



Universidade de Brasília
Faculdade de Economia, Administração, Contabilidade e Gestão Pública
Departamento de Ciências Contábeis e Atuariais
Programa de Pós-Graduação em Ciências Contábeis

Noel Kentinus Olokodana

**Fraudes em relatórios financeiros:
efeitos nos preços das ações e detecção por meio da máquina de suporte vetorial**

Brasília
2023

Noel Kentinus Olokodana

**Fraudes em relatórios financeiros:
efeitos nos preços das ações e detecção por meio da máquina de suporte vetorial**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Contábeis da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas da Universidade de Brasília como requisito parcial para a obtenção do título de Mestre em Ciências Contábeis.

Área de concentração: Mensuração Contábil.

Orientador: Prof. Dr. Rodrigo de Souza Gonçalves.

Brasília
2023

Noel Kentinus Olokodana

**Fraudes em relatórios financeiros:
efeitos nos preços das ações e detecção por meio da máquina de suporte vetorial**

*Financial statement frauds:
Effect on stock price and detection via support vector machine*

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Contábeis da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas da Universidade de Brasília como requisito parcial para a obtenção do título de Mestre em Ciências Contábeis.

Área de concentração: Mensuração Contábil.

Orientador: Prof. Dr. Rodrigo de Souza Gonçalves.

Data de aprovação: 30 de janeiro de 2023.

Banca examinadora

Prof. Dr. Rodrigo de Souza Gonçalves – UnB
Presidente da banca

Prof^a. Dr^a. Michele Rílany Rodrigues Machado – UFG
Membro externo

Prof. Dr. Bruno Vinícius Ramos Fernandes – UnB
Membro interno

Resumo

As ocorrências de fraudes aumentaram muito nos últimos anos, em muitos países (PricewaterhouseCoopers, 2020). Isto, provavelmente, tem afetado negativamente a eficiência dos mercados de capitais, tornando-os em um ambiente de desconfiança, causando fugas de investimentos e hesitação na entrada de novos investidores. Segundo Hung et al. (2015), as fraudes corporativas não afetam somente os credores ou os acionistas das firmas fraudadoras, mas afetam também clientes, fornecedores, o mercado financeiro como um todo, o governo, ou seja, toda a economia. De acordo com Dyck et al. (2010), as fraudes financeiras nos relatórios contábeis são, geralmente, detectadas muito tempo depois, após as suspeições dos impactos do crime. Ocorre que a atuação *post factum* (detecção de fraude) não é eficiente para reduzir ou eliminar as ocorrências; nesse sentido, uma ação preventiva é relevante e necessária. Há vários métodos para prevenção de fraudes em relatórios financeiros, e esta pesquisa tem por objetivo analisar a reação do mercado às fraudes, sem e com o conhecimento deste. Para isso, o estudo comparou o comportamento dos preços das ações ao redor da data em que a fraude foi cometida e ao redor da data em que foi descoberta. Em segundo lugar, o estudo analisou a precisão da máquina de suporte vetorial (SVM) na previsão de fraudes em relatórios financeiros considerando índices financeiros e índices não financeiros. A pesquisa fez uso da metodologia de estudos de eventos para alcançar seu primeiro objetivo e da SVM para alcançar seu segundo objetivo. Retornos anormais positivos foram encontrados um dia depois e do quarto até o décimo dia depois da realização de fraudes em relatórios financeiros. Sem seu conhecimento, o mercado reage positivamente à realização de fraudes em relatórios financeiros, gerando ganhos anormais às entidades fraudadoras. Provavelmente, pelo tempo demorado até a suspeita de fraude ser julgada pela Comissão de Valores Mobiliários (CVM), o mercado ignora a decisão do órgão regulador ao reconhecer oficialmente a prática ou manobra suspeita como fraudulenta, por infringir uma lei. Na pesquisa, variáveis sensíveis para detecção de riscos de fraudes foram identificadas, e o modelo polinomial da SVM utilizado conseguiu prever riscos de fraudes em pelo menos 64% dos relatórios financeiros fraudulentos. Os achados contribuem para a academia no que diz respeito às consequências das fraudes em relatórios financeiros no mercado, incentivando, assim, outras pesquisas na área. O estudo contribui aos trabalhos dos auditores e reguladores, recomendando o uso do modelo polinomial de SVM para prever fraudes em relatórios financeiros. O estudo avança no tratamento de dados financeiros trazendo à atenção dos pesquisadores que os métodos clássicos como a regressão linear podem não ser adequadas, em consonância com as conclusões de Adepoju et al. (2019) e Gonçalves (2019).

Palavras-chave: fraudes em relatórios financeiros; estudo de eventos; reação do mercado; máquina de suporte vetorial; detecção de fraudes.

Abstract

Fraud occurrences have increased significantly in recent years in many countries (PricewaterhouseCoopers, 2020). Probably, this has negatively affected the efficiency of capital markets, turning them into an environment of distrust, causing investment flight and hesitation in the entry of new investors. According to Hung et al. (2015), corporate fraud does not only affect creditors or shareholders of defrauding firms, but also it affects customers, suppliers, the financial market as a whole, the government; that is, the entire economy. According to Dyck et al. (2010), financial fraud in accounting reports is usually detected much later, after suspicions of the impacts of the crime. It turns out that *post factum* action (fraud detection) is not efficient in reducing or eliminating occurrences; in this sense, preventive action is relevant and necessary. There are several methods for preventing fraud in financial reporting, and this research aims to analyze the market reaction to fraud without and with its knowledge. For this, the study compared stock price behavior around the date the fraud was committed and around the date it was discovered. Secondly, the study analyzed the accuracy of the support vector machine (SVM) in predicting fraud in financial reports, considering financial indices and non-financial indices. The research used event studies methodology to achieve its first objective and SVM to achieve its second objective. Positive abnormal returns were found one day after and from the fourth to the tenth day after financial reporting fraud. Without knowledge of the fraud, the market reacts positively to the performance of fraud in financial reports, generating abnormal gains for the fraudsters. Probably, due to the time it takes for the suspected fraud to be judged by the Securities and Exchange Commission of Brazil (CVM), the market ignores the regulatory agency's decision when officially recognizing the suspected practice or maneuver as fraudulent, for breaking the law. Sensitive variables to detect fraud risks were identified and the SVM polynomial model was capable to predict fraud risks in at least 64% of fraudulent financial reports. The findings contribute to the academy about the consequences of fraud in financial reports in the market and also encourage other research. The study also contributes to the work of auditors and regulators, recommending the SVM polynomial model to predict fraud in financial reports. The study advances in the treatment of financial data, bringing to the attention of researchers that classic methods such as linear regression may not be adequate, in line with the conclusions of Adepoju et al. (2019) and Gonçalves (2019).

Keywords: financial statement frauds; event study; market reaction; support vector machine; fraud detection.

Lista de figuras

Figura 1 – Estratégias/valores da CVM para enfrentar fraudes ou irregularidades no mercado	33
Figura 2 – Estudo de evento à realização da fraude	39
Figura 3 – Distribuição da frequência de retornos compostos	40
Figura 4 – Distribuição da frequência de retornos contínuos ou logarítmicos.....	41
Figura 5 – Máquina de suporte vetorial linear.....	53
Figura 6 – Máquina de suporte vetorial não linear	56
Figura 7 – Máquina de suporte vetorial polinomial	56
Figura 8 – Máquina de suporte vetorial radial/gaussiano.....	57
Figura 9 – Distribuição dos relatórios fraudulentos por setor de atividade.....	59

Lista de tabelas

Tabela 1 – Amostra/processos	34
Tabela 2 – Amostra/Relatórios fraudulentos	35
Tabela 3 – Retornos reais e esperados das ações um dia depois da data de realização de fraudes	61
Tabela 4 – Retornos anormais dos quatro primeiros dias depois da realização de fraudes	62
Tabela 5 – Estimadores e retornos esperados das ações um dia depois da data de detecção de fraudes	66
Tabela 6 – Retornos anormais dos quatro primeiros dias depois da detecção de fraudes	67
Tabela 7 – Resultados dos modelos univariados	69
Tabela 8 – Soma acumulada dos retornos anormais à realização de fraudes (CAAR_Real) ..	73
Tabela 9 – Soma acumulada dos retornos anormais à detecção de fraudes (CAAR_Det)	76
Tabela 10 – Acurácia de detecção de fraudes em relatórios fraudulentos (amostra 2)	78
Tabela 11 – Comparação dos modelos do SVM	78

Lista de quadros

Quadro 1 – Parâmetros para a execução do modelo de estudo de eventos	43
Quadro 2 – Anos dos relatórios fraudulentos e não fraudulentos	46
Quadro 3 – Indicadores contábeis e sua relação com a realização de fraudes	48
Quadro 4 – Relação entre relatórios fraudulentos à realização e à detecção da fraude	64

Lista de abreviaturas e siglas

ANN: *Artificial neural network* – Rede neural artificial

AR: Retorno anormal

B3: Brasil Balcão Bolsa

CVM: Comissão de Valores Mobiliários

DT: *Decision tree* – Árvore de decisão

E(R): Retorno esperado

EBIT: *Earnings before interest and taxes* – Lucros antes de juros e taxas

EBITDA: *Earnings before interest, taxes, depreciation and amortizations* – Lucros antes de juros, taxas, depreciação e amortização

EDC1: Dívida total líquida

EDC10: EBIT / Despesa financeira líquida

EDC11: Dívida bruta / EBITDA

EDC12: Dívida líquida / EBITDA

EDC2: Dívida total bruta

EDC3: Índice de endividamento (Dívida bruta / Ativo total)

EDC4: (Dívida bruta / Patrimônio líquido)

EDC5: Dívida líquida / Patrimônio líquido

EDC6: Estrutura de capital (Dívida / Dívida + Patrimônio líquido)

EDC7: EBIT / Dívida bruta

EDC8: EBIT / Dívida líquida

EDC9: EBIT / Despesa financeira bruta

IPCA: Índice de preços ao consumidor amplo

INF1: Pagamento de dividendo

INF2: Participação minoritária

INF3: Ciclo financeiro

INF4: Ciclo operacional

INF5: Giro do ativo

INF6: Giro do patrimônio líquido

KNN: *K-nearest neighbor* – K vizinho mais próximo

LIQ1: Liquidez geral

LIQ2: Liquidez corrente

LIQ3: Liquidez seca

LIQ4: Capital de giro

LIQ5: *Capital employed*

OUT1: *Capital expenditure* (Despesa de capital incluindo os investimentos)

OUT2: Depreciação e amortização

OUT3: *Capital expenditure* / Depreciação

OUT4: Depreciação / Imobilizado e intangíveis

OUT5: Fluxo de caixa livre

R: Retorno

REN1: EBITDA

REN10: ROE rentabilidade do patrimônio (patrimônio inicial)

REN11: ROIC (*Invested capital* final)

REN12: ROIC (*Invested capital* médio)

REN13: ROIC (*Invested capital* inicial)

REN14: *Invested capital*

REN15: Rentabilidade de investimento de subsidiários

REN16: Alavancagem financeira

REN17: Alavancagem operacional

REN2: LAIR + Despesa financeira líquida

REN3: Margem bruta

REN4: Margem EBIT (margem operacional)

REN5: Margem líquida

REN6: Margem EBITDA

REN7: ROA rentabilidade do ativo

REN8: ROE rentabilidade do patrimônio (patrimônio final)

REN9: ROE rentabilidade do patrimônio (patrimônio médio)

RF: *Random forest* – Floresta aleatória

R_M: Retorno do mercado

SAS: Statement of Auditing Standards

SEC: The Securities and Exchange Commission

SVM: *Support vector machine* – máquina de suporte vetorial

Sumário

1	Introdução	14
1.1	Contextualização	14
1.2	Objetivo geral e objetivos específicos	17
1.3	Justificativa.....	17
2	Referencial teórico	19
2.1	Hipótese do mercado eficiente (HEM) e a reação do mercado à ocorrência de fraudes em relatórios financeiros.....	20
2.2	Fraudes.....	22
2.3	Fraudes em relatórios financeiros e retornos anormais	23
2.4	Fatores de riscos e detecção de fraudes em relatórios financeiros	24
2.5	Fraude nas demonstrações financeiras e sua repercussão no mercado financeiro .	30
3	Metodologia	32
3.1	Amostra para a análise das consequências de fraudes em relatórios financeiros no mercado	34
3.2	Coleta e análise de dados para o modelo de estudo de eventos para a análise das consequências de fraudes em relatórios financeiros no mercado	35
3.3	Definição das variáveis para a análise das consequências das fraudes em relatórios financeiros	37
3.4	Desenvolvimento do modelo de estudo de eventos	38
3.5	Coleta e análise de dados para a análise da detecção de fraudes em relatórios financeiros	44
3.6	Variáveis para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes.....	47
3.6.1	Variável dependente (RF_RNF) para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes	48
3.6.2	Variáveis independentes para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes	48
3.7	Máquina de suporte vetorial (SVM).....	51
3.8	Máquina de suporte vetorial linear	52
3.9	Máquina de suporte vetorial não linear	55
3.1	Métricas de avaliação da eficiência dos modelos de detecção de fraudes	58
4	Tratamento dos dados	58
4.1	Análises do estudo de eventos.....	59

4.1.1	Análise do modelo do estudo de eventos à realização de fraudes	60
4.1.2	Análise do modelo do estudo de eventos à detecção de fraudes	63
4.2	Análise do modelo de detecção de fraudes (SVM)	67
5	Análise dos resultados.....	71
5.1	Análise dos resultados do estudo de eventos.....	72
5.1.1	Análise dos resultados do estudo de eventos depois da realização de fraudes sem conhecimento do mercado	72
5.1.2	Análise dos resultados do estudo de eventos depois da detecção de fraudes ..	75
5.2	Análise dos resultados da SVM.....	77
5.3	Teste da escolha do melhor modelo de detecção de fraude	78
6	Considerações finais.....	79
	Referências	83
	Apêndice	93

1 Introdução

1.1 Contextualização

As fraudes contábeis são a origem de enriquecimento ilícito de várias instituições/agentes, o que provoca, por consequência, a perda do capital de vários investidores e pessoas/instituições que se relacionam direta ou indiretamente. As fraudes corporativas não afetam somente os credores ou os acionistas das firmas fraudadoras, mas afetam também clientes, fornecedores, investidores, governo; portanto, toda a economia (Hung et al., 2015).

Segundo Wuerges e Borba (2014), é muito difícil identificar fraudes em relatórios financeiros, e ainda, para Murcia e Borba (2007) e Sousa (2016) a dificuldade na identificação de fraudes em relatórios financeiros se agrava quando se referem a omissões de transações. Essa situação piora, pois, de acordo com Silva Filho (2019), as normas subjetivas deixam os fraudadores em uma zona de conforto, o que dificulta as possíveis denúncias ou a detecção de fraudes ou irregularidades.

Dyck et al. (2010) e Alves (2016) afirmam que são os impactos (consequências) das fraudes, os quais aparecem muito tempo depois, que levantam a suspeita de fraudes e levam os reguladores a proceder às investigações.

Conforme Gao e Srivastava (2007), as fraudes financeiras podem ter as mais diversas origens, tais como: (1) receitas fictícias, como o reconhecimento prematuro de receitas; (2) superavaliação de ativos e despesas subestimadas; (3) omissão ou subestimação de despesas/passivo; (4) ativos fictícios; (5) falsa classificação; (6) omissão na divulgação ou divulgação inapropriada; (7) estornos de lançamentos contábeis, entre outras origens. A presença desses elementos nas demonstrações financeiras traz aos investidores e credores (as principais vítimas) consequências relevantes, que podem chegar a uma escala nacional ou mesmo internacional, o que minimiza a confiabilidade nos mercados de capitais, deixando-os menos eficientes (Vlad et al., 2011).

De acordo com a PricewaterhouseCoopers (2020), as ocorrências de fraudes aumentaram muito nos últimos anos, em muitos países passando de 30% em 2009 para 47% em 2020 o número de ocorrências em nível global. Para o mesmo período, nos Estados Unidos, a experiência com fraude nas entidades passou de 35% para 56%. Na Tailândia e na Austrália, em 2020, essa porcentagem é de 33%; 35%, na Austrália; 39%, na Bélgica; 43%, na Malásia; 47%, na França, na Alemanha e no Luxemburgo; 51%, na Irlanda; 56%, no Reino Unido; 58%, no Quênia; e 60%, na África do Sul e na China. Entretanto, apesar da incerteza econômica que o mundo está enfrentando desde 2020, a PricewaterhouseCoopers (2022) encontrou uma leve

queda (de 47%, em 2020, para 46%, em 2022) nas ocorrências de fraudes em nível global, cujo percentual de ocorrência ainda pode ser considerado elevado.

Para Eryigit (2019), tais ocorrências corroem a confiabilidade da informação contábil, a qual é relevante para o mercado de capitais e para os investidores em seus processos de tomada de decisões. O autor ainda identificou ausência de retornos anormais significativos no período de divulgação de informações não verídicas, o que justifica que os investidores mantêm suas decisões inalteradas antes e depois da divulgação de relatórios considerados fraudulentos.

Segundo Fama (1991), Palmrose et al. (2001), Finnerty et al. (2016) e Karajian e Ullah (2021), a fraude pode causar consequências no mercado, gerando retornos anormais. De acordo com Rezaee (2002), os benefícios devem valer a pena, quando comparados aos custos, para que o fraudador tome sua decisão, no que se refere a fraudes ainda não descobertas e realizadas no mercado financeiro por meio das demonstrações contábeis.

Contudo, o anúncio de fraudes pode gerar a existência de retornos anormais positivos ou negativos. De acordo com Tsai e Huang (2020), Palmrose et al. (2001) e Karajian e Ullah (2021), investidores reagem negativamente aos anúncios de fraudes em relatórios financeiros; enquanto que, para Du (2017) e Frendy e Hu (2014), os anúncios de fraudes causam retornos anormais positivos, aspecto este muito controverso. Com base nesse cenário, segundo Eryigit (2019), também se observa que o mercado pode ser indiferente aos anúncios de fraudes.

Tais aspectos observados nas pesquisas ora mencionadas podem decorrer, conforme Tay et al. (2016), da metodologia empregada, muito embora haja consenso entre os pesquisadores de que a metodologia de estudos de eventos é normalmente a mais indicada na identificação de retornos anormais aos eventos como anúncios de fraudes em relatórios financeiros.

A metodologia de estudo de eventos proposta por Ball e Brown (1968) e Fama et al. (1969) vale-se apenas do preço das ações como variáveis para identificar os retornos anormais. A parte dos preços das ações, das características (a data do evento, o intervalo de tempo da observação do evento, o intervalo de tempo necessário para as estimações) dos eventos em estudo (fraudes) é utilizada para identificar possíveis retornos anormais. Dos preços das ações, os retornos reais são obtidos na janela de observação (janela do evento). Ainda nesta mesma janela, a partir dos preços do período de estimações, retornos esperados são apurados. A diferença entre os retornos reais e os retornos esperados representa os retornos anormais.

No caso de variáveis que predizem o risco de fraude, existem diversos estudos que se valem das informações financeiras divulgadas para poder realizar a mensuração do referido

risco (como Altman, 1968; Ohlson, 1980; Beneish, 1999; Persons, 1995; Beaver et al., 2005; Wuerges & Borba, 2014; entre outros).

Para Altman (1968), uma entidade cuja lucratividade é baixa e o índice de endividamento é alto, há maior risco de falência, e, nesse contexto, há maior propensão ao cometimento de fraudes. Ocorre que, em todos os estudos ora mencionados, a precisão dos índices é dúbia, e o uso, por exemplo, da análise discriminatória múltipla (MDA) pode contribuir para uma melhor estimativa do risco de fraude (Altman, 1968).

Ainda na busca de se melhorar os percentuais de predição de risco de fraude, Beneish (1999) propõe em seu modelo variáveis independentes para captar variações dos estados financeiros, tais como: a variação do prazo de recebimento; a variação da margem de lucro bruto; o índice de qualidade do ativo (a porcentagem do ativo não circulante, menos imóveis e equipamentos, no ativo total); a variação de vendas; a variação de depreciação; a variação de despesas administrativas e despesas de vendas; e a variação de capital de giro/ativo total.

Já Ohlson (1980) se vale da regressão logística para determinar a probabilidade de ocorrência de fraudes nos relatórios financeiros, utilizando-se das seguintes variáveis: log do ativo total (LOGTA); passivo total/ativo total (TLTA); capital de giro/ativo total (WCTA); ativo circulante/ativo total (CLCA); lucro líquido/ativo total (NITA); e fundo advindo de operações/passivo total (FUTL). O autor conclui que para ter uma previsão exitosa de falência de uma entidade, dados recentes (até um ano) devem ser utilizados.

Persons (1995) se vale de dois grupos de relatórios – fraudulentos e não fraudulentos – , sendo que os dois grupos de relatórios advêm de empresas diferentes, mas do mesmo setor de atividade e do mesmo período. Por meio de um procedimento logístico, o método é usado para prever a probabilidade de um relatório ser fraudulento. Seu modelo é um modelo de regressão logística com uma análise avançada, ao contrário do modelo de Ohlson (1980), que conta com a determinação do “escore de corte” para reduzir os erros do tipo 1 e do tipo 2. A equação de Pearson (1995) usa as seguintes variáveis independentes: ativo circulante/ativo total (CATA); vendas/ativo total (SATA); e duas outras variáveis de controle, para as quais se determinou que o escore ideal seja de 0,6018 para identificar fraude no relatório analisado.

O modelo de Beaver et al. (2005) utiliza-se também do modelo de regressão logística, assim como o modelo de Pearson (1995), levando em consideração as seguintes variáveis: passivo total/ativo total (LTA); EBITDA (ETL); e lucro líquido/ativo total (ROA). O modelo de Lenard et al. (2007) é baseado na regressão logística e, além de usar as mesmas variáveis que Pearson (1995), acrescenta uma variável chamada “variável de fraude”.

Conforme Lenard e Alam (2009), a inclusão dessa variável aumenta a robustez do modelo, reduzindo então os erros (tipo 1 e 2). A mencionada variável é gerada por meio de algoritmos estatísticos (lógica *fuzzy*) que identificam quão próximas as variáveis advindas de relatórios financeiros estão das recomendações do SAS 82 e SAS 99.

Mais recentemente, na busca por melhoria da predição do risco de fraude, os modelos de aprendizado de máquina e de inteligência artificial, com o auxílio de algoritmos como árvore de decisão (*decision tree* – DT), floresta aleatória (*random forest* – RF), redes neurais artificiais (*artificial neural network* – ANN), K vizinho mais próximo (*K-nearest neighbor* – KNN) e máquina de suporte vetorial (*support vector machine* – SVM), vêm sendo utilizados em pesquisas e na prática, a fim de solucionar os problemas de detecção de fraudes (Yazici, 2020).

Ante o cenário exposto, o problema da pesquisa deste trabalho consiste em: **empresas que empregam o uso de fraudes nas suas demonstrações financeiras apresentam retornos anormais no período de divulgação de suas informações?**

1.2 Objetivo geral e objetivos específicos

Esta pesquisa tem como objetivo geral analisar a existência de retornos anormais nas demonstrações financeiras tipificadas como fraudulentas em empresas brasileiras listadas. Para tanto, tem-se os seguintes objetivos específicos:

- Avaliar a existência de retornos anormais durante o período de realização e de detecção de fraudes em relatórios financeiros; e
- Analisar a precisão da SVM na previsão de fraudes em relatórios financeiros, tendo por finalidade identificar os riscos de fraudes em relatórios financeiros.

1.3 Justificativa

Segundo Eryigit (2019), a veracidade das informações contábeis de uma entidade, suas operações e a elaboração dos seus relatórios de acordo com as normas vigentes são importantes para a tomada de decisão racional dos investidores, a fim de possibilitar que se faça uma comparação justa entre as entidades no mercado de capitais. Assim, qualquer entidade que divulgasse qualquer informação falsa estaria buscando alguma vantagem ilícita. Segundo a PricewaterhouseCoopers (2020), existe uma grande variedade de informações falsas tipificadas como fraudes das quais entidades podem fazer uso em busca de vantagens ilícitas; entre elas, estão as fraudes em relatórios financeiros, que podem, em virtude de sua característica, gerar indevidamente retornos anormais.

Nesse sentido, Du (2017) e Frendy e Hu (2014) afirmam que o anúncio de fraudes (quando a fraude é descoberta e o mercado é informado pelo órgão regulador) pode causar retornos anormais positivos, representando um ganho ilícito para a entidade fraudadora. Por outro lado, de acordo com Tsai e Huang (2020), Palmrose et al. (2001) e Karajian e Ullah (2021), a realização de fraudes em relatórios financeiros deve prejudicar a entidade fraudadora, gerando retornos anormais negativos.

Ocorre que as fraudes em relatórios financeiros não causam consequências somente à entidade fraudadora, mas normalmente afetam também outras entidades. De acordo com Beatty et al. (2013), de um modo geral, os anúncios de fraudes em relatórios financeiros de uma entidade podem causar retornos anormais positivos a outras entidades em razão do deslocamento de investidores que deixariam de investir na entidade fraudadora e passariam a investir em outras entidades.

Em uma outra perspectiva, Darrough et al. (2020) afirmam que as entidades do mesmo país de uma entidade fraudadora são afetadas negativamente pelo anúncio em virtude do receio de que outras entidades do mesmo país de origem de uma entidade acusada de fraude em relatório financeiro estejam também cometendo alguma irregularidade, situação esta também identificada por Eckert et al. (2019).

Visto que as consequências de fraudes em relatórios financeiros não afetam somente a entidade fraudadora, mas podem afetar outras entidades do mesmo setor de atividades – ou entidades do mesmo país de origem no mercado internacional – ou causar a fuga de investimentos tanto nacionais como internacionais, esta pesquisa se justifica ao buscar suprir a lacuna das reais consequências de fraudes em relatórios financeiros no mercado de capitais brasileiro.

Estudos anteriores sobre fraudes em relatórios financeiros se dedicaram na detecção destas por meio de técnicas convencionais (regressão), de técnicas de aprendizado de máquina ou de mineração de dados através de informações obtidas de mercados de capitais (Song et al., 2014; Kanapickiene & Grundiene, 2015; Lin et al., 2015; Karpoff et al., 2017; Hajek & Henriques, 2017; Bao et al., 2019). Algumas pesquisas se dedicaram à taxonomia de fraudes em relatórios financeiros (Gupta & Gupta, 2015; Ikbal et al., 2020), buscando classificar diferentes tipos de fraudes.

Outras pesquisas, como a de Gepp et al. (2021), dedicaram-se a estudar variáveis que são sensíveis à detecção de fraudes em relatórios financeiros. Em outras palavras, tais pesquisas tentaram identificar as variáveis que podem compor um modelo de detecção de fraude,

enquanto outras analisaram as variáveis que podem afetar a ocorrência de fraudes, como é o caso de Yuwono e Marlina (2021) e Gupta e Mehta (2020).

Não obstante os fatos identificados, no que tange à reação do mercado financeiro às fraudes, os estudos ainda são controversos, dado que alguns encontraram reações negativas (Palmrose et al., 2001; Tsai & Huang, 2020; Karajian & Ullah, 2021), enquanto outros relataram reações positivas do mercado (Frendy & Hu, 2014; Du, 2017) ou a ausência de mudança de comportamento nos preços das ações (Eryigit, 2019), o que pode, em princípio, ser explicado em razão dos procedimentos metodológicos, como o uso de comunicados jornalísticos (anúncio de envolvimento da entidade em esquemas de fraude), que muitas vezes são feitos na fase inicial das investigações, tendo apenas um carácter informativo e não definitivo (logo, podem ser retratados depois de influenciar as decisões dos investidores).

Nesse contexto, a fim de evitar anúncios viesados e/ou tendenciosos, uma fonte de informações neutra (livre de tendências) e confiável (sem possibilidade de retratamento) seria mais adequada ao estudo da reação do mercado às fraudes, a exemplo dos resultados de apuração realizada por agentes de supervisão. É nesse sentido que este estudo busca avançar na análise das ocorrências de fraudes em relatórios financeiros e na compreensão do mercado de capitais em torno do resultado de investigação realizada por um órgão regulador, adicionando-se, o uso da SVM de modo a testar se há melhorias na capacidade preditiva do risco de fraude.

Por fim, o estudo avança no tratamento de dados financeiros. Os dados financeiros, ainda que sejam de uma mesma entidade, não são lineares, por tanto, uma análise linear pode não ser adequada.

2 Referencial teórico

Harvey et al. (2016) afirmam que as anomalias do mercado causam a existência de retornos anormais, sendo uma das causas em decorrência de fraudes. A detecção e/ou a previsão de fraudes é uma forma de evitar os efeitos (existência de retornos anormais) da anomalia (fraude) no mercado. Partindo da hipótese de mercados eficientes, o estudo citado apresenta as possíveis consequências da anomalia/fraude.

A detecção de fraudes não é uma tarefa fácil; o erro do tipo 1 (rejeitar que um relatório fraudulento enquanto, na verdade, ele é) teria um custo alto, o que favorece a perpetuação da prática de fraudes no mercado. O erro do tipo 2 (acusar um relatório não fraudulento de fraudulento) também tem um custo, ainda que eventualmente seja menor; todavia, o erro no processo de detecção de fraudes deve ser evitado.

A dificuldade de se distinguir uma fraude contábil de um erro contábil ou uma fraude de um risco de fraude pode estar na origem dos erros de detecção. Considerando tais pontos, este capítulo apresenta a base teórica quanto à atividade de detecção de fraudes e à ocorrência de retornos anormais.

2.1 Hipótese do mercado eficiente (HEM) e a reação do mercado à ocorrência de fraudes em relatórios financeiros

Segundo Fama (1970), a hipótese do mercado eficiente supõe que os mercados de capitais são totalmente eficientes; assim, os preços dos ativos incorporam totalmente as informações disponíveis. Os pressupostos da teoria de portfólio e do modelo de precificação de ativos (*capital asset pricing model* – CAPM) que justificam que o mercado é um jogo justo constituem-se como a base para a hipótese de eficiência (HEM).

Ainda com Fama (1970), a ausência de custos de transação nos preços das ações (1), o acesso de todos os participantes a todas as informações divulgadas no mercado (2) e a concordância de todos na implicação de informações nos preços correntes e na distribuição dos preços futuros de cada ação (3) são as três condições suficientes para que um mercado seja considerado eficiente.

Segundo a HEM, o mercado será considerado eficiente se um número suficiente de investidores tiver pronto acesso às informações disponíveis, impossibilitando a obtenção de retornos anormais por alguns investidores e assumindo as condições mencionadas. No entanto, é quase impossível evidenciar ou provar que os preços disponíveis refletem todas as informações disponíveis e, por isso, Fama (1970) argumenta que a eficiência do mercado pode ser testada em três níveis (eficiência fraca, semiforte e forte).

De acordo com o autor citado, a eficiência fraca é testada quando os preços ou índices históricos são usados para estimar preços ou índices futuros. Nesse teste, o interesse está apenas na aleatoriedade ou não dos preços ou dos retornos das ações. A eficiência semiforte é testada quando se responde à pergunta sobre o tempo que leva para os preços refletirem as informações divulgadas publicamente. Assim, o que diferencia a eficiência fraca da eficiência semiforte é a sustentação dos resultados. Por sua vez, a eficiência forte é testada quando se busca o uso de informações privilegiadas.

Na busca das consequências de fraudes, no que diz respeito à existência de retornos anormais, este estudo busca identificar o tempo que leva para que a notícia de realização e de detecção de fraudes reflitam nos preços das ações; assim, o foco consiste basicamente no tipo de eficiência semiforte do mercado.

Para Da Silva e Da Silva (2021), vários autores apontaram algumas limitações da hipótese do mercado eficiente, entre elas a de que a HEM foi desenvolvida considerando que os investidores apresentam uma racionalidade ilimitada (pressuposto sustentado pela teoria de utilidade), característica contestada originalmente por Kahneman e Tversky (1979).

Ao desenvolverem a teoria do prospecto, Kahneman e Tversky (1979) argumentam que, em situações de riscos, o indivíduo não teria uma racionalidade ilimitada, mas sim limitada ao seu apetite ao risco. Adicionalmente, de acordo com Grossman e Stiglitz (1980), nem todas as informações estão disponíveis no mercado, e os investidores têm que adquirir as informações que não são divulgadas no mercado para tomar suas decisões, portanto, se isso ocorre, tal aspecto é inconsistente com o pressuposto de que todas as informações estão disponíveis a todos os investidores.

Grossman e Stiglitz (1980) ainda argumentam que o equilíbrio competitivo não existe no mercado, pois, se existisse, aqueles investidores que adquirem informações além daquelas disponíveis no mercado não seriam compensados por isso; dessa forma, apenas uma parte da informação necessária estaria se refletindo nos preços dos ativos.

Para defender sua hipótese, Fama (1991) argumenta que as limitações da hipótese de mercado eficiente, na verdade, residem no fato de que a hipótese não pode ser testada por si mesma, pois a evidência de sua ocorrência está presente nos modelos de precificação de ativos. Portanto, a hipótese de mercado eficiente é baseada nos modelos de precificação de ativos e depende deles, dessa forma, a eficiência do mercado depende dos retornos esperados e da racionalidade limitada ao risco aceitável pelos investidores.

Isso motivou o desenvolvimento da hipótese dos mercados adaptáveis (Lo, 2005), segundo a qual o comportamento do investidor não seria racional o tempo todo, assim como as oportunidades de lucratividade não são permanentes, assim como a relação risco/retorno também não é constante (portanto, varia) ao longo do tempo. Segundo Lo (2005), a melhor maneira de se alcançar um nível consistente (alto) de retornos esperados é se adaptando às mudanças nas condições do mercado.

Para Fama (1991), há situações em que os modelos de precificação de ativos (CAPM) não conseguirão prever o preço dos ativos; nessas situações, o mercado seria ineficiente. Essas situações são chamadas de anomalias do mercado.

Ainda, de acordo com Harvey et al. (2016), as anomalias podem ser macroeconômicas (estendendo-se ao mercado todo) ou específicas a cada entidade. Em cada esfera, ela pode ser financeira, microestrutural, contábil ou de outras características, sendo a fraude uma das anomalias contábeis que teria o potencial de gerar retornos anormais.

2.2 Fraudes

Depreende-se da teoria da agência, segundo Segrestin e Hatchuel (2011), que o interesse dos acionistas é o foco principal nas decisões gerenciais pois os gestores atuam em nome dos acionistas. De acordo com a teoria (teoria da agência) e ainda segundo Segrestin e Hatchuel (2011) neste tipo de contrato (contrato entre gestores e acionistas), os gestores não têm autonomia. Esta falta de autonomia leva os gestores a tomarem decisões que agradem os acionistas. Supõe-se, então, que para agradar os acionistas, os gestores fazem uso de fraudes para apresentar uma situação financeira melhor que a situação real da entidade.

De acordo com Zahra, Priem e Rasheed (2007), fraudes cometidas pela gerência (fraudes como apropriação indébita, negociação com informações privilegiadas, autonegociação, mentira sobre fatos, falha em divulgar fatos, corrupção e encobrimentos) não trazem consequências, só para a entidade mas também para seus funcionários e a sociedade. As consequências da fraude afetam, então, vários stakeholders e quando a fraude é descoberta (detectada) as punições aplicadas à entidade recaem ou afetam os funcionários também. De acordo com Hung et al. (2015), as fraudes corporativas não afetam somente os credores ou os acionistas das firmas fraudadoras, mas afetam também clientes, fornecedores, investidores, governo; portanto, toda a economia. Segundo Zahra, Priem e Rasheed (2007) crimes corporativos como fraudes em relatórios financeiros proporcionam ganhos à entidade e também aos funcionários que contribuíram com a realização da fraude.

Apesar da função da auditoria não é detectar fraudes, mas, sim assegurar que os relatórios financeiros estão livres de distorções materiais, os Sistemas de auditoria interna e externa falharam em identificar distorções (fraudes). Provavelmente, é por isso que Sharma e Panigrahi (2012) afirmaram que a falha dos sistemas de auditoria interna em identificar fraudes levou ao uso de métodos da contabilidade investigativa com as técnicas de mineração de dados.

Segundo Soares (2004), a teoria econômica de crimes, relata a probabilidade de um indivíduo se envolver em atividades criminosas com os custos e benefícios dessas atividades, quando comparadas a ocupações legais. De acordo com o autor, há evidências (segundo Wellford, 1974) de que quanto mais desenvolvido é o país, maior (em termos financeiros e o impacto da fraude no país) são as fraudes. Fazendo uma analogia (substituindo a o país pela entidade), é provável que quanto maior é uma entidade (classificação pelo faturamento), maior podem ser a magnitude das fraudes e potencialmente, os impactos nos retornos.

2.3 Fraudes em relatórios financeiros e retornos anormais

Ozili (2015) considera que as fraudes em relatórios financeiros têm o objetivo de esconder a real situação econômica da entidade. Por meio delas, a entidade demonstra uma capacidade econômico-financeira melhor do que sua situação real a fim de captar mais recursos dos investidores, provavelmente com custos de captação mais baixos.

Oliveira (2020) destaca que as ações das entidades que evidenciam, por meio de relatórios financeiros, seu crescimento econômico são mais atrativos aos investidores e, conseqüentemente, mais valorizadas. Para manter sua credibilidade perante os investidores, as entidades tentam disfarçar qualquer sinal que demonstre quedas ou estagnação nos seus indicadores econômicos ou financeiros, recorrendo, entre outros mecanismos, a fraudes em relatórios financeiros, com o objetivo de manter sua capacidade de captação de recursos a custos mais atrativos do que realmente seria, caso seu verdadeiro desempenho econômico fosse demonstrado.

