evidenciando que esse tema é bastante recente.

1.4. Justificativa

Fan e Palaniswami (2001) foram os primeiros a aplicar as MSVs em um contexto de finanças, utilizando-as para classificar ações e formar portfólios. Cao e Tay (2001), Kim (2003), Huang et al. (2005), Rosillo et al. (2014), empregam a MSV para prever a movimentação de índices de mercado. Nesses casos, as MSVs obtiveram resultados promissores, sugerindo aos previsores financeiros e *traders* que esse caminho pode levar a ganhos de capital. Na América Latina esse tipo de estudo segue inédito.

A relevância teórica desse artigo está relacionada a uma abordagem inovadora no contexto de previsão de séries temporais financeiras e à aplicação de MSV nos Mercados Financeiros, principalmente, no Mercado de Ações Brasileiro. No Brasil não há artigos publicados relacionados a MSV no contexto de Mercado de Capitais, até o momento e há poucos estudos sobre o tema.

Capítulo 2

Referencial teórico

Desde o início da sociedade, a humanidade produz, distribui, negocia e consome bens e serviços. Constituinto um sistema econômico que, ao longo do tempo, foi se tornando cada vez maior e mais complexo, tanto geograficamente como numericamente. Ao longo desse processo de desenvolvimento, que é contínuo, criou-se várias ferramentas, uma delas para facilitar as transações comerciais, chamada de moeda. Desde a criação da moeda tivemos casos curiosos, como a crise das tulipas na Holanda e a crise dos subprimes em 2008 (Neal e D., 2003; Stephens et al., 2013). Esses eventos influenciaram a busca por uma melhor previsão e compreensão das tendências de mercado.

Para apresentar o referencial que dá base a este trabalho, será introduzido primeiramente o conceito da Hipótese dos Mercados Eficientes. Em seguida serão debatidos o conceito de AT, seus indicadores e sua dinâmica com a Hipótese dos Mercados Eficientes. Por último as MSVs no contexto de finanças.

2.1 Hipótese dos Mercados Eficientes

No começo da década de 1960, Fama desenvolveu a Hipótese dos Mercados Eficientes como parte de sua tese de Ph.D na Universidade de Chicago. Essa hipótese, daqui em

Capítulo 2. Referencial teórico

diante chamada de HME, requer, além da utilidade normal de maximização dos lucros dos agentes econômicos, que os agentes tenham expectativas racionais e sempre que nova informação aparecer, que o agente atualize suas expectativas apropriadamente. Note que não é requerido que os agentes sejam racionais: a HME permite que, quando diante de uma nova informação, alguns investidores possam “sobre reagir” e outros “sub reagir”. Tudo que é necessário pela HME é que as reações dos investidores sejam aleatórias e sigam um padrão de distribuição normal para que o efeito de rede no mercado de preços não possa ser explorado de forma confiável a ponto de criar lucro anormal, especialmente quando considerados os custos de transação. Como Fama (1965) afirma: “qualquer pessoa pode estar equivocada sobre o mercado, mas o mercado como um todo está sempre certo”.

O ponto principal da HME é que é impossível obter lucros acima do mercado porque a eficiência do mercado de ações faz com que os preços existentes sempre incorporem e reflitam toda informação relevante. De acordo com a HME, ações são sempre negociadas ao seu preço justo na bolsa de valores. Consequentemente, é impossível para os investidores comprarem ações subvalorizadas e/ou vender ações por preços inflacionados. As flutuações ocasionais nos preços das ações são decorrentes do processo de absorção de novas informações pelos investidores e seguem um padrão aleatório, não previsível. Portanto, deveria ser impossível superar o mercado integralmente seja por meio de hábil seleção de ações ou por tempo de mercado (“*Market timing*”). Consequentemente, o único meio pelo qual um investidor pode possivelmente obter retornos maiores é através da compra de investimentos mais arriscados.

Em 1965, Fama publica uma dissertação que dá suporte à sua teoria de que o mercado se comporta de forma aleatória, chamada de hipótese do passeio aleatório (*Random Walk Hypothesis*). O termo passeio aleatório, entretanto, foi popularizado por Burton Malkiel em seu livro “*Random Walk Down Wall Street*” de 1973. Essa dialética entre aleatoriedade e movimento de preço de ações foi formulada, pela primeira vez, por

2.1. Hipótese dos Mercados Eficientes

Louis Bachelier (1900), em sua tese de doutorado intitulada “A Teoria da Especulação”. Bachelier modela o processo estocástico que atualmente é conhecido como movimento Browniano, sendo ele o pioneiro no estudo das análises de ações e opções ([Courtault et al., 2000](#)).

A HME foi revisada em 1970 por Fama, que adicionou à hipótese anteriormente desenvolvida a definição de três formas de eficiência de mercado: fraca, semiforte e forte. Na forma fraca de eficiência os preços futuros não podem ser preditos analisando preços do passado. Retornos excedentes não podem ser adquiridos no longo prazo utilizando estratégias de investimento baseadas em preços históricos das cotas ou outros dados históricos. Não há padrão nos preços de ativos. Isso implica que os movimentos dos preços futuros são determinados inteiramente por informação não contida nas séries de preço. Por isso, os preços devem seguir o passeio aleatório. Essa forma de eficiência não requer que os preços se mantenham próximos ao equilíbrio, mas somente que os participantes do mercado não sejam capazes de sistematicamente lucrar das “ineficiências” de mercado.

Na forma semiforte de eficiência, os preços das cotas ou ações se ajustam a novas informações publicamente disponíveis rapidamente e de uma forma imparcial, tal que nenhum retorno excedente pode ser ganho por negociar essa informação. Já na forma forte de eficiência, o preço das cotas reflete toda informação existente, pública e privada, ninguém pode ter retornos excedentes. Se houver barreiras legais para informação privada se tornar pública, como ocorre com leis de *insider trading*, a forma forte de eficiência é impossível – exceto no caso em que as leis sejam universalmente ignoradas ([Fama, 1970](#)).

Deve se notar que a HME e o passeio aleatório não são conceitos correspondentes. O passeio aleatório dos preços das ações não implica que o mercado de ações é eficiente, com investidores racionais. O passeio aleatório é definido pelo fato de que os preços mudam aleatoriamente. O mesmo se aplica ao modelo de Martingale: ambos, Martingale

Capítulo 2. Referencial teórico

e passeio aleatório, se refutados não implicam em refutar a HME.

Investidores e pesquisadores tem contestado a HME empiricamente e teoricamente. A HME é rebatida por diversas fontes, [Keim e Stambaugh \(1986\)](#) relatam em seu estudo o encontro de sazonalidade em algumas medidas de risco e nos retornos médios. [Chen \(1991\)](#) acredita que o mercado é previsível até certo grau. Os resultados de seu estudo oferecem interpretações diretas de evidências recentes sobre as previsões do excesso de retorno do mercado por utilização de fatores macroeconômicos como produção industrial, taxa “*t-bill*”, “*Default Spread*”, entre outras. Outros estudos econômicos contestam a aleatoriedade de mercado bem como a HME, um destes é o artigo de [Lo e MacKinley \(1988\)](#), dois economistas que tentam provar que a hipótese do passeio aleatório está equivocada.

No Brasil a HME foi estudada por pesquisadores estrangeiros como [Mobarek e Fiorante \(2014\)](#), que em seu estudo determinam se os mercados de ações do Brasil, Rússia, Índia e China (BRIC) podem ser considerados eficientes na forma fraca nos anos recentes. A amostra do estudo abrange dados que vão de setembro de 1995 a março de 2010, Os autores concluíram que os resultados dão suporte à crença de que esses mercados estão se aproximando a um estado de razoável eficiência na forma fraca. [Ojah e Karamera \(1999\)](#) estudam a HME e o passeio aleatório nos mercados de ações da América Latina, Argentina, Brasil, Chile e México, de 1987 a 1997. O estudo concluiu que esses mercados seguem o passeio aleatório e geralmente são eficientes na forma fraca.

2.2 HME e sua dinâmica com a AT

A HME não é unânime entre os pesquisadores e agentes financeiros, embora esse pensamento seja o mais aceito no âmbito das finanças. Em parte, porque é difícil incorporar toda a complexidade e a dinâmica do mercado em uma teoria. A HME

reflete muito bem o comportamento básico do mercado, mas, como será visto adiante, o mercado não é eficiente e racional o tempo todo (Tversky e Kahneman, 1986), o que colabora com o argumento da ineficiência do mercado que vai contra a HME e traz ao debate a questão da possibilidade de ganhar do mercado de forma sistemática.

2.2.1 Análise Técnica

AT é uma metodologia de análise de títulos para previsão da direção dos preços de mercado por intermédio do estudo de dados passados, como preço e volume negociado. O principal objetivo da AT é ser capaz de identificar mudanças nas tendências em um estágio inicial e manter uma estratégia de investimento até que o peso da evidência indique que a tendência se inverteu (Gençay, 1998). A AT possui alguns aspectos do investimento ativo, que entram em contradição com grande parte da HME e a teoria moderna do portfólio. A eficácia da AT é desafiada pela HME, que afirma que os preços do mercado de ações são essencialmente imprevisíveis.

A AT surge da Teoria de Dow. Segundo Brown et al. (1998), a maior parte do que conhecemos por Teoria de Dow, não vem do editor fundador do “*The Wall Street Journal*”, Charles Henry Dow mas, de William Peter Hamilton, que assumiu a edição do jornal depois da morte de Dow, em 1902. Ainda segundo os autores, Hamilton no livro “*The Stock Market Barometer*” (1922), fundamentou a Teoria de Dow, a primeira teoria para os grafistas. Os autores afirmam que Robert Rhea popularizou a ideia na década de 1930 com seu livro intitulado “*The Dow Theory*” (1932).

AT se divide no estudo de gráficos e seus padrões (análise qualitativa) e no estudo de indicadores (análise quantitativa) (Menkhoff e Taylor, 2007) que podem ser divididos em indicadores de tendência como Médias Móveis Convergentes Divergentes (MMCD) e osciladores como o Índice de Força Relativa (IFR). Há mais de 100 indicadores técnicos que podem ser usados para tentar ter uma visão sobre o comportamento do mercado, o problema é que esses indicadores são altamente dependentes dos preços (Ince e Trafalis,

2006).

Curcio et al. (1997) fizeram um estudo avaliando a rentabilidade das estratégias de negociação baseadas na aplicação da técnica de suporte e resistência nos mercados de câmbio “*intraday*”. Os resultados, na média, não geraram retorno, principalmente levando em conta os custos de transação. Taylor (2014), examinou a performance de um grupo de IATs aplicados a todas as ações que compõem o índice “*Dow Jones Industrial Average*” (DJIA) de 1928 a 2012, usando um conjunto de modelos econométricos com características particulares. O autor concluiu que a rentabilidade não é constante e que varia ao longo do tempo, evoluindo lentamente, restritos a episódios particulares, principalmente dos meados da décadas de 1960 a meados de 1980.

Um documento com o nome de “*Support for Resistance: Technical Analysis and Intraday Exchange Rates*” (Osler, 2004) do “*Federal Reserve Bank of New York*” sobre os níveis de suporte e resistência nas taxas de câmbio de curto prazo, oferece uma forte evidência de que os níveis ajudam a prever interrupções de tendência “*intraday*”, embora o “poder preditivo” desses níveis varia entre as taxas de câmbio e as empresas examinadas.

Neftci (1991) mostra, sob a hipótese de as séries temporais econômicas serem gaussianas, que indicadores de mercado não são capazes de prever os preços futuros, mas que, quando os processos levados em consideração são não-lineares, as regras de análise técnica podem capturar algumas informações ignoradas pela teoria de predição de Wiener-Kolmogorov, mostrando algum valor preditivo. Já Brock et al. (1992), afirmam ser prematuro dizer que análise técnica é inútil. Estes autores constataram que 26 regras de negociação da AT aplicadas ao *Dow Jones Industrial Average* tem performance significativamente superior a estratégia de *Buy-and-hold* (comprar e esperar). Os resultados chamam atenção pois todas as regras de negociação de AT estudadas foram capazes de bater o *benchmark* de *Buy-and-hold*. Os autores ainda afirmam que os preços de transação devem ser cuidadosamente considerados antes de implementar

tais estratégias. [Sullivan et al. \(1999\)](#) dão continuidade a esse trabalho expandindo o universo das 26 regras de negociação, aplicando-as a 100 anos de dados diários do *Dow Jones Industrial Average* e determinando os efeitos de *data-snooping*. O resultado corrobora com os achados de [Brock et al. \(1992\)](#).

A evidência empírica de que a AT funciona não foi provada de forma definitiva. As avaliações acadêmicas muitas vezes acham que a AT tem pouco poder preditivo. De 95 estudos modernos sobre o assunto, 56 concluem que a AT pode ter resultados positivos, embora vieses de *data-snooping* e outros problemas fazem a análise difícil ([Park e Irwin, 2007](#)), uma continuação do estudo de [Park e Irwin \(2004\)](#) citado nesse estudo anteriormente.

No Brasil há estudos referentes à relevância das Médias Móveis Convergentes Divergentes, segundo [Vidotto et al. \(2009\)](#) a rentabilidade média acumulada de cinco ações, operadas utilizando esse indicador, foi de 26,7% no ano de 2006, contra uma rentabilidade média acumulada de 0,90% do Ibovespa. Já [Carvalho et al. \(2008\)](#) analisaram 28 ações da Bolsa de Valores de São Paulo no período de julho de 1994 a julho de 2003 não encontrando resultados significantes para o uso de Médias Móveis Convergentes Divergentes (MMCD). Porém o resultado desse estudo dá suporte à ideia de que quanto mais eficiente é o mercado de ações, menos eficiente é a estratégia de MMCD. [de Paula e Triches \(2010\)](#) empregaram um indicador técnico semelhante, chamado de Média Móvel, na gestão de uma carteira criada a partir das ações com maior participação na carteira teórica do Ibovespa. As evidências encontradas mostraram que a gestão utilizando Média Móvel apresentou um rendimento de 69,27%, tal desempenho foi superior ao rendimento alcançado pela estratégia comparativa de *Buy-and-hold* da mesma carteira de ações, que foi de 44,61%.

Segundo [Lo et al. \(2000\)](#), a AT, tem sido uma parte da prática financeira por muitas décadas, mas não tem recebido o mesmo nível de escrutínio acadêmico e aceitação de abordagens mais tradicionais, como análise fundamentalista. Os autores ressaltam que

um dos maiores obstáculos à AT é a alta subjetividade na sua natureza qualitativa – a presença de formas geométricas nos gráficos de histórico de preços, frequentemente, só existe aos olhos do observador. Em seu artigo os autores propõem uma abordagem sistemática e automática para o reconhecimento de padrões técnicos usando regressão não-paramétrica com *kernels*. Esse método foi aplicado a um grande número de ações no mercado dos EUA de 1962 a 1996, de modo a avaliar a efetividade da AT. Os autores concluem o trabalho afirmando que a AT, em muitos períodos, fornece informações incrementais, mas que isso não implica na geração de retornos anormais de negociação e sim na possibilidade da AT adicionar valor ao processo de investimento.

2.3 MSV no contexto de finanças

Como mencionado antes, [Fan e Palaniswami \(2001\)](#) foram os primeiros a aplicar as MSVs em um contexto de finanças. Os autores utilizaram dados contábeis e informações de preço de ações negociadas na Bolsa de Valores Australiana, para identificar ações que podem ter retornos excepcionais e porventura, superar o mercado. Depois de classificar as ações, as que possuem o melhor desempenho, são selecionadas para compor um portfólio com pesos igualmente ponderados. Por fim, os autores obtiveram um desempenho de 208% de retorno ao longo de 5 anos, o *benchmark* do mesmo período foi de 71%.

[Cao e Tay \(2001\)](#) utilizam uma variação de MSV para prever o movimento diário do índice S&P 500, comparando seu desempenho com redes neurais. Seus dados de entrada são *Relative Difference Percentage* de 5, 10, 15 e 20 dias, média móvel de 15 dias, MMCD e a volatilidade do *On Balance Volume* (OBV). Os dados do S&P 500 da pesquisa vão de 01/04/1993 até o final de dezembro de 1994 para treinamento e de 01/03/1995 até o final de dezembro de 1995 para teste. Os autores concluirão que a MSV fornece uma alternativa promissora para a previsão de séries temporais, tendo

2.3. MSV no contexto de finanças

desempenho superior ao algoritmo de rede neural e sugeriram que a MSV ainda pode ser aprimorada para essa aplicação.

Kim (2003) emprega a MSV para prever o KOPSI, índice da bolsa coreana, semanalmente, comparando o desempenho da MSV com um tipo de rede neural, chamada *Back-Propagation Neural Network* (BPN), e raciocínio baseado em casos (RBC). O autor categoriza as direções do índice em 0 e 1, 0 significa que o valor do índice no próximo dia é mais baixo que o de hoje, e 1 significa que o valor do índice no próximo dia é mais alto que o de hoje.

A amostra do estudo de Kim (2003) vai de janeiro de 1989 a dezembro de 1998 totalizando 2928 dias de negociação. 80% da amostra foi utilizada para treinamento e 20% para teste. Os dados de entrada utilizados nesse estudo foram estocástico %K, estocástico %D (uma média móvel de %K), *Slow %D*, *Momentum*, *Rate-of-change* (ROC), %R de William, Oscilador de acumulação e distribuição, Disparidade de 5 dias, Disparidade de 10 dias, Oscilador de preço, *Commodity channel index*, e IFR. Os autores afirmam que a performance de previsão da MSV é sensível aos parâmetros C e σ^2 , ressaltam a importância de achar os valores ótimos desses parâmetros e por último que nos resultados experimentais a MSV superou o RBC e o BPN.