Essa assimetria informacional, segundo Stiglitz (2000), pode gerar conseqüências relevantes no mercado, podendo levá-lo ao desequilíbrio. Em outras palavras, pequenas informações distorcidas, quando publicadas no mercado, podem gerar conseqüências inesperadas; entre elas, tem-se a possibilidade de existência de retornos inesperados ou considerados acima do normal. Esses retornos anormais podem se dar em razão de fraudes nas demonstrações financeiras.

Sob o aspecto das motivações, como a obtenção de retornos anormais para a entidade fraudadora, Eryigit (2019) apresenta evidências de que os retornos anormais causados pelo anúncio de fraudes em relatórios financeiros não são significantes para inferir a sua existência ou não, dado que o anúncio de envolvimento de entidades em fraudes ou irregularidades contábeis foi considerado como não significativo estatisticamente.

Tsai e Huang (2020), Palmrose et al. (2001) e Karajian e Ullah (2021) encontraram retornos anormais negativos no mercado americano nos períodos de 2000 a 2018, de 1995 a 1999 e de 1997 a 2011, respectivamente. No entanto, de acordo com Karajian e Ullah (2021), a reação negativa do mercado ao anúncio de envolvimento de entidades em esquemas de fraudes acaba por ser revertida a longo prazo. Como conseqüência, Zahra, Priem e Rasheed (2007) afirma que quando o público tem conhecimento da fraude realizada pela entidade, imediatamente, o valor do mercado das ações da entidade começa a registrar quedas. Segundo Zahra, Priem e Rasheed (2007) a queda do valor das ações não afeta só a entidade fraudadora, mas, afeta também entidades credoras e parceiras (como clientes e fornecedores).

Os resultados encontrados por Du (2017) relatam a existência de retornos anormais positivos em empresas que se envolveram com fraudes em suas demonstrações financeiras. No entanto, em momento posterior, a entidade sofreu outras consequências, como a falta de confiança na tentativa de contratação de novos créditos.

Já no que diz respeito à periodicidade de retornos para as empresas com o uso de fraudes, Finnerty et al. (2016) afirmam que os retornos anormais podem existir em um horizonte de um, dois e até cinco meses antes e ao redor da data em que a fraude foi realizada ou ao redor da data em que a fraude foi detectada (data de divulgação da fraude pelo órgão regulador). Segundo os autores, os gestores se engajam na fraude com o objetivo de camuflar os problemas da entidade, o que pode ocasionar ações deliberadas que alterem as informações financeiras não somente no curto prazo, mas ao longo de anos, trazendo vieses no processo decisório dos usuários externos.

De acordo com Kothari e Warner (2007), os retornos diários são melhores para capturar comportamentos anormais do que os retornos mensais ou anuais. O uso de retornos que não sejam diários busca variações nos comportamentos dos preços das ações, a longo ou médio prazo, ao redor da data do evento. É evidente que períodos maiores permitem observar o evento em um espaço de tempo maior; no entanto, essa forma de observação pode sofrer influências de outros eventos.

É provável que o mercado tenha reações desde o início das investigações até a decisão condenatória, sendo esta última a data de comunicação oficial ao mercado do envolvimento da entidade em esquemas de fraudes. Sob o aspecto de análise, essa data pode ser considerada como adequada para um estudo de eventos, dado que as variações observadas ao redor de uma data antes da condenação oficial da entidade estariam se baseando em rumores, e não em comunicados ao mercado.

Assim, a data do evento é relevante para o sucesso de um estudo de eventos. Quando se trata de fraudes, a data do evento pode variar, a depender do ponto de vista do pesquisador. De acordo com Eckert et al. (2019), a data do anúncio da fraude é a data em que a fraude foi anunciada oficialmente no mercado.

2.4 Fatores de riscos e detecção de fraudes em relatórios financeiros

Visto que os contratos (entre entidades e investidores) no mercado de capitais são imperfeitos devido à assimetria informacional (situação em que algumas entidades não comunicam todas as informações ao mercado), esta última deve ser mitigada, de forma a eliminar ou pelo menos reduzi-la entre os agentes do mercado de capitais.

De acordo com Segal (2016), as fraudes em relatórios financeiros resultam das falhas dos sistemas de controle interno, por isso o trabalho dos auditores é indispensável. No entanto, nem os auditores nem os princípios contábeis conseguem eliminar as realizações de fraudes. Segundo Sharma e Panigrahi (2012) a falha dos sistemas de auditoria interna em identificar fraudes levou ao uso de métodos da contabilidade investigativa.

Segundo as normas de auditoria International Standard on Auditing (ISA) 205, de 2018 (International Federation of Accountants, 2018), Auditing Standard (AS) 1001, de 2002 (Public Company Accounting Oversight Board, 2002), a Norma Brasileira de Auditoria (NBC TA) 240, de 2009 (Conselho Federal de Contabilidade, 2009a), e a Resolução CFC nº 1.207, de 2009 (Conselho Federal de Contabilidade, 2009b), o auditor é responsável por obter segurança razoável de que as demonstrações contábeis, como um todo, não contêm distorções relevantes, causadas por fraude ou erro. Todas as normas supracitadas, lidam com a responsabilidade do auditor em considerar normas, instruções normativas e regulações na auditoria de demonstrações financeiras. Estas normas não se aplicam quando o auditor está, especialmente, testando a aplicação de uma determinado norma, instrução normativa ou regulação.

Segal (2016) ressalta que, apesar dos esforços normativos antes e depois dos casos de fraudes, como da companhia de energia Enron, as ferramentas de detecção de fraudes em relatórios financeiros são ainda insuficientes para eliminar ou evitar o fenômeno. Contudo, o uso de ferramentas tecnológicas pode contribuir para trazer alertas no que tange ao risco contábil de fraudes de informações (Aghghaleh & Mohamed, 2014). Para o referido autor, a apresentação de relatórios no formato XBRL, o uso de softwares ERP e os métodos de mineração de dados são algumas formas para detectar, de forma eficiente, fraudes em relatórios financeiros.

Segundo Roohani et al. (2009), o chamado Extensible Business Reporting Language Software (XBRL), que é um software de linguagem de relatório corporativo extensivo, seria uma forma de padronização de relatórios, permitindo que os usuários possam ler, analisar e comparar relatórios em diversos países e línguas e de modo eficiente. Esse formato de relatório possibilita reunir os elementos das demonstrações em uma base de dados, facilitando apurar os índices no formato desejado. O XBRL permite recalcular eficientemente o resultado do exercício, a depender de sua composição, de acordo com a norma vigente.

De acordo com Christensen e Byington (2003), as ferramentas utilizadas para cometer fraudes também podem ser usadas para detectá-las. Nesse sentido, Khan et al. (2010) propuseram o sistema de registro de atividades diários (*enterprise resource planning software*

– ERP) para detectar atividades suspeitas de fraudes, pois, em princípio, o ERP detectaria as fraudes antes de elas irem para os relatórios financeiros.

Segundo Segal (2016), a mineração de dados é uma técnica que trata um grande número de dados e que filtra informações de interesse, a exemplo dos modelos de aprendizado de máquina e de inteligência artificial. Por sua vez, Yazici (2020) ressalta que, além de detectar fraudes, os algoritmos DT, RF, ANN, KNN e SVM podem ser usados para extrair informações de bases de dados grandes para detectar se determinado relatório financeiro é fraudulento ou não.

Combinado às técnicas de mineração de dados, para Wilks e Zimbelman (2004), é necessário compreender o ambiente organizacional para tipificá-lo como de maior ou menor risco da ocorrência de fraude. Nesse sentido, os autores (Wilks & Zimbelman, 2004, p. 232) classificaram as situações a seguir como essenciais nessa avaliação:

- Proporções significativas de compensações da gerência (bônus e participações acionárias) dependem de metas agressivas atingidas, dos preços das ações, do resultado operacional, da posição financeira da entidade no mercado ou dos fluxos de caixa;
- Garantias pessoais dos gestores na contratação de dívidas da entidade;
- Parte significativa de transações de controladas classificada como atividades não ordinárias, transações de controladas não auditadas ou auditadas por outra empresa de auditoria;
- Subjetividade ou incerteza na mensuração de ativos, passivos, receitas ou despesas;
- Transações significativas, raras e altamente complexas, sobretudo aquelas que acontecem no fechamento do ano; e
- Ausência ou fraca justificativa às atividades de contas bancárias, entidades subsidiárias, ou às áreas operacionais em paraísos fiscais.

Observa-se que as situações de riscos de fraudes estudadas por Wilks e Zimbelman (2004) são características relativas à entidade, cuja execução não depende apenas de riscos de fraudes relativas a ela, mas depende também de alguns comportamentos individuais que têm participação na realização de fraudes e constituem, então, um risco de fraude. Todos os riscos de fraudes são denominados *red flags*, ou seja, bandeiras vermelhas, que devem ser consideradas na avaliação de riscos de fraudes a fim de detectá-las, prevê-las e discerni-las de erros.

Tanto as *red flags* institucionais quanto comportamentos individuais favorecem a realização de fraudes nos relatórios financeiros, sendo eles caracterizados pelo não atendimento às características qualitativas fundamentais de representação fidedigna e relevância.

De acordo com Rezaee (2002), fraudes em relatórios financeiros são caracterizadas pela omissão ou ocultação intencional de fatos materiais ou dados contábeis e são capazes de

afetar a decisão do usuário da informação contábil quando essas omissões ou ocultações estiverem presentes. Segundo o autor, geralmente, esse tipo de fraude acontece em situações de pressão econômica, como o declínio do desempenho da organização ou do setor de atividade da entidade, ou ainda em situações de recessão.

No entanto, de acordo com Wilks e Zimbelman (2004) e Sandhu (2020), a presença de riscos de fraudes é suficiente para o acontecimento de fraudes em relatórios financeiros, e não depende necessariamente da conjuntura econômica. Os riscos de fraudes, para Rezaee (2002, p. 6) são caracterizados por:

- Falsificação, alteração ou manipulação de lançamentos contábeis, de documentos suportes ou de transações;
- Ocultação ou omissão intencional e material de eventos, transações, contas ou outra informação do relatório;
- Aplicação deliberadamente errônea de princípios ou padrões contábeis, regras ou procedimentos usados para mensuração, reconhecimento, preparação e *disclosure* de relatórios financeiros; e
- Omissão intencional de divulgação ou apresentação inadequada de valores financeiros, em desacordo com as normas contábeis.

Em razão de suas consequências, bem como por meio de indicadores que sinalizam o risco de fraude, para Amiram et al. (2018), as fraudes em relatórios contábeis devem ser detectadas a tempo e prevenidas para ser garantida a eficiência do mercado, a fim de se evitar a perpetuação da anomalia, assim como das suas consequências.

Entre as possíveis consequências, segundo Rezaee (2005), tem-se, como exemplos, a perda de valor das ações no mercado, o pagamento de multas, a perda de confiança e, conseqüentemente, de investidores dispostos a alocar recursos.

Ainda sob a perspectiva do mercado, para Beatty et al. (2013), Joe e Oh (2018), Eckert et al. (2019) e Darrough et al. (2020), a fraude realizada por uma entidade pode gerar consequências para seus concorrentes pelo fato de os investidores reagirem com desconfiança de que outras entidades do mesmo setor ou mercado possam também se utilizar desse expediente, com isso, ocasionar fuga de capitais.

Sob a ótica da reação do mercado às informações tipificadas como fraudulentas, para Fama (1991) e Latif et al. (2011), os retornos anormais são evidências de uma anomalia do mercado (por exemplo, fraude) à qual o mercado reage de uma forma inesperada às informações que não foram precificadas adequada e tempestivamente. Assim, a depender da forma pela qual os investidores entendem a anomalia e suas possíveis consequências, eles podem reagir negativa ou positivamente, comprando e vendendo as ações (da entidade cuja ação é o epicentro

da anomalia) a um preço abaixo ou acima (respectivamente) do usual (normal), o que se reflete na ocorrência de retornos anormais negativos ou positivos.

Segundo Palmrose et al. (2001) e Karajian e Ullah (2021), a fraude é uma das anomalias que tem, como consequência, reações negativas do mercado, o que gera retornos anormais negativos. De acordo com Fama (1991), por ser uma anomalia em que o CAPM não conseguiria prever o preço dos ativos de forma eficiente, pode ser usado para testar a eficiência do mercado. Isto é, verifica-se se o anúncio de fraude está incorporado aos preços dos ativos.

Quando a fraude não é detectada até a divulgação dos relatórios financeiros, é provável que as consequências (presença de retornos anormais) iniciem antes do anúncio da fraude. Sob esse aspecto, Gupta e Gill (2012) entendem que, quanto mais cedo a fraude é detectada, menor seriam suas consequências.

Nesse sentido, o desenvolvimento e o uso de métodos de mineração de dados vêm sendo utilizados para se detectar com maior acurácia e predição a probabilidade de se identificar demonstrações financeiras fraudulentas.

Song et al. (2014), por exemplo, estudaram os índices que podem ser usados na tarefa de detecção de fraudes e identificaram 23 índices financeiros por meio de cinco modelos de classificação: (1) regressão logística (LR); (2) redes neurais de retropropagação (*back-propagation neural network* – BPNN); (3) árvore de decisão (*decision tree* – DT); (4) máquina de suporte vetorial (*support vector machine* – SVM) com modelo radial; e (5) um modelo que junta os quatro primeiros à SVM. Tal estudo demonstrou melhor performance entre os modelos considerados individualmente; mas o modelo constituído pelo conjunto dos quatro modelos (*ensembling model*) se mostrou mais eficiente.

Kanapickiene e Grundiene (2015) investigaram variáveis sensíveis à detecção de fraudes por meio da análise de 51 índices financeiros, identificando que os índices “inventários/ativo total”, “vendas/ativo total”, “obrigações totais/ativo total” e “caixa (disponível)/passivo circulante” são mais sensíveis na detecção de fraudes em relatórios financeiros.

Por sua vez, Lin et al. (2015) discutiram vários aspectos de fraudes em relatórios financeiros com a aplicação das dimensões do triângulo de fraudes e a exploração de técnicas para a detecção de fraudes, considerando também a opinião dos profissionais sobre as novas técnicas de detecção. Segundo os autores, a dimensão “pressão ou incentivo” das fraudes em relatórios financeiros é definida por situações que remetem à necessidade de atingir as metas ou às situações de baixo desempenho da entidade. A dimensão “oportunidade” é explicada por situações como transações com controladores, patrimônio líquido complexo, diminuição da

frequência do controle interno ou fatos anteriores que conferem um status desejável (como ser auditada por uma Big 4, o tempo de existência da entidade, entre outros fatores). A dimensão “atitude ou racionalização” é explicada por situações que remetem às *red flags*, como a rotatividade de gestão ou de CEO.

Ainda na aplicação das dimensões do triângulo de fraudes, Lin et al. (2015) identificaram 35 variáveis (financeiras e não financeiras) e três técnicas de mineração de dados (*logistic regression*, *artificial neural network* e *decision tree*) para a atividade de detecção. Em vez de se usar todas as variáveis explanatórias, estas são classificadas por ordem de importância (por meio de questionário aos profissionais), e as mais importantes são incluídas nos modelos.

Ao identificar as métricas usadas na detecção de fraudes em relatórios financeiros, Karpoff et al. (2017) apontaram que os índices financeiros (capital de giro por ativo, novas emissões de títulos de dívida por ativo, números de ações compradas por ativo, retorno por ativo, despesas de capital por ativo, despesas de pesquisa e desenvolvimento por ativo) são significantes na detecção de fraudes.

Bao et al. (2019) estudaram as combinações entre os métodos de aprendizado de máquina (logit, SVM e RUSBoost) e os tipos de dados (índices financeiros e valores dos elementos dos relatórios financeiros) que apresentam mais precisão na detecção de fraude em relatórios financeiros. Os autores concluíram que o método de regressão logit usando os valores (sem nenhuma transformação) dos elementos dos relatórios financeiros apresentou um resultado melhor que o uso de seus índices contábeis. Segundo eles, os índices financeiros causam uma perda informacional que os valores reais não causariam, enquanto que o método unificado do modelo da SVM apresentou também o mesmo resultado. Os autores argumentam que, ao se usar os índices contábeis – como se usa tradicionalmente – como variáveis de análise, muitas informações importantes se perdem, e a análise acaba não sendo muito fiel à amostra. Nesta pesquisa, os índices financeiros e os valores dos elementos das demonstrações financeiras foram explorados para identificar se as referidas demonstrações são fraudulentas ou não.

De acordo com Reurink (2018), quando as fraudes em relatórios financeiros de uma determinada entidade são anunciadas ou detectadas, oficialmente ou não oficialmente, os usuários das informações contábeis, sobretudo os investidores, preocupam-se acerca da forma e do modo que a realização da fraude e/ou a sua detecção impactou, impacta ou impactará seus investimentos.

2.5 Fraude nas demonstrações financeiras e sua repercussão no mercado financeiro

Segundo Sutherland (1945), fraude é uma transgressão, violação ou desobediência a uma lei, uma norma ou um regramento. De acordo com Cressey (1953), três fatores (pressão, oportunidade e racionalização) levam os indivíduos (sejam eles pessoas físicas ou jurídicas) a cometerem uma fraude; esses três fatores constituem, desde então, o tradicional triângulo de fraude.

A Association of Certified Fraud Examiners (2020) classifica as fraudes em três grandes categorias: (1) apropriação indébita de ativos; (2) corrupção; e (3) fraudes em relatórios financeiros. Nas duas primeiras categorias do crime, as entidades são as vítimas e os fraudadores são os empregados ou outras pessoas físicas. Na última categoria, as vítimas podem se apresentar nas mais diversas dimensões, demonstrando a consequência nefasta da fraude em relatórios financeiros.

Não por acaso, segundo dados da citada associação (Association of Certified Fraud Examiners, 2020), as fraudes em relatórios financeiros geraram, em média, uma perda de USD 954.000 entre janeiro 2018 e setembro 2019, enquanto as duas primeiras categorias (apropriação indébita e corrupção) geraram, respectivamente, em média, uma perda USD 100.000 e USD 200.000.

Conforme Suh et al. (2018), a fraude nas demonstrações financeiras engloba as fraudes realizadas pela própria entidade a seu favor ou realizadas pelos funcionários em favor da entidade, que, para sua realização, utilizam artifícios como: (1) superestimação de receita; (2) receitas fictícias; (3) aquisição de títulos por um valor superior ao valor justo; (4) desvio de fundos para entidades offshore; (5) capitalização de despesas operacionais; (6) superestimação de ativos; (7) não divulgação de transações; (8) transações com partes relacionadas; (9) subestimação de despesas; e (10) subestimação de ativos. Kostova (2013), por sua vez, aponta cinco meios de fraudes em relatórios financeiros: (1) vendas fictícias; (2) reconhecimento antecipado de receitas, antes de cumprir todas as condições de uma venda; (3) vendas sob condições; (4) relatórios falsificados de vendas do período; e (5) entregas não autorizadas e vendas em consignação.

O estudo de caso da fraude da Kmart por Condé, Almeida e Quintal (2015) demonstra que a fraude em relatórios financeiros acontece, também, por omissão ou não divulgação de fatos relevantes. A omissão ou não divulgação de fatos relevantes viola a característica qualitativa sobre a fidedignidade da informação contábil e é considerado como uma fraude em relatórios financeiros (contábeis).

Segundo Kostova (2013), fraudes em relatórios financeiros têm o objetivo de apresentar situações financeiras inapropriadas, que podem ser expressas por palavras tendenciosas e ambíguas ou figuras forjadas (isto é, o aspecto qualitativo de fraudes em relatórios financeiros), ou, ainda, a omissão, a falsa apresentação ou a divulgação imprecisa de dados numéricos (isto é, o aspecto quantitativo de fraudes em relatórios financeiros).

Observa-se que, entre as três grandes categorias de fraudes – de acordo com a classificação da Association of Certified Fraud Examiners (2020) –, não há possibilidade de erro para uma apropriação indébita e uma corrupção enquanto as fraudes em relatórios financeiros são passíveis de serem confundidas com um erro (quando a irregularidade não é intencional). Logo, enquanto a fraude em relatório financeiro ocorre de modo intencional e voluntário, o erro não é intencional. De acordo com Cressey (1953), a realização de fraudes é causada por situações que podem ser problemas financeiros não compartilhados, que são favorecidos por outras situações que podem ser chamadas de oportunidades, sendo, por fim, a fraude justificada por argumentos que podem ser chamados de racionalização. Vários fatores de riscos podem ser investigados pelos auditores e, posteriormente, ser usados como ponto de partida para que a liderança da gestão da entidade detecte as fraudes.

Considerando que o mercado é eficiente e que as informações divulgadas nela incorporam os preços das ações, supõe-se que as informações não divulgadas não devem afetar o mercado e, portanto, não podem ser incorporadas nos preços das ações. O estudo espera que o mercado esteja indiferente às fraudes cometidas em relatórios financeiros até elas sejam detectadas.

De acordo com Stein (2012), a fraude proporciona algum benefício para a entidade fraudadora, já que, pelo menos, mantém a credibilidade que tinha ou aumenta (fortalece) sua credibilidade. Para manter seu status quo, ao realizar a fraude, a entidade não espera retornos anormais. Todavia, quando o objetivo consiste em aumentar o nível de confiança dos investidores na entidade, são esperados retornos anormais positivos enquanto o mercado não possuir conhecimento do evento fraudulento.

As fraudes da alta gerência (que agrupa as fraudes em relatórios financeiros) têm um efeito devastador nos acionistas da entidade, nos funcionários e podem ruinar a reputação e a credibilidade da entidade (Zahra, Priem e Rasheed, 2007).

Apesar das divergências a respeito das inferências sobre a reação do mercado aos anúncios de fraudes ou de irregularidades contábeis, considerando que o mercado é eficiente e que as informações são incorporadas nos preços das ações e que as fraudes ou irregularidades contábeis são anomalias do mercado, este estudo considera como hipótese que o mercado

brasileiro reaja negativamente depois do julgamento definitivo pela Comissão de Valores Mobiliários. Portanto, tem-se a seguinte hipótese de pesquisa:

H₁: Fraudes em relatórios financeiros causam retornos anormais negativos no preço dos ativos quando são de conhecimento público do mercado.

Em razão da hipótese ora descrita, adicionalmente se busca testar a eficiência da SVM na detecção de fraudes, de modo a apresentar um índice de assertividade na previsão de fraudes em relatórios financeiros no mercado brasileiro em torno de 81%, como evidenciaram Severino e Peng (2021). Assim, este estudo visa contribuir para o desenvolvimento do monitoramento das demonstrações financeiras pelos usuários externos e pelos *gatekeepers*.

3 Metodologia

Com o objetivo de identificar empresas brasileiras que incorreram em fraudes, foram analisados os processos administrativos sancionadores (PAS) conduzidos pela Comissão de Valores Mobiliários de 01/01/2017 a 31/12/2021. Nesse sentido, a amostra da pesquisa é composta pelos relatórios fraudulentos das empresas brasileiras listadas na B3, identificados por meio dos PAS e julgados pela CVM no referido período. Destaca-se que as fraudes julgadas nesse período foram cometidas entre 1998 e 2018.

Vale também ressaltar que os processos administrativos sancionadores são tramitados pelo rito ordinário, julgados pelas áreas técnicas da CVM e apreciados em reunião com o colegiado (Comissão de Valores Mobiliários, 2021). Assim, os relatórios financeiros julgados e que resultam em condenações são considerados, neste estudo, como relatórios financeiros fraudulentos porque há evidências atestadas de fraudes. A figura 1 apresenta as estratégias da CVM para enfrentar fraudes e/ou irregularidades na B³

Figura 1 – Estratégias/valores da CVM para enfrentar fraudes ou irregularidades no mercado



Fonte: Comissão de Valores Mobiliários (2020).

Depreende-se do propósito e valores da Comissão de Valores Mobiliários (CVM,2020) a intenção de deixar a B³ mais eficiente o regulando de forma a evitar qualquer anomalia. Neste mesmo espírito, o órgão definiu seu primeiro resultado que consiste em garantir a integridade do mercado e seu segundo objetivo consiste em estimular a eficiência do mercado. Seus dois primeiros objetivos finalísticos (“Aumentar a eficiência da supervisão com uso de inteligência e novas tecnologia” e “Aumentar a eficiência da atuação sancionadora de acordo com as prioridades da supervisão”) levam o regulador a atingir seu primeiro objetivo organizacional que consiste em ampliar a capacidade de estruturação e análise de dados para cumprimentos dos mandatos legais.

Segundo a Comissão de Valores Mobiliários (CVM, 2020) sua cadeia de valores consiste em normatizar o mercado, supervisionar o mercado, sancionar o mercado, fornecer orientações ao mercado e à sociedade e elaborar estratégias de gestão e suportes para sustentar os valores. No que diz respeito ao enfrentamento às fraudes (não somente fraudes em relatórios financeiros), seus três primeiros valores (figura 1) são importantes para inibir as consequências das fraudes ou irregularidades em relatórios financeiros no mercado.

3.1 Amostra para a análise das consequências de fraudes em relatórios financeiros no mercado

A coleta de dados da primeira etapa foi realizada seguindo o procedimento utilizado por Nascimento (2020), que consiste nos processos inteiramente julgados e com acusação pela CVM de fraude, risco de fraude, erro ou com uma determinação final de correções nos relatórios da entidade acusada. De acordo com Nascimento (2020, p. 34), o acesso aos processos julgados pela CVM é feito no site do órgão, seguindo o procedimento a seguir: site da CVM > Processos > Pesquisa avançada. O campo “Termo” é deixado em branco; “período”, de 01/01/2017 a 31/12/2021; e, por fim, o campo “Tipo” deve estar marcado como “processos sancionadores julgados”.

A partir da primeira consulta, foram identificados 283 processos, dos quais 167 estão relacionados a instituições financeiras ou a entidades de investimentos, 4 tratam de fraudes cometidas por pessoas físicas (investidores) sem vínculo com nenhuma entidade e 2 processos cujos acusados foram inocentados. Dado que a análise se concentrou na ocorrência de fraudes em demonstrações contábeis de empresas não financeiras, a amostra final representou 110 processos analisados.

Em alguns casos, o mesmo processo possui como objeto o julgamento de mais de um relatório (fraudes cometidas em relatórios de vários anos), e, desse modo, os 110 processos ora analisados dizem respeito a um total de 129 relatórios contábeis julgados como fraudulentos, sendo assim distribuídos: 99 processos trataram apenas de um relatório (investigação de 99 relatórios financeiros); 7 processos trataram de 2 relatórios de anos distintos (investigação de 14 relatórios financeiros); 2 processos trataram de 3 relatórios de anos distintos (investigação de 6 relatórios financeiros); e, por fim, 2 processos trataram de 5 relatórios de anos distintos (investigação de 10 relatórios financeiros). A Tabela 1 apresenta os critérios de exclusão, o número de processos e de relatórios analisados por grupo amostral.

Tabela 1 – Amostra/processos

Grupos da amostra	Número de processos por grupo amostral	Número de relatórios analisados
Amostra inicial	283	
Processos excluídos	(173)	
Instituição financeira (bancos e instituições de investimento)	167	
Julgamento de pessoas físicas	4	
Processos em que os réus foram inocentados	2	
Processos analisados	110	
Número de processos que trataram de 1 relatório	99	99

Número de processos que trataram de 2 relatórios	7	14
Número de processos que trataram de 3 relatórios	2	6
Número de processos que trataram de 5 relatórios	2	10
Total	110	129

Fonte: elaboração própria.

Dos 129 relatórios fraudulentos identificados, 53 pertencem a entidades que não possuem ações negociadas na B3, 3 relatórios pertencem a entidades cujas ações foram canceladas, 1 processo faz referência a 2 relatórios de uma entidade em que não houve divulgação de relatórios por dois anos consecutivos. Pelo fato de mais de um processo poder tratar de um mesmo relatório, 5 relatórios foram tratados em 2 processos e 1 relatório foi tratado em 3 processos. Estes 66 relatórios foram retirados da amostra.

Após o tratamento realizado, tem-se uma amostra de 63 relatórios, os quais provêm de 36 entidades diferentes, sendo que, no período deste estudo: 20 entidades emitiram apenas 1 relatório fraudulento; 10 entidades diferentes emitiram 2 relatórios fraudulentos, cada uma; 2 entidades emitiram 3 relatórios fraudulentos; 1 entidade emitiu 4 relatórios fraudulentos; 1 entidade emitiu 5 relatórios fraudulentos; e 1 entidade emitiu 7 relatórios fraudulentos.

Portanto, a amostra para a análise do primeiro objetivo do estudo consiste em 63 fraudes em relatórios financeiros (eventos a serem estudados), distribuídas entre 36 ativos listadas na B3 e da seguinte forma: (Tabela 2)¹.

Tabela 2 – Amostra/Relatórios fraudulentos

Nº de relatórios fraudulentos emitidos por entidade	Nº de entidade	Relatórios fraudulentos
1	20	21
2	10	20
3	2	6
4	1	4
5	1	5
7	1	7
Total	36	63

Fonte: elaboração própria.

3.2 Coleta e análise de dados para o modelo de estudo de eventos para a análise das consequências de fraudes em relatórios financeiros no mercado

A segunda etapa consiste na coleta que foi realizada por meio da base de dados Economatica®. Nessa etapa da pesquisa, foram extraídos os preços diários de fechamento dos

¹ Para maior detalhamento acerca da amostra final dos relatórios fraudulentos, vide Apêndice 5.

63 relatórios das entidades acusadas de fraude (entidades identificadas no período de análise) e identificadas na primeira etapa da coleta. Essa variável foi utilizada para testar a eficiência do mercado no que diz respeito à realização de fraudes. O modelo de estudo de eventos requer a série histórica dos preços das ações, a ser analisada, e a do mercado (Índice Bovespa). Os códigos das ações da amostra analisada são apresentados no Quadro 1.

De acordo com Oliveira e Lemes (2011), a convergência contábil tornou-se obrigatória a partir de 1º de janeiro de 2010. Isso fez com que a composição do resultado do exercício mudou, a magnitude das depreciações aumentou. Por esta razão, a análise do estudo de eventos foi feita em dois períodos diferentes. No primeiro momento, a análise de estudo de eventos foi feita considerando apenas os relatórios emitidos até 2009. No segundo momento, a análise foi feita considerando apenas os relatórios emitidos a partir de 2010.

Além do preço das ações e do mercado, a execução do modelo de estudo de eventos requer também os preços da ação do mercado, as datas dos eventos, a forma de agrupamento de dados, a data do início da janela de eventos, a data final da janela de eventos, a data do início da janela de estimação e o tamanho da janela de estimação.

Os preços das ações analisadas, assim como os preços da ação do mercado, foram extraídos do site <http://yahoo.finance.com>. Aos códigos de identificação das ações, como elas não são identificadas na B3, “.SA” teve de ser adicionado, a fim de identificá-las. A ação do mercado é identificada por “^IBOV” no yahoo.finance.

Assim, dois eventos são estudados nesta pesquisa. Em primeiro lugar, foi estudado o efeito do evento “fraude” nos preços das ações. Nesta análise, o comportamento dos preços das ações antes e depois da data em que a fraude foi realizada foi estudado. A tentativa de identificação desta data não foi totalmente exitosa, pois apenas a data exata da realização de fraudes de 25 dos 63 relatórios fraudulentos pesquisados (ou seja, 39%) foi identificada nos processos, enquanto que, na maior parte dos casos (61%), os processos não mencionam a data de realização da fraude, já que estas podem ocorrer várias vezes durante um exercício contábil.

Considerando a ausência da data exata de realização da fraude, como proxy a esta data, foi utilizada a data do fechamento do demonstrativo, objetivando, assim, padronizar as datas. Adicionalmente, pode-se afirmar que o uso da data de fechamento como a data de realização da fraude é adequado aos objetivos da pesquisa, pois, entre a data exata da realização da fraude até a data de fechamento do relatório, a entidade teria como corrigir o lançamento fraudulento fazendo um lançamento de ajuste e evitando a realização da fraude.

Portanto, a data do primeiro evento é a data da realização da fraude. Por se tratar de fraudes em relatórios financeiros, infere-se que as consequências da anomalia iniciariam no

mercado desde o momento em que o relatório foi divulgado, sendo este o marco para a identificação do primeiro evento e a determinação de sua janela.

O segundo evento a ser analisado é a data do julgamento final pela CVM, ou seja, da acusação de fraude. A data da detecção é, então, a data do processo administrativo sancionador (PAS) da CVM. Dessa maneira, ressalta-se que o evento de realização de fraudes e o evento de detecção de fraudes são independentes.

Conforme Soares et al. (2002), o período em que os retornos das ações serão testados é a janela de eventos, sendo esta constituída, cronologicamente, pela janela de estimação, pela janela do evento propriamente dito e pela janela pós-evento. Segundo Peterson (1989), quando se trata de observações diárias, o número de observações para calcular os estimadores que foram utilizados para o cálculo de retornos estimados na janela do evento pode ser entre 100 e 300. No estudo de Reese Jr. e Robins (2017), foram tomados como base os procedimentos utilizados por Peterson (1989) ao se considerar 100 observações na janela de estimação, 20 na janela de eventos (sendo 10 observações antes e depois da data do evento) e 100 observações na janela pós-evento. Apesar de Eryigit (2019) mencionar sobre a ausência de consenso quanto ao número de observações necessárias na janela de estimação, nesta pesquisa também foram utilizados os procedimentos adotados por Peterson (1989), considerando 150 observações na janela de estimação, com uma janela ajustada de eventos com 11 observações (sendo 5 antes e depois e a observação do dia do evento) e 21 observações (sendo 10 antes e depois e a observação do dia do evento), procedimento este também adotado por Eryigit (2019).

3.3 Definição das variáveis para a análise das consequências das fraudes em relatórios financeiros

De acordo com Eryigit (2019), o processo da análise de um estudo de eventos consiste, em um primeiro momento, em calcular retornos esperados na janela do evento escolhida, baseando-se nos retornos reais observados na janela de estimação. As variáveis necessárias para a obtenção dos retornos esperados são o preço das ações das entidades selecionadas para a análise e o preço da ação do mercado.

Com o objetivo de diminuir os possíveis erros de manipulação de planilhas com séries históricas de ações, os códigos das ações foram colocados no formato do yahoo.finance a fim de se extrair as séries históricas das cotações diretamente do yahoo.finance para o R Studio, visando executar as análises. A codificação do yahoo.finance consiste em adicionar “.SA” à codificação original da B3. Por exemplo, a ação “PETR3” da Petrobras (como é codificada na

B3) é recodificada como “PETR3.SA” no yahoo.finance. A ação do mercado B3 codificada como “IBOV” é recodificada como “^BVSP” no yahoo.finance.

Partindo dessa variável (preços das ações), as variáveis como retornos reais observados ($R_t = \ln \frac{P_t}{P_{t-1}}$), retornos esperados ($ER_t = \alpha + \beta R_{Mt}$), retornos anormais ($AR_t = R_t - ER_t$) e as médias dos retornos anormais acumulados ($CAAR_t = \sum_{i=-1}^t \frac{1}{N} \sum_{i=1}^N AR_{i,t}$) são calculadas como parte da análise do modelo de estudo de eventos.

As médias dos retornos anormais acumulados são utilizadas para identificar, em geral, o efeito do evento estudado nos preços das ações. Os detalhes das variáveis estão apresentados na seção 3.4 que trata do desenvolvimento do modelo de estudo de eventos.

Segundo Reese Jr. e Robins (2017), dos preços das ações (P_t) são determinadas os retornos (r) reais observados. Baseando-se nos retornos (R_t) observados na janela de estimação (período antes da janela de evento, conforme apresentado na Figura 2, adiante), os estimadores de retornos esperados (α e β) são calculados. Usando os retornos do mercado (R_{Mt}) como variáveis e os estimadores α e β como parâmetros, os retornos estimados (ER_t) são calculados na janela de evento. Neste período, os retornos reais observados (R_t) são confrontados com os retornos estimados (ER_t) para calcular os possíveis retornos anormais (AR_t). As somas das médias dos retornos anormais ($CAAR_t$) são calculadas e testadas com base na estatística t de Student.

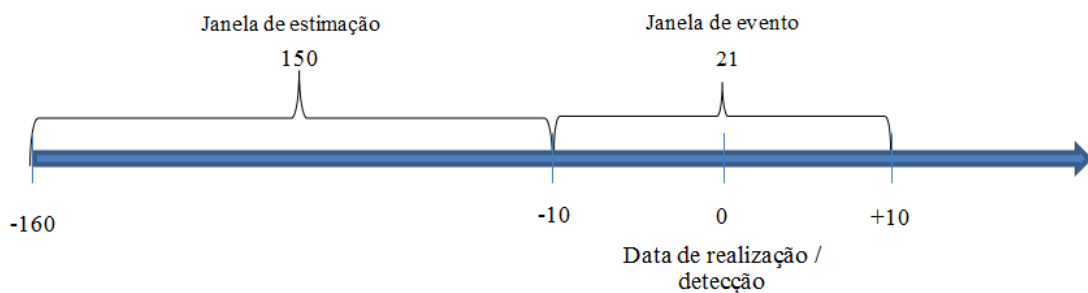
3.4 Desenvolvimento do modelo de estudo de eventos

Segundo MacKinlay (1997), apesar da metodologia de análise de estudo de eventos ter sido criada e utilizada desde 1933, ela violava muitas premissas estatísticas, como a não preocupação com eventos simultâneos. Somente quase no fim da década de 1960 que Ball e Brown (1968) e Fama et al. (1969) trouxeram como procedimento retirar dos preços das ações, a serem analisados sob um evento, os efeitos de outros eventos que aconteceriam simultaneamente com o evento em análise. Brown e Warner (1980, 1985), ao estudarem a temporalidade dos preços das ações no estudo de eventos e a sua capacidade de detectar retornos anormais, revelaram que a temporalidade não afeta a detecção de retornos anormais.

De acordo com MacKinlay (1997), o estudo de eventos pode ser utilizado em finanças para estudar o comportamento dos preços acerca de eventos como mudanças de gestão, aquisições, anúncios de distribuição de eventos, contratações de uma nova dívida e anúncios dos agregados econômicos.