Huang et al. (2005) empregam a MSV para prever a movimentação semanal do índice NIKKEI 225, comparando sua performance com análise linear discriminante, análise discriminante quadrática e *Elman Backpropagation Neural Networks*. Os dados de entrada dos modelos são taxa de juros, taxa de juros de curto prazo, taxa de juros de longo prazo, índice de preço do consumidor, produção industrial, consumo do governo, consumo privado, produto nacional bruto, produto interno bruto.

A amostra da pesquisa de Huang et al. (2005) vai de 01/01/1990 a 31/12/2002, totalizando 676 pares de observação, sendo 640 pares de observação usados para determinar as especificações do modelo e seus parâmetros. Os últimos 36 pares de observação são reservados para avaliação e comparação de performance entre os vários modelos de

Capítulo 2. Referencial teórico

predição. Para avaliar a habilidade de previsão da MSV eles usaram um modelo de Passeio-Aleatório (PA) como uma *benchmark* para comparação. Neste caso, os autores afirmam que a MSV obteve um resultado superior à análise discriminante linear, análise discriminante quadrática e redes neurais, sugerindo aos previsores financeiros e traders que esse caminho pode levar a ganhos de capital.

Rosillo et al. (2014), por sua vez, analisam a efetividade do índice de volatilidade do “*Chicago Board Options Exchange*” usado com a MSV para prever as mudanças semanais no índice S&P 500. A amostra vai de 03/01/2000 a 30/12/2011, os dados de entrada dessa pesquisa são MMCD padrão (*fast* = 12 dias e *slow* = 26 dias), RSI padrão (14 dias) e o índice de volatilidade (VIX). Para complementar as medidas de performance econômicas, os autores implementaram uma simulação de negociação utilizando essas previsões, concluindo que a MSV consegue atingir bons resultados de negociação e, sobretudo, a estratégia da MSV reduz o risco global do investimento.

Capítulo 3

Metodologia

3.1 Métodos e técnicas de pesquisa

Essa pesquisa segue a metodologia de previsão por intermédio de MSV, proposta por [Dunis et al. \(2013\)](#) aplicada aos índices IBOV, IBrX-50, IEE,INDX, IFNC e SMLL da BOVESPA.

3.2 Tipo e descrição geral da pesquisa

A presente pesquisa é considerada Correlacional Quantitativa Exploratória. A amostra consiste dos valores diários dos índices da BOVESPA. Foi utilizada a MSV com o objetivo de prever o movimento dos índices e a probabilidade de ocorrência dessa previsão para o intervalo de 1 e 5 dias.

Este artigo replica o método utilizado por [Dunis et al. \(2013\)](#) com algumas modificações. A ideia, em ambos, é que a MSV classificadora analise os valores do MMCD, IRF e Log-Retorno diários de 5 dias de negociação passados e classifique os valores dos índices em duas classes: alta e baixa. A predição feita pela MSV, neste trabalho, tem o horizonte de 1 e 5 dias de negociação futuros.

A MSV mostra as condições de alta e de baixa como a probabilidade de que a previsão será cumprida. A taxa de acerto da MSV é calculada a partir de uma validação cruzada chamada de *K-fold* (sendo o $k = 5$). A Máquina calcula se o valor do índice, no horizonte previsto, apresenta uma alta (1) ou uma baixa (0), em relação ao dia da previsão, e a probabilidade de ocorrência da previsão ($0 \leq p \leq 1$).

Esse estudo utiliza dois indicadores: Médias Móveis Convergentes Divergentes (MMCD) e Índice de Força Relativa (IFR). MMCD e IFR foram escolhidos por serem os IATs mais utilizados pelos analistas quantitativos. O IFR é calculado utilizando os Log-Retornos (LR). Nesse artigo olhamos para eles com uma perspectiva diferente – eles são apenas variáveis de entrada para o método MSV.

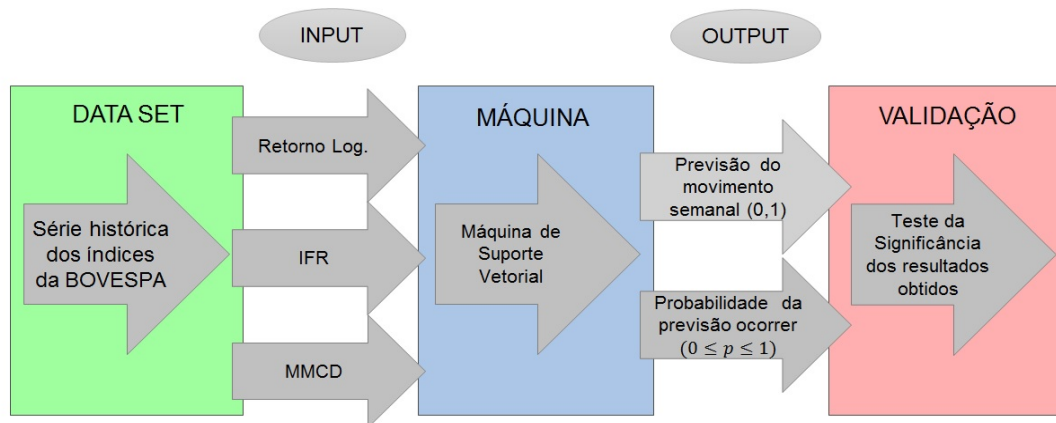


Figura 3.1: Fluxograma do método da pesquisa

3.3 Caracterização da organização, setor ou área

Esta pesquisa estudou a aplicação da MSV para a previsão da movimentação do IBOVESPA, IBrX-50, IEE, IFNC, INDX e SMLL que são índices do mercado de ações da BM&FBOVESPA, que é fruto da fusão da BOVESPA (Bolsa de valores de São Paulo) com a BM&F (Bolsa de Valores, Mercadorias e Futuros) em 2008, formando a

BM&FBOVESPA.SA.

A BM&FBOVESPA.SA é a bolsa de valores do Brasil, a maior da América Latina e é uma das maiores do mundo em valor de mercado. Promovendo a negociação de ações, derivativos de ações, derivativos financeiros, títulos de renda fixa, títulos públicos federais, moedas à vista e commodities.

3.4 População e amostra

A população da pesquisa é composta pelas séries históricas dos índices abordados tendo dimensão infinita, sendo impossível, por definição, coletar toda a série de dados. Sendo formada a partir de cálculos expostos nos próximos tópicos.

As amostras tem características distintas, mas com a base de cálculo semelhante. Por isso ela são caracterizadas individualmente em tópicos adiante.

3.4.1 Cálculo dos índices

Fórmula geral de cálculo do preço “*ex-teórico*”

$$P_{ex} = \frac{P_c + (S \times Z) - D - J - V_{et}}{1 + B + S}. \quad (3.1)$$

O preço teórico P_{ex} é igual à razão entre: o último preço "com-direito" ao provento P_c , mais o produto entre percentual de subscrição, em número de índice S com o valor de emissão do ativo a ser subscrito, em moeda corrente Z , menos o valor recebido a título de dividendo em moeda corrente D , menos juros sobre o capital em moeda corrente J , menos o valor econômico teórico resultante do recebimento de provento em outro tipo V_{et} ; dividido por 1 mais o percentual de bonificação (ou desdobramento) B , mais S . (BMF&BOVESPA, 2013)

Algumas observações devem ser tomadas. Primeiramente o V_{et} é calculado considerando-se o montante financeiro que seria apurado com a venda dos ativos do outro tipo e/ou

Capítulo 3. Metodologia

outro ativo (debêntures, ativos de outra empresa etc.) recebidos. Um exemplo seria a empresa A distribuí-se gratuitamente, aos seus acionistas um ativo da empresa B, que é avaliado em \$5,00 por ativo, portanto, para cada dois ativos da empresa A possuídos, o V_{et} será igual a \$2,50. A segunda observação é no caso de subscrições, o preço *ex-teórico* somente será apurado se os atuais acionistas tiverem preferência para aquisição dos novos valores mobiliários, e se essa operação resultar em vantagem econômica para o investidor. (BMF&BOVESPA, 2013)

$$IN = \sqrt[3]{\frac{n_i}{N}} \times \sqrt[3]{\left(\frac{v_i}{V}\right)^2} \times \frac{p_i}{P}. \quad (3.2)$$

O índice de negociabilidade IN , segundo a equação 3.2, é igual ao produto entre: raiz cúbica da razão do número de negócios com o ativo i no mercado a vista (lote-padrão) n_i , sobre o número total de negócios no mercado a vista da BMF&BOVESPA (lote-padrão) N ; raiz cúbica do quadrado da razão do volume financeiro gerado pelos negócios com o ativo i no mercado a vista (lote-padrão) v_i , sobre o volume financeiro total do mercado a vista BM&FBOVESPA (lote-padrão) V ; e a razão do número de pregões em que o ativo foi negociado p_i , sobre o número de pregões total do período analisado P . (BMF&BOVESPA, 2013)

Lembrando que no cálculo do Índice de Negociabilidade, não são considerados os negócios diretos. E os valores do Índice de Negociabilidade de cada ativo são apurados em base diária, somados e divididos pelo número de pregões total do período analisado. (BMF&BOVESPA, 2013)

A BM&FBOVESPA calcula os índices durante o período de negociação, em um intervalo de 30 segundos, considerando os preços dos últimos negócios efetuados no mercado a vista com ativos que compõem a carteira teórica do índice. Nesse estudo utilizamos o valor de fechamento diários dos índices da BOVESPA. O cálculo dos índices é feito da seguinte forma:

$$\text{Índice}_{(t)} = \frac{\text{Valor total da carteira}}{\text{reductor}} = \frac{\sum_{i=1}^n (P_{it} \times Q_{it})}{\alpha}. \quad (3.3)$$

Conforme a equação 3.3, $\text{Índice}_{(t)}$ é igual ao valor do índice no instante t . n é o número total de ativos integrantes da carteira teórica do índice. P_{it} é o último preço do ativo i no instante t . Q_{it} é a quantidade do ativo i na carteira teórica no instante t . α é o reductor utilizado para adequar o valor total da carteira ao valor de divulgação do índice. (BMF&BOVESPA, 2013)

IBOVESPA

A amostra do IBOVESPA é referente ao período de 22/01/2001 até 21/08/2015. O IBOVESPA é calculado a partir de uma carteira teórica que inclui todas as ações negociadas na BOVESPA com base no desempenho do IN (eq.: 3.2), sendo ajustada a cada quadrimestre. Lembrando que a BMF&BOVESPA alterou o cálculo do IBOVESPA ao final de 2013, valendo a partir do início de 2014. Todos os dados contidos nessa pesquisa seguem a nova forma de cálculo do IBOVESPA.

Como a carteira teórica que compõem o IBOVESPA é alterada a cada quadrimestre, as empresas que compõem a carteira no início da série mudam ao longo do tempo como sua participação também, por influência do IN (eq.: 3.2). A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice A, válida para a data: 06/07/2015 (BMF&BOVESPA, 2015f).

IBrX-50

A amostra do IBrX-50 é referente ao período de 22/01/2001 até 02/09/2015. O IBrX-50 é calculado a partir de uma carteira teórica que inclui todas as ações negociadas na BOVESPA com base no desempenho do IN (eq.: 3.2), formada pelos 50 primeiros colocados de forma decrescente, ajustada a cada quadrimestre.

Capítulo 3. Metodologia

O IBrX-50 é o resultado de uma carteira teórica de ativos, elaborada de acordo com os seguintes critérios:

1. Estar entre os ativos elegíveis que, no período de vigência, ocupem as 50 primeiras posições no IN (eq.: 3.2).
2. Ter presença de pregão de 95% no período de vigência das 3 carteiras anteriores.
3. Não ser classificado como *Penny Stock*.

([BMF&BOVESPA, 2015d](#))

Um ativo que seja objeto de Oferta Pública realizada durante o período de vigência das 3 carteiras anteriores ao rebalanceamento será elegível, mesmo sem estar listado todo o período, desde que: Oferta Pública de distribuição de ações ou units, conforme o caso, tenha sido realizada antes do rebalanceamento imediatamente anterior; possua 95% (noventa e cinco por cento) de presença desde seu início de negociação; atenda cumulativamente aos critérios 1 e 3 do tópico anterior ([BMF&BOVESPA, 2015d](#)).

Sendo excluído da carteira ativos que deixarem de atender a qualquer um dos critérios de inclusão acima citados. E durante a vigência da carteira passem a ser listados em situação especial (para mais informações a respeito da situação especial ver [BMF&BOVESPA \(2013\)](#)).

No IBrX 50, os ativos são ponderados pelo valor de mercado do “*free float*” (ativos que se encontram em circulação) da espécie pertencente à carteira.

A carteira teórica que compõe o IBrX-50 é alterada a cada quadrimestre, da mesma forma que o IBOVESPA. Sendo assim as empresas que compõem a carteira no início da série mudam ao longo do tempo bem como sua participação. A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice B, válida para a data: 26/10/2015 ([BMF&BOVESPA, 2015g](#)).

IEE

A amostra do IEE é referente ao período de 22/01/2001 até 02/09/2015. O IEE é um índice de retorno total com reinvestimento no próprio ativo. O ajuste é efetuado considerando que o investidor vendeu as ações pelo último preço de fechamento anterior ao início da negociação "*ex-provento*" e utilizou os recursos na compra das mesmas ações sem o provento distribuído. Tendo como objetivo ser o indicador do desempenho médio das cotações dos ativos de maior negociabilidade e representatividade do setor de energia elétrica (BMF&BOVESPA, 2015a).

A fórmula utilizada para alteração da quantidade teórica do ativo quando da distribuição de proventos é a seguinte:

$$Q_n = \frac{Q_o \times P_c}{P_{ex}}, \quad (3.4)$$

Onde Q_n é a quantidade nova, Q_o a quantidade antiga, P_c é o último preço de fechamento anterior ao início da negociação "*ex-provento*", P_{ex} é o preço "*ex-teórico*", calculado com base em P_c (BMF&BOVESPA, 2015a).

O IEE é o resultado de uma carteira teórica de ativos, elaborada de acordo os seguintes critérios:

1. Ter participação em termos de volume financeiro maior ou igual a 0,01% no mercado a vista (lote-padrão) no período de vigência das 3 carteiras anteriores.
2. Ter presença em pregão de 80% no período de vigência das 3 carteiras anteriores.
3. Ter apresentado no mínimo 2 negócios/dia em pelo menos 80% dos pregões em que foi negociada.
4. Não ser classificado como "Penny Stock".
5. Pertencer ao setor de energia elétrica.

Capítulo 3. Metodologia

As empresas participam da carteira somente com seu ativo mais líquido, sendo o mais líquido, aquele que apresenta o maior índice de negociabilidade no período (BMF&BOVESPA, 2015a).

Os ativos são excluídos da carteira caso: Eles deixem de atender aos critérios de inclusão 1, 3 ou 4 acima listados; tiverem menos do que 70% de presença em pregão no período de vigência das 3 carteiras anteriores; durante a vigência da carteira passem a ser listados em situação especial. Serão excluídos ao final de seu primeiro dia de negociação nesse enquadramento; forem emitidos por uma empresa que, durante a vigência da carteira, altere sua área de atuação principal para um setor diferente do objetivo do índice (BMF&BOVESPA, 2015a).

O IEE é um índice em que as ações têm igual participação em termos de valor na carteira ("*equal value weighted*"), na época de sua compilação e em seus ajustes periódicos. Isso faz com que as participações sejam iguais no início de cada quadrimestre, variando suas participações, durante o período quadrimestral, ao longo da evolução dos preços individuais de cada ação. O índice tem características peculiares como, ser calculado utilizando o método de "igual ponderação por empresa em termos de Real aplicado", para assegurar que cada ação componente do IEE seja igualmente representada, em termos de valor, na carteira do índice. (BMF&BOVESPA, 2015a).

O segundo a BMF&BOVESPA (2015a) IEE tem um procedimento diferenciado para o início das divulgações diárias do índice: sua difusão tem início somente após a abertura de, pelo menos, 60% dos ativos componentes de sua carteira. As quantidades teóricas dos ativos na carteira do IEE são arredondadas para seus lotes-padrão ou múltiplos.

A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice C, válida para a data: 26/10/2015 (BMF&BOVESPA, 2015h).

IFNC

A amostra do IFNC é referente ao período de 20/01/2005 (data de início do índice) até 22/09/2015. O IFNC é um índice de retorno total, que tem como objetivo ser o indicador do desempenho médio das cotações dos ativos de maior negociabilidade e representatividade dos setores de intermediários financeiros, serviços financeiros diversos, previdência e seguros (BMF&BOVESPA, 2015b).

O IFNC é composto das ações e units exclusivamente de ações de companhias listadas na BM&FBOVESPA que, segundo o BMF&BOVESPA (2015b), atendam cumulativamente aos seguintes critérios de inclusão:

1. Estar entre os ativos elegíveis que, no período de vigência das 3 carteiras anteriores, em ordem decrescente de IN (eq.: 3.2), representem em conjunto 99% do somatório total desses indicadores.
2. Ter presença em pregão de 95% no período de vigência das 3 carteiras anteriores.
3. Não ser classificado como *Penny Stock*.
4. Pertencer aos setores de intermediários financeiros, serviços financeiros diversos, previdência e seguros.

Um ativo que seja objeto de Oferta Pública realizada durante o período de vigência das 3 carteiras anteriores ao rebalanceamento será elegível, mesmo sem estar listado todo o período, desde que: a Oferta Pública de distribuição de ações ou units, conforme o caso, tenha sido realizada antes do rebalanceamento imediatamente anterior; possua 95% de presença desde seu início de negociação; atenda cumulativamente aos critérios 1, 3 e 4 acima citados BMF&BOVESPA (2015b).

Os ativos são excluídos da carteira caso: Eles deixem de atender aos critérios de inclusão indicados acima; durante a vigência da carteira passem a ser listados em situação especial. Serão excluídos ao final de seu primeiro dia de negociação nesse en-

Capítulo 3. Metodologia

quadramento; forem emitidos por uma empresa que, durante a vigência da carteira, altere sua área de atuação principal para um setor diferente do objetivo do índice (BMF&BOVESPA, 2015b).

Segundo o BMF&BOVESPA (2015b) no IFNC, os ativos são ponderados pelo valor de mercado do “*free float*” da espécie pertencente à carteira. A participação dos ativos de uma companhia no índice (considerando todas as espécies e classes de ações ou units que tenham como lastro tais ações da companhia, conforme o caso) não poderá ser superior a 20%, quando de sua inclusão ou nas reavaliações periódicas. Caso isso ocorra, serão efetuados ajustes para adequar o peso dos ativos da companhia a esse limite, redistribuindo-se o excedente proporcionalmente aos demais ativos da carteira.

A carteira teórica que compõe o IFNC é alterada a cada quadrimestre, da mesma forma que os outros índices. Sendo assim as empresas que compõem a carteira no início da série mudam ao longo do tempo bem como sua participação. A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice D, válida para a data: 26/10/2015 (BMF&BOVESPA, 2015i)

INDX

A amostra do INDX é referente ao período de 22/01/2001 até 02/09/2015. O INDX é um índice de retorno total, que tem como objetivo de medir o desempenho das ações mais representativas do setor industrial. Esse índice é resultado de um convênio entre a FIESP (Federação das Indústrias do Estado de São Paulo) e a BM&FBOVESPA (BMF&BOVESPA, 2015e).

O INDX é composto das ações e units exclusivamente de ações de companhias listadas na BM&FBOVESPA que, segundo o BMF&BOVESPA (2015e), atendam cumulativamente aos seguintes critérios de inclusão:

1. Estar entre os ativos elegíveis que, no período de vigência das 3 carteiras anteriores, em ordem decrescente de IN (eq.: 3.2), ocupem as 150 primeiras posições.

3.4. População e amostra

2. Ter presença em pregão de 95% no período de vigência das 3 carteiras anteriores.
3. Não ser classificado como *Penny Stock*.
4. Pertencer aos setor industrial.

Um ativo que seja objeto de Oferta Pública realizada durante o período de vigência das 3 carteiras anteriores ao rebalanceamento será elegível, mesmo sem estar listado todo o período, desde que: a Oferta Pública de distribuição de ações ou units, conforme o caso, tenha sido realizada antes do rebalanceamento imediatamente anterior; possua 95% de presença desde seu início de negociação; atenda cumulativamente aos critérios 1, 3 e 4 acima citados [BMF&BOVESPA \(2015e\)](#).

Os ativos são excluídos da carteira caso: Eles deixem de atender aos critérios de inclusão indicados acima; durante a vigência da carteira passem a ser listados em situação especial. Serão excluídos ao final de seu primeiro dia de negociação nesse enquadramento; forem emitidos por uma empresa que, durante a vigência da carteira, altere sua área de atuação principal para um setor diferente do objetivo do índice ([BMF&BOVESPA, 2015e](#)).

Segundo o [BMF&BOVESPA \(2015e\)](#) no INDX, os ativos são ponderados pelo valor de mercado do “*free float*” da espécie pertencente à carteira. A participação dos ativos de uma companhia no índice não poderá ser superior a 20%, quando de sua inclusão ou nas reavaliações periódicas. Caso isso ocorra, serão efetuados ajustes para adequar o peso dos ativos da companhia a esse limite, redistribuindo-se o excedente proporcionalmente aos demais ativos da carteira.

A carteira teórica que compõe o INDX é alterada a cada quadrimestre, como os demais índices. Sendo assim as empresas que compõem a carteira no início da série mudam ao longo do tempo bem como sua participação. A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice [E](#), válida para a data: 26/10/2015 ([BMF&BOVESPA, 2015j](#))

SMLL

A amostra do SMLL é referente ao período de 21/05/2008 (data de início do índice) até 16/09/2015. O SMLL é um índice de retorno total, que tem como objetivo ser o indicador do desempenho médio das cotações dos ativos de uma carteira composta pelas empresas de menor capitalização. (BMF&BOVESPA, 2015c).

O SMLL é composto das ações e units exclusivamente de ações de companhias listadas na BM&FBOVESPA que, segundo o BMF&BOVESPA (2015c), atendam cumulativamente aos seguintes critérios de inclusão:

1. Estar entre os ativos que, em ordem decrescente, estejam classificados fora da lista dos que representam 85% do valor de mercado de todas as empresas listadas no mercado a vista (lote-padrão) da BM&FBOVESPA.
2. Estar entre os ativos elegíveis que, no período de vigência das 3 carteiras anteriores, em ordem decrescente de IN (eq.: 3.2), representem em conjunto 99% do somatório total desses indicadores.
3. Ter presença em pregão de 95% no período de vigência das 3 carteiras anteriores.
4. Não ser classificado como *Penny Stock*.

Um ativo que seja objeto de Oferta Pública realizada durante o período de vigência das 3 carteiras anteriores ao rebalanceamento será elegível, mesmo sem estar listado todo o período, desde que: a Oferta Pública de distribuição de ações ou units, conforme o caso, tenha sido realizada antes do rebalanceamento imediatamente anterior; possua 95% de presença desde seu início de negociação; atenda cumulativamente aos critérios 1, 3 e 4 acima citados BMF&BOVESPA (2015c).

Os ativos são excluídos da carteira caso: Eles deixem de atender aos critérios de inclusão indicados acima; durante a vigência da carteira passem a ser listados em situação

3.5. Caracterização dos instrumentos de pesquisa

especial. Serão excluídos ao final de seu primeiro dia de negociação nesse enquadramento;

Segundo o [BMF&BOVESPA \(2015c\)](#) no SMLL, os ativos são ponderados pelo valor de mercado do “*free float*” da espécie pertencente à carteira. A participação dos ativos de uma companhia no índice não poderá ser superior a 20%, quando de sua inclusão ou nas reavaliações periódicas. Caso isso ocorra, serão efetuados ajustes para adequar o peso dos ativos da companhia a esse limite, redistribuindo-se o excedente proporcionalmente aos demais ativos da carteira.

A carteira teórica que compõe o SMLL é alterada a cada quadrimestre. Sendo assim as empresas que compõem a carteira no início da série mudam ao longo do tempo bem como sua participação. A carteira teórica que compõem a data final da nossa amostra está presente na tabela do apêndice [F](#), válida para a data: 26/10/2015 ([BMF&BOVESPA, 2015k](#)).

3.5 Caracterização dos instrumentos de pesquisa

Para a formação da base de dados e o cálculo do IFR e MMCD foi utilizado o Excel. Já na análise dos dados pelo método MSV foi feito no software livre R, criado por Robert Gentleman e Ross Ihaka do Departamento de Estatística da Universidade de Auckland, Nova Zelândia. O R é uma interface computacional gratuita de linguagem própria, que opera de forma colaborativa – seus usuários compartilham módulos em bibliotecas on-line.

3.6 Procedimentos de coleta e de análise dos dados

A base de dados da pesquisa foi obtida através de consulta a Economática. Em seguida os dados carregados foram transferidos para uma planilha do Excel onde foram organizados de forma temporal. Os LRs foram calculados nessa planilha. Após a

Capítulo 3. Metodologia

obtenção dos LRs foi feita uma programação em VBA (Visual Basic Applications) para calcular os IATs (IFR e MMCD, códigos conforme apêndice G.1 e G.2, para o IFR e MMCD respectivamente), seguindo a metodologia abaixo.

Por último essa tabela foi importada no software R, que, com o auxílio das bibliotecas de extensão, aplicou a metodologia MSV.

3.6.1 LR, IFR e MMCD

O Log-Retorno ou Retorno Logarítmico (LR) é dado por:

$$LR_i = \ln \left(\frac{VF}{VP} \right), \quad (3.5)$$

LR_i é Log-Retorno no dia i , sendo VF o valor futuro e VP o valor presente. Podendo ser representado da seguinte forma, para o caso do IBOVESPA:

$$LR_n = \ln \left(\frac{IBOVESPA_n}{IBOVESPA_{n-1}} \right), \quad (3.6)$$

LR_n é o retorno no dia n , $IBOVESPA_n$ é o valor do índice no dia n e $IBOVESPA_{n-1}$ é o valor do índice no dia anterior. Essa forma de calcular o retorno foi aplicada a todos os índices dessa pesquisa.

O IFR , criado por Wilder Jr. (1978), é um oscilador que mostra a força e a velocidade do preço de um ativo pela comparação dos movimentos de alta ou de baixa dos preços de fechamento consecutivos.

$$FR = \frac{\text{Média de } n \text{ dias de fechamento em alta}}{\text{Média de } n \text{ dias de fechamento em baixa}}, \quad (3.7)$$

onde FR significa Força Relativa.

$$IFR = 100 - 100 \frac{1}{1 + FR}. \quad (3.8)$$

3.6. Procedimentos de coleta e de análise dos dados

Neste estudo foi utilizado $n = 14$. [Wilder Jr. \(1978\)](#) o criador do índice, atribui o valor 14 a n e [Dunis et al. \(2013\)](#) também utiliza $n = 14$ em seu trabalho. O IFR varia de 0 a 100, mas esse intervalo foi normalizado entre 0 e 1, a fim de se usar na MSV .

A $MMCD$, criada por Gerald Appel na segunda metade do século XX, é definida como a diferença de duas Médias Móveis Exponenciais (MME) e é comumente utilizada para prever tendências de mercado. Introduzimos primeiramente o conceito de MME :

$$MME_n(i) = \alpha \times p(i) + (1 - \alpha) \times MME_n(i - 1), \quad (3.9)$$

$$\alpha = \frac{2}{1 + n}, \quad (3.10)$$

sendo $p(i)$ o preço do ativo no dia i e n é o número de períodos.

Devido à contínua evolução nos mercados, os dados antigos podem representar padrões de comportamento que não possuem relevância no presente, por isso a MME dá pesos maiores para dados mais recentes e pesos menores para dados mais antigos.

$MMCD(i)$ é a subtração de uma MME de período curto de uma de período longo:

$$MMCD(n) = \sum_{i=1}^n MME_f(i) - \sum_{i=1}^n MME_s(i), \quad (3.11)$$

$f(fast)$ é usualmente igual a 12 períodos (curto) e $s(slow)$ e igual a 26 períodos (longo) ([Kara et al., 2011](#); [Rosillo et al., 2014](#)). Neste estudo é utilizado o histograma da $MMCD(i)$, denominado $MMCD_h(i)$ e calculado a partir do cruzamento da $MMCD(i)$ com uma $MME_g(i)$, chamada de gatilho:

$$MMCD_h(n) = \sum_{i=1}^n MME_f(i) - \sum_{i=1}^n MME_s(i) - \sum_{i=1}^n MME_g(i), \quad (3.12)$$

onde o gatilho, g , é igual a 9. A subtração do $MMCD(i)$ com o gatilho, tem o intuito de evitar uma falsa inversão de tendência. O intervalo do $MMCD_h(i)$ foi

normalizado entre 0 e 1, assim como no caso do IFR.

3.6.2 Metodologia da MSV

A metodologia do MSV começa com um problema simples, porém não trivial. Consiste em separar uma amostra de dados em dois grupos. Para isso criamos duas classes A e B, onde:

$$\begin{cases} y_i = 1 & \text{se } x_i \in \text{classe A} \\ y_i = -1 & \text{se } x_i \in \text{classe B.} \end{cases} \quad (3.13)$$

Para cada dado da amostra x_n temos uma classificação y_i , sua combinação é dada por (x_n, y_i) . Onde a amostra é representada da seguinte forma:

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_p, y_p) \quad (3.14)$$

No caso da Figura 3.2 as classes A e B seriam os pontos vermelhos e azuis. Agora dentre as possibilidades que podemos ver na Figura 3.2, como podemos estimar qual seria o melhor separador?

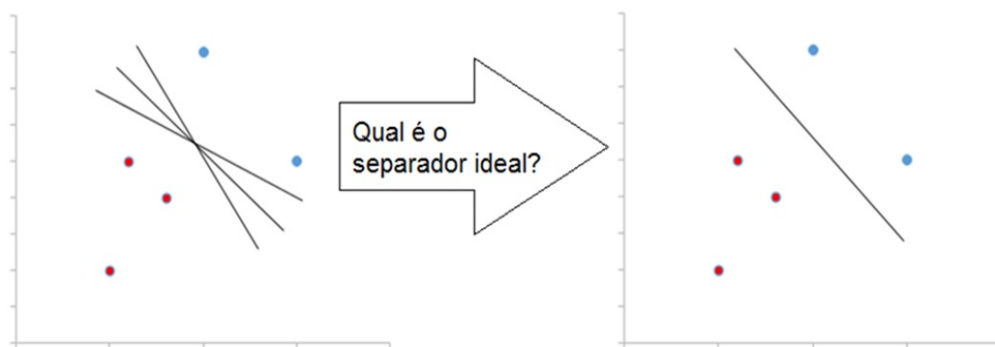


Figura 3.2: Separador ideal

3.6. Procedimentos de coleta e de análise dos dados

Respondendo a última pergunta, o melhor separador linear é aquele que maximiza a margem entre o separador e os pontos da amostra vide a Figura 3.3. A maximização da área é necessária para se reduzir a probabilidade de cometer um erro de classificação que veremos mais adiante. A margem é dada por M que é formada por w .

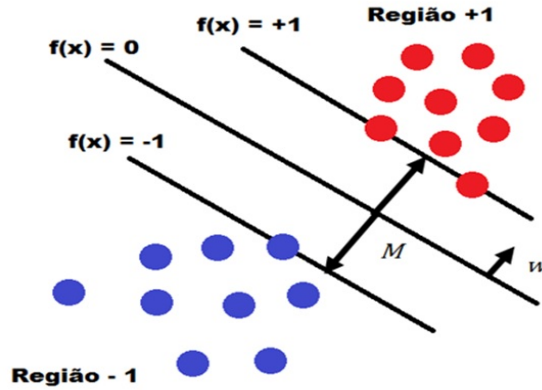


Figura 3.3: Separador de máxima margem

Com base no que foi dito, temos as seguintes formulações para a quantificação do problema. A margem é igual à:

$$M = \frac{2}{\|w\|}. \quad (3.15)$$

O problema de maximização é dado por:

$$w^* = \max_w \frac{2}{\|w\|}, \quad (3.16)$$

podendo ser escrito da seguinte forma:

$$w^* = \min_w \frac{1}{2} \|w\|^2, \quad (3.17)$$

onde o problema de minimização 3.17 é restrito a:

$$\begin{cases} y_i = 1 \Rightarrow \bar{w} \cdot \bar{x}_- + b \leq -1 \\ y_i = -1 \Rightarrow \bar{w} \cdot \bar{x}_+ + b \geq +1, \end{cases} \quad (3.18)$$

generalizando 3.18 temos

$$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 \geq 0, \quad (i = 1, 2, \dots, N),$$

$$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 = 0. \quad (3.19)$$

Os suportes vetoriais da máquina são os dados da amostra que restringem o tamanho da margem M . Na Figura 3.4 são os pontos onde $f(x) = +1$ e $f(x) = -1$.

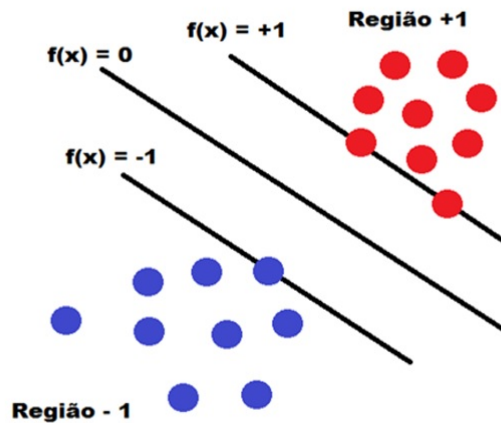


Figura 3.4: Suportes vetoriais

Para otimização do problema dentro das restrições dadas, aplicamos o Lagrangeano:

$$L(\bar{w}, b, \alpha) = \frac{1}{2} \|\bar{w}\|^2 - \sum \alpha_i [y_i(\bar{w} \cdot \bar{x}_i + b) - 1] \quad (3.20)$$

Dada a restrição 3.19 para o problema primal, temos:

3.6. Procedimentos de coleta e de análise dos dados

$$\begin{aligned}\frac{\partial L}{\partial \bar{w}} &= \bar{w} - \sum \alpha_i y_i x_i = 0, \\ \frac{\partial L}{\partial b} &= - \sum \alpha_i y_i = 0,\end{aligned}$$

tal que,

$$0 = \sum \alpha_i y_i, \quad (3.21)$$

$$\bar{w} = \sum \alpha_i y_i x_i. \quad (3.22)$$

Substituindo 3.22 em 3.20

$$= \frac{1}{2} \left(\sum \alpha_i y_i \bar{x}_i \right) \cdot \left(\sum \alpha_j y_j \bar{x}_j \right) - \left(\sum \alpha_i y_i x_i \right) \cdot \left(\sum \alpha_j y_j x_j \right) - \sum \alpha_i y_i b + \sum \alpha_i. \quad (3.23)$$

Resolvendo e simplificando temos a forma Dual de Wolfe,

$$L = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T \cdot x_j \quad (3.24)$$

3.6.3 MSV com Margem Suave

Outro fator importante para o funcionamento da MSV é o custo pelo erro de classificação. A MSV pode penalizar igualmente, ou não, cada tipo de erro de classificação, a exemplo de classificar um item como de classe A quando na verdade ele é de classe B. Isso é representado por $C \sum_{i=1}^N \xi_i$, onde C é uma constante positiva usada para expressar o *tradeoff* entre os erros de treinamento e o rigor da margem e $C \sum_{i=1}^N \xi_i$ é a quantidade de erros de treinamento. O problema de maximização da margem com a

Capítulo 3. Metodologia

adição do custo do erro é dado por:

$$\text{dual } w^* = \max_{\substack{\alpha \geq 0 \\ w, b, \xi \\ r \geq 0}} \min L(w, b, \xi, \alpha, r), \quad (3.25)$$

restrito:

$$y_i[w^T x_i + b] \geq \xi_i, \quad (\xi_i \geq 0, i = 1, 2, \dots, N). \quad (3.26)$$

Para otimização do problema dentro das restrições dadas, aplicamos o Lagrangeano:

$$L(w, b, \xi, \alpha, r) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum \alpha_i [y_i(w^T \cdot x_i + b) - 1 + \xi_i] - \sum r_i \xi_i \quad (3.27)$$

Dentro das restrições da equação 3.25 onde r_i e α_i , são os multiplicadores Lagrangeanos não-negativos. Resolvendo o Lagrangeano de Margem Suave temos:

$$\begin{aligned} \frac{\partial L}{\partial \bar{w}} &= \bar{w} - \sum \alpha_i y_i x_i = 0, \\ \frac{\partial L}{\partial b} &= - \sum \alpha_i y_i = 0, \\ \frac{\partial L}{\partial \xi_i} &= C - \alpha_i - \xi_i = 0, \end{aligned}$$

$$\xi_i \geq 0, \quad \alpha_i \geq 0, \quad r_i \geq 0. \quad (3.28)$$

Dadas a restrição 3.26 e as condições KKT (Karush-Kuhn-Tucker) 3.28 para o problema primal, temos

$$0 = \sum \alpha_i y_i, \quad (3.29)$$

3.6. Procedimentos de coleta e de análise dos dados

$$\bar{w} = \sum \alpha_i y_i x_i, \quad (3.30)$$

$$\alpha_i = C - r_i, \forall i. \quad (3.31)$$

Substituindo 3.30 em 3.27, resolvendo e simplificando temos a forma Dual de Wolfe de Margem Suave:

$$\max_{\alpha \geq 0} = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T \cdot x_j, \quad (3.32)$$

$$0 = \sum \alpha_i y_i, \quad (3.33)$$

$$0 \leq \alpha_i \leq C \forall i, C \in \mathbb{R}^+, \quad (3.34)$$

onde o parâmetro C limita a variação de α_i de zero a $+\infty$.