Conforme Sorescu et al. (2017), o estudo de eventos pode ser utilizado para investigar a reação do mercado aos anúncios voluntários da entidade, como os anúncios de introdução de novos produtos, a formação de alianças, a reestruturação de cadeia, ou, ainda, os anúncios de outras entidades que tenham uma relação com a entidade em estudo, como os anúncios dos órgãos reguladores ou dos competidores. O modelo de estudo de eventos pode ser esquematizado como segue na Figura 2.

Figura 2 – Estudo de eventos à realização da fraude



Fonte: elaboração própria.

O procedimento metodológico adotado nesta pesquisa segue o que foi utilizado por Reese Jr. e Robins (2017) considerando apenas o modelo de mercado, denominado de modelo ajustado ao risco, pois, de acordo com Soares et al. (2002), este modelo de apuração de retornos anormais é o mais usado e demonstra mais eficácia à medida que remove a influência do mercado em geral (Corrado, 2011).

Conforme Reese Jr. e Robins (2017), nessa janela, para cada dia, o retorno da ação é regredido com o retorno do mercado, conforme equação a seguir:

$$E(R) = \alpha + \beta(R_M) \quad (1)$$

Onde:

- $E(R)$ é o retorno esperado da ação;
- R_M é o retorno do mercado;
- α é o intercepto da regressão; e
- β é o coeficiente da regressão.

Segundo Soares et al. (2002), há duas formas de se obter o retorno de uma ação: a forma tradicional e a logarítmica. A forma tradicional parte do princípio de que o preço de uma ação em um período dado é uma capitalização discreta do preço da mesma ação no período anterior a uma taxa de retorno dada, conforme equação a seguir.

$$P_t = P_{t-1}(1 + r) \quad (2)$$

Onde:

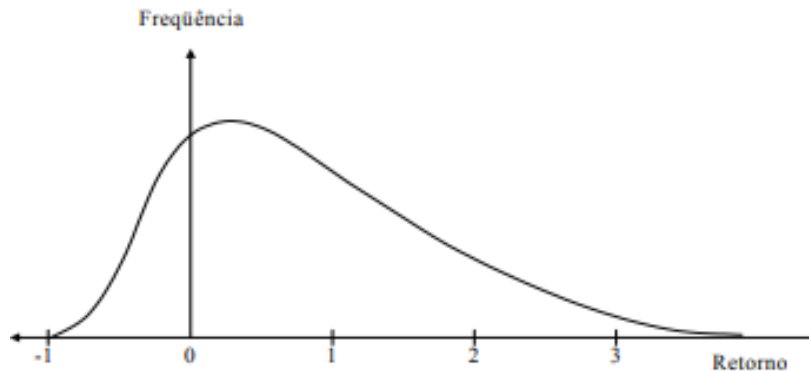
- r é a taxa de retorno;
- P_t é o preço da ação na data; e
- P_{t-1} é o preço da ação na data $t-1$.

Assim, a Equação 3 apresenta o retorno de uma ação:

$$r = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (3)$$

De acordo com Soares et al. (2002), com o retorno calculado desta forma, observa-se uma assimetria na distribuição. Visto que há duas possibilidades, ou seja, retornos negativos e positivos com probabilidades iguais e considerando que não pode haver preços negativos e que os preços são aleatórios, quando P_t é maior que P_{t-1} , os retornos se distribuem no intervalo $[0; +\infty]$. Entretanto, quando P_t é menor que P_{t-1} , os retornos se distribuem no intervalo $[-1; 0]$. Quando P_t é igual a P_{t-1} , o retorno é igual a 0. A distribuição da frequência dos retornos obtidos pela Equação 3 se apresenta na Figura 3.

Figura 3 – Distribuição da frequência de retornos compostos



Fonte: Soares et al. (2002, p. 5).

A segunda forma de se calcular o retorno de uma ação, segundo Soares et al. (2002), é a forma logarítmica. Em vez de uma capitalização composta, ela considera uma capitalização contínua. Assim:

$$P_t = P_{t-1}e^{Rt} \quad (4)$$

Onde:

- r é a taxa de retorno;
- P_t é o preço da ação na data; e
- P_{t-1} é o preço da ação na data $t-1$.

A Equação 4 pode ser reescrita por:

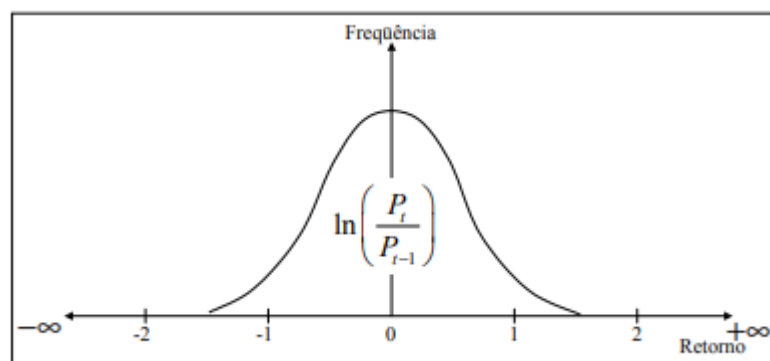
$$e^{Rt} = \frac{P_t}{P_{t-1}} \quad (5)$$

Isolando o (r) , obtém-se:

$$R_t = \ln \frac{P_t}{P_{t-1}} \quad (6)$$

Observa-se uma simetria na curva representativa da distribuição de frequência de retornos logarítmicos. Visto que há duas possibilidades, ou seja, retornos negativos e positivos com probabilidades iguais e considerando que não pode haver preços negativos e que os preços são aleatórios, quando P_t é menor que P_{t-1} , $\ln \frac{P_t}{P_{t-1}}$ tende a $-\infty$, e, quando P_t é maior que P_{t-1} , $\ln \frac{P_t}{P_{t-1}}$ tende a $+\infty$. Assim, a curva da distribuição da frequência de retornos contínuos se apresenta como segue na Figura 4. Essa forma de retorno foi considerada e utilizada neste estudo.

Figura 4 – Distribuição da frequência de retornos contínuos ou logarítmicos



Fonte: Soares et al. (2002, p. 6).

Tendo os retornos do mercado, os estimadores (α_i e β_i), das ações, obtidos na janela de estimação, são utilizados para estimar os retornos na janela do evento. Segundo Park (2004), existem vários modelos – como o modelo do mercado, o modelo do mercado ajustado, o modelo do mercado com o beta (β) estimado de Scoules-Williams, o modelo do mercado com Garch e Egarch, o modelo do fator de Fama-French e o modelo do fator de impulso de Fama-French – para calcular os retornos estimados na janela do evento. De acordo com Corrado (2011), o

retorno estimado na janela do evento, calculado pelo modelo do mercado, é uma das estruturas básicas para a modelagem de estudo de eventos, assim como propõem Fama et al. (1969).

De acordo com Corrado (2011), Reese Jr. e Robins (2017) e Soares et al. (2002), o retorno estimado na janela do evento, pelo modelo do mercado, é obtido pela Equação 7.

$$ER_t = \alpha + \beta(R_{Mt}) \quad (7)$$

Onde:

- R_t é o retorno esperado da ação no dia t ;
- α é o intercepto obtido pela Equação 1 na janela de estimação;
- β é o coeficiente obtido pela Equação 1 na janela de estimação; e
- R_{Mt} é o retorno do mercado do mercado no dia t .

De acordo com Corrado (2011), Reese Jr. e Robins (2017) e Soares et al. (2002), o retorno anormal, pelo modelo do mercado, é obtido pela Equação 8, subtraindo o retorno esperado do retorno real observado. A fim de testar estatisticamente os retornos anormais, estes são calculados não somente na janela do evento, mas também na janela de estimação.

$$\begin{aligned} AR_t &= R_t - (\alpha + \beta(R_{Mt})) \\ AR_t &= R_t - ER_t \end{aligned} \quad (8)$$

Onde:

- AR_t é o retorno anormal no dia t ;
- R_t é o retorno real da ação na data t ;
- α é o intercepto obtido pela Equação 1 na janela de estimação;
- β é o coeficiente obtido pela Equação 1 na janela de estimação; e
- R_{Mt} é o retorno do mercado do mercado no dia t .

O cálculo retorno anormal (AR_t) é suficiente para identificar a presença ou não de retornos anormais individuais (ou seja, para cada ação analisada). No entanto, uma análise cumulativa dos eventos traria uma visão global dos efeitos do evento em estudo; segundo Liu et al. (2020), a média acumulada dos retornos ($CAAR_t$) anormais é a melhor medida para identificar o impacto de eventos nos preços das ações. Visto que várias entidades cometeram mais de uma fraude no período estudado, uma análise cumulativa, reunindo os efeitos de vários períodos de uma mesma ação, foi feita.

$$CAAR_t = \sum_{t=-1}^t \frac{1}{N} \sum_{i=1}^N AR_{i,t} \quad (9)$$

Onde:

- $CAAR_t$ é a média acumulada dos retornos;

- $t = (0, 1, 2, 3, \dots, 10)$ são os dias depois dos eventos estudados; e
- N é o número total de ações analisadas.

Segundo Corrado (2011), os retornos anormais obtidos são testados pelo teste t de Student, assumindo que são distribuídos normalmente. De acordo com Iew (2020), o teste dos retornos anormais antes e depois do evento é suportado na hipótese nula de que os retornos anormais são iguais a zero e na hipótese alternativa de retornos anormais diferentes de zero².

$$H_0: CAAR_t = 0 \quad (10)$$

$$H_1: CAAR_t \neq 0 \quad (11)$$

O Quadro 1 apresenta os parâmetros ao se executar o modelo de estudo de eventos um dia antes e depois da data de realização da fraude.

Quadro 1 – Parâmetros para a execução do modelo de estudo de eventos

Identificação do evento	Ação das firmas	Ação do mercado	Data do evento	Agrupamento	Início da janela do evento	Fim da janela do evento	Fim da janela de estimação	Tamanho da janela de estimação
Antes da convergência contábil (até 2009)								
10010	CEBR6.SA	^BVSP	31/12/2009	Addition	-1	1	-2	150
10013	IMBI3.SA	^BVSP	31/12/2002	Addition	-1	1	-2	150
10017	EMAE3.SA	^BVSP	31/12/1998	Addition	-1	1	-2	150
10018	EMBR3.SA	^BVSP	31/12/2009	Addition	-1	1	-2	150
10030	LIXC3.SA	^BVSP	31/12/2001	Addition	-1	1	-2	150
10044	PETR3.SA	^BVSP	30/12/2005	Addition	-1	1	-2	150
10045	PETR3.SA	^BVSP	31/12/2008	Addition	-1	1	-2	150
10047	PTNT3.SA	^BVSP	31/12/2009	Addition	-1	1	-2	150
10053	RJCP3.SA	^BVSP	31/12/2009	Addition	-1	1	-2	150
10061	SULT3.SA	^BVSP	31/12/2001	Addition	-1	1	-2	150
Depois da convergência contábil (a partir de 2010)								
10001	BAHI3.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150
10002	BMYB34.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10003	BUET3.SA	^BVSP	30/12/2016	Addition	-1	1	-2	150
10004	CBMA3.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150
10005	CBMA3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10006	CBMA3.SA	^BVSP	31/12/2015	Addition	-1	1	-2	150
10007	CEBR6.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10008	CEBR6.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10009	CEBR6.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150
10011	CGRA3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10012	CMIG3.SA	^BVSP	30/12/2016	Addition	-1	1	-2	150
10014	IMBI3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10015	DUQE3.SA	^BVSP	31/12/2010	Addition	-1	1	-2	150
10016	DUQE3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10019	GGBR3.SA	^BVSP	31/12/2015	Addition	-1	1	-2	150
10020	HETA3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10021	I1HG34.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10022	IGUA3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10023	IMBI3.SA	^BVSP	31/12/2015	Addition	-1	1	-2	150

² No Apêndice 1 são apresentadas questões técnicas relativas às variáveis necessárias, à disposição de dados e às funções dos pacotes necessários para a execução do estudo de eventos no R.

10024	INEP3.SA	^BVSP	31/12/2018	Addition	-1	1	-2	150
10025	INEP3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10026	JBSS3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10027	JBSS3.SA	^BVSP	29/12/2017	Addition	-1	1	-2	150
10028	JHSF3.SA	^BVSP	30/12/2016	Addition	-1	1	-2	150
10029	LIXC3.SA	^BVSP	30/12/2016	Addition	-1	1	-2	150
10031	MNDL3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10032	MPLU3.SA	^BVSP	29/12/2017	Addition	-1	1	-2	150
10033	MSPA3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10034	MSPA3.SA	^BVSP	29/12/2017	Addition	-1	1	-2	150
10035	OGXP3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10036	OGXP3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10037	OSXB3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10038	PDGR3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10039	PETR3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10040	PETR3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10041	PETR3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10042	PETR3.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150
10043	PETR3.SA	^BVSP	31/12/2010	Addition	-1	1	-2	150
10046	PTNT3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10048	RCSL3.SA	^BVSP	31/12/2015	Addition	-1	1	-2	150
10049	RCSL3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10050	RCSL3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10051	RCSL3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10052	RCSL3.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150
10054	RJCP3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10055	RSID3.SA	^BVSP	31/12/2010	Addition	-1	1	-2	150
10056	SCLO3.SA	^BVSP	31/12/2012	Addition	-1	1	-2	150
10057	SCLO3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10058	SCLO3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10059	SPRI3.SA	^BVSP	31/12/2014	Addition	-1	1	-2	150
10060	SULT3.SA	^BVSP	31/12/2015	Addition	-1	1	-2	150
10062	TASA3.SA	^BVSP	31/12/2013	Addition	-1	1	-2	150
10063	TENE3.SA	^BVSP	30/12/2016	Addition	-1	1	-2	150
10001	BAHI3.SA	^BVSP	30/12/2011	Addition	-1	1	-2	150

Fonte: dados da pesquisa.

A forma de agrupamento condiciona a forma como os retornos anormais acumulados ($CAAR_t$) foram calculados. O agrupamento “Addition” consiste em calcular o ($CAAR_t$) para todas as ações analisadas.

3.5 Coleta e análise de dados para a análise da detecção de fraudes em relatórios financeiros

Para esta análise, além de relatórios fraudulentos, a amostra constitui-se também de relatórios não fraudulentos. Seguindo os procedimentos de Ravisankar et al. (2011), Mohammadi et al. (2020) e Shen et al. (2021), a amostra, para a análise do segundo objetivo deste estudo, foi construída a partir de relatórios financeiros fraudulentos e não fraudulentos em quantidades iguais. Dos 63 relatórios fraudulentos identificados, não foi possível obter informações financeiras de 7 ações, finalizando, portanto, em uma amostra de 56 relatórios tipificados como fraudulentos e em igual quantidade como não fraudulentos.

Para cada relatório tipificado como fraudulento (RF) foi adicionado à base um relatório não fraudulento (RNF) da mesma entidade, seguindo os procedimentos de Ravisankar et al. (2011), Mohammadi et al. (2020) e Shen et al. (2021).

Martins e Ventura Júnior (2020) ajustaram com dados brasileiros o modelo *probit*, de Beneish (1999), da probabilidade de um relatório conter distorções (fraudes) e chegaram ao seguinte resultado:

$$M = \frac{1}{1 + e^{-(1,635 + 0,251DSRI + 0,462GMI - 0,231AQI + 0,113SGI - 1,075DEPI + 0,139SGAI + 0,014LVGI + 0,058TAT)}}$$

Onde:

- **DSRI** = $\frac{(\text{Contas a receber}_{it}/\text{Receita}_{it})}{(\text{Contas a receber}_{it-1}/\text{Receita}_{it-1})}$ é o índice de recebíveis por vendas;
- **GMI** = $\frac{\text{Margem Bruta}_{it}}{\text{Margem Bruta}_{it-1}}$ é o índice de margem bruta;
- **AQI** = $\frac{1 - ((\text{Ativo Circulante}_{it} + \text{Imobilizado}_{it})/\text{Ativo Total}_{it})}{1 - ((\text{Ativo Circulante}_{it-1} + \text{Imobilizado}_{it-1})/\text{Ativo Total}_{it-1})}$ é o índice de qualidade dos ativos;
- **SGI** = $\frac{\text{Receita}_{it}}{\text{Receita}_{it-1}}$ é o índice de crescimento das vendas;
- **DEPI** = $\frac{(\text{Depreciação}_{it}/(\text{Depreciação}_{it} + \text{Imobilizado}_{it}))}{(\text{Depreciação}_{it-1}/(\text{Depreciação}_{it-1} + \text{Imobilizado}_{it-1}))}$ é o índice de depreciação;
- **SGAI** = $\frac{\text{Despesas com vendas e administrativas}_{it}/\text{Receita}_{it}}{\text{Despesas com vendas e administrativas}_{it-1}/\text{Receita}_{it-1}}$ é o índice de despesas de vendas, gerais e administrativas;
- **LVGI** = $\frac{\text{Exigível Total}_{it}/\text{Ativo Total}_{it}}{\text{Exigível Total}_{it-1}/\text{Ativo Total}_{it-1}}$ é o índice de alavancagem; e
- **TATA** = $\frac{(\text{Lucro Líquido}_{it} - \text{Caixa das operações}_{it})/\text{Ativo Total}_{it}}{(\text{Lucro Líquido}_{it-1} - \text{Caixa das operações}_{it-1})/\text{Ativo Total}_{it-1}}$ é o índice de *accruals* totais pelos ativos totais.

De acordo com Martins e Ventura Júnior (2020), quando o M-score de um relatório é superior a 80%, é muito provável que a entidade tenha manipulado o resultado. Para uma mesma entidade, foi considerado como relatório não fraudulento o relatório imediatamente posterior ao seu relatório fraudulento e que tenha um M-score inferior a 80%.

Para identificar os relatórios não fraudulentos, o M-score dos relatórios posteriores ao relatório fraudulento foi calculado, fazendo uso dos coeficientes do modelo de Martins e Ventura Júnior (2020) até que o M-score seja inferior a 80%.

O relatório imediatamente posterior ao relatório fraudulento foi emitido no período entre a publicação do relatório fraudulento e a investigação da CVM. É provável que, na sua tarefa de investigação de possíveis distorções em um relatório de um exercício dado, a CVM inclua pelo menos o relatório do exercício seguinte. Se o relatório em questão (relatório imediatamente depois de um relatório fraudulento) não foi indiciado como fraudulento e se seu

M-score for inferior a 80%, ele estaria livre de distorções materiais e, portanto, seria considerado como não fraudulento. Suponhamos o relatório fraudulento de 2011 da ação BAH13, que foi julgado pela CVM, em 2019; todos os relatórios entre 2012 e 2019 com um M-score inferior a 80% são tipificados como não fraudulentos, a menos que alguma fraude em relatório fraudulento seja identificada em um exercício.

Considerando as explicações dos parágrafos anteriores desta seção, interpreta-se como um relatório não fraudulento o relatório do ano consecutivo ao ano de um relatório fraudulento e com um M-score inferior a 80%.

Como mencionado anteriormente, de acordo com Oliveira e Lemes (2011), a convergência contábil tornou-se obrigatória a partir de 1º de janeiro de 2010. A partir de 2010, as entidades listadas no mercado brasileiro são obrigadas a apresentar suas demonstrações de acordo com o formato internacional. Com efeito, a composição do resultado mudou alguns índices financeiros. Por isso, a amostra foi dividida em duas partes: a primeira parte apresenta os relatórios fraudulentos e não fraudulentos antes da convergência; a segunda parte, os dados depois da convergência. O Quadro 2 apresenta, por ação, os relatórios tipificados como fraudulentos e não fraudulentos.

Uma exceção foi aplicada em situações em que o relatório não fraudulento, correspondente a um relatório fraudulento do período antes da convergência contábil (até 2009), tenha sido emitido depois de 2009. Nessas situações, com o objetivo de manter relatórios fraudulentos e não fraudulentos no mesmo período (antes da convergência ou depois da convergência), o relatório anterior ao relatório fraudulento foi considerado como não fraudulento (desde que seu M-score seja inferior a 80%).

Quadro 2 – Anos dos relatórios fraudulentos e não fraudulentos

Código da ação	Ano do relatório fraudulento (RF)	Ano do relatório não fraudulento (RNF)	M-score calculado
Antes da convergência contábil (até 2009)			
EMAE3	1998	1999	0.1030
LIXC3	2001	2002	0.0840
SULT3	2001	2002	0.0853
DOCA3	2002	2003	0.0836
PETR3	2005	2006	0.1176
PETR3	2008	2009	0.1428
CEBR6	2009	2008	0.2983

EMBR3	2009	2008	0.1801
PTNT3	2009	2008	0.2793
Depois da convergência contábil (a partir de 2010)			
PETR3	2010	2016	0.0941
RSID3	2010	2011	0.1342
BAHI3	2011	2012	0.1602
CBMA3	2011	2019	0.1487
CEBR6	2011	2014	0.2983
PETR3	2011	2017	0.1440
RCSL3	2011	2016	0.0609
CEBR6	2012	2015	0.2908
IGUA3	2012	2013	0.0000
JBSS3	2012	2013	0.1496
OGXP3	2012	2014	0.0468
OSXB3	2012	2014	0.0334
PDGR3	2012	2013	0.0343
PETR3	2012	2018	0.0059
PTNT3	2012	2013	0.1229
RCSL3	2012	2017	0.0000
SCLO3	2012	2015	0.0470
CEBR6	2013	2016	0.2991
DUQE3	2013	2014	0.1631
HETA3	2013	2015	0.2825
INEP3	2013	2014	0.0801
MNDL3	2013	2014	0.0797
OGXP3	2013	2015	0.1339
PETR3	2013	2015	0.1298
RCSL3	2013	2019	0.1501
SCLO3	2013	2016	0.4510
TASA3	2013	2014	0.1396
CBMA3	2014	2017	0.1422
CGRA3	2014	2015	0.1258
MSPA3	2014	2015	0.0660
PETR3	2014	2010	0.1432
RCSL3	2014	2010	0.2195
RJCP3	2014	2015	0.1362
SCLO3	2014	2017	0.3141
SPRI3	2014	2015	0.2118
CBMA3	2015	2018	0.1447
GGBR3	2015	2016	0.1052
RCSL3	2015	2017	0.0000
SULT3	2015	2016	0.0929
CMIG3	2016	2017	0.1631
JHSF3	2016	2017	0.1101
LIXC3	2016	2017	0.0436
TENE3	2016	2017	0.1346
JBSS3	2017	2018	0.1200
MPLU3	2017	2019	0.1631
MSPA3	2017	2018	0.1143
INEP3	2018	2019	0.4410

Fonte: dados da pesquisa.

3.6 Variáveis para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes

Ravisankar et al. (2011), Song et al. (2014), Kanapickiene e Grundiene (2015) e Bao et al. (2019) estudaram modelos de detecção de fraudes em relatórios financeiros com o uso de

observações binárias (0,1) como variável dependente, em que a observação 1 identifica que o relatório é fraudulento e 0 identifica que o relatório é não fraudulento.

3.6.1 Variável dependente (RF_RNF) para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes

Segundo Ravisankar et al. (2011), Song et al. (2014), Kanapickiene e Grundiene (2015) e Bao et al. (2019), a variável dependente em um modelo de detecção de fraudes em relatórios financeiros, usando os modelos tradicionais de regressão – como a regressão logística – e os métodos de aprendizado de máquina ou de mineração de dados, é binária. Esta variável identifica se o relatório é fraudulento (1) ou não (0), sendo neste estudo a variável dependente codificada como “RF_RNF”.

3.6.2 Variáveis independentes para o modelo de máquina de suporte vetorial (SVM) para a detecção de fraudes

No desenvolvimento do modelo de máquina de suporte vetorial para a detecção de fraudes em relatórios financeiros, Song et al. (2014), Kanapickiene e Grundiene (2015), Karpoff et al. (2017) e Bao et al. (2019) usaram indicadores contábeis como variáveis explanatórias.

Ressalta-se que as variáveis independentes (indicadores financeiros) foram extraídas da plataforma Economatica®, de forma atualizada pelo índice de preço ao consumidor amplo (IPCA).

De acordo com Song et al. (2014), Kanapickiene e Grundiene (2015), Karpoff et al. (2017) e Bao et al. (2019), os indicadores de rentabilidade, eficiência, liquidez e fluxo de caixa estão correlacionados com fraudes em relatórios financeiros. Por esta razão, os indicadores financeiros da Economatica® foram agrupados em três categorias, segundo a classificação da Economatica® (estrutura de capital; liquidez; rentabilidade), sendo adicionada mais uma categoria, denominada “indicadores não financeiros”³. As variáveis coletadas, no primeiro momento, e como elas representam um risco de fraude estão apresentadas no Quadro 3.

Quadro 3 – Indicadores contábeis e sua relação com a realização de fraudes

Categoria	Indicador	Relação com RF_RNF	Autor
Estrutura de Capital	Dívida total bruta (EDC2)	A necessidade de mais endividamento ou de recursos para financiar as atividades é um risco de fraude.	Song et al. (2014)

³ Os indicadores não financeiros utilizados nesta dissertação não são uma classificação categórica de indicadores da Economatica®.

	Dificuldades financeiras aumentam o risco de fraude. O índice de endividamento traduz a proporção do ativo total financiado por recursos de terceiros. Quanto maior este índice, maior é o endividamento da entidade, maior é o risco de fraude.	Kanapickiene e Grundiene (2015) e Ravisankar et al. (2011)
Índice de endividamento (Dívida bruta / Ativo total) (EDC3)	A alavancagem total do patrimônio líquido indica a vulnerabilidade da entidade a um risco de endividamento. Quanto maior este índice, maior é o endividamento da entidade, maior o risco de fraude.	Ravisankar et al. (2011) e Kanapickiene e Grundiene (2015)
(Dívida bruta / Patrimônio líquido) (EDC4)	A alavancagem líquida do patrimônio líquido indica a vulnerabilidade da entidade a um risco de endividamento. Quanto maior este índice, maior é o endividamento da entidade, maior o risco de fraude.	Ravisankar et al. (2011) e Kanapickiene e Grundiene (2015)
Dívida líquida / Patrimônio líquido (EDC5)	O índice de liquidez geral ((Ativo circulante + realizável a longo prazo)/(Passivo circulante + passivo não circulante)) traduz a capacidade de a entidade arcar com suas dívidas a curto prazo. Quando o índice é menor que 1, a entidade não tem capacidade de arcar com suas dívidas, tornando a situação um risco de fraude.	Ravisankar et al. (2011) e Kanapickiene e Grundiene (2015)
Liquidez geral (LIQ1)	O índice de liquidez corrente (Ativo circulante/Passivo circulante) indica a capacidade de a entidade cobrir suas atividades a curto prazo. Quando o índice é menor que 1, a entidade não tem capacidade para cobrir seu passivo a curto prazo, tornando a situação um risco de fraude.	Ravisankar et al. (2011) e Kanapickiene e Grundiene (2015)
Liquidez corrente (LIQ2)	O índice de liquidez corrente ((Ativo circulante – Estoques)/(Passivo circulante)) indica a capacidade financeira de a entidade cobrir seu passivo a curto prazo. Quando o índice é menor que 1, a entidade não tem capacidade para cobrir seu passivo a curto prazo, tornando a situação um risco de fraude.	Ravisankar et al. (2011) e Kanapickiene e Grundiene (2015)
Liquidez seca (LIQ3)	Ainda chamado de capital circulante líquido, o capital de giro (Ativo circulante – Passivo circulante) indica a capacidade de a entidade cobrir seu passivo a curto prazo. Quanto menor o valor, maior o risco de fraude.	Ravisankar et al. (2011), Kanapickiene e Grundiene (2015) e Karpoff et al. (2017)
Capital de giro (LIQ4)	Uma rentabilidade do ativo (Lucro operacional / Ativo total) menor indica um mau uso do ativo ou um risco de fraude (superestimação do ativo). Um indicador muito alto indica que a entidade não está investindo em ativos, o que pode ser um indicador de risco de fraude (Superestimação de receitas).	Ravisankar et al. (2011), Kanapickiene e Grundiene (2015) e Karpoff et al. (2017)
ROA Rentabilidade do ativo (REN7)	Uma rentabilidade do patrimônio líquido (Lucro operacional / Patrimônio líquido) menor indica um mau uso do patrimônio líquido ou um risco de fraude.	Ravisankar et al. (2011), Kanapickiene e Grundiene (2015) e Karpoff et al. (2017)
ROE Rentabilidade do patrimônio (patrimônio final) (REN8)		

Fonte: elaboração própria.

Além dessas variáveis, a sensibilidade de outras variáveis, disponibilizadas pela Economatica®, para detectar fraudes em relatórios financeiros foi testada. Essas variáveis são agrupadas pela plataforma em categorias como “estrutura de capital”, “rentabilidade”, “indicadores não financeiro” e “outras”, sendo elas: “Dívida total líquida” (EDC1); “Estrutura

de capital (Dívida/Dívida + Patrimônio líquido)” (EDC6); “EBIT / Dívida bruta” (EDC7); “EBIT / Dívida líquida” (EDC8); “EBIT / Despesa financeira bruta” (EDC9); “EBIT / Despesa financeira líquida” (EDC10); “Dívida bruta / EBITDA” (EDC11); “Dívida líquida / EBITDA” (EDC12); “*Capital employed*” (LIQ5); “EBITDA” (REN1); “LAIR + Despesa financeira líquida” (REN2); “Margem bruta” (REN3); “Margem EBIT (margem operacional)” (REN4); “Margem líquida” (REN5); “Margem EBITDA” (REN6); “ROE Rentabilidade do patrimônio (patrimônio médio)” (REN9); “ROE Rentabilidade do patrimônio (patrimônio inicial)” (REN10); “ROIC (*Invested capital* final)” (REN11); “ROIC (*Invested capital* médio)” (REN12); “ROIC (*Invested capital* inicial)” (REN13); “*Invested capital* (REN14)”; “Rentabilidade de investimento de subsidiários” (REN15); “Alavancagem financeira” (REN16); “Alavancagem operacional” (REN17); “Pagamento de dividendo” (INF1); “Participação minoritária” (INF2); “Ciclo financeiro” (INF3); “Ciclo operacional” (INF4); “Giro do ativo” (INF5); “Giro do patrimônio líquido” (INF6); “*Capital expenditure* (Despesa de capital incluindo os investimentos)” (OUT1); “Depreciação e amortização” (OUT2); “*Capital expenditure* / Depreciação” (OUT3); “Depreciação / Imobilizado e intangíveis” (OUT4); e “Fluxo de caixa livre” (OUT5).

As variáveis da subcategoria de dados por ação (DPA) de indicadores financeiros informam, em valor monetário, a participação de cada ação em grandezas estratégicas. Quanto maiores forem estes indicadores, melhor é a situação econômica da entidade. Para passar uma boa impressão, uma entidade fraudadora pode manipular suas grandezas estratégicas.

Segundo Ravisankar et al. (2011), quanto maiores são as dívidas, maior é a probabilidade de fraudes em relatórios financeiros. Em razão disso, variáveis de endividamento foram inseridas no modelo de detecção de fraudes, sendo elas: “Log(Dívida total líquida)” (EDC1); “Log(Dívida total bruta)” (EDC2); “Dívida bruta / Ativo total” (EDC3); “Dívida bruta / Patrimônio líquido” (EDC4); e “Dívida líquida / Patrimônio líquido” (EDC6).

De acordo com Dalnial et al. (2014), o risco de fraude em contratos de crédito é maior quando os índices de alavancagem financeira são altos. A alavancagem pode ser medida por “Estrutura de capital (Dívida/Dívida + Patrimônio líquido)” (EDC11), “Dívida bruta / EBITDA” (EDC11) e “Dívida líquida / EBITDA” (EDC12), além da própria “Log(Alavancagem financeira)” (REN16) e da “Alavancagem operacional” (REN17).

Segundo Kanapickiene e Grundiene (2015), os gestores esperam uma rentabilidade maior dos investimentos. Assim, quanto menor a rentabilidade das despesas financeiras, maior seria a probabilidade de acontecimento de fraudes em relatórios financeiros. Os indicadores de rentabilidade considerados e utilizados neste estudo são “EBIT / Dívida bruta” (EDC7), “EBIT

/ Dívida líquida” (EDC8), “EBIT / Despesas financeiras brutas” (EDC9) e “EBIT / Despesas financeira líquida” (EDC10).

Ainda segundo Kanapickiene e Grundiene (2015), problemas de liquidez podem incentivar a realização de fraudes em relatórios financeiros. Para tanto, foram inseridas ao modelo as seguintes variáveis: “Liquidez geral” (LIQ1), “Liquidez corrente” (LIQ2), “Liquidez seca” (LIQ3) e “Log(Capital de giro)” (LIQ4).

Kanapickiene e Grundiene (2015) usaram vários indicadores no seu modelo de detecção de fraudes em relatórios financeiros que motivam a considerar nesta pesquisa os indicadores de rentabilidade da Economatica®, tais como “Log(EBITDA)” (REN1), “Log(LAJIR+Despesas financeiras líquidas)” (REN2), “Margem bruta” (REN3), “Margem de EBIT operacional” (REN4), “Margem líquida” (REN5), “Margem do EBITDA” (REN6), “Rentabilidade do ativo” (REN7), “Rentabilidade do patrimônio final” (REN8), “Rentabilidade do patrimônio médio” (REN9), “Rentabilidade do patrimônio inicial” (REN10), “Retorno do capital final investido” (REN11), “Retorno do capital médio investido” (REN12), “Retorno do capital inicial investido” (REN13) e “Rentabilidade dos investimentos em subsidiários” (REN15).

A Economatica® fornece quatro outros indicadores cuja importância no modelo de detecção de fraudes em relatórios financeiros foi julgada neste estudo. As variáveis são: (1) “Despesa de capital na sua forma logarítmica” (OUT1); (2) “Índice despesa de capital por depreciação” (OUT2); (3) “Índice depreciação por imobilizados e intangíveis” (OUT3); e (4) “Fluxo de caixa livre na sua forma logarítmica” (OUT4).

Como indicadores não financeiros, este estudo utiliza o “Ciclo financeiro” (INF3) e o “Ciclo operacional” (INF4), que, segundo Ramos (2011), trazem noções sobre os prazos das atividades financeiras e operacionais da entidade. O “Giro do ativo” (INF5) e o “Giro do patrimônio líquido” (INF6) trazem noções sobre a quantia de giros do ativo e do patrimônio, respectivamente, para financiar as atividades da entidade.

“Pagamento de dividendo” (INF2) capta a distribuição de “Pagamento de dividendos” (INF1) durante o período do relatório. É uma variável dicotômica que receberá o valor “0” na ausência de distribuição de dividendos e “1” quando houve distribuição.

3.7 Máquina de suporte vetorial (SVM)

A SVM é uma função do método de aprendizagem da inteligência artificial, proposta por Vapnik desde 1995. Segundo Vapnik (2000), com o propósito de distinguir observações em várias classes, ela permite fazer previsões e classificações baseando-se em metodologias de divisibilidade linear e não linear.

A atividade de detecção de fraudes em relatórios financeiros é uma atividade de classificação binária pelo fato de classificar os relatórios em fraudulentos ou não fraudulentos. Conforme Chauhan et al. (2019), existem vários modelos de SVM, tais como Hard SVM, Soft SVM, One-versus-Rest, Transductive SVM, Weston Watkins SVM, One-versus-One, Least Squares SVM, ν -SVM, Directed Acyclic Graph SVM, Smooth SVM, Lagrangian SVM, Proximal SVM, Crammer Singer SVM, Fuzzy SVM, Ev-SVM, Structured SVM, Multi-Task SVM, Fuzzy Proximal SVM, Cascade SVM, GEP SVM, Multi-View SVM, Twin SVM, Sloppy SVM, Capped SVM, C-SVM e Kernel SVM. O modelo Kernel SVM foi considerado e utilizado neste estudo pelo fato de permitir a escolha do modelo de classificação desejado, assim como de outros parâmetros.

O modelo Kernel SVM pode ser executado com quatro tipos de classificação. De acordo com Zhang (2020), o Kernel SVM pode ser linear, polinomial, radial e sigmoide. Cada um desses tipos de classificação foi testado a fim de identificar o tipo de classificação que detectaria melhor as fraudes em relatórios financeiros no Brasil.

O modelo Kernel SVM foi escolhido também por ser flexível à escolha dos parâmetros. De acordo com Rakotomamonjy (2003), o pesquisador tem autonomia de penalizar o modelo por uma falsa classificação (classificar um relatório fraudulento como não fraudulento ou um relatório não fraudulento como fraudulento).

3.8 Máquina de suporte vetorial linear

Segundo Chauhan et al. (2019), o modelo linear da SVM é uma técnica de classificação binária no aprendizado de máquina, o que permite separar as classes com maior margem entre instâncias de fronteiras, chamadas de suportes vetoriais. O problema de tal modelo consiste em otimizar a margem de separação entre as classes. Em outras palavras, o problema de classificação pela SVM procura um hiperplano que separe as observações de forma eficiente. Para isso, de acordo com Chauhan et al. (2019), o problema consiste em maximizar o conjunto de equações do hiperplano. De acordo com Deepa e Dhanapal (2012) e Bashir et al. (2021), a SVM foi desenvolvida a partir da teoria estrutural de minimização de risco, e sua aplicação consiste em resolver a Função 9:

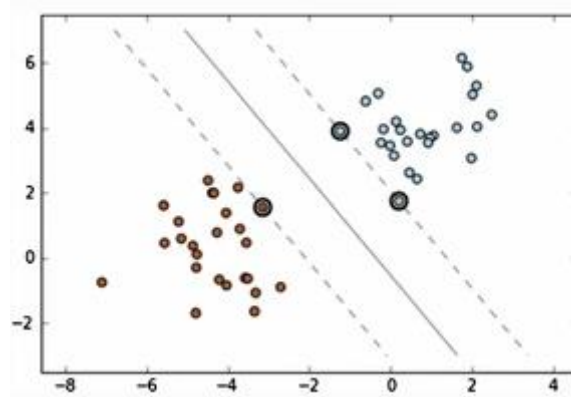
$$\mathbf{f}(\mathbf{x}) = \mathbf{Sinal}(\mathbf{x} \cdot \mathbf{w}) + \mathbf{b} \quad (9)$$

Onde:

- \mathbf{x} representa a matriz das variáveis explanatórias, e
- \mathbf{w} e \mathbf{b} são os parâmetros aprendidos pela SVM na fase de treinamento.