3.6.4 Método Kernel

O diferencial da MSV é a possibilidade de separar dados não linearmente separáveis em seu plano de origem a partir de uma técnica chamada *Kernel Trick*. Esta consiste em construir um hiperplano no qual a separação linear dos dados seja possível e a margem de separação entre as classes seja maximizada (Figura 3.5), isso é feito através de uma transformação $\varphi(x)$ tal que $\Phi : X \rightarrow F$ onde F é um espaço característico de X , sendo a função Kernel: $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$, $K : X \times X \rightarrow F$.

Qualquer função simétrica, positiva e semi-definida que satisfaça a condição de Mercer (1909) pode ser usada como uma função Kernel (Cortes e Vapnik, 1995; Smola e Scholkopf, 2004).

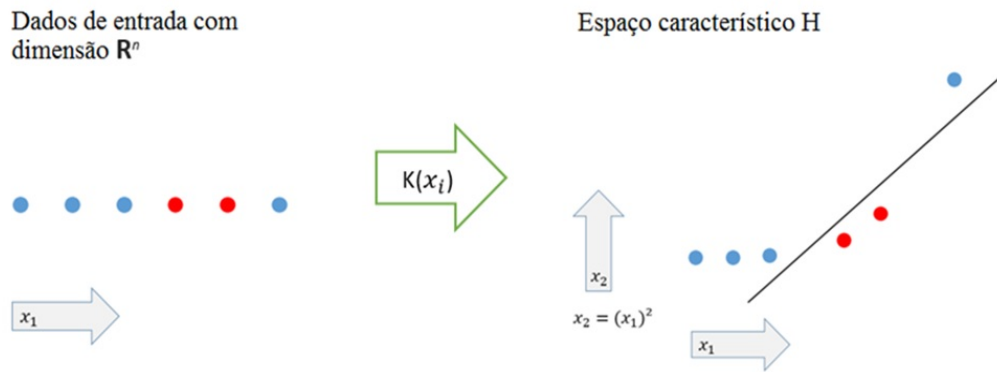


Figura 3.5: Dados de entrada não são linearmente separáveis, após a aplicação da função Kernel isso é possível

3.6.5 MSV Não-Linear com Margem Suave

Partindo da explicação acerca da função Kernel e de seu papel na MSV, o problema de maximização da margem é dado por:

$$\max \phi(w) = \frac{2}{\|w\|^2}, \quad (3.35)$$

que pode ser escrito da seguinte forma

$$\min \phi(w) = \frac{1}{2} \|w\|^2, \quad (3.36)$$

que é restrito a

$$\begin{cases} w^T \varphi(x_i) + b \geq 1 \text{ se } y_i = 1 \\ w^T \varphi(x_i) + b \leq -1 \text{ se } y_i = -1, \end{cases} \quad (3.37)$$

generalizando

$$y_i [w^T \varphi(x_i) + b] \geq \xi_i, \quad (\xi_i \geq 0, i = 1, 2, \dots, N)$$

3.6. Procedimentos de coleta e de análise dos dados

e adicionando o custo pelo erro

$$\min \phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i. \quad (3.38)$$

A formulação do problema de otimização com a restrição apresentada acima é dada pelo Lagrangeano:

$$L(w, b, \xi, \alpha, r) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum \alpha_i [y_i (w^T \cdot \phi(x_i) + b) - 1 + \xi_i] - \sum r_i \xi_i. \quad (3.39)$$

Onde α_i e r_i são multiplicadores de Lagrange não-negativos. Resolvendo e substituído em 3.10 os resultados obtidos no Lagrangeano, temos:

$$\begin{aligned} \frac{\partial L}{\partial \bar{w}} &= \bar{w} - \sum \alpha_i y_i x_i = 0, \\ \frac{\partial L}{\partial b} &= - \sum \alpha_i y_i = 0, \\ \frac{\partial L}{\partial \xi_i} &= C - \alpha_i - \xi_i = 0, \end{aligned}$$

$$\xi_i \geq 0, \alpha_i \geq 0, r_i \geq 0. \quad (3.40)$$

$$y_i [w^T \phi(x_i) + b] \geq 1 - \xi_i.$$

$$\alpha [y_i [w^T \phi(x_i) + b] - 1 + \xi_i] = 0. \quad (3.41)$$

$$r_i \xi_i = 0. \quad (3.42)$$

Capítulo 3. Metodologia

Consequentemente

$$\bar{w} = \sum_i \alpha_i y_i \varphi(x_j). \quad (3.43)$$

Para determinarmos o valor de b , usamos as equações 3.41, e 3.42 e as condições de complementaridade KKT

$$b = y_j - w^T \cdot \varphi(x_j). \quad (3.44)$$

É numericamente sensato tomar o valor médio de todos os b resultantes do cálculo,

$$b = \frac{1}{N_s} \sum_{0 < \alpha_i < C} [y_i - W^T \cdot \varphi(x_j)], \quad (3.45)$$

onde N_s é o número de Suportes Vetoriais. Para cada novo dado x , a função de classificação é dada por

$$f(x) = \text{Sinal}(w^T \phi(x) + b). \quad (3.46)$$

Substituindo as equações 3.43 e 3.45 em 3.46, temos a função de classificação final

$$f(x) = \text{Sinal} \left(\sum_{i=1}^N \alpha_i y_i \varphi(x_i)^T \varphi(x_j) + \frac{1}{N_s} \sum_{0 < \alpha_i < C} \left(y_i - \sum_{i=1}^N \alpha_i y_i \varphi(x_i)^T \varphi(x_j) \right) \right) \quad (3.47)$$

Dado que $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$, substituímos na equação 3.47

$$f(x) = \text{Sinal} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + \frac{1}{N_s} \sum_{0 < \alpha_i < C} \left(y_j - \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) \right) \right). \quad (3.48)$$

Para visualizar melhor a transformação φ a Figura 3.6 exemplifica essa situação

com uma amostra não-linearmente separável em seu espaço de origem.

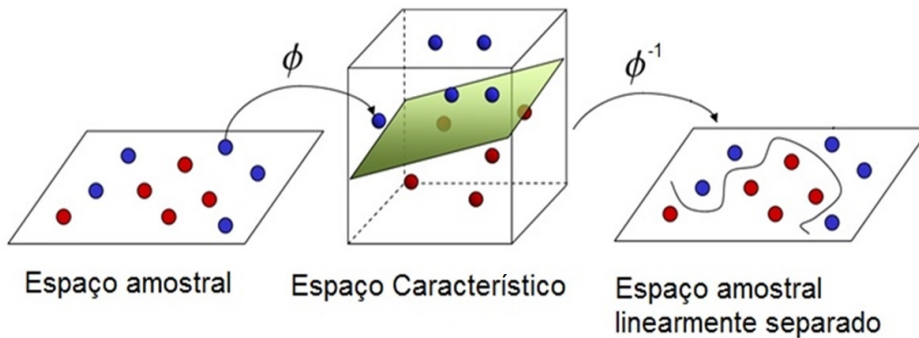


Figura 3.6: Truque Kernel

3.7 Parâmetros do SVM

Como discutido na seção 2.3 desse estudo, a performance de previsão da MSV é sensível aos parâmetros C e σ^2 (Kim, 2003). σ^2 é um parâmetro da função Kernel, neste caso, uma Função de Base Radial Gaussiana (FBR-G). O pacote "*kernlab*" do programa R criado por Karatzoglou et al. (2015) tem 8 Kernels inclusos, conforme a tabela 3.1.

3.7.1 Kernel e seus parâmetros

No começo foram escolhidos dois Kernels para serem utilizados, o Laplaciano e o FBR-G, em testes preliminares o Kernel FBR-G teve uma performance melhor e foi escolhido. Kim (2003), em seu trabalho, testa preliminarmente dois Kernels o FBR-G e o Polinomial, onde o primeiro fornece melhores resultados. Cao e Tay (2003) usam o Kernel FBR-G por culpa da sua não-linearidade e Smola e Scholkopf (2004) justificam seu uso em uma situação onde se conhece muito pouco sobre os dados fornecidos e se o tempo de computar a MSV é importante. Com isso em mente o Kernel escolhido para esse estudo foi o FBR-G.

Tabela 3.1: Tabela de Kernels do pacote "*kernlab*"

Nome do Kernel	Fórmulas
FBR-G	$k(x, x') = \exp(-\sigma \ x - x'\ ^2)$
Polinomial	$k(x, x') = (\text{escala} \langle x, x' \rangle + \text{offset})^{\text{grau}}$
Linear	$k(x, x') = \langle x, x' \rangle$
Tangente Hiperbólica	$k(x, x') = \tanh(\text{escala} \langle x, x' \rangle + \text{offset})$
Laplaciano	$k(x, x') = \exp(-\sigma \ x - x'\)$
Bessel	$k(x, x') = (-\text{Bessel}_{(v+1)}^n) \sigma \ x - x'\ ^2$
FBR ANOVA	$k(x, x') = \sum_{1 \leq i_1 \dots < i_D \leq N} \prod_{d=1}^D k(x_{i_d}, x'_{i_d})$, onde $k(x, x')$ é um kernel FBR-G.
Spline	$\prod_{d=1}^D \left(1 + x_i x_j + x_i x_j \min(x_i, x_j) - \frac{x_i + x_j}{2} \min(x_i, x_j)^2 + \frac{\min(x_i, x_j)^3}{3} \right)$

Retirado de [Karatzoglou et al. \(2015\)](#).

O parâmetro do Kernel FBR-G é o σ . Para descobrir o seu valor ótimo, foi criada uma grade de variação de parâmetros no código do R (apêndice [H.1](#)). Foram selecionados 100 números equidistantes entre si, pertencentes ao intervalo $0,000000001 \leq \sigma \leq 5$.

3.7.2 Parâmetro C

O parâmetro C, como vimos na seção [3.6.3](#) do presente estudo, é o custo pelo erro. [Cao e Tay \(2003\)](#) dizem em seu trabalho que um C muito grande afeta a MSV trazendo um problema de sobre-adaptação (*overfit*) a base de treinamento, levando a uma deterioração no desempenho de generalização e previsão da máquina. Na grade de variação de parâmetros foram selecionados 10 números equidistantes entre si, pertencentes ao intervalo $1 \leq C \leq 1000$.

3.8 Validação Cruzada

A função *ksvm* do pacote "*kernelab*" do programa R (Karatzoglou et al., 2015), possui um argumento chamado *cross*. Se um valor $K > 0$ é especificado, uma validação cruzada *k-fold* é realizada na base de treinamento para avaliar a qualidade do modelo: a taxa de precisão para a classificação e o erro médio quadrático para a regressão (Karatzoglou et al., 2015).

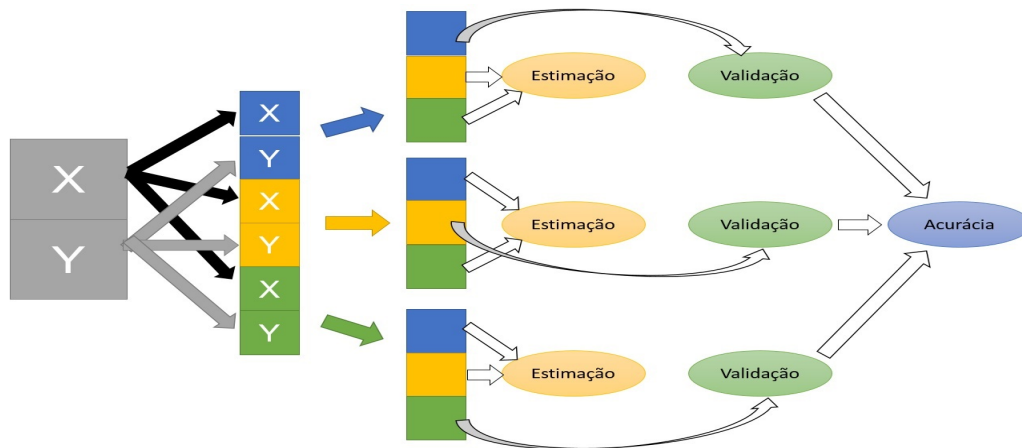


Figura 3.7: *K-fold* com $k=3$

Modificado de Argoncalves (2015).

Rodríguez et al. (2010) recomenda o uso de $k = 5$ ou $k = 10$, porque eles são menos tendencioso do que $k = 2$ e têm menos custo computacional que $k = n$. A *performance*, para este estudo, é calculada a partir de um *k-fold* onde o k escolhido é 5.

3.9 Dados de entrada da MSV

Os dados de entrada da MSV, como visto anteriormente, são os LR, *IFR* e *MMCD_h*. A MSV utiliza 5 dias de dados para realizar a previsão, por isso o número de observações é igual ao número de dias da base dividido por 5. A Tabela 3.2 apresenta o número de dias que a base possui e o número de observações.

Capítulo 3. Metodologia

Tabela 3.2: Números da base de dados

Nome do índice	número de dias	número de observações
IBOVESPA	3620	724
IBrX-50	3620	724
IEE	3620	724
INDX	3620	724
IFNC	2640	528
SMLL	1815	363

Como são 5 dias de LRs , $IFRs$ e $MMCD_{h,s}$ temos 15 variáveis para cada observação. A resposta para cada observação é dada pelo o retorno que o índice apresenta com o horizonte de 1 dia (H-1), ou com o horizonte de 5 dias (H-5). A Figura 3.8 exemplifica melhor a questão da resposta e dos índices de entrada.

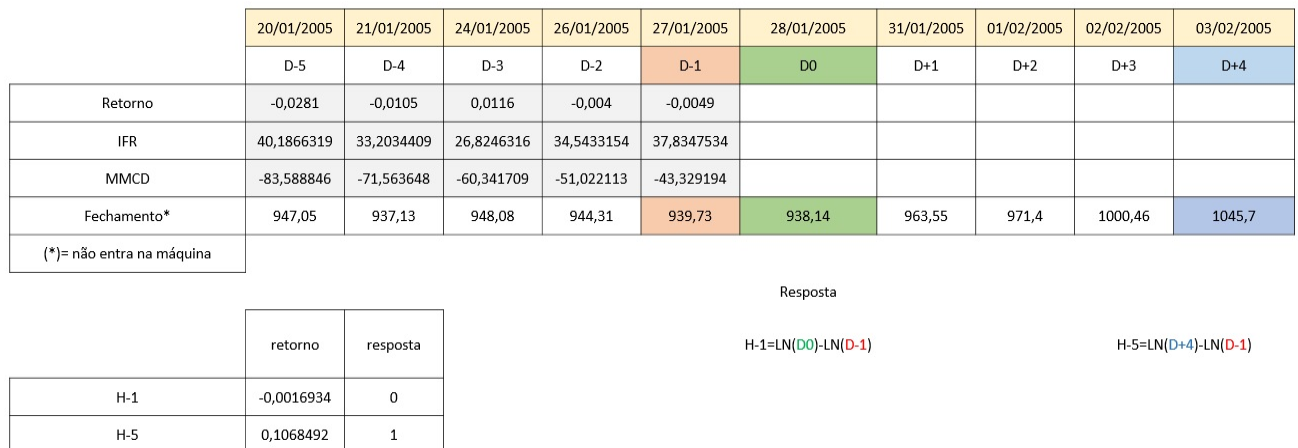


Figura 3.8: Organização dos *inputs* da MSV

A Figura 3.9 é um exemplo de como fica a forma final de entrada dos dados da MSV, para as 10 primeiras observações, onde a classe é a resposta que aquela observação tem de fornecer. Lembrando que a resposta de H-1 é diferente da de H-5.

3.9. Dados de entrada da MSV

	R1	R2	R3	R4	R5	IFR1	IFR2	IFR3	IFR4	IFR5	MMCD1	MMCD2	MMCD3	MMCD4	MMCD5	classe
1	-0.0281	-0.0105	0.0116	-0.0040	-0.0049	40.18663	33.20344	26.82463	34.54332	37.83475	-83.58884571	-71.5636478	-60.3417094	-51.02211296	-43.3291937	0
2	-0.0017	0.0267	0.0081	0.0295	0.0442	40.94087	40.53741	53.05476	57.83359	69.06167	-36.76731931	-29.3501741	-22.8184649	-15.68264253	-7.3332058	1
3	0.0499	-0.0039	-0.0096	0.0017	0.0273	73.49880	79.64024	78.87801	74.93560	73.25872	2.03051969	7.7641442	10.4128335	11.71663248	13.9551124	1
4	0.0188	-0.0218	0.0125	-0.0153	-0.0221	85.20883	90.06753	81.82815	83.95091	80.71213	16.09254186	14.9602553	14.2808014	11.80511372	7.7396269	1
5	0.0095	-0.0031	0.0438	-0.0033	-0.0168	75.04758	73.48339	71.84107	73.26604	67.38450	5.12416336	2.6431986	3.8855786	3.97486782	2.2642800	1
6	0.0083	0.0037	0.0267	0.0199	0.0152	54.21554	56.97895	60.37737	64.61482	63.45095	1.34292132	0.6357247	1.8797278	3.87119068	5.9147681	0
7	-0.0111	-0.0168	-0.0006	-0.0068	-0.0109	62.85697	66.03903	58.93508	63.26446	68.48100	5.77533225	3.7417341	1.8271741	-0.42420682	-3.1099026	0
8	-0.0102	0.0107	0.0075	-0.0202	-0.0274	62.89810	60.60469	52.48411	55.72014	54.59327	-5.90060976	-7.0215286	-7.2105908	-8.91773154	-12.0438261	0
9	-0.0121	-0.0244	0.0263	-0.0277	-0.0241	44.61343	40.80868	27.51306	29.82784	20.92631	-14.78071851	-17.9990105	-17.6744265	-18.91794744	-20.8039501	1
10	0.0194	0.0113	-0.0026	-0.0195	-0.0017	19.71962	27.98282	31.46590	32.03218	30.89584	-19.91839426	-17.7540959	-15.8082970	-15.26605535	-14.3791262	1

Figura 3.9: Forma final dos *inputs* da MSV

Capítulo 4

Resultados

4.1 *Performance* da MSV

As bases dos índices foram divididas conforme a Tabela 4.1. O *k-fold* e a *performance* medida, foram executados no intervalo de treinamento. Os gráficos a seguir (fig.4.1 e fig.4.2) mostram a relação entre os parâmetros C e σ e a *performance* atingida. As figuras estão organizadas conforme o índice (linhas) e o horizonte (colunas) em análise. Cumpre destacar que $0,000000001 \leq \sigma \leq 5$, $1 \leq C \leq 1000$ e $0 \leq performance \leq 1$

Tabela 4.1: Intervalos da base de dados

Índice	Interv. Treinamento	Interv. Validação	Interv. Teste	Total
IBOVESPA	[1, 362] = 50%	[363, 603] = 33%	[604, 724] = 17%	724 = 100%
IBrX-50	[1, 362] = 50%	[363, 603] = 33%	[604, 724] = 17%	724 = 100%
IEE	[1, 362] = 50%	[363, 603] = 33%	[604, 724] = 17%	724 = 100%
INDX	[1, 362] = 50%	[363, 603] = 33%	[604, 724] = 17%	724 = 100%
IFNC	[1, 264] = 50%	[265, 438] = 33%	[439, 528] = 17%	528 = 100%
SMLL	[1, 182] = 50%	[183, 274] = 25%	[275, 363] = 25%	363 = 100%

o Intervalo de Validação foi utilizado para treinar a máquina utilizando os melhores parâmetros (Tabela 4.2 e 4.3) para em seguida prever o Intervalo de Teste. A previsão é considerada da classe 1 caso a probabilidade de ocorrência dada pela máquina seja maior que 60%, caso contrário ela é classificada como da classe 0. Os resultados podem

Capítulo 4. Resultados

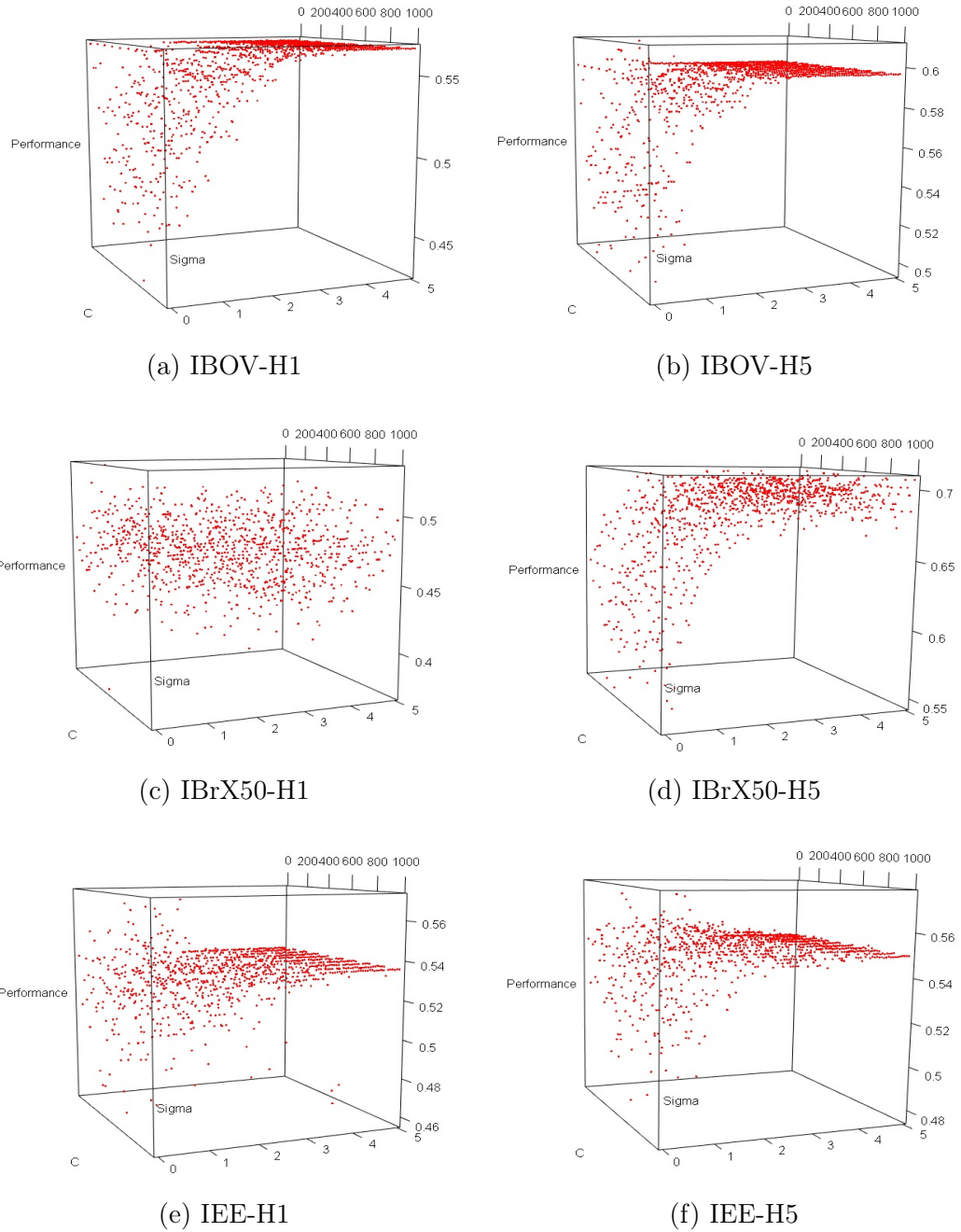
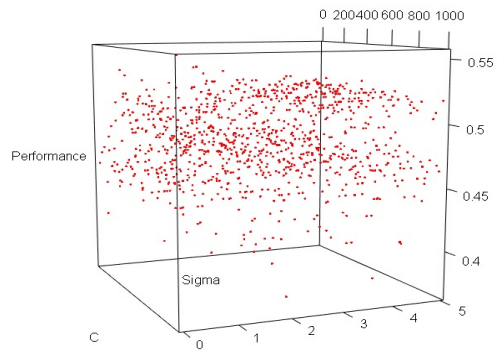


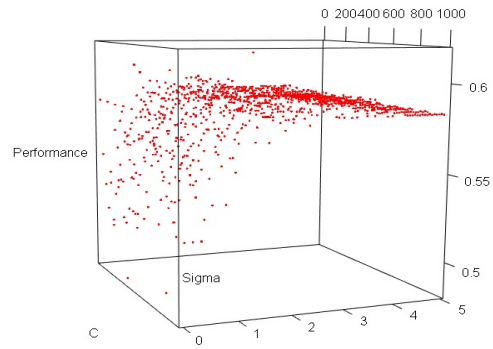
Figura 4.1: Gráficos de performance da MSV em relação aos parâmetros σ e C

ser vistos na Tabela de matrizes de confusão 4.4. Os valores preditos, pela máquina, são as linhas e os valores reais (ou a resposta) as colunas.

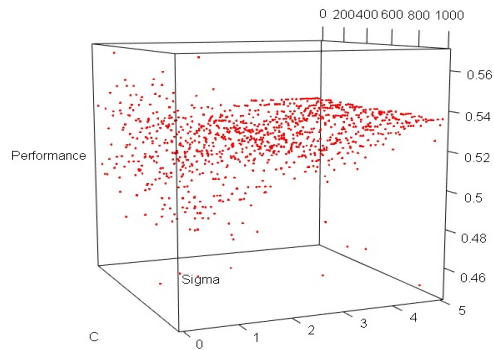
4.1. Performance da MSV



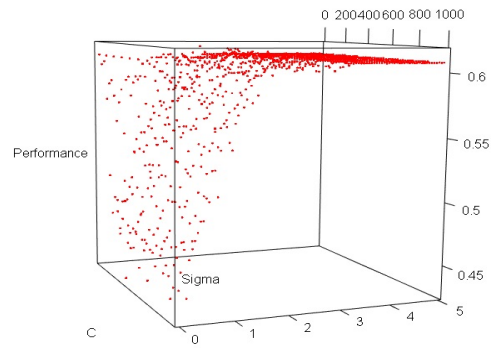
(a) IFNC-H1



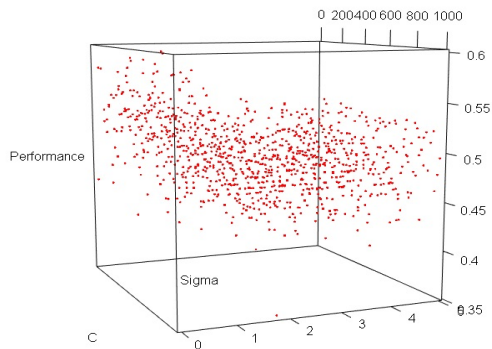
(b) IFNC-H5



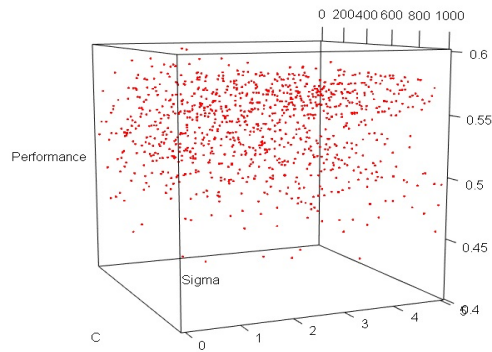
(c) INDX-H1



(d) INDX-H5



(e) SMLL-H1



(f) SMLL-H5

Figura 4.2: Gráficos de performance da MSV em relação aos parâmetros σ e C

Capítulo 4. Resultados

Tabela 4.2: Melhores parâmetros de C e σ para H-1

Índice	Melhor C	Melhor σ	<i>Performance</i>
IBOVESPA	667	3,4343	0,5667
IBrX-50	112	0,5050	0,5334
IEE	889	0,7070	0,5718
INDX	1000	0,3535	0,5689
IFNC	778	0,0505	0,5648
SMLL	445	0,6565	0,5992

Tabela 4.3: Melhores parâmetros de C e σ para H-5

Índice	Melhor C	Melhor σ	<i>Performance</i>
IBOVESPA	112	1,2121	0,61605
IBrX-50	556	2,8282	0,7076
IEE	1000	0,2525	0,5772
INDX	556	1,7171	0,6160
IFNC	1000	1,2626	0,6177
SMLL	223	1,4141	0,5987

Tabela 4.4: Matriz de confusão da predição da MSV

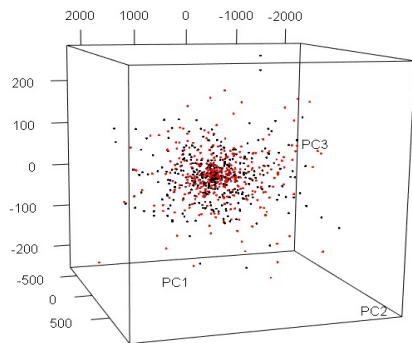
		H-1		H-5	
IBOVESPA		0	1	0	1
	0	63	58	58	62
IBrX-50	1	0	0	1	0
		0	1	0	1
IEE	0	67	54	58	61
	1	0	0	0	2
INDX		0	1	0	1
	0	60	61	54	43
IFNC	1	0	0	8	16
		0	1	0	1
SMLL	0	49	55	61	60
	1	11	6	0	0
		0	1	0	1
	0	43	44	49	39
	1	0	2	0	1

4.2 Análise de Componente Principal e Hiperplano Separador

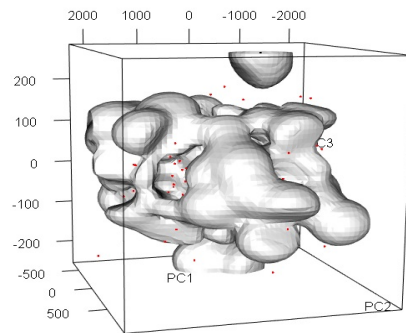
Como as observações possuem 15 variáveis, seria impossível plotar um gráfico para auxiliar a visualização do problema. Para contornar esse problema foi feita uma Análise de Componente Principal (ACP), com o intuito de reduzir o número de variáveis para um número que fosse possível plotar os gráficos. Feita a análise constatamos que a perda de informação após o terceiro Componente Principal, para todas as séries, era irrelevante.

A partir da ACP e da escolha dos primeiros 3 Componentes Principais, podemos plotar o gráfico da distribuição dos pontos de descida (0 - preto) subida (1 - vermelho) e a função separadora (hiperplano separador) da MSV, conforme as Figuras [4.3](#) e [4.4](#). Os parâmetros C e σ foram escolhidos conforme as Tabelas [4.2](#) e [4.3](#)

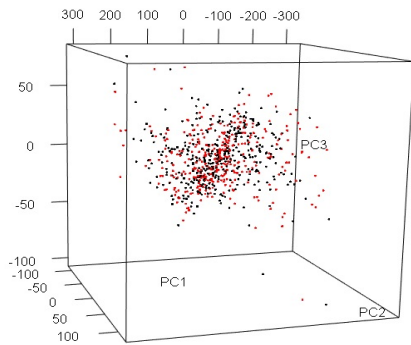
Capítulo 4. Resultados



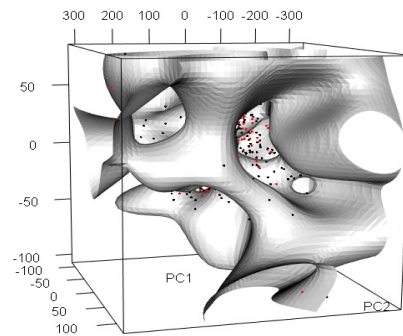
(a) IBOV-H1



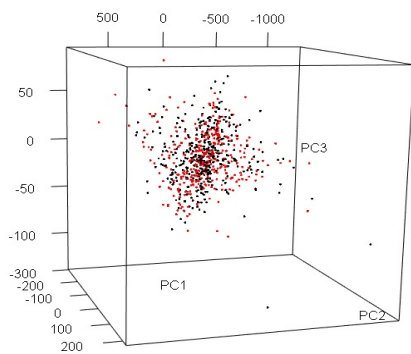
(b) MSV IBOV-H1



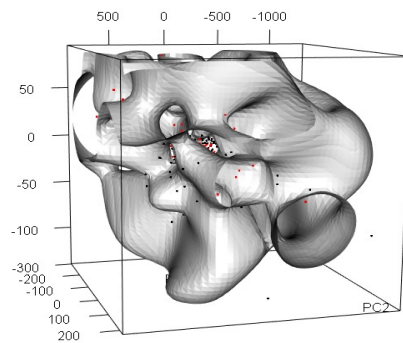
(c) IBRX-H1



(d) MSV IBRX-H1



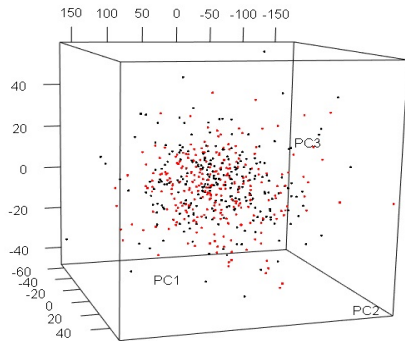
(e) IEE-H1



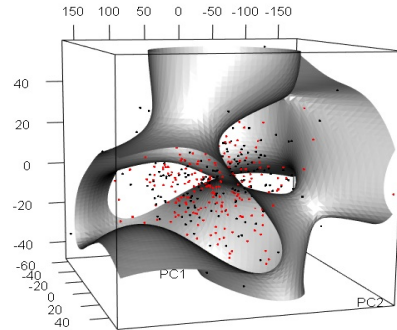
(f) MSV IEE-H1

Figura 4.3: Gráficos de pontos de subida (vermelho) e descida (preto) e seus respectivos separadores lineares (Hiperplano)