A Figura 5 apresenta a forma pela qual a SVM linear separa as classes. As linhas pontilhadas partindo dos pontos 0 e 4 do eixo horizontal são as fronteiras, e o espaço entre elas é o hiperplano ou, ainda, o suporte vetorial. O separador real entre as duas classes é a linha vertical que divide o espaço entre as fronteiras em duas partes iguais.

Figura 5 – Máquina de suporte vetorial linear



Fonte: Chauhan et al. (2019).

Segundo a metodologia linear, considerando as observações $(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)$, tal que $x \in \mathbb{R}^n$, $y \in \{1; -1\}$, onde os x representam as variáveis independentes e y a variável dependente, o vetor que marca a fronteira das observações positivas (observações coloridas em azul, ou observações do lado direito) é representado pela Equação 12:

$$w \cdot x + b = 1 \quad (12)$$

Por sua vez, a reta que marca a fronteira das observações negativas (observações coloridas em vermelho, ou observações do lado esquerdo) é representada pela Equação 13:

$$w \cdot x + b = -1 \quad (13)$$

A reta que faz a real classificação dos elementos divide o espaço entre as duas fronteiras em duas partes iguais e é representada pela Equação 14:

$$w \cdot x + b = 0 \quad (14)$$

As Equações 12, 13 e 14 formam a margem que separa (classifica) as observações. A Equação 12 divide o hiperplano em duas partes.

Sabe-se que a distância entre o ponto $(x_0; y_0)$ e um reta $A_x + B_y + C$ é $d = \sqrt{A^2 + B^2}$. Então a distância entre a Equação 12 e a Equação 14 é $d = \frac{|w \cdot x + b|}{\|w\|} = \frac{1}{\|w\|}$ e a distância entre

a Equação 13 e a Equação 14 é $\mathbf{d}^- = \frac{|\mathbf{w} \cdot \mathbf{x} + \mathbf{b}|}{\|\mathbf{w}\|} = \frac{1}{\|\mathbf{w}\|}$. A distância total da margem (hiperplano) é $\mathbf{d} = \mathbf{d}^+ + \mathbf{d}^- = \frac{1}{\|\mathbf{w}\|} + \frac{1}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|}$.

Para diminuir os erros de classificação acerca da variável ξ_i , a Função 9 pode ser reescrita por:

$$\mathbf{y}_i(\mathbf{x} \cdot \mathbf{w} + \mathbf{b}) \geq 1 - \xi_i \quad (15)$$

Onde:

$$- \quad i = 1, 2, \dots, n.$$

O problema de otimização para o cálculo \mathbf{w} e \mathbf{b} pode ser definido por:

$$\left\{ \begin{array}{l} \text{Min } \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \\ \text{sob restrição} \\ \mathbf{y}_i(\mathbf{x} \cdot \mathbf{w} + \mathbf{b}) \geq 1 - \xi_i \text{ com } \xi_i \geq 0 \end{array} \right. \quad (15)$$

Usando a função de Lagrange, a Função 5 pode ser reescrita por:

$$\mathbf{L}(\mathbf{w}, \mathbf{b}, \xi, \alpha, \beta) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i \{ \mathbf{y}_i[\mathbf{w} \cdot \mathbf{x} - \mathbf{b}] - 1 - \xi_i \} - \sum_{i=1}^n \beta_i \xi_i \quad (17)$$

A resolução do problema de maximização consiste, desde então, em minimizar \mathbf{w} , \mathbf{b} & ξ e em maximizar α & β . Segundo Deepa e Dhanapal (2012), para resolver o problema, é melhor introduzir a formulação dupla.

$$\max_{\alpha, \beta} \mathbf{w}(\alpha, \beta) = \max_{\alpha, \beta} \left\{ \min_{\mathbf{w}, \mathbf{b}, \xi} \mathbf{w}, \mathbf{b}, \xi, \alpha, \beta \right\} \quad (18)$$

Substituindo, o problema se transforma na seguinte formulação dupla:

$$\left\{ \begin{array}{l} \max \{ \sum_{i=1}^n \alpha - \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbf{y}_i \mathbf{y}_j \langle \mathbf{x}_i, \mathbf{y}_j \rangle \} \\ \text{Sob restrição} \\ \sum_{i=1}^n \alpha_i \mathbf{y}_i = \mathbf{o} \text{ e } \mathbf{o} \leq \alpha_i \leq C \text{ para } i = 1, 2, \dots, n. \end{array} \right. \quad (19)$$

De acordo com Deepa e Dhanapal (2012), aplicando a condição de Kuhn-Tucker à Equação 9, tem-se que:

$$\alpha_i \{ \mathbf{y}_i[\mathbf{w} \cdot \mathbf{x} - \mathbf{b}] - 1 - \xi_i \} = 0 \quad (20)$$

Onde:

$$- \quad i = 1, 2, \dots, n.$$

Os vetores do Lagrange são os suportes vetoriais esperados para descrever o hiperplano. A decisão é determinada pela Função 21:

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{NS} \alpha_i y_i \langle \mathbf{x}, \mathbf{x}_i \rangle + \mathbf{b} \quad (21)$$

Onde:

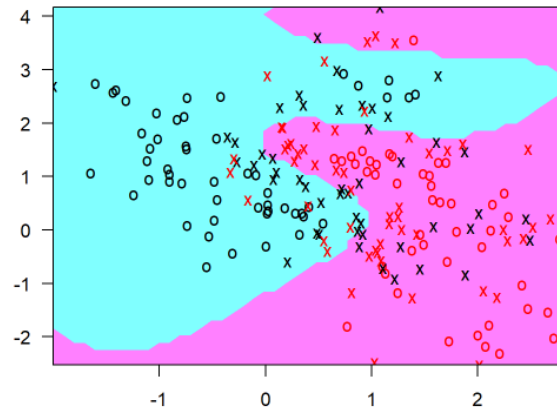
- \mathbf{x} é o vetor das variáveis explanatórias;
- $\langle \mathbf{x}, \mathbf{x}_i \rangle$ é o i -ésimo produto entre as variáveis explanatórias;
- NS é o número de suportes vetoriais; e
- \mathbf{b} é o termo de erro.

O i -ésimo produto entre as variáveis explanatórias ($\langle \mathbf{x}_i, \mathbf{x}_j \rangle$) é chamado de *kernel*. Segundo Chauhan et al. (2019), os *kernels* são funções matemáticas que transformam dados de um espaço dado (espaço de entrada) em um novo espaço de uma dimensão maior (dimensão de saída), onde os dados podem ser separados por superfícies (hiperplanos). Em outras palavras, o *kernel* é uma função que, considerando dois argumentos, retorna o produto dos mapeamentos dos argumentos para o *kernel* linear. As superfícies de separação podem ser lineares, como é o caso de um *kernel* linear. De acordo com Meyer (2021), considerando as matrizes das variáveis \mathbf{x}_i e \mathbf{x}_j , o *kernel* linear seria o produto da transposta da matriz \mathbf{x}_i com a matriz \mathbf{x}_j . Então: *kernel* linear $K(\langle \mathbf{x}_i, \mathbf{x}_j \rangle) = \mathbf{x}_i^T \mathbf{x}_j$. As superfícies de separação podem ser também não lineares; trata-se de um *kernel* não linear.

3.9 Máquina de suporte vetorial não linear

Segundo Chauhan et al. (2019), a SVM linear e a não linear têm a mesma formulação, exceto o espaço no qual a classificação é executada. Enquanto a SVM linear executa a classificação no espaço dos dados em análise, a SVM não linear executa a classificação no espaço das características das observações. No modelo não linear, a atividade consiste em encontrar uma função que classifique melhor as observações de acordo com as características. A Figura 6 apresenta uma classificação não linear:

Figura 6 – Máquina de suporte vetorial não linear

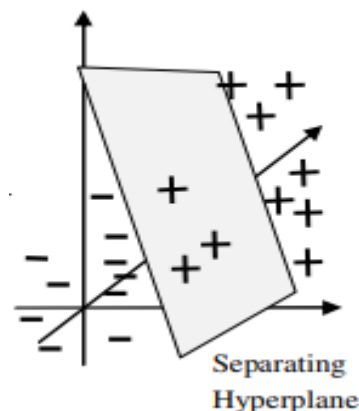


Fonte: Chauhan et al. (2019).

De acordo com Meyer et al. (2021), a execução da SVM não linear no R requer a instalação e a ativação do pacote `e1071`. O pacote `e1071` analisa as possíveis classes identificáveis em um conjunto de dados. Assim, os algoritmos utilizados neste estudo são apresentados no Apêndice 4. Meyer (2021) identificou três tipos de classificação não linear com a SVM. São os modelos de classificação pela SVM com uma função *kernel* polinomial, radial ou sigmoide.

De acordo com Anthony (1995), em vez de um separador linear (como demonstra a Figura 5), a classificação pode ser feita por meio de um separador quadrático, cúbico, em geral, polinomial. Segundo Prajapati e Patle (2010), o *kernel* polinomial baseia-se na direção dos vetores das variáveis para transformá-las em uma dimensão maior. Conforme Meyer (2021), considerando as matrizes das variáveis \mathbf{x}_i e \mathbf{x}_j , o *kernel* polinomial aplica os parâmetros (γ , \mathbf{d} , \mathbf{C}_0) da seguinte forma: $(\gamma \mathbf{x}_i^T \mathbf{x}_j + \mathbf{C}_0)^d$. A Figura 7 apresenta uma classificação polinomial:

Figura 7 – Máquina de suporte vetorial polinomial



Fonte: Prajapati e Patle (2010).

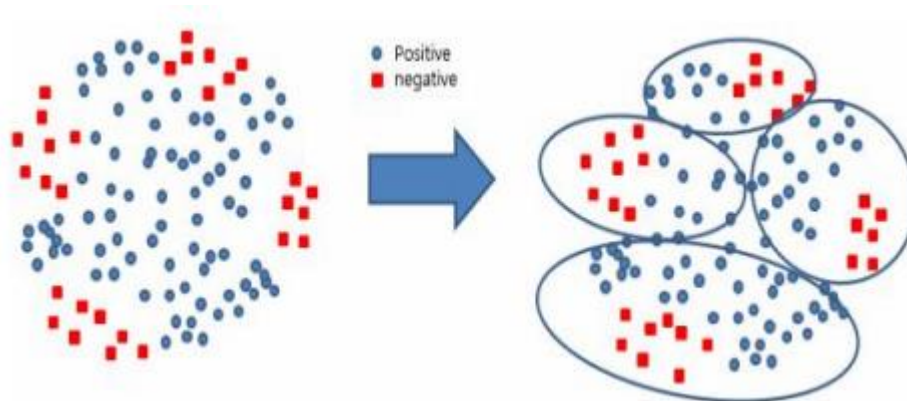
Segundo Wang et al. (2018), em situações em que os dados são muito aglutinados, não é possível fazer uma classificação considerando todos os dados juntos. Uma forma eficiente de realizar a classificação é dividir a base em pequenos grupos, executar a classificação em cada um dos grupos (como apresentado na Figura 8) e, por fim, juntar os resultados. Este é o tipo de classificação radial ou, ainda, chamado de gaussiano. De acordo com Meyer (2021), o *kernel* radial aplica apenas o parâmetro γ da seguinte forma:

$$\text{Exp} \left\{ -\gamma |\mathbf{x}_i - \mathbf{x}_j|^2 \right\}$$

Onde:

- γ é o número de subgrupos.

Figura 8 – Máquina de suporte vetorial radial/gaussiano



Fonte: Wang et al. (2018).

Por último, de acordo com Camps-Valls et al. (2004), o *kernel* sigmoide é uma série de funções vizinhas. A classificação pode ser feita de forma linear, onde é possível, em seguida, que uma função não linear tangencie a função linear para definir outra zona de classificação e vice-versa, até percorrer toda a base em análise. O *kernel* sigmoide aplica os parâmetros (γ , \mathbf{C}_0) seguinte da forma: $\text{Tanh}\{\gamma \mathbf{x}_i^T \mathbf{x}_j + \mathbf{C}_0\}$.

De acordo com Patle e Chouhan (2013) a escolha do *kernel* correto é crucial para a obtenção de bons resultados. Por esse motivo, esta pesquisa explora todos os quatro modelos de *kernel* com a finalidade de indicar o *kernel* ou o modelo de SVM mais adequado à detecção de fraudes em relatórios financeiros do mercado brasileiro.

3.10 Métricas de avaliação da eficiência dos modelos de detecção de fraudes

Para verificar a eficiência de cada modelo, a taxa de acurácia foi levada em consideração, seguindo os procedimentos de Ravisankar et al. (2011), Song et al. (2014), Mohammadi et al. (2020) e Severino e Peng (2021).

Com o objetivo de apurar a eficiência estatística da acurácia, o teste de eficiência de McNemar foi considerado e utilizado, seguindo a recomendação de De Leeuw et al. (2006). O teste de eficiência de McNemar baseia-se na hipótese nula de igualdade nos erros de classificação (erros do tipo 1 e erros do tipo 2). O número do erro do tipo 1 cometido é igual ao número do erro do tipo 2 cometido. Em outras palavras, o número de relatórios fraudulentos classificados como não fraudulento é igual ao número de relatórios não fraudulentos classificados como fraudulento.

Segundo De Leeuw et al. (2006), a hipótese nula do teste de eficiência de McNemar pode ser formulada da seguinte forma:

$$H_0: \Psi = \frac{f_{12}}{f_{21}} = 1$$

Onde:

- f_{12} é o número de relatórios fraudulentos classificados como não fraudulentos; e
- f_{21} é o número de relatórios não fraudulentos classificados como fraudulentos.

Além do teste de eficiência de McNemar, o teste de Vuong (1988) foi empregado para trazer mais robustez à análise. De acordo com Vuong (1988), o teste consiste em calcular a média da razão de verossemelhança de cada modelo e em seguida, comparar os valores. As médias da razão de verossemelhança dos modelos são chamadas de estatísticas - z de Vuong. O teste é feito sob uma hipótese nula e duas hipóteses alternativas.

Considerando a comparação entre dois modelos M1 e M2, a hipótese nula é expressa por:

$H_0: z = 0$ o que significa que dois modelos que estão sendo comparados (M1 e M2) são iguais ou um explica o fenômeno estudado quanto outro ($M1 = M2$).

A primeira hipótese alternativa expressa por:

$H_1: z > 0$ significa que o primeiro modelo é melhor que o segundo modelo. Em outras palavras, M1 explica melhor o modelo estudado do que M2 ($M1 > M2$).

Por fim, a segunda hipótese alternativa

$H_1: z < 0$ significa que o primeiro modelo não é melhor que o segundo modelo. Em outras palavras, M2 explica melhor o modelo estudado do que M1 ($M1 < M2$).

4 Tratamento dos dados

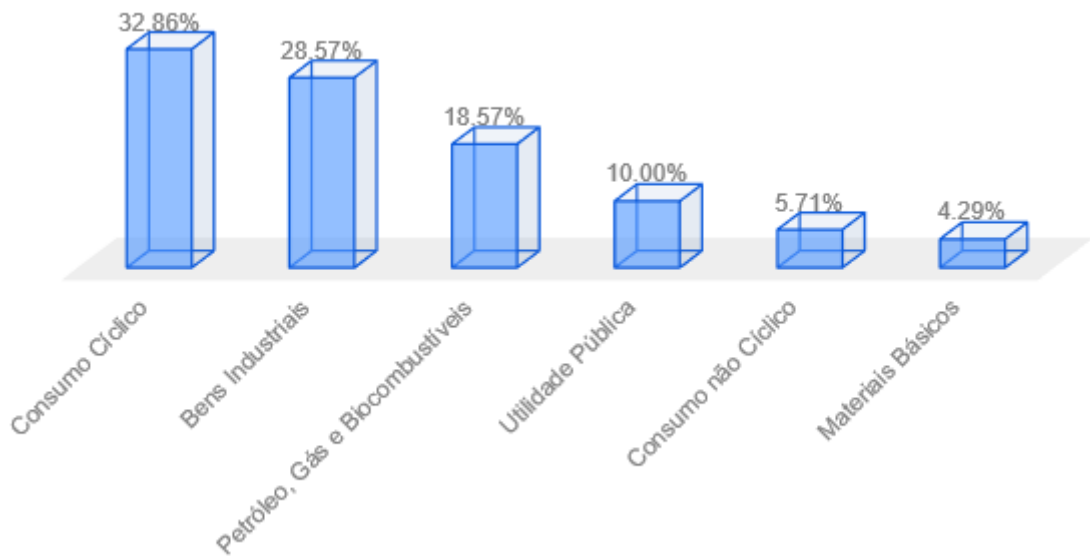
Nesta seção, além das estatísticas descritivas dos dados, as análises realizadas são explicadas. Na primeira subseção, as análises sobre o estudo de eventos são realizadas com

objetivo de testar se empresas, ao se utilizarem de fraudes nas demonstrações financeiras, obtêm retornos anormais. As análises sobre a detecção de fraudes são feitas na segunda subseção.

4.1 Análises do estudo de eventos

A execução do estudo de eventos para apurar as consequências de fraudes no mercado (existência ou não de retornos anormais) requer como variável de entrada os preços de cotação das ações ordinárias das entidades que emitiram relatórios fraudulentos. A estatística descritiva, por setor de atividade, das entidades da amostra é apresentada na Figura 9:

Figura 9 – Distribuição dos relatórios fraudulentos por setor de atividade



Fonte: elaboração própria.

Por ordem decrescente, seguindo a classificação de setores de atividade da B3, o setor de consumo cíclico é o setor no qual mais fraudes foram identificadas em suas demonstrações financeiras, compondo 33% da amostra. Em seguida está o setor de bens industriais, com 28%; depois, o setor de petróleo, gás e biocombustíveis, com 18%; logo, o setor de utilidade pública, com 10%. Os setores que realizam menos fraudes nas demonstrações financeiras são os setores de consumo não cíclico e de materiais básicos, com 6% e 4%, respectivamente.

A frequência de fraudes identificadas no setor de consumo cíclico pode estar relacionada ao fato de que há mais entidades registradas neste setor na CVM. Na coleta dos dados, realizada em 1º de maio de 2022, acerca da classificação setorial das empresas negociadas na B3, o setor de consumo cíclico representou a maior proporção, com 21%.

A análise do estudo de eventos foi realizada em quatro partes, sendo as duas primeiras partes referentes à realização da fraude (com a amostra de até 2009 e a amostra que se inicia a partir de 2010) e as duas últimas referentes à detecção da fraude (com a amostra antes da convergência contábil e a amostra depois da convergência contábil).

4.1.1 Análise do modelo do estudo de eventos à realização de fraudes

O estudo de eventos relativo à realização da fraude foi feito 10 dias ao redor da data de realização da fraude. Cada execução do estudo de eventos representa o número de dias ao redor da data de fraude em análise. Durante o período de análise – de 1999 até 2009 e de 2010 até 2019 –, foram identificados, respectivamente, 10 e 63 eventos de realização de fraudes. Na janela de estimação, os estimadores para obtenção do retorno anormal acumulado (α e β) foram obtidos considerando os 150 dias de observação.

Dos 10 eventos iniciais do grupo amostral antes da convergência contábil (de 1999 até 2009), 5 não foram processados por insuficiência de cotações, sendo eles: 10013 (IMBI3.SA); 10017 (EMAE3.SA); 10030 (LIXC3.SA); 10053 (RJCP3.SA); e 10061 (SULT3.SA). Dos 53 eventos do grupo amostral depois da convergência contábil (de 2010 até 2019), 18 eventos não foram processados por insuficiência de cotações, sendo eles: 10003 (BUET3.SA); 10014 (IMBI3.SA); 10015 (DUQE3.SA); 10016 (DUQE3.SA); 10020 (HETA3.SA); 10021 (I1HG34.SA); 10022 (IGUA3.SA); 10029 (LIXC3.SA); 10032 (MPLU3.SA); 10037 (OSXB3.SA); 10046 (PTNT3.SA); 10054 (RJCP3.SA); 10056 (SCLO3.SA); 10057 (SCLO3.SA); 10058 (SCLO3.SA); 10059 (SPRI3.SA); 10060 (SULT3.SA); e 10063 (TENE3.SA). Nessas situações, os ativos deixaram de ser negociados, no mínimo, por um ano, ou pelo fato de haver ausência de cotação na janela de estimação, ou pelo fato de a janela de estimação conter cotações constantes.

Quando as cotações são constantes, os retornos são iguais a zero, os estimadores são indetermináveis, assim como os retornos estimados e anormais. Em razão dessas características, a análise foi realizada com 40 eventos de demonstrações financeiras tipificadas como fraudulentas.

Os retornos reais das ações (r), assim como o retorno do mercado (R_M), obtidos por meio da Equação 3, bem como os estimadores α e β obtidos por meio da Equação 7, os resíduos decorrentes dessa regressão e os retornos esperados das ações obtidos por meio da Equação 7 estão apresentados na Tabela 3.

Tabela 3 – Retornos reais e esperados das ações um dia depois da data de realização de fraudes

Identificação das ações	Ação das firmas	Ação do mercado	Data de realização da fraude	Retorno real das ações	Retorno real do mercado	Intercepto (α)	Coefficiente (β)	Resíduo (μ)	Retorno esperado das ações
Antes da convergência contábil (até 2009)									
10010	CEBR6.SA	^BVSP	2009-12-31	0.0000	0.0043	0.0000	-0.0023	0.0025	0.0000
10018	EMBR3.SA	^BVSP	2009-12-31	0.0000	0.0043	-0.0014	0.7471	0.0213	0.0018
10044	PETR3.SA	^BVSP	2005-12-30	0.0134	0.0096	0.0010	0.7346	0.0158	0.0080
10045	PETR3.SA	^BVSP	2008-12-31	0.0147	0.0131	-0.0006	1.1091	0.0232	0.0139
10047	PTNT3.SA	^BVSP	2009-12-31	0.0000	0.0043	-0.0072	0.7031	0.0605	-0.0042
Depois da convergência contábil (a partir de 2010)									
10001	BAHI3.SA	^BVSP	2011-12-30	0.0000	0.0039	-0.0007	0.1491	0.0225	-0.0001
10002	BMYB34.SA	^BVSP	2013-12-31	0.0000	0.0047	0.0015	-0.0565	0.0160	0.0012
10004	CBMA3.SA	^BVSP	2011-12-30	0.0000	0.0039	-0.0048	0.3972	0.0553	-0.0032
10005	CBMA3.SA	^BVSP	2014-12-31	0.0000	-0.0117	-0.0015	-0.0983	0.0640	-0.0004
10006	CBMA3.SA	^BVSP	2015-12-31	0.0000	-0.0070	-0.0049	-0.1821	0.1160	-0.0037
10007	CEBR6.SA	^BVSP	2013-12-31	0.0000	0.0047	-0.0015	0.0139	0.0210	-0.0015
10008	CEBR6.SA	^BVSP	2012-12-31	0.0000	0.0088	0.0008	0.8020	0.0291	0.0079
10009	CEBR6.SA	^BVSP	2011-12-30	0.0000	0.0039	-0.0018	0.1232	0.0394	-0.0013
10011	CGR3.SA	^BVSP	2014-12-31	0.0006	-0.0117	0.0000	0.1920	0.0170	-0.0023
10012	CMIG3.SA	^BVSP	2016-12-30	0.0166	0.0074	0.0007	1.1700	0.0213	0.0094
10019	GGBR3.SA	^BVSP	2015-12-31	0.0381	-0.0070	-0.0031	1.1087	0.0294	-0.0109
10023	IMBI3.SA	^BVSP	2015-12-31	0.0000	-0.0070	-0.0002	0.0619	0.0278	-0.0006
10024	INEP3.SA	^BVSP	2018-12-31	0.0163	0.0280	0.0006	0.5538	0.0584	0.0161
10025	INEP3.SA	^BVSP	2013-12-31	0.0817	0.0047	-0.0029	0.2775	0.0362	-0.0016
10026	JBSS3.SA	^BVSP	2012-12-31	-0.0066	0.0088	-0.0006	1.1446	0.0252	0.0095
10027	JBSS3.SA	^BVSP	2017-12-29	0.0000	0.0000	0.0020	0.9041	0.0333	0.0020
10028	JHSF3.SA	^BVSP	2016-12-30	-0.0141	0.0074	0.0001	0.9160	0.0351	0.0069
10031	MNDL3.SA	^BVSP	2013-12-31	0.0465	0.0047	-0.0030	0.4290	0.0287	-0.0010
10033	MSPA3.SA	^BVSP	2014-12-31	0.0000	-0.0117	0.0002	0.0047	0.0030	0.0002
10034	MSPA3.SA	^BVSP	2017-12-29	0.0000	0.0000	0.0023	0.0699	0.0131	0.0023
10035	OGXP3.SA	^BVSP	2012-12-31	0.0138	0.0088	-0.0076	1.8655	0.0397	0.0088
10036	OGXP3.SA	^BVSP	2013-12-31	0.0426	0.0047	-0.0121	4.1111	0.1048	0.0071
10038	PDGR3.SA	^BVSP	2012-12-31	0.0030	0.0088	-0.0011	1.5101	0.0295	0.0122
10039	PETR3.SA	^BVSP	2014-12-31	-0.0288	-0.0117	-0.0028	1.9352	0.0170	-0.0254
10040	PETR3.SA	^BVSP	2013-12-31	0.0050	0.0047	-0.0004	1.1637	0.0191	0.0050
10041	PETR3.SA	^BVSP	2012-12-31	0.0026	0.0088	-0.0005	1.0722	0.0159	0.0089
10042	PETR3.SA	^BVSP	2011-12-30	0.0092	0.0039	-0.0001	0.9605	0.0131	0.0036
10043	PETR3.SA	^BVSP	2010-12-31	0.0148	0.0051	-0.0010	0.9636	0.0157	0.0039
10048	RCSL3.SA	^BVSP	2015-12-31	0.1803	-0.0070	-0.0001	1.2323	0.1623	-0.0088
10049	RCSL3.SA	^BVSP	2014-12-31	0.0000	-0.0117	-0.0038	-0.1417	0.0917	-0.0021
10050	RCSL3.SA	^BVSP	2013-12-31	0.1542	0.0047	-0.0029	0.7357	0.0857	0.0005
10051	RCSL3.SA	^BVSP	2012-12-31	0.1178	0.0088	-0.0081	-0.4851	0.0729	-0.0124
10052	RCSL3.SA	^BVSP	2011-12-30	0.0000	0.0039	-0.0025	0.4195	0.0652	-0.0009
10055	RSID3.SA	^BVSP	2010-12-31	0.0184	0.0051	-0.0004	1.5258	0.0215	0.0074
10062	TASA3.SA	^BVSP	2013-12-31	0.0000	0.0047	-0.0009	0.0493	0.0247	-0.0006

Fonte: dados da pesquisa.

No primeiro momento, para cada ação, a cada evento, um dia após o evento, o retorno real (observado) das ações (r) foi estimado de acordo com a Equação 3.

No segundo momento, o objetivo foi obter os retornos estimados com base nas observações (retornos) anteriores. Para isso, foram considerados os 150 retornos das ações das firmas e do mercado antes da data do evento. Para atingir o objetivo deste momento (apurar o retorno esperado das ações), foi calculado, primeiramente, o retorno real (observado) do

mercado (BVSP) de acordo com a Equação 3. Os estimadores α e β e o resíduo μ foram obtidos considerando os cálculos estatísticos, onde $\hat{\beta}_i = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)}$, $\hat{\alpha} = \hat{\mu}_{R_i} - \hat{\beta}_i \hat{\mu}_{R_m}$, com $\hat{\mu}_{R_i} = \frac{1}{T_n - T_0} (\sum_{s=T_0}^{T_n=150} R_{is})$ e $\hat{\mu}_{R_m} = \frac{1}{T_n - T_0} (\sum_{s=T_0}^{T_n=150} R_{ms})$.

Em seguida, os retornos anormais foram apurados de acordo com a Equação 8 e são apresentados na Tabela 4, que representa os quatro primeiros dias depois da data de realização de fraudes. Os retornos anormais completos depois da realização de fraudes são apresentados no Apêndice 6, e os retornos anormais completos depois da detecção de fraudes são apresentados no Apêndice 7.

Tabela 4 – Retornos anormais dos quatro primeiros dias depois da realização de fraudes

Ação das firmas	AR (1)	Valor-t (1)	AR (2)	Valor-t (2)	AR (3)	Valor-t (3)	AR (4)	Valor-t (4)
Antes da convergência contábil (até 2009)								
CEBR6.SA	-0,0032	-0,1468	-0,0042	-0,1927	-0,0017	-0,078	-0,002	-0,0917
EMBR3.SA	-0,0235	-10981	-0,0075	-0,3505	-0,008	-0,3738	0,0024	0,1121
PETR3.SA	0,0217	13478	-0,0025	-0,1553	0,0027	0,1677	-0,0041	-0,2547
PETR3.SA	0,0058	0,2522	-0,0126	-0,5478	-0,0046	-0,2	-0,0134	-0,5826
PTNT3.SA	0,0054	0,0891	0,0845	13944	0,0099	0,1634	0,009	0,1485
Depois da convergência contábil (a partir de 2010)								
BAHI3.SA	-0,0049	-0,2121	0,0007	0,0303	0,0028	0,1212	0,0008	0,0346
BMYB34.SA	-0,0028	-0,1687	-0,0009	-0,0542	-0,0016	-0,0964	-0,0022	-0,1325
CBMA3.SA	0,1018	1.7891*	0,0019	0,0334	0,0086	0,1511	0,1076	1.891*
CBMA3.SA	-0,0032	-0,0514	-0,0015	-0,0241	0,0037	0,0595	0,0071	0,1141
CBMA3.SA	-0,0079	-0,0611	0,0024	0,0186	-0,0041	-0,0317	-0,0073	-0,0565
CEBR6.SA	0,002	0,0985	0,0019	0,0936	0,002	0,0985	0,002	0,0985
CEBR6.SA	-0,0237	-0,75	-0,0108	-0,3418	0,0116	0,3671	0,0094	0,2975
CEBR6.SA	-0,0036	-0,0914	0,0015	0,0381	0,0034	0,0863	0,0016	0,0406
CGRA3.SA	0,0041	0,2547	-0,0442	-2.745***	-0,0317	-1.9689*	0,0044	0,2733
CMIG3.SA	0,0398	1.7848*	-0,0316	-1,417	-0,0187	-0,8386	-0,0011	-0,0493
GGBR3.SA	-0,0093	-0,3153	-0,0467	-15,831	-0,0749	-2.539**	0,0009	0,0305
IMBI3.SA	0,002	0,0752	-0,1301	-4.891***	0,0015	0,0564	0,0019	0,0714
INEP3.SA	-0,0199	-0,3144	0,0172	0,2717	0,0296	0,4676	-0,0493	-0,7788
INEP3.SA	-0,0837	-2.3058**	0,051	1,405	-0,0501	-13,802	0,0053	0,146
JBSS3.SA	-0,0084	-0,3158	-0,0042	-0,1579	0,0477	1.7932*	-0,0312	-11,729
JBSS3.SA	-0,0213	-0,4712	-0,0058	-0,1283	0,0039	0,0863	-0,0004	-0,0088
JHSF3.SA	0,0315	0,8678	0,0073	0,2011	0,0432	11,901	0,0554	15,262
MNDL3.SA	-0,0319	-0,8961	0,0198	0,5562	-0,0373	-10,478	0,0233	0,6545
MSPA3.SA	0,0049	0,2311	0,0038	0,1792	0,0003	0,0142	-0,002	-0,0943
MSPA3.SA	-0,0015	-0,0617	-0,003	-0,1235	-0,0024	-0,0988	-0,0021	-0,0864
OGXP3.SA	0,0419	10,554	0,0133	0,335	0,0442	11,134	0,0403	10,151
OGXP3.SA	0,0623	0,5973	0,0009	0,0086	0,1652	15,839	-0,0587	-0,5628
PDGR3.SA	0,014	0,4531	0,0169	0,5469	-0,0168	-0,5437	0,0349	11,294
PETR3.SA	-0,0033	-0,2129	-0,0427	-2.755***	-0,0424	-2.735***	-0,0071	-0,4581
PETR3.SA	0,0165	0,8684	-0,0345	-1.8158*	0,0129	0,6789	-0,0187	-0,9842
PETR3.SA	-0,012	-0,7692	0,0242	15,513	0,0166	10,641	-0,0031	-0,1987
PETR3.SA	0,0097	0,7293	0,0047	0,3534	0,0072	0,5414	-0,0016	-0,1203
PETR3.SA	-0,0154	-0,9872	-0,0112	-0,7179	0,0102	0,6538	0,0021	0,1346
RCSL3.SA	-0,1384	-0,8793	-0,0012	-0,0076	0,0923	0,5864	0,1268	0,8056
RCSL3.SA	-0,0039	-0,0437	-0,0014	-0,0157	0,0064	0,0717	0,2346	2.63***
RCSL3.SA	0,0168	0,2002	-0,0079	-0,0942	0,001	0,0119	0,0083	0,0989
RCSL3.SA	-0,0999	-14,416	0,1309	1.8889*	0,11	15,873	-0,0997	-14,387
RCSL3.SA	-0,0151	-0,2323	0,0023	0,0354	0,0088	0,1354	0,0026	0,04
RSID3.SA	-0,0104	-0,4502	-0,0333	-14,416	-0,0202	-0,8745	-0,0032	-0,1385

TASA3.SA	-0,0606	-2.4048**	0,015	0,5952	0	0	-0,0073	-0,2897
----------	---------	-----------	-------	--------	---	---	---------	---------

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Na maioria das vezes, os retornos anormais individuais (ou seja, para cada evento) foram detectados depois da realização da fraude; porém, em poucas situações esses retornos foram estatisticamente significativos, de acordo com a estatística t de Student. Em poucas situações os retornos anormais se igualam a 0, o que significaria a ausência real de retornos anormais.

No Apêndice 6, observa-se que, antes da convergência contábil (até 2009), 48% dos retornos anormais são positivos e quase todos são estatisticamente insignificativos. Depois da convergência contábil, observa-se que 53% dos retornos anormais são positivos e apenas 9% são significativos. Analisando individualmente os retornos anormais, conclui-se que, depois da realização de fraudes em relatórios financeiros, a maioria dos retornos anormais é não significativa.

A maioria dos retornos anormais individuais é estatisticamente não significativa, o que sugere a não existência de retornos anormais devido ao uso de informações viesadas para a realização de práticas fraudulentas nas demonstrações financeiras.

4.1.2 Análise do modelo do estudo de eventos à detecção de fraudes

O estudo de eventos para detecção de fraude em relatórios financeiros é realizado 10 dias ao redor da data do evento. Cada execução do estudo de evento representa o número de dias ao redor da data de fraude em análise.

Pelo fato de vários relatórios fraudulentos serem detectados em um único processo administrativo da CVM, estes teriam a mesma data de evento. Contudo, para uma execução adequada do modelo de estudo de eventos, cada evento (dia de detecção da fraude) deve ser único para cada ativo, e, por esta razão, os 63 relatórios fraudulentos iniciais foram reduzidos para 51, sendo compostos por 6 relatórios do período antes da convergência e 45 do período depois da convergência. Observou-se que todas as fraudes foram detectadas no período depois da convergência. Entretanto, pelo fato de a composição de alguns elementos dos relatórios ter mudado depois da convergência contábil (a partir de 2010), os relatórios emitidos antes da convergência e depois dela foram analisados separadamente.

O Quadro 4 evidencia a relação entre os relatórios fraudulentos na realização da fraude e os relatórios fraudulentos na detecção. Pelo fato de mais de um relatório do mesmo ativo ser detectado em um mesmo processo administrativo, 14 relatórios foram retirados da amostra de

relatórios fraudulentos na detecção da fraude por estarem com a mesma data de detecção e pertencerem a um mesmo ativo.

Por exemplo, os relatórios fraudulentos de 2011, 2012 e 2013 do ativo CEBR6 foram detectados no processo administrativo de 02/03/2021 da CVM. Esses três relatórios, da mesma entidade, têm a mesma data de detecção (que é a data de 02/03/2021). De acordo com os procedimentos de execução do modelo de estudo de eventos, esses três eventos têm as mesmas características e, portanto, devem se resumir em apenas um evento. Nesse sentido, apenas o relatório mais recente foi considerado com a data de detecção na amostra de relatórios fraudulentos na detecção de fraudes, dado que é este o relatório mais atualizado da entidade, em relação à fraude.