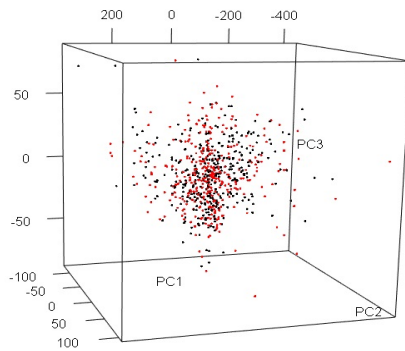
4.2. Análise de Componente Principal e Hiperplano Separador



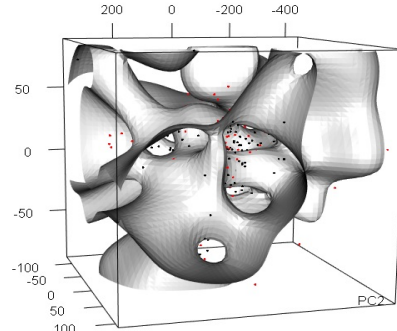
(a) IFNC-H1



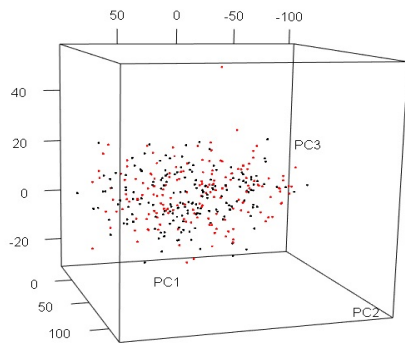
(b) MSV IFNC-H1



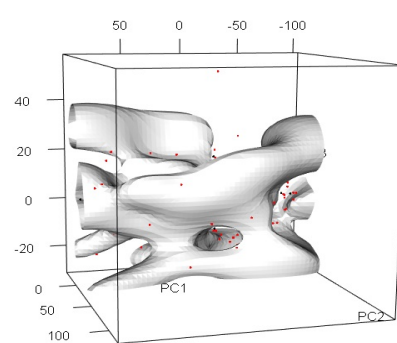
(c) INDX-H1



(d) MSV INDX-H1



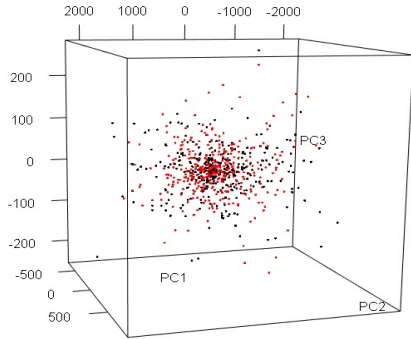
(e) SMLL-H1



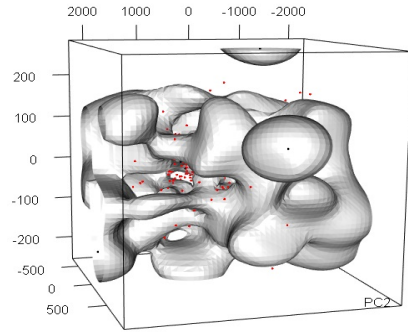
(f) MSV SMLL-H1

Figura 4.4: Gráficos de pontos de subida (vermelho) e descida (preto) e seus respectivos separadores lineares (Hiperplano)

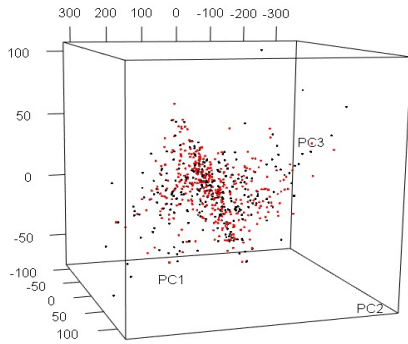
Capítulo 4. Resultados



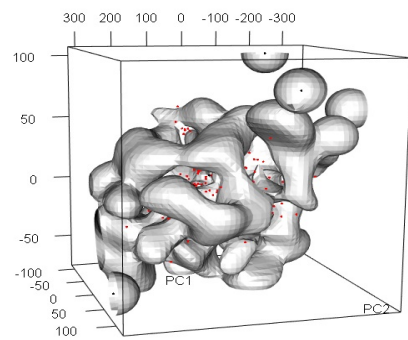
(a) IBOV-H5



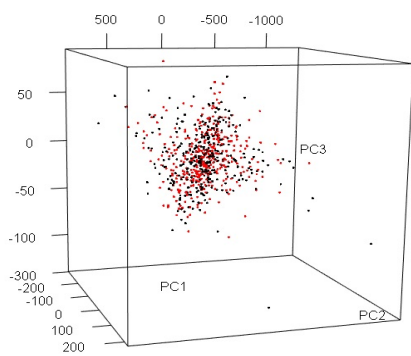
(b) MSV IBOV-H5



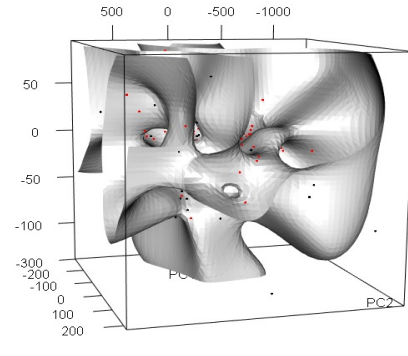
(c) IBRX-H5



(d) MSV IBRX-H5



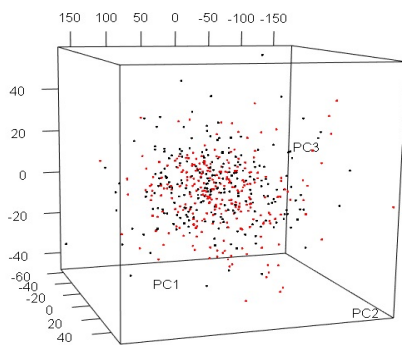
(e) IEE-H5



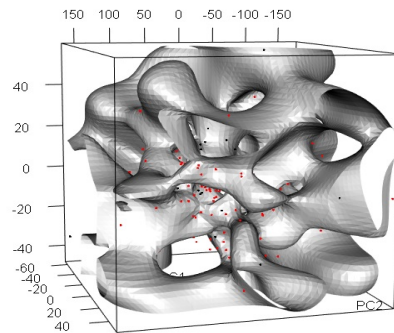
(f) MSV IEE-H5

Figura 4.5: Gráficos de pontos de subida (vermelho) e descida (preto) e seus respectivos separadores lineares (Hiperplano)

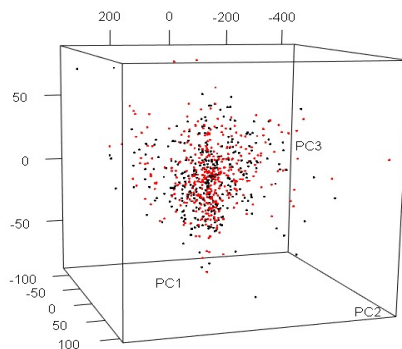
4.2. Análise de Componente Principal e Hiperplano Separador



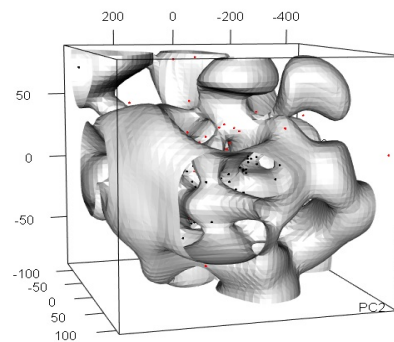
(a) IFNC-H5



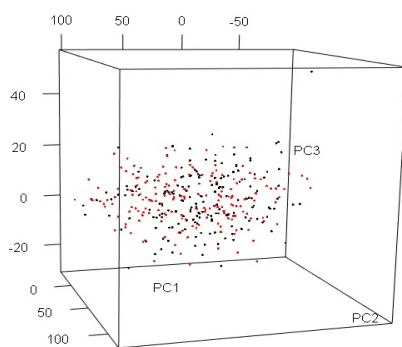
(b) MSV IFNC-H5



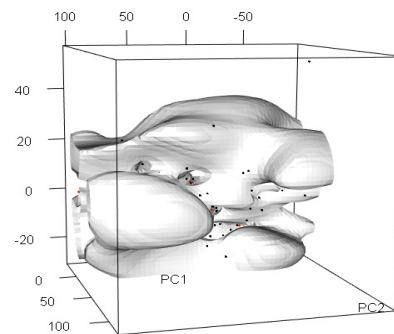
(c) INDX-H5



(d) MSV INDX-H5



(e) SMLL-H5



(f) MSV SMLL-H5

Figura 4.6: Gráficos de pontos de subida (vermelho) e descida (preto) e seus respectivos separadores lineares (Hiperplano)

Capítulo 5

Considerações finais

Esta pesquisa replicou, com alterações, o modelo de Máquinas de Suporte Vetorial proposto por [Dunis et al. \(2013\)](#), para a previsão do movimento de índices, aplicada ao contexto brasileiro. As MSVs foram utilizadas para apurar até que ponto é possível prever a movimentação dos índices.

A amostra desta pesquisa é formada pelos índices IBOVESPA, IBrX-50, IEE, IFNC, INDX e SMLL. Os recortes temporais vão de 22 de janeiro de 2001 até 07 de agosto de 2015 para o IBOVESPA, 22 de Janeiro de 2001 até 01 de setembro de 2015 para o IBrX-50, IEE e INDX, 20 de janeiro de 2005 até 21 de setembro de 2015 para o IFNC e 21 de maio de 2008 até 22 de setembro de 2015 para o SMLL. Os dados dos Índices foram coletados no sistema Economática. A base de dados foi dividida em três conjuntos com recortes temporais diferentes, porém em sequência, conforme a [Tabela 4.1](#).

A MSV foi utilizada para classificar o movimento dos índices em duas classes: classe 0 ($y = -1$), movimento de queda, e classe 1 ($y = +1$), movimento de alta. Na primeira etapa o Intervalo de Treinamento foi utilizado para calcular os Melhores parâmetros utilizando um *k-fold* de $k = 5$ para medir a acurácia. Na segunda etapa os Intervalos de Validação e Teste serviram para testar a previsibilidade da máquina. Lembrando

Capítulo 5. Considerações finais

que a previsão, nessa segunda etapa, é considerada da classe 1 caso a probabilidade de ocorrência dada pela máquina seja maior que 60%, caso contrário ela é classificada como da classe 0.

Esta pesquisa favorece a utilização de MSVs para a previsão de séries temporais financeiras. A melhor performance obtida para o modelo foi de 70,76% para o IBrX-50, como um horizonte de 5 dias, onde o $C = 556$ e $\sigma = 2,8282$. A máquina utiliza dados de IFR, MMCD e LR de 5 dias anteriores a data focal da previsão, totalizando 15 variáveis para cada dia previsto.

Os resultados desta pesquisa vão ao encontro dos seguintes estudos: [Kim \(2003\)](#) que em seu trabalho obteve uma taxa de acerto que varia na casa de 60%; [Dunis et al. \(2013\)](#) obtêm resultados superiores ao deste estudo, próximos a 70% de acerto, contudo eles utilizam um kernel diferente além de testarem diversos intervalos de treinamento e teste; [Huang et al. \(2005\)](#) conseguem um resultado de 73% de taxa de acerto, utilizando outros tipos de dados de entrada, que não IATs; [Kara et al. \(2011\)](#) alcançam uma taxa de acerto de 71,52%, utilizando períodos de teste mais curtos e testando o kernel FBR-G e o polinomial.

Os melhores parâmetros obtidos nas Tabelas 4.2 e 4.3, não tiveram um bom desempenho para prever o intervalo de Teste, como pode ser observado na Tabela 4.4, isso pode ser devido a forma como foi calculado o melhor parâmetro. Como podemos observar nas Figuras 4.1 e 4.2 existem *outliers* que podem ter sido escolhidos como os melhores parâmetros. Podemos afirmar que com a metodologia e os parâmetros adotados, não é possível prever o movimento dos índices com exatidão.

Para estudos futuros, uma saída para o problema dos *outliers*, seria calcular os melhores parâmetros utilizando Otimização por Enxame de Partículas (em inglês *Particle Swarm Optimization*), que leva em conta um conjunto de pontos para selecionar os melhores parâmetros. Os Intervalos de Treinamento, possuem influencia segundo [Dunis et al. \(2013\)](#). Outras sugestões para estudos futuros seria explorar a utilização de

Intervalos mais curtos, que podem resultar em maiores taxas de acerto, e a aplicação dessa metodologia a vários Kernels distintos, para observar suas respectivas *Performances*. Lembrando que a função Kernel é fator determinante para o desempenho da MSV, podendo aumentar significativamente a performance da máquina.

Referências Bibliográficas

Abu-Mostafa, Y. S. e Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, 6(3):205–213.

Argoncalves (2015). k-fold. <https://pt.wikipedia.org/wiki/Ficheiro:Kfold.pdf>. Acesso em: 03/10/2015.

BMF&BOVESPA (2013). Manual de definições e procedimentos dos índices da bm&f bovespa. <http://www.bmfbovespa.com.br/Indices/download/Manual-de-Definicoes-e-Procedimentos-dos-Indices-da-BMFBOVESPA-R.pdf>. Acesso em: .

BMF&BOVESPA (2015a). Metodologia do índice bm&fbovespa energia elétrica (iee). <http://www.bmfbovespa.com.br/Indices/download/IEE-Metodologia-pt-br.pdf>. Acesso em: 25/10/2015.

BMF&BOVESPA (2015b). Metodologia do índice bm&fbovespa financeiro (ifnc). <http://www.bmfbovespa.com.br/Indices/download/IFNC-Metodologia-pt-br.pdf>. Acesso em: 25/10/2015.

BMF&BOVESPA (2015c). Metodologia do índice bm&fbovespa small cap (smll). <http://www.bmfbovespa.com.br/Indices/download/SMLL-Metodologia-pt-br.pdf>. Acesso em: 25/10/2015.

- BMF&BOVESPA (2015d). Metodologia do índice brasil 50 (ibrx 50). <http://www.bmfbovespa.com.br/Indices/download/IBXL-Metodologia-pt-br.pdf>. Acesso em: 25/10/2015.
- BMF&BOVESPA (2015e). Metodologia do índice do setor industrial (indx). <http://www.bmfbovespa.com.br/Indices/download/INDX-Metodologia-pt-br.pdf>. Acesso em: 25/10/2015.
- BMF&BOVESPA (2015f). Tabela de composição da carteira teórica que compõe o ibovespa, 2º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoCarteiraTeorica.aspx?Indice=Ibovespa&idioma=pt-br>. Acesso em: 08/07/2015.
- BMF&BOVESPA (2015g). Tabela de composição da carteira teórica que compõe o ibrx-50, 3º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoCarteiraTeorica.aspx?Indice=IBrX50&idioma=pt-br>. Acesso em: 25/10/2015.
- BMF&BOVESPA (2015h). Tabela de composição da carteira teórica que compõe o iee, 3º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoCarteiraTeorica.aspx?Indice=IEE&idioma=pt-br>. Acesso em: 25/10/2015.
- BMF&BOVESPA (2015i). Tabela de composição da carteira teórica que compõe o ifnc, 3º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx?Indice=IFNC&Idioma=pt-br>. Acesso em: 25/10/2015.
- BMF&BOVESPA (2015j). Tabela de composição da carteira teórica que compõe o indx, 3º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx?Indice=INDX&Idioma=pt-br>. Acesso em: 25/10/2015.

Referências Bibliográficas

- BMF&BOVESPA (2015k). Tabela de composição da carteira teórica que compõe o smll, 3º quadrimestre de 2015. <http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx?Indice=SMLL&Idioma=pt-br>. Acesso em: 25/10/2015.
- Boser, B. E., Guyon, I. M., e Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*.
- Brock, W., Lakonishok, J., e Lebaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, 47(5):1731–1764.
- Brown, S. J., Goetzmann, W. N., e Kumar, A. (1998). The dow theory: William peter hamilton’s track record re-considered. Working paper series, New York University.
- Cao, L. e Tay, F. E. H. (2001). Financial forecasting using support vector machines. *Neural Computing Applications*, 10(2):184–192.
- Cao, L. e Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE transactions on neural networks*, 14(6):1506–1518.
- Carvalho, L. G. P. d., Costa Jr., N. C. A. d., e Goulart, M. A. d. O. V. (2008). Análise técnica versus hipótese dos mercados eficientes: Um estudo utilizando o indicador macd. *Revista Alcance*, 15(3):398–416.
- Chen, N. (1991). Financial investment opportunities and the macroeconomy. *The Journal of Finance*, 46(2):529–554.
- Cortes, C. e Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.
- Courtault, J.-M., Kabanov, Y., Bru, B., Crepel, P., Lebon, I., e Marchand, A. L. (2000).

- Louis bachelier on the centenary of théorie de la spéculation. *Mathematical Finance*, 10(3):339–353.
- Curcio, R., Goodhart, C., Guillaume, D., e Payne, R. (1997). Do technical trading rules generate profits? conclusions from the intra-day foreign exchange market. *International Journal of Finance & Economics*, 2(4):267–280.
- de Paula, A. e Triches, D. (2010). O mercado acionário e o desempenho dos ativos financeiros no brasil com emprego da análise técnica. *Perspectiva Econômica*, 6(1):37–62.
- Dunis, C. L., Rosillo, R., delaFuente, D., e R., P. (2013). Forecasting ibex-35 moves using support vector machines. *Neural Computing and Applications*, 23:229–236.
- Fama, E. (1965). Behavior of stock market price. *The Journal of Business*, 38(1):34–105.
- Fama, E. (1970). Capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):28–30.
- Fan, A. e Palaniswami, M. (2001). Stock selection using support vector machines. International Joint Conference on Neural Networks.
- Gençay, R. (1998). The predictability of security returns with simple technical trading rules. *Journal of Empirical Finance*, 5(4):347–359.
- Gold, S. C. e Lebowitz, P. (1999). Computerized stock screening rules for portfolio selection. *Financial Services Review*, 8(2):61–70.
- Huang, W., Nakamori, Y., e Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machines. *Computes & Operations Research*, 32(10):2513–2522.

Referências Bibliográficas

- Ince, H. e Trafalis, T. B. (2004). Kernel principal component analysis and support vector machines for stock price prediction. *Institute of Industrial Engineers*, 39(6):1793–1798.
- Ince, H. e Trafalis, T. B. (2006). Kernel methods for short-term portfolio management. *Expert Systems with Applications*, 30(3):535–542.
- Ince, H. e Trafalis, T. B. (2008). Short term forecasting with support vector machines and applications to stock price prediction. *International Journal of General Systems*, 37(6):677–687.
- Jain, P. K. (2005). Financial market design and the equity premium: Electronic versus floor trading. *The Journal of Finance*, 60(6):2955–2985.
- Kara, Y., Boyacioglu, M. A., e Baykan, Ö. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert Systems with Applications*, 38(5):5311–5319.
- Karatzoglou, A., Smola, A., e Hornik, K. (2015). Package ‘kernlab’. <https://cran.r-project.org/web/packages/kernlab/kernlab.pdf>. Acesso em: 05/10/2015.
- Keim, D. B. e Stambaugh, R. F. (1986). Predicting returns in the stock and bonds markets. *Journal of Financial Economics*, 17(2):357–390.
- Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1):307–319.
- Lo, A. W. e MacKinley, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specifications test. *The Review of Financial Studies*, 1(1):41–66.

- Lo, A. W., Mamaysky, H., e Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *the Journal of Finance*, 55(4):1705–1765.
- Menkhoff, L. e Taylor, M. P. (2007). The obstinate passion of foreign exchange professionals: Technical analysis. *Journal of Economic Literature*, 45:936–972.
- Mobarek, A. e Fiorante, A. (2014). The prospects of brics countries: Testing weak-form market efficiency. *Reaserch in International Business and Finance*, 30(1):217–232.
- Neal, L. D. e D., W. M. (2003). Crises in the global economy from tulips to today. In of Economic Research, N. B., editor, *Globalization in Historical Perspective*. University of Chicago Press.
- Neftci, S. N. (1991). Naive trading rules in financial markets and wiener-kolmogorov prediction theory: A study of "technical analysis". *The Journal of Business*, 64(4):549–571.
- Ojah, K. e Karamera, D. (1999). Random walks and market efficiency tests of latin american emerging equity markets: A revisit. *The Financial Review*, 34(2):52–72.
- Osler, C. (2004). Support for resistance: Technical analysis and intraday exchange rates. *Federal Reserve Bank of New York Economic Policy Review*, 6(2):53–68.
- Park, C. H. e Irwin, S. H. (2004). The profitability of technical analysis: A review. Agmas project research report, University of Illinois.
- Park, C. H. e Irwin, S. H. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, 21(4):786–826.
- Rodríguez, J. D., Pérez, A., e Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 32(3):569–575.

Referências Bibliográficas

- Rosillo, R., Giner, J., e delaFuente D. (2014). The effectiveness of the combined use of vix and support vector machines on the prediction of s&p 500. *Neural Computing & Applications*, 25:321–332.
- Smola, A. J. e Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222.
- Stephens, A. A., Atwater, J. B., e Kannan, V. R. (2013). From tulip bulbs to sub-prime mortgages examining the sub-prime crisis: The case for a systemic approach. *The learning Organization*, 20(1):65–84.
- Sullivan, R., Timmermann, A., e White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance*, 54(5):1647–1691.
- Taylor, N. (2014). The rise and fall of technical trading rule success. *Journal of Banking & Finance*, 40(1):286–302.
- Tversky, A. e Kahneman, D. (1986). Part 2: The behavioral foundations of economic theory. *The journal of Business*, 59(4):251–278.
- Vidotto, R. S., Migliato, A. L. T., e Zambon, A. C. (2009). O moving average convergence-divergence como ferramenta para a decisão de investimentos no mercado de ações. *Revista de Administração Contemporânea*, 13(2):291–309.
- Wilder Jr., J. W. (1978). The relative strength index. *Stock & Commodities*, 4(9):343–346.

Apêndice A

Composição da carteira teórica

IBOVESPA

Tabela A.1: Tabela de composição da carteira teórica que compõe o IBOVESPA 2º quadrimestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ABEV3	AMBEV S/A	ON	3.608.712.145	7,992
2	BBAS3	BRASIL	ON NM	828.582.888	2,264
3	BBDC3	BRADESCO	ON EJ N1	458.568.908	1,476
4	BBDC4	BRADESCO	PN EJ N1	2.431.329.422	8,006
5	BBSE3	BBSEGURIDADE	ON NM	675.000.000	2,653
6	BRAP4	BRADESPAR	PN N1	222.476.686	0,258
7	BRFS3	BRF SA	ON EJ NM	557.128.268	4,213
8	BRKM5	BRASKEM	PNA N1	264.665.622	0,38
9	BRML3	BR MALLS PAR	ON NM	459.540.434	0,756
10	BRPR3	BR PROPERT	ON NM	298.228.434	0,353
11	BVMF3	BMFBOVESPA	ON NM	1.797.987.343	2,377

Apêndice A. Composição da carteira teórica IBOVESPA

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
12	CCRO3	CCR SA	ON NM	861.282.756	1,52
13	CESP6	CESP	PNB N1	181.791.331	0,423
14	CIEL3	CIELO	ON NM	772.547.017	3,998
15	CMIG4	CEMIG	PN N1	758.933.452	1,017
16	CPFE3	CPFL ENERGIA	ON NM	302.642.915	0,665
17	CPLE6	COPEL	PNB N1	100.948.359	0,412
18	CRUZ3	SOUZA CRUZ	ON	378.068.610	1,059
19	CSAN3	COSAN	ON NM	132.898.602	0,365
20	CSNA3	SID NACIONAL	ON	702.056.615	0,387
21	CTIP3	CETIP	ON EJ NM	253.479.395	1,025
22	CYRE3	CYRELA REALT	ON NM	250.463.857	0,284
23	DTEX3	DURATEX	ON NM	261.597.936	0,223
24	ECOR3	ECORODOVIAS	ON NM	198.534.559	0,174
25	ELET3	ELETROBRAS	ON N1	223.455.547	0,151
26	ELET6	ELETROBRAS	PNB N1	219.731.566	0,216
27	EMBR3	EMBRAER	ON NM	734.970.461	1,973
28	ENBR3	ENERGIAS BR	ON NM	232.602.924	0,313
29	ESTC3	ESTACIO PART	ON NM	274.067.082	0,562
30	FIBR3	FIBRIA	ON NM	222.307.680	1,098
31	GFSA3	GAFISA	ON NM	365.371.997	0,098
32	GGBR4	GERDAU	PN N1	876.090.759	0,674
33	GOAU4	GERDAU MET	PN N1	268.577.413	0,172
34	GOLL4	GOL	PN N2	105.770.630	0,082
35	HGTX3	CIA HERING	ON NM	127.537.447	0,178
36	HYPE3	HYPERMARCAS	ON NM	371.533.847	0,935

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
37	ITSA4	ITAUSA	PN N1	3.427.012.618	3,518
38	ITUB4	ITAUUNIBANCO	PN ED N1	2.680.430.933	10,68
39	JBSS3	JBS	ON NM	1.653.729.368	3,059
40	KLBN11	KLABIN S/A	UNT N2	382.961.180	0,844
41	KROT3	KROTON	ON NM	1.349.648.918	1,847
42	LAME4	LOJAS AMERIC	PN	428.103.759	0,882
43	LREN3	LOJAS RENNER	ON NM	126.711.043	1,726
44	MRFG3	MARFRIG	ON NM	346.968.736	0,213
45	MRVE3	MRV	ON NM	272.744.763	0,244
46	MULT3	MULTIPLAN	ON EJ N2	75.442.416	0,417
47	NATU3	NATURA	ON NM	168.827.352	0,544
48	OIBR4	OI	PN N1	323.368.383	0,206
49	PCAR4	P.ACUCAR-CBD	PN N1	155.497.607	1,287
50	PETR3	PETROBRAS	ON	2.708.485.740	4,05
51	PETR4	PETROBRAS	PN	4.037.414.651	5,468
52	POMO4	MARCOPOLO	PN N2	538.014.578	0,141
53	QUAL3	QUALICORP	ON NM	218.349.295	0,542
54	RENT3	LOCALIZA	ON EJ NM	149.142.340	0,522
55	RUMO3	RUMO LOG	ON NM	1.237.835.430	0,19
56	SANB11	SANTANDER BR	UNT	400.076.102	0,768
57	SBSP3	SABESP	ON NM	339.985.611	0,636
58	SMLE3	SMILES	ON EJ NM	56.172.745	0,339
59	SUZB5	SUZANO PAPEL	PNA N1	461.173.399	0,895
60	TBLE3	TRACTEBEL	ON NM	203.919.927	0,815
61	TIMP3	TIM PART S/A	ON NM	807.753.067	0,961

Apêndice A. Composição da carteira teórica IBOVESPA

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
62	UGPA3	ULTRAPAR	ON NM	341.884.973	2,597
63	USIM5	USIMINAS	PNA N1	481.427.177	0,224
64	VALE3	VALE	ON N1	1.262.807.257	2,578
65	VALE5	VALE	PNA N1	1.878.758.554	3,303
66	VIVT4	TELEF BRASIL	PN	347.269.914	1,772
	Total			47.641.400.743	100
	Redutor			16.490.815,14	

(BMF&BOVESPA, 2015f)

Apêndice B

Composição da carteira teórica

IBrX-50

Tabela B.1: Tabela de composição da carteira teórica que compõe o IBrX-50 3º quadrimestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ABEV3	AMBEV S/A	ON	4.394.000.710,00	10,817
2	BBAS3	BRASIL	ON NM	839.653.556,00	1,743
3	BBDC3	BRADESCO	ON N1	663.958.238,00	1,985
4	BBDC4	BRADESCO	PN N1	2.431.329.422,00	6,524
5	BBSE3	BBSEGURIDADE	ON NM	674.972.311,00	2,391
6	BRAP4	BRADESPAR	PN N1	221.328.102,00	0,236
7	BRFS3	BRF SA	ON NM	832.081.344,00	7,192
8	BRKM5	BRASKEM	PNA N1	264.585.622,00	0,68
9	BRML3	BR MALLS PAR	ON NM	462.652.489,00	0,68
10	BVMF3	BMFBOVESPA	ON NM	1.778.324.029,00	2,668
11	CCRO3	CCR SA	ON NM	861.253.436,00	1,377

Apêndice B. Composição da carteira teórica IBrX-50

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
12	CIEL3	CIELO	ON NM	772.547.017,00	3,725
13	CMIG4	CEMIG	PN N1	758.373.518,00	0,708
14	CSAN3	COSAN	ON NM	152.149.931,00	0,449
15	CSNA3	SID NACIONAL	ON	601.218.004,00	0,368
16	CTIP3	CETIP	ON NM	252.173.097,00	1,084
17	CYRE3	CYRELA REALT	ON NM	241.564.476,00	0,282
18	EMBR3	EMBRAER	ON NM	734.705.564,00	2,489
19	EQTL3	EQUATORIAL	ON NM	197.784.403,00	0,842
20	ESTC3	ESTACIO PART	ON NM	276.966.273,00	0,586
21	FIBR3	FIBRIA	ON NM	222.307.680,00	1,514
22	GGBR4	GERDAU	PN N1	856.830.567,00	0,615
23	GOAU4	GERDAU MET	PN N1	268.558.664,00	0,094
24	HYPE3	HYPERMARCAS	ON NM	372.960.637,00	0,785
25	ITSA4	ITAUSA	PN N1	3.456.819.165,00	3,215
26	ITUB4	ITAUNIBANCO	PN N1	2.921.797.735,00	9,86
27	JBSS3	JBS	ON NM	1.676.252.804,00	3,162
28	KLBN11	KLABIN S/A	UNT N2	508.711.805,00	1,432
29	KROT3	KROTON	ON NM	1.614.466.975,00	2,047
30	LAME4	LOJAS AMERIC	PN	473.472.344,00	1,055
31	LREN3	LOJAS RENNER	ON NM	632.324.010,00	1,506
32	MRVE3	MRV	ON NM	272.917.996,00	0,247
33	MULT3	MULTIPLAN	ON N2	78.663.321,00	0,435
34	NATU3	NATURA	ON NM	171.984.796,00	0,472
35	OIBR3	OI	ON N1	297.854.617,00	0,122
36	PCAR4	P.ACUCAR-CBD	PN N1	155.869.376,00	0,986

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
37	PETR3	PETROBRAS	ON	2.708.739.869,00	3,268
38	PETR4	PETROBRAS	PN	4.037.695.282,00	4,017
39	QUAL3	QUALICORP	ON ER NM	216.515.885,00	0,426
40	RADL3	RAIADROGASIL	ON NM	198.046.365,00	1,047
41	RENT3	LOCALIZA	ON NM	149.287.167,00	0,476
42	RUMO3	RUMO LOG	ON NM	168.809.915,00	0,162
43	SMLE3	SMILES	ON NM	56.503.753,00	0,214
44	SUZB5	SUZANO PAPEL	PNA N1	463.298.896,00	1,017
45	TIMP3	TIM PART S/A	ON NM	808.255.142,00	0,814
46	UGPA3	ULTRAPAR	ON NM	542.593.819,00	4,699
47	USIM5	USIMINAS	PNA N1	481.380.035,00	0,184
48	VALE3	VALE	ON ED N1	1.469.201.955,00	3,403
49	VALE5	VALE	PNA ED N1	1.945.837.689,00	3,63
50	VIVT4	TELEF BRASIL	PN	473.532.117,00	2,27
	Total			45.113.111.923,00	100
	Redutor			98.550.857,55	

(BMF&BOVESPA, 2015g)

Apêndice C

Composição da carteira teórica IEE

Tabela C.1: Tabela de composição da carteira teórica que compõe o IEE 3º quadrimestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ALUP11	ALUPAR	UNT N2	4.200,00	6,659
2	CESP6	CESP	PNB N1	3.900,00	6,748
3	CMIG4	CEMIG	PN N1	8.000,00	6,385
4	COCE5	COELCE	PNA	1.600,00	6,641
5	CPFE3	CPFL ENERGIA	ON NM	3.900,00	6,416
6	CPLE6	COPEL	PNB N1	2.000,00	7,181
7	ELET6	ELETROBRAS	PNB N1	7.000,00	7,255
8	ELPL4	ELETROPAULO	PN N2	5.100,00	5,601
9	ENBR3	ENERGIAS BR	ON NM	5.200,00	6,48
10	EQTL3	EQUATORIAL	ON NM	1.800,00	6,549
11	GETI4	AES TIETE	PN	4.100,00	6,226
12	LIGT3	LIGHT S/A	ON NM	5.100,00	6,974
13	TAEE11	TAESA	UNT N2	3.600,00	7,631

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
14	TBLE3	TRACTEBEL	ON NM	1.800,00	6,436
15	TRPL4	TRAN PAULIST	PN N1	1.600,00	6,818
	Total			58.900,00	100
	Redutor			35,13	

([BMF&BOVESPA, 2015h](#))

Apêndice D

Composição da carteira teórica

IFNC

Tabela D.1: Tabela de composição da carteira teórica que compõe o IFNC 3º trimestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ABCB4	ABC BRASIL	PN N2	51.116.447,00	0,208
2	BBAS3	BRASIL	ON NM	839.653.556,00	6,385
3	BBDC3	BRADESCO	ON N1	399.635.555,00	4,378
4	BBDC4	BRADESCO	PN N1	1.463.413.851,00	14,387
5	BBSE3	BBSEGURIDADE	ON NM	674.972.311,00	8,76
6	BRSR6	BANRISUL	PNB N1	174.315.074,00	0,484
7	BVMF3	BMFBOVESPA	ON NM	1.778.324.029,00	9,776
8	CIEL3	CIELO	ON NM	772.547.017,00	13,646
9	CTIP3	CETIP	ON NM	252.173.097,00	3,97
10	ITSA4	ITAUSA	PN N1	3.456.819.165,00	11,78
11	ITUB3	ITAUNIBANCO	ON N1	144.641.467,00	1,692

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
12	ITUB4	ITAUUNIBANCO	PN N1	1.469.057.858,00	18,162
13	PSSA3	PORTO SEGURO	ON NM	94.329.291,00	1,485
14	SANB11	SANTANDER BR	UNT	400.076.102,00	2,649
15	SULA11	SUL AMERICA	UNT N2	240.708.419,00	2,238
	Total			12.211.783.239,00	100
	Redutor			47.250.787,85	

(BMF&BOVESPA, 2015i)

Apêndice E

Composição da carteira teórica

INDX

Tabela E.1: Tabela de composição da carteira teórica que compõe o INDX 3º quadrimestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ABEV3	AMBEV S/A	ON	2.540.193.731,00	20,418
2	ALPA4	ALPARGATAS	PN N1	164.577.781,00	0,545
3	BEEF3	MINERVA	ON NM	128.878.572,00	0,682
4	BEMA3	BEMATECH	ON NM	40.557.887,00	0,178
5	BRFS3	BRF SA	ON NM	704.721.811,00	19,887
6	BRKM5	BRASKEM	PNA N1	264.585.622,00	2,22
7	CSAN3	COSAN	ON NM	152.149.931,00	1,465
8	CSNA3	SID NACIONAL	ON	601.218.004,00	1,202
9	CYRE3	CYRELA REALT	ON NM	241.564.476,00	0,919
10	DIRR3	DIRECIONAL	ON NM	73.232.619,00	0,113
11	DTEX3	DURATEX	ON NM	261.597.936,00	0,73