Dois relatórios fraudulentos foram adicionados a esta relação por serem detectados por diferentes fraudes em relatórios financeiros em dois processos administrativos distintos. Ressalta-se que, para um ativo cujo relatório de um ano dado foi detectado em mais de um processo, este relatório foi duplicado, mas com datas de detecções diferentes. É o caso do relatório fraudulento de 2012 do ativo OGXP3, que foi citado nos processos administrativos de 27/05/2019 e 14/12/2017. Este relatório fraudulento de 2012 foi considerado duas vezes, com suas respectivas datas de detecção.

fraudulentos à realização e à detecção da fraude

Realização de fraudes		Nº de relatórios retirados	Nº de relatórios adicionados	Detecção de fraudes	
Ativo	Data de fraude			Ativo	Data de fraude
BAHI3	12/31/2011		BAHI3	9/24/2019	
BMYB34	12/12/2013		BMYB34	12/10/2019	
BUET3	8/30/2016		BUET3	12/10/2019	
CBMA3	12/31/2011		CBMA3	10/30/2018	
CBMA3	12/31/2014		CBMA3	12/12/2017	
CBMA3	12/31/2015	-1			
CEBR6	12/31/2013		CEBR6	3/2/2021	
CEBR6	12/31/2012	-1			
CEBR6	12/31/2011	-1			
CEBR6	4/20/2009		CEBR6	3/26/2019	
CGRA3	12/31/2014		CGRA3	10/23/2018	
CMIG3	4/29/2016		CMIG3	12/18/2018	
DOCA3	31/09/2002		DOCA3	2/19/2019	
DOCA3	12/31/2014		DOCA3	12/18/2018	
DUQE3	10/31/2010		DUQE3	2/26/2019	
DUQE3	12/31/2013		DUQE3	7/12/2018	
EMAE3	4/8/1998		EMAE3	1/29/2019	
EMBR3	1/1/2009		EMBR3	9/11/2018	
GGBR3	7/14/2015		GGBR3	12/3/2019	
HETA3	12/31/2013		HETA3	12/14/2017	
IHGG34	12/12/2013		IHGG34	12/26/2018	
IGBR3	12/31/2011		IGUA3	6/30/2020	
IMBI3	12/31/2015		IMBI3	4/2/2019	
INEP3	5/7/2018		INEP3	12/3/2019	

INEP3	12/31/2013		INEP3	12/18/2018
JBSS3	12/31/2012		JBSS3	7/21/2020
JBSS3	5/11/2017	-1		
JHSF3	11/10/2016		JHSF3	12/4/2018
LIXC3	12/31/2016		LIXC3	8/27/2019
LIXC3	12/31/2001		LIXC3	10/30/2018
MNDL3	12/31/2013		MNDL3	12/14/2017
MPLU3	12/31/2017		MPLU3	12/10/2019
MSPA3	12/31/2014		MSPA3	12/11/2018
MSPA3	12/31/2017	-1		
OGXP3	7/31/2013		OGXP3	6/25/2019
OGXP3	2/1/2012		OGXP3	5/27/2019
			OGXP3	12/14/2017
OSXB3	9/24/2012		OSXB3	6/25/2019
PDGR3	5/28/2012		PDGR3	11/13/2018
PETR3	12/31/2014		PETR3	11/3/2020
PETR3	12/31/2012	-1		
PETR3	12/31/2011	-1		
PETR3	12/31/2010	-1		
PETR3	12/31/2005	-1		
PETR3	12/31/2008	-1		
PETR3	9/5/2013		PETR3	12/16/2019
PTNT3	6/30/2012		PETR3	12/13/2018
PTNT3	10/29/2009		PTNT3	8/13/2019
RCSL3	12/31/2015		PTNT3	12/4/2018
RCSL3	12/31/2014		RCSL3	9/29/2020
RCSL3	12/31/2013		RCSL3	7/30/2019
RCSL3	12/31/2012	-1		
RCSL3	12/31/2011	-1		
RJCP3	6/30/2009		RJCP3	6/11/2019
RJCP3	9/10/2014		RJCP3	12/4/2018
			RJCP3	10/30/2018
RSID3	6/30/2010		RSID3	8/20/2019
SCLO3	12/31/2012	-1		
SCLO3	12/31/2013	-1		
SCLO3	12/31/2014		SCLO3	12/12/2017
SPRI3	3/31/2014		SPRI3	7/7/2020
SULT3	8/18/2015		SULT3	12/10/2019
SULT3	12/31/2001		SULT3	2/26/2019
TASA3	7/16/2013		TASA3	7/13/2018
TENE3	3/31/2016		TENE3	12/26/2018
	63	-14	+2	51

Fonte: dados da pesquisa.

Na janela de estimação, os estimadores α e β foram obtidos considerando os 150 dias de observação.

Dos 6 relatórios do período antes da convergência contábil, 5 não foram processados por insuficiência de cotações, sendo eles: 10010 (DOCA3.SA); 10014 (EMAE3.SA); 10026 (LIXC3.SA); 10042 (RJCP3.SA); e 10049 (SULT3.SA). Dos 45 relatórios do período depois da convergência contábil, 21 não foram processados por insuficiência de cotações, sendo eles: 10003 (BUET3.SA); 10004 (CBMA3.SA); 10005 (CBMA3.SA); 10011 (DOCA3.SA); 10012 (DUQE3.SA); 10013 (DUQE3.SA); 10017 (HETA3.SA); 10018 (I1HG34.SA); 10019 (IGUA3.SA); 10020 (IMBI3.SA); 10025 (LIXC3.SA); 10028 (MPLU3.SA); 10030 (OGXP3.SA); 10031 (OGXP3.SA); 10033 (OSXB3.SA); 10043 (RJCP3.SA); 10044

(RJCP3.SA); 10046 (SCLO3.SA); 10047 (SPRI3.SA); 10048 (SULT3.SA); e 10051 (TENE3.SA). Nessas situações, os ativos pararam de ser negociados, no mínimo, por um ano, ou pelo fato de haver ausência de cotação na janela de estimação, ou pelo fato da janela de estimação conter cotações constantes. Nesses casos, os possíveis retornos anormais não seriam referentes ao período esperado; por isso, foram excluídos da análise.

Quando as cotações são constantes, os retornos são iguais a 0 e os estimadores são indetermináveis, assim como os retornos estimados e anormais.

Os retornos reais das ações (r), assim como o retorno do mercado (R_M), obtidos por meio da Equação 3, bem como os estimadores α e β obtidos por meio da Equação 7, os resíduos decorrentes dessa regressão e os retornos esperados das ações obtidos por meio da Equação 7 estão apresentados na Tabela 5.

Tabela 5 – Estimadores e retornos esperados das ações um dia depois da data de detecção de fraudes

Identificação das ações	Ação das firmas	Ações do mercado	Data de detecção da fraude	Retorno real das ações	Retorno real do mercado	Intercepto (α)	Coefficiente (β)	Resíduo	Retorno esperado das ações
Antes da convergência contábil (até 2009)									
10015	EMBR3.SA	^BVSP	2018-09-11	0.006	-0.0235	0.0002	0.529	0.024	-0.0123
Depois da convergência contábil (a partir de 2010)									
10001	BAHI3.SA	^BVSP	2019-09-24	0.0068	-0.0073	-0.003	0.219	0.026	-0.0048
10002	BMYB34.SA	^BVSP	2019-12-10	0	-0.0028	0.0017	0.054	0.015	0.0016
10006	CEBR6.SA	^BVSP	2021-03-02	0	0.0109	0.0056	-0.03	0.038	0.0053
10007	CEBR6.SA	^BVSP	2019-03-26	0.1004	0.0174	-0.001	-0.11	0.018	-0.0026
10008	CGRA3.SA	^BVSP	2018-10-23	0.0215	-0.0035	-0.001	0.347	0.022	-0.002
10009	CMIG3.SA	^BVSP	2018-12-18	0.0138	0.0024	0.0031	1.143	0.022	0.0058
10016	GGBR3.SA	^BVSP	2019-12-03	-0.020	-0.001	-0.001	0.842	0.015	-0.0013
10021	INEP3.SA	^BVSP	2019-12-03	0	-0.001	-0.004	0.496	0.044	-0.0047
10022	INEP3.SA	^BVSP	2018-12-18	0	0.0024	0.0045	0.276	0.063	0.0051
10023	JBSS3.SA	^BVSP	2020-07-21	-0.017	-0.0011	-0.001	0.936	0.033	-0.0018
10024	JHSF3.SA	^BVSP	2018-12-04	-0.027	-0.0134	0.0005	0.995	0.029	-0.0129
10027	MNDL3.SA	^BVSP	2017-12-14	0	-0.0067	0.0037	0.496	0.035	0.0004
10029	MSPA3.SA	^BVSP	2018-12-11	0	0.0059	0.0006	0.004	0.019	0.0006
10032	OGXP3.SA	^BVSP	2017-12-14	-0.012	-0.0067	-0.001	0.401	0.045	-0.003
10034	PDGR3.SA	^BVSP	2018-11-13	-0.079	-0.0072	-0.002	0.528	0.056	-0.0059
10035	PETR3.SA	^BVSP	2020-11-03	0.0387	0.0173	0.0003	1.302	0.022	0.0228
10036	PETR3.SA	^BVSP	2019-12-16	-0.008	-0.006	-0.001	1.118	0.012	-0.0073
10037	PETR3.SA	^BVSP	2018-12-13	0.0053	0.0099	0.0004	1.542	0.027	0.0156
10038	PTNT3.SA	^BVSP	2019-08-13	0	0.0135	0.0001	0.003	0.009	0.0001
10039	PTNT3.SA	^BVSP	2018-12-04	0	-0.0134	0.0001	-0.01	0.002	0.0002
10040	RCSL3.SA	^BVSP	2020-09-29	0.0101	-0.0115	-0.002	0.838	0.040	-0.0113
10041	RCSL3.SA	^BVSP	2019-07-30	0.0144	-0.0053	-0.006	0.372	0.038	-0.0078
10045	RSID3.SA	^BVSP	2019-08-20	-0.088	-0.0025	-0.001	0.644	0.023	-0.0023
10050	TASA3.SA	^BVSP	2018-07-13	0	0.0097	-0.000	0.221	0.035	0.002

Fonte: dados da pesquisa.

Em seguida, os retornos anormais foram apurados de acordo com a Equação 8 e são apresentados na Tabela 6, que representa os quatro primeiros dias depois da data da detecção

de fraudes. Os retornos anormais completos depois da detecção de fraudes são apresentados no Apêndice 7.

Tabela 6 – Retornos anormais dos quatro primeiros dias depois da detecção de fraudes

Ação das firmas	AR (1)	Valor-t (1)	AR (2)	Valor-t (2)	AR (3)	Valor-t (3)	AR (4)	Valor-t (4)
Antes da convergência contábil (até 2009)								
EMBR3.SA	-0.0369	-1.5311	-0.0079	-0.3278	-0.007	-0.2905	-0.0054	-0.2241
Depois da convergência contábil (a partir de 2010)								
BAHI3.SA	0.0087	0.3258	0.0348	1.3034	0.0483	1.809*	-0.0087	-0.3258
BMYB34.SA	0.0452	3.0133***	0.0172	1.1467	0.0214	1.4267	-0.0204	-1.36
CEBR6.SA	-0.0057	-0.1473	-0.0052	-0.1344	-0.0025	-0.0646	-0.0067	-0.1731
CEBR6.SA	-0.0032	-0.1778	0.0036	0.2	0.0019	0.1056	0.0014	0.0778
CGRA3.SA	0.0269	1.2227	-0.0203	-0.9227	-0.0063	-0.2864	-0.0051	-0.2318
CMIG3.SA	0.0297	1.3879	0.0758	3.5421***	-0.0032	-0.1495	0.0075	0.3505
GGBR3.SA	0.0042	0.2763	0.0043	0.2829	0.0111	0.7303	-0.0114	-0.75
INEP3.SA	0.0078	0.1769	0.0055	0.1247	0.0088	0.1995	-0.0062	-0.1406
INEP3.SA	-0.0648	-1.0335	-0.0276	-0.4402	-0.0418	-0.6667	-0.0045	-0.0718
JBSS3.SA	-0.0085	-0.2607	0.0323	0.9908	-0.0131	-0.4018	-0.0223	-0.684
JHSF3.SA	0.032	1.0997	0.0017	0.0584	0.0684	2.3505**	-0.0001	-0.0034
MNDL3.SA	-0.0097	-0.2779	0.0281	0.8052	-0.0303	-0.8682	0.0367	1.0516
MSPA3.SA	-0.0006	-0.0317	-0.0006	-0.0317	-0.0005	-0.0265	-0.0005	-0.0265
OGXP3.SA	-0.0261	-0.5865	0.0104	0.2337	-0.0069	-0.1551	-0.0034	-0.0764
PDGR3.SA	-0.0136	-0.2416	0.1781	3.1634***	-0.0896	-1.5915	0.0936	1.6625*
PETR3.SA	-0.0257	-1.1735	-0.0341	-1.5571	-0.0041	-0.1872	0.0574	2.621***
PETR3.SA	0.0072	0.6154	0.0098	0.8376	-0.0089	-0.7607	-0.0108	-0.9231
PETR3.SA	-0.0012	-0.0446	0.0075	0.2788	-0.0386	-1.4349	-0.0001	-0.0037
PTNT3.SA	-0.0000	-0.0011	-0.0001	-0.0116	-0.0001	-0.0116	-0.0001	-0.0116
PTNT3.SA	-0.0001	-0.0556	-0.0001	-0.0556	-0.0002	-0.1111	-0.0003	-0.1667
RCSL3.SA	0.0024	0.0596	-0.0061	-0.1514	-0.0105	-0.2605	-0.0372	-0.9231
RCSL3.SA	-0.0117	-0.3095	0.0476	1.2593	0.0039	0.1032	-0.0131	-0.3466
RSID3.SA	0.0474	2.017**	0.0274	1.166	-0.0298	-1.2681	0.0065	0.2766
TASA3.SA	0.0056	0.1564	-0.0098	-0.2737	0.0413	1.1536	-0.0336	-0.9385

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Ao redor da data de detecção, retornos anormais individuais (ou seja, para cada evento) foram detectados na maioria das vezes, mas eles não foram estatisticamente significativos para inferir a sua existência.

Acerca do Apêndice 7, observou-se que, antes da convergência contábil, apenas 30% dos retornos anormais eram positivos e nenhum apresentou significância. Depois da convergência, apenas 40% dos retornos anormais eram positivos e apenas 9% são significativos. Depois da detecção de fraudes, a maioria dos retornos anormais é negativa, porém não significativa.

4.2 Análise do modelo de detecção de fraudes (SVM)

Dos 63 relatórios fraudulentos identificados inicialmente na seção 3.1, 7 foram retirados da amostra. Na busca automatizada dos indicadores financeiros e não financeiros, os relatórios

dos ativos RJCP3, DUQE3, BMYB34, I1HG34, DOCA3, IMBI3 e BUET3 não retornaram nenhum valor; por este motivo, eles foram retirados da amostra para a análise de detecção de fraudes, resultando, então, em um efetivo de 56 relatórios fraudulentos e 56 relatórios não fraudulentos.

Em razão da convergência contábil iniciada em 2010 no Brasil, a amostra dos indicadores financeiros e não financeiros usados como variáveis explanatórias no modelo de SVM foi dividida em duas partes. A primeira contém as variáveis de 1998 até 2009 (amostra 1), com um efetivo de 9 relatórios fraudulentos e 9 relatórios não fraudulentos, totalizando 18 relatórios analisados. A segunda parte contém as variáveis a partir de 2010 (amostra 2), com um efetivo de 47 relatórios fraudulentos e 47 relatórios não fraudulentos, totalizando 94 relatórios analisados. No total, a amostra coletada contém 56 relatórios fraudulentos e 56 relatórios não fraudulentos, totalizando 112 relatórios analisados.

A análise da detecção de fraudes em relatórios fraudulentos foi feita em dois momentos. No primeiro momento, a análise foi feita com a amostra 1 e, no segundo momento, a análise foi feita com a amostra 2.

Ressalta-se que o modelo de detecção de fraude, por meio da SVM, testa a eficiência do processo de identificar um relatório fraudulento. Baseando-se na amostra estudada, a eficiência de identificação de um relatório fraudulento é a porcentagem de uma classificação bem-sucedida (ou seja, classificar um relatório fraudulento como fraudulento).

Com o objetivo de identificar riscos de fraudes nos relatórios financeiros, os modelos da SVM buscam padrões de reconhecimento nos relatórios financeiros por meio das variáveis apresentadas no Quadro 3. Essas variáveis foram escolhidas com base na literatura sobre o tema da pesquisa. Outras variáveis disponíveis no plataforma Economatica® foram exploradas a fim de identificar a possibilidade de elas serem capazes de revelar fraudes em relatórios financeiros (ou seja, valendo-se da análise por meio de mineração dos dados).

Antes de executar o modelo de detecção de fraude propriamente dito, modelos univariados foram executados. A cada execução do modelo univariado, a taxa de precisão na detecção de relatórios financeiros fraudulentos é apurada, analisada e considerada ou não nos modelos da SVM. As variáveis explanatórias dos modelos univariados com uma taxa de detecção superior ou igual a 50% e com significância no mínimo de 10% foram escolhidas como variáveis sensíveis que devem compor o modelo de previsão de fraudes.

A significância dos modelos de detecção de fraudes foi obtida de acordo com o teste de eficiência de McNemar apresentado na seção 3.10, em que o valor calculado é a própria porcentagem de assertividade.

Com o objetivo de selecionar as variáveis mais sensíveis à detecção de fraudes, modelos univariados foram executados. Em outras palavras, por meio dos modelos univariados, variáveis que podem ser úteis para identificar riscos de fraudes em relatórios fraudulentos foram selecionadas para compor o modelo final de detecção de fraudes. Para cada amostra e para cada modelo da SVM, os modelos univariados de detecção de fraudes foram executados. Os resultados estão compilados na Tabela 7.

A porcentagem de acurácia representa a capacidade da variável unilateralmente classificar com sucesso os relatórios fraudulentos. Em outras palavras, a porcentagem de detecção é a capacidade para cada variável identificar sozinha um relatório fraudulento. Ou, ainda, é a chance de uma variável solitariamente identificar se um relatório é fraudulento ou não. A significância da proporção de detecção foi apurada de acordo com o teste de eficiência de McNemar, em que o valor calculado é a própria proporção de detecção.

Tabela 7 – Resultados dos modelos univariados

Variável	Amostra 1				Amostra 2			
	% de detecção linear	% de detecção polinomial	% de detecção o radial	% de detecção sigmoide	% de detecção o linear	% de detecção polinomial	% de Detecção o radial	% de detecção Sigmoide
INF1	50%	75%	50%	25%	61%***	50%	61%*	54%
INF2	50%	50%	50%	25%	46%***	71%*	46%	54%
INF3	33%	50%	50%	50%	50%***	50%***	61%	50%
INF4	33%	50%	50%	50%	50%***	50%***	50%*	43%
INF5	33%	25%	75%	50%	46%***	50%***	39%*	50%***
INF6	17%	50%	50%	50%	50%***	54%***	36%	54%
DPA1	33%	50%	25%	50%	50%***	54%**	46%	54%
DPA2	50%	50%	25%	50%	43%***	50%***	46%***	57%
DPA3	50%	25%	25%	25%	50%***	54%***	57%	46%
EDC1	33%	50%	25%	25%	50%***	61%**	61%***	54%
EDC2	50%	50%	25%	25%	50%***	61%*	50%*	61%
EDC3	67%	50%	100%	25%	54%***	46%***	43%	50%***
EDC4	50%	50%	50%	25%	50%***	46%***	43%	61%
EDC5	33%	25%	50%	75%	50%***	50%***	50%	57%
EDC6	50%	50%	50%	75%	50%***	36%	46%	64%
EDC7	33%	50%	50%	50%	46%***	43%	43%	57%
EDC8	33%	50%	50%	50%	50%***	46%***	54%*	39%
EDC9	50%	50%	50%	25%	50%***	50%***	43%	57%
EDC10	50%	50%	50%	50%	46%***	46%***	39%**	36%
EDC11	67%	50%	50%	50%	46%***	50%***	57%	46%
EDC12	50%	50%	50%	25%	50%***	50%***	54%**	50%
LIQ1	50%	50%	25%	50%	46%***	50%***	54%	50%***
LIQ2	33%	50%	50%	25%	50%***	50%***	46%**	46%
LIQ3	67%	50%	50%	25%	50%***	54%***	46%**	54%
LIQ4	50%	75%	50%	25%	46%***	61%**	61%***	50%
LIQ5	33%	50%	25%	25%	46%***	61%**	61%***	61%***
REN1	50%	25%	25%	25%	46%***	50%	71%**	64%
REN2	50%	25%	25%	25%	46%***	61%**	61%***	61%***
REN3	50%	50%	50%	50%	50%***	50%***	50%	57%
REN4	33%	25%	50%	50%	50%***	50%***	50%*	57%
REN5	67%	25%	50%	50%	50%***	54%***	46%***	57%***
REN6	33%	25%	50%	50%	50%***	50%***	50%*	57%
REN7	50%	25%	25%	75%	54%***	54%***	39%	46%***

REN8	50%	25%	50%	50%	43%***	54%***	39%	46%
REN9	50%	25%	25%	50%	43%***	50%***	43%	57%
REN10	50%	25%	25%	50%	50%***	54%***	39%	57%
REN11	33%	25%	25%	50%	50%***	46%***	32%	50%***
REN12	50%	25%	25%	50%	46%***	54%***	50%	54%
REN13	50%	25%	25%	25%	50%***	50%***	50%	46%***
REN14	50%	50%	25%	25%	46%***	61%**	61%***	61%
REN15	50%	25%	50%	50%	50%***	50%***	46%***	50%***
REN16	33%	50%	50%	50%	50%***	57%***	57%	50%
REN17	50%	50%	50%	50%	54%***	46%	43%	50%
OUT1	33%	50%	25%	25%	46%***	61%**	68%	54%
OUT2	50%	50%	25%	25%	50%***	46%	64%	57%
OUT3	50%	50%	50%	0%	50%***	57%***	68%	57%
OUT4	17%	50%	25%	25%	43%***	50%***	75%	64%
OUT5	33%	50%	50%	25%	50%***	61%*	64%	54%

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Com o objetivo de identificar o modelo de detecção de fraudes com maior acurácia, as variáveis-referências de riscos de fraudes foram escolhidas. Variáveis com uma acurácia superior ou igual a 50% e com uma significância de no máximo 10% foram escolhidas para compor os modelos finais de detecção de fraudes.

Recorda-se que a amostra 1 é formada pela série temporal das variáveis de 1998 até 2009, com um efetivo total de 18 relatórios. Observa-se que, para o modelo linear, a acurácia mínima e a acurácia máxima registradas são, respectivamente, de 16,67% e 66,67%. A acurácia com maior frequência é de 50%; isto significa que, com a amostra 1, a maioria das variáveis tem 50% de chance de detectar fraudes em relatórios financeiros com o modelo linear. Para o modelo polinomial, a acurácia mínima registrada é de 25% e a acurácia máxima é de 75%. A moda das acurácias é de 17%; isto significa que, com a amostra 1, a maioria das variáveis tem apenas 17% de chance de detectar fraudes em relatórios financeiros com o modelo polinomial. Para o modelo radial, a acurácia mínima registrada é de 25% e a acurácia máxima é de 100%. A moda das acurácias é de 33%; isto significa que, com a amostra 1, a maioria das variáveis tem apenas 33% de chance de detectar fraudes em relatórios financeiros com o modelo radial. Para o modelo sigmoide, a acurácia mínima registrada é de 0% e a acurácia máxima é de 75%. A moda das acurácias é de 50%; isto significa que, com a amostra 1, a maioria das variáveis tem apenas 50% de chance de detectar fraudes em relatórios financeiros com o modelo sigmoide. Assim, nenhum dos modelos univariados foi significativo, de acordo com o teste de eficiência de McNemar.

Analisando a amostra 2, observou-se que, para o modelo linear, a acurácia mínima e acurácia máxima registradas são, respectivamente, de 42,86% e 60,71%. A acurácia com maior frequência é de 50%; isto significa que, com a amostra 2, a maioria das variáveis tem 50% de

chance de detectar fraudes em relatórios financeiros com o modelo linear e 100% com os modelos univariados significantes. Para o modelo polinomial, a acurácia mínima registrada é de 35,71% e a acurácia máxima é de 71,43%. A moda das acurácias é de 50%; isto significa que, com a amostra 2, a maioria das variáveis tem apenas 50% de chance de detectar fraudes em relatórios financeiros com o modelo polinomial e 87,50% com os modelos univariados significantes. Para o modelo radial, a acurácia mínima registrada é de 32,14% e a acurácia máxima é de 75%. A moda das acurácias é de 46,42%; isto significa que, com a amostra 2, a maioria das variáveis tem 46,42% de chance de detectar fraudes em relatórios financeiros com o modelo radial e 41,67% com os modelos univariados significantes. Para o modelo sigmoide, a acurácia mínima registrada é de 35,71% e a acurácia máxima é de 64,29%. A moda das acurácias é de 57,14%; isto significa que, com a amostra 2, a maioria das variáveis tem apenas 57,14% de chance de detectar fraudes em relatórios financeiros com o modelo sigmoide e 20,83% com os modelos univariados significantes.

Visto que os modelos univariados com a amostra 1 não foram significativos, seus modelos multivariados não foram executados, pois nenhum modelo univariado atendeu aos critérios de seleção de variáveis (ter uma acurácia superior ou igual a 50% e com uma significância de no máximo 10%).

Seguindo os critérios de seleção de variáveis para cada amostra e para cada modelo da SVM, as variáveis escolhidas para compor os modelos multivariados da amostra 2 para o modelo da SVM-linear foram: INF3; INF4; INF6; DPA1; DPA3; EDC1; EDC2; EDC3; EDC4; EDC5; EDC6; EDC8; EDC9; EDC12; LIQ2; LIQ3; REN3; REN4; REN5; REN6; REN7; REN10; REN11; REN13; REN15; REN16; REN17; OUT2; OUT3; e OUT5. Para o modelo da SVM-polinomial, as variáveis selecionadas foram: INF2; INF3; INF4; INF5; INF6; DPA1; DPA2; DPA3; EDC1; EDC2; EDC5; EDC9; EDC11; EDC12; LIQ1; LIQ2; LIQ3; LIQ4; LIQ5; REN2; REN3; REN4; REN5; REN6; REN7; REN8; REN9; REN10; REN12; REN13; REN14; REN15; REN16; OUT1; OUT3; OUT4; e OUT5. Para o modelo da SVM-radial, as variáveis selecionadas foram: INF1; INF4; EDC1; EDC2; EDC8; EDC12; LIQ4; LIQ5; REN1; REN2; REN4; REN6; e REN14. Para o modelo da SVM-sigmoide, as variáveis selecionadas foram: INF5; EDC3; EDC12; LIQ1; LIQ3; LIQ5; REN2; REN5; REN11; REN15; e OUT5.

5 Análise dos resultados

Nesta seção, busca-se executar os testes estatísticos para o cumprimento dos objetivos propostos e a fim de responder aos problemas da pesquisa. A primeira parte trata do modelo de estudo de eventos e a segunda parte trata do modelo para a detecção da fraude. A primeira

parte é subdividida em duas partes. A primeira subdivisão da primeira parte trata dos comportamentos dos preços das ações após a realização das fraudes em relatórios financeiros e a segunda subdivisão trata dos comportamentos dos preços das ações à detecção de fraudes em relatórios financeiros.

5.1 Análise dos resultados do estudo de eventos

No primeiro momento, analisou-se o efeito da realização de fraudes em relatórios financeiros no mercado, considerando um intervalo de 10 dias depois do evento. No segundo momento, analisou-se o efeito da detecção de fraudes em relatórios financeiros no mercado, considerando um intervalo de 10 dias depois do evento. Por fim, analisaram-se as variáveis que podem auxiliar a detectar fraudes em relatórios financeiros (em outras palavras, buscou-se identificar as variáveis que podem servir para identificar os riscos de fraudes em relatórios financeiros) e a capacidade da SVM para detectar eficientemente fraudes em relatórios financeiros no Brasil.

5.1.1 Análise dos resultados do estudo de eventos depois da realização de fraudes sem conhecimento do mercado

A análise do estudo de eventos à realização de fraudes foi feita observando os retornos anormais individuais significativos, a fim de se buscar identificar a existência de retornos anormais individualmente. A inferência final é feita baseando-se na soma acumulada dos retornos anormais.

Analisando individualmente cada evento, conclui-se que, na maioria das vezes, os retornos anormais à realização de fraudes em relatórios financeiros não foram estatisticamente significativos para inferir a sua existência. Acerca do Apêndice 6, que apresenta os retornos anormais, observa-se que, no primeiro dia depois da realização de fraudes, apenas os retornos anormais das entidades Cobrasma S/A (CBMA3), Companhia Energética de Minas Gerais S.A (CMIG3), Inepar SA Indústria e Construções (INEP3) e Taurus Armas SA (TASA3) são significativos, com, pelo menos, 90% de confiança. Dois dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Grazziotin SA (CGRA3), Doc Imbituba (IMBI3), 2 eventos da Petrobras (PETR3) e Recrusul SA (RCSL3) foram significativos, com, pelo menos, 90% de confiança. Três dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Grazziotin SA (CGRA3), Gerdau (GGBR3), JBS (JBSS3) e Petrobras (PETR3) foram significativos, com, pelo menos, 90% de confiança. Quatro dias depois da realização de fraudes em relatórios financeiros, apenas os

retornos anormais de Cobrasma S/A (CBMA3) e Recrusul SA (RCSL3) foram significativos, com, pelo menos, 90% de confiança. Cinco dias depois da realização de fraudes em relatórios financeiros, apenas o retorno anormal da Petrobras (PETR3) foi significativo, com 99% de confiança. Seis dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Embraer (EMBR3) e Petrobras (PETR3) foram significativos, com, pelo menos, 90% de confiança. Sete dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Cobrasma S/A (CBMA3), Petrobras (PETR3) e Recrusul SA (RCSL3) foram significativos, com, pelo menos, 90% de confiança. Oito dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Inepar SA Indústria e Construções (INEP3), Recrusul SA (RCSL3) e Taurus Armas SA (TASA3) foram significativos, com, pelo menos, 90% de confiança. Nove dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Cobrasma S/A (CBMA3), Companhia Energética de Minas Gerais S.A (CMIG3), Inepar SA Indústria e Construções (INEP3) e Recrusul SA (RCSL3) foram significativos, com, pelo menos, 90% de confiança. Dez dias depois da realização de fraudes em relatórios financeiros, apenas os retornos anormais de Cobrasma S/A (CBMA3), Inepar SA Indústria e Construções (INEP3) e Petrobras (PETR3) foram significativos, com, pelo menos, 90% de confiança.

Dos 40 ativos analisados, 11 (CBMA3, CMIG3, INEP3, TASA3, CGRA3, IMBI3, PETR3, RCSL3, GGBR3, JBSS3 e EMBR3) tiveram, pelo menos, um retorno anormal significativo, com, pelo menos, 90% de confiança.

Apesar da impossibilidade de inferir individualmente (ou seja, para cada evento separadamente) a existência de retornos anormais na maioria das vezes, a análise da soma cumulativa dos retornos anormais (ou seja, considerando todos os eventos conjuntamente) relata a existência de retornos anormais significativos a partir do quarto dia depois da realização de fraudes em relatórios financeiros, conforme apresentado na Tabela 8.

Tabela 8 – Soma acumulada dos retornos anormais à realização de fraudes (CAAR_Real)

Período	Antes da convergência contábil (até 2009)		Depois da convergência contábil (a partir de 2010)	
	CAAR	Estatística t de Student	CAAR	Estatística t de Student
1	0.0161	1.9382	0.0169	1.7622*
2	0.0106	1.4253	0.0151	1.1117
3	0.0337	1.7725	0.0297	1.4814
4	0.0356	1.7097	0.0483	1.7491*
5	0.0195	0.6838	0.0470	2.4673**
6	0.0433	1.2113	0.0524	2.3998**
7	0.0464	1.2794	0.0507	2.3122**
8	0.0712	1.6479	0.0635	2.4657**
9	0.0786	1.7318	0.0610	2.4952**

10	0.0902	1.7848	0.0535	2.1084**
----	--------	--------	--------	----------

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Antes da convergência contábil, a realização de fraudes causou retornos anormais positivos, porém estatisticamente não significativos. Observou-se que, durante os 10 dias de observação, os retornos anormais aumentaram com o passar do tempo.

Depois da convergência contábil, a realização de fraudes causou retornos anormais positivos e estatisticamente significativos na maioria das vezes. Um dia depois da realização de fraudes, constatou-se um retorno anormal positivo de 0.0169, com uma significância de 10% (com uma confiança de 90%). No segundo e terceiro dias após a realização de fraudes, os retornos anormais positivos 0.0151 e 0.0297 e não foram estatisticamente significativos. Quatro dias após a realização de fraudes, constatou-se um retorno anormal positivo de 0.0483, com uma significância de 10% (com uma confiança de 90%). A partir do quinto dia depois da realização de fraudes, verificaram-se retornos anormais positivos, respectivamente, de 0.0470, 0.0524, 0.0507, 0.0635, 0.0610 e 0.0535, com uma significância de 5% (com uma confiança de 95%).

De forma global, a realização de fraude causou retornos anormais positivos desde o primeiro dia até o décimo dia depois do evento, e, portanto, a partir das evidências, é possível concluir pela existência de retornos anormais positivos à realização de fraudes em relatórios financeiros no período de 2010 até 2021.

O aumento nos retornos anormais é o reflexo do aumento nos preços das ações e, portanto, do aumento do valor de mercado das entidades fraudadoras, demonstrando que as empresas obtêm vantagens com a manipulação das informações contábeis frente aos investidores, em períodos em que a referida manipulação (fraude) não é de conhecimento público.

A evidência desta análise, está em consonância com as evidências de Du (2017) segundo o qual retornos de swaps de default de crédito mais positivos estão associados a rerepresentações envolvendo fraude. Em outras palavras, fraudes causam retornos positivos inesperados (anormais).

Esse aumento do valor de mercado da entidade confirma o argumento de Segrestin e Hatchuel (2011) de que, de acordo com a teoria da agência, os gestores tomam decisões para agradar aos acionistas. Provavelmente, a situação econômica das entidades não estavam como os acionistas esperavam e os gestores os gestores tiveram que cometer uma fraude e isso resultou no aumento de valor das entidades.

Este resultado confirma o argumento de Oliveira (2020) segundo o qual para manter sua credibilidade perante os investidores, as entidades tentam disfarçar qualquer sinal que demonstre quedas ou estagnação nos seus indicadores econômicos ou financeiros, recorrendo, entre outros mecanismos, as fraudes em relatórios financeiros, com o objetivo de manter sua capacidade de captação de recursos a custos mais atrativos do que realmente seria, caso seu verdadeiro desempenho econômico fosse demonstrado. Os retornos anormais positivos evidenciados após a realização de fraudes, confirmam o aumento de capacidade de capção de recursos ou, pelo menos, a capacidade de manter o status quo.

5.1.2 Análise dos resultados do estudo de eventos depois da detecção de fraudes

A análise do estudo de eventos à detecção de fraudes foi feita observando os retornos anormais individuais significativos depois da detecção de fraudes em relatórios financeiros, a fim de tentar identificar a possibilidade de identificar a existência de retornos anormais individualmente. A inferência sobre a presença de retornos anormais à detecção de fraudes em relatórios financeiros é feita baseando-se na soma acumulada dos retornos anormais.

Ao se analisar individualmente os eventos de detecção de fraudes em relatórios financeiros, identificou-se a não significância na maioria das vezes de retornos anormais negativos, inferindo a sua inexistência. Acerca do Apêndice 7, que apresenta os retornos anormais individuais à detecção de fraudes em relatórios financeiros, foi observado que, no primeiro dia depois da detecção de fraudes, apenas os retornos anormais das entidades Bristol-Myers Squibb Company (BMYB34.SA) e Rossi Residencial (RSID3.SA) foram significativos, com, pelo menos, 90% de confiança. No segundo dia depois da detecção de fraudes, apenas os retornos anormais das entidades Companhia Energética de Minas Gerais (CMIG3.SA) e PDG (PDGR3.SA) foram significativos, com, pelo menos, 90% de confiança.

No terceiro dia depois da detecção de fraudes, apenas o retorno anormal da entidade JHSF (JHSF3.SA) foi significativo, com 95% de confiança. No quarto dia depois da detecção de fraudes, apenas os retornos anormais das entidades PDG (PDGR3.SA) e Petrobras (PETR3.SA) foram significativos, com, pelo menos, 90% de confiança. No quinto dia depois da detecção de fraudes, apenas os retornos anormais das entidades Bahema SA (BAHI3.SA), Petrobras (PETR3.SA) e Pettenati SA Indústria Têxtil (PTNT3.SA) foram significativos, com, pelo menos, 90% de confiança. No sexto dia depois da detecção de fraudes, apenas os retornos anormais das entidades Bahema SA (BAHI3.SA) e Companhia Energética de Brasília (CEBR6.SA) foram significativos, com, pelo menos, 90% de confiança.

No sétimo dia depois da detecção de fraudes, apenas o retorno anormal da entidade Companhia Energética de Brasília (CEBR6.SA) foi significativo, com 99% de confiança. No oitavo dia depois da detecção de fraudes, apenas o retorno anormal da entidade Gerdau (GGBR3.SA) foi significativo, com 90% de confiança. No nono dia depois da detecção de fraudes, apenas os retornos anormais das entidades Companhia Energética de Brasília (CEBR6.SA) e Rossi Residencial (RSID3.SA) foram significativos, com, pelo menos, 90% de confiança. Por fim, no décimo dia depois da detecção de fraudes, apenas os retornos anormais das entidades Bristol-Myers Squibb Company (BMYB34.SA), Companhia Energética de Brasília (CEBR6.SA), Companhia Energética de Minas Gerais (CMIG3.SA) e Rossi Residencial (RCSL3.SA) foram significativos, com, pelo menos, 90% de confiança.

Analisando de forma global os retornos anormais por meio da soma cumulativa deles (ou seja, considerando todos os eventos conjuntamente), a inexistência de retornos anormais significativos estatisticamente depois da detecção de fraudes foi constatada, conforme apresentado na Tabela 9.

Tabela 9 – Soma acumulada dos retornos anormais à detecção de fraudes (CAAR_Det)

Período	Antes da convergência contábil (até 2009)		Depois da convergência contábil (a partir de 2010)	
	CAAR	Estatística t de Student	CAAR	Estatística t de Student
1	-0.0083	-0.198	-0.0052	-0.7022
2	-0.0216	-0.3992	0.0063	0.6678
3	-0.0426	-0.6653	-0.0056	-0.3620
4	-0.0154	-0.213	0.0066	0.4832
5	-0.0122	-0.1526	0.0156	0.8792
6	-0.0173	-0.1991	0.0000	-0.0019
7	-0.0435	-0.466	0.0173	0.9641
8	-0.0663	-0.6672	0.0239	1.2971
9	-0.0537	-0.5112	0.0130	0.7051
10	-0.0537	-0.5112	-0.0056	-0.2313

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Antes da convergência contábil, a detecção de fraudes causou retornos anormais negativos, porém estatisticamente não significativos, enquanto que, depois da convergência, a detecção de fraudes causou retornos anormais estatisticamente não significativos negativos por alguns períodos e positivos em outros períodos. No primeiro, terceiro, sexto e décimo dia depois da detecção de fraudes, os retornos anormais foram negativos, enquanto que, no resto do período (ou seja, no segundo, quarto, quinto, sétimo, oitavo e nono dia), os retornos anormais foram positivos.

De forma global, a detecção de fraude não causou retornos anormais, pois estes são não significativos, permitindo inferir que, no Brasil e no período analisado, depois da descoberta

de fraudes em relatórios financeiros, é possível que haja retornos anormais, porém estes não seriam significativos. Desse modo, o mercado não consegue – de forma uniforme, ou, pelo menos, consistente – agir em resposta favorável à decisão do órgão regulador, o que evidencia que o mercado tende a ser indiferente à detecção da fraude, considerando-se os parâmetros utilizados nesta pesquisa.

Essas evidências apontam, de forma combinada, que as empresas que praticaram fraudes em suas demonstrações financeiras tiveram ganhos em dois momentos, dado que, no período em que introduziram a manipulação em suas informações contábeis, obtiveram retornos anormais positivos e significativos, ao passo que, em momento posterior – quando a fraude é detectada e divulgada publicamente –, o mercado age de forma indiferente, não havendo retornos anormais significativos, de acordo com os achados deste estudo.

Era esperado que, assim que o mercado souber da fraudes, retornos anormais negativos sejam registrados conforme evidências de Tsai e Huang (2020), Palmrose et al. (2001) e Karajian e Ullah (2021). No entanto, foi a evidência de Eryigit (2019) que foi confirmada nesta análise. Eryigit (2019) apresenta evidências de que os retornos anormais causados pelo anúncio de fraudes em relatórios financeiros não são significantes para inferir a sua existência ou não, dado que o anúncio de envolvimento de entidades em fraudes ou irregularidades contábeis foi considerado como não significativo estatisticamente.

5.2 Análise dos resultados da SVM

De acordo com Sharma e Panigrahi (2012) os sistemas de auditoria estão falhando na sua tarefa de assegurar que os relatórios financeiros estão livres de distorções materiais. Como evidenciado na sessão 5.1, fraudes em relatórios financeiros podem causar retornos anormais positivos, sobretudo logo após a realização da fraude (momento em que a fraude não foi descoberta). Isso leva a recorrer, segundo Sharma e Panigrahi (2012), às técnicas da contabilidade investigativa como a SVM (uma das técnicas de mineração de dados).

Baseando-se nos indicadores financeiros e não financeiros que podem sinalizar risco de fraudes nos relatórios financeiros, os modelos da SVM foram estimados com o objetivo de identificar a acurácia do modelo que representa a proporção de relatórios fraudulentos classificados adequadamente. Quanto maior a especificidade, menor seria a chance de um relatório fraudulento não ser detectado. Assim, o modelo com maior acurácia seria aquele indicado para detecção de fraudes em relatórios fraudulentos no Brasil.

Ressalta-se que, provavelmente, por insuficiência amostral, os modelos finais da amostra 1 (subdivisão da amostra contendo os relatórios de 1998 até 2009) não foram

executados. Os resultados das análises dos modelos finais de detecção de fraudes em relatórios financeiros por meio da máquina de suporte vetorial, executados com a amostra 2, estão apresentados na Tabela 10.

Tabela 10 – Acurácia de detecção de fraudes em relatórios fraudulentos (amostra 2)

Modelo de SVM	Acurácia	<i>p-value</i> do teste de McNemar
Linear	60.71%	0.2278
Polinomial	64.29%	0.0269
Radial	60.71%	0.2278
Sigmoide	57.14%	0.1489

Fonte: dados da pesquisa.

Analisando os modelos finais multivariados da amostra 2 com variáveis que registraram uma acurácia maior ou igual a 50% nos modelos univariados e que foram significantes com uma significância de, pelo menos, 10%, de acordo com o teste de eficiência de McNemar (conforme os resultados da Tabela 10), foi observado que o modelo linear consegue detectar até 61% dos relatórios fraudulentos, porém não significativo estatisticamente. O modelo polinomial consegue detectar 64% dos relatórios fraudulentos, com significância de 1%, enquanto o modelo radial consegue detectar 60% dos relatórios fraudulentos, mas não é significativo. Por fim, o modelo sigmoide consegue detectar 57% dos relatórios, sendo não significativo. Dessa amostra, com esse critério de seleção de variáveis que podem sinalizar riscos de fraudes e prezando pela qualidade (significância) de detectar fraudes em relatórios fraudulentos, o modelo polinomial utilizado nesta pesquisa foi o que encontrou melhor ajuste.

Em geral, a título de conclusão, baseando-se nas variáveis que podem sinalizar riscos de fraudes, o modelo polinomial da SVM é melhor para detectar fraudes em relatórios financeiros no Brasil, podendo, de acordo com a amostra estudada, detectar até 64% dos relatórios fraudulentos.

5.3 Teste da escolha do melhor modelo de detecção de fraude

Usando as acurácias dos modelos de SVM, o modelo polinomial foi escolhido como o melhor modelo de detecção de fraudes. Com objetivo de testar estatisticamente essa escolha, o teste de Vuong (1989) foi empregado.

Com os modelos de classificação da SVM, não é possível obter os coeficientes de regressão das variáveis o máximo verossimelhança necessários para executar o teste de Vuong (1989). Por isso, as variáveis sensíveis à detecção de fraudes sugeridas por cada modelo da SVM foram usadas para executar regressões logit. A regressão logit foi escolhida por ser um modelo similar aos modelos de classificação da SVM e um modelo de regressão capaz de ser

utilizado no teste de Vuong (1989). Os quatros modelos foram comparados dois a dois e a tabela 11 apresenta os resultados do teste.

Tabela 11 – Comparação dos modelos do SVM

	Linear		Polinomial		Radial		Sigmoide	
	Estatística-Z do Vuong	<i>p-value</i>	Estatística-Z do Vuong	<i>p-value</i>	Estatística-Z do Vuong	<i>p-value</i>	Estatística-Z do Vuong	<i>p-value</i>
Linear	-	-	-1.8835	0.0298	3.0279	0.0012	3.0517	0.0011
Polinomial	1.8835	0.0298	-	-	4.0682	0.0000	3.9636	0,0001
Radial	-3.0279	0.0012	-4.0682	0,0000	-	-	-0.3243	0.3729
Sigmoid	-3.0517	0.0011	-3.9636	0,0000	0.3243	0.3729	-	-

Fonte: dados da pesquisa.

O teste de Vuong (1989) foi executado comparando, dois a dois cada modelo com os outros três modelos. Ao comparar o modelo linear com os outros modelos, observa-se que o modelo polinomial é melhor que o modelo linear com uma significância de 5%, enquanto que o modelo linear é melhor que os modelos radial e sigmoide com uma significância de 1%.

Ao comparar o modelo polinomial com os outros modelos, observa-se que o modelo polinomial é melhor que o modelo linear com uma significância de 5% e melhor que os modelos radial e sigmoide com uma significância de 1%. Conclui-se desta análise que o modelo polinomial é melhor para detectar fraudes.

Ao comparar o modelo radial com os outros modelos, observa-se que o modelo linear é melhor que o modelo radial com uma significância de 1%, enquanto o modelo polinomial é melhor que o modelo radial com uma significância de 1%. O modelo sigmoide é melhor que o modelo radial, mas sem significância, os dois modelos teriam desempenhos iguais na detecção de fraudes.

Ao comparar o modelo sigmoide com os outros modelos, observa-se que os modelos linear e polinomial são melhores que o modelo sigmoid com uma significância de 1%. O modelo radial é melhor que o modelo sigmoid, mas sem significância, os dois modelos teriam desempenhos iguais na detecção de fraudes.

Por meio do teste de Vuong (1989), conclui-se que o modelo de classificação polinomial do SVM é melhor na detecção de fraudes. O resultado do teste de Vuong (1989) corrobora a conclusão das análises feitas baseando-se nas acurácias de classificação dos próprios modelos.

6 Considerações finais

Conforme os argumentos de Kostova (2013), as fraudes em relatórios financeiros têm o objetivo de apresentar uma imagem equivocada e, sobretudo, ambígua, sinalizando condições favoráveis às entidades. De acordo com Stiglitz (2000), a finalidade das fraudes é de fornecer às empresas ganhos suplementares. Respondendo ao primeiro objetivo do estudo, qual seja, de se analisar a existência de possíveis retornos anormais à realização de fraudes em relatórios financeiros, pode-se afirmar que retornos anormais positivos foram identificados ao longo dos 10 primeiros dias depois da realização da fraude.

As consequências (retornos anormais positivos) das fraudes em relatórios financeiros surgem logo depois do primeiro dia depois da realização da fraude, enquanto que, nos dois seguintes dias, as consequências são insignificantes, mas retornam a partir do quarto dia e permanecem até, no mínimo, o décimo dia depois da realização de fraudes.

A respeito do segundo objetivo da dissertação, que foi analisar a existência de retornos anormais depois da detecção de fraudes em relatórios financeiros, nenhuma significância foi encontrada nos retornos anormais depois do dia em que a CVM julga oficialmente, por meio de processos administrativos sancionadores, a prática ou manobra suspeita como fraudulenta por estar infringindo um determinado artigo de uma determinada lei. Isto significa que o mercado ignora a decisão da CVM nos processos administradores sancionadores.

Supõe-se que a decisão da CVM confirma oficialmente a realização da fraude, trazendo ao conhecimento do público o evento fraudulento. Conforme as análises realizadas, foi rejeitada a hipótese do estudo (H_1 : Fraudes em relatórios financeiros causam retornos anormais negativos no preço dos ativos quando são de conhecimento público do mercado).

Nesse sentido, é possível que o mercado já tenha reagido negativamente aos rumores não oficiais de que uma determinada entidade tenha cometido fraudes em relatórios financeiros dentro de um período considerado anterior à data oficial de divulgação da detecção de fraudes. Assim, existe a hipótese de que comportamentos anormais nos preços das ações possam ser registrados a partir de tal reação negativa. Dessa forma, com a finalidade de não deixar o mercado ser impactado por rumores não oficiais, o órgão regulador precisa prever, com muita antecedência, fraudes em relatórios financeiros.

Respondendo ao terceiro objetivo da dissertação, qual seja, analisar a acurácia da SVM em identificar fraudes em relatórios financeiros, foram considerados 34 indicadores financeiros e não financeiros que podem ser observados para a identificação de riscos de fraudes. Esses indicadores são: “Participação minoritária”; “Dívida total líquida”; “Dívida total bruta”; “Liquidez corrente”; “Capital de giro”; “EBITDA”; “LAIR + Despesa financeira líquida”; “ROE Rentabilidade do patrimônio (patrimônio médio)”; “ROIC (*Invested Capital* final)”; “*Invested capital*”; “Depreciação e amortização”; “*Capital expenditure / Depreciação*”; “Ciclo financeiro”; “Ciclo operacional”; “Giro do ativo”; “Valor patrimonial por ação”; “Venda por ação”; “EBIT / Dívida líquida”; “EBIT / Despesa financeira bruta”; “EBIT / Despesa financeira líquida”; “Margem líquida”; “ROA Rentabilidade do ativo”; “ROE Rentabilidade do patrimônio (patrimônio final)”; “ROE Rentabilidade do patrimônio (patrimônio inicial)”; “Rentabilidade de investimento de subsidiários”; “Alavancagem operacional”; “Fluxo de caixa livre”; “Dívida líquida / EBITDA”; “Liquidez geral”; “*Capital*

employed”; “*Capital expenditure* (Despesa de capital incluindo os investimentos)”; “Depreciação / Imobilizado e intangíveis”; “Índice de endividamento (Dívida bruta / Ativo total)”; e “ROIC (*Invested capital* médio)”.

Com base nos achados desta pesquisa, ressalta-se que o *support vector machine*, por meio do seu modelo polinomial, consegue detectar até 64% dos relatórios fraudulentos. Quanto à investigação sobre a suspeita de fraude, assim que os relatórios são enviados ao órgão regulador, este procede a uma atividade de mitigação a fim de tentar prever os relatórios fraudulentos. Considera-se que, ao identificar possíveis relatórios fraudulentos, o órgão regulador não deveria permitir que estes sejam publicados ao mercado. Já que isto evitaria ganhos anormais da entidade fraudadora.

No âmbito do tema aqui pesquisado, estudos futuros podem ter como objetivo a análise da constância das consequências das fraudes em relatórios financeiros no Brasil. Tais estudos seriam capazes de identificar até quando os retornos anormais, vinculados à realização de fraudes em relatórios financeiros, continuariam existindo. Isto poderia justificar a magnitude dos ganhos que as entidades fraudadoras obtêm com a prática ilícita.

Com base nos resultados aqui obtidos, visto que os investidores não reagiram nem negativamente nem positivamente, de forma significativa à detecção oficial de fraudes em relatórios financeiros pela CVM, pesquisas futuras podem estudar a data próxima em que o mercado tinha reagido e antes da data oficial de detecção de fraudes. Simplesmente, a data em que os rumores (não oficiais) iniciaram, a data em que a CVM recebeu a denúncia, ou, ainda, a data de anúncios nos jornais sobre a suspeita de fraudes podem ser estudadas.

Além do mais, visto que os investidores não reagiram nem negativamente nem positivamente, de forma significativa, à detecção oficial de fraudes em relatórios financeiros pela CVM, pesquisas futuras podem estudar o que o mercado espera do órgão regulador sobre a detecção de fraudes.

Com o objetivo de otimizar os modelos de detecção de fraudes em relatórios financeiros no Brasil, pesquisas futuras podem estudar o efeito da penalização (custo C) nos modelos não lineares da SVM a fim de identificar o custo que otimize a acurácia de detecção de relatórios fraudulentos. Rakotomamonjy (2003) sugeriu um custo $C = 1000$ como penalidade do modelo por uma falsa classificação. Por se tratar de um estudo incipiente e de caráter exploratório sobre a detecção de fraudes em relatórios financeiros no Brasil, esta dissertação considerou um custo simbólico de $C = 1$, com o objetivo de não penalizar muito o modelo. Nesse sentido, é possível que estudos futuros investiguem o efeito da penalidade do modelo no mercado brasileiro.

Quatro pontos foram encontrados como limitações desta pesquisa. A primeira limitação está no fato de considerar apenas as empresas listadas no mercado brasileiro (B3). Logo, as inferências dos resultados do estudo se referem somente às entidades listadas na B3.

A segunda limitação da pesquisa consiste no fato de considerar apenas as entidades não financeiras. Por isso, não é possível inferir os mesmos resultados às entidades do setor financeiro.

A terceira limitação do estudo consiste no fato de serem analisados todos os setores de forma homogênea. Uma análise setorial seria mais específica, o que possibilita inferir os resultados de acordo com o setor da entidade.

Por fim, há de se mencionar a não discricionariedade da amostra pelo tamanho (seja ele pelo faturamento ou pelo número de funcionários empregados) das entidades, pelo número de integrantes de seus conselhos administrativos e pelo gênero de seus membros. Estes fatores, aqui desconsiderados, também podem ser variáveis relevantes para a investigação do potencial de realização de fraudes nas entidades.

Referências

- Adepoju, O., Wosowei, J., & Jaiman, H. (2019, October). Comparative evaluation of credit card fraud detection using machine learning techniques. In 2019 Global Conference for Advancement in Technology (GCAT) (pp. 1-6). IEEE. DOI: [10.1109/GCAT47503.2019.8978372](https://doi.org/10.1109/GCAT47503.2019.8978372)
- Aghghaleh, S. F., & Mohamed, Z. M. (2014). Fraud risk factors of fraud triangle and the likelihood of fraud occurrence: Evidence from Malaysia. *Information Management and Business Review*, 6(1), 1–7. doi.org/10.22610/imbr.v6i1.1095
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. [doi:10.1111/j.1540-6261.1968.tb00843.x](https://doi.org/10.1111/j.1540-6261.1968.tb00843.x)
- Alves, C. P. T. L. (2016). *Determinantes da denúncia da fraude por parte dos auditores portugueses* [Dissertação de mestrado, Universidade de Lisboa]. Repositório da Universidade de Lisboa. <https://www.repository.utl.pt/bitstream/10400.5/12643/1/DM-CPTLA-2016.pdf>.
- Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial reporting fraud and other forms of misconduct: A multidisciplinary review of the literature. *Review of Accounting Studies*, 23(2), 732–783. [doi:10.1007/s11142-017-9435-x](https://doi.org/10.1007/s11142-017-9435-x)
- Anthony, M. (1995). Classification by polynomial surfaces. *Discrete Applied Mathematics*, 61(2), 91–103. [doi.org/10.1016/0166-218X\(94\)00008-2](https://doi.org/10.1016/0166-218X(94)00008-2)
- Association of Certified Fraud Examiners. (2020). *Report to the nations on occupational fraud and abuse*. ACFE. <https://acfe-public.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf>
- Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 6(2), 159–178. [doi:10.2307/2490232](https://doi.org/10.2307/2490232)
- Bao, Y., Ke, B., Li, B., Yu, J., & Zhang, J. (2019). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Social Science Research Network*, 58(1), 199–235. <http://dx.doi.org/10.2139/ssrn.2670703>
- Bashir, A. K., Khan, S., Prabadevi, B., Deepa, N., Alnumay, W. S., Gadekallu, T. R., & Maddikunta, P. K. R. (2021). Comparative analysis of machine learning algorithms for prediction of smart grid stability. *International Transactions on Electrical Energy Systems*, 31(9), e12706. doi.org/10.1002/2050-7038.12706.
- Beatty, A., Liao, S., & Yu, J. J. (2013). The spillover effect of fraudulent financial reporting on peer firms' investments. *Journal of Accounting and Economics*, 55(2–3), 183–205. doi.org/10.1016/j.jacceco.2013.01.003
- Beaver, W. H., McNichols, M. F., & Rhie, J.-W. (2005). Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies*, 10(1), 93–122. [doi:10.2139/ssrn.634921](https://doi.org/10.2139/ssrn.634921)
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. doi.org/10.2469/faj.v55.n5.2296

- Bolsa Balcão Brasil. (2022). *Classificação setorial das empresas negociadas na B3*. B3. <https://bvmf.bmfbovespa.com.br/InstDados/InformacoesEmpresas/ClassifSetorial.zip>
- Brasil. (1976). *Lei 6.385*. Dispõe sobre o mercado de valores mobiliários e cria a Comissão de Valores Mobiliários. Presidência da República. https://www.planalto.gov.br/ccivil_03/leis/l6385compilada.htm
- Brasil. (1976). *Lei 6.404*. Dispõe sobre as Sociedades por Ações. Presidência da República. https://www.planalto.gov.br/ccivil_03/leis/l6404consol.htm
- Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics*, 8(3), 205–258. [doi:10.1016/0304-405x\(80\)90002-1](https://doi.org/10.1016/0304-405x(80)90002-1)
- Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics*, 14(1), 3–31. [doi.org/10.1016/0304-405X\(85\)90042-X](https://doi.org/10.1016/0304-405X(85)90042-X)
- Bryan, J., Wickham, H., & Barrett, M. (2021). *Usethis*. R Studio. <https://usethis.r-lib.org/>
- Camps-Valls, G., Martín-Guerrero, J. D., Rojo-Alvarez, J. L., & Soria-Olivas, E. (2004). Fuzzy sigmoid kernel for support vector classifiers. *Neurocomputing*, 62(C), 501–506. [10.1016/j.neucom.2004.07.004](https://doi.org/10.1016/j.neucom.2004.07.004)
- Chauhan, V. K., Dahiya, K., & Sharma, A. (2019). Problem formulations and solvers in linear SVM: A review. *Artificial Intelligence Review*, 52(2), 803–855. <https://doi.org/10.1007/s10462-018-9614-6>
- Chen, R.-C., Chen, T.-S., & Lin, C.-C. (2006). A new binary support vector system for increasing detection rate of credit card fraud. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(2), 227–239. doi.org/10.1142/S0218001406004624
- Christensen, J. A., & Byington, J. R. (2003). The computer: An essential fraud detection tool. *Journal of Corporate Accounting & Finance*, 14(5), 23–27. <https://doi.org/10.1002/jcaf.10179>
- Comissão de Valores Mobiliários. (2020). *Missão, valores e objetivos estratégicos: Cadeia de valor. Mapa estratégico. Propósito e valores*. CVM. <https://www.gov.br/cvm/pt-br/aceso-a-informacao-cvm/institucional/missao-valores-e-objetivos-estrategicos>
- Comissão de Valores Mobiliários. (2021). *Processos sancionadores julgados*. CVM.
- Condé, R. A. D., de Almeida, C. O. F., & Quintal, R. S. (2015). Fraude Contábil: Análise empírica à luz dos pressupostos teóricos do triângulo da fraude e dos escândalos corporativos. *Gestão & Regionalidade*, 31(93). doi.org/10.13037/gr.vol31n93.2929.
- Conselho Federal de Contabilidade. (2009a). *Norma Brasileira de Contabilidade (NBC TA) 240: Responsabilidade do auditor em relação a fraude, no contexto da auditoria de demonstrações contábeis*. CFC. https://cfc.org.br/wp-content/uploads/2016/02/NBC-TA_110716.pdf
- Conselho Federal de Contabilidade. (2009b). *Resolução CFC nº 1.207.09: Aprova a NBC TA 240 – Responsabilidade do auditor em relação a fraude, no contexto da auditoria de demonstrações contábeis*. CFC. https://www.oas.org/juridico/portuguese/res_1207.pdf

- Corrado, C. J. (2011). Event studies: A methodology review. *Accounting & Finance*, 51(1), 207–234. [doi:10.1111/j.1467-629x.2010.00375.x](https://doi.org/10.1111/j.1467-629x.2010.00375.x)
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.
- Dalnial, H., Kamaluddin, A., Sanusi, Z. M., & Khairuddin, K. S. (2014). Accountability in financial reporting: detecting fraudulent firms. *Procedia-Social and Behavioral Sciences*, 145, 61-69. doi.org/10.1016/j.sbspro.2014.06.011.
- Da Silva, A. M. C., & Da Silva, M. D. O. P. (2021). Hipótese do mercado eficiente e teoria de precificação de ativos. In K. K. N. Almeida, & R. D. França (Orgs.), *Teorias aplicadas à pesquisa em contabilidade: Uma introdução às teorias econômicas, organizacionais e comportamentais* (pp. 195–219). Editora UFPB. <http://www.editora.ufpb.br/sistema/press5/index.php/UFPB/catalog/book/758>
- Darrough, M., Huang, R., & Zhao, S. (2020). Spillover effects of fraud allegations and investor sentiment. *Contemporary Accounting Research*, 37(2), 982–1014. doi.org/10.1111/1911-3846.12541
- De Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., & Skidmore, A. K. (2006). Comparing accuracy assessments to infer superiority of image classification methods. *International Journal of Remote Sensing*, 27(1), 223–232. [doi:10.1080/01431160500275762](https://doi.org/10.1080/01431160500275762).
- Dheepa, V., & Dhanapal, R. (2012). Behavior based credit card fraud detection using support vector machines. *ICTACT Journal on Soft computing*, 2(4), 391-397. <https://doaj.org/article/3ee05fa8ecc5411496717c349747d26d>.
- Du, L. (2017). The CDS market reaction to restatement announcements. *Journal of Business Finance & Accounting*, 44(7–8), 1015–1035. [doi:10.1111/jbfa.12250](https://doi.org/10.1111/jbfa.12250)
- Dyck, A., Morse, A., & Zingales, L. (2010). Who blows the whistle on corporate fraud? *The Journal of Finance*, 65(6), 2213–2253. [doi:10.1111/j.1540-6261.2010.01614.x](https://doi.org/10.1111/j.1540-6261.2010.01614.x)
- Eckert, C., Gatzert, N., & Pisula, A. (2019). Spillover effects in the European financial services industry from internal fraud events: Comparing three cases of rogue trader scandals. *The Journal of Risk Finance*, 20(3), 249–266. [doi:10.1108/jrf-07-2018-0117](https://doi.org/10.1108/jrf-07-2018-0117)
- Eryigit, M. (2019). Short-term performance of stocks after fraudulent financial reporting announcement. *Journal of Financial Crime*, 26(2), 464–476. [10.1108/JFC-11-2016-0076](https://doi.org/10.1108/JFC-11-2016-0076)
- EventStudyTools. (2021). *Expected return models*. EST. <https://www.eventstudytools.com/expected-return-models>
- EventStudyTools. (2021). *Instructions R-Package “EventStudy”*. EST. <https://www.eventstudytools.com/R-package>
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417. [doi:10.1111/j.1540-6261.1970.tb00518.x](https://doi.org/10.1111/j.1540-6261.1970.tb00518.x)
- Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance*, 46(5), 1575–1617. [doi:10.2307/2328565](https://doi.org/10.2307/2328565)

- Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustments of stock prices to new information. *International Economic Review*, 10(1), 1–21. [doi:10.2307/2525569](https://doi.org/10.2307/2525569)
- Finnerty, J. D., Hegde, S., & Malone, C. B. (2016). Fraud and firm performance: Keeping the good times (apparently) rolling. *Managerial Finance*, 42(2), 151–172. [doi:10.1108/MF-01-2015-0009](https://doi.org/10.1108/MF-01-2015-0009)
- Frendy, & Hu, D. (2014). Japanese stock market reaction to announcements of news affecting auditors' reputation: The case of the Olympus fraud. *Journal of Contemporary Accounting & Economics*, 10(3), 206–224. [doi:10.1016/j.jcae.2014.08.004](https://doi.org/10.1016/j.jcae.2014.08.004)
- Gao, L., & Srivastava, R. P. (2007). The anatomy of management fraud schemes: Analyses and implications. *Indian Accounting Review*, 15(1), 1–23. https://www.academia.edu/21001048/The_anatomy_of_management_fraud_schemes_Analyses_and_implications
- Gepp, A., Kumar, K., & Bhattacharya, S. (2021). Lifting the numbers game: Identifying key input variables and a best-performing model to detect financial statement fraud. *Accounting & Finance*, 61(3), 4601–4638. doi.org/10.1111/acfi.12742
- Gonçalves, C. V. (2019). SVM aplicado a criptomoeda Ethereum (Dissertation, Universidade de São Paulo). <https://www.teses.usp.br/teses/disponiveis/55/55137/tde-04022020-121437/publico/ClaudioViniciusGoncalves.pdf>.
- Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3), 393–408. [doi:10.2307/1805228](https://doi.org/10.2307/1805228)
- Gupta, P. K., & Gupta, S. (2015). Corporate frauds in India – perceptions and emerging issues. *Journal of Financial Crime*, 22(1), 79–103. doi.org/10.1108/JFC-07-2013-0045
- Gupta, R., & Gill, N. S. (2012). A data mining framework for prevention and detection of financial statement fraud. *International Journal of Computer Applications*, 50(8), 7–14. <https://research.ijcaonline.org/volume50/number8/pxc3880889.pdf>
- Gupta, S., & Mehta, S. K. (2020). Feature selection for dimension reduction of financial data for detection of financial statement frauds in context to Indian Companies. *Global Business Review*, 1–26. doi.org/10.1177/0972150920928663
- Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud: A comparative study of machine learning methods. *Knowledge-Based Systems*, 128, 139–152. [doi:10.1016/j.knosys.2017.05.001](https://doi.org/10.1016/j.knosys.2017.05.001)
- Harvey, C. R., Liu, Y., & Zhu, H. (2016). ...and the cross-section of expected returns. *The Review of Financial Studies*, 29(1), 5–68. [doi:10.1093/rfs/hhv059](https://doi.org/10.1093/rfs/hhv059)
- Hester, J., & Wickham, C. (2021). *Devtools*. R Studio. <https://devtools.r-lib.org/>
- Hung, M., Wong, T. J., & Zhang, F. (2015). The value of political ties versus market credibility: Evidence from corporate scandals in China. *Contemporary Accounting Research*, 32(4), 1641–1675. [doi:10.2139/ssrn.2607783](https://doi.org/10.2139/ssrn.2607783)

- Ikbal, M., Irwansyah, I., Paminto, A., Ulfah, Y., & Darma, D. C. (2020). Explores the specific context of financial statement fraud based on empirical from Indonesia. *Universal Journal of Accounting and Finance*, 8(2), 29-40. [doi:10.13189/ujaf.2020.080201](https://doi.org/10.13189/ujaf.2020.080201).
- International Federation of Accountants. (2009). *International Standard on Auditing (ISA) 205: Consideration of laws and regulations in an audit of financial statements*. IFAC. <https://www.ifac.org/system/files/downloads/a013-2010-iaasb-handbook-isa-250.pdf>
- Joe, D. Y., & Oh, F. D. (2018). Spillover effects within business groups: The case of Korean chaebols. *Management Science*, 64(3), 1396–1412. doi.org/10.1287/mnsc.2016.2596
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292. [doi:10.2307/1914185](https://doi.org/10.2307/1914185)
- Kanapickiene, R., & Grundiene, Ž. (2015). The model of fraud detection in financial statements by means of financial ratios. *Procedia – Social and Behavioral Sciences*, 213, 321–327. [doi:10.1016/j.sbspro.2015.11.545](https://doi.org/10.1016/j.sbspro.2015.11.545)
- Karajian, S., & Ullah, S. (2021). Consequences of fraud and overcoming negative market reaction. *Global Finance Journal*, 52, 100635. [doi:10.1016/j.gfj.2021.100635](https://doi.org/10.1016/j.gfj.2021.100635)
- Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and databases in financial misconduct research. *The Accounting Review*, 92(6), 129–163. [doi:10.2308/accr-51766](https://doi.org/10.2308/accr-51766)
- Khan, R., Corney, M., Clark, A., & Mohay, G. (2010). Transaction mining for fraud detection in ERP Systems. *Industrial Engineering and Management Systems*, 9(2), 141–156. <https://doi.org/10.7232/iems.2010.9.2.141>
- Kostova, S. (2013). Audit procedures for disclosure of errors and fraud in financial statements. *Ekonomске Teme*, 51(2), 355–375. http://ekonomsketeme.rs/pdf/et2013en2_07.pdf
- Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In B. E. Eckbo (Ed.), *Handbook of Empirical Corporate Finance* (Vol. 1, pp. 3–36). North Holland. <https://doi.org/10.1016/B978-0-444-53265-7.50015-9>
- Kuhn, M. (2019). *The caret Package*. GitHub. <https://topepo.github.io/caret/>
- Latif, M., Arshad, S., Fatima, M., & Farooq, S. (2011). Market efficiency, market anomalies, causes, evidences, and some behavioral aspects of market anomalies. *Research Journal of Finance and Accounting*, 2(9), 1–13. <https://www.iiste.org/Journals/index.php/RJFA/article/view/1287/1208>
- Lenard, M. J., & Alam, P. (2009). An historical perspective on fraud detection: From bankruptcy models to most effective indicators of fraud in recent incidents. *Journal of Forensic & Investigative Accounting*, 1(1), 1–27. https://www.researchgate.net/publication/253950844_An_Historical_Perspective_on_Fraud_Detection_From_Bankruptcy_Models_to_Most_Effective_Indicators_of_Fraud_in_Recent_Incidents
- Lenard, M. J., Watkins, A. L., & Alam, P. (2007). Effective use of integrated decision making: An advanced technology model for evaluating fraud in service-based

- computer and technology firms. *The Journal of Emerging Technologies in Accounting*, 4(1), 123–137. doi.org/10.2308/jeta.2007.4.1.123
- Liew, V. K. S. (2020). Abnormal returns on tourism shares in the Chinese stock exchanges amid the COVID-19 pandemic. doi.org/10.2139/ssrn.3863889.
- Lin, C.-C., Chiu, A.-A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 89, 459–470. doi.org/10.1016/j.knosys.2015.08.011
- Liu, H., Manzoor, A., Wang, C., Zhang, L., & Manzoor, Z. (2020). The COVID-19 outbreak and affected countries stock markets response. *International Journal of Environmental Research and Public Health*, 17(8), 2800. doi.org/10.3390/ijerph17082800
- Lo, A. W. (2005). Reconciling efficient markets with behavioral finance: The adaptive markets hypothesis. *Journal of Investment Consulting*, 7(2), 21–44. <https://ssrn.com/abstract=1702447>
- MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13–39. [doi:10.2307/2729691](https://doi.org/10.2307/2729691)
- Martins, O. S., & Ventura Júnior, R. (2020). Influência da governança corporativa na mitigação de relatórios financeiros fraudulentos. *Revista Brasileira de Gestão de Negócios*, 22(1), 65–84. doi.org/10.7819/rbgn.v22i1.4039
- Meyer, D. (2021). *Support vector machines: The interface to libsvm in package e1071*. The Comprehensive R Archive Network. <https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf>
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.-C., & Lin, C.-C. (2021). *Package 'e1071'*. The Comprehensive R Archive Network. <https://cran.r-project.org/web/packages/e1071/e1071.pdf>
- Mohammadi, M., Yazdani, S., Khanmohammadi, M. H., & Maham, K. (2020). Financial reporting fraud detection: An analysis of data mining algorithms. *International Journal of Finance & Managerial Accounting*, 4(16), 1–12. https://ijfma.srbiau.ac.ir/article_15385.html
- Mueller, S. (2019). *EventStudy: Event study analysis*. R Studio. License: GPL (>= 2). <https://rdrr.io/cran/EventStudy/>
- Murcia, F. D. R., & Borba, J. A. (2007). Estrutura para detecção do risco de fraude nas demonstrações contábeis: Mapeando o ambiente fraudulento. *Brazilian Business Review*, 4(3), 171–190. <https://www.redalyc.org/pdf/1230/123016560001.pdf>
- Nascimento, M. R. D. (2020). *Contribuições de red flags para detecção de fraudes corporativas*. [Dissertação de mestrado, Faculdade de Administração, Ciências Contábeis e Ciências Econômicas]. Biblioteca Digital de Teses e Dissertações da Universidade Federal de Goiás. <https://repositorio.bc.ufg.br/tede/handle/tede/10523>.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109–131. doi.org/10.2307/2490395.

- Oliveira, R. X. D. (2020). *Risco de fraudes contábeis e ganho informacional na análise do risco de crédito para o setor bancário brasileiro*. [Dissertação de mestrado, Faculdade de Administração, Ciências Contábeis e Ciências Econômicas]. Repositório Institucional da Universidade de Brasília. https://repositorio.unb.br/bitstream/10482/41038/1/2020_RafaelXavierdeOliveira.pdf.
- Oliveira, V. A., & Lemes, S. (2011). Nível de convergência dos princípios contábeis brasileiros e norte-americanos às normas do IASB: Uma contribuição para a adoção das IFRS por empresas brasileiras. *Revista Contabilidade & Finanças*, 22(56), 155–173. doi.org/10.1590/S1519-70772011000200003.
- Ozili, P. K. (2015). Forensic accounting and fraud: A review of literature and policy implications. *International Journal of Accounting and Economics Studies*, 3(1), 63–68. <http://dx.doi.org/10.2139/ssrn.2628554>.
- Palmrose, Z.-V., Richardson, V. J., & Scholz, S. W. (2001). Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics*, 37(1), 59–89. [doi:10.2139/ssrn.265009](https://doi.org/10.2139/ssrn.265009).
- Park, N. K. (2004). A guide to using event study methods in multi-country settings. *Strategic Management Journal*, 25(7), 655–668. <https://doi.org/10.1002/smj.399>.
- Patle, A., & Chouhan, D. S. (2013, January). SVM kernel functions for classification. In 2013 International Conference on Advances in Technology and Engineering (ICATE) (pp. 1-9). IEEE. [doi:10.1109/ICAdTE.2013.6524743](https://doi.org/10.1109/ICAdTE.2013.6524743).
- Persons, O. (1995). Using financial statement data to identify factors associated with fraudulent financial reporting. *Journal of Applied Business Research*, 11(3), 38–46. doi.org/10.19030/jabr.v11i3.5858.
- Peterson, P.P. (1989). Event studies: A review of issues and methodology. *Quarterly Journal of Business and Economics*, 28(3), 36–66. <https://www.jstor.org/stable/40472954>.
- Prajapati, G. L., & Patle, A (2010). On performing classification using SVM with radial basis and polynomial kernel functions. *Proceedings of the 3rd International Conference on Emerging Trends in Engineering and Technology*, 512–515. [doi:10.1109/icetet.2010.134](https://doi.org/10.1109/icetet.2010.134).
- PricewaterhouseCoopers. (2020). *Global Economic Crime and Fraud Survey 2020: Fighting fraud: A never-ending battle*. PwC. https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/PwC_Global_Economic_Crime_and_Fraud_Survey_2020.pdf.
- PricewaterhouseCoopers. (2022). *Global Economic Crime and Fraud Survey 2022: Fighting fraud: A never-ending battle*. PwC. <https://www.pwc.com/gx/en/forensics/gecsm-2022/pdf/PwC%20%80%99s-Global-Economic-Crime-and-Fraud-Survey-2022.pdf>.
- Public Company Accounting Oversight Board. (2002). *Auditing Standard – AS 1001: Responsibilities and Functions of the Independent Auditor*. PCAOB. <https://pcaobus.org/oversight/standards/auditing-standards/details/AS1001>.
- Rakotomamonjy, A. (2003). Variable selection using SVM-based criteria. *Journal of Machine Learning Research*, 3, 1357–1370. [doi/10.5555/944919.944977](https://doi.org/10.5555/944919.944977).