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
12	EMBR3	EMBRAER	ON NM	734.705.564,00	8,128
13	ETER3	ETERNIT	ON NM	151.213.054,00	0,127
14	EVEN3	EVEN	ON NM	205.645.961,00	0,325
15	EZTC3	EZTEC	ON NM	53.227.551,00	0,281
16	FIBR3	FIBRIA	ON NM	222.307.680,00	4,944
17	GFS3	GAFISA	ON NM	365.768.901,00	0,373
18	GGBR3	GERDAU	ON N1	100.345.473,00	0,184
19	GGBR4	GERDAU	PN N1	856.830.567,00	2,007
20	GOAU4	GERDAU MET	PN N1	268.558.664,00	0,308
21	GRND3	GRENDENE	ON NM	81.976.022,00	0,583
22	HBOR3	HELBOR	ON NM	104.154.863,00	0,09
23	HGTX3	CIA HERING	ON NM	127.749.285,00	0,785
24	HYPE3	HYPERMARCAS	ON NM	372.960.637,00	2,563
25	JBSS3	JBS	ON NM	1.676.252.804,00	10,325
26	JHSF3	JHSF PART	ON NM	116.679.501,00	0,093
27	KEPL3	KEPLER WEBER	ON	26.285.325,00	0,247
28	KLBN11	KLABIN S/A	UNT N2	508.711.805,00	4,674
29	LEVE3	METAL LEVE	ON NM	38.492.519,00	0,351
30	MAGG3	MAGNESITA SA	ON NM	162.343.363,00	0,185
31	MDIA3	M.DIASBRANCO	ON NM	28.909.897,00	0,826
32	MRFG3	MARFRIG	ON NM	357.340.458,00	0,939
33	MRVE3	MRV	ON NM	272.917.996,00	0,806
34	MYPK3	IOCHP-MAXION	ON NM	70.989.042,00	0,456
35	NATU3	NATURA	ON NM	171.984.796,00	1,541
36	PMAM3	PARANAPANEMA	ON NM	318.776.136,00	0,295

Apêndice E. Composição da carteira teórica INDX

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
37	POMO4	MARCOPOLO	PN N2	538.761.121,00	0,407
38	RAPT4	RANDON PART	PN N1	153.348.656,00	0,191
39	RSID3	ROSSI RESID	ON NM	74.108.766,00	0,031
40	SMT03	SAO MARTINHO	ON NM	44.547.723,00	0,832
41	SUZB5	SUZANO PAPEL	PNA N1	463.298.896,00	3,319
42	TCSA3	TECNISA	ON NM	76.051.070,00	0,096
43	TECN3	TECHNOS	ON NM	66.578.654,00	0,114
44	TUPY3	TUPY	ON NM	63.470.970,00	0,516
45	USIM3	USIMINAS	ON N1	121.966.457,00	0,432
46	USIM5	USIMINAS	PNA N1	481.380.035,00	0,601
47	WEGE3	WEG	ON NM	568.838.135,00	3,666
	Total			15.220.756.717,00	100
	Redutor			18.220.541,20	

(BMF&BOVESPA, 2015j)

Apêndice F

Composição da carteira teórica

SMLL

Tabela F.1: Tabela de composição da carteira teórica que compõe o SMLL 3º quadri-mestre de 2015

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
1	ABCB4	ABC BRASIL	PN N2	51.116.447	0,471
2	ALPA4	ALPARGATAS	PN N1	164.577.781	1,386
3	ALSC3	ALIANSC	ON NM	79.678.412	0,913
4	ALUP11	ALUPAR	UNT N2	73.633.600	1,133
5	ANIM3	ANIMA	ON NM	36.571.481	0,593
6	ARTR3	ARTERIS	ON NM	105.881.135	1,035
7	ARZZ3	AREZZO CO	ON NM	41.822.802	0,912
8	BBRK3	BR BROKERS	ON NM	165.451.038	0,301
9	BEEF3	MINERVA	ON NM	128.878.572	1,732
10	BRAP4	BRADSPAR	PN N1	221.328.102	1,954
11	BRIN3	BR INSURANCE	ON NM	74.674.350	0,052

Apêndice F. Composição da carteira teórica SMLL

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
12	BRML3	BR MALLS PAR	ON NM	462.652.489	5,638
13	BRPR3	BR PROPERT	ON NM	295.259.171	3,422
14	BRSR6	BANRISUL	PNB N1	174.315.074	1,096
15	BTOW3	B2W DIGITAL	ON NM	110.382.954	1,904
16	CESP6	CESP	PNB N1	181.964.945	3,056
17	CSMG3	COPASA	ON NM	58.136.685	0,739
18	CSNA3	SID NACIONAL	ON	601.218.004	3,055
19	CVCB3	CVC BRASIL	ON NM	39.918.030	0,634
20	CYRE3	CYRELA REALT	ON NM	241.564.476	2,336
21	DIRR3	DIRECIONAL	ON NM	73.232.619	0,287
22	DTEX3	DURATEX	ON NM	261.597.936	1,853
23	ECOR3	ECORODOVIAS	ON NM	198.834.559	1,366
24	ELPL4	ELETROPAULO	PN N2	92.735.426	0,988
25	ENBR3	ENERGIAS BR	ON NM	232.602.876	2,814
26	ESTC3	ESTACIO PART	ON NM	276.966.273	4,86
27	EVEN3	EVEN	ON NM	205.645.961	0,826
28	EZTC3	EZTEC	ON NM	53.227.551	0,715
29	GFSA3	GAFISA	ON NM	365.768.901	0,948
30	GOAU4	GERDAU MET	PN N1	268.558.664	0,782
31	GOLL4	GOL	PN N2	106.237.845	0,398
32	GRND3	GRENDENE	ON NM	81.976.022	1,482
33	HBOR3	HELBOR	ON NM	104.154.863	0,228
34	HGTX3	CIA HERING	ON NM	127.749.285	1,992
35	IGTA3	IGUATEMI	ON NM	84.039.084	1,851
36	LEVE3	METAL LEVE	ON NM	38.492.519	0,893

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
37	LIGT3	LIGHT S/A	ON NM	97.629.475	1,296
38	LINX3	LINX	ON NM	32.932.383	1,667
39	MEAL3	IMC S/A	ON NM	50.177.986	0,241
40	MGLU3	MAGAZ LUIZA	ON NM	6.292.751	0,077
41	MILS3	MILLS	ON NM	79.839.249	0,396
42	MRFG3	MARFRIG	ON NM	357.340.458	2,384
43	MRVE3	MRV	ON NM	272.917.996	2,046
44	MYPK3	IOCHP-MAXION	ON NM	70.989.042	1,158
45	ODPV3	ODONTOPREV	ON NM	269.986.765	2,788
46	OIBR3	OI	ON N1	421.692.886	1,437
47	POMO4	MARCOPOLO	PN N2	538.761.121	1,035
48	PRIO3	PETRORIO	ON NM	29.726.533	0,063
49	QGEP3	QGEP PART	ON NM	71.577.609	0,468
50	QUAL3	QUALICORP	ON ER NM	216.515.885	3,533
51	RAPT4	RANDON PART	PN N1	153.348.656	0,485
52	RENT3	LOCALIZA	ON NM	149.287.167	3,947
53	RLOG3	COSAN LOG	ON NM	151.098.726	0,218
54	RSID3	ROSSI RESID	ON NM	74.108.766	0,078
55	RUMO3	RUMO LOG	ON NM	168.809.915	1,342
56	SEDU3	SOMOS EDUCA	ON NM	111.496.900	1,514
57	SEER3	SER EDUCA	ON NM	36.288.702	0,353
58	SLCE3	SLC AGRICOLA	ON NM	46.561.624	0,848
59	SMTO3	SAO MARTINHO	ON NM	44.547.723	2,113
60	SULA11	SUL AMERICA	UNT N2	240.708.419	5,067
61	TCSA3	TECNISA	ON NM	76.051.070	0,244

Apêndice F. Composição da carteira teórica SMLL

	Código	Ação	Tipo	Qtde. Teórica	Part.(%)
62	TOTS3	TOTVS	ON NM	133.894.682	4,793
63	USIM3	USIMINAS	ON N1	121.966.457	1,098
64	USIM5	USIMINAS	PNA N1	481.380.035	1,526
65	VLID3	VALID	ON NM	64.658.764	3,14
	Total			10.451.433.677	100
	Redutor			103.278.249,18	

(BMF&BOVESPA, 2015k)

Apêndice G

Programação VBA

G.1 Programação VBA RSI

```
Dim p As Long
```

```
Dim N As Long
```

```
Dim medG As Double
```

```
Dim medL As Double
```

```
Dim RS As Double
```

```
Dim IFR As Double
```

```
Sub RSI()
```

```
p = Worksheets("BVSP").Cells(2, 4).Value
```

```
N = Worksheets("BVSP").Cells(4, 1).Value
```

Apêndice G. Programação VBA

```
For i = (p + 5) To N
```

```
medG = 0
```

```
medL = 0
```

```
RS = 0
```

```
For j = 1 To p
```

```
medG = (medG + Worksheets("BVSP").Cells(i - j, 12).Value)
```

```
medL = (medL + Worksheets("BVSP").Cells(i - j, 13).Value)
```

```
Next j
```

```
RS = medG / medL
```

```
Worksheets("BVSP").Cells(i, 14).Value = 100 - 100 / (1 + RS)
```

```
'Calculate
```

```
Next i
```

```
End Sub
```

G.2 Programação VBA MMCD

```
Dim STF As Double
Dim fast As Double
Dim slow As Double
Dim trigger As Double
Dim N1 As Double
Dim N2 As Double
Dim N3 As Double

Sub MACD()

STF = Worksheets("BVSP").Cells(5, 7).Value 'Preço fechamento
STL = STF
sinal = STF

N = Worksheets("BVSP").Cells(4, 1).Value 'Número de obs

N1 = Worksheets("BVSP").Cells(2, 5).Value 'parametro EMA fast
N2 = Worksheets("BVSP").Cells(2, 6).Value 'parametro EMA slow
N3 = Worksheets("BVSP").Cells(2, 7).Value 'parametro EMA fast

fast = 2 / (N1 + 1)
```

Apêndice G. Programação VBA

```
slow = 2 / (N2 + 1)
```

```
trigger = 2 / (N3 + 1)
```

```
For i = 1 To N
```

```
preco = Worksheets("BVSP").Cells(i + 4, 7).Value 'calcula fechamento
```

```
STF = fast * preco + (1 - fast) * STF
```

```
STL = slow * preco + (1 - slow) * STL
```

```
HMACD = STF - STL
```

```
Worksheets("BVSP").Cells(i + 4, 15).Value = HMACD ' coluna MACD
```

```
sinal = trigger * HMACD + (1 - trigger) * sinal
```

```
Worksheets("BVSP").Cells(i + 4, 16).Value = sinal ' Coluna Sinal
```

```
Next i
```

```
End Sub
```

Apêndice H

Programação R

H.1 MSV de previsão

```
setwd("C:\\Users\\Samsung\\OneDrive\\Documentos\\UnB\\ETC\\BASE DE DADOS H-1")
dados<-read.csv2("DADOS IBOV H-1.csv")
dados<-dados[1:3620,]

Ret.5<-as.data.frame(c())
conte<-1
for(i in 1:724)
{
  x<-c(dados$Retorno.LN.Dia[which(dados$Ret5==dados$Ret5[conte])])
  Ret.5<-rbind(Ret.5,x)
  conte<-conte+5
}

IFR.5<-as.data.frame(c())
conte<-1
```

Apêndice H. Programação R

```
for(i in 1:724)
{
  x<-c(dados$IFR[which(dados$Ret5==dados$Ret5[conte])])
  IFR.5<-rbind(IFR.5,x)
  conte<-conte+5
}

MMCD.5<-as.data.frame(c())
conte<-1
for(i in 1:724)
{
  x<-c(dados$MMCD..Histo.[which(dados$Ret5==dados$Ret5[conte])])
  MMCD.5<-rbind(MMCD.5,x)
  conte<-conte+5
}

### unindo tabelas
Todaldados<-cbind(Ret.5,IFR.5,MMCD.5)
colnames(Todaldados)<-c("R1","R2","R3","R4","R5","IFR1","IFR2",
"IFR3","IFR4","IFR5","MMCD1","MMCD2","MMCD3","MMCD4","MMCD5")
Todaldados$classe<-as.numeric(dados$resposta[seq(1,3620,by=5)])-2

##### Separação de bases - normalização
Todaldados[,1:15]<-scale(Todaldados[,1:15])
treino1<-Todaldados[1:362,]
valida1<-Todaldados[363:603,]
teste1<-Todaldados[604:nrow(Todaldados),]
```


H.1. MSV de previsão

```
#####  
table(Todaldados$classe)  
  
#####          CRIAÇÃO DA SEQUENCIAS DO SVM          #####  
library(kernlab)  
#Carrega os pacotes necessários para realizar o paralelismo  
library(foreach)  
library(doParallel)  
library(doSNOW)  
####              #####          Configuração de nucleos para Otimização  
#Checa quantos núcleos existem  
ncl<-detectCores()  
ncl  
#Registra os clusters a serem utilizados  
cl <- makeCluster(ncl)  
registerDoParallel(cl)  
  
##### Grade de variação dos parametros #####  
sigma<- seq(0.000000001,5,length.out=100)  
c<-seq(1,1000,length.out=10)  
parametro<-as.data.frame(expand.grid(sigma=sigma,c=c))  
performace<-as.data.frame(rep(0.0,nrow(parametro)))  
#####  
ptm <- proc.time()  
acuracia <-foreach(i=(1:nrow(parametro)),.packages="kernlab", .combine='rbind') %dopar% {  
#####  
x<- parametro[i,1]
```

Apêndice H. Programação R

```
y<-parametro[i,2]

svm<-ksvm(classe~.,data=treino1,type="C-svc",C=y,kernel="rbfdot"
          ,kpar=list(sigma=x),scale=T,prob.model=TRUE, cross=5)
#####
performace[i,1]<-(1-svm@cross)
}
proc.time() - ptm
parametro$performace<-unname(acuracia)
stopCluster(cl)
##seleção de parametros
performace<-na.omit(performace)
parametro<-na.omit(parametro)
melhorparametro<-parametro[which(parametro$performace==max(parametro$performace))[1],]
melhorparametro

#####
x<- as.numeric(melhorparametro[1])
y<-as.numeric(melhorparametro[2])

svmot<-ksvm(classe~.,data=valida1,type="C-svc",C=y,kernel="rbfdot"
            ,kpar=list(sigma=x),scale=T,prob.model=TRUE)

respostapred2<-predict(svmot,teste1,type="probabilities")

### definir linha de corte para classificação.
previ2<-ifelse(respostapred2[,2]>0.60,1,0)
```

```
table(previ2, teste1$classe)
```

```
##### Plot do gráfico performance #####
```

```
library(rgl)
```

```
y<-parametro$sigma
```

```
x<-parametro$c
```

```
z<-parametro$performace
```

```
plot3d(x, y, z, col="red", size=3, type = "p", xlab="C", ylab="Sigma", zlab="Performance")
```

```
#$
```

H.2 ACP e MSV

```
setwd("C:\\Users\\Samsung\\OneDrive\\Documentos\\UnB\\ETC\\BASE DE DADOS H-1")
```

```
dados<-read.csv2("DADOS IBOV H-1.csv")
```

```
dados<-dados[1:3620,]
```

```
Ret.5<-as.data.frame(c())
```

```
conte<-1
```

```
for(i in 1:724)
```

```
{
```

```
  x<-c(dados$Retorno.LN.Dia[which(dados$Ret5==dados$Ret5[conte])])
```

```
  Ret.5<-rbind(Ret.5,x)
```

```
  conte<-conte+5
```

```
}
```

```
IFR.5<-as.data.frame(c())
```

Apêndice H. Programação R

```
conte<-1
for(i in 1:724)
{
  x<-c(dados$IFR[which(dados$Ret5==dados$Ret5[conte])])
  IFR.5<-rbind(IFR.5,x)
  conte<-conte+5
}

MMCD.5<-as.data.frame(c())
conte<-1
for(i in 1:724)
{
  x<-c(dados$MMCD..Histo.[which(dados$Ret5==dados$Ret5[conte])])
  MMCD.5<-rbind(MMCD.5,x)
  conte<-conte+5
}

### unindo tabelas
Totaldados<-cbind(Ret.5,IFR.5,MMCD.5)
colnames(Totaldados)<-c("R1","R2","R3","R4","R5","IFR1","IFR2",
"IFR3","IFR4","IFR5","MMCD1","MMCD2","MMCD3","MMCD4","MMCD5")
### unindo tabelas
Totaldados<-cbind(Ret.5,IFR.5,MMCD.5)

PCA<-prcomp(Totaldados)
base<-PCA$x
```

```

plot(PCA,type="l")

basepca<-as.data.frame(base[,1:3])
basepca$resposta<-as.numeric(dados$resposta[seq(1,3620,by=5)])-2
basepca$resposta<-basepca$resposta+1
dat<-basepca[,c(4,1:3)]
colnames(dat)<-c("group","X1","X2","X3")
dat$group<-as.factor(dat$group)
library(e1071)
library(rgl)
library(misc3d)
nnew <- 50

# Fit SVM
fit <- svm(group ~. , data=dat, cost=667, gamma=3.4343)

# Plot original data
plot3d(dat[,,-1], col=dat$group, xlab="PC1", ylab="PC2",zlab="PC3")

# Get decision values for a new data grid
newdat.list <- lapply(dat[,,-1], function(x) seq(min(x), max(x), len=nnew))
newdat<-expand.grid(newdat.list)
newdat.pred<-predict(fit, newdata=newdat, decision.values=T)
newdat.dv<-attr(newdat.pred, 'decision.values')
newdat.dv<-array(newdat.dv, dim=rep(nnew, 3))

# Fit/plot an isosurface to the decision boundary

```

Apêndice H. Programação R

```
contour3d(newdat.dv, level=0, x=newdat.list$X1, y=newdat.list$X2,  
z=newdat.list$X3, add=T)
```