- Ramos, M. V. M. (2011). Utilizando o ciclo operacional para fazer o orçamento de caixa e calcular a Necessidade de Capital de Giro (NCG). *Connexio*, 1(1), 81–90. <https://repositorio.unp.br/index.php/connexio/article/view/10>.
- Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, 50(2), 491–500. doi.org/10.1016/j.dss.2010.11.006.
- Reese Jr., W. A., & Robins, R. P. (2017). Performing an event study: An exercise for finance students. *The Journal of Economic Education*, 48(3), 206–215. [doi:10.1080/00220485.2017.1320603](https://doi.org/10.1080/00220485.2017.1320603).
- Reurink, A. (2018). Financial fraud: A literature review. *Journal of Economic Surveys*, 32(5), 1292–1325. doi.org/10.1111/joes.12294.
- Rezaee, Z. (2002). *Financial statement fraud: Prevention and detection*. John Wiley & Sons.
- Rezaee, Z. (2005). Causes, consequences, and deterrence of financial statement fraud. *Critical Perspectives on Accounting*, 16(3), 277–298. [doi.org/10.1016/S1045-2354\(03\)00072-8](https://doi.org/10.1016/S1045-2354(03)00072-8).
- Roohani, S., Furusho, Y., & Koizumi, M. (2009). XBRL: Improving transparency and monitoring functions of corporate governance. *International Journal of Disclosure and Governance*, 6(4), 355-369. [doi.:10.1057/jdg.2009.17](https://doi.org/10.1057/jdg.2009.17).
- RStudio Team. (2020). *R Studio: Integrated development for R*. RStudio. <http://www.rstudio.com/>.
- Saxena, R. (2017). *Support vector machine classifier implementation in R with the caret package*. Dataaspirant. <https://dataaspirant.com/support-vector-machine-classifier-implementation-r-caret-package/>.
- Segal, S. (2016). Accounting frauds – review of advanced technologies to detect and prevent frauds. *Economics and Business Review*, 2(4), 45–64. [doi:10.18559/ebr.2016.4.3](https://doi.org/10.18559/ebr.2016.4.3).
- Segrestin, B., & Hatchuel, A. (2011). Beyond agency theory, a post-crisis view of corporate law. *British Journal of Management*, 22(3), 484-499. doi.org/10.1111/j.1467-8551.2011.00763.x
- Severino, M. K., & Peng, Y. (2021). Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata. *Machine Learning with Applications*, 5, 100074. doi.org/10.1016/j.mlwa.2021.100074.
- Soares, R. R. (2004). Development, crime and punishment: accounting for the international differences in crime rates. *Journal of development Economics*, 73(1), 155-184. doi.org/10.1016/j.jdeveco.2002.12.001.
- Financial feature embedding with knowledge representation learning for financial statement fraud detection. *Procedia Computer Science*, 187, 420-425. doi.org/10.1016/j.procs.2021.04.110.
- Sharma, A., & Panigrahi, P. K. (2012). A review of financial accounting fraud detection based on data mining techniques. *International Journal of Computer Applications*. 39(1): 37 – 47. doi.org/10.48550/arxiv.1309.3944

- Silva Filho, G. M. D. (2019). *Fatores determinantes da intenção de denúncia ou do silêncio conivente da fraude contábil: Um estudo à luz das teorias do comportamento planejado e do triângulo de fraude*. [Tese de doutorado, Centro de Ciências Sociais e Aplicadas]. Repositório Institucional da Universidade Federal da Paraíba. <https://repositorio.ufpb.br/jspui/handle/123456789/19001>.
- Soares, R. O., Rostagno, L. M., & Soares, K. T. C. (2002). Estudo de evento: O método e as formas de cálculo do retorno anormal. *Anais do XXVI Encontro Nacional dos Programas de Pós Graduação em Administração*, 27, 30–33. http://www.anpad.org.br/admin/pdf/enanpad_2002-fin-1440.pdf.
- Song, X.-P., Hu, Z.-H., Du, J.-G., & Sheng, Z.-H. (2014). Application of machine learning methods to risk assessment of financial statement fraud: Evidence from China. *Journal of Forecasting*, 33(8), 611–626. [doi:10.1002/for.2294](https://doi.org/10.1002/for.2294).
- Sorescu, A., Warren, N. L., & Ertekin, L. (2017). Event study methodology in the marketing literature: An overview. *Journal of the Academy of Marketing Science*, 45(2), 186–207. [doi:10.1007/s11747-017-0516-y](https://doi.org/10.1007/s11747-017-0516-y).
- Sousa, N. C. (2016). *A importância da Auditoria na detecção e prevenção da fraude*. [Dissertação de mestrado, Instituto Superior de Contabilidade e Administração do Porto]. Repositório Científico do Instituto Politécnico do Porto. https://recipp.ipp.pt/bitstream/10400.22/9668/1/Natacha_Sousa_MA_2016%20Corrigido.pdf.
- Stein, A. (2012). Corrupt Intentions: Bribery, Unlawful Gratuity, and Honest-Services Fraud. *Law & Contemp. Probs.*, 75, 61. <https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1668&context=lcp>.
- Stiglitz, J. E. (2000). Capital market liberalization, economic growth, and instability. *World development*, 28(6), 1075-1086. [doi:10.1016/S0305-750X\(00\)00006-1](https://doi.org/10.1016/S0305-750X(00)00006-1).
- Suh, J. B., Shim, H. S., & Button, M. (2018). Exploring the impact of organizational investment on occupational fraud: Mediating effects of ethical culture and monitoring control. *International Journal of Law, Crime and Justice*, 53, 46-55. [doi:10.1016/j.ijlcrj.2018.02.003](https://doi.org/10.1016/j.ijlcrj.2018.02.003).
- Sutherland, E. H. (1945). Is “white collar crime” crime? *American Sociological Review*, 10(2), 132–139. <https://doi.org/10.2307/2085628>.
- Tay, L. M., Pua, C. H., Brahmana, R. K., & Malek, N. I. A. (2016). The effect of white collar crime announcement on stock price performance: Evidence from Malaysian stock market. *Journal of Financial Crime*, 23(4), 1126–1139. <https://doi.org/10.1108/JFC-03-2015-0016>.
- Tsai, Y.-C., & Huang, H.-W. (2020). Internal control material weakness opinions and the market’s reaction to securities fraud litigation announcements. *Finance Research Letters*, 41, 101833. [doi:10.1016/j.frl.2020.101833](https://doi.org/10.1016/j.frl.2020.101833).
- Vapnik, V. N. (2000). *The nature of statistical learning theory*. Springer.
- Vlad, M., Tulvinschi, M., & Chiriță, S. I. (2011). The consequences of fraudulent financial reporting. *The USV Annals of Economics and Public Administration*, 11(1(13)), 216–268. <http://annals.seap.usv.ro/index.php/annals/article/viewArticle/392>.

- Wang, Z., Yang, C., Oh, S.-K., & Fu, Z. (2018). Multi-radial basis function SVM classifier: Design and analysis. *Journal of Electrical Engineering and Technology*, 13(6), 2511–2520. doi.org/10.5370/JEET.2018.13.6.2511.
- Wickham, H. (2020). *rshape2*. R Studio. <https://cran.r-project.org/web/packages/reshape2/reshape2.pdf>.
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (1991). *ggplot2*. R Studio. <https://ggplot2.tidyverse.org/>
- Wilks, T. J., & Zimbelman, M. F. (2004). Decomposition of fraud-risk assessments and auditors' sensitivity to fraud cues. *Contemporary Accounting Research*, 21(3), 719–745. [doi:10.1506/hgxp-4dbh-59d1-3fhj](https://doi.org/10.1506/hgxp-4dbh-59d1-3fhj).
- Wuerges, A. F. E., & Borba, J. A. (2014). Fraudes contábeis: Uma estimativa da probabilidade de detecção. *Revista Brasileira de Gestão de Negócios*, 26(52), 466–483. doi.org/10.7819/rbgn.v16i52.1555.
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, 307-333. doi.org/10.2307/1912557.
- Yazici, Y. (2020). Approaches to fraud detection on credit card transactions using artificial intelligence methods. *Computer Science & Information Technology*, 235–244. [doi:10.5121/csit.2020.101018](https://doi.org/10.5121/csit.2020.101018).
- Yuwono, Y. P., & Marlina, M. A. E. (2021). Peran Fraud Triangle dalam Mendeteksi Financial Statement Fraud di Perusahaan Perbankan ASEAN. *E-Journal of Akuntansi*, 31(3), 713–730. <https://dspace.uc.ac.id/handle/123456789/3306>.
- Zahra, S. A., Priem, R. L., & Rasheed, A. A. (2007). Understanding the causes and effects of top management fraud. *Organizational Dynamics*, 36(2), 122-139. [doi:10.1016/j.orgdyn.2007.03.002](https://doi.org/10.1016/j.orgdyn.2007.03.002).
- Zhang, X.-D. (2020). Support vector machines. In X.-D. Zhang, *A matrix algebra approach to artificial intelligence* (pp. 617–679). Springer Singapore.

Apêndice

Apêndice 1 – Questões técnicas para desenvolver um estudo de eventos

De acordo com Mueller (2019), um desenvolvimento bem-sucedido de uma análise de estudo de eventos do programa R depende do tratamento dos dados e da forma de dispô-los. Isto é melhor apresentado na seção 4 desta dissertação.

A instalação e a ativação dos pacotes *usethis*, *devtools*, *ggplot2*, *reshape2* e *EventStudy* são necessárias para um desenvolvimento bem-sucedido da análise. O pacote *usethis* é um pacote de fluxo de trabalho que automatiza as tarefas repetitivas ou as tarefas que eram realizadas manualmente (Bryan et al., 2021). Segundo Hester e Wickham (2021), o pacote *devtools* possibilita o desenvolvimento de pacotes de modo mais fácil, fornecendo funções do programa R que simplificam e agilizam as tarefas comuns. Este pacote permite a combinação de várias funções em um mesmo argumento. Ele é útil na seleção de alguns vetores da base de dados primária e a sua transformação em um arquivo em *formatocomma separated values* (csv), por exemplo, o que permite a leitura no formato adequado pelo programa que executará a tarefa. De acordo com Wickham et al. (1991), o pacote *ggplot2* é um sistema para a criação de gráficos. Este pacote é necessário, pois ele é um prerequisite para o pacote principal, que executa o estudo de eventos (*EventStudy*). O pacote *reshape2* é também um prerequisite para o pacote *EventStudy*. Conforme Wickham (2020), o pacote *reshape2* remodela um quadro de dados entre o formato “largo” com medições repetidas em colunas separadas do mesmo registro, e o formato “longo” com as medições repetidas em registros separados. Por fim, o último pacote é o *EventStudy*; ele fornece funcionalidades para executar estudos de eventos a partir de R usando a interface *EventStudyTools.com* API, analisando resultados e possibilitando sua visualização (Mueller, 2019).

O Apêndice 2 apresenta os algoritmos para a instalação e a ativação dos pacotes necessários para executar o estudo de eventos. Ressalta-se que o pacote *readxl* é um pacote básico do R Studio; ele já vem instalado, mas precisa ser ativado. Sua ativação é importante, pois é comum que os pesquisadores importam seus dados para o programa R no formato Excel (xl). O referido pacote permite a leitura de arquivos nesse formato.

O Apêndice 2 também apresenta os algoritmos para ativar e autenticar a chave pessoal adquirida no *EventStudytools* (<https://www.eventstudytools.com/R-package>). Ressalta-se que, sem a ativação da chave, o pacote principal *EventStudy* não executa as atividades.

Antes de importar os dados no programa R, é importante se assegurar do conteúdo dos três arquivos. O primeiro arquivo deve conter nove colunas, trazendo informações como a identificação numérica do evento de cada título, a identificação nominal (código) de cada título,

a identificação nominal (código) do mercado, a data do evento, a forma de agrupamento de variáveis (*Addition* por padrão), o início da janela de eventos, o fim da janela de eventos, o fim da janela de estimação e o tamanho da janela de estimação.

O segundo arquivo deve conter apenas três colunas, trazendo informações sobre a identificação nominal (código) dos títulos, a data de observação e o preço de fechamento nesta data (para uma análise de retornos) ou o volume de negociação (para uma análise de liquidez). O terceiro arquivo traz as mesmas informações que o segundo, mas com relação ao título do mercado.

Depois de importar os dados, é importante convertê-los para o formato adequado (csv), a fim de permitir a leitura pelo *EventStudy*. O Apêndice 2 apresenta os algoritmos para garantir a leitura correta dos dados pela função *EventStudy*.

A função *EventStudy* vem com o retorno contínuo e o modelo de regressão de mercado como padrão. As especificidades devem ser ativadas. O Apêndice 2 apresenta os algoritmos da escolha dos parâmetros. Ressalta-se que, caso o pesquisador opte pelos parâmetros-padrão (que é o caso desta pesquisa), não há necessidade de ativá-los. Outros parâmetros podem ser escolhidos na página do pacote *EventStudy* (https://rdrr.io/cran/EventStudy/f/vignettes/parameters_eventstudy.Rmd).

Antes de se executar o estudo de eventos propriamente dito, é necessário juntar os três arquivos convertidos no quarto passo em um único arquivo. O Apêndice 2 apresenta os algoritmos. Por fim, o Apêndice 2 apresenta a execução propriamente dita do estudo de eventos. Antes disso, caso o pesquisador queira verificar se o arquivo único está no formato adequado para o sucesso da análise. Ressalta-se que este passo não é necessário, porém recomendável. O Apêndice 2 apresenta o algoritmo de verificação do arquivo único.

Apêndice 2 – Execução de um estudo de eventos no R e ativação de pacotes

```
#####ATIVAR OS PACOTES NECESSÁRIOS#####
library(readr)
library(readxl)
library(usethis)
library(devtools)
library(Rtools)
library(ggplot2)
library(EventStudy)
library(reshape2)

#####ESCOLHER UMA JANELA EM QUE OS DADOS SERÃO COLETADAS#####
#####DE 1982 ATE 2022#####

startDate <- "1982-01-01"
endDate <- "2022-06-01"

#####IMPORTAR O ARQUIVO DE EVENTOS: 01requestfile #####

X01_RequestFile_real_Addiction_01d <- read_excel("endereço da pasta/nome do arquivo
(RequestFile_real_Addiction_01d.xlsx)")
View(X01_RequestFile_real_Addiction_01d)

###TRANSFORMAR O ARQUIVO “X01_RequestFile_real_Addiction_01d” EM UM DOCUMENTO
CSV###

X01_RequestFile_real_Addiction_01d %>%
  as.data.frame() %>%
  readr::write_delim("01_RequestFile_real_Addiction_01d.csv", delim = ";", col_names = F)

#####IMPORTAR DE YAHOO FINANCE, OS DADOS DAS ENTIDADES: 02firmdata#####

firmSymbols <- c("BAHI3.SA", "BMYB34.SA", "BUET3.SA", "CBMA3.SA", "CEBR6.SA", "CGRA3.SA",
"CMIG3.SA", "DUQE3.SA", "EMAE3.SA", "EMBR3.SA", "GGBR3.SA", "HETA3.SA", "I1HG34.SA",
"IGUA3.SA", "IMBI3.SA", "INEP3.SA", "JBSS3.SA", "JHSF3.SA", "LIXC3.SA", "MNDL3.SA",
"MPLU3.SA", "MSPA3.SA", "OGXP3.SA", "OSXB3", "PDGR3.SA", "PETR3.SA", "PTNT3.SA",
"RCSL3.SA", "RJCP3.SA", "RSID3.SA", "SCLO3.SA", "SPRI3.SA", "SULT3.SA", "TASA3.SA",
"TENE3.SA")

firmNames <- c("BAHI3", "BMYB34", "BUET3", "CBMA3", "CEBR6", "CGRA3", "CMIG3", "DUQE3",
"EMAE3", "EMBR3", "GGBR3", "HETA3", "I1HG34", "IGUA3", "IMBI3", "INEP3", "JBSS3", "JHSF3",
"LIXC3", "MNDL3", "MPLU3", "MSPA3", "OGXP3", "OSXB3", "PDGR3", "PETR3", "PTNT3", "RCSL3",
"RJCP3", "RSID3", "SCLO3", "SPRI3", "SULT3", "TASA3", "TENE3")
firmSymbols %>%
  tidyquant::tq_get(from = startDate, to = endDate) %>%
  dplyr::mutate(date = format(date, "%d.%m.%Y")) -> firmData
knitr::kable(head(firmData), pad=0)

str(firmData)

#####SELECIONAR AS COLUNAS "symbol", "date" E "adjusted" E TRANSFORMAR EM CSV###

firmData %>%
  dplyr::select(symbol, date, adjusted) %>%
  readr::write_delim(path = "02_firmData.csv",
```

```

delim = ";",
col_names = F)

#####IMPORTAR DE YAHOO FINANCE, OS DADOS DO MERCADO: 03marketdata

marketSymbol <- c("^BVSP")
marketName <- c("BVSP")
marketSymbol %>%
  tidyquant::tq_get(from = startDate, to = endDate) %>%
  dplyr::mutate(date = format(date, "%d.%m.%Y")) -> marketData
marketData_real_5d$symbol <- "BVSP"
knitr::kable(head(marketData), pad=0)

str(marketData)

#####SELECIONAR AS COLUNAS "symbol", "date" E "adjusted" E TRANSFORMAR EM CSV###

marketData %>%
  dplyr::select(symbol, date, adjusted) %>%
  readr::write_delim(path = "03_marketData.csv",
                    delim = ";",
                    col_names = F)

#####ATIVAR A CHAVE EVENT STUDY

Key <- "c6c92bd2dbf486d1f1fa357e9cf2af5c"
est <- EventStudyAPI$new()
est$authentication(apiKey = Key)

#####ATIVAR OS PARAMETROS DA SVM

esaParams <- EventStudy::ARCAApplicationInput$new()
esaParams$resultFileType("xlsx")
esaParams$setBenchmarkModel("mm")
esaParams$setNonTradingDays("later")

#####JUNTAR OS DADOS EM UM ARQUIVO UNICO

dataFiles_real_Addition_01d <- c("request_file" = "01_RequestFile_real_Addiction_01d.csv",
                              "firm_data" = "02_firmData.csv",
                              "market_data" = "03_marketData.csv")

#####VERIFICAR O ARQUIVO UNICO

EventStudy::checkFiles(dataFiles_real_Addition_01d)

```



```
#####EXECUTAR O ESTUDO DE EVENTO  
  
arEventStudy <- est$performEventStudy(estParams = esaParams,  
                                     dataFiles = dataFiles_real_Addition_01d,  
                                     downloadFiles = T)
```

Fonte: elaboração própria por meio do R Studio.

Apêndice 3 – Execução da SVM linear no R

```

### DOWNLOAD E CARREGAMENTO DE PACOTES NECESSÁRIOS ###
library(readxl)
library(caret)

### IMPORTAÇÃO E VISUALIZAÇÃO DE DADOS ####
dados <- read_excel("caminho para o arquivo_nome do arquivo.xlsx")

### DIVISÃO DOS DADOS EM 2 SUBGRUPOS: "TRAINING" E TESTING" ###
intrain <- createDataPartition(y = nome do arquivo$"variável dependente", p= 0.66, list = FALSE)
training <- dadosr[intrain,]
testing <- dados[-intrain,]

dim(training); dim(testing)

training[["variável dependente"]]= factor(training[["variável dependente"]])
trctrl1 <- trainControl(method = "repeatedcv", number = 10, repeats = 3)

###TREINAMENTO DO MODELO svm (CLASSIFICAÇÃO LINEAR) COM O SUBGRUPO DA DA
AMOSTRA "TRAINING"###
svm_Linear <- train(variável dependente ~., data = training, method = "svmLinear", trControl=trctrl1,
preProcess = c("center", "scale"), tuneLength = 10)
svm_Linear

###APLICAÇÃO DO MODELO SVM (CLASSIFICAÇÃO LINEAR) COM O SUBGRUPO DA
AMOSTRA "TESTING"###
test_pred<- predict(svm_Linear1, newdata = testing)
test_pred

##### VISUALIZAÇÃO DOS RESULTADOS#####
previsao_realidade1<- table(test_pred, testing$variável dependente)
confusionMatrix(previsao_realidade)

```

Fonte: adaptado de Saxena (2017).

Apêndice 4 – Execução da SVM não linear no R

```

### DOWNLOAD E CARREGAMENTO DE PACOTES NECESSÁRIOS ###
library(readxl)
library(e1071)

### IMPORTAÇÃO E VISUALIZAÇÃO DE DADOS ####
dados <- read_excel("caminho para o arquivo_nome do arquivo.xlsx")

### DIVISÃO DOS DADOS EM 2 SUBGRUPOS: "TRAINING" E TESTING" ###
intrain <- createDataPartition(y = nome do arquivo$"variável dependente", p= 0.66, list = FALSE)
training <- dadosr[intrain,]
testing <- dados[-intrain,]

dim(training); dim(testing)

training[["variável dependente"]]= factor(training[["variável dependente"]])
trctrl1 <- trainControl(method = "repeatedcv", number = 10, repeats = 3)

###TREINAMENTO DO MODELO svm (CLASSIFICAÇÃO NÃO LINEAR) COM O SUBGRUPO DA
DA AMOSTRA "TRAINING"###
svm_Training_NãoLinear <- svm(variável dependente ~., data = training, scale=FALSE, kernel="radial",
cost=5)
svm_Training_NãoLinear

svmfit=svm(factor(y)~.,data=dat,scale=FALSE, kernel="radial",cost=5)

###APLICAÇÃO DO MODELO SVM (CLASSIFICAÇÃO NÃO LINEAR) COM O SUBGRUPO DA
AMOSTRA "TESTING"###
svm_Test_NãoLinear <- svm(variável dependente ~., data = testing, scale=FALSE, kernel="radial", cost=5)

##### VISUALIZAÇÃO DOS RESULTADOS#####
previsao_realidade1<- table(test_pred, testing$variável dependente)
confusionMatrix(previsao_realidade)

```

Fonte: elaboração própria por meio do R Studio.

Apêndice 5 – Processos julgados pela CVM com condenação de fraude/irregularidade (amostra final)

Link do processo	Data da decisão CVM	Data da fraude	Nome da companhia	Setor de atividade	Código da ação	Causa
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190924-PAS-SP2017-493.html	24/09/2019	31/12/2011	Bahema S.A.	Consumo cíclico	BAHI3	Omissão de informações sobre transações com partes relacionadas. Falhas informacionais nos formulários de referência e nas demonstrações financeiras. Infração ao art. 177, §3º, da Lei nº 6.404/1976, c/c o item 18 do Pronunciamento Técnico CPC 05 (R1). Infração aos arts. 14, 24 e 45 da Instrução CVM nº 480/2009.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20191210_PAS_CVM_RJ2016_8155_SEI_19957_007937_2016_119_voto_diretora_flavia_perlingeiro.pdf	10/12/2019	12/12/2013	Empreendimento Bristol Convention Hotel	Consumo cíclico	BYMB34 retirar	Oferta pública de CIC hoteleiro sem a obtenção de registro ou dispensa. Infração do art. 19 da Lei nº 6.385/1976 e do art. 2º da ICVM nº 400/2003, e sem a dispensa prevista no art. 19, §5º, I, da Lei nº 6.385/1976 e no art. 4º da ICVM nº 400/2003.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20191210_PAS_CVM_RJ2019_3127_SEI_19957_004986_2019_34_voto_presidente_marcelo_barbosa.pdf	10/12/2019	30/08/2016	Buettner SA Ind. e Comércio	Bens industriais	BUET3 retirar	Não envio das contas demonstrativas da administração e da prestação de contas da continuidade provisória da falência. Infração ao art. 39, II, da Instrução CVM nº 480/2009.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181030_PAS_RJ201513006.html	30/10/2018	31/12/2011	Cobrasma S.A.	Bens industriais	CBMA3 retirar	Inobservância de normas emanadas pelo Conselho Federal de Contabilidade e da regra de rodízio de auditores independentes, em infração aos arts. 20 e 31 da Instrução CVM nº 308/1999.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ20162159_Cobrasma.pdf	12/12/2017	31/12/2014	Cobrasma S.A.	Bens industriais	CBMA3	Irregularidades contábeis na elaboração das demonstrações financeiras da companhia. Abstenção de opinião do auditor independente. Descumprimento dos deveres de diligência e de fiscalização. Subestimação de provisões.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ20162159_Cobrasma.pdf	12/12/2017	31/12/2015	Cobrasma S.A.	Bens industriais	CBMA3	Irregularidades contábeis na elaboração das demonstrações financeiras da companhia. Abstenção de opinião do auditor independente. Descumprimento dos deveres de diligência e de fiscalização.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ20162159_Cobrasma.pdf	12/12/2017	31/12/2014	Cobrasma S.A.	Bens industriais	CBMA3	Determinação de provisão em desacordo com as normas.

http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2021/SEI_199570092_92_2017_21.pdf	02/03/2021	31/12/2012	Companhia Energética de Brasília	Utilidade pública	CEBR6	Não constituição de provisão face à inadimplência de clientes, que resultou em uma superestimação de ativos (créditos a receber). Responsabilidade de membros do conselho de administração e do conselho fiscal da Companhia Energética de Brasília (CEB) por infrações relativas às demonstrações financeiras dos exercícios de 2011, 2012 e 2013, em razão do descumprimento de normas contábeis dos Pronunciamentos Técnicos CPC 38 e CPC 40, c/c o art. 176, §5º, III, da Lei nº 6.404/1976, quanto a créditos inadimplidos detidos em face do acionista controlador.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2021/SEI_199570092_92_2017_21.pdf	02/03/2021	31/12/2013	Companhia Energética de Brasília	Utilidade pública	CEBR6	Não constituição de provisão face à inadimplência de clientes, que resultou em uma superestimação de ativos (créditos a receber). Responsabilidade de membros do conselho de administração e do conselho fiscal da Companhia Energética de Brasília (CEB) por infrações relativas às demonstrações financeiras dos exercícios de 2011, 2012 e 2013, em razão do descumprimento de normas contábeis dos Pronunciamentos Técnicos CPC 38 e CPC 40, c/c o art. 176, §5º, III, da Lei nº 6.404/1976, quanto a créditos inadimplidos detidos em face do acionista controlador.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2021/SEI_199570092_92_2017_21.pdf	02/03/2021	31/12/2011	Companhia Energética de Brasília	Utilidade pública	CEBR6	Não constituição de provisão face à inadimplência de clientes, que resultou em uma superestimação de ativos (créditos a receber). Responsabilidade de membros do conselho de administração e do conselho fiscal da Companhia Energética de Brasília (CEB) por infrações relativas às demonstrações financeiras dos exercícios de 2011, 2012 e 2013, em razão do descumprimento de normas contábeis dos Pronunciamentos Técnicos CPC 38 e CPC 40, c/c o art. 176, §5º, III, da Lei nº 6.404/1976, quanto a créditos inadimplidos detidos em face do acionista controlador.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20190326_PAS_072014.html	26/03/2019	20/04/2009	Duke Energy International Geração Paranapanema S.A.	Utilidade pública	CEBR6	Descumprimento do dever de diligência no processo de fixação da remuneração dos administradores da Companhia Duke Energy International Geração Paranapanema S.A. no ano de 2009. Infração aos arts. 152 e 153, ambos da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181023_PAS_RJ20168769.html	23/10/2018	31/12/2014	Grazziotin S.A.	Consumo cíclico	CGRA3	Não consolidação de controlados nas demonstrações. Inobservância ao item 24 do Pronunciamento Técnico CPC 19 (R2) e ao item 1.b da Interpretação Técnica ICPC 09 (R2), incorrendo no descumprimento do art. 4º da Instrução CVM nº 457/2007, do art. 177, §§ 3º e 5º, da Lei nº 6.404/1976, do item 6.a da NBC TA 705, dos itens 2 e 11 da NBC TA 200 e do art. 20 da Instrução CVM nº 308/1999.

http://www.cvm.gov.br/sancionadores/sancionador/2018/20181218_PAS_RJ2017_2553.html	18/12/2018	29/04/2016	Companhia energética de Minas Gerais CEMIG	Utilidade pública	CMIG3	Permissão, como presidente das assembleias gerais ordinária e extraordinária, da eleição pelos acionistas minoritários com direito a voto de conselheiro de administração em inobservância ao <i>quorum</i> legal de 15% do total das ações com direito a voto, de que trata o art. 141, §4º, I, c/c art. 128, ambos da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20190219_PAS_16-2013.html	19/02/2019	31/09/2002	Docas de Imbituba S.A	Bens industriais	DOCA3	Prática de atos negociais desvantajosos para a Companhia Docas de Imbituba S.A (CDI) por seu controlador. Descumprimento do dever de diligência. Ausência de livro de ata das reuniões da diretoria. Infração aos arts. 117, §1º, “a”, 153 e 100, VI, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181218_PAS_RJ2016_5649.html	18/12/2018	31/12/2014	Companhia Docas de Imbituba S.A	Bens industriais	DOCA3	Elaboração e divulgação de demonstrações financeiras da Companhia Docas de Imbituba S.A. em desacordo com os pronunciamentos técnicos do Comitê de Pronunciamentos Contábeis (CPC). Não fornecimento de elementos e condições necessários para o desempenho das funções do auditor independente. Embaraço à fiscalização. Infração aos arts. 176 e 177, §§ 3º e 5º, e art. 142, III e V, da Lei nº 6.404/1976. Infração ao art. 14 da Instrução CVM nº 480/2009. Infração ao art. 26 da Instrução CVM nº 308/1999 e incidência do art. 1º, III, da Instrução CVM nº 491/2011.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20190226_PAS_RJ2014-4077.html	26/02/2019	31/10/2010	Metalúrgica Duque S/A	Bens industriais	DUQE3	Exercício abusivo do poder de controle, desvio de poder, irregularidades nas demonstrações financeiras divulgadas pela Metalúrgica Duque S/A, atuação do administrador da companhia em situação de conflito de interesses e contra os interesses sociais da empresa e negligência na fiscalização da gestão dos diretores por membros do conselho de administração da companhia. Infrações contábeis nos documentos divulgados. Infração ao art. 117, §1º, “a” e “f”, da Lei nº 6.404/1976. Infração ao art. 154, <i>caput</i> e §2º, “b”, da Lei nº 6.404/1976. Infração ao art. 177, <i>caput</i> e §3º, c/c o art. 176, <i>caput</i> , da Lei nº 6.404/1976 e aos arts. 26, I, e 29, I, da Instrução CVM nº 480/2009. Infração ao art. 156, <i>caput</i> , da Lei nº 6.404/1976. Suposta infração ao art. 142, III e IV, c/c art. 153 da Lei 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20180712_PAS_20167808.html	12/07/2018	31/12/2013	Metalúrgica Duque S.A	Bens industriais	DUQE3	Não elaboração e não envio à CVM das informações periódicas obrigatórias – não convocação de assembleias gerais ordinárias –, não elaboração e não submissão à revisão de auditores independentes dos Formulários de Informações Trimestrais (ITRs), não elaboração de demonstrações financeiras.

http://www.cvm.gov.br/sancionadores/sancionador/2019/20190129_PAS_RJ2016-802.html	29/01/2019	08/04/1998	EMAE – Empresa Metropolitana de Águas e Energia	Utilidade pública	EMAE3	Responsabilidade do Estado de São Paulo, na qualidade de acionista controlador da Empresa Metropolitana de Águas e Energia (EMAE), pela suposta utilização, a título gratuito, de serviços de controle de cheias prestados pela companhia. Infração ao art. 116, § único, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20180911_PAS_RJ20151760.html	11/09/2018	01/01/2009	Embraer S.A	Bens industriais	EMBR3	Apuração da responsabilidade de administradores da Embraer S.A por descumprimento do dever de atuar no interesse social da companhia. Infração ao art. 154, <i>caput</i> , da Lei nº 6.404/1976. Inabilitação temporária.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20191203_PAS_CVM_SEI_1995_7004309_2016_73_voto_diretora_flavia_perlingeiro.pdf	03/12/2019	14/07/2015	Gerdau S.A.	Materiais básicos	GGBR3	Irregularidades em operação de aquisição de participações minoritárias em empresas controladas. Infração aos arts. 116, § único, 154, c/c 245, 156 e 157, §4º, todos da Lei nº 6.404/1976, art. 3º, §5º, da Instrução CVM nº 358/2002 e art. 14 da Instrução CVM nº 480/2009.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ2017565_Mundial_SA.pdf	14/12/2017	31/12/2013	Hercules S.A.	Consumo cíclico	HETA3	Emissão, pela Hercules S.A., de debêntures e subscrição pela Mundial S.A em condições não comutativas, vistas as duas entidades estarem em mútuo.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181226_PAS_RJ2017_4780.html	26/12/2018	12/12/2013	Intercontinental Hotels Group do Brasil Ltda e seu administrador Francisco Cesar Garcia Diez	Consumo cíclico	IHG34	Oferta pública irregular de contratos de investimento coletivo hoteleiro referentes ao empreendimento hoteleiro Hotel Holiday Inn Belo Horizonte, sem a obtenção de registro e sem a dispensa prévia prevista em lei. Infração ao disposto no art. 19 da Lei nº 6.385/1976 e no art. 2º da Instrução CVM nº 400/2003.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2020/20200630-PAS-19957.001575_2020_21.html	30/06/2020	31/12/2012	Companhia Iguçu de Café Solúvel	Consumo não cíclico	IGUA3	Indícios de fraude envolvendo uma das controladas. Descumprimento do disposto nos arts. 153 e 142, III e V, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20190402-PAS-RJ2017-6169.html	02/04/2019	31/12/2015	Companhia Docas de Imbituba	#N/A	IMBI3	Não elaboração de demonstrações financeiras da Companhia Docas de Imbituba. Não realização de assembleia geral ordinária. Não envio de informações periódicas à CVM. Infração ao art. 132 e ao art. 142, IV, da Lei nº 6.404/1976. Infração ao art. 176, <i>caput</i> , da Lei nº 6.404/1976. Infração ao art. 21, I, e ao art. 23, § único, da Instrução CVM nº 480/2009. Infração ao art. 21, V, e ao art. 29, II, da Instrução CVM nº 480/2009.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20191202_PAS_CVM_SEI_1995	03/12/2019	07/05/2018	INEPAR	Bens industriais	INEP3	Não divulgação de fato relevante. Infração ao art. 3º, c/c art. 6º, § único, ambos da Instrução CVM nº 358/2002, e art. 157, §4º, da Lei nº 6.404/1976.

7009116_2018_71_relatorio_presidente_marcelo_barbosa.pdf						
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181218_PAS_RJ2014_7351.html	18/12/2018	31/12/2013	Inepar S.A. Indústrias e Construções	Bens industriais	INEP3	Elaboração das demonstrações financeiras da Inepar Equipamentos e Montagens S.A., relativas ao exercício social encerrado em 31/12/2013 e ao período trimestral encerrado em 31/03/2014, em desacordo com o disposto nos itens 4.15, 4.17, QC 26 e QC 28 do Pronunciamento Técnico CPC 00, aprovado por meio da Deliberação CVM nº 675/2011. Não manutenção de conselho de administração instalado em sua subsidiária integral, a Inepar Equipamentos e Montagens S.A. Infração ao art. 177, <i>caput</i> , e ao §3º, <i>c/c</i> o <i>caput</i> do art.176 da Lei nº 6.404/1976, bem como ao art. 26, I, e ao art. 29, I, ambos da Instrução CVM nº 480/2009. Infração ao disposto no art. 138, §2º, da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/sancionadores/sancionador/2020/20200721-PAS-19957.010904_2018_18.html	21/07/2020	31/12/2012	JBS S/A	Consumo não cíclico	JBSS3	Responsabilidade de Wesley Batista e Joesley Batista pelo descumprimento do art. 154, §2º, “b”, da Lei nº 6.404/1976 e, especificamente no caso de Wesley Batista, também por violação ao art. 153 da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/sancionadores/sancionador/2020/20200721-PAS-19957.010904_2018_18.html	21/07/2020	11/05/2017	JBS S/A	Consumo não cíclico	JBSS3	Responsabilidade de Wesley Batista e Joesley Batista pelo descumprimento do art. 154, §2º, “b”, da Lei nº 6.404/1976 e, especificamente no caso de Wesley Batista, também por violação ao art. 153 da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2020/20200720_PAS_CVM_SEI_19957_010904_2018_18_relatorio_presidente_marcelo_barbosa.pdf	21/07/2020	26/05/2017	JBS S/A	Consumo não cíclico	JBSS3	Uso dos ativos da companhia para fins pessoais. Infração do art. 153 e ao art. 154, §2º, “b”, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181204_PAS_RJ2017-863.html	12/04/2018	10/11/2016	JHSF Participações S.A	Consumo cíclico	JHSF3	Uso indevido de informação privilegiada. Infração ao art. 155, §4º, da Lei nº 6.404/1976, <i>c/c</i> o art. 13, §1º, da Instrução CVM nº 358/2002.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190827-PAS-RJ2018-3372.html	27/08/2019	31/12/2016	Lix da Cunha S.A.	Consumo cíclico	LIXC3	Não convocação da assembleia geral ordinária e não realização de auditoria. Infração ao art. 132, <i>c/c</i> o art. 142, IV, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181030_PAS_19957010233_2018_87.html	30/10/2018	31/12/2001	Construtora Lix da Cunha S.A	Consumo cíclico	LIXC3	Destinação de dividendos às ações ordinárias de emissão da Construtora Lix da Cunha S.A. inferiores aos que fariam jus nos exercícios sociais de 2001 e de 2002. Pagamento dos dividendos relativos ao exercício social de 2001 fora do prazo legal. Constituição de reserva especial para dividendos não distribuídos

						em relação aos exercícios sociais de 2002, 2003, 2005, 2006 e 2007. Cálculo equivocado dos dividendos mínimos relativos aos exercícios sociais de 2005, 2006 e 2007. Não comunicação à CVM de justificativa para o não pagamento do dividendo obrigatório relativo ao exercício de 2003. Registro da destinação do lucro líquido do exercício nas demonstrações financeiras relativas a 2001, 2002, 2003 e 2005. Não absorção dos prejuízos do exercício social de 2004 pelas reservas de lucro. Atraso na elaboração das demonstrações financeiras da companhia relativas aos exercícios sociais de 2004 e 2005 e atraso na convocação e na realização de assembleias gerais ordinárias relativas aos exercícios sociais de 2004 e de 2005. Preliminar de extinção de punibilidade.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ2017565_Mundial_SA.pdf	14/12/2017	31/12/2013	Mundial S.A	Consumo cíclico	MNDL3	Emissão, pela Hercules S.A., de debêntures e subscrição pela Mundial S.A em condições não comutativas, vistas as duas entidades estarem em mútuo.
https://www.gov.br/cvm/pt-br/assuntos/noticias/anexos/2019/20191209_PAS_CVM_RJ2016_8155_SEI_19957_007937_2016_19_relatorio_diretora_flavia_perlingeiro.pdf-fa279689eabc4a60a697832874bdb00b	10/12/2019	31/12/2017	Multiplus S.A.	Consumo cíclico	MPLU3	Irregularidades praticadas na oferta pública primária de ações. Recebimento e tratamento de reservas em desacordo com o que foi previsto na oferta.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20181211_PAS_RJ2016-4426.html	11/12/2018	31/12/2014	Companhia Melhoramentos de São Paulo	Materiais básicos	MSPA3	Apuração da responsabilidade de administradores da Companhia Melhoramentos de São Paulo por terem recebido remuneração em montante superior ao autorizado em assembleia e contrariamente aos critérios legais no período de 2010 a 2014. Infração ao disposto no art. 152, c/c o 154 da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20181211_PAS_RJ2017-3871.html	11/12/2018	31/12/2017	Companhia Melhoramentos de São Paulo	Materiais básicos	MSPA3	Apuração da responsabilidade de administradores da Companhia Melhoramentos de São Paulo por terem recebido remuneração contrariamente aos critérios legais no período de 2015 a 2017. Infração ao disposto no art. 152 da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190625_PAS_RJ2014-12838.html	25/06/2019	30/09/2012	OGX Petróleo e Gás Participações S.A	Petróleo, gás e biocombustíveis	OGXP3	Divulgação de fatos relevantes como instrumento de manipulação de preços da OGX Petróleo e Gás Participações S.A. Divulgação de fato relevante com omissão material, induzindo investidores a erro. Infração ao inciso I, na forma prevista pelo item II, “b”, da Instrução CVM nº 08/1979. Infração ao art. 153 da Lei nº 6.404/1976, c/c o art. 14 da Instrução CVM nº 480/2009.

https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190625_PAS_RJ2014-12838.html	25/06/2019	30/09/2012	OGX Petróleo e Gás Participações S.A.	Petróleo, gás e biocombustíveis	OGXP3	Omissão de informações relevantes para a compreensão da situação financeira e patrimonial da OGX Petróleo e Gás Participações S.A. por ex-administradores da companhia. Infração aos arts. 142, III, 153, 176, c/c o 177, §3º, da Lei nº 6.404/1976 e aos arts. 14, 26 e 29 da Instrução CVM nº 480/2009.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190527_PAS_RJ2014-0578.html	27/05/2019	31/07/2013	OGX Petróleo e Gás Participações S.A. e da OSX Brasil S.A	Petróleo, gás e biocombustíveis	OGXP3	Utilização de informações privilegiadas na negociação de ações de emissão da OGX Petróleo e Gás Participações S.A. e da OSX Brasil S.A. Manipulação de preços de ações de emissão das companhias. Infração ao art. 155, §1º, da Lei nº 6.404/1976, c/c art. 13, <i>caput</i> , da Instrução CVM nº 538/2002. Infração ao item II, "b", vedada pelo item I da Instrução CVM nº 08/1979.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/SP2013094_Iwa_o_Jouti_2.pdf	14/12/2017	01/02/2012	OGX	Petróleo, gás e biocombustíveis	OGXP3	Negociação de valores mobiliários de posse de informações relevantes ainda não divulgadas ao mercado.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190625_PAS_RJ2014-12838.html	25/06/2019	24/09/2012	OSX Brasil S.A	Petróleo, gás e biocombustíveis	OSXB3	Não adoção de providências por ex-administradores da OSX Brasil S.A. para que as demonstrações financeiras evidenciassem informações relevantes para compreensão de sua situação financeira e patrimonial. Infração aos arts.142, III e V, e 153 da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2019/20181113_PAS_RJ2014-3616.h	13/11/2018	28/05/2012	PDG Realty S.A. Empreendimentos e Participações	Consumo cíclico	PDGR3	Negociação de valores mobiliários de emissão da PDG Realty S.A. Empreendimentos e Participações de posse de informações relevantes sobre a sociedade antes de sua divulgação ao mercado.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_19957005789_2017_71.pdf	03/11/2020	31/12/2014	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento contábil da redução ao valor recuperável de ativos (<i>impairment</i>). Arts. 153, 176 e 177, §3º, da Lei nº 6.404/1976 e arts. 14 e 26 da ICVM nº 480/2009; para membros do conselho de administração: arts. 142, III e V, e 153, e também art. 160, da Lei nº 6.404/1976 para integrantes do comitê de auditoria; e para membros do conselho fiscal, art. 163, VII, e 165 da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_19957005789_2017_71.pdf	03/11/2020	31/12/2013	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento contábil da redução ao valor recuperável de ativos (<i>impairment</i>). Arts. 153, 176 e 177, §3º, da Lei nº 6.404/1976 e arts. 14 e 26 da ICVM nº 480/2009; para membros do conselho de administração: arts. 142, III e V, e 153, e também art. 160, da Lei nº 6.404/1976, para integrantes do comitê de auditoria; e para membros do conselho fiscal, art. 163, VII, e 165 da Lei nº 6.404/1976.

http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570057_89_2017_71.pdf	03/11/2020	31/12/2012	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento contábil da redução ao valor recuperável de ativos (<i>impairment</i>). Arts. 153, 176 e 177, §3º, da Lei nº 6.404/1976 e arts. 14 e 26 da ICVM nº 480/2009; para membros do conselho de administração: arts. 142, III e V, e 153, e também art. 160, da Lei nº 6.404/1976, para integrantes do comitê de auditoria; e para membros do conselho fiscal, art. 163, VII, e 165 da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570057_89_2017_71.pdf	03/11/2020	31/12/2011	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento contábil da redução ao valor recuperável de ativos (<i>impairment</i>). Arts. 153, 176 e 177, §3º, da Lei nº 6.404/1976 e arts. 14 e 26 da ICVM nº 480/2009; para membros do conselho de administração: arts. 142, III e V, e 153, e também art. 160, da Lei nº 6.404/1976, para integrantes do comitê de auditoria; e para membros do conselho fiscal, art. 163, VII, e 165 da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570057_89_2017_71.pdf	03/11/2020	31/12/2010	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento contábil da redução ao valor recuperável de ativos (<i>impairment</i>). Arts. 153, 176 e 177, §3º, da Lei nº 6.404/1976 e arts. 14 e 26 da ICVM nº 480/2009; membros do conselho de administração: art. 142, III e V, e 153, e também art. 160, da Lei nº 6.404/1976, para integrantes do comitê de auditoria; e para membros do conselho fiscal, art. 163, VII, e 165 da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/PAS-09-2016.html	16/12/2019	31/12/2005	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento dos deveres fiduciários em contratações por troca de vantagens. Infração ao art. 155, <i>caput</i> , da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/PAS-09-2016.html	16/12/2019	31/12/2008	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Descumprimento dos deveres fiduciários em contratações por troca de vantagens. Infração ao art. 155, <i>caput</i> , da Lei nº 6.404/77.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181213_PAS_RJ2014_3402.html	13/12/2018	05/09/2013	Petróleo Brasileiro S.A.	Petróleo, gás e biocombustíveis	PETR3	Não divulgação de fato relevante sobre nova metodologia de precificação de combustíveis em 05/09/2013 pelo diretor de relações com investidores da Petróleo Brasileiro S.A. Infração ao art. 6º, § único, da Instrução CVM nº 358/2002, c/c o art. 157, §4º, da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/sancionadores/sancionador/2019/20190813-PAS-RJ2018-4328.html	13/08/2019	30/06/2012	Pettenati S.A.	Consumo cíclico	PTNT3	Retenção indevida de lucro. Descumprimento do §1º do art. 115 e do §1º do art. 134, ambos da Lei nº 6.404/1976.

http://www.cvm.gov.br/sancionadores/sancionador/2018/20181204_PAS_RJ2014-2426-.html	13/04/2018	29/10/2009	Pettenati S/A Indústria Têxtil	Consumo cíclico	PTNT3	Destinação de resultados para reserva de lucros supostamente irregular por membros conselho de administração e pelo acionista controlador da Pettenati S/A Indústria Têxtil. Não manifestação contrária à constituição e manutenção de reserva de retenção de lucros supostamente irregular por membros do conselho fiscal da Pettenati S/A Indústria Têxtil. Infração aos arts. 116, § único, 153, 196 e 163, III e IV, da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570075_52_2016_43.pdf	29/09/2020	31/12/2015	Recrusul S.A.	Bens industriais	RCSL3	Infração aos arts. 156 (“Proibição de investimentos de administradores de uma companhia em atividades de interesse ou que conflitam com as atividades da companhia”) e 157 (“Divulgação de todos os fatos relevantes”), §4º, ambos da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570075_52_2016_43.pdf	29/09/2020	31/12/2014	Recrusul S.A.	Bens industriais	RCSL3	Infração aos arts. 156 (“Proibição de investimentos de administradores de uma companhia em atividades de interesse ou que conflitam com as atividades da companhia”) e 157 (“Divulgação de todos os fatos relevantes”), §4º, ambos da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570075_52_2016_43.pdf	29/09/2020	31/12/2013	Recrusul S.A.	Bens industriais	RCSL3	Infração aos arts. 156 (“Proibição de investimentos de administradores de uma companhia em atividades de interesse ou que conflitam com as atividades da companhia”) e 157 (“Divulgação de todos os fatos relevantes”), §4º, ambos da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570075_52_2016_43.pdf	29/09/2020	31/12/2012	Recrusul S.A.	Bens industriais	RCSL3	Infração aos arts. 156 (“Proibição de investimentos de administradores de uma companhia em atividades de interesse ou que conflitam com as atividades da companhia”) e 157 (“Divulgação de todos os fatos relevantes”), §4º, ambos da Lei nº 6.404/1976.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2020/SEI_199570075_52_2016_43.pdf	29/09/2020	31/12/2011	Recrusul S.A.	Bens industriais	RCSL3	Infração aos arts. 156 (“Proibição de investimentos de administradores de uma companhia em atividades de interesse ou que conflitam com as atividades da companhia”) e 157 (“Divulgação de todos os fatos relevantes”), §4º, ambos da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20190730_PAS_CVM_RJ2016_2245_voto_diretor_carlos_rebello.pdf	30/07/2019	31/12/2013	Recrusul S.A.	Bens industriais	RCSL3	Reconhecimento de crédito fiscal sem evidências. Ausência de constituição de provisão para clientes duvidosos.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20190611_PAS_CVM_RJ2013_8	11/06/2019	30/06/2009	RJCP Equity S.A	Consumo cíclico	RJCP3	Manipulação de preços das ações da RJCP Equity S.A. e descumprimento do dever de diligência. Infração definida pelo inciso II, “b”, e vedada pelo inciso I, ambos da Instrução CVM nº

880_voto_diretor_henrique_machado.pdf						08/1979. Infração ao art. 153, c/c o art. 142, III, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181204_PAS_RJ2014-14161.html	04/12/2018	13/01/2014	RJCP Equity S.A	Consumo cíclico	RJCP3	Obstrução dos trabalhos do conselho fiscal e omissão na divulgação de fato relevante da RJCP Equity S.A. Infração ao art. 154 da Lei nº 6.404/1976. Infração ao art. 3º da Instrução CVM nº 358/2002, c/c o art. 157, §4º, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181030_PAS_RJ20165160.html	30/10/2018	10/09/2014	RJ Capital Partners S.A	Consumo cíclico	RJCP3	Não convocação e não realização de assembleia geral destinada à aprovação de bens utilizados em integralização de aumento de capital da RJ Capital Partners S.A, em infração aos arts. 8º e 170, §3º, c/c o art. 142, IV, da Lei nº 6.404/1976.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20190820_PAS_CVM_RJ2015_13127_voto_diretora_flavia_perlingeiro.pdf	20/08/2019	30/06/2010	Rossi Residencial S.A.	Consumo cíclico	RSID3	Antecipação de receitas.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ201412056_Cia_Industria_I_Schlosser.pdf	12/12/2017	31/12/2012	Companhia Indústria I Schlosser S.A.	#N/A	SCLO3	Entidade não apresentou o teste de recuperabilidade tendo a recuperação judicial aprovada. Elaboração de relatórios sem constar as depreciações sobre ativos.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ201412056_Cia_Industria_I_Schlosser.pdf	12/12/2017	31/12/2013	Companhia Indústria I Schlosser S.A.	#N/A	SCLO3	Realização de inventário na ausência de auditores, estando em recuperação judicial.
http://conteudo.cvm.gov.br/export/sites/cvm/sancionadores/sancionador/anexos/2017/RJ201412056_Cia_Industrial_Schlosser.pdf	12/12/2017	31/12/2014	Companhia Indústria I Schlosser S.A.	#N/A	SCLO3	Realização de inventário na ausência de auditores, estando em recuperação judicial.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2020/20200706_PAS_CVM_SEI_19957_010686_2017_22_relatorio_diretor_gustavo_gonzalez.pdf	07/07/2020	31/03/2014	Springer S.A	Consumo cíclico	SPRI3	Irregularidades em transações entre partes relacionadas. Infração aos arts. 154, <i>caput</i> , e 177, §3º, da Lei nº 6.404/1976, e aos arts. 14 e 24 da Instrução CVM nº 480/2009.
https://conteudo.cvm.gov.br/export/sites/cvm/noticias/anexos/2019/20191210_PAS_CVM_RJ2016_7162_SEI_19957_005934_2016_32_voto_diretora_flavia_perlingeiro.pdf	10/12/2019	18/08/2015	Construtora Sultepa S.A	Consumo cíclico	SULT3	Omissão na adoção de medidas para atender o regulamento, não divulgação ao mercado de valores mobiliários da suspensão de negociação das ações da companhia. Infração aos arts. 153 e 157, §4º, da Lei nº 6.404/1976, c/c art. 3º, <i>caput</i> , da ICVM nº 358/2002.

http://www.cvm.gov.br/sancionadores/sancionador/2019/20190226_PAS_16-2010.html	26/02/2019	31/12/2001	Sultepa Participações S.A	Consumo cíclico	SULT3	Irregularidades na contratação e na execução de mútuos com partes relacionadas e na forma de contabilização de créditos detidos nas demonstrações financeiras Exercício abusivo do poder de controle por parte dos administradores da Construtora Sultepa S.A. Omissão por parte dos administradores na proteção dos direitos e interesses da companhia. Descumprimento do dever de lealdade. Infração ao art. 245 da Lei nº 6.404/1976. Infração ao art. 117, §1º, “F”, da Lei nº 6.404/1976. Infração ao art. 176, <i>caput</i> , c/c o art. 179, <i>caput</i> e incisos I e II, da Lei nº 6.404/1976. Infração ao art. 155, II, da Lei nº 6.404/1976.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20180713_PAS_RJ20155813.html	13/07/2018	16/07/2013	Forjas Taurus S.A	Bens industriais	TASA3	Negociação com valores mobiliários de posse de informações relevantes ainda não divulgadas ao mercado.
http://www.cvm.gov.br/sancionadores/sancionador/2018/20181226_PAS_RJ2016-8900.html	26/12/2018	31/03/2016	TECBLU – Tecelagem Blumenau S/A	Utilidade pública	TENE3	Inobservância da regra de rotatividade de auditores, prevista no art. 31 da Instrução CVM nº 308/1999.

Fonte: elaboração própria.

Apêndice 6 – Retornos anormais individuais por evento 10 dias depois da realização da fraude

Ação das firmas	AR (1)	Valor-t (1)	AR (2)	Valor-t (2)	AR (3)	Valor-t (3)	AR (4)	Valor-t (4)	AR (5)	Valor-t (5)
Antes da convergência contábil (até 2009)										
CEBR6.SA	-0,0032	-0,1468	-0,0042	-0,1927	-0,0017	-0,078	-0,002	-0,0917	-0,0031	-0,1422
EMBR3.SA	-0,0235	-10981	-0,0075	-0,3505	-0,008	-0,3738	0,0024	0,1121	0,0099	0,4626
PETR3.SA	0,0217	13478	-0,0025	-0,1553	0,0027	0,1677	-0,0041	-0,2547	0,0192	11925
PETR3.SA	0,0058	0,2522	-0,0126	-0,5478	-0,0046	-0,2	-0,0134	-0,5826	0,0201	0,8739
PTNT3.SA	0,0054	0,0891	0,0845	13944	0,0099	0,1634	0,009	0,1485	0,0056	0,0924
Depois da convergência contábil (a partir de 2010)										
BAHI3.SA	-0,0049	-0,2121	0,0007	0,0303	0,0028	0,1212	0,0008	0,0346	-0,0001	-0,0043
BMYB34.SA	-0,0028	-0,1687	-0,0009	-0,0542	-0,0016	-0,0964	-0,0022	-0,1325	-0,0014	-0,0843
CBMA3.SA	0,1018	1.7891*	0,0019	0,0334	0,0086	0,1511	0,1076	1.891*	-0,0009	-0,0158
CBMA3.SA	-0,0032	-0,0514	-0,0015	-0,0241	0,0037	0,0595	0,0071	0,1141	0,0036	0,0579
CBMA3.SA	-0,0079	-0,0611	0,0024	0,0186	-0,0041	-0,0317	-0,0073	-0,0565	-0,0001	-0,0008
CEBR6.SA	0,002	0,0985	0,0019	0,0936	0,002	0,0985	0,002	0,0985	0,002	0,0985
CEBR6.SA	-0,0237	-0,75	-0,0108	-0,3418	0,0116	0,3671	0,0094	0,2975	0,0127	0,4019
CEBR6.SA	-0,0036	-0,0914	0,0015	0,0381	0,0034	0,0863	0,0016	0,0406	0,0007	0,0178
CGRA3.SA	0,0041	0,2547	-0,0442	-2.74***	-0,0317	-1.969*	0,0044	0,2733	0,0015	0,0932
CMIG3.SA	0,0398	1.7848*	-0,0316	-1,417	-0,0187	-0,8386	-0,0011	-0,0493	-0,0078	-0,3498
GGBR3.SA	-0,0093	-0,3153	-0,0467	-15,831	-0,0749	-2.54**	0,0009	0,0305	0,0129	0,4373
IMBI3.SA	0,002	0,0752	-0,1301	-4.89***	0,0015	0,0564	0,0019	0,0714	0,0009	0,0338
INEP3.SA	-0,0199	-0,3144	0,0172	0,2717	0,0296	0,4676	-0,0493	-0,7788	-0,0036	-0,0569
INEP3.SA	-0,0837	-2.3058**	0,051	1,405	-0,0501	-13,802	0,0053	0,146	0,0014	0,0386
JBSS3.SA	-0,0084	-0,3158	-0,0042	-0,1579	0,0477	1.7932*	-0,0312	-11,729	0,0182	0,6842
JBSS3.SA	-0,0213	-0,4712	-0,0058	-0,1283	0,0039	0,0863	-0,0004	-0,0088	-0,0242	-0,5354
JHSF3.SA	0,0315	0,8678	0,0073	0,2011	0,0432	11,901	0,0554	15,262	0,0127	0,3499
MNDL3.SA	-0,0319	-0,8961	0,0198	0,5562	-0,0373	-10,478	0,0233	0,6545	0,0072	0,2022
MSPA3.SA	0,0049	0,2311	0,0038	0,1792	0,0003	0,0142	-0,002	-0,0943	0,0003	0,0142
MSPA3.SA	-0,0015	-0,0617	-0,003	-0,1235	-0,0024	-0,0988	-0,0021	-0,0864	0,0002	0,0082
OGXP3.SA	0,0419	10,554	0,0133	0,335	0,0442	11,134	0,0403	10,151	-0,0135	-0,3401
OGXP3.SA	0,0623	0,5973	0,0009	0,0086	0,1652	15,839	-0,0587	-0,5628	-0,0016	-0,0153
PDGR3.SA	0,014	0,4531	0,0169	0,5469	-0,0168	-0,5437	0,0349	11,294	-0,0198	-0,6408
PETR3.SA	-0,0033	-0,2129	-0,0427	-2.75***	-0,0424	-2.7***	-0,0071	-0,4581	0,0495	3.193***
PETR3.SA	0,0165	0,8684	-0,0345	-1.8158*	0,0129	0,6789	-0,0187	-0,9842	-0,0021	-0,1105
PETR3.SA	-0,012	-0,7692	0,0242	15,513	0,0166	10,641	-0,0031	-0,1987	-0,0151	-0,9679
PETR3.SA	0,0097	0,7293	0,0047	0,3534	0,0072	0,5414	-0,0016	-0,1203	0,0121	0,9098
PETR3.SA	-0,0154	-0,9872	-0,0112	-0,7179	0,0102	0,6538	0,0021	0,1346	-0,0067	-0,4295
RCSL3.SA	-0,1384	-0,8793	-0,0012	-0,0076	0,0923	0,5864	0,1268	0,8056	-0,2297	-14,593
RCSL3.SA	-0,0039	-0,0437	-0,0014	-0,0157	0,0064	0,0717	0,2346	2.63***	0,0063	0,0706
RCSL3.SA	0,0168	0,2002	-0,0079	-0,0942	0,001	0,0119	0,0083	0,0989	-0,0012	-0,0143
RCSL3.SA	-0,0999	-14,416	0,1309	1.8889*	0,11	15,873	-0,0997	-14,387	0,0044	0,0635
RCSL3.SA	-0,0151	-0,2323	0,0023	0,0354	0,0088	0,1354	0,0026	0,04	-0,0005	-0,0077
RSID3.SA	-0,0104	-0,4502	-0,0333	-14,416	-0,0202	-0,8745	-0,0032	-0,1385	0,0089	0,3853
TASA3.SA	-0,0606	-2.4048**	0,015	0,5952	0	0	-0,0073	-0,2897	-0,0002	-0,0079

Retornos anormais individuais por evento 10 dias depois da realização da fraude (continuação)

Ação das firmas	AR (6)	Valor-t (6)	AR (7)	Valor-t (7)	AR (8)	Valor-t (8)	AR (9)	Valor-t (9)	AR (10)	Valor-t (10)
Antes da convergência contábil (até 2009)										
CEBR6.SA	-0,0032	-0,1468	-0,0042	-0,1927	-0,0017	-0,078	-0,002	-0,0917	-0,0031	-0,1422
EMBR3.SA	-0,0235	-10981	-0,0075	-0,3505	-0,008	-0,3738	0,0024	0,1121	0,0099	0,4626
PETR3.SA	0,0217	13478	-0,0025	-0,1553	0,0027	0,1677	-0,0041	-0,2547	0,0192	11925
PETR3.SA	0,0058	0,2522	-0,0126	-0,5478	-0,0046	-0,2	-0,0134	-0,5826	0,0201	0,8739
PTNT3.SA	0,0054	0,0891	0,0845	13944	0,0099	0,1634	0,009	0,1485	0,0056	0,0924
Depois da convergência contábil (a partir de 2010)										
BAHI3.SA	-0,0007	-0,0303	-0,0118	-0,5108	0,0036	0,1558	0,0027	0,1169	-0,0009	-0,039
BMYB34.SA	-0,003	-0,1807	-0,0011	-0,0663	-0,0019	-0,1145	-0,0013	-0,0783	-0,0011	-0,0663
CBMA3.SA	-0,0026	-0,0457	0,0968	1.7012*	0,09	15,817	0,0083	0,1459	-0,1905	-3.35***
CBMA3.SA	-0,0018	-0,0289	-0,0005	-0,008	0,0016	0,0257	-0,2871	-4.6***	0,0034	0,0547
CBMA3.SA	-0,0044	-0,034	-0,0028	-0,0217	-0,0039	-0,0302	0,0047	0,0363	-0,0066	-0,051
CEBR6.SA	0,002	0,0985	0,0019	0,0936	0,002	0,0985	0,0019	0,0936	0,0019	0,0936
CEBR6.SA	-0,0064	-0,2025	-0,001	-0,0316	0,0032	0,1013	-0,0084	-0,2658	0,0058	0,1835
CEBR6.SA	-0,0002	-0,0051	0,0014	0,0355	0,0018	0,0457	0,0033	0,0838	0,0001	0,0025
CGRA3.SA	0,0034	0,2112	0,026	16,149	0,0003	0,0186	0,0012	0,0745	-0,0072	-0,4472
CMIG3.SA	-0,0044	-0,1973	0,0087	0,3901	-0,0026	-0,1166	0,0752	3.37***	-0,0051	-0,2287
GGBR3.SA	0,0252	0,8542	0,0014	0,0475	-0,0242	-0,8203	-0,0121	-0,4102	-0,0199	-0,6746
IMBI3.SA	0,0015	0,0564	0,0013	0,0489	0,0014	0,0526	0,0002	0,0075	0,0018	0,0677
INEP3.SA	-0,0093	-0,1469	-0,0229	-0,3618	-0,039	-0,6161	0,0025	0,0395	-0,003	-0,0474
INEP3.SA	-0,0124	-0,3416	0,0001	0,0028	0,0779	2.146**	0,0601	1.6556*	-0,0905	-2.493**
JBSS3.SA	0,01	0,3759	0,0107	0,4023	-0,0011	-0,0414	-0,0089	-0,3346	0,0072	0,2707
JBSS3.SA	0,0148	0,3274	-0,0054	-0,1195	0,0067	0,1482	-0,0069	-0,1527		
JHSF3.SA	-0,006	-0,1653	0,006	0,1653	-0,0282	-0,7769	0,0025	0,0689	0,0227	0,6253
MNDL3.SA	0,0168	0,4719	-0,0037	-0,1039	-0,0232	-0,6517	0,0002	0,0056	-0,0098	-0,2753
MSPA3.SA	0,004	0,1887	0,0031	0,1462	0,0017	0,0802	0,0024	0,1132	0,0005	0,0236
MSPA3.SA	0,0006	0,0247	-0,0044	-0,1811	-0,0012	-0,0494	-0,0023	-0,0947		
OGXP3.SA	-0,009	-0,2267	0,0384	0,9673	0,0477	12,015	0,0234	0,5894	-0,0111	-0,2796
OGXP3.SA	0,1533	14,698	-0,0207	-0,1985	0,0328	0,3145	-0,0125	-0,1198	0,015	0,1438
PDGR3.SA	-0,0423	-13,689	-0,0185	-0,5987	0,0073	0,2362	0,039	12,621	-0,0002	-0,0065
PETR3.SA	0,0745	4.8065***	-0,0276	-1.7806*	0,0133	0,8581	-0,0198	-12,774	0,0721	4.652***
PETR3.SA	0,0025	0,1316	0,0059	0,3105	-0,0133	-0,7	-0,01	-0,5263	0,0228	1,2
PETR3.SA	0,0026	0,1667	0,005	0,3205	0,0057	0,3654	-0,0156	-1	0,0096	0,6154
PETR3.SA	0,0023	0,1729	-0,0012	-0,0902	0,0132	0,9925	0,0209	15,714	0,0105	0,7895
PETR3.SA	0,0075	0,4808	0,0071	0,4551	0,0172	11,026	-0,0182	-11,667	0,0074	0,4744
RCSL3.SA	-0,1307	-0,8304	0,0145	0,0921	0,2356	14,968	-0,0137	-0,087	0,0289	0,1836
RCSL3.SA	-0,0018	-0,0202	0,0002	0,0022	0,0033	0,037	0,0017	0,0191	0,0059	0,0661
RCSL3.SA	0,0184	0,2193	-0,0044	-0,0524	0,0046	0,0548	-0,003	-0,0358	-0,0047	-0,056
RCSL3.SA	-0,1063	-15,339	0,1273	1.8369*	0,1133	16,349	-0,0932	-13,449	0,007	
RCSL3.SA	-0,0021	-0,0323	0,0317	0,4877	-0,1219	-1.875*	0,1482	2.28**	-0,0027	-0,0415
RSID3.SA	-0,0106	-0,4589	-0,0054	-0,2338	0,0161	0,697	0,0162	0,7013	-0,0217	-0,9394
TASA3.SA	-0,0024	-0,0952	-0,0005	-0,0198	-0,0529	-2.10**	-0,0004	-0,0159	-0,0006	-0,0238

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.

Apêndice 7 – Retornos anormais individuais por evento 10 dias depois da detecção da fraude

Ação das firmas	AR (1)	Valor-t (1)	AR (2)	Valor-t (2)	AR (3)	Valor-t (3)	AR (4)	Valor-t (4)	AR (5)	Valor-t (5)
Antes da convergência contábil (até 2009)										
EMBR3.SA	-0.0369	-1.5311	-0.0079	-0.3278	-0.007	-0.2905	-0.0054	-0.2241	0.0105	0.4357
Depois da convergência contábil (a partir de 2010)										
BAHI3.SA	0.0087	0.3258	0.0348	1.3034	0.0483	1.809*	-0.0087	-0.3258	0.2006	7.513***
BMYB34.SA	0.0452	3.0133***	0.0172	1.1467	0.0214	1.4267	-0.0204	-1.36	-0.0001	-0.0067
CEBR6.SA	-0.0057	-0.1473	-0.0052	-0.1344	-0.0025	-0.0646	-0.0067	-0.1731	-0.0122	-0.3152
CEBR6.SA	-0.0032	-0.1778	0.0036	0.2	0.0019	0.1056	0.0014	0.0778	0	0
CGRA3.SA	0.0269	1.2227	-0.0203	-0.9227	-0.0063	-0.2864	-0.0051	-0.2318	-0.0318	-1.4455
CMIG3.SA	0.0298	1.3925	0.0758	3.54***	-0.0032	-0.1495	0.0075	0.3505	-0.0012	-0.0561
GGBR3.SA	0.0042	0.2763	0.0043	0.2829	0.0111	0.7303	-0.0114	-0.75	0.0227	1.4934
INEP3.SA	0.0078	0.1769	0.0055	0.1247	0.0088	0.1995	-0.0062	-0.1406	0.0069	0.1565
INEP3.SA	-0.0648	-1.0335	-0.0276	-0.4402	-0.0418	-0.6667	-0.0045	-0.0718	-0.0055	-0.0877
JBSS3.SA	-0.0085	-0.2607	0.0323	0.9908	-0.0131	-0.4018	-0.0223	-0.684	0.0163	0.5
JHSF3.SA	0.032	1.0997	0.0017	0.0584	0.0684	2.351**	-0.0001	-0.0034	-0.0266	-0.9141
MNDL3.SA	-0.0097	-0.2779	0.0281	0.8052	-0.0303	-0.8682	0.0367	1.0516	-0.0155	-0.4441
MSPA3.SA	-0.0006	-0.0317	-0.0006	-0.0317	-0.0005	-0.0265	-0.0005	-0.0265	-0.0006	-0.0317
OGXP3.SA	-0.0261	-0.5865	0.0104	0.2337	-0.0069	-0.1551	-0.0034	-0.0764	-0.0028	-0.0629
PDGR3.SA	-0.0136	-0.2416	0.1781	3.16***	-0.0896	-1.5915	0.0936	1.6625*	-0.0461	-0.8188
PETR3.SA	-0.0257	-1.1735	-0.0341	-1.5571	-0.0041	-0.1872	0.0574	2.62***	0.0583	2.662***
PETR3.SA	0.0072	0.6154	0.0098	0.8376	-0.0089	-0.7607	-0.0108	-0.9231	0.008	0.6838
PETR3.SA	-0.0012	-0.0446	0.0075	0.2788	-0.0386	-1.4349	-0.0001	-0.0037	-0.0177	-0.658
PTNT3.SA	-0.0000	-0.0011	-0.0001	-0.0116	-0.0001	-0.0116	-0.0001	-0.0116	0.0644	7.488***
PTNT3.SA	-0.0001	-0.0556	-0.0001	-0.0556	-0.0002	-0.1111	-0.0003	-0.1667	-0.0001	-0.0556
RCSL3.SA	0.0024	0.0596	-0.0061	-0.1514	-0.0105	-0.2605	-0.0372	-0.9231	0.0057	0.1414
RCSL3.SA	-0.0117	-0.3095	0.0476	1.2593	0.0039	0.1032	-0.0131	-0.3466	-0.0017	-0.045
RSID3.SA	0.0474	2.017**	0.0274	1.166	-0.0298	-1.2681	0.0065	0.2766	0.0147	0.6255
TASA3.SA	0.0056	0.1564	-0.0098	-0.2737	0.0413	1.1536	-0.0336	-0.9385	-0.003	-0.0838

Retornos anormais individuais por evento 10 dias depois da realização da fraude (continuação)

Ação das firmas	AR (6)	Valor-t (6)	AR (7)	Valor-t (7)	AR (8)	Valor-t (8)	AR (9)	Valor-t (9)	AR (10)	Valor-t (10)
Antes da convergência contábil (Até 2009)										
EMBR3.SA	-0,0235	-10981	-0,0075	-0,3505	-0,008	-0,3738	0,0024	0,1121	0,0099	0,4626
Depois da convergência contábil (A partir de 2010)										
BAHI3.SA	-0.113	-4.23***	-0.0338	-1.2659	0.0132	0.4944	0.0075	0.2809	0.0104	0.3895
BMYB34.SA	-0.0208	-1.3867	-0.0021	-0.14	-0.0017	-0.1133	0.014	0.9333	-0.0256	-1.7067*
CEBR6.SA	-0.0049	-0.1266	0.0012	0.031	-0.0065	-0.168	-0.0028	-0.0724	-0.006	-0.155
CEBR6.SA	-0.2024	-11.2***	0.0562	3.12***	0.0016	0.0889	-0.0446	-2.47**	-0.195	-10.8***
CGRA3.SA	-0.0106	-0.4818	-0.017	-0.7727	0.0056	0.2545	0.0111	0.5045	-0.0165	-0.75
CMIG3.SA	0.0029	0.1355	-0.0058	-0.271	-0.0314	-1.4673	0.0168	0.785	-0.0358	-1.6729*
GGBR3.SA	-0.0112	-0.7368	0.0041	0.2697	0.029	1.9079*	0.0068	0.4474	-0.0108	-0.7105
INEP3.SA	-0.0223	-0.5057	0.0044	0.0998	0.0249	0.5646	-0.0138	-0.3129	0.0052	0.1179
INEP3.SA	0.0041	0.0654	-0.0195	-0.311	0.0171	0.2727	0.0295	0.4705	-0.0495	-0.7895
JBSS3.SA	-0.0162	-0.4969	-0.003	-0.092	-0.0001	-0.0031	0.0451	1.3834	-0.0087	-0.2669
JHSF3.SA	-0.0069	-0.2371	-0.0103	-0.354	0.0293	1.0069	-0.0036	-0.1237	0.0072	0.2474
MNDL3.SA	-0.0041	-0.1175	-0.007	-0.2006	0.0295	0.8453	-0.0059	-0.1691	-0.0037	-0.106
MSPA3.SA	-0.0005	-0.0265	-0.0005	-0.0265	-0.0006	-0.0317	-0.0005	-0.0265	-0.0006	-0.0317
OGXP3.SA	-0.0032	-0.0719	-0.0088	-0.1978	0.0015	0.0337	0.0114	0.2562	0.0177	0.3978
PDGR3.SA	0.0489	0.8686	-0.041	-0.7282	-0.0122	-0.2167	-0.0472	-0.8384	-0.0091	-0.1616
PETR3.SA	0.0090	0.411	-0.0125	-0.5708	-0.0066	-0.3014	0.0079	0.3607	0.0012	0.0548
PETR3.SA	-0.0034	-0.2906	-0.009	-0.7692	0.0008	0.0684	0.0004	0.0342	-0.0162	-1.3846
PETR3.SA	-0.0073	-0.2714	0.0282	1.0483	-0.012	-0.4461	-0.0051	-0.1896	-0.0063	-0.2342
PTNT3.SA	-0.0002	-0.0233	-0.0001	-0.0116	0	0	-0.0001	-0.0116	-0.0001	-0.0116
PTNT3.SA	-0.0001	-0.0556	-0.0001	-0.0556	-0.0002	-0.1111	-0.0002	-0.1111	-0.0001	-0.0556
RCSL3.SA	0.0178	0.4417	0.011	0.273	0.0346	0.8586	0.0072	0.1787	0.1108	2.749***
RCSL3.SA	-0.0036	-0.0952	0.0155	0.4101	-0.0081	-0.2143	0.0278	0.7354	0.015	0.3968
RSID3.SA	0.0091	0.3872	0.0301	1.2809	-0.0009	-0.0383	0.0443	1.8851*	-0.0335	-1.4255

TASA3.SA	0.0017	0.0475	0.0411	1.148	-0.0028	-0.0782	0.0343	0.9581	-0.017	-0.4749
----------	--------	--------	--------	-------	---------	---------	--------	--------	--------	---------

Fonte: dados da pesquisa.

Legenda: (*), (**) e (***) representam uma significância, respectivamente, a 10%, 5% e 1%